



OPEN

SUBJECT AREAS:

COMPARATIVE  
GENOMICS

INNATE IMMUNITY

Received  
10 July 2014Accepted  
16 October 2014Published  
5 November 2014

Correspondence and  
requests for materials  
should be addressed to  
Q.-H.W.  
(qiu hongwan@zju.  
edu.cn)

\* These authors  
contributed equally to  
this work.

# The first report of a Pelecaniformes defensin cluster: Characterization of $\beta$ -defensin genes in the crested ibis based on BAC libraries

Hong Lan\*, Hui Chen\*, Li-Cheng Chen\*, Bei-Bing Wang, Li Sun, Mei-Ying Ma, Sheng-Guo Fang & Qiu-Hong Wan

The Key Laboratory of Conservation Biology for Endangered Wildlife of the Ministry of Education, State Conservation Centre for Gene Resources of Endangered Wildlife, College of Life Sciences, Zhejiang University, Hangzhou 310058, P. R. China.

Defensins play a key role in the innate immunity of various organisms. Detailed genomic studies of the defensin cluster have only been reported in a limited number of birds. Herein, we present the first characterization of defensins in a Pelecaniformes species, the crested ibis (*Nipponia nippon*), which is one of the most endangered birds in the world. We constructed bacterial artificial chromosome libraries, including a 4D-PCR library and a reverse-4D library, which provide at least 40 equivalents of this rare bird's genome. A cluster including 14  $\beta$ -defensin loci within 129 kb was assigned to chromosome 3 by FISH, and one gene duplication of *AvBD1* was found. The ibis defensin genes are characterized by multiform gene organization ranging from two to four exons through extensive exon fusion. Splicing signal variations and alternative splice variants were also found. Comparative analysis of four bird species identified one common and multiple species-specific duplications, which might be associated with high GC content. Evolutionary analysis revealed birth-and-death mode and purifying selection for avian defensin evolution, resulting in different defensin gene numbers among bird species and functional conservation within orthologous genes, respectively. Additionally, we propose various directions for further research on genetic conservation in the crested ibis.

The crested ibis (*Nipponia nippon*) is one of the most endangered avian species in the world. It has suffered a catastrophic decline since the late nineteenth century and was thought to be extinct until 7 wild birds were rediscovered in China in 1981<sup>1</sup>. Since then, this tiny population has stably expanded under great conservation efforts and is currently the only wild population in the world<sup>1</sup>. However, several captive populations were established to help its recovery. By 2013, the crested ibis population in China increased to more than 1500<sup>2</sup>. However, as a bottlenecked species with intensive inbreeding, recent populations suffer from low genetic diversity<sup>3</sup>. The mortality of ibis embryos and nestlings is also high<sup>1,4,5</sup>, and many pathogens can fatally infect the bird<sup>2,6-8</sup>. Therefore, the genetic conservation and immunological function of the crested ibis require rigorous investigation.

Construction of a genomic library for the crested ibis can conserve its DNA resources and thereby promote detailed genomic research on this rare species and development of strategies for better protection. Bacterial artificial chromosomes (BACs) provide a capacity to clone genomic fragments larger than 100 kb<sup>9</sup>, which is sufficient for most research on a genomic scale. Moreover, BACs can be stably transformed into bacterial hosts with low occurrence of chimerism and are more easily purified than bacteriophage  $\lambda$  cosmid vectors or yeast artificial chromosomes<sup>10</sup>. Given these advantages, BACs have been widely used in the construction of DNA libraries and in genomic studies<sup>10</sup>.

Defensins are a collection of small peptides (commonly less than 100 amino acids) enriched in hydrophobic and cationic residues<sup>11,12</sup> that play a critical role in regulating the innate immune systems<sup>13</sup> of a wide range of organisms including fungi, plants, invertebrates, and vertebrates<sup>12,14-16</sup>. These antimicrobial peptides can protect the host against a broad spectrum of pathogens, such as bacteria, enveloped viruses, certain fungi, and parasitic protozoa<sup>13,17</sup>, mainly by binding to and disturbing pathogen membranes, blocking viral replication cycles, or changing host cell recognition sites<sup>13,18</sup>. Moreover, defensins are chemotactic for T cells, macrophages, mast cells, and immature dendritic cells<sup>19</sup>, thereby serving as a bridge between innate and adaptive immunity.



Defensins contain 6–8 conserved cysteine residues forming 3–4 intramolecular disulfide bonds. Based on the positions of the cysteine residues and the pairing of the disulfide bridge, they are categorized into five groups: plant (C1–C8, C2–C5, C3–C6, and C4–C7), invertebrate (C1–C4, C2–C5, and C3–C6),  $\alpha$ -defensins (C1–C6, C2–C4, and C3–C5),  $\beta$ -defensins (C1–C5, C2–C4, and C3–C6), and  $\theta$ -defensins (C1–C6, C2–C5, and C3–C4), with latter three unique to vertebrates<sup>15,16,20,21</sup>.  $\beta$ -defensins are found in a majority of vertebrates<sup>12,16,22,23</sup>, while  $\alpha$ -defensins are found only in mammals, including marsupials, but are absent in some Laurasiatheria species, such as dogs and cattle<sup>24,25</sup>. The cyclic  $\theta$ -defensins, derived from the posttranslational ligation of two truncated  $\alpha$ -defensin-like precursor peptides<sup>21</sup>, have only been found in Old World monkeys and some apes and are defunct in gorillas, chimpanzees, and humans<sup>26</sup>. Based on their prevalence, the major  $\beta$ -defensins subfamilies are believed to have arisen before the bird–mammal split, followed by the appearance of  $\alpha$ -defensins after the divergence of mammals and finally  $\theta$ -defensins in the primate lineage<sup>26,27</sup>. In the human genome, at least 6  $\alpha$ -defensins<sup>13</sup> and 40 potential  $\beta$ -defensins were found within 5 clusters scattered on chromosome 8, 6, and 20, and the number of  $\beta$ -defensins may be much higher due to rapid duplication<sup>28</sup>. Typically, the human  $\beta$ -defensins contain two exons: exon 1 corresponds to the 5' untranslated region (UTR) and signal peptide, and exon 2 corresponds to the short anionic propiece followed by the mature peptide and 3'UTR<sup>28</sup>. Post-translational modifications include proteolytic cleavage of the signal peptide and propiece from the prepeptide to yield the mature peptide<sup>28,29</sup>.

The first avian defensin genes were isolated from chicken and turkey leukocytes in the mid 1990s<sup>30</sup>. Since then, defensins have been increasingly studied in other avian species<sup>31,32</sup>, but comprehensive genomic information on defensin clusters has only been reported in chicken (14 genes)<sup>27,33</sup>, duck (19 genes including a pseudogene)<sup>34</sup>, and zebra finch (22 genes)<sup>35</sup>, all suggesting that only  $\beta$ -defensins are present in birds and revealing various gene duplications. Phylogenetic analysis based on the three birds revealed that the avian  $\beta$ -defensins (AvBDs) could be classified into 12 subfamilies<sup>34</sup>; nine subfamilies (*AvBD2*, *AvBD4–5*, and *AvBD8–13*) maintain one-to-one orthologous genes in all species; two subfamilies including three

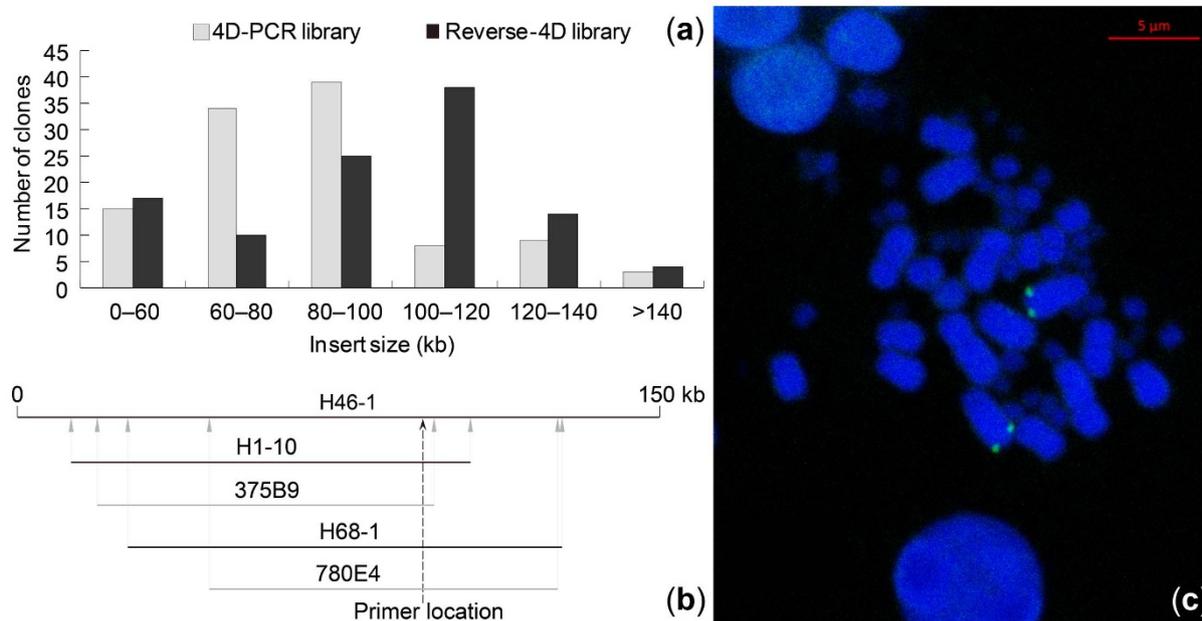
members (*AvBD1/AvBD3* and *AvBD7*) show different lineage-specific duplications among species; and one subfamily (*AvBD14*) is absent in zebra finch. Accordingly, most avian defensins evolved before the bird species split from each other. In contrast to the human  $\beta$ -defensins, the avian  $\beta$ -defensins typically contain four exons: exon 1 for the first part of the 5' UTR, exon 2 for the rest of the 5' UTR and the signal peptide, exon 3 for the short propiece (sometimes absent) and majority of the mature peptide, and exon 4 for the C-terminus of the mature peptide and the 3'UTR<sup>12,27</sup>.

