Ovodefensins, an Oviduct Specific Antimicrobial Gene Family Have Evolved in Birds and Reptiles to Protect the Egg by Both Sequence and Intra Six Cysteine Sequence Motif Spacing

Ovodefensins, an oviduct specific antimicrobial gene family have evolved in birds and reptiles to protect the egg by both sequence and intra six cysteine sequence motif spacing[1].

Running title: Oviduct ovodefensin evolution

Keywords: Defensin, egg, antimicrobial, chicken, oviduct, *E. coli*, steroid, evolution, *S. aureus.*

Natasha Whenham[2,3], Tian Chee Lu[3], Maisarah B. M. Maidin[3], Peter W. Wilson[3], Maureen M. Bain[4], Lynn Stevenson[4], Mark P. Stevens[3], Michael R. Bedford[5] and Ian C. Dunn[3]

[3]Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, Midlothian, Scotland, UK. [4]College of Medical, Veterinary and Life Sciences (MVLS), University of Glasgow, Glasgow, Scotland, UK. [5]AB Vista Feed Ingredients, Marlborough, Wiltshire, England, UK

Summary sentence: Bird and reptile oviduct specific antimicrobial gene family, the ovodefensins, have evolved both their sequence and the amino acid spacing between a conserved cysteine sequence motif.

Some of the data in the paper has been presented at the Avian Immunology Research Group Meeting, July 16-19 2014, Guelph, Canada: Ovodefensins; egg specific antimicrobial peptides of birds and reptiles.

[1]Funded by a CASE award to NW from the BBSRC and AB Vista. The Roslin Institute is funded by a core strategic grant from the BBSRC.

[2]Correspondence: Natasha Whenham, Roslin Institute and Royal (Dick) School of Veterinary Studies, University of Edinburgh, Easter Bush Campus, The Roslin Building, Midlothian EH25 9RG, Scotland, UK, E-mail: natasha.whenham@roslin.ed.ac.uk

Abstract

Ovodefensins are a novel beta defensin related family of antimicrobial peptides containing conserved glycine and six cysteine residues. Originally thought to be restricted to the albumen producing region of the avian oviduct, expression was found in chicken, turkey, duck and zebra finch in large quantities in many parts of the oviduct, but this varied between species and between gene forms in the same species. Using new search strategies the ovodefensin family now has 35 members including reptiles, but no representatives outside birds and reptiles have been found. Analysis of their evolution shows that ovodefensins divide into 6 groups based on the intra cysteine amino acid spacing, representing a unique mechanism alongside traditional evolution of sequence. The groups have been used to base a nomenclature for the family. Antimicrobial activity for three ovodefensins from chicken and duck was confirmed against *E. coli* and a pathogenic *E. coli* strain as well as a gram +ve organism, *S. aureus*, for the first time. However, activity varied greatly between peptides, with *Gallus gallus* OvoDA1 being the most potent, suggesting a link with the different structures. Expression of *Gallus gallus* OvoDA1 (gallin) in the oviduct was increased by oestrogen and progesterone and in the reproductive state. Overall the results support the hypothesis that ovodefensins evolved to protect the egg but they are not necessarily restricted to the egg white. There divergent motif structure and sequence present an interesting area of research for antimicrobial peptide design and understanding protection of the *cleidoic* egg.

Introduction

The avian egg has many mechanisms in place to protect itself from bacterial invasion as does the reproductive tract; these can effectively be considered as either physical or chemical [1]. An important part of the eggs chemical defense is provided by antimicrobial peptides (AMPs), otherwise known as host defense peptides (HDPs). A particular group of these, the ovodefensins, was recently shown to be a new family of egg specific defensins [2]. The family had been shown to be conserved across divergent avian species and was thought to be avian specific [2]. Proteomic methods had confirmed the presence of the chicken, turkey and duck members of the ovodefensin family in the egg [3-5] and quantitative reverse transcriptase PCR confirmed that the expression of the chicken member, *gallin* was restricted to the oviduct of the hen [2]. In contrast classical defensins are widely distributed across many tissues and can be found in all vertebrates [6]. Currently, only three of the classical chicken defensins are found in the egg; AvBD9, 10 and 11 [7]. It has yet to be determined if the expression of other avian members of the ovodefensin family are also restricted to the oviduct, which would imply that the whole family is likely to be expressed principally for inclusion in the egg. If this is the case ovodefensins are likely to be influenced by gonadal steroids and would be expected to show expression patterns in response to steroids typical of egg specific genes such as ovalbumin [8] and 'transiently expressed in neural precursors' (TENP) [9]. Although the connection had not previously been made with classical vertebrate defensins, ovodefensins were classed as a new branch of this family largely because of the conservation of a characteristic 6 cysteine sequence motif and a common glycine residue between C1 and C2 attributed to all vertebrate β-defensins [2]. Their position in the genome was also close to the β-defensin cluster on chromosome 3 [2]. The ovodefensins differ from classical vertebrate defensins in the spacing of amino acids within

the 6 cysteine sequence motif and are slightly shorter in length, the mature peptides ranging from only 39-41 amino acids, although they are still highly cationic as is expected with defensins. Two cysteine sequence motifs were observed in ovodefensins; C-X5-C-X3-C-X11-C-X3-CC and C-X3-C-X3-C-X11-C-X4-CC [2] which may be due to the fact that antimicrobial peptides are often under high selective pressure to evolve due to the ongoing arms race between pathogen and host, such as observed in classical vertebrate β-defensins [10]. Interestingly the 3D structure of the chicken ovodefensin gallin has recently been solved which confirmed the presence of the three-stranded antiparallel β-sheet observed in all classical β-defensins reinforcing its relationship with the β-defensins [7]. However, gallin contains an additional short two-stranded β-sheet [7], this five-stranded arrangement supports the hypothesis that gallin, and presumably the other ovodefensins, form a structurally distinct sub-family of β-defensins.

Host defense peptides such as the defensins have been suggested previously as an interesting template for new classes of antimicrobial drugs as they often possess a wide spectrum of antimicrobial activity [11, 12]. Cationic host defense peptides are small, typically containing a high abundance of positively charged and hydrophobic residues [13]. Gallin, the chicken representative of the ovodefensin family was shown to be highly antibacterial against a laboratory-adapted strain of *E. coli* [2]. It has since been suggested that its direct antimicrobial actions are limited to *E. coli* [7], however few studies have been carried out relating to the antimicrobial capabilities of gallin and indeed this novel family of defensins.

Our overall aim of this study was therefore; to determine the extent of the ovodefensin family using the possibility that evolution acting on spacing as well as sequence may reveal the presence of further molecules with novel cysteine sequence motifs; to determine if ovodefensins are avian specific, as was previously speculated, and to determine if ovodefensins are restricted

to the oviduct and are therefore egg specific as was the case with gallin. Finally we wished to determine the antibacterial capabilities of gallin, a newly discovered chicken ovodefensin member (OvoDB1) and duck BPS_2 and the pH and salt sensitivity of gallin.

