

RESEARCH ARTICLE

Open Access



# Evolution of the avian $\beta$ -defensin and cathelicidin genes

Yuanyuan Cheng<sup>1\*†</sup>, Michael Dennis Prickett<sup>2†</sup>, Weronika Gutowska<sup>3</sup>, Richard Kuo<sup>3</sup>, Katherine Belov<sup>1</sup> and David W. Burt<sup>3\*</sup>

## Abstract

**Background:**  $\beta$ -defensins and cathelicidins are two families of cationic antimicrobial peptides (AMPs) with a broad range of antimicrobial activities that are key components of the innate immune system. Due to their important roles in host defense against rapidly evolving pathogens, the two gene families provide an ideal system for studying adaptive gene evolution. In this study we performed phylogenetic and selection analyses on  $\beta$ -defensins and cathelicidins from 53 avian species representing 32 orders to examine the evolutionary dynamics of these peptides in birds.

**Results and conclusions:** Avian  $\beta$ -defensins are found in a gene cluster consisting of 13 subfamilies. Nine of these are conserved as one to one orthologs in all birds, while the others (AvBD1, AvBD3, AvBD7 and AvBD14) are more subject to gene duplication or pseudogenisation events in specific avian lineages. Avian cathelicidins are found in a gene cluster consisting of three subfamilies with species-specific duplications and gene loss. Evidence suggested that the propiece and mature peptide domains of avian cathelicidins are possibly co-evolving in such a way that the cationicity of the mature peptide is partially neutralised by the negative charge of the propiece prior to peptide secretion (further evidence obtained by repeating the analyses on primate cathelicidins). Negative selection (overall mean  $d_N < d_S$ ) was detected in most of the gene domains examined, conserving certain amino acid residues that may be functionally crucial for the avian  $\beta$ -defensins and cathelicidins, while episodic positive selection was also involved in driving the diversification of specific codon sites of certain AMPs in avian evolutionary history. These findings have greatly improved our understanding of the molecular evolution of avian AMPs and will be useful to understand their role in the avian innate immune response. Additionally, the large dataset of  $\beta$ -defensin and cathelicidin peptides may also provide a valuable resource for translational research and development of novel antimicrobial agents in the future.

## Background

Defensins and cathelicidins are two families of cationic small peptides that have broad-spectrum antimicrobial activities against a wide range of bacterial, fungal or viral pathogens. These peptides are produced in a large variety of invertebrate and vertebrate organisms, representing an ancient form of host defense against microbes. In addition to their antimicrobial function, defensins and cathelicidins have also been found to exhibit diverse immunomodulatory activities, rendering them important components of both innate and adaptive immune systems [1–3]. With a

key role in host defense against rapidly evolving pathogens, defensin and cathelicidin gene families provide an ideal system for studying adaptive molecular evolution [4]. Previous studies in mammalian lineages have revealed positive selection driving rapid divergence of these host defense peptides [4, 5]. Recent whole-genome sequence analysis of 48 bird species [6] has enabled us to perform a comprehensive comparative analysis on the avian lineages of defensins and cathelicidins, which will not only greatly improve our understanding of the evolutionary diversification of these ancient peptides over the past 100 million years through the avian radiation, but also provide a valuable resource for developing novel antibiotics to treat microbial infections in birds and other vertebrates [7, 8].

Among the three defensin subfamilies ( $\alpha$ ,  $\beta$  and  $\theta$ ) identified in vertebrates, only  $\beta$ -defensins have been found in birds [9, 10]. Avian  $\beta$ -defensins (AvBDs) have

\* Correspondence: [yuanyuan.cheng@sydney.edu.au](mailto:yuanyuan.cheng@sydney.edu.au); [Dave.Burt@roslin.ed.ac.uk](mailto:Dave.Burt@roslin.ed.ac.uk)

<sup>†</sup>Equal contributors

<sup>1</sup>RMC Gunn Building B19, Faculty of Veterinary Science, University of Sydney, Camperdown 2006 NSW, Australia

<sup>3</sup>The Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, Easter Bush, Midlothian EH25 9RG, UK

Full list of author information is available at the end of the article

been previously described in the chicken (*Gallus gallus*), zebra finch (*Taeniopygia guttata*), mallard (*Anas platyrhynchos*) and crested ibis (*Nipponia nippon*) with a cluster of 14, 22, 18 and 14 closely linked  $\beta$ -defensin genes found in each genome, respectively [9–13]. AvBD genes usually consist of 3–4 exons [10]. The first exon contains the 5'-untranslated region and the others contain the open reading frame encoding three peptide domains, including a hydrophobic, leucine-rich signal peptide, a short propiece (absent in *AvBD12*) and the mature peptide. The mature  $\beta$ -defensin forms a characteristic  $\beta$  sheet-rich fold, which contains three disulfide bridges (Cys<sup>1</sup>-Cys<sup>5</sup>, Cys<sup>2</sup>-Cys<sup>4</sup> and Cys<sup>3</sup>-Cys<sup>6</sup>) within a conserved six-cysteine framework. AvBDs have been shown to be expressed in a number of tissues with varying levels of expression [14, 15] and play a role in the innate host defense against viruses, Gram-positive and Gram-negative bacteria, and parasites in the digestive system [16–21] and reproductive system [22–26]. Evidence has also indicated that they may have similar regulatory functions as mammalian  $\beta$ -defensins in various immune processes, such as chemotaxis, wound-healing, and inducing or suppressing inflammatory responses [reviewed in 1]. Within-species genetic diversity and trans-species polymorphisms of AvBDs have been observed in passerine bird species [27, 28].

Cathelicidins are characterised by having a conserved cathelin domain and a highly variable mature peptide domain. Four avian cathelicidin genes have been identified in the chicken, designated *CATHL1*, *CATHL2*, *CATHL3* and *CATHB1* [2, 29–31]. These genes share similar structures with mammalian cathelicidins, each comprising of four exons, encoding a propiece consisting of a signal peptide, the cathelin-like domain (propiece) and the mature peptide. One unusual feature found in the propiece of chicken *CATHB1*, is the presence of nine octamer repeats (PGLDGSXS) N-terminal to the cathelin domain [30], which is not seen in other cathelicidin genes. Mature avian cathelicidins are activated before secretion with the propiece cleaved off by serine proteases [32] and form an alpha-helix predominant structure with a kink induced by a glycine or proline close to the centre [33–35]. All four chicken cathelicidins show high levels of expression in immune organs and gastrointestinal, respiratory and urogenital tracts, with *CATHL1*, 2 and 3 most highly expressed in the bone marrow and lung, and *CATHB1* in the bursa of Fabricius [36]. In addition to broad antimicrobial activity against bacteria and fungi, avian cathelicidins also play a range of immunoregulatory roles, such as blocking lipopolysaccharide-induced cytokine expression [37] or inducing specific chemokine production [31].

In this paper, we studied the evolution of  $\beta$ -defensins and cathelicidins in 53 avian species (Table 1) and discussed the following issues: conservation and lineage-

specific duplication/deletion of genes, conserved genomic organisation of  $\beta$ -defensin and cathelicidin gene clusters, coevolution between pro- and mature peptides in cathelicidins, and amino acid sites and evolutionary branches under selection.

## Results and discussion

### Avian $\beta$ -defensins and cathelicidins

A total of 758 genes, including 714  $\beta$ -defensins and 44 cathelicidins, were annotated in 53 avian genomes (scaffold number and gene coordinates provided in Additional file 1; nucleotide sequences in Additional file 2). The selected species represent 32 different avian orders spanning 52 genera (Table 1), last sharing a common ancestor around 114 million years ago (Fig. 1) [38].

Annotated genes were designated based on their orthologous relationships with chicken  $\beta$ -defensins (*AvBD1*–*AvBD14*) and cathelicidins (*CATHL1*, *CATHL2*, *CATHL3* and *CATHB1*). Phylogenetic analysis revealed 13 distinct AvBD groups, namely AvBD1, AvBD2, AvBD3, AvBD4, AvBD5, AvBD6 and AvBD7, AvBD8, AvBD9, AvBD10, AvBD11, AvBD12, AvBD13, and AvBD14 (Fig. 2a), while cathelicidin genes cluster into three clades, *CATHL1* and *CATHL3*, *CATHL2*, and *CATHB1* (Fig. 2b).

Characteristics of AvBD gene subfamilies are summarised in Table 2 (alignments shown in Additional file 3). Ten genes (AvBD2, 4, 5, 7, 8, 9, 10, 11, 12 and 13) are conserved as one-to-one orthologues in most surveyed species, suggesting a high level of conservation of these genes for over 100 million years (Fig. 3). *AvBD1* and *AvBD3* are subject to lineage-specific expansion, with up to three *AvBD1* paralogues found in the killdeer (*Charadrius vociferus*), saker falcon (*Falco cherrug*), medium ground finch (*Geospiza fortis*) and zebra finch; extensive *AvBD3* duplications occurred in Passeriformes with up to 14 paralogues found in the white-throated sparrow (*Zonotrichia albicollis*). *AvBD6* is a duplication of *AvBD7*, which has arisen within Galliformes after the divergence of family Odontophoridae and before Phasianidae, as it is present in chicken and turkey but absent in the northern bobwhite (*Colinus virginianus*). *AvBD7* has degenerated into a pseudogene in the sunbittern (*Eurypyga helias*) and is missing in all three psittacines. *AvBD14* was found in 23 species and has been degraded into a pseudogene in the orders Falconiformes, Passeriformes and Psittaciformes, and species *Columbia livia*, *Gavia stellata*, and *Calypte anna*.

The spacing between cysteines within the defensin motif is conserved within subfamilies with the overall consensus being C (4-7) C (3-6) C (7-10) C (5-6) CC, numbers representing the number of residues between cysteines (Table 2). The majority of subfamilies have 9–11 residues preceding the first cysteine of the defensin motif. Exceptions include the group of passeriforme AvBD3s immediately preceding AvBD5 which have 2