In this work, we constructed genomic libraries for the crested ibis using BACs, including a routine 4D-PCR library and a simplified reverse-4D library. Based on the target BAC clones screened from the libraries, we characterized the content and organization of the defensin gene cluster, which is the first to be reported in Pelecaniformes. The genomic and evolutionary information about defensins revealed by this work will increase our understanding of the immunity of this endangered species and the historic events in avian defensin gene evolution.

## Results

**Construction and characterization of BAC libraries.** Since the crested ibis is an endangered species, it required a routinely constructed BAC genomic DNA library, a 4D-PCR BAC library<sup>36</sup>, to store its genetic resource. The library contains 129312 clones conserved on 1347 96-well plates and was arrayed into 27 superpools (49 × 96 clones). The average insert size of the library was 86.5 kb, with an empty vector rate of 6.5% (Figure 1a). Through comparison with the C-value of other threskiornithidae species available in the Animal Genome Size Database (<http://www.genomesize.com/>), we estimated the crested ibis genome at about 1.35 Gb. Thus, the genomic coverage of this library was about 7.8 equivalents of the crested ibis genome.

Considering the relatively small insert size in the 4D-PCR library, we decided to construct a second BAC library with greater coverage. However, 4D-PCR BAC library construction is time-consuming and labor-intensive. Recently, a new type of simply arrayed BAC library, a reverse-4D BAC library, was developed for birds<sup>37</sup>; this method can significantly reduce the time investment and is efficient enough to



**Figure 1 | Quality and usability of BAC libraries.** (a) Distribution of insert sizes for randomly selected BAC library clones. (b) BAC overlapping cluster of defensin genes. H46-1, H1-10, 375B9, H68-1, and 780E4 correspond to the five positive BACs. Grey arrows: end sequences of BACs; black arrow: location of primers (*AvBD7*) used to screen BAC libraries. (c) Chromosome localization of the crested ibis  $\beta$ -defensin cluster. The H46-1 probe was mapped to crested ibis chromosome 3 (green signals; numbering described in Methods). Chromosome scales are at the upper right corner.



acquire positive clones covering the target genomic region. Therefore, we adjusted library construction conditions (Table 1) and built a reverse-4D BAC library for the crested ibis consisting of 1300 sub-libraries, each with an average of 400 clones. Randomly selected clones had average insert sizes of 100.9 kb and an empty vector rate of 5.6% (Figure 1a). The reverse-4D library therefore provides at least 35-fold coverage of the crested ibis genome.

**Validation of the usability of the BAC libraries.** We obtained two pieces of evidence to verify the usability of the crested ibis BAC libraries. First, we employed the primer pair from defensin gene *AvBD7* to screen the BAC libraries and achieved five positive BACs (Figure 1b), of which 375B9 and 780E4 are from the routine 4D-PCR library while the other three (H46-1, H1-10, and H68-1) are from the simplified reverse-4D library. The overlap among BACs shows that the big BAC clones are all from the reverse-4D library, validating large inserts in the reverse-4D library. The biggest BAC, H46-1, contains a 150 kb insert spanning the target genomic region and covers all defensin genes of the crested ibis as determined from a BLAST search against the relevant genomic region of chicken defensins, providing an opportunity to characterize the crested ibis defensin cluster in detail.

Second, we used the plasmid DNA of BAC H46-1 as a probe to perform fluorescence in situ hybridization (FISH). The result shows that the crested ibis defensins physically map to chromosome 3, the same chromosome that contains the defensin clusters in the chicken and zebra finch genomes<sup>27,35</sup>, without any non-specific binding (Figure 1c). This means that, in contrast to the human and mouse defensins, which are scattered on several different chromosomes<sup>29</sup>, the avian defensins are probably restricted to only one chromosome.

**Characterization of the crested ibis defensin genes.** Fourteen  $\beta$ -defensin loci clustering within a region of about 129 kb were identified in the clone H46-1 (Table 2). We also obtained the full-length cDNA sequences of these defensin genes except *AvBD11* and *AvBD14*, which may not be expressed in the selected organs or may be only inductively expressed when exposed to certain pathogens. The cDNA sequences obtained range from 378 bp to 813 bp. By aligning the deduced amino acid sequences of the 14 crested ibis defensins, we detected high sequence similarity in the signal peptides but a very low level of conservation in the remaining regions, including the propeptide and mature peptides (Figure 2a). This pattern was also observed in many other vertebrates<sup>23,27,35</sup>, and the high sequence divergence concentrated in the mature peptide among different defensin genes was suggested to be driven by accelerated adaptive selection after successive rounds of gene duplication<sup>23</sup>. Nevertheless, all the defensin genes in crested ibis exhibit six conserved cysteine residues located in the mature peptides (Figure 2a). Moreover, we noticed splicing signal variations in intron 2 (AG/CT) from *AvBD12* and in intron 3 (both GT/CA) from *AvBD2* and *AvBD1 $\alpha$* , which coincidentally correspond to the two hypervariable regions of *AvBDs* (i.e., the propeptide and mature C-terminal, respectively) (Figure 2b–2c). Interestingly, *AvBD11* showed two tandem copies of the special six-cysteine motif in the C-terminal segment with a low level of sequence identity (Figure 2a). Similar patterns were also observed

**Table 2 |  $\beta$ -defensin genes in the crested ibis**

Gene	Orientation	Location <sup>a</sup>	Exon size (bp)			
			Exon 1	Exon 2	Exon 3	Exon 4
<i>AvBD13</i>	–	14688–16253	/ <sup>b</sup>	113	265	/ <sup>b</sup>
<i>AvBD12</i>	–	22992–24182	/ <sup>b</sup>	151	457	/ <sup>b</sup>
<i>AvBD11<sup>c</sup></i>	–	32767–35049	N/A <sup>d</sup>	58	129	128
<i>AvBD10</i>	+	68066–71073	/ <sup>b</sup>	133	133	381
<i>AvBD9</i>	+	77585–81637	/ <sup>b</sup>	131	130	232
<i>AvBD8</i>	–	83676–86995	35	91	134	553
<i>AvBD7</i>	+	93568–95300	54	86	127	245
<i>AvBD2</i>	+	101437–104896	66	79	121	232
<i>AvBD1<math>\beta</math></i>	–	109609–112581	84	88	127	241
<i>AvBD1<math>\alpha</math></i>	–	114639–117374	131	88	127	248
<i>AvBD3</i>	–	119977–123234	80	88	109	246
<i>AvBD5<sup>e</sup></i>	–	125566–127120	224	87	127	222
		125566–126799	/ <sup>b</sup>	116	127	222
<i>AvBD4</i>	–	132130–135514	99	87	127	147
<i>AvBD14<sup>f</sup></i>	–	142664–143208	N/A <sup>d</sup>	55	125	N/A <sup>d</sup>

<sup>a</sup>Gene locations were annotated by sequenced full-length mRNA.

<sup>b</sup>Indicates exon fusion with the adjacent exon.

<sup>c</sup>The *AvBD11* and *AvBD14* mRNA sequences were not obtained from the liver and spleen so that only their coding sequences were predicted using GeneWise and chicken *AvBD11* and *AvBD14* amino acid sequences.

<sup>d</sup>N/A means not available or not applicable (i.e., without untranslated regions).

<sup>e</sup>Two *AvBD5* splice variants were obtained from rapid amplification of cDNA ends.