Materials and Methods

Discovery of new ovodefensin family members

Available genome databases Ensembl [14], PreEnsembl [15] and UCSC [16] were searched using TBLASTN and BlastP [17] to locate potential homologs using the 41 amino acid mature peptide sequence of gallin (GenBank: CBE70283.1) and the previously published 39 amino acid mature peptide sequence of taeniopygin 2 [2]. Further iterative searches were made with the homologues discovered. Protein database Uni-prot [18] was also searched using BlastP to identify peptide sequences previously unidentified as ovodefensins.

*Phylogenetic analysis of the evolution of spacing between conserved residues***.**

A distance matrix based on the amino acid sequence length between each cysteine and the conserved glycine residue was built with the statistical computing software R [19] for all known and newly discovered ovodefensins (Supplementary Data 1). Hierarchical clustering for each distance matrix was calculated using R [19] which was then used to produce a cladogram of the phylogenetic relationships using the R 'hclust' function (Figure 1A). Similarly cladograms based on spacing were produced separately for avian species (Figure 1C) and reptiles (Figure 1D). The resulting ovodefensin 'sub-families' were used to propose new nomenclature for existing and newly discovered ovodefensin molecules (Table 1) which currently have trivial names based on a range of criteria determined by the discoverer. We propose that each gene has the prefix OvoD to identify it as an ovodefensin and is attributed a letter from A-F to identify the

sub-family to which it belongs. Within a sub-family each gene is given a numerical identifier allowing multiple forms of the same gene to be identified. For example *gallin* would become *Gallus gallus OvoDA1* and the additional copies *OvoDA1_2* and *OvoDA1_3*. Through the use of this nomenclature *meleagrin* is *Meleagris gallopavo OvoDA1* allowing it to be easily identified as a member of the same sub-family and a *gallin* ortholog aiding comparison across species. This nomenclature will be used throughout the remainder of this manuscript to aid clarity.

In addition to this a phylogram was constructed using Mega5 [20] and the core peptide sequence from the conserved glycine until the fourth cysteine residue inclusively for sequences where the number of amino acids within this region was identical (Figure 1B). Mega5 was also used to construct phylograms based on the whole mature peptide sequences for OvoDA family members (Figure 1E) and also OvoDB (Figure 1F).

ClustalW alignments of each of the 6 cysteine sequence motifs (OvoDA-OvoDF) and a global alignment of all ovodefensin peptides can be found in supplementary data 2.

Animals and tissue collection

To determine *Gallus gallus OvoDA1* and *OvoDB1*expression the following tissues were taken from four sexually mature White Leghorn LSL hens (Lohmann): Oviduct (magnum, isthmus, shell gland vagina), ovarian stroma, crop, duodenal loop, gizzard, caeca, cloaca, lung, adrenals, cerebellum, retina, spleen, liver, kidney, and heart.

For analysis of *Anas platyrhynchos OvoDB1* (*dBPS$_1$)* and *OvoDA1* (*dBPS$_2$)* expression, tissues were collected from three sexually mature Pekin ducks (Cherry Valley): Oviduct (magnum, isthmus, shell gland, vagina), ovarian stroma, crop, proventriculus, small intestine, duodenal

loop, gizzard, large intestine, caeca, cloaca, gall bladder, lung, trachea, pituitary, adrenals, cerebellum, hypothalamus, tongue, spleen, breast muscle, liver, kidney and heart.

Combined expression of Meleagris gallopavo OvoDA1 (Meleagrin) and *OvoDA2* and expression of *OvoDB1* was determined using tissues from four sexually mature turkeys: Oviduct (magnum, isthmus, shell gland, vagina), ovarian stroma, oesophagus, crop, duodenal loop, gizzard, caeca, cloaca, lung, adrenals, cerebellum, tongue, spleen, breast muscle, skin, liver, kidney and heart.

Taeniopygia guttata OvoDA1 (Taeniopygin 1) and *OvoDB1 (Taeniopygin 2)* expression was measured using five adult female zebra finches courtesy of Dr Karen Spencer, University of St Andrews, Scotland: Oviduct (magnum, isthmus shell gland), ovarian stroma, small intestine, duodenal loop, gizzard, lung, spleen, breast muscle, skin, liver, kidney and heart.

After dissection all tissue was placed in RNA later (Ambion) and stored at 4°C overnight before storage at -80°C. Samples weighed no more than 100mg.

Time of oviposition

Magnum tissue was obtained from sexually mature white leghorn hens with an ovum at various positions in the oviduct, see Gong *et al* [2] for details. Briefly, magnum tissue was collected either when the egg was in the magnum (n=5), in the shell gland where the stage of calcification was determined by electron microscopy and recorded as early (n=8), mid (n=9) and late (n=10) calcification or during a pause day (n=11) when there was no evidence of ovulation.

Effect of oviduct development

To determine if *Gallus gallus OvoDA1* (*gallin)* and *Gallus gallus OvoDB1* expression differed between in lay hens (n=11) and those where the oviduct had regressed due to incubation behavior (n=11) magnum and shell gland tissue were collected from hens of a Silkie x White

Leghorn cross, see Whenham, *et al* [9]. After dissection tissue was frozen in liquid N_2 and stored at -80°C. Samples weighed no more than 100mg.

Induction with steroid hormones

The protocol for examining induction of *Gallus gallus OvoDA1* (*gallin)* with steroid hormones was adapted from a method described by Kunnas, *et al* [21]. Three week old ISA brown chicks (n=60) were given an intramuscular injection of 0.5mg diethylstilbestrol (DES; Sigma-Aldrich D4268)) in 0.5 ml propylene glycol (vehicle) daily for seven days (primary stimulation) and then split into two groups. Ten days following the primary DES treatment one group of birds (n=30) were re-stimulated on two consecutive days with DES (primed group) and group two (non-primed) birds (n=30) received vehicle. The next day within both groups a subset were given a single intramuscular injection of progesterone (Sigma-Aldrich P3972, 20mg/kg) (n=10), estradiol (Sigma-Aldrich E8515, 10mg/kg) (n=10) or vehicle (propylene glycol; 1ml/kg) (n=10). All chicks were killed 12-16 hours after the single injection; magnum tissue was removed and immediately frozen in liquid N_2 then stored at -80°C.

Ethics: All animals were killed in accordance with Schedule 1 of the Animals (Scientific Procedures) Act 1986, UK under project license PPL 60/3964. All procedures were approved by the institute's ethics committee.

RNA preparation

Tissues were homogenized in Ultraspec II total RNA isolation reagent (AMS Biotechnology) using an Ultraturrax homogenizer (IKA®-Werke GmbH & Co. KG). Samples were then processed as per manufacturer's protocol.