**Table 1** List of 53 examined bird species

Scientific name	Common name	Abbreviation	Order	Family	Genus	GenBank Assembly ID	AvBD1 paralogues	AvBD3 paralogues	Pseudogenes	Genes not found	Cluster status
<i>Acanthisitta chloris</i>	Rifleman	ACACH	Passeriformes	Acanthisittidae	Acanthisitta	GCA_000695815	2	5	AvBD11, AvBD14		fragmented
<i>Amazona vittata</i>	Puerto Rican parrot	AMAVI	Psittaciformes	Psittacidae	Amazona	GCA_000332375	2	1	AvBD7, AvBD14		fragmented
<i>Anas platyrhynchos</i>	American Pekin duck	ANAPL	Anseriformes	Anatidae	Anas	GCA_000355885	1	6			fragmented
<i>Apaloderma vittatum</i>	Bar-tailed trogon	APAVI	Trogoniformes	Trogonidae	Apaloderma	GCA_000703405	1	2		AvBD14	fragmented
<i>Aptenodytes forsteri</i>	Emperor penguin	APTFO	Sphenisciformes	Spheniscidae	Aptenodytes	GCA_000699145	1	1			intact
<i>Balearica regulorum gibbericeps</i>	Grey crowned crane	BALRE	Gruiformes	Gruidae	Balearica	GCA_000709895	1	2			fragmented
<i>Buceros rhinoceros silvestris</i>	Rhinoceros hornbill	BUCRH	Bucerotiformes	Bucerotidae	Buceros	GCA_000710305	1	0		AvBD3, AvBD12	fragmented
<i>Calypte anna</i>	Anna's hummingbird	CALAN	Trochiliformes	Trochilidae	Calypte	GCA_000699085	1	1	AvBD7, AvBD14	AvBD11	fragmented
<i>Caprimulgus carolinensis</i>	Chuck-will's-widow	CAPCA	Caprimulgiformes	Caprimulgidae	Antristomus	GCA_000700745	2	1			fragmented
<i>Cariama cristata</i>	Red-legged seriema	CARCR	Gruiformes	Cariamidae	Cariama	GCA_000690535	2	1			fragmented
<i>Cathartes aura</i>	Turkey vulture	CATAU	Cathartiformes	Cathartidae	Cathartes	GCA_000699945	1	0		AvBD7	fragmented
<i>Chaetura pelagica</i>	Chimney swift	CHAPE	Apodiformes	Apodidae	Chaetura	GCA_000747805	1	3	AvBD7, AvBD11		fragmented
<i>Charadrius vociferus</i>	Killdeer	CHAVO	Charadriiformes	Charadriidae	Charadrius	GCA_000708025	3	2			fragmented
<i>Chlamydotis undulata</i>	Houbara bustard	CHLUN	Gruiformes	Otididae	Chlamydotis	GCA_000695195	2	1		AvBD14	fragmented
<i>Colinus virginianus</i>	Northern bobwhite	COLVI	Galliformes	Odontophoridae	Colinus	GCA_000599465	1	1			fragmented <sup>a</sup>
<i>Colius striatus</i>	Speckled mousebird	COLST	Coliiformes	Coliidae	Colius	GCA_000690715	1	1		AvBD13	fragmented <sup>a</sup>
<i>Columba livia</i>	Rock dove	COLLI	Columbiformes	Columbidae	Columba	GCA_000337935	1	2	AvBD14		fragmented <sup>a</sup>
<i>Corvus brachyrhynchos</i>	American crow	CORBR	Passeriformes	Corvidae	Corvus	GCA_000691975	2	6	AvBD14		fragmented
<i>Cuculus canorus</i>	Common cuckoo	CUCCA	Cuculiformes	Cuculidae	Cuculus	GCA_000709325	2	2			fragmented <sup>a</sup>
<i>Egretta garzetta</i>	Little egret	EGRGA	Ciconiiformes	Ardeidae	Egretta	GCA_000687185	2	1		AvBD14	fragmented

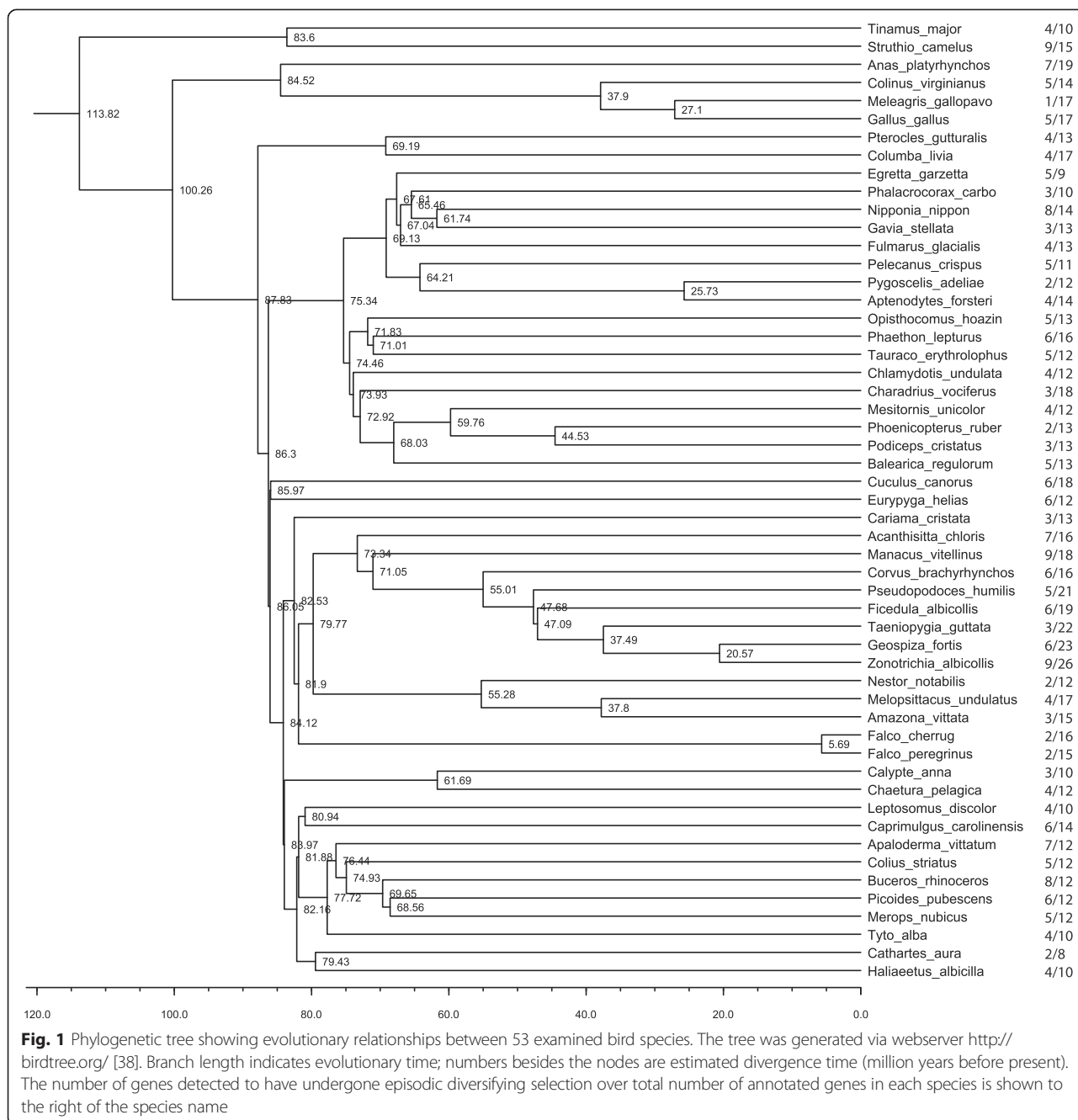
**Table 1** List of 53 examined bird species (Continued)

<i>Eurypyga helias</i>	Sunbittern	EURHE	Gruiformes	Eurypygidae	Eurypyga	GCA_000690775	1	1	AvBD7	AvBD14	fragmented
<i>Falco cherrug</i>	Saker falcon	FALCH	Falconiformes	Falconidae	Falco	GCA_000337975	3	1	AvBD14		fragmented <sup>a</sup>
<i>Falco peregrinus</i>	Peregrine falcon	FALPE	Falconiformes	Falconidae	Falco	GCA_000337955	2	1	AvBD14	AvBD3.3	fragmented
<i>Ficedula albicollis</i>	Collared flycatcher	FICAL	Passeriformes	Muscicapidae	Ficedula	GCA_000247815	2	6	AvBD14		fragmented
<i>Fulmarus glacialis</i>	Northern fulmar	FULGL	Procellariiformes	Procellariidae	Fulmarus	GCA_000690835	2	1	AvBD1.1	Exon2 AvBD13	fragmented
<i>Gallus gallus</i>	Chicken	GALGA	Galliformes	Phasianidae	Gallus	GCA_000002315	1	1			intact
<i>Gavia stellata</i>	Red-throated loon	GAVST	Gaviiformes	Gaviidae	Gavia	GCA_000690875	1	1	AvBD14		fragmented
<i>Geospiza fortis</i>	Medium ground finch	GEOFO	Passeriformes	Fringillidae	Geospiza	GCA_000277835	3	9	AvBD14		intact
<i>Haliaeetus albicilla</i>	White-tailed eagle	HALAL	Accipitriformes	Accipitridae	Haliaeetus	GCA_000691405	1	1		AvBD13	fragmented
<i>Leptosomus discolor</i>	Cuckoo roller	LEPDI	Coraciiformes	Leptosomidae	Leptosomus	GCA_000691785	1	1		AvBD2	fragmented
<i>Manacus vitellinus</i>	Golden-collared manakin	MANVI	Passeriformes	Pipridae	Manacus	GCA_000692015	1	7	AvBD14		fragmented
<i>Meleagris gallopavo</i>	Turkey	MELGA	Galliformes	Phasianidae	Meleagris	GCA_000146605	1	1			intact
<i>Melopsittacus undulatus</i>	Budgerigar	MELUN	Psittaciformes	Psittacidae	Melopsittacus	GCA_000238935	2	1	AvBD7, AvBD14		intact
<i>Merops nubicus</i>	Carmine bee-eater	MERNU	Coraciiformes	Meropidae	Merops	GCA_000691845	1	1			fragmented
<i>Mesitornis unicolor</i>	Brown roatelo	MESUN	Gruiformes	Mesitornithidae	Mesitornis	GCA_000695765	2	3		AvBD14	fragmented
<i>Nestor notabilis</i>	Kea	NESNO	Psittaciformes	Psittacidae	Nestor	GCA_000696875	1	1	AvBD7	AvBD14	fragmented
<i>Nipponia nippon</i>	Crested ibis	NIPNI	Ciconiiformes	Threskiornithidae	Nipponia	GCA_000708225	2	1			intact
<i>Opisthocomus hoazin</i>	Hoatzin	OPIHO	Opisthocomiformes	Opisthocomidae	Opisthocomus	GCA_000692075	1	1	AvBD9		fragmented <sup>a</sup>
<i>Pelecanus crispus</i>	Dalmatian pelican	PELCR	Pelecaniformes	Pelecanidae	Pelecanus	GCA_000687375	1	1		Exon 2 AvBD13	fragmented <sup>a</sup>
<i>Phaethon lepturus</i>	White-tailed tropicbird	PHALE	Phaethontiformes	Phaethontidae	Phaethon	GCA_000687285	2	2			fragmented
<i>Phalacrocorax carbo</i>	Great cormorant	PHACA	Pelecaniformes	Phalacrocoracidae	Phalacrocorax	GCA_000708925	1	1		AvBD4, AvBD11, AvBD14	fragmented
<i>Phoenicopterus ruber</i>	American flamingo	PHORU	Phoenicopteriformes	Phoenicopteridae	Phoenicopterus	GCA_000687265	2	2		AvBD14	fragmented

**Table 1** List of 53 examined bird species (Continued)

<i>Picoides pubescens</i>	Downy woodpecker	PICPU	Piciformes	Picidae	Picoides	GCA_000699005	1	1	AvBD7	fragmented	
<i>Podiceps cristatus</i>	Great crested grebe	PODCR	Podicipediformes	Podicipedidae	Podiceps	GCA_000699545	2	1	AvBD14	fragmented	
<i>Pseudopodoces humilis</i>	Tibetan ground-tit	PSEHU	Passeriformes	Paridae	Pseudopodoces	GCA_000331425	2	7	AvBD14	intact	
<i>Pterocles gutturalis</i>	Yellow-throated sandgrouse	PTEGU	Ciconiiformes	Pteroclididae	Pterocles	GCA_000699245	2	1		fragmented	
<i>Pygoscelis adeliae</i>	Adelie penguin	PYGAD	Sphenisciformes	Spheniscidae	Pygoscelis	GCA_000699105	1	1		intact	
<i>Struthio camelus</i>	African ostrich	STRCA	Struthioniformes	Struthionidae	Struthio	GCA_000698965	2	2		fragmented	
<i>Taeniopygia guttata</i>	Zebra finch	TAEGU	Passeriformes	Estrildidae	Taeniopygia	GCA_000151805	3	9	AvBD14	intact	
<i>Tauraco erythrolophus</i>	Red-crested turaco	TAUER	Musophagiformes	Musophagidae	Tauraco	GCA_000709365	2	2	AvBD7	AvBD14	fragmented
<i>Tinamus major</i>	Great tinamou	TINMA	Tinamiformes	Tinamidae	Tinamus	GCA_000705375	1	1		fragmented	
<i>Tyto alba</i>	Barn owl	TYTAL	Strigiformes	Tytonidae	Tyto	GCA_000687205	1	1	AvBD5, AvBD13, AvBD14	fragmented	
<i>Zonotrichia albicollis</i>	White-throated sparrow	ZONAL	Passeriformes	Fringillidae	Zonotrichia	GCA_000385455	2	14	AvBD14	fragmented	