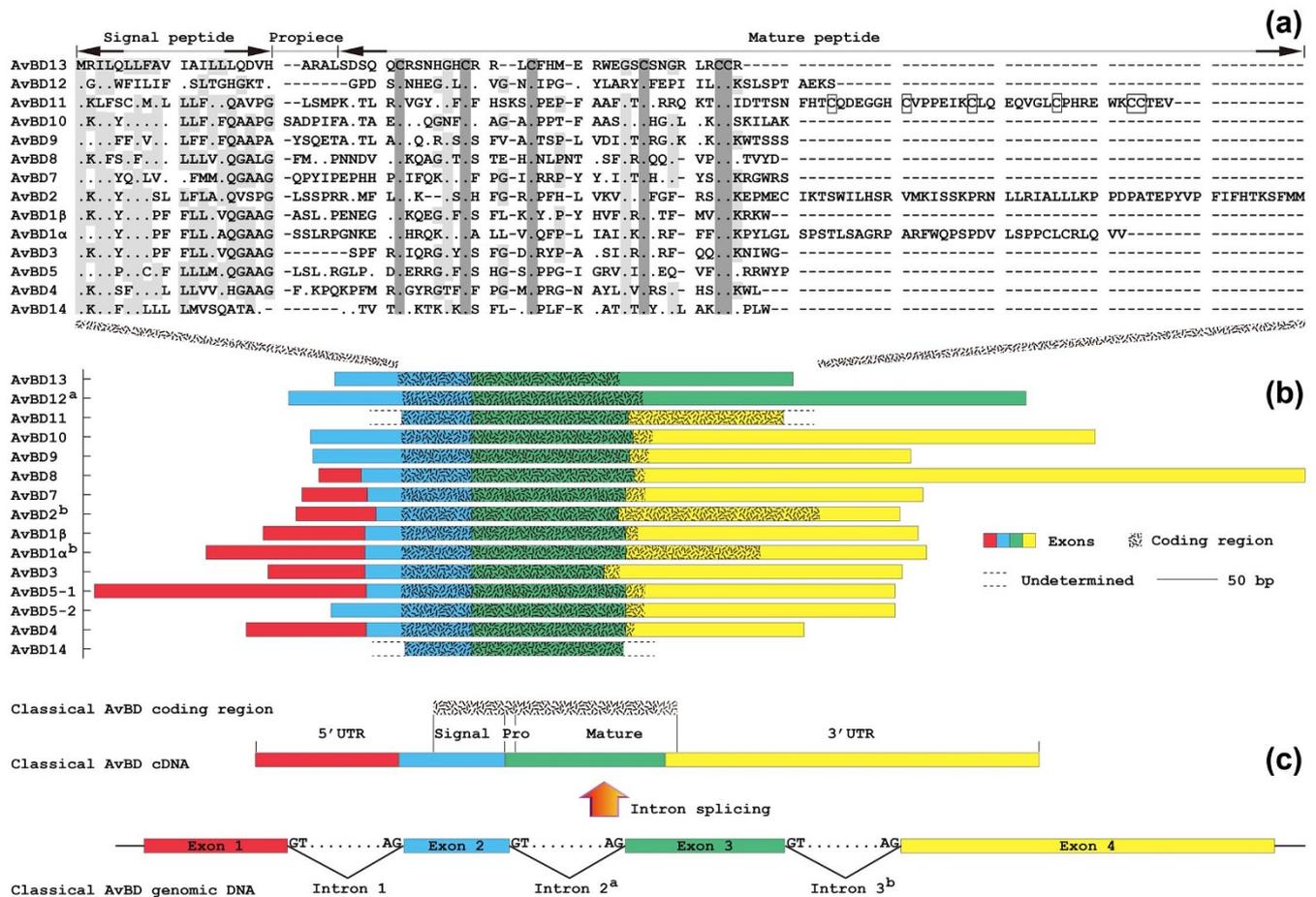
for chicken *AvBD11*<sup>27</sup> and lizard *Anolis carolinensis* *AcBD14*<sup>22</sup>. Obviously, these defensins have six, instead of three, intramolecular disulfide bridges, but the functional significance remains to be studied.

Notably, we also observed extensive fusion of adjacent exons in the crested ibis  $\beta$ -defensins (Table 2), creating three different gene organization patterns (including the numbers and locations of introns/exons) ranging from two exons (typical for human  $\beta$ -defensins) to four exons (typical for the avian  $\beta$ -defensins), as illustrated in Figure 2b. Classical *AvBD* genes are constituted by four exons (four-exon organization; Figure 2c), in which exon 1 corresponds to the 5'UTR, the two internal exons encode the signal, propeptide (absent in some genes), and mature peptides, and exon 4 includes the C-terminal sequence of the mature peptide and the 3'UTR. In some defensins (*AvBD5*, 9, 10), the first two exons fused, leaving only three exons (three-exon organization). Moreover, for *AvBD12* and 13, not only the first two but also the last two exons fused, resulting in “two-exon organization,” which perfectly resembles that of typical human  $\beta$ -defensins<sup>28</sup>. Additionally, we found one alternative mRNA splicing for *AvBD5* (Table 2 and Figure 2b). Compared with the transcript *AvBD5-1*, the other transcript *AvBD5-2* lacks the corresponding exon 1, but its exon 2 instead extends towards intron 1 for 29 bp spliced in *AvBD5-1*. Nevertheless, the two splice variants only differ in 5'UTR and thus share the same translated sequence.

**Comparative genomic analyses.** The genomic organizations of whole  $\beta$ -defensin gene clusters in the chicken<sup>27</sup>, duck<sup>34</sup>, zebra finch<sup>35</sup>, and crested ibis are shown in Figure 3a–3d, and dot plot comparisons between the crested ibis and other birds are shown in

**Table 1 | Conditions for constructing the 4D-PCR and reverse-4D libraries**

Phase	4D-PCR library	Reverse-4D library
Cell concentration in low melting point agarose	$5 \times 10^8$ cells/mL	$2.5 \times 10^8$ cells/mL
Parameters in pulse field gel electrophoresis	pulse time 1–40 s, 6 V/cm for 13 h	1 <sup>st</sup> step: pulse time 1–40 s, 6 V/cm, 13 h 2 <sup>nd</sup> step: pulse time 3–5 s, 5 V/cm, 8 h
DNA concentration in ligation	~6 ng/ $\mu$ L gDNA, 0.5 ng/ $\mu$ L vector	~2 ng/ $\mu$ L gDNA, 0.5 ng/ $\mu$ L vector
Parameters in electrotransformation	100 $\Omega$ , 1.5 kV, 25 $\mu$ F	200 $\Omega$ , 1.5 kV, 25 $\mu$ F



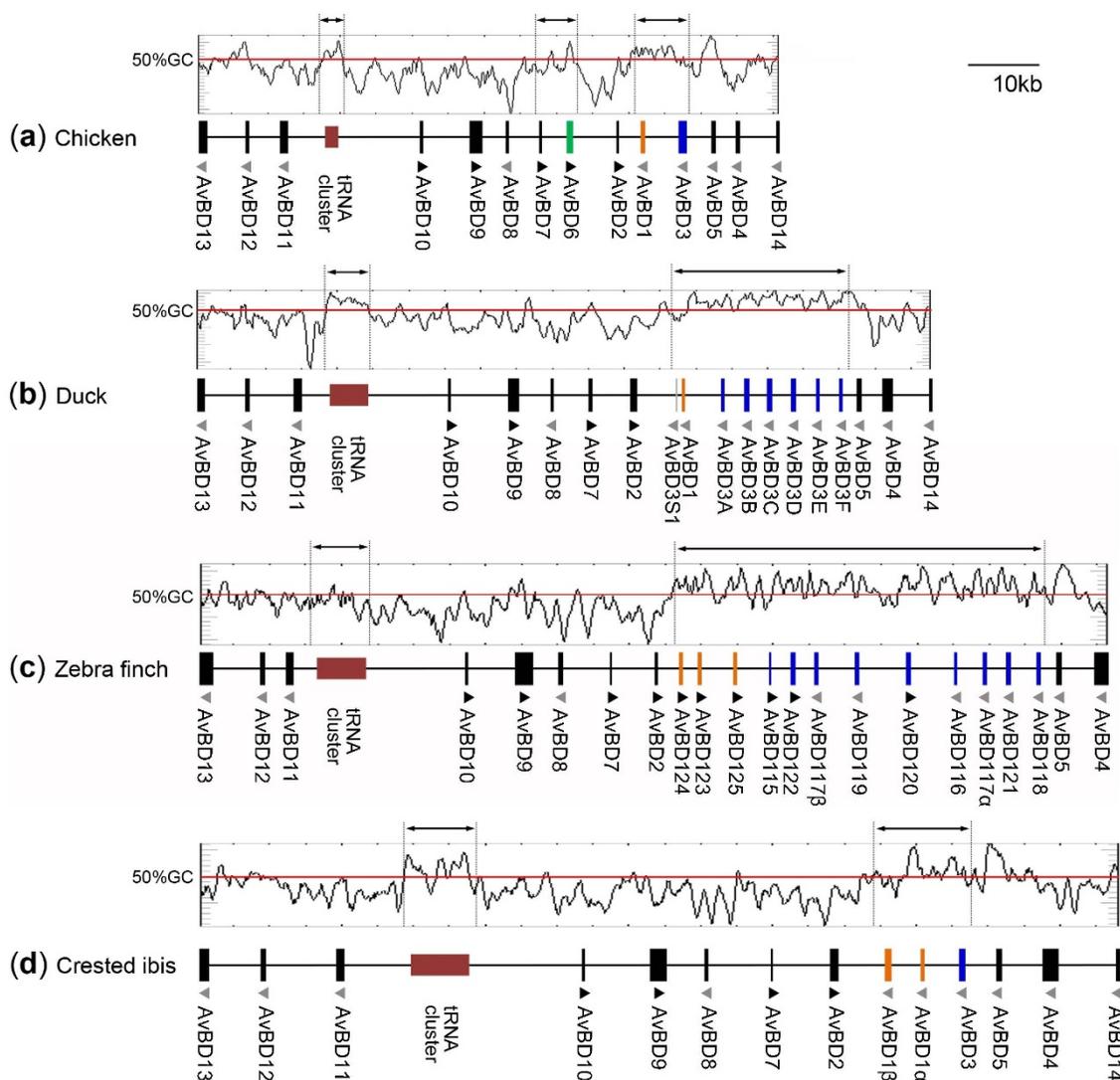
**Figure 2 | Amino acid sequence alignment and structural variation of mRNA/coding sequences.** (a) Dots and dashes represent identity and deletion to *AvBD13*, respectively. Regions corresponding to signal, propiece, and mature peptides are indicated. Light-grey: conserved residues (frequency  $\geq 50\%$ ); dark grey: six highly conserved cysteine motifs; boxes: six additional cysteines in *AvBD11*. (b) Multi-form gene organizations ranging from two to four exons are distinguished by different colors. Two-exon organizations contain fusions of the first and last two exons; three-exon organizations contain fusion of the first two exons (Table 2). “Undetermined” indicates prediction of coding sequences without UTRs (*AvBD11* and *AvBD14*). Gene regions are drawn to scale (lower right corner). Coding regions correspond to sequence alignments in (a). *AvBD5* has two splice variants (5-1 and 5-2) with the same coding sequence. (c) Classical four-exon structure of avian  $\beta$ -defensins. <sup>a</sup>Intron 2 of *AvBD12* shows splicing signal variation (GT/AG to AG/CT). <sup>b</sup>Intron 3 of *AvBD2* and *AvBD1α* exhibits splicing signal variation (GT/AG to GT/CA).