RT-PCR transcript detection and sequencing characterization

1µg samples of chicken, duck, turkey and zebra finch magnum RNA were reverse transcribed using a high capacity cDNA reverse transcription kit (Applied Biosystems) according to the manufacturer's protocol. Primers (Table 2) were designed to ensure complete coverage of each ovodefensin's coding sequence. PCR amplification was performed using the following conditions: an initial denaturation at 95°C for 4 minutes, followed by 40 cycles of 30s at 95°C, 30s at 58°C and 30s at 72°C, followed by an extension of 7 minutes at 72°C. All products were separated by 2% agarose-gel electrophoresis and visualized using SYBR Safe DNA gel stain (Invitrogen).

The amplified PCR fragments were sequenced with their respective forward and reverse primers. Sequences were assembled by Staden [22] to produce consensus sequences.

5'/3' RACE

Rapid amplification of cDNA ends (RACE) was carried out using a 5'/3' RACE 2nd Generation kit (Roche Diagnostics) to determine the number of exons encoding *Gallus gallus OvoDA1* (*gallin*). Briefly, for 5'RACE, synthesis of first strand cDNA was carried out on magnum RNA using primer gallinSP1 and the mRNA template degraded. cDNA was purified using a High Pure PCR Product Purification kit (Roche Diagnostics) and polyA tailed at the 3' end. The tailed cDNA was amplified by PCR using the oligo (dt)-anchor primer provided and a further nested primer OvoDA1SP2. The product from this PCR was run using 3% agarose-gel electrophoresis and visualized using SYBR Safe DNA gel stain (Invitrogen). PCR product was excised from the gel and cleaned from the agarose using a QIAEX II Gel Extraction Kit (Qiagen) and sequenced using primer OvoDA1SP2.

3'RACE used the oligo(dT)-anchor primer to initiate cDNA synthesis at the poly(A)-tail of magnum RNA. Amplification using a PCR anchor primer and OvoDA1SP5 was then performed directly. PCR product was excised from the gel and cleaned from the agarose using a QIAEX II Gel Extraction Kit (Qiagen) and sequenced using primer OvoDA1 SP5. Primer sequences for RACE can be found in Table 2.

Reverse transcription quantitative polymerase chain reaction (RT-QPCR).

A first strand cDNA synthesis kit (GE Healthcare Life Sciences) was used for reverse transcription of a 0.5µg sample of total RNA as per manufacturer's instructions. Reverse transcribed samples were diluted by a factor of 10 with H_2O prior to use. Primer3 was used to design primers for quantification, sequences can be found in Table 2. RT-QPCR was carried out with 8µl of the diluted cDNA and a primer concentration of 20mM according to Agilent Brilliant III SYBR® Green QPCR master mix (Agilent) instructions. The following conditions were used for RT-QPCR; 95°C for 2 min, then 40 cycles of 95°C for 15s, 60°C for 30s using an MX3000 (Stratagene). Reactions containing no template were run as a control. Standard PCR conditions were used to obtain products for the construction of a standard curve and Lamin B-receptor (LBR) expression was measured in the same way to normalize concentrations (as described previously) [9] in order to determine the absolute concentration of the different ovodefensin transcripts. Products were run on an agarose gel to confirm only products of the correct length with no primer-dimer were amplified as well as ensuring that there was only a single peak dissociation curve, correct amplification was also confirmed through sequencing of the PCR product.

Production and titres of polyclonal antibodies

Production of antibodies was carried out by Dundee Cell Products Ltd; epitopes for raising antibodies were chosen on their surface probability, hydrophilicity and antigenicity with consideration of the peptide solubility. Briefly, two rabbits per peptide were immunized four times at three week intervals by intramuscular injection of synthesized *Gallus gallus* OvoDB1 (R108 and R109) epitope (CNKKDEWSFHQ) and *Taeniopygia guttata* OvoDB1 (R112 and R113) epitope (KGEREEHTED) synthesized by Dundee Cell Products Ltd. and emulsified in Freunds adjuvant. Serum was collected after each immunization. Antiserum was purified via a two-step affinity purification using cognate peptides coupled to beads.

To measure the titres of anti-*Gallus gallus* OvoDB1 and anti-*Taeniopygia guttata* OvoDB1 in their respective antisera, the synthesized epitopes were diluted with 50mM Na_2CO_3 (pH9.6) to a final concentration of 1ng/µl (0.5ng/µl of each epitope) and 50µl of the solution was added to each well of a 96 well plate. The plate was covered and stored overnight at 4°C. This was aspirated and the wells incubated for 2 hours at room temperature with 200µl tris-buffered saline (pH7.5), 0.5% Tween 20 (TBST), 1% bovine serum albumin (BSA) to block unsaturated binding sites. Pre-immune (null) sera and antisera were serially diluted 1/1000 to 1/32000 with TBST, 1% BSA, pH7.5. To each well, 10µl of diluted null sera or antisera were added and the plate incubated for 2 hours at room temperature. The plate was washed 5 times with TBST. Horse radish peroxidase (HRP) conjugated Anti-rabbit IgG- (SAPU) diluted 1/2000 with TBST, 1% BSA, pH7.5. 100µl was applied to each well and the plate incubated at room temperature for 2 hours. After five washings with TBST, peroxidase activity was detected by adding 100µl detection solution (100mM citric acid, 200mM Na_2HPO_4, O-Phenylenediamine (OPD), H_2O_2)).

The reaction was stopped with 50µl 2M H_2SO_4 and absorbance (490nm) measured spectrophotometrically.

Immunohistochemistry (IHC)

Wax embedded tissues were sectioned at 3 microns using a Thermoshandon finesse microtome, lifted onto Vectabond slides (Vector Laboratories) and incubated at 60^0C for 1 hour before they were de-waxed and taken down to water. Heat induced epitope retrieval (HIER) was carried out using a Menarini Access Retrieval Unit, in buffer (Sodium Citrate pH6) for 1 minute 40 seconds at 125ºC full pressure. Each section was then loaded onto a Dako Autostainer (Dako). A standard IHC protocol was then applied; optimal staining was achieved at a 1:500 dilution for the polyclonal anti-*Taeniopygia guttata* OvoDB1 (113_KGE _2.1) and 1:1000 for the anti-*Gallus gallus* OvoDB1 (108 CNK-1.3). The sections were viewed using a Leica DM 4000 B microscope and images captured using a Leica DC480 camera with Qwin program for PC (Leica).