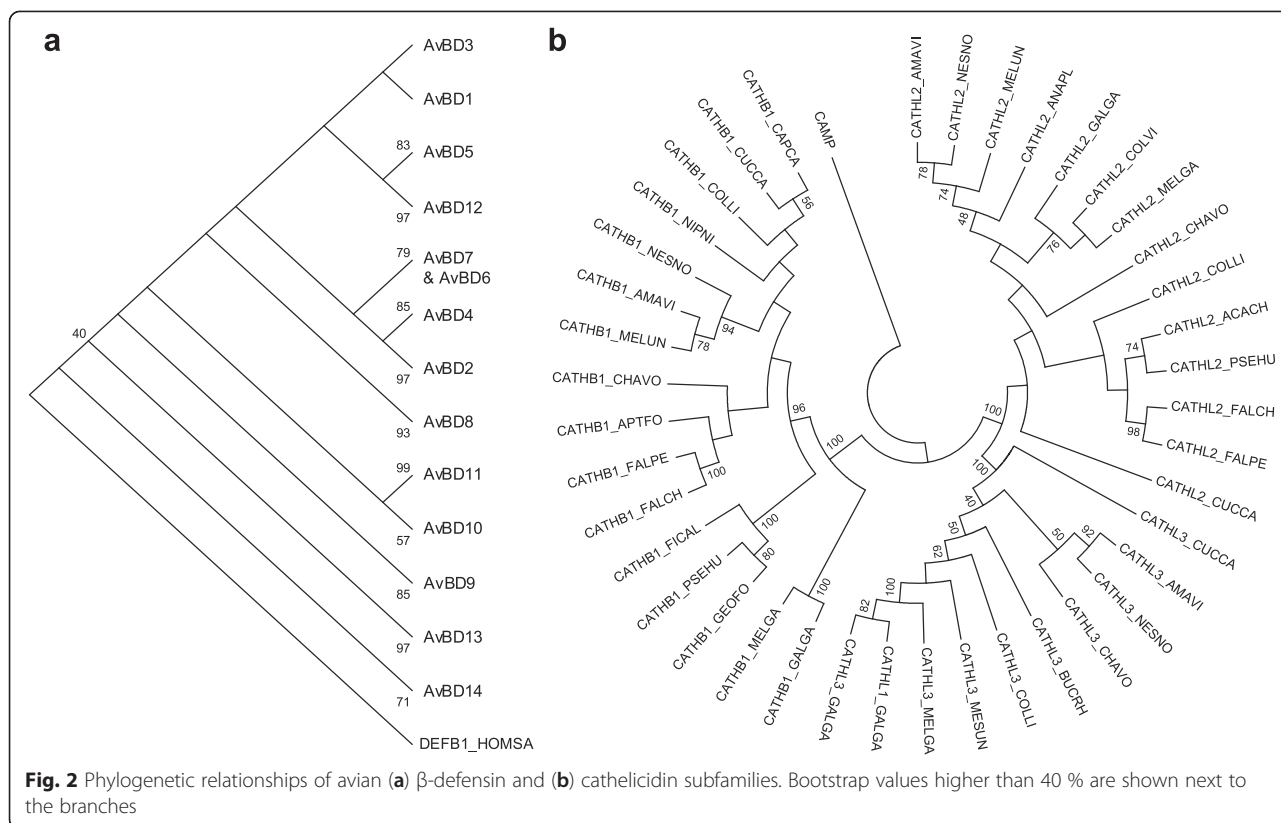
<sup>a</sup> indicates an intact region between AvBD2 and AvBD5 in the genomic assembly, thus the actual number and position of AvBD1(s) and AvBD3(s) can be determined



residues prior to the first cysteine in addition to having only 2 exons. AvBD1 and AVBD3 have the same spacing of cysteines but differ in the number of residues prior to the first cysteine.

Cathelicidins were only found in 21 surveyed species due to low assembly quality of the genomic regions. Similar to the 10 relatively conserved AvBDs, *CATHL2*, *CATHL3* and *CATHB1* have been conserved across a variety of avian orders (Fig. 4). Evidence suggested that *CATHL3* has been reduced to a pseudogene in Falconiformes and lost in Passeriformes, yet duplicated to

give rise to *CATHL1* in Galliformes. Annotation results also suggested that *CATHL2* and *CATHL3* may have been lost in Sphenisciformes and Ciconiiformes, represented by the emperor penguin (*Aptenodytes forsteri*) and crested ibis (*Nipponia nippon*), respectively (Fig. 4), though this could be an artefact caused by high degree of assembly gaps in the genomic region. The octamer repeats feature found in chicken *CATHB1* appeared to be unique to Galliformes (or Phasianidae), as it has only been seen in the chicken and turkey.



**Conserved synteny**

Avian  $\beta$ -defensins cluster in a syntenic region flanked by cathepsin B (*CTSB*) and translocation associated membrane protein 2 (*TRAM2*) genes (Fig. 3). The gene order is highly conserved as *CTSB*, AvBD13, 12, 11, 10, 9, 8, 7, 2, 1, 3, 5, 4, 14, and *TRAM2*, with duplicated AvBD1

and 3 interspersed in the region between AvBD2 and 5. The strand orientations of AvBDs are also strictly maintained, with the exception of several inverted AvBD1 and 3 gene duplicates in Passeriformes. However, it should be noted that most of the genome assemblies surveyed were of draft quality with defensins found on

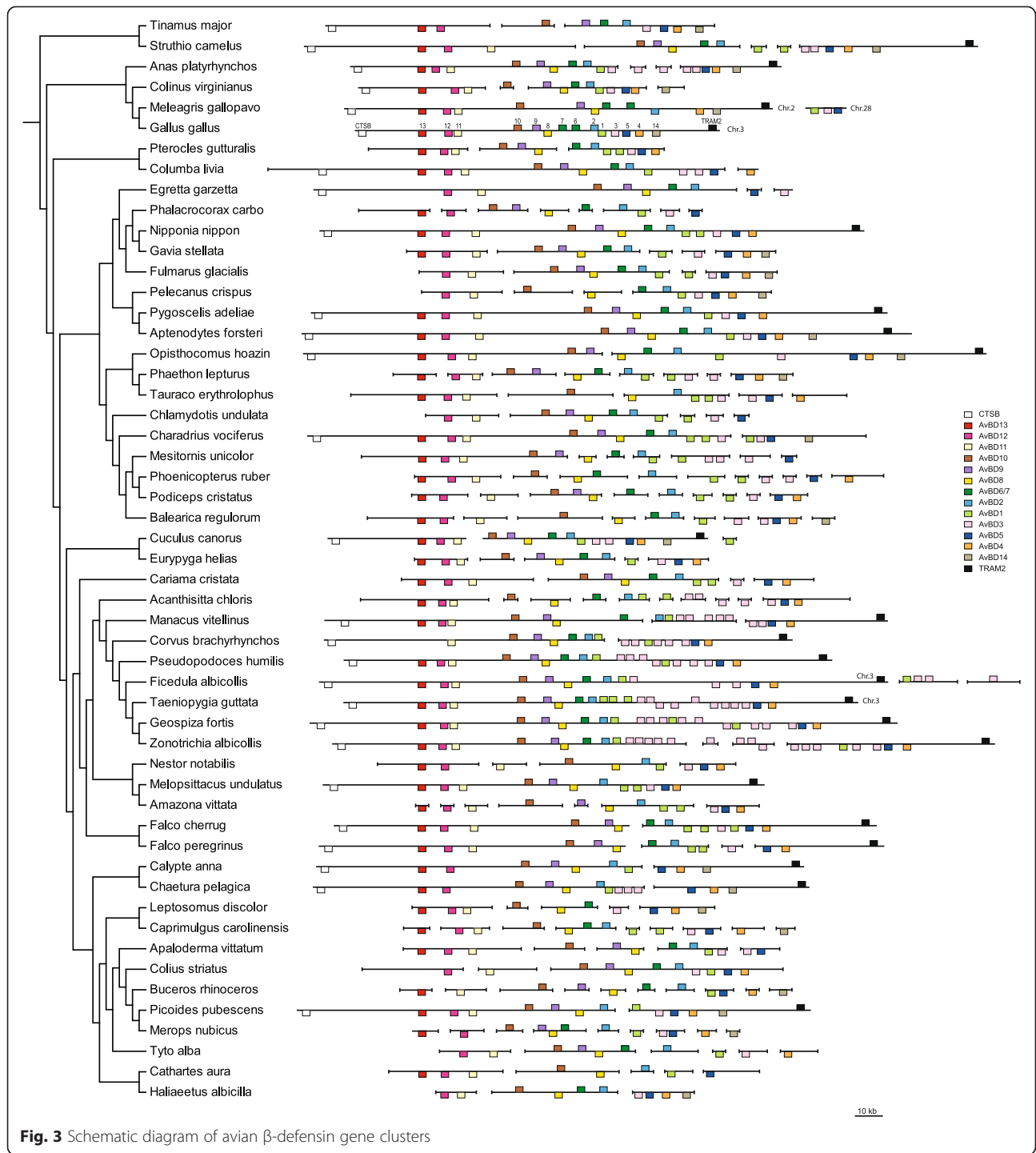
**Table 2** Characteristics of avian  $\beta$ -defensin subfamilies

Gene <sup>a</sup>	Type	Defensin motif <sup>b</sup>	Coding exons <sup>c</sup>
AvBD13	one to one ortholog	(9) C (6) C (3) C (9) C (6) CC	2 (3 in Galliformes)
AvBD12	one to one ortholog	(4–6) C (6) C (4) C (9) C (6) CC (10–11)	2
AvBD11	one to one ortholog	(9) C (6) C (5) C (9) C (6) CC(9) C (6) C (6) C (7) C (6) CC	3 (defensin motif duplicated)
AvBD10	one to one ortholog	(11) C (6–7) C (4) C (9) C (6) CC	3
AvBD9	one to one ortholog	(10) C (6) C (4) C (9) C (6) CC	3
AvBD8	one to one ortholog	(11) C (6) C (4) C (10) C (5) CC	3
AvBD7	Duplicated as AvBD6 in Galliformes	(11) C (6) C (4) C (9) C (5) CC	3
AvBD2	one to one ortholog	(11) C (4) C (4) C (9) C (5) CC	3 (3rd exon: 14 residues in Passeriformes)
AvBD1	Up to 3 paralogs in a single species	(11) C (6) C (4) C (9) C (5) CC	3
AvBD3	Up to 14 paralogs in a single species	(5–7) C (6) C (4) C (9) C (5) CC	3 (3rd exon: 12–14 residues in Galliformes)
AvBD5	one to one ortholog	(11) C (6) C (4) C (9) C (5) CC	3
AvBD4	one to one ortholog	(9–14) C (6) C (4) C (9) C (5) CC	2–3
AvBD14	One to one ortholog	(6) C (6) C (4) C (9) C (6) CC	2