Figure 4a–4c. First, compared to the duck (19 genes within 103 kb, 0.184 genes/kb)<sup>34</sup>, zebra finch (22 genes within 127 kb, 0.173 genes/kb)<sup>35</sup>, and chicken (14 genes within 82 kb, 0.171 genes/kb)<sup>27</sup>, the gene density within the  $\beta$ -defensin cluster is remarkably low in crested ibis (14 genes within 128 kb, 0.109 genes/kb).

Second, a series of gene duplication events were identified in these species (replicated blocks, Figure 4a–4c), including several species-specific duplications and one common duplication (*AvBD1* vs. *AvBD3*), also supported by the relative gene locations (Figure 3a–3d). The detection of two *AvBD7* copies (*AvBD6* and *AvBD7*) in chicken (Figure 4a), seven *AvBD3* copies (*AvBD3S1*, *AvBD3A–3F*) in duck (Figure 4b), and three *AvBD1* copies (*AvBD123–AvBD125*) and nine *AvBD3* copies (*AvBD115–AvBD122*) in zebra finch (Figure 4c) were all in accordance with previous reports<sup>27,34,35</sup>. Similar to the zebra finch, duplication of *AvBD1* was also detected in the crested ibis (Figure 4a–4c), leading to two closely linked gene copies (*AvBD1α* and *AvBD1β*) between *AvBD2* and *AvBD3*. Besides these gene duplications unique to each species, the dot plots also first revealed a common duplication between *AvBD1* and *AvBD3* in all the four birds. This result shows that a historical duplication event preceding the speciation of these species gave rise to the two ancestral AvBD lineages with adjacent locations, which can be dated back to 87–135 Mya<sup>38</sup>.

Third, *AvBD14*, which is absent from zebra finch<sup>35</sup>, was identified in the chicken<sup>27</sup> and duck<sup>34</sup> as well as the crested ibis. In fact, we found that Hwamei (Passeriforme) also lost *AvBD14* (unpublished data). Furthermore, the phylogenetic analysis in the next section shows that *AvBD14* is a single copy orthologous gene among these three birds (Figure 5). Thus, we predicted that the absence of this gene may be exclusive to Passeriformes. Collectively, this study reveals that gene gain and loss have occurred widely during the evolution of the avian defensin gene clusters. Except for the different gene duplications found in the four birds,  $\beta$ -defensin gene clusters are generally conserved with nearly identical gene order and orientation ranging from *AvBD13* to *AvBD14* on the gene map (Figure 3a–3d) with considerable sequence homology (unbroken diagonal lines, Figure 4a–4c).

Interestingly, we noticed that the GC content of the regions corresponding to the subfamily *AvBD1/AvBD3* in all species and the subfamily *AvBD7* in chicken, where gene duplication events occurred, is generally greater than in non-duplicated areas of each species, especially in duck (Figure 3b) and zebra finch (Figure 3c). We further compared the average GC content between the duplicated and non-duplicated regions in each bird and found significant differences ( $P < 0.005$ ) (Supplementary Table S1). Earlier studies reported a positive correlation between duplication/recombination and GC content<sup>39,40</sup>, which implies that the frequent duplication



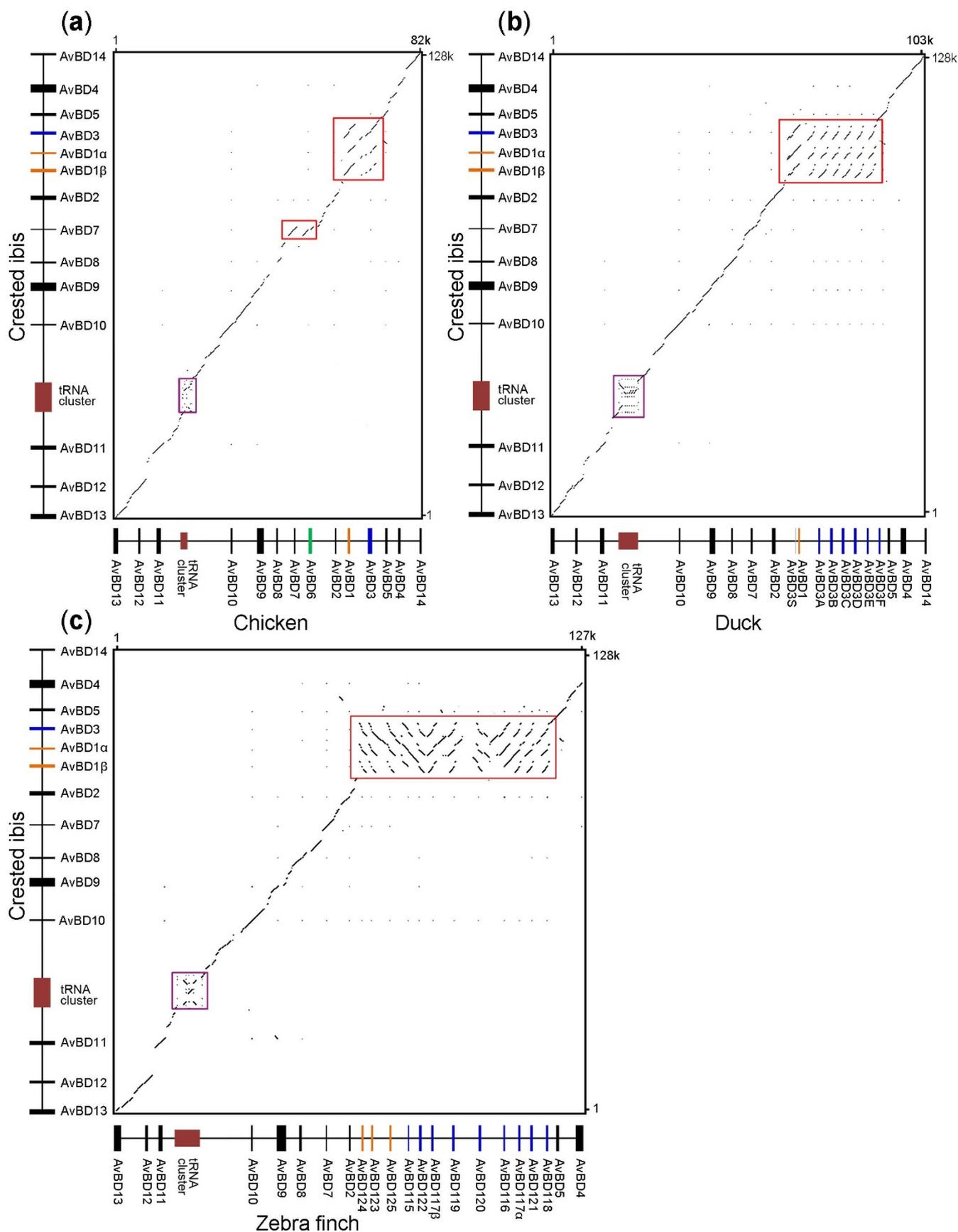
**Figure 3 | Organization of  $\beta$ -defensin gene clusters.**  $\beta$ -defensin clusters from (a) chicken, (b) duck, (c) zebra finch, and (d) crested ibis. Vertical bars: gene locations; lengths drawn to scale (upper right corner). Paralogous genes for *AvBD1* (orange) and *AvBD3* (blue) and chicken-specific *AvBD6* (green) are shown. Red: tRNA clusters. Small triangles: transcriptional orientations of defensin genes. GC contents of gene clusters are shown above the genomic maps. Duplicated regions are bracketed by double-sided arrows and dotted lines.