Peptide production

Gallus gallus OvoDA1, OvoDB1 and *Anas platyrhynchos* OvoDA1 were commercially obtained from Almac Group (East Lothian, Scotland). The peptides were synthesized on a 0.2mmol scale using an automated Applied Biosystems 433 peptide synthesizer and fluorenylmethoxy (Fmoc) solid phase peptide synthesis protocols. Each amino acid was coupled after activation with diisopropylcarbodiimide/Oxyma pure. Upon completion of the synthesis the peptide was cleaved from the resin and the side chain protecting groups removed using a cocktail of trifluoroacetic acid, ethanedithiol triisopropylsilane and H_2O. The peptide was folded in the

presence of oxidised and reduced glutathione at pH8 and the final product isolated using preparative HPLC and a gradient of H_2O, acetonitrile and 0.1% trifluoroacetic acid. Freeze dried peptides were reconstituted in 10% dimethyl sulfoxide (DMSO).

Antimicrobial activity assay

The antimicrobial assay, adapted from methods described previously [2, 24, 25] was used to determine the efficacy of *Gallus gallus* OvoDA1, OvoDB1 and *Anas platyrhynchos* OvoDB1. Five different strains were used to assess antimicrobial activity: *E. coli* K-12 strain DH5α, avian pathogenic *E. coli* (APEC) O78:H9 strain χ7122, *Salmonella enterica* serovars Enteritidis (SE125109) and Typhimurium (ST4/74) and *Staphylococcus aureus* (8325-4). *E. coli* DH5α and *S. aureus* 8325-4 are non-pathogenic laboratory strains of bacteria and APEC, SE125109 and ST4/74 are pathogenic strains known ability to colonise chickens. Bacteria were cultured overnight at 37°C in Luria broth (LB) or tryptone soya broth (TSB) (*S. aureus*); two hundred and fifty µl of overnight culture was sub-cultured into 20ml of LB or TSB and incubated at 37°C for 3 hours. After the second incubation 20µl of culture was diluted with 2ml of phosphate buffered saline (PBS), pH7.4. Ten µl of *Gallus gallus* OvoDA1, OvoDB1 or *Anas platyrhynchos* OvoDB1 peptide (3, 6, 15, 30, 60, 150, 300 and 600µM) or DMSO (control) or PBS was added to 50µl of diluted culture to produce final concentrations of 0.5, 1, 2.5, 5, 10, 25, 50 and 100µM respectively. For *S.aureus* final concentrations of 10, 25, 50, 100, 150 and 200µM were used. After vortexing this was incubated at 37°C for 3 hours and then the suspensions were serially diluted to $1x10^{-4}$ with PBS, all dilutions were plated on LB or tryptone soya agar plates. Plates were incubated overnight at 37°C and the colonies were counted.

Results are expressed as a reduction in colony forming units per ml (CFU/ml) and where possible the effective dose 50 (ED50) was calculated using the DRC program in R [26].

Effect of pH and salt on antimicrobial activity

The antimicrobial assay as outlined above (2.9.1) was adapted to test the effect of pH on the efficacy of *Gallus gallus* OvoDA1 against *E. coli* DH5α. The assay was carried out as before but using PBS at pH 6.4, 7.4 and 8.4. *Gallus gallus* OvoDA1 was tested at final concentrations of 1.9, 5.6, 16.7 and 50μM.The antimicrobial assay (2.9.1) was adapted to test the effect of salt sensitivity on *Gallus gallus* OvoDA1 efficacy. PBS with NaCl concentrations of 50, 100 and 150mM, pH 6.4 were used.

Sequencing and database submission

All sequencing was carried out by GATC biotech and consensus sequences were submitted to EMBL, *Gallus gallus OvoDB1* (EMBL accession no. LN717248), *Meleagris gallopavo OvoDB1* (EMBL accession no. LN717250), *Taeniopygia guttata OvoDA1* (EMBL accession no. LN717251), *OvoDA1_2* (EMBL accession no. LN717252) and *OvoDB1* (EMBL accession no. LN717253). Putative budgerigar, medium ground finch, anole lizard, american alligator, collared flycatcher, painted turtle and chinese soft-shelled turtle sequences were not submitted because they remain predicted but can be found in the supplementary data (Supplementary Data 1).

Statistical analysis

One way or two way ANOVA and least significant difference to test between the means were used as appropriate for statistical analysis using Genstat (13th edition, VSN International Ltd). Log transformed data was used where appropriate to equalize the variance.

Results

Bioinformatic analysis and transcript confirmation

TBLASTN similarity searches of available genomes located twenty four new ovodefensin homologues (Supplementary Data 1, Table 3). Hierarchical clustering based on the distance between each cysteine identifies six specific sub-families termed OvoDA-OvoDF (Figure 1A). In the chicken the first representative of OvoDB, a cysteine sequence motif with a shorter spacing between C1-C2 and a longer spacing between C4-C5 (C-X3-C-X3-C-X11-C-X4-CC) was identified on chromosome 3, where *OvoDA1* (C-X5-C-X3-C-X11-C-X3-CC) and the classical chicken beta defensins are also located. This was named *OvoDB1* in accordance with the proposed nomenclature outlined in 2.1.2 (Figure 1A, C, Table 3). The first turkey representative of the OvoDB1 motif (Figure 1A, C, Table 3) was discovered on chromosome 2 the same chromosome as *Meleagris gallopavo OvoDA1* (*meleagrin*). A further potential paralog of *Meleagris gallopavo OvoDA1* was also located on this chromosome; the mature peptide sequence shares a 95% identity with *OvoDA1* and was named *Meleagris gallopavo OvoDA2* (Figure 1A, C, Table 3). A budgerigar representative of sub-family B, *Melopsittacus undulatus OvoDB1* (Figure 1A, C, Table 3) was discovered during this analysis however a representative from the same subfamily (A) as *gallin* has yet to be located for this species. Three putative copies of a sub-family A ovodefensin were found in the April 2012 assembly of the medium ground finch genome; *Geospiza fortis OvoDA1*, *OvoDA1_2* and *OvoDA1_3* (Figure 1A, C, Table

3). A search of the flycatcher genome discovered a member from both the A and B sub-families, these were named *Ficedula albicollis OvoDA1* and *OvoDB1* respectively (Figure 1A, C, Table 3). For the first time reptile representatives of the ovodefensin family were identified. Two copies of *Pelodiscus sinensis OvoDB1* (Figure 1A, D, Table 3) were identified in the Chinese soft-shell turtle genome. In the painted turtle genome a sub-family B member was also located; *Chrysemys picta bellii OvoDB1* (Figure 1A, D, Table 3) in addition to this a new cysteine spacing motif (C-X3-C-X3-C-X13-C-X3-CC), which we have termed sub-family C was uncovered. Within this the painted turtle sub-family C contains two copies of the *OvoDC1* gene and an *OvoDC2* gene (Figure 1A, D, Table 3). TBLASTN searches of the American alligator genome identified two members of yet another new cysteine sequence motif, sub-family D (C-X3-C-X3-C-X11-C-X3-CC); *Alligator mississippiensis OvoDD1* and *OvoDD2* (Figure 1A, D, Table 3). Finally two further motifs were uncovered in the anole lizard genome; sub-family E (C-X6-C-X3-C-X11-C-X2-CC) and F (C-X6-C-X4-C-X11-C-X2-CC). Within sub-family E one copy of *Anolis carolinensis OvoDE1* was located, three copies of *OvoDE2* and one copy of *OvoDE3* (Figure 1A, D, Table 3). One copy of each of the sub-family F members was identified, *OvoDF1*, *OvoDF2* and *OvoDF3* (Figure 1A, D, Table 3). Genome build and chromosome/scaffold locations for each ovodefensin are outlined in Tables 1 and 3.