<sup>a</sup> Genes are listed in the order from the flanking gene cathepsin B to the flanking gene TRAM2

<sup>b</sup> The number of residues between cysteines are noted in parenthesis

<sup>c</sup> Third exons only code for 2–5 residues unless otherwise noted



**Fig. 3** Schematic diagram of avian  $\beta$ -defensin gene clusters

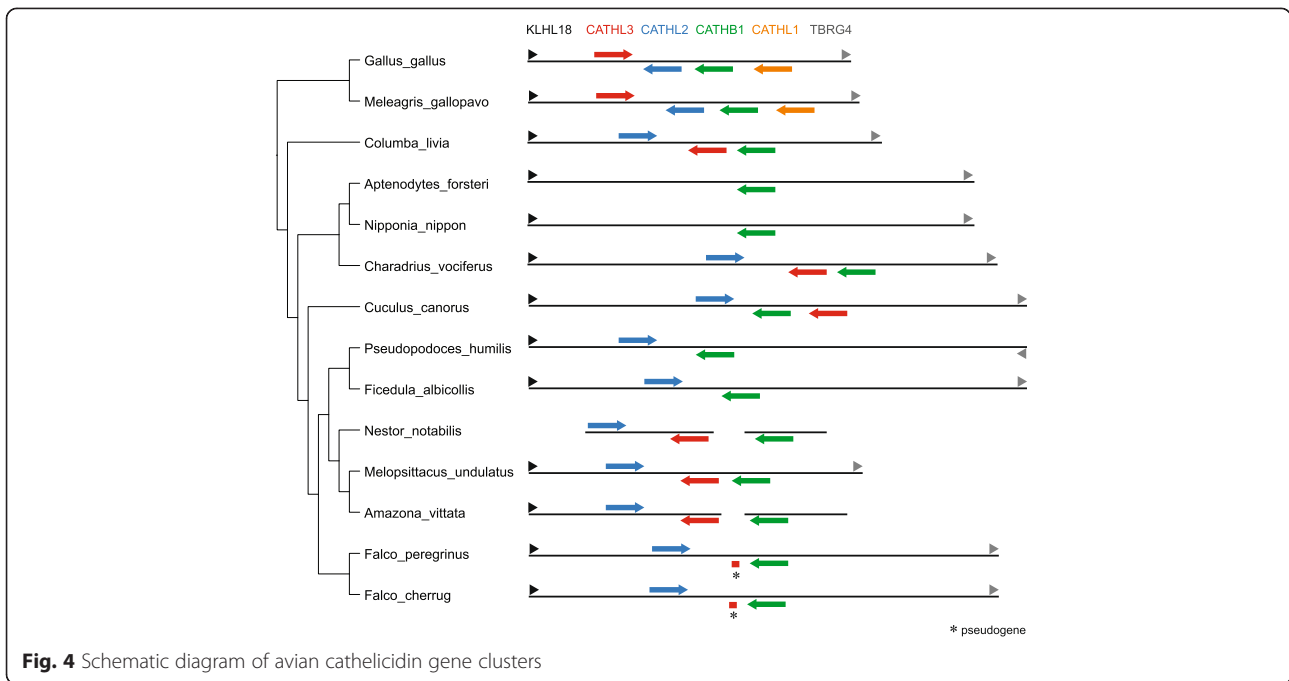
up to 11 different scaffolds/contigs, thus there may be unrevealed lineage-specific gene rearrangements.

Similar to AvBDs, avian cathelicidins also form a conserved gene cluster, which is flanked by kelch-like family member 18 (*KLHL18*) and transforming growth factor beta regulator 4 (*TBRG4*) genes (Fig. 4). The majority of species share a conserved gene order of *KLHL18*, *CATHL2*, *CATHL3*, *CATHB1* and *TBRG4*. Exceptions

include an inversion of the region containing *CATHL3* and *CATHL2* in Galliformes, and *CATHL3* and *CATHB1* being arranged in the reverse order in the common cuckoo (*Cuculus canorus*).

Such clustering of homologous genes in a tightly linked fashion is rather common for immune genes. Other well-known examples include the major histocompatibility complex (MHC), immunoglobulins (Ig),



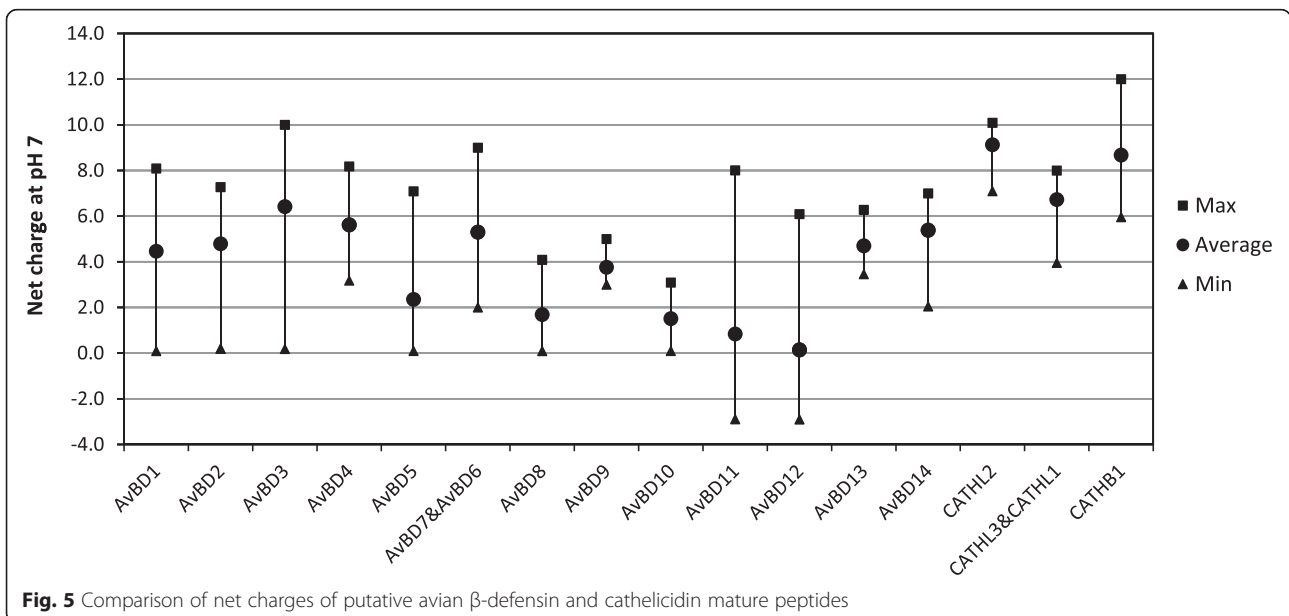


**Fig. 4** Schematic diagram of avian cathelicidin gene clusters

Fc receptors (FcR) and killer-cell Ig-like receptors (KIR). These immune gene families are believed to be regularly refreshed via *cis* duplication, resulting in related genes lying next to each other in linked array in the genome [39]. It has been suggested that immune genes clustering together may be biologically significant in that it may facilitate the coordinated expression of functionally related loci, and therefore has been selectively maintained [39, 40].

**Net charge of mature peptides of  $\beta$ -defensins and cathelicidins**

Disruption of microbial membranes is a major mechanism underlying antimicrobial activity of defensins and cathelicidins [41] and it has been demonstrated that the net charge of the mature peptide directly influences its antimicrobial potency [42]. The net charge of the mature peptide ranges from -2.9 to +10.0 in avian  $\beta$ -defensins and +4.0 to +12.0 in cathelicidins (Fig. 5). On average,



**Fig. 5** Comparison of net charges of putative avian  $\beta$ -defensin and cathelicidin mature peptides

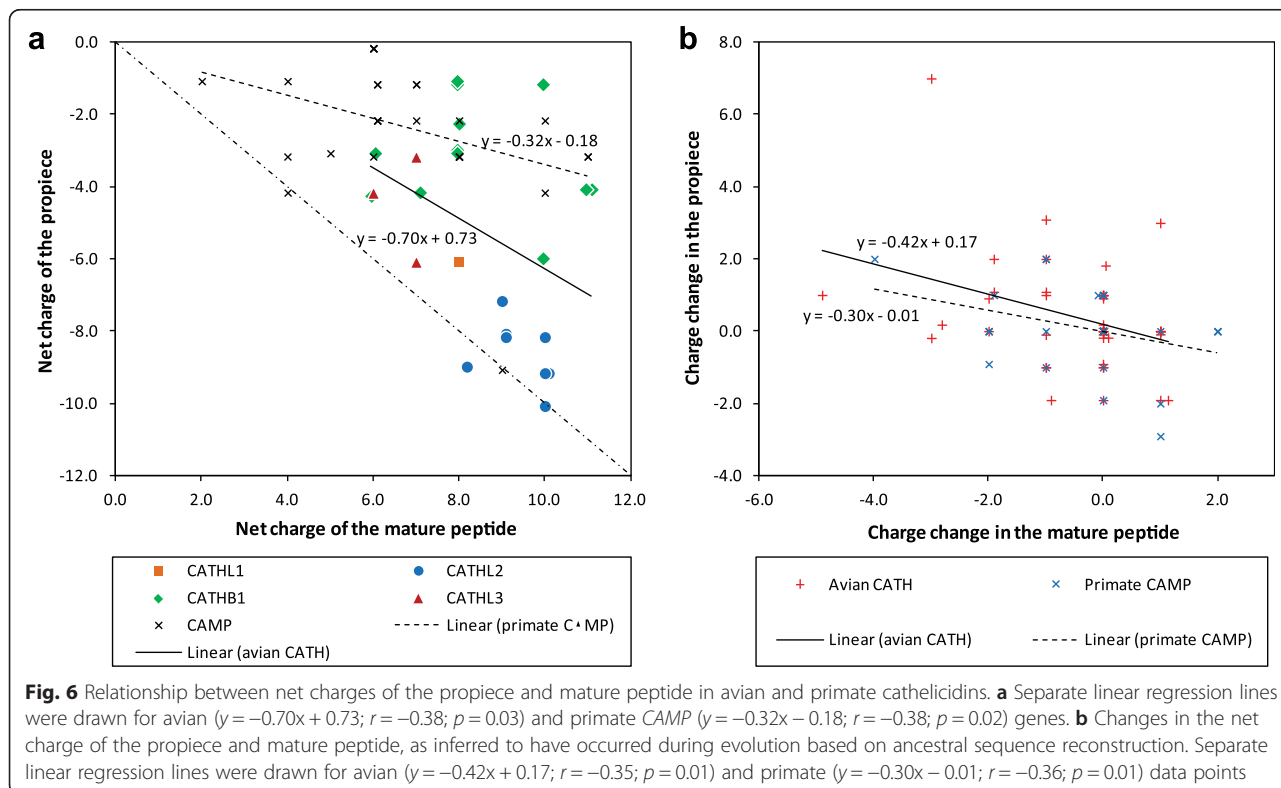
cathelicidins have higher charges than the defensins. Six defensin subfamilies (AvBD5, 8, 9, 10, 11 and 12) have an average net charge lower than +4.0, with AvBD12 showing the lowest average charge (+0.1). AvBD3 and all cathelicidin subfamilies have an average charge higher than +6.0. The low net charge of certain AvBDs indicates that they may have lower activities in terms of direct killing microbes. By contrast, the highly cationic peptides, such as AvBD3 in the emperor penguin, Adélie penguin (*Pygoscelis adeliae*) and yellow-throated sandgrouse (*Pterocles gutturalis*) with a net charge of +10.0, and CATHB1 in the kea (*Nestor notabilis*) with a net charge of +12.0, may provide valuable templates for developing new antimicrobial agents. The mature peptide sequences of all avian  $\beta$ -defensins and cathelicidins are provided in Additional file 4 and net charges shown in Additional file 5.