events in certain avian defensins, especially in the *AvBD1/AvBD3* subfamily, might be associated with the high GC content in these regions. The *AvBD5* region shows the highest GC content in chicken, zebra finch, and crested ibis but has an obviously low GC percentage in duck, suggesting this segment has species-specific GC content variation. Actually, besides the duplicated defensin gene regions, we also detected significantly higher ( $P < 0.005$ ) GC content in the tRNA clusters located between *AvBD10* and *AvBD11* in the four species (Figure 3a–3d) (Supplementary Table S1), within which extensive duplications were also detected by dot plots (Figure 4a–4c). According to the predicted results, these tRNA genes (4 in chicken, 8 in duck, 5 in zebra finch, and 7 in crested ibis) are mainly responsible for transferring tryptophan (hydrophobic) and arginine (cationic), which are typical for defensins<sup>12,20</sup>, indicating the coevolution of defensin and tRNA genes via physical proximity.

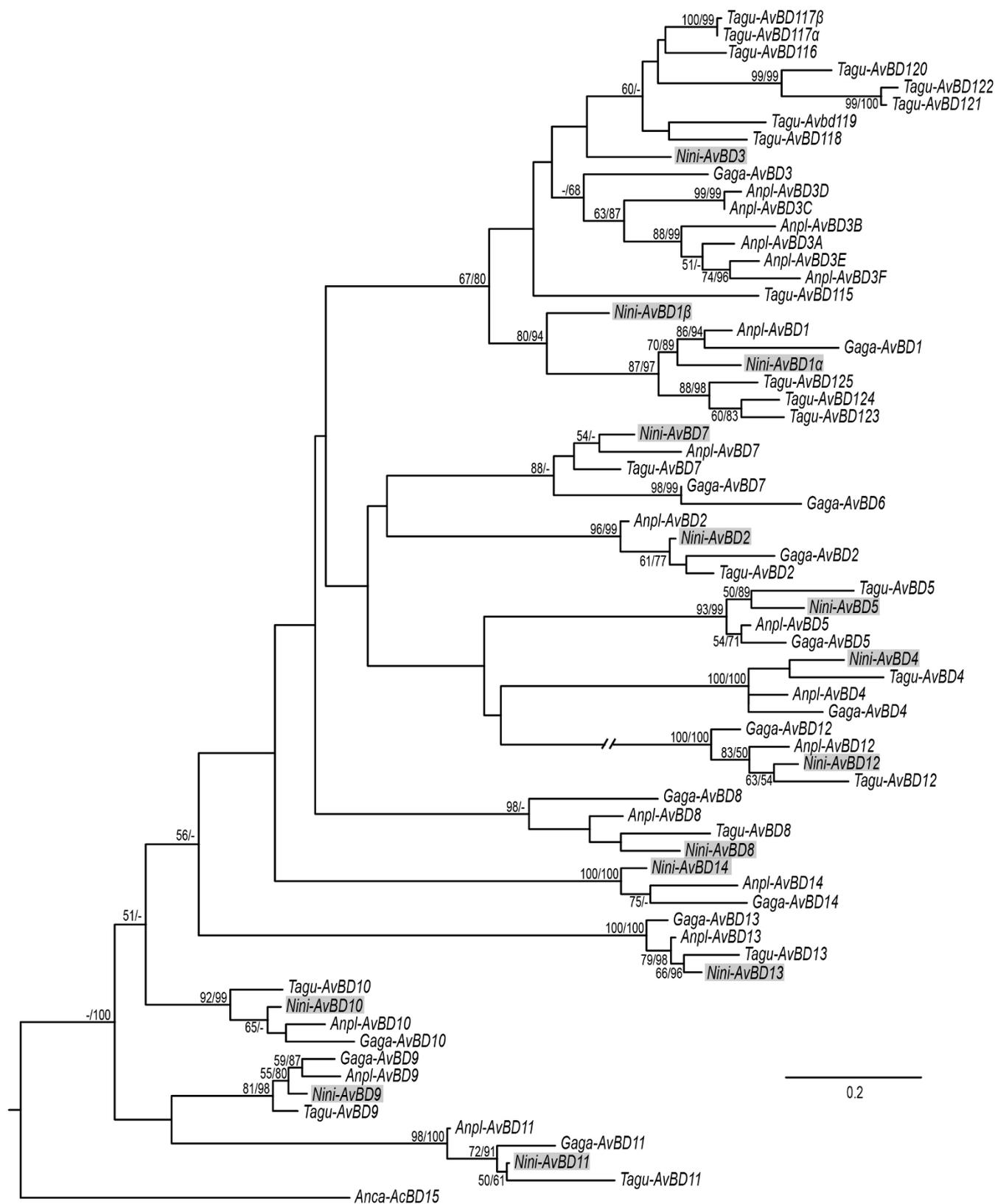
**Evolutionary analyses.** In Bayesian and maximum-likelihood (ML) phylogenetic analyses (Figure 5), genes from the same subfamilies were clustered with relatively high posterior probabilities (or bootstrap values), most of which were  $>90\%$  (bootstrap values  $> 90$ ). Meanwhile, except for 1 : 1 : 1 : 1 orthologs in the *AvBD2*, 4, 5, 8, 9, 10, 11, 12, and 13 clusters, the other clades exhibit single-copy

orthologs among 2–3 birds and/or extra paralogs in the remaining species (Figure 5). Based on the strong orthologous signals and the frequent gene gain and loss, we propose that the avian defensin genes evolved by the birth-and-death process, which was suggested to be an important evolutionary mode for multigene families, especially for immune system genes<sup>41</sup>, and has been reported for the mammalian<sup>42</sup> and reptilian  $\beta$ -defensins<sup>22</sup>.

In both phylogenetic trees, the *AvBD1* and *AvBD3* lineages are grouped into a large subfamily, which further supports the historical gene duplication event detected in Figure 4a–4c. For *AvBD1*, chicken and duck present one copy each while crested ibis and zebra finch show two and three copies, respectively (Figure 5). Nonetheless, the three zebra finch *AvBD1*s are grouped into one intra-species cluster while the crested ibis *AvBD1 $\alpha$*  and *1 $\beta$*  are separated into one basal primitive branch of the all-*AvBD1* clade and another group with the chicken and duck *AvBD1*s (Figure 5), suggesting one ancient duplication in the crested ibis *AvBD1* and two recent duplications in the zebra finch *AvBD1*. For *AvBD3*, we noticed a sharp contrast in gene distribution pattern among species; chicken and crested ibis have only one copy *AvBD3* while duck and zebra finch have many duplicated *AvBD3*s (Figure 5). In addition, *AvBD6* is unique to chicken and shows a high degree of pairing with *Gaga-AvBD7*



**Figure 4** | Dot plot analysis of defensin clusters. Crested ibis vs. (a) chicken, (b) duck, and (c) zebra finch. Positions of the genes for each species are indicated on the corresponding axes (colors as in Figure 3). Defensin (red boxes) and tRNA cluster (purple boxes) duplications are depicted as replicated blocks between syntenic regions.



**Figure 5 | Phylogenetic relationship of avian  $\beta$ -defensins.** AvBD genes from chicken (*Gallus gallus*, *Gaga*), zebra finch (*Taeniopygia guttata*, *Tagu*), duck (*Anas platyrhynchos*, *Anpl*), and crested ibis (*Nipponia nippon*, *Nini*; grey). Numbers beside nodes indicate maximum-likelihood bootstrap supports/Bayesian posterior probabilities (values > 50 only). The AvBD12 lineage branch is truncated by a dash. Green lizard (*Anolis carolinensis*, *Anca*) AcBD15 is included as the outgroup.



(Figure 5), indicating that the *AvBD6* gene was duplicated from *AvBD7* recently and exclusively in Galliformes<sup>35</sup>. As a result, gene duplication further enlarged and diversified the repertoire of  $\beta$ -defensins in different birds.

For orthologous genes among the four birds, the pattern of purifying selection was supported by a mean  $\omega$  (overall  $d_N/d_S$  ratio)  $< 1$  in all defensin loci and further supported by all the pairwise calculated  $\omega < 1$ , with exception for comparison of *AvBD2* between the crested ibis and zebra finch (Table 3). This implies that the  $\beta$ -defensins may have mainly evolved under relaxed purifying selection in Aves<sup>35</sup>. Furthermore, the signal peptide coding regions generally have higher  $\omega$  values than the mature peptides, indicating that the mature peptides have relatively greater sequence conservation within each gene lineage, which may be critical for maintaining specific structure or function for each orthologous defensin gene.