Three homologs not previously classified as ovodefensins were also identified during this study. Shapiro *et al* [27] produced a rock pigeon reference genome from which a putative sub-family A member was identified. This had been named cygnin due to its homology with black swan cygnin, we now propose the name *Columba livia OvoDA1* (Figure 1A, C, Table 3). From this genome a putative sub-family B member, named small basic protein was also located, the name *Columba livia OvoDB1* (Figure 1A, C, Table 3) is proposed for this ovodefensin homolog. In

addition to this TEWP, a loggerhead turtle peptide was isolated from egg white and shown to be a defensin [28], this peptide has the motif of a sub-family B ovodefensin and we therefore suggest the name *Caretta caretta OvoDB1*(Figure 1A, C, Table 3).

5'/3' RACE

5'/3' RACE using magnum RNA and primers resulted in a PCR product with 100% identity to the published *Gallus gallus OvoDA1* (*gallin*) sequence (ENSGALG00000028311.1), the 5' and 3' sequence of which was derived by prediction. Alignment with the chicken genome confirmed that *Gallus gallus OvoDA1* is encoded for by two exons.

Tissue expression

Gallus gallus OvoDB1 expression was restricted to magnum and isthmus as previously documented with *OvoDA1* (Figure 2). In contrast to this both *Anas platyrhynchos OvoDA1* and *OvoDB1* expression was greatest in the shell gland, although some expression was observed in both the magnum and the isthmus (Figure 2) and no expression was detected outside of the oviduct (data not shown). The combined expression of *Meleagris gallopavo OvoDA1* and *OvoDA1_2* and expression of *OvoDB1* was high across all oviduct tissues, the greatest expression for these turkey ovodefensins was measured in the shell gland, magnum and isthmus (Figure 2). *Taeniopygia guttata OvoDA1* and *OvoDB1* expression was also restricted to the oviduct with expression being observed most highly in the shell gland for both genes (Figure 2). It should be noted that vagina tissue was not available for the zebra finch. For all the ovodefensin transcripts analyzed, no expression was detected in tissues outside the reproductive tract (data not shown).

There was no significant effect of the position of the egg in the oviduct or the lack of an egg in the oviduct on *Gallus gallus OvoDB1* expression in the magnum (Figure 3) (ANOVA = P=0.269). The level of *Gallus gallus OvoDA1* and *OvoDB1* expression in the magnum was higher (P<0.001) in birds in-lay than in those with an oviduct in the regressed state (Figure 4). When estrogen and progesterone were administered to juvenile hens *Gallus gallus OvoDA1* expression (Figure 5) was higher in birds treated with the two steroids (P<0.001) and where priming with an estrogenic compound had been performed overall expression increased (P<0.001).

Immunohistochemistry

The chicken anti-OvoDB1 antiserum (108 CNK-1.3) produced positive staining in the tubular gland cells of the magnum (Figure 6A, C). The ciliated and non-ciliated cells lining the magnum region of the oviduct did not react to the primary antibody. No staining was observed in the isthmus, shell gland or caecum (data not shown). In contrast to this the zebra finch anti-OvoDB1 antiserum (113_KGE _2.1) produced positive staining in the tubular gland cells and surface epithelium of the magnum, isthmus and shell gland (Figure 7A, C, E). No convincing staining was observed in the breast muscle (Figure 7G).

Antimicrobial activity

Gallus gallus OvoDA1 (gallin) peptide achieved a relatively large effect on *E. coli* DH5α with around a 98% reduction in CFU/ml at 100μM (Figure 8A, Table 4), this is comparable to the results reported previously with BL21 [2]. *Gallus gallus* OvoDA1 achieved around a 40% reduction in viability of avian pathogenic *E. coli* O78:H9 strain χ7122 when used at 50 or

100μM (Figure 8A, Table 4) and a >90% reduction in viability against *S. aureus* 8325-4 a Gram-positive organism when used at 100 or 200μM (Figure 8B, Table 4). No reduction was observed against either of the *Salmonella* strains used in this study (data not shown). A 35% reduction in viability of *E. coli* DH5α was achieved with 100μM *Gallus gallus* OvoDB1 (Table 4), and a very small reduction was observed in APEC numbers (~3%). No reduction was observed against *S. aureus* the *Salmonella* strains tested (data not shown). A duck representative of the ovodefensin family, *Anas platyrhynchos* OvoDB1 ($dBPS_2$) showed a >80% reduction in viability of *E. coli* DH5α (Table 4) but no convincing activity was seen against any of the other strains tested.

Gallus gallus OvoDA1 potency as measured by ED50 did not differ significantly due to pH (P=0.42), the ED50 (± SEM) of *Gallus gallus* OvoDA1 at pH 6.4, 7.4 and 8.4 was 7.38μM (±15.3), 3.57μM (±1.4) and 3.67μM (±0.96) respectively (Figure 9A). The potency of *Gallus gallus* OvoDA1 at ED50 level was also not effected by salt concentration (P=0.49), the ED50 of *Gallus gallus* OvoDA1 at 50, 100 and 150mM NaCl was 3.55μM (±2.39), 3.88μM (±29.8), and 3.27μM (±14.9) respectively (Figure 9B). However at higher concentrations of *Gallus gallus* OvoDA1 the maximum antibacterial activity appeared to be diminished by higher salt concentration (Figure 9B).