**Avian cathelicidins and the “charge balance hypothesis”**

In mammalian  $\alpha$ -defensins, Michaelson et al. [43] proposed that the anionic propiece plays a role in preventing autocytotoxicity by neutralising the cationicity of the mature peptide. The authors demonstrated a linear relationship between the net negative charge of the propiece and the positive charge of the mature peptide of seven  $\alpha$ -defensins (two human and five rabbit genes). Hughes and Yeager [4, 44] provided further supportive evidence of this relationship ( $r = -0.742$ ;  $p < 0.001$ ) using 28  $\alpha$ -

defensins from five mammalian species (mouse, rat, guinea pig, rabbit and human). This “charge balance hypothesis” is unlikely to apply to the  $\beta$ -defensins due to the short length of the propiece (0–7 amino acids in AvBDs). However, we explored the hypothesis in avian cathelicidins and revealed a similar, though weaker ( $r = -0.38$ ;  $p = 0.03$ ) association between the electrostatic charges of the propiece and mature peptides (Fig. 6a). All annotated avian cathelicidins with an intact coding sequence were used in the analysis ( $n = 24$ ). While the mature form of all cathelicidins is highly cationic ( $8.51 \pm 0.61$ ), the propiece has an anionic character ( $-5.24 \pm 1.13$ ). The inactive form of CATHL2 before secretion has the lowest mean net charge ( $0.81 \pm 0.62$ ) among avian cathelicidins (overall mean =  $3.27 \pm 1.06$ ), suggesting a better neutralising effect of the propiece in this subfamily. It should be noted that changes in cytoplasmic pH also affect the electrostatic charges of peptides, with higher pH resulting in lower charges (e.g. at pH 7.5, overall mean charge of presecretory avian cathelicidins =  $2.79 \pm 1.05$ ). Similar level of correlation ( $r = -0.38$ ;  $p = 0.02$ ) was also found in 28 primate cathelicidin genes (Fig. 6a; accession numbers of primate genes listed in Additional file 6).

Further analysis was performed on reconstructed ancestral sequences to infer changes that may have occurred in avian and primate cathelicidins over evolutionary time (Fig. 6b, Additional file 7) [44]. A significant



negative correlation was observed between electrostatic property changes in the propiece and mature peptide in both groups (aves:  $r = -0.35$ ,  $p = 0.01$ ; primate:  $r = -0.36$ ,  $p = 0.01$ ). When substitutions increasing peptide charge occurred in the mature cathelicidin, charge in the propiece tended to either decrease or remain unchanged, whereas when the mature peptide became less cationic, the propiece tended to become less anionic. This is highly similar to what Hughes and Yeager observed in mammalian defensins [44].

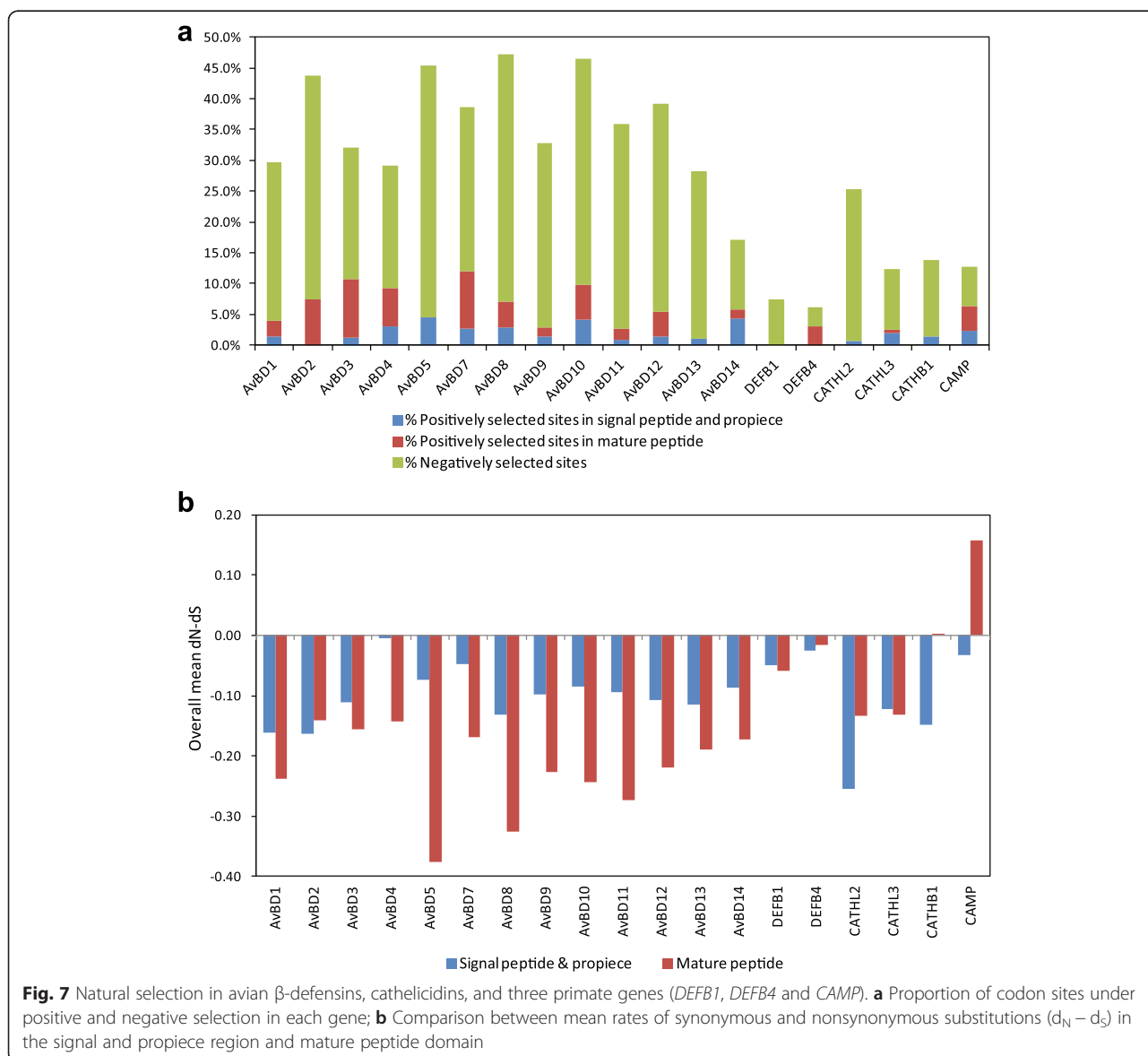
Moreover, evidence of intra-molecular amino acid residue co-evolution was detected between two pairs of sites with electrostatic properties in the propiece and mature peptide domains of CATHB1 (highlighted in Additional file 3). Residue 122L/G/E/K and 226G/K/R, and 160R/Q/P and 222E/D/G/N showed 96.9 and 80.2 % probability of

having been co-evolving, respectively (residue positions are based on chicken CATHB1).

These observations indicate that, similar to mammalian  $\alpha$ -defensins, the propiece and mature peptide of avian and primate cathelicidins may have co-evolved in such a way that amino acid substitutions in both regions are selected and accumulated to balance the charge. Further experimental evidence will be needed to validate this hypothesis and elucidate the role of cathelicidin propiece in preventing autotoxicity.

**Amino acid sites under selection**

Negative selection was detected in a large proportion of amino acid sites in the examined avian  $\beta$ -defensin (11.4–40.9 %) and cathelicidin (9.7–24.7 %) genes (Fig. 7a, Additional file 3). In AvBD1, 2, 3, 8, 9, 11, 12 and 13, the



overall mean rate of nonsynonymous nucleotide substitutions ( $d_N$ ) was significantly (at 0.05 nominal level) lower than that of synonymous substitutions ( $d_S$ ) in both mature peptide and the signal domains, suggesting an overall effect of negative selection on these genes (Fig. 7b, Table 3). In the other five AvBDs significant negative selection was detected only in the mature peptide, whereas the opposite pattern was observed in the avian cathelicidin genes with only the signal and propiece domain showing significantly higher  $d_S$  than  $d_N$ .

Despite the strong background of negative selection, evidence indicating specific amino acid sites subject to episodic diversifying selection was found in most studied avian genes (Fig. 8), though the effect of such selection was very weak in AvBD9, AvBD11 and AvBD13 and the three cathelicidins, with less than 4 % of total sites inferred to be positively selected. To reduce chances of false positive detections, only codon sites that were detected by multiple selection test methods were considered significant (see Methods section) [45, 46]. In the examined avian AMPs, most positively selected sites were found inside the mature peptide domain, except for AvBD5, AvBD13 and AvBD14 and all cathelicidins, in which more were found in the signal and propiece

regions (Figs. 7a and 8). Within mature defensins, diversifying selection appears to mainly (68.3 % cases) affect those residues that are close to (within two residues) the conserved cysteines. For comparison, same analyses were performed on two primate  $\beta$ -defensin subfamilies, DEFB1 ( $n = 24$ ) and DEFB4 ( $n = 10$ ), and the primate cathelicidin CAMP ( $n = 28$ ) (gene accession numbers provided in Additional file 6). Consistent with previous reports, our results suggested that positive selection has involved in driving the evolution of primate DEFB4 (coding for  $\beta$ -defensin 2) [47], whereas DEFB1 is highly conserved with no evidence of diversifying selection in primate lineages [48]. Taken together both positive and negative selection, AvBDs appear to have evolved under higher selective pressures and restraint, while the two examined mammalian  $\beta$ -defensins have evolved more neutrally in primates (Fig. 7a). Contrary to what was observed in avian cathelicidins, in primate CAMP a majority (58.3 %) of positively selected residues are located inside the mature peptide, resulting in a significantly higher overall  $d_N$  than  $d_S$  in the gene domain (Fig. 7).