## Discussion

The present study constructed a 4D-PCR library with a relatively small insert size and a reverse-4D library with an elevated average insert size, which is attributed to our condition optimization (Table 1). First, we decreased the number of cells embedded in the low melting point agarose in order to reduce the DNA concentration and increase separation efficiency for small DNA fragments by pulse field gel electrophoresis (PFGE), thus decreasing the amount of short fragments and facilitating inclusion of large DNA fragments into vectors. Second, compared to the PFGE process in the construction of 4D-PCR library, we performed an additional PFGE step for 8 hours (3–5 s pulses at 5 V/cm), which further eliminated small DNA fragments. Third, the electrical resistance was increased to 200  $\Omega$ , thereby elongating the pulse time for electrotransformation, which may promote the uptake of large BACs.

Due to the difficulty of acquiring samples from endangered and rare species, it is important to store genetic resources in genomic DNA libraries. Thus, this study provides optimized conditions of library construction and increases the convenience of building large-insert BAC libraries for other species. Because very few molecular or genomic studies have been performed in the crested ibis, the BAC libraries presented here are not only key for this work but also provide a solid foundation for future species protection based on genomic features of the crested ibis.

In this study, we observed extensive fusion in the crested ibis  $\beta$ -defensins (Figure 2b and Table 2), ranging from the classical four-exon structure (Figure 2c) to mammal-like two-exon organization. In chicken, only *AvBD12* presents fusion of the last two exons (resultant three-exon structure), while all the other defensin genes consist of four exons<sup>27</sup>. In the green lizard, at least seven different gene

organization patterns also ranging from two to four exons were reported<sup>22</sup>, including the exact “four-exon organization” (*AcBD7*, *12*, and *26*), “three-exon organization” (most *AcBDs*) and “two-exon organization” (*AcBD17*, *21*, *22*, *31*, and *30b*). Such variability of gene organization has also been reported in many invertebrates and was proven to be a result of exon-shuffling during evolution<sup>14</sup>. Exon-shuffling by intronic recombination plays an important role in rapid construction of new genes encoding multidomain proteins during evolution<sup>43</sup>. It was previously speculated that the two-exon human  $\beta$ -defensins may evolve through exon-shuffling, and this method is presumably adaptive since it allows  $\beta$ -defensins to mobilize faster to efficiently defend against pathogens<sup>27</sup>.

We also observed two splice variants for the crested ibis *AvBD5* (Figure 2a–2b). Such alternative splice patterns have been revealed in the green lizard (*AcBD13a/b* and *AcBD18a/b*), and the authors suggested that the first exons in both variants probably associate with different proximal promoters, which may contribute to tissue- or time-specific gene expression regulation<sup>22</sup>. Actually, four different alternative splice patterns were found for the green lizard  $\beta$ -defensins, which is undoubtedly a critical mechanism for generating new molecular diversity in  $\beta$ -defensin genes. Thus, exon fusion and alternative splicing might play special roles in defensin evolution and innate immunity of the crested ibis.

The Bayesian and ML trees show that the crested ibis *AvBD1 $\alpha$*  has a closer phylogenetic relationship to the *AvBD1* genes of the other three birds, while *AvBD1 $\beta$*  forms a basal primitive branch of the whole *AvBD1* lineage (Figure 5). This implies that after the duplication between ancestral *AvBD1* and *AvBD3* lineages, the *AvBD1* was further duplicated before the four birds diverged, which generated the two copies *AvBD1 $\alpha$*  and *AvBD1 $\beta$* . However, *AvBD1 $\beta$*  was later lost during the subsequent avian evolution history and has so far only been identified in the crested ibis genome. Consequently, the crested ibis defensins provide insight into ancestral avian defensin structure; in addition to the previously suggested 13 ancestral avian  $\beta$ -defensins including *AvBD1–5* and *AvBD7–14*<sup>34</sup>, we propose that the ancient Neognathae may also have *AvBD1 $\beta$* , which needs verification in other avian orders.

The crested ibis  $\beta$ -defensins, first characterized in this study, will be helpful for protecting this endangered bird in the two following aspects. In recent years, there has been an increasing number of reports regarding infectious diseases in the crested ibis, including septicemia<sup>6</sup>, new castle<sup>2</sup>, *Salmonella* infection<sup>8</sup>, *Escherichia coli* infection<sup>5</sup>, avian influenza<sup>7</sup>, and others, which implies that this rare bird is at increasingly high risk under variable pathogen pressures. Identification of defensin genes responsive to certain pathogens has been conducted in many avian studies<sup>12,34</sup> but never in crested

**Table 3** |  $d_N/d_S$  values calculated for orthologous  $\beta$ -defensin genes in chicken, zebra finch, duck, and crested ibis

Gene	Crested ibis vs. Chicken	Crested ibis vs. Duck	Crested ibis vs. Zebra finch	Chicken vs. Duck	Chicken vs. Zebra finch	Duck vs. Zebra finch	Overall $d_N/d_S$		
							All	Signal peptide	Mature peptide
<i>AvBD2</i>	0.253	0.167	<b>1.925<sup>b</sup></b>	0.197	0.191	0.323	0.260	0.244	0.239
<i>AvBD4</i>	0.533	0.480	0.732	0.707	0.484	0.661	0.586	0.984	0.461
<i>AvBD5</i>	0.339	0.278	0.190	0.426	0.407	0.208	0.279	<b>1.062<sup>b</sup></b>	0.135
<i>AvBD7</i>	0.357	0.412	0.480	0.157	0.464	0.469	0.353	0.541	0.273
<i>AvBD8</i>	0.200	0.200	0.285	0.242	0.238	0.394	0.245	<b>1.053<sup>b</sup></b>	0.167
<i>AvBD9</i>	0.256	0.824	0.158	0.482	0.184	0.322	0.303	0.122	0.371
<i>AvBD10</i>	0.264	0.096	0.248	0.285	0.414	0.162	0.240	0.624	0.168
<i>AvBD11</i>	0.325	0.433	0.171	0.518	0.222	0.390	0.302	0.609	0.232
<i>AvBD12</i>	0.411	0.148	0.655	0.301	0.309	0.191	0.294	0.798	0.188
<i>AvBD13</i>	0.291	0.443	0.331	0.290	0.384	0.386	0.345	0.395	0.311
<i>AvBD14<sup>a</sup></i>	0.292	0.513	/	0.423	/	/	0.410	0.644	0.204
Mean	0.320	0.363	0.517	0.366	0.330	0.351	0.329	0.643	0.250

<sup>a</sup> $d_N/d_S$  values of *AvBD14* were only calculated among chicken, duck, and crested ibis.

<sup>b</sup>The  $d_N/d_S$  values greater than 1 are in bold.



ibis. This study determines crested ibis  $\beta$ -defensin genes in detail and thus makes it possible to conduct pathogen-associated analysis and further help this endangered bird better cope with pathogen invasion.

Although the number of crested ibis individuals has expanded greatly, the severe threat of high mortality rates during the pre- and post-hatch periods, largely caused by pathogen infection, has emerged due to extreme inbreeding<sup>1,4,5</sup>. This could be mainly attributed to the immature immune system of the developing embryos and newly hatched chicks, since components of the adaptive response are in a naïve state<sup>44</sup>. Many avian studies have demonstrated that avian defensins can protect embryos and nestlings from potential pathogenic assault and promote the early transition from an innate immune response to an adaptive one<sup>11,12,45</sup>. Thus, some studies utilized defensin genes as genetic markers to select for resistance to certain infections and thereby enhance the innate immune response via selective breeding in poultry<sup>45,46</sup>, while some others adopted artificially regulated defensin expression via maternal diet, for example, to control the selective colonization of birds and thereby suppress pathological microflora<sup>11,47</sup>. These findings all provide conservation implications for scientists and managers, e.g., methods for using defensin-based genetic management and dietary regulation to reduce the early mortality rates of crested ibis, especially in captive populations.

## Methods

**Ethics statement.** The samples used in this study, including blood, liver and spleen, were collected from two captive individuals with permission from the Crested Ibis Breeding Centers (CIBC) and the Department of Wildlife Conservation, State Forestry Administration. The blood sample was obtained from the wing vein of a female individual bred in CIBC, Louguantai County, Shaanxi Province, China and was performed in cooperation with technical staff from the center with utmost care. The liver and spleen tissues were acquired from another dead nestling provided by CIBC, Deqing County, Zhejiang Province, China, which was immediately dissected after its death that was probably caused by a birth defect. All experiments were approved by the ethics committees of these two CIBCs and were carried out in accordance with the approved guidelines.