Discussion

Seven ovodefensins had previously been identified within divergent avian species and it had been observed that two cysteine sequence motifs exist [2], suggesting within the ovodefensins there may be further divisions in structure and function (Table 1). This study identified a further 25 ovodefensin members (Table 3) through genome analysis, and attributed a further 3

previously known sequences to the group expanding the ovodefensin family to include reptile species for the first time. It was first suggested by Gong *et al.* that the ovodefensins appeared to be a new family of defensin molecules, most likely belonging to the β-defensin group [2]. However it was predicted that ovodefensins share the genomic organization of mammalian β-defensins, being encoded for by two exons [2] rather than the four exons which encode for classical avian defensins [6]. This study confirmed the exon arrangement of *Gallus gallus OvoDA1* by defining the transcription start site thus confirming the two exons predictions. Where available, predictions of the newly discovered avian and reptilian ovodefensins also support the 2 exon arrangement. In addition to an altered exon arrangement the ovodefensins differ to classical avian defensins in the spacing within the cysteine sequence motif. When the 3D structure and specific cysteine pairing was solved by Hervé *et al* it was observed that *Gallus gallus* OvoDA1 (gallin) had the classical disulphide pairing (C1-C5, C2-C4, C3-C6) of β-defensins. It was however demonstrated that *Gallus gallus* OvoDA1 contained an additional two stranded parallel β-sheet compared to known beta-defensin structures, and no amino-terminal helix [7], it is therefore possible that the ovodefensins form a new structural sub-family of defensins. Now phylogenetic analysis of both the spacing and sequence of the ovodefensins suggests that multiple sub-families may exist within the family (Figure 1A, C, D). These sub-families include four new ovodefensin-like motifs OvoDC (C-X3-C-X3-C-X13-C-X3-CC), OvoDD (C-X3-C-X3-C-X11-C-X3-CC), OvoDE (C-X6-C-X3-C-X11-C-X2-CC) and OvoDF (C-X6-C-X4-C-X11-C-X2-CC) revealing that the ovodefensin family is much more diverse than was first thought. All four new cysteine spacing motifs appear to be reptile specific (Figure 1A, D) whereas of the originally identified motifs OvoDA, (C-X5-C-X3-C-X11-C-X3-CC) was

avian specific (Figure 1A, C) and OvoDB (C-X3-C-X3-C-X11-C-X4-C) contained both avian and reptilian counterparts (Figure 1A, C, D).

Phylogenetic analysis of spacing (Figure 1A) suggests that a common ancestor gave rise to two progenitor molecules which appear in turn to have independently evolved three cysteine sequence motifs. However if you view this in the context of sequence analysis (Figure 1B) it would appear that the OvoDB cysteine sequence motif has in fact evolved separately within both avian and reptilian lineages suggesting spacing is an important feature of the ovodefensin molecules. In a sense this seems like a form of convergent evolution, albeit the available repertoire of spacing is relatively limited. In order to assess all the ovodefensin molecules together either spacing (Figure 1A) or a core region (Figure 1B) was used however we recognize that this may influence the outcome of analysis depending on the evolutionary constraints of each selection. In particular analysis based on the core region of the molecule (Figure 1B) produced some unexpected outliers, for example within the OvoDA cluster we see that the turkey molecules appear closer to the duck and swan than the chicken. However when the full length peptides of the OvoDA sub-family are analyzed separately (Figure 1E) a more classical species arrangement is observed. This demonstrates the complex nature of what appears to be the co-evolution of sequence and spacing and the need for both aspects to be studied in combination. Overall the evolution within birds, and now in this study within another egg laying clade the reptiles, of a large repertoire of peptides that not only vary in sequence but the spacing between conserved cysteine residues suggests that the spacing between these cysteine residues may be critical for function. The region between the 1st and 2nd cysteine which varies from 3-6 amino acids in length for example has been demonstrated in *Gallus gallus* OvoDA1 to be important because of the basic residue in an otherwise hydrophobic region [7]. Interestingly the

spacing between the 3rd and 4th cysteines residues is constant (X11) in each ovodefensin sub-family with the exception of OvoDC which has 13 residues in this region. Spacing between all cysteine residues is quite variable between ovodefensin sub-families, indeed it forms the basis of the proposed nomenclature however it should be noted that as this is an otherwise unvarying region and OvoDC has so far only been identified in reptiles we may be observing a distinct ovodefensin-like family of peptides. It has been difficult to find examples in the literature of situations where the amino acid distance between conserved residues of a motif are altering or under selection, possibly because this is rare or possibly because the methods of finding homologous genes rely heavily on the conservation of sequence, not pattern recognition. Some flexibility is observed within the general sequence of vertebrate β-defensins with a relaxed consensus of C-X(4-8)-C-X(3-5)-C-X(9-13)-C-X(4-7)-CC being observed [29] yet TBLASTN searches with these β-defensins do not identify ovodefensin genes. A study by Maxwell *et al* used hidden Markov models to identify novel mouse and human β-defensin genes and suggested that murine and human β-defensin families could be divided into two sub-groups based on the strong sequence conservation of exon 1 [10]. Regions that vary in their spacing between residues however are within the mature peptide and are encoded for by exon 2; this exon shows substantial divergence which is consistent with rapidly evolving genes under positive selection [10]. Other approaches have been developed to get round the problem in aligning large proteins where there were differences in spacing between conserved features important for protein secondary structure [30] and these alignments appeared to have a better agreement with the accepted view of evolution than if this was not undertaken. In large proteins this has been used to demonstrate evolution by insertion of new domains in molecules which have effects on structure and the authors conclude that using structure is likely to be more robust than sequence when

molecules cannot be unambiguously aligned [31]. Whilst appreciating that this is not comparable directly to the situation described in this paper of a short peptide, it does offer an analogy that may be useful in terms of understanding the ways in which evolution can work on gene products. In the case of the relatively small ovodefensins it also seems that this is the case.

This study demonstrated very different levels of activity between the ovodefensin peptides tested. However both *Gallus gallus* and *Anas platyrhynchos* OvoDA1 are from the same sub-family suggesting that spacing alone does not define direct antimicrobial activity. Another possibility for differing levels of activity is charge which varies from +4 to +10 within the currently identified ovodefensin members. However charge alone is unlikely to explain all of the difference in activity as in this study *Gallus gallus* OvoDA1 and *Anas platyrhynchos* OvoDA1 had a very similar charge of +7 and +6 respectively yet differed greatly in their activity. Additionally an increased charge of +10 did not result in increased activity for *Gallus gallus* OvoDB1.

RT-qPCR analysis on chicken, duck, turkey and zebra finch representatives of the ovodefensin family demonstrate that in all cases expression is restricted to the oviduct of the bird. However, interestingly, levels and patterns of expression within the oviduct vary between genes and species. As previously seen with *Gallus gallus OvoDA1* [2], *Gallus gallus OvoDB1* was expressed more highly in the magnum of the oviduct where the egg white is formed (Figure 2), however expression of *Gallus gallus OvoDA1* was more than 40 times that of *Gallus gallus OvoDB1* (Figure 4). In contrast to this both the duck and zebra finch ovodefensins were expressed most highly in the shell gland and the turkey ovodefensins had very high levels of expression in the magnum, isthmus and shell gland regions of the oviduct (Figure 2). Immunohistochemistry confirmed the expression of *Gallus gallus* OvoDB1 (Figure 6) and

Taeniopygia guttata OvoDB1 (Figure 7) at a peptide level in specific regions of the oviduct. *Gallus gallus* OvoDB1 peptide distribution was restricted to the tubular gland cells of the magnum suggesting it is secreted into the egg white as previously seen with *Gallus gallus* OvoDA1[2]. In contrast *Taeniopygia guttata* OvoDB1 was expressed in both the tubular gland cells and surface epithelium of the magnum, isthmus and shell gland suggesting this may play a greater role in local protection of the oviduct as well as the innate defense of the egg.