Due to the important role of electrostatic charge on antimicrobial potency of the mature peptide, positively charged amino acid residues are expected to have

**Table 3** Mean rates of synonymous ( $d_S \pm S.E.$ ) and nonsynonymous ( $d_N \pm S.E.$ ) nucleotide substitutions and test for overall selection in  $\beta$  defensin and cathelicidin genes

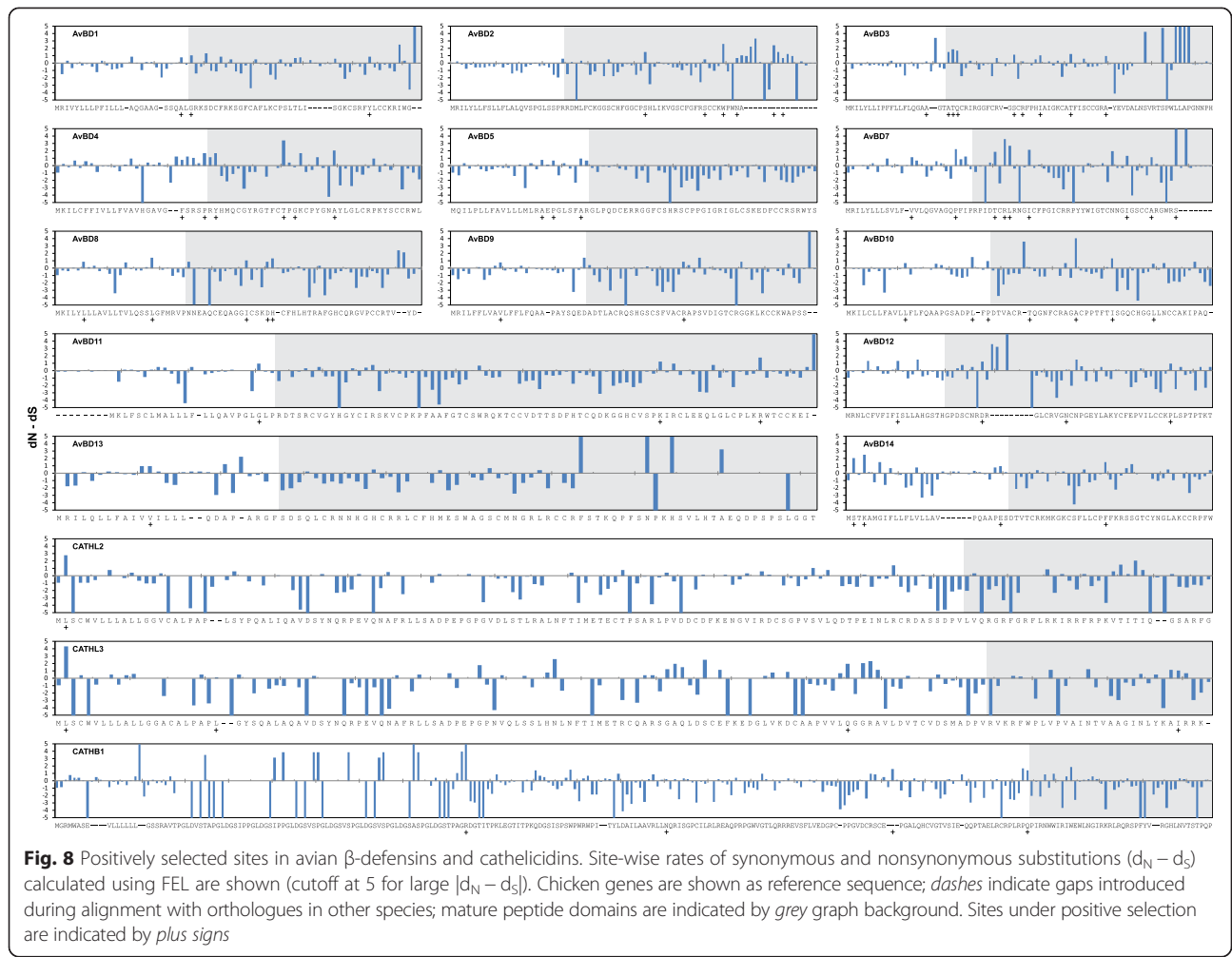
Gene	Signal peptide & propiece					Mature peptide				
	$d_S$	$d_N$	$d_N - d_S$	Stat <sup>b</sup>	P <sup>c</sup>	$d_S$	$d_N$	$d_N - d_S$	Stat <sup>b</sup>	P <sup>c</sup>
AvBD1	0.338 ± 0.061	0.175 ± 0.035	-0.162 ± 0.062	2.623	<b>0.005</b>	0.585 ± 0.073	0.347 ± 0.068	-0.237 ± 0.094	2.494	<b>0.007</b>
AvBD2	0.223 ± 0.033	0.060 ± 0.015	-0.163 ± 0.034	4.450	<b>0.000</b>	0.288 ± 0.032	0.147 ± 0.030	-0.141 ± 0.046	3.102	<b>0.001</b>
AvBD3	0.250 ± 0.052	0.139 ± 0.034	-0.111 ± 0.045	2.495	<b>0.007</b>	0.690 ± 0.048	0.534 ± 0.088	-0.155 ± 0.092	1.699	<b>0.046</b>
AvBD4	0.210 ± 0.039	0.205 ± 0.036	-0.005 ± 0.040	0.117	0.453	0.319 ± 0.061	0.176 ± 0.042	-0.143 ± 0.068	2.183	<b>0.016</b>
AvBD5	0.216 ± 0.045	0.143 ± 0.029	-0.073 ± 0.053	1.401	0.082	0.473 ± 0.055	0.098 ± 0.022	-0.376 ± 0.060	6.476	<b>0.000</b>
AvBD7	0.194 ± 0.049	0.147 ± 0.037	-0.047 ± 0.066	0.708	0.240	0.396 ± 0.057	0.227 ± 0.046	-0.169 ± 0.072	2.377	<b>0.010</b>
AvBD8	0.226 ± 0.045	0.095 ± 0.021	-0.131 ± 0.052	2.529	<b>0.006</b>	0.446 ± 0.043	0.121 ± 0.024	-0.325 ± 0.049	6.715	<b>0.000</b>
AvBD9	0.152 ± 0.038	0.053 ± 0.017	-0.098 ± 0.040	2.552	<b>0.006</b>	0.317 ± 0.061	0.090 ± 0.018	-0.227 ± 0.056	4.025	<b>0.000</b>
AvBD10	0.191 ± 0.044	0.106 ± 0.024	-0.085 ± 0.059	1.461	0.073	0.361 ± 0.051	0.117 ± 0.026	-0.244 ± 0.054	4.612	<b>0.000</b>
AvBD11	0.164 ± 0.054	0.069 ± 0.015	-0.095 ± 0.059	1.670	<b>0.049</b>	0.362 ± 0.037	0.088 ± 0.013	-0.274 ± 0.040	6.450	<b>0.000</b>
AvBD12	0.281 ± 0.048	0.171 ± 0.043	-0.108 ± 0.056	1.895	<b>0.030</b>	0.334 ± 0.053	0.115 ± 0.022	-0.219 ± 0.056	3.752	<b>0.000</b>
AvBD13	0.222 ± 0.050	0.108 ± 0.023	-0.114 ± 0.057	2.135	<b>0.017</b>	0.249 ± 0.044	0.059 ± 0.013	-0.190 ± 0.042	4.547	<b>0.000</b>
AvBD14	0.351 ± 0.071	0.263 ± 0.044	-0.087 ± 0.079	1.099	0.137	0.309 ± 0.054	0.137 ± 0.025	-0.172 ± 0.055	3.029	<b>0.002</b>
DEFB1 <sup>a</sup>	0.072 ± 0.039	0.023 ± 0.008	-0.049 ± 0.040	1.261	0.105	0.097 ± 0.041	0.040 ± 0.020	-0.058 ± 0.039	1.467	0.072
DEFB4 <sup>a</sup>	0.054 ± 0.034	0.029 ± 0.016	-0.025 ± 0.041	0.622	0.268	0.158 ± 0.049	0.143 ± 0.030	-0.016 ± 0.057	0.283	0.389
CATHL2	0.417 ± 0.051	0.162 ± 0.023	-0.255 ± 0.053	4.756	<b>0.000</b>	0.401 ± 0.082	0.267 ± 0.054	-0.134 ± 0.095	1.416	0.080
CATHL3	0.310 ± 0.047	0.187 ± 0.026	-0.123 ± 0.052	2.448	<b>0.008</b>	0.379 ± 0.095	0.247 ± 0.051	-0.132 ± 0.095	1.372	0.086
CATHB1	0.448 ± 0.039	0.300 ± 0.024	-0.148 ± 0.043	3.362	<b>0.001</b>	0.508 ± 0.074	0.508 ± 0.057	0.001 ± 0.090	0.006 <sup>d</sup>	0.497 <sup>d</sup>
CAMP <sup>a</sup>	0.095 ± 0.015	0.063 ± 0.008	-0.032 ± 0.017	1.852	<b>0.033</b>	0.066 ± 0.018	0.216 ± 0.039	0.157 ± 0.041	3.906 <sup>d</sup>	<b>0.000<sup>d</sup></b>

<sup>a</sup> Primate genes

<sup>b</sup> Z-test statistic for purifying selection ( $d_S > d_N$ )

<sup>c</sup> Values of  $P < 0.05$  are considered significant (in bold)

<sup>d</sup> Tested for positive selection ( $d_S < d_N$ )



evolved under positive selection. However, in the examined AvBDs and cathelicidins, a number of such sites showed strong negative selection or neutral results. One possible explanation is that these cationic residues may have been selectively accumulated prior to the divergence of the studied bird lineages (~114 million years ago), and then have been conserved by negative selection through the long evolutionary history due to their significant role on peptide function. Further studies to include more distant taxa (such as reptiles) will help elucidate the evolutionary dynamics of the peptides.

**Lineages subject to episodic diversifying selection**

Evolutionary branches that were indicated to have experienced diversifying selection episodes at each gene are highlighted in Additional file 8. On average, 20.9 and 26.6 % of nodes over the course of evolution of AvBDs and cathelicidins, respectively, were detected with episodes of positive selection, with the highest percentage observed in AvBD7 (41.4 %) and lowest in AvBD13 (7.1 %). The later stages of evolution within avian lineages seem to

have involved relatively stronger diversifying selection as compared to the more ancient branches, which is consistent with the observation that the examined gene families are generally well conserved across all avian orders and families. This result contradicts to what was previously detected in mammalian, particularly primate  $\beta$ -defensins, which involved more positive selection episodes in more ancient branches due to duplication and diversification of  $\beta$ -defensins in the early stages of mammalian evolution [5].

Several evolutionary nodes were detected to have undergone diversifying selection at multiple genes (Additional file 8). For example, AvBD5, 7, 10, 12 and 14 of Galliformes were inferred to be under diversifying selection between 84.5 and 37.9 million years before present (evolutionary time estimated based on branch length in Fig. 1), prior to the divergence between family Phasianidae (chicken and turkey) and Odontophoridae (northern bobwhite *Colinus virginianus*). Within the order Psittaciformes, AvBD3, 4 and 8 were estimated to have been subject to episodic positive selection between 79.8 and 55.3 million years ago before the divergence of the Puerto

Rican parrot (*Amazona vittata*), budgerigar (*Melopsittacus undulatus*) and kea (*Nestor notabilis*), all belonging to the Psittacidae family. Similarly, positive selection appeared to have been involved in driving the diversification of AvBD1, 3, 4, 5, 8, 10 and 12 during the early radiations of Passeriformes. For each species, the number of genes indicated to have been affected by episodic diversifying selection is shown in Fig. 1.

## Conclusion

In this study, we investigated the evolution of  $\beta$ -defensins and cathelicidins in 53 bird species. Both gene families form a generally conserved gene cluster in avian genomes with certain genes being more prone to duplication (AvBD1, AvBD3 and AvBD7) or pseudogenisation (AvBD14) events. Intense negative selection was detected in a majority of examined gene domains, likely accounting for the conservation of certain amino acid residues that are essential for the functioning of  $\beta$ -defensins and cathelicidins in birds. Evidence indicated that episodic positive selection also played a role in driving the diversification of specific residues of certain antimicrobial peptides in avian evolutionary history, contributing to high variability of gene sequences and electrostatic property of the peptides. Our results also revealed that selection may have acted on cathelicidins to maintain a balanced charge between the anionic propeptide and cationic mature peptide over evolutionary time. This work not only has greatly improved our understanding of the molecular evolution of these host defense peptides, but also provides a valuable resource for potential translational research and development of novel antimicrobial agents.