**Construction of the BAC library.** A routine BAC genomic DNA library was generated with DNA from the blood sample. The construction of this library was based on a previous protocol<sup>36</sup> with a few modifications using a pCCIBAC vector (Epicentre) and MboI restriction enzyme. The library was arrayed for rapid screening by 4D-PCR<sup>36</sup>. A reverse-4D BAC library was generated and characterized following a previous protocol<sup>37</sup>, with some conditions optimized (Table 1). Three different restriction enzymes (MboI, EcoRI, and HindIII) were used separately to form DNA inserts with different cohesive ends in order to maximize the complementation of sublibraries.

**BAC screening and sequencing.** Nine pairs of universal primers<sup>48</sup> were used to perform PCRs with the crested ibis genomic DNA as template. Agarose gel electrophoresis on PCR products indicated that four primer pairs (primers for *AvBD4*, 7, 11, and 12) had derived products, among which the product of primers for *AvBD7* (5' to 3' sequences, F: TTAGTCACTTCATGGTGAATGAC, R: TCACCTTTGCGAGCAAGAATA) produced a single band of the expected size. To further confirm this identification, the product amplified by *AvBD7* primers was processed by (1) isolation by agarose gel electrophoresis and purification by DNA gel extraction kit (Axygen), (2) ligation into pMD18-T vector (TaKaRa) and transformation into *E. coli* DH5 $\alpha$  competent cells (TaKaRa), and (3) sequencing of the cloned products. Sequences were identified by BLAST as *AvBD7* from the crested ibis.

BAC clones containing  $\beta$ -defensin loci were subsequently screened from both BAC libraries using the *AvBD7* primers above. The PCR screening processes were performed as previously described<sup>36,37</sup>. Insert sizes of positive clones were estimated by electrophoresis, and BAC end sequencing<sup>36,37</sup> was performed. The end sequences of these  $\beta$ -defensin-containing clones were aligned with the available chicken  $\beta$ -defensin genomic sequences from NCBI to preliminarily determine whether the BAC clone contained the entire defensin cluster. Finally, the BAC was shotgun subcloned, sequenced, and assembled by a commercial sequencing service (Majorbio). The full BAC sequence was deposited in GenBank (accession number: KM098134).

**Chromosome localization of the crested ibis  $\beta$ -defensins.** Metaphase chromosomes were prepared from the peripheral blood of the crested ibis according to the method described by Moorhead et al.<sup>49</sup>, and the BAC clone H46-1 was used as a probe. FISH was performed following the previously reported protocol<sup>50</sup>. One hundred well-spread and distinct metaphases were karyotyped by screening 15 slides, and the chromosome sizes were measured using ImageJ<sup>51</sup>, including the total length of the

macrochromosomes and the arm ratio (long arm/short arm, q/p). Chromosome numbering was performed according to the reported karyotype ideogram of crested ibis<sup>52</sup> and the criteria recommended by Levan et al.<sup>53</sup>.

**Sequence analyses.** According to the order and orientation as well as the clustering results of phylogenetic analysis described below, the crested ibis defensin genes were named following the nomenclature proposed previously<sup>33</sup>. The mRNA was extracted from the liver and spleen, and the full-length sequences of  $\beta$ -defensin cDNA were obtained by GeneRacer™ Kit (Invitrogen) using the primers listed in Supplementary Table S2. All the cDNA sequences were used for gene annotation and deposited in GenBank (accession number: KM272304–KM272316). The defensin genes undetected in the liver and spleen tissues were analyzed with BLAST based on homologies to known chicken genes and the coding regions were predicted by GeneWise<sup>54</sup>.

The chicken (chromosome 3: 107,022,000–107,104,000 bp) and zebra finch (chromosome 3: 110,733,400–110,860,340 bp) defensin cluster sequences were both obtained from Ensembl (<http://asia.ensembl.org/index.html>), while the duck defensin sequences were obtained from NCBI (NW\_004678274.1, NW\_004683518.1, KB742693.1, and NW\_004716371.1). The several gaps in and between these four duck defensin scaffolds were filled by applying long-range PCR (Supplementary Table S3), cloning, and sequencing. We predicted tRNA genes in defensin clusters for four birds by tRNAscan-SE<sup>55</sup>, analyzed GC content with 200 bp windows using Isochore<sup>56</sup> ([http://www.ebi.ac.uk/Tools/seqstats/emboss\\_isochore/](http://www.ebi.ac.uk/Tools/seqstats/emboss_isochore/)), and further conducted Pearson's chi-squared test in SPSS 20.0 (SPSS Inc., Chicago, IL, US) to identify any significant differences between each duplicated region and the non-duplicated region. Finally, we drew dot plots using PipMaker<sup>57</sup> to perform syntenic analysis between the crested ibis and other bird defensin clusters.

**Evolutionary analyses.** Amino acid sequences were aligned by MUSCLE in MEGA5.0<sup>58</sup>. According to the amino acid sequence alignment, nucleotide sequences were aligned and subsequently used for phylogenetic analysis with chicken, zebra finch, and duck  $\beta$ -defensin genes. Nucleotide sequences encoding the signal and mature peptides were used for phylogenetic reconstruction in MrBayes 3.2.2<sup>59</sup> with the green lizard *AcBD15* (GenBank accession number: FR850158) as the outgroup. Two independent runs of 600,000 generations with a sample frequency of 100 were performed, and the first 25% of the samples were discarded as "burn-in." An ML tree was also constructed using the PhyML web service Phylogeny.fr<sup>60</sup> with default settings. The rates of synonymous and non-synonymous substitutions ( $d_S$  and  $d_N$ , respectively) were calculated in MEGA5.0. Since the saturated non-synonymous sites in paralogous genes could bias the comparisons<sup>35</sup>, the  $d_N/d_S$  ratios ( $\omega$ ) were calculated pairwise for orthologous genes among the four birds (*AvBD14* genes were only compared among the chicken, duck, and crested ibis). The overall  $d_N/d_S$  ratios were calculated separately for signal peptide and mature peptide coding regions.

- Ding, C. Q. *Research on the crested ibis*. 1–388 (Shanghai Scientific and Technological Publishing House, 2004) (in Chinese).
- Chen, S. L. et al. Phylogenetic and pathogenic analyses of two virulent Newcastle disease viruses isolated from Crested Ibis (*Nipponia nippon*) in China. *Virus Genes* **46**, 447–453 (2013).
- Zhang, B., Fang, S. G. & Xi, Y. M. Low genetic diversity in the endangered crested ibis *Nipponia nippon* and implications for conservation. *Bird Conserv Int* **14**, 183–190 (2004).
- Li, X. H. & Li, D. M. Current state and the future of the crested ibis (*Nipponia nippon*): a case study by population viability analysis. *Ecol Res* **13**, 323–333 (1998).
- Fan, G. L. et al. The histopathological observations of young crested ibis infected with *Escherichia coli*. *Chinese J Zool* **39**, 44–46 (2003).
- Xi, Y. M., Wood, C., Lu, B. Z. & Zhang, Y. M. Prevalence of a septicemia disease in the crested ibis (*Nipponia nippon*) in China. *Avian Dis* **51**, 614–617 (2007).
- Li, L. J. et al. Isolation and identification of LPAIV from crested ibis. *Prog Vet Med* **32**, 42–44; 10.3969/j.issn.1007-5038.2011.02.010 (2011) (in Chinese).
- He, Y. L. et al. The pathological observation of crested ibis death of *Salmonellosis*. *Chinese Agr Sci Bull* **21**, 15–17 (2005) (in Chinese).
- Shizuya, H. et al. Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector. *P Natl Acad Sci USA* **89**, 8794–8797 (1992).
- Osoegawa, K. et al. An improved approach for construction of bacterial artificial chromosome libraries. *Genomics* **52**, 1–8 (1998).
- Zhang, G. L. & Sunkara, L. T. Avian antimicrobial host defense peptides: from biology to therapeutic applications. *Pharmaceuticals* **7**, 220–247 (2014).
- Cuperus, T., Coorens, M., van Dijk, A. & Haagsman, H. P. Avian host defense peptides. *Dev Comp Immunol* **41**, 352–369 (2013).
- Klotman, M. E. & Chang, T. L. Defensins in innate antiviral immunity. *Nat Rev Immunol* **6**, 447–456 (2006).
- Froy, O. & Gurevitz, M. Arthropod and mollusk defensins—evolution by exon-shuffling. *Trends Genet* **19**, 684–687 (2003).
- Thomma, B. P., Cammue, B. P. & Thevissen, K. Plant defensins. *Planta* **216**, 193–202 (2002).
- Lehrer, R. I. & Ganz, T. Defensins of vertebrate animals. *Curr Opin Immunol* **14**, 96–102 (2002).
- Chen, H. Q. et al. Recent advances in the research and development of human defensins. *Peptides* **27**, 931–940 (2006).