Gallus gallus OvoDB1 expression in the magnum of the oviduct did not differ significantly in relation to the position of the egg in the oviduct at the time of sampling (Figure 3); this was also the case for *Gallus gallus OvoDA1* [2]. This expression profile is typical of egg proteins such as TENP [9] or ovalbumin [8]. Both *Gallus gallus OvoDA1* and *OvoDB1* expression in the magnum was significantly higher when the hen is in-lay compared to when the oviduct is regressed (Figure 4), suggesting that these genes are under the control of gonadal steroids, therefore being specifically up-regulated during egg production when steroid levels are elevated [32]. This observation was confirmed for *Gallus gallus OvoDA1* by measurement of expression after the administration of estrogen and progesterone to juvenile hens, with estrogen producing the largest increase in expression. The increase in *Gallus gallus OvoDA1* expression was greatest when the hens had previously been primed with an estrogenic compound, showing the synergistic activity of estrogen and progesterone expected of a gene controlled by gonadal steroids in the oviduct [21, 33]. We cannot conclude if this is a direct effect of steroids on the promoter or an indirect effect and indeed no classical ERE were observed in the proximal promoter.

In support of an egg defense role, *Gallus gallus* OvoDA1 was previously shown by Gong *et al* to be antimicrobial against a strain of *E. coli* [2], an observation that was confirmed in another three

strains of *E. coli* by Hervé *et al* [7]. The data outlined in this study also demonstrated antimicrobial activity of *Gallus gallus* OvoDA1 against laboratory-adapted and pathogenic *E. coli* (Figure 8A), and in agreement with the former publication [7], found no activity against *Salmonella* serovars Enteritidis and Typhimurium. However this study did demonstrate antimicrobial activity of *Gallus gallus* OvoDA1 against *S. aureus* (Figure 8B). Although it had been previously documented that *Gallus gallus* OvoDA1 did not possess antimicrobial activity against *S. aureus* [7] both the method (solid vs liquid phase) used to measure antimicrobial activity and the strain of *S. aureus* differ. This is the first time activity has been recorded for a Gram-positive organism with an ovodefensin, and indeed for an organism other than *E. coli* demonstrating the need for more in depth analysis of the spectrum of activity of the ovodefensins.

For the first time this study examined the antimicrobial activity of two other avian members of the ovodefensin family, *Gallus gallus* OvoDB1 and *Anas platyrhynchos* OvoDA1. *Gallus gallus* OvoDB1 represents a member of the sub-family B cysteine sequence motif (C-X3-C-X3-C-X11-C-X4-CC) whereas *Anas platyrhynchos* OvoDA1 contains the same sub-family A cysteine sequence motif as *Gallus gallus* OvoDA1 (C-X5-C-X3-C-X11-C-X3-CC). Although *Gallus gallus* OvoDB1 possessed antimicrobial activity against *E.coli* DH5α (35% reduction in CFU/ml) it was not as potent as OvoDA1 from the same species at the same concentration (100μM) and no activity was recorded against APEC, *S. aureus* or either of the *Salmonella* strains tested in this study. *Anas platyrhynchos* OvoDA1 demonstrated good activity against *E.coli* DH5α (>80% reduction) yet no activity was recorded against any of the other strains used in this study. It should be noted that a study was previously carried out on duck ovodefensins which observed no antimicrobial activity [34], however this study did not report on the method

used to assess activity or whether the bacterial strains used were laboratory or pathogenic. We were only able to observe activity for *Anas platyrhynchos* OvoDA1 (dBPS2) against the laboratory strain of *E. coli* are therefore the strains used in the previous study would be of interest. The antimicrobial results from this study suggest that the family is as diverse in its activity as it is in sequence and raises the question of whether the sequence or perhaps spacing within the cysteine sequence motifs of these molecules affects their ability to kill microorganisms, perhaps evolving to counter the specific challenges each organism faces.

Overall *Gallus gallus* OvoDA1 with motif C-X5-C-X3-C-X11-C-X3-CC was most active in the assay used and was therefore assessed in further detail, but neither pH or salt significantly affected the ED50 in the range used in this study (Figure 9A). Overall potency of the peptide did diminish slightly at pH 6.4. The pH of egg white ranges from approximately 7.6-8.5 immediately after lay and becomes increasingly more alkaline (~9.6) as CO2 is lost [35]; this natural variation in pH may account for the apparent insensitivity of *Gallus gallus* OvoDA1 to the pH range used in this study. Although the ED50 value did not differ significantly between salt concentrations it is clear that salt concentration did significantly affect the maximal potency of the peptide (Figure 9B). Salt sensitivity of classical mammalian defensins such as mouse β-defensin 1 [36] has been well documented and it is postulated that this is a feature common to all defensin molecules [37]. Salt concentration in the egg is relatively low (~50mM) [38] and the data presented in this paper would suggest that ovodefensins are relatively unaffected by increased salt but share with defensin a sensitivity to salt concentration.

If ovodefensins behave as some classical vertebrate defensins it is possible to speculate that both the direct ability to kill microorganisms and also the ability to modulate the immune system may be dependent on the structure of the molecule. Further investigation will be required to

determine whether the sub-families possess differing antimicrobial activity related to characteristics such as inter cysteine spacing, sequence or charge because the naturally occurring peptides tested have too many factors varying between them to draw a conclusion about each factor in isolation. However the large natural diversity within the ovodefensin family makes it an interesting group to study in order to understand which ovodefensin properties are important for function and may provide a resource for novel antimicrobials as well as aiding understanding when engineering new, more potent derivatives, something of great value as the threat of antibiotic resistance intensifies.

So to conclude, the ovodefensins are avian and reptilian specific members of the β-defensin family. Expression of avian members has been demonstrated to be restricted to the oviduct and in the chicken is up-regulated in laying hens versus those with a regressed oviduct suggesting gonadal steroids control expression. This is confirmed by the increased expression of *Gallus gallus OvoDA1* after administration of estrogen and progesterone in juveniles. Although the chicken ovodefensins show the classic signature of an egg specific gene, the pattern across the range of species examined is more of an oviduct specific family. This coupled with the antimicrobial activity demonstrated in this study and others suggests that ovodefensins have specifically evolved for a role in egg defense as a component of the eggs innate chemical defense; however they may also contribute to maintaining sterility in the oviduct through local tissue activity. There is a large diversity within the ovodefensin family, with six motifs relating to spacing of the conserved cysteines discovered so far. This suggests that evolution is acting not only on amino acid sequence but also spacing of the molecule. This novel finding offers an additional avenue of investigation for the design of new antimicrobial compounds.