## Methods

### Database search and gene nomenclature

Fifty-three bird genomes were searched for  $\beta$ -defensin and cathelicidin genes (GenBank Assembly IDs provided in Table 1). For each genome, four steps were taken to identify genes of interest: 1) An initial search was performed on a predicted protein/CDS database with BLAST programs using chicken genes as query sequences. Hits with  $E$ -value  $<0.1$  were extracted from the database, aligned to chicken sequences with ClustalW [49], and manually examined to exclude false positives. 2) Then for each gene, a profile hidden Markov model (HMM) [50] was built from a peptide sequence alignment that includes all orthologues found in the previous step. The profile HMMs were searched against the predicted protein databases with HMMER3.1 programs [51] on a GALAXY platform. Hits with both  $E$ -values (full sequence and best 1 domain)  $<1$  were extracted, aligned with previously found sequences, and manually checked to confirm real homologues. 3) Sequences identified in the previous steps were then used to BLAST search the whole genome to find any genes, gene

fragments, or pseudogenes that are not included in the protein/CDS database. 4) All genomic scaffolds and contigs containing  $\beta$ -defensin or cathelicidin genes were extracted to study the genomic organisation of these genes. In addition, sequences containing *CTSB*, *TRAM2*, *KLHL18* or *TBRG4*, the flanking genes of the  $\beta$ -defensin or cathelicidin gene clusters, were also extracted. Scaffolds and contigs were manually curated using Artemis [52].

Annotated sequences were named by tagging the gene name with a five-letter abbreviation as a suffix that distinguishes the species. For example, AvBD1\_ACACH refers to gene *AvBD1* of the rifleman (*Acanthisitta chloris*). Duplications of AvBD1 and AvBD3 were numbered 1.n and 3.n from AvBD2. Some AvBD1 and AvBD3 duplicates were found on isolated scaffolds and numbering of these defensins may not represent the actual position in the intact cluster. Previously used identification references for zebra finch duplicates [9] are included in the identification.

### Mature peptide prediction and net charge estimation

Sequence features, such as signal peptide, propeptide and mature peptide, within annotated genes were speculated based on functional domains in chicken defensins [10] and cathelicidins [30, 31, 33, 34]. The net charge of mature peptide was estimated as  $\sum_i N_i \frac{1}{1+10^{pH-pK_{a_i}}} - \sum_j N_j \frac{1}{1+10^{pK_{a_j}-pH}}$  with  $N_i$  and  $pK_{a_i}$  being the number and pKa values of histidine (H), lysine (K) and arginine (R) residues and the N-terminus, and  $N_j$  and  $pK_{a_j}$  the number and pKa of aspartic acid (D), glutamic acid (E), cysteine (C) and tyrosine (Y) residues and the C-terminus [53]. Lehninger's set of pKa values were used [54] and intramolecular disulfide bond formation was taken into account in calculation.

### Evolutionary analyses

Overall phylogenetic analyses between all avian  $\beta$ -defensins and cathelicidins (Fig. 2) were conducted in MEGA5 with a Maximum Likelihood method [55, 56] using four discrete categories for the Gamma distribution to model evolutionary rate differences among sites and 100 bootstrap replicates to infer the level of confidence on the phylogeny (support values lower than 40 % are not shown in the consensus tree) [57].

A range of evolutionary analyses were performed on each  $\beta$ -defensin and cathelicidin gene via the Datamonkey webserver [58], including: 1) Negative selection sites were detected using Fixed Effect Likelihood (FEL) [59] and Fast Unconstrained Bayesian Approximation for Inferring Selection (FUBAR) [60]. 2) Individual sites under positive selection were detected using three test methods, including Mixed Effects Model of Evolution (MEME) [61], FEL, and FUBAR. Codon sites found to be significant for positive or

negative selection by more than two methods (MEME  $p < 0.05$ , FEL  $p < 0.1$ , and FUBAR posterior probability  $> 0.9$ ) were included in the analyses [45, 46]. 3) Individual branches with episodic diversifying selection were inferred by combining results from two analyses—MEME (empirical bayes factor  $> 20$ ) and branch-site REL ( $p < 0.05$ ) [62]. 4) Intramolecular co-evolution of amino acid sites in cathelicidins were detected using the Spidermonkey/Bayesian Graphical Model [63]. Only sites that are involved in electrostatic properties and have more than three branches with nonsynonymous substitutions were included in the analysis for covariation. Assessment of overall mean rates of nonsynonymous ( $d_N$ ) and synonymous ( $d_S$ ) nucleotide substitutions in the signal and propeptide region and the mature peptide domain, and the significance test of overall selection were calculated in MEGA5 [55] using the Kumar model [64] and 1000 bootstrap replicates to estimate standard errors.

Ancestral cathelicidin sequences were inferred using the Maximum Likelihood method [64] under the Whelan And Goldman model [65] in MEGA5. Rates among sites were treated as a Gamma distribution using five Gamma categories.

### Ethics

No human, human data, or animal was used in this study.

### Availability of supporting data

GenBank assembly IDs of all 53 examined avian genomes are provided in Table 1. Scaffold and coordinates information of annotated genes are detailed in Additional file 1. Nucleotide sequences of all annotated genes are available in Additional file 2. Full-length amino acid sequence alignments and putative mature peptide sequence of each gene are provided in Additional files 3 and 4, respectively. GenBank accession numbers of primate genes used in analyses are listed in Additional file 6.

### Additional files

**Additional file 1: List of annotated  $\beta$ -defensin and cathelicidin genes in 53 avian species with scaffold and coordinates information.** (XLSX 47 kb)

**Additional file 2: Nucleotide sequences of annotated avian  $\beta$ -defensin and cathelicidin genes (in fasta format).** (FAS 204 kb)

**Additional file 3: Sequence alignments of avian  $\beta$ -defensins and cathelicidins.** The ruler is based on residue positions of chicken sequences. Dots represent identical residues in comparison with the chicken sequence in the top; dashes are gaps or missing data. Functionally significant conserved residues are boxed. Plus and minus signs on top of the alignment indicate sites under positive and negative selection, respectively. Letters 'a' and 'b' above the CATHB1 alignment indicate two pairs of amino acid residues with evidence of intramolecular co-evolution. A sequence logo was generated at the bottom of the alignment to show the frequency of certain amino acid

substitution at a position, indicated by the height of the letter [66]. The labelled alignments and sequence logo were produced with TEXshade [67]. (PDF 1538 kb)

**Additional file 4: Putative mature peptide sequences from annotated avian  $\beta$ -defensin and cathelicidin genes (in fasta format).** (FAS 42 kb)

**Additional file 5: Net charge of putative avian  $\beta$ -defensin and cathelicidin mature peptides.** (XLSX 30 kb)

**Additional file 6: List of primate genes used in analyses with Genbank accession numbers.** (XLSX 11 kb)

**Additional file 7: Reconstructed avian and primate ancestral cathelicidin sequences and inferred charge changes between nodes.** (PDF 299 kb)

**Additional file 8: Phylogenetic trees of avian  $\beta$ -defensin and cathelicidin genes.** Branches with evidence suggesting episodic diversifying selection are highlighted in red. (PDF 1444 kb)

### Abbreviations

AMP: Antimicrobial peptide; AvBD: Avian  $\beta$ -defensin; Abbreviations for species names are listed in Table 1.

### Competing interests

The authors declare that they have no competing interests.

### Authors' contributions

DWB and KB had the idea, guided and supervised analyses and revised drafts. YC and MDP both carried out gene annotation, data analyses and draft writing. YC also prepared figures, Table 3 and additional files. MDP also made Table 1 and 2 and proof checked the figures. WG generated protein/CDS sequences from the genomes and prepared databases for BLAST and HMMER programs. RK set up the GALAXY platform, built HMM sequence models, and optimised workflows for database searches.

### Acknowledgements

DWB and WG are supported by Biotechnology and Biological Science Research Council (BBSRC), UK. RK is supported by the Wellcome Trust, UK. YC and KB are supported by the Australian Research Council (ARC).

### Author details

<sup>1</sup>RMC Gunn Building B19, Faculty of Veterinary Science, University of Sydney, Camperdown 2006 NSW, Australia. <sup>2</sup>Dipartimento di Scienze della Vita-Edif. C11, Università di Trieste, Via Licio Giorgieri 1, 34127 Trieste, Italy. <sup>3</sup>The Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, Easter Bush, Midlothian EH25 9RG, UK.

Received: 27 April 2015 Accepted: 21 August 2015

Published online: 15 September 2015

### References

- Cuperus T, Coorens M, van Dijk A, Haagsman HP. Avian host defense peptides. *Dev Comp Immunol*. 2013;41:352–69.
- van Dijk A, Molhoek EM, Bikker FJ, Yu PL, Veldhuizen EJA, Haagsman HP. Avian cathelicidins: paradigms for the development of anti-infectives. *Vet Microbiol*. 2011;153:27–36.
- Sample F, Dorin JR.  $\beta$ -defensins: multifunctional modulators of infection, inflammation and more? *Journal of Innate Immunity*. 2012;4:337–48.
- Hughes AL. Evolutionary diversification of the mammalian defensins. *CMLS, Cell Mol Life Sci*. 1999;56:94–103.
- Sample CAM, Taylor K, Eastwood H, Barran PE, Dorin JR. b-defensin evolution: selection complexity and clues for residues of functional importance. *Biochem Soc Trans*. 2006;34:257–62.
- Jarvis ED, Mirarab S, Aberer AJ, Li B, Houde P, Li C, et al. Whole-genome analyses resolve early branches in the tree of life of modern birds. *Science*. 2014;346:1320–31.
- Higgs R, Lynn D, Cahalane S, Alaña I, Hewage C, James T, et al. Modification of chicken avian  $\beta$ -defensin-8 at positively selected amino acid sites enhances specific antimicrobial activity. *Immunogenetics*. 2007;59:573–80.