18. Brogden, K. A. Antimicrobial peptides: pore formers or metabolic inhibitors in bacteria? *Nat Rev Microbiol* **3**, 238–250 (2005).
19. Grigat, J., Soruri, A., Forssmann, U., Riggert, J. & Zwirner, J. Chemoattraction of macrophages, T lymphocytes, and mast cells is evolutionarily conserved within the human alpha-defensin family. *J Immunol* **179**, 3958–3965 (2007).
20. Selsted, M. E. & Ouellette, A. J. Mammalian defensins in the antimicrobial immune response. *Nat Immunol* **6**, 551–557; 10.1038/ni1206 (2005).
21. Tang, Y. Q. *et al.* A cyclic antimicrobial peptide produced in primate leukocytes by the ligation of two truncated  $\alpha$ -defensins. *Science* **286**, 498–502 (1999).
22. Dalla Valle, L., Benato, F., Maistro, S., Quinzani, S. & Alibardi, L. Bioinformatic and molecular characterization of beta-defensins-like peptides isolated from the green lizard *Anolis carolinensis*. *Dev Comp Immunol* **36**, 222–229; 10.1016/j.dci.2011.05.004 (2012).
23. Maxwell, A. I., Morrison, G. M. & Dorin, J. R. Rapid sequence divergence in mammalian  $\beta$ -defensins by adaptive evolution. *Mol Immunol* **40**, 413–421; 10.1016/s0161-5890(03)00160-3 (2003).
24. Bruhn, O., Paul, S., Tetens, J. & Thaller, G. The repertoire of equine intestinal  $\alpha$ -defensins. *BMC Genomics* **10**, 631; 10.1186/1471-2164-10-631 (2009).
25. Belov, K. *et al.* Characterization of the opossum immune genome provides insights into the evolution of the mammalian immune system. *Genome Res* **17**, 982–991 (2007).
26. Nguyen, T. X., Cole, A. M. & Lehrer, R. I. Evolution of primate  $\theta$ -defensins: a serpentine path to a sweet tooth. *Peptides* **24**, 1647–1654 (2003).
27. Xiao, Y. J. *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* **5**, 56; 10.1186/1471-2164-5-56 (2004).
28. Pazgier, M., Hoover, D. M., Yang, D., Lu, W. & Lubkowski, J. Human  $\beta$ -defensins. *Cell Mol Life Sci* **63**, 1294–1313 (2006).
29. Schutte, B. C. *et al.* Discovery of five conserved  $\beta$ -defensin gene clusters using a computational search strategy. *P Natl Acad Sci USA* **99**, 2129–2133 (2002).
30. Evans, E. W., Beach, G. G., Wunderlich, J. & Harmon, B. G. Isolation of antimicrobial peptides from avian heterophils. *J Leukocyte Biol* **56**, 661–665 (1994).
31. Wang, R. Q. *et al.* Identification and characterization of an avian  $\beta$ -defensin orthologue, avian  $\beta$ -defensin 9, from quails. *Appl Microbiol Biot* **87**, 1395–1405 (2010).
32. Ma, D. Y. *et al.* Functional analysis and induction of four novel goose (*Anser cygnoides*) avian  $\beta$ -defensins in response to *salmonella enteritidis* infection. *Comp Immunol Microbiol Infect Dis* **35**, 197–207 (2012).
33. Lynn, D. J. *et al.* Avian beta-defensin nomenclature: a community proposed update. *Immunol Lett* **110**, 86–89 (2007).
34. Huang, Y. H. *et al.* The duck genome and transcriptome provide insight into an avian influenza virus reservoir species. *Nat Genet* **45**, 776–783 (2013).
35. 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* **6**, 3; 10.1186/1745-7580-6-3 (2010).
36. Asakawa, S. *et al.* Human BAC library: construction and rapid screening. *Gene* **191**, 69–79 (1997).
37. Ye, Q., He, K., Wu, S. Y. & Wan, Q. H. Isolation of a 97-kb minimal essential MHC B locus from a new reverse-4D BAC library of the golden pheasant. *PLoS One* **7**, e32154; 10.1371/journal.pone.0032154 (2012).
38. Hedges, S. B. & Kumar, S. *The timetree of life*. (Oxford University Press, 2009).
39. Fullerton, S. M., Carvalho, A. B. & Clark, A. G. Local rates of recombination are positively correlated with GC content in the human genome. *Mol Biol Evol* **18**, 1139–1142 (2001).
40. Zhang, L. Q., Lu, H. H. S., Chung, W. Y., Yang, J. & Li, W. H. Patterns of segmental duplication in the human genome. *Mol Biol Evol* **22**, 135–141 (2005).
41. Nei, M. & Rooney, A. P. Concerted and birth-and-death evolution of multigene families. *Annu Rev Genet* **39**, 121–152 (2005).
42. Semple, C. A. M., Gautier, P., Taylor, K. & Dorin, J. R. The changing of the guard: Molecular diversity and rapid evolution of  $\beta$ -defensins. *Mol Divers* **10**, 575–584; 10.1007/s11030-006-9031-7 (2006).
43. Patthy, L. Genome evolution and the evolution of exon-shuffling—a review. *Gene* **238**, 103–114 (1999).
44. Firth, M. A., Shewen, P. E. & Hodgins, D. C. Passive and active components of neonatal innate immune defenses. *Anim Health Res Rev* **6**, 143–158 (2005).
45. Hasenstein, J. R., Zhang, G. L. & Lamont, S. J. Analyses of five gallinacin genes and the *Salmonella enterica* serovar Enteritidis response in poultry. *Infect Immun* **74**, 3375–3380 (2006).
46. Hasenstein, J. R. & Lamont, S. J. Chicken gallinacin gene cluster associated with *Salmonella* response in advanced intercross line. *Avian Dis* **51**, 561–567 (2007).
47. Trebichavský, I. & Šplíchal, I. Probiotics manipulate host cytokine response and induce antimicrobial peptides. *Folia Microbiol* **51**, 507–510 (2006).
48. Hellgren, O. & Sheldon, B. C. 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* **11**, 686–692 (2011).
49. Moorhead, P. S., Nowell, P. C., Mellman, W. J., Battips, D. M. & Hungerford, D. A. Chromosome preparations of leukocytes cultured from human peripheral blood. *Exp Cell Res* **20**, 613–616 (1960).
50. Wan, Q. H. *et al.* Genome analysis and signature discovery for diving and sensory immune genes (antimicrobial peptides:  $\beta$ -defensins) across endangered Chinese alligator. *Cell Res* **23**, 1091–1105 (2013).
51. Abramoff, M. D., Magalhães, P. J. & Ram, S. J. Image processing with ImageJ. *Biophotonics Int* **11**, 36–43 (2004).
52. Liu, L. Y., Wang, B. G., Guo, X. C. & Li, F. L. chromosomal identification of sex and karyotype analysis in *Nipponia nippon*. *J B Norm Uni* **28**, 554–557 (1992) (in Chinese).
53. Levan, A., Fredga, K. & Sandberg, A. A. Nomenclature for centromeric position on chromosomes. *Hereditas* **52**, 201–220 (1964).
54. Birney, E. & R, D. Using GeneWise in the Drosophila annotation experiment. *Genome Res* **10**, 547–548 (2000).
55. Lowe, T. M. & Eddy, S. R. tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res* **25**, 0955–0964; 10.1093/nar/25.5.0955 (1997).
56. Bernardi, G. Isochores and the evolutionary genomics of vertebrates. *Gene* **241**, 3–17 (2000).
57. Schwartz, S. *et al.* PipMaker—a web server for aligning two genomic DNA sequences. *Genome Res* **10**, 577–586 (2000).
58. Tamura, K. *et al.* MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* **28**, 2731–2739 (2011).
59. Ronquist, F. *et al.* MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol* **61**, 539–542 (2012).
60. Dereeper, A. *et al.* Phylogeny. fr: robust phylogenetic analysis for the non-specialist. *Nucleic Acids Res* **36**, W465–W469; 10.1093/nar/gkn180 (2008).

## Acknowledgments

We would like to thank Dr. Zhang SL (Zhejiang University) for his help in improving the sensitivity of FISH. We appreciate Dr. Huang YH (China Agriculture University) for providing amino acid sequences of duck defensins to design primers for filling gaps. This work was supported by a grant from the National Natural Science Foundation of China (No. 31270423), a special grant from the State Forestry Administration, and the Fundamental Research Funds for the Central Universities of the P. R. China.

## Author contributions

Q.H.W. conceived and designed the research. L.C.C., H.C., B.B.W., L.S. and M.Y.M. performed the experiments. S.G.F. provided the samples. L.C.C. drafted the earlier version of this manuscript. H.C. prepared the Figures and Tables of both drafted and revised manuscripts. H.L. and Q.H.W. revised the manuscript. All authors read and approved the final manuscript.

## Additional information

**Supplementary information** accompanies this paper at <http://www.nature.com/scientificreports>

**Competing financial interests:** The authors declare no competing financial interests.

**How to cite this article:** Lan, H. *et al.* The first report of a Pelecaniformes defensin cluster: Characterization of  $\beta$ -defensin genes in the crested ibis based on BAC libraries. *Sci. Rep.* **4**, 6923; DOI:10.1038/srep06923 (2014).



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder in order to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>