References

1. Kovacs-Nolan J, Phillips M, Mine Y. Advances in the value of eggs and egg components for human health. Journal of Agricultural and Food Chemistry 2005; 53:8421-8431.
2. Gong D, Wilson PW, Bain MM, McDade K, Kalina J, Herve-Grepinet V, Nys Y, Dunn IC. Gallin; an antimicrobial peptide member of a new avian defensin family, the ovodefensins, has been subject to recent gene duplication. BMC Immunology 2010; 11:Article No.: 12.
3. Mann K. The chicken egg white proteome. Proteomics 2007; 7:3558-3568.
4. Odani S, Koide T, Ono T, Takahashi Y, Suzuki J. Covalent structure of a low-molecular-mass protein, meleagrin, present in a turkey (Meleagris-gallopavo) ovomucoid preparation. Journal of Biochemistry 1989; 105:660-663.
5. Naknukool S, Hayakawa S, Sun Y, Ogawa M. Structural and physicochemical characteristics of novel basic proteins isolated from duck egg white. Bioscience Biotechnology and Biochemistry 2008; 72:2082-2091.
6. van Dijk A, Veldhuizen EJA, Haagsman HP. Avian defensins. Veterinary Immunology and Immunopathology 2008; 124:1-18.
7. Herve V, Meudal H, Labas V, Rehault Godbert S, Gautron J, Berges M, Guyot N, Delmas AF, Nys Y, Landon C. 3D NMR structure of hen egg gallin (chicken ovo-defensin) reveals a new variation of the beta-defensin fold. J Biol Chem 2014.
8. Muramatsu T, Hiramatsu H, Okumura J. Ovalbumin messenger-RNA in the magnum of laying hens changes during egg formation cycle. British Poultry Science 1994; 35:457-461.
9. Whenham N, Wilson PW, Bain MM, Stevenson L, Dunn IC. Comparative biology and expression of TENP, an egg protein related to the bacterial permeability-increasing family of proteins. Gene 2014; 538:99-108.
10. Maxwell AI, Morrison GM, Dorin JR. Rapid sequence divergence in mammalian beta-defensins by adaptive evolution. Molecular Immunology 2003; 40:413-421.
11. Hancock REW. Peptide antibiotics. Lancet 1997; 349:418-422.
12. Hancock REW, Lehrer R. Cationic peptides: a new source of antibiotics. Trends in Biotechnology 1998; 16:82-88.
13. Hancock REW, Sahl HG. Antimicrobial and host-defense peptides as new anti-infective therapeutic strategies. Nature Biotechnology 2006; 24:1551-1557.
14. ENSEMBL GENOME BROWSER: [Internet]. http://www.ensembl.org. Accessed 23 August 2013.
15. ENSEMBL PRE-RELEASE GENOME BROWSER: [Internet]. http://pre.ensembl.org. Accessed 23 August 2013.
16. CHICKEN (*GALLUS GALLUS*) GENOME BROWSER GATEWAY: [Internet]. http://genome.ucsc.edu/. Accessed 23 August 2013.
17. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. BLAST plus : architecture and applications. Bmc Bioinformatics 2009; 10.
18. UNIPROT: [Internet]. http://www.uniprot.org/. Accessed 23 August 2013.
19. R: [Internet]. http://www.r-project.org/. Accessed 10 May 2014.
20. MEGA5: [Internet]. http://www.megasoftware.net. Accessed 24 June 2014.

21. Kunnas TA, Joensuu TK, Viitala KK, Sopanen P, Tuohimaa P, Kulomaa MS. Induction of avidin messenger-ribonucleic-acid in the chick oviduct by progesterone and other steroids. Endocrinology 1992; 130:3421-3426.
22. Staden R. The Staden sequence analysis package. Molecular Biotechnology 1996; 5:233-241.
23. PRIMER3: [Internet]. http://frodo.wi.mit.edu/primer3/. Accessed 23 August 2013.
24. Nagaoka I, Hirota S, Yomogida S, Ohwada A, Hirata M. Synergistic actions of antibacterial neutrophil defensins and cathelicidins. Inflammation Research 2000; 49:73-79.
25. Townes CL, Michailidis G, Nile CJ, Hall J. Induction of cationic chicken liver-expressed antimicrobial peptide 2 in response to Salmonella enterica infection. Infection and Immunity 2004; 72:6987-6993.
26. BIOASSAY: [Internet]. http://www.bioassay.dk/. Accessed 24 June 2014.
27. Shapiro MD, Kronenberg Z, Li C, Domyan ET, Pan HL, Campbell M, Tan H, Huff CD, Hu HF, Vickrey AI, Nielsen SCA, Stringham SA, et al. Genomic Diversity and Evolution of the Head Crest in the Rock Pigeon. Science 2013; 339:1063-1067.
28. Chattopadhyay S, Sinha NK, Banerjee S, Roy D, Chattopadhyay D, Roy S. Small cationic protein from a marine turtle has beta-defensin-like fold and antibacterial and antiviral activity. Proteins-Structure Function and Bioinformatics 2006; 64:524-531.
29. Xiao YJ, Hughes AL, Ando J, Matsuda Y, Cheng JF, Skinner-Noble D, Zhang GL. 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.
30. Zhu ZY, Sali A, Blundell TL. A variable gap penalty-function and feature weights for protein 3-D structure comparisons. Protein Engineering 1992; 5:43-51.
31. Jiang H, Blouin C. Insertions and the emergence of novel protein structure: a structure-based phylogenetic study of insertions. BMC Bioinformatics 2007; 8.
32. Lague PC, Vanthienhoven A, Cunningham FJ. Concentrations of estrogens, progesterone and lh during ovulatory cycle of laying chicken gallus-domesticus. Biology of Reproduction 1975; 12:590-598.
33. Omalley BW, McGuire WL, Korenman SG. Estrogen stimulation of synthesis of specific proteins and RNA polymerase activity in immature chick oviduct. Biochimica Et Biophysica Acta 1967; 145:204-&.
34. Naknukool S, Hayakawa S, Ogawa M. Multiple Biological Functions of Novel Basic Proteins Isolated from Duck Egg White: Duck Basic Protein Small 1 (dBPS(1)) and 2 (dBPS(2)). Journal of Agricultural and Food Chemistry 2011; 59:5081-5086.
35. HEATH JL. Chemical and Related Osmotic Changes in Egg Albumen During Storage. Poultry Science 1977; 56:822-828.
36. Bals R, Goldman MJ, Wilson JM. Mouse beta-defensin 1 is a salt-sensitive antimicrobial peptide present in epithelia of the lung and urogenital tract. Infection and Immunity 1998; 66:1225-1232.
37. Ganz T. Defensins: Antimicrobial peptides of innate immunity. Nature Reviews Immunology 2003; 3:710-720.
38. Bell