8. van der Does AM, Bergman P, Agerberth B, Lindbom L. Induction of the human cathelicidin LL-37 as a novel treatment against bacterial infections. *J Leukoc Biol.* 2012;92:735–42.
9. Hellgren O, Ekblom R. Evolution of a cluster of innate immune genes (beta-defensins) along the ancestral lines of chicken and zebra finch. *Immunome Res.* 2010;6:1745–7580.
10. Xiao Y, Hughes A, Ando J, Matsuda Y, Cheng J-F, Skinner-Noble D, et al. A genome-wide screen identifies a single beta-defensin gene cluster in the chicken: implications for the origin and evolution of mammalian defensins. *BMC Genomics.* 2004;5:56.
11. Huang Y, Li Y, Burt DW, Chen H, Zhang Y, Qian W, et al. The duck genome and transcriptome provide insight into an avian influenza virus reservoir species. *Nat Genet.* 2013;45:776–83.
12. Lynn DJ, Higgs R, Lloyd AT, O'Farrelly C, Hervé-Grépinet V, Nys Y, et al. Avian beta-defensin nomenclature: a community proposed update. *Immunol Lett.* 2007;110:86–9.
13. Lan H, Chen H, Chen L-C, Wang B-B, Sun L, Ma M-Y, et al. The first report of a pelecaniformes defensin cluster: characterization of b-defensin genes in the crested ibis based on BAC libraries. *Scientific Reports.* 2014;4.
14. van Dijk A, Veldhuizen EJA, Haagsman HP. Avian defensins. *Vet Immunol Immunopathol.* 2008;124:1–18.
15. Zhang G, Sunkara L. Avian antimicrobial host defense peptides: from biology to therapeutic applications. *Pharmaceuticals.* 2014;7:220–47.
16. Akbari MR, Haghighi HR, Chambers JR, Brisbin J, Read LR, Sharif S. Expression of antimicrobial peptides in cecal tonsils of chickens treated with probiotics and infected with salmonella enterica serovar typhimurium. *Clin Vaccine Immunol.* 2008;15:1689–93.
17. Chranova M, Hradecka H, Faldynova M, Matulova M, Havlickova H, Sisak F, et al. Immune response of chicken gut to natural colonization by gut microflora and to salmonella enterica serovar enteritidis infection. *Infect Immun.* 2011;79:2755–63.
18. Derache C, Esnault E, Bonsergent C, Le Vern Y, Quéré P, Lalmanach A-C. Differential modulation of  $\beta$ -defensin gene expression by Salmonella Enteritidis in intestinal epithelial cells from resistant and susceptible chicken inbred lines. *Dev Comp Immunol.* 2009;33:959–66.
19. Hong YH, Song W, Lee SH, Lillehoj HS. Differential gene expression profiles of  $\beta$ -defensins in the crop, intestine, and spleen using a necrotic enteritis model in 2 commercial broiler chicken lines. *Poult Sci.* 2012;91:1081–8.
20. Milona P, Townes CL, Bevan RM, Hall J. The chicken host peptides, gallinacins 4, 7, and 9 have antimicrobial activity against Salmonella serovars. *Biochem Biophys Res Commun.* 2007;356:169–74.
21. van Dijk A, Veldhuizen EJA, Kalkhove SIC, Tjeerdsma-van Bokhoven JLM, Romijn RA, Haagsman HP. The  $\beta$ -defensin gallinacin-6 is expressed in the chicken digestive tract and has antimicrobial activity against food-borne pathogens. *Antimicrob Agents Chemother.* 2007;51:912–22.
22. Abdel Mageed AM, Isobe N, Yoshimura Y. Immunolocalization of avian  $\beta$ -defensins in the hen oviduct and their changes in the uterus during eggshell formation. *Reproduction.* 2009;138:971–8.
23. Das SC, Isobe N, Yoshimura Y. Expression of toll-like receptors and avian  $\beta$ -defensins and their changes in response to bacterial components in chicken sperm. *Poult Sci.* 2011;90:417–25.
24. Hervé-Grépinet V, Réhault-Godbert S, Labas V, Magallon T, Derache C, Lavergne M, et al. Purification and characterization of avian  $\beta$ -defensin 11, an antimicrobial peptide of the hen egg. *Antimicrob Agents Chemother.* 2010;54:4401–9.
25. Michailidis G, Avdi M, Argiriou A. Transcriptional profiling of antimicrobial peptides avian  $\beta$ -defensins in the chicken ovary during sexual maturation and in response to Salmonella enteritidis infection. *Res Vet Sci.* 2012;92:60–5.
26. Yoshimura Y, Ohashi H, Subedi K, Nishibori M, Isobe N. Effects of age, egg-laying activity, and Salmonella-inoculation on the expressions of gallinacin mRNA in the vagina of the hen oviduct. *J Reprod Dev.* 2006;52:211–8.
27. Hellgren O. Allelic variation at innate immune genes (avian  $\beta$ -defensins), within a natural population of great tits. *J Avian Biol.* 2015;46:113–8.
28. Hellgren O, Sheldon BC. Locus-specific protocol for nine different innate immune genes (antimicrobial peptides:  $\beta$ -defensins) across passerine bird species reveals within-species coding variation and a case of trans-species polymorphisms. *Mol Ecol Resour.* 2011;11:686–92.
29. Bhunia A, Mohanram H, Bhattacharjya S. Lipopolysaccharide bound structures of the active fragments of fowlicidin-1, a cathelicidin family of antimicrobial and antiendotoxic peptide from chicken, determined by transferred nuclear overhauser effect spectroscopy. *Pept Sci.* 2009;92:9–22.
30. Goitsuka R, Chen C-LH, Benyon L, Asano Y, Kitamura D, Cooper MD. Chicken cathelicidin-B1, an antimicrobial guardian at the mucosal M cell gateway. *Proc Natl Acad Sci.* 2007;104:15063–8.
31. van Dijk A, Molhoek EM, Veldhuizen EJA, Bokhoven JLMT-V, Wagendorp E, Bikker F, et al. Identification of chicken cathelicidin-2 core elements involved in antibacterial and immunomodulatory activities. *Mol Immunol.* 2009;46:2465–73.
32. Zanetti M. The role of cathelicidins in the innate host defenses of mammals. *Current Issues in Molecular Biology.* 2005;7:179–96.
33. Bommineni YR, Dai H, Gong Y-X, Soulages JL, Fernando SC, DeSilva U, et al. Fowlicidin-3 is an  $\alpha$ -helical cationic host defense peptide with potent antibacterial and lipopolysaccharide-neutralizing activities. *FEBS J.* 2007;274:418–28.
34. Xiao Y, Dai H, Bommineni YR, Soulages JL, Gong Y-X, Prakash O, et al. Structure–activity relationships of fowlicidin-1, a cathelicidin antimicrobial peptide in chicken. *FEBS J.* 2006;273:2581–93.
35. Xiao Y, Herrera AI, Bommineni YR, Soulages JL, Prakash O, Zhang G. The central kink region of fowlicidin-2, an  $\alpha$ -helical host defense peptide, is critically involved in bacterial killing and endotoxin neutralization. *Journal of Innate Immunity.* 2009;1:268–80.
36. Achanta M, Sunkara L, Dai G, Bommineni Y, Jiang W, Zhang G. Tissue expression and developmental regulation of chicken cathelicidin antimicrobial peptides. *Journal of Animal Science and Biotechnology.* 2012;3:15.
37. Xiao Y, Cai Y, Bommineni YR, Fernando SC, Prakash O, Gilliland SE, et al. Identification and functional characterization of three chicken cathelicidins with potent antimicrobial activity. *J Biol Chem.* 2006;281:2858–67.
38. Jetz W, Thomas GH, Joy JB, Hartmann K, Mooers AO. The global diversity of birds in space and time. *Nature.* 2012;491:444–8.
39. Trowsdale J, Parham P. Mini-review: defense strategies and immunity-related genes. *Eur J Immunol.* 2004;34:7–17.
40. Rached LA, McDermott MF, Pontarotti P. The MHC big bang. *Immunol Rev.* 1999;167:33–45.
41. Shai Y. Mode of action of membrane active antimicrobial peptides. *Biopolymers.* 2002;66:236–48.
42. Klüver E, Schulz-Maronde S, Scheid S, Meyer B, Forssmann W-G, Adermann K. Structure–activity relation of human  $\beta$ -defensin 3: influence of disulfide bonds and cysteine substitution on antimicrobial activity and cytotoxicity. *Biochemistry.* 2005;44:9804–16.
43. Michaelson D, Rayner J, Couto M, Ganz T. Cationic defensins arise from charge-neutralized propeptides: a mechanism for avoiding leukocyte autotoxicity? *J Leukoc Biol.* 1992;51:634–9.
44. Hughes AL, Yeager M. Coordinated amino acid changes in the evolution of mammalian defensins. *J Mol Evol.* 1997;44:675–82.
45. Grueber CE, Wallis GP, Jamieson IG. Episodic positive selection in the evolution of avian toll-like receptor innate immunity genes. *PLoS One.* 2014;9, e89632.
46. Azarian T, Lo Presti A, Giovanetti M, Cella E, Rife B, Lai A, et al. Impact of spatial dispersion, evolution, and selection on Ebola Zaire Virus epidemic waves. *Scientific Reports.* 2015;5:10170.
47. Boniotti M, Tossi A, DelPero M, Sgubin S, Antcheva N, Santon D, et al. Evolution of the beta defensin 2 gene in primates. *Genes and Immunity.* 2003;4:251–7.
48. Del Pero M, Boniotti M, Zuccon D, Cervella P, Spanò A, Amoroso A, et al.  $\beta$ -Defensin 1 gene variability among non-human primates. *Immunogenetics.* 2002;53:907–13.
49. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 1994;22:4673–80.
50. Eddy SR. Profile hidden Markov models. *Bioinformatics.* 1998;14:755–63.
51. Eddy SR. Accelerated profile HMM searches. *PLoS Comput Biol.* 2011;7, e1002195.
52. Carver T, Berriman M, Tivey A, Patel C, Böhme U, Barrell BG, et al. Artemis and ACT: viewing, annotating and comparing sequences stored in a relational database. *Bioinformatics.* 2008;24:2672–6.
53. Kozłowski LP. Calculation of protein isoelectric point. <http://isoelectric.ovh.org/index.html>. Accessed March 2015.
54. Lehninger AL, Nelson DL, Cox MM. Lehninger principles of biochemistry. 6th ed. New York: W.H. Freeman; 2013.
55. Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood,



- evolutionary distance, and maximum parsimony methods. *Mol Biol Evol.* 2011;28:2731–9.
56. Hasegawa M, Fujiwara M. Relative efficiencies of the maximum likelihood, maximum parsimony, and neighbor-joining methods for estimating protein phylogeny. *Mol Phylogenet Evol.* 1993;2:1–5.
  57. Felsenstein J. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution.* 1985;39:783–91.
  58. Delpont W, Poon AFY, Frost SDW, Kosakovsky Pond SL. Datamonkey 2010: a suite of phylogenetic analysis tools for evolutionary biology. *Bioinformatics.* 2010;26:2455–7.
  59. Kosakovsky Pond SL, Frost SDW. Not So different after All: a comparison of methods for detecting amino acid sites under selection. *Mol Biol Evol.* 2005;22:1208–22.
  60. Murrell B, Moola S, Mabona A, Weighill T, Sheward D, Kosakovsky Pond SL, et al. FUBAR: a fast, unconstrained bayesian approximation for inferring selection. *Mol Biol Evol.* 2013;30:1196–205.
  61. Murrell B, Wertheim JO, Moola S, Weighill T, Scheffler K, Kosakovsky Pond SL. Detecting individual sites subject to episodic diversifying selection. *PLoS Genet.* 2012;8, e1002764.
  62. Kosakovsky Pond SL, Murrell B, Fourment M, Frost SDW, Delpont W, Scheffler K. A random effects branch-site model for detecting episodic diversifying selection. *Mol Biol Evol.* 2011.
  63. Poon AFY, Lewis FI, Pond SLK, Frost SDW. An evolutionary-network model reveals stratified interactions in the V3 loop of the HIV-1 envelope. *PLoS Comput Biol.* 2007;3, e231.
  64. Nei M, Kumar S. *Molecular evolution and phylogenetics.* New York: Oxford University Press; 2000.
  65. Whelan S, Goldman N. A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Mol Biol Evol.* 2001;18:691–9.
  66. Schneider TD, Stephens RM. Sequence logos: a new way to display consensus. *Nucleic Acids Res.* 1990;18:6097–100.
  67. Beitz E. TEXshade: shading and labeling multiple sequence alignments using LATEX2e. *Bioinformatics.* 2000;16:135–9.

**Submit your next manuscript to BioMed Central  
and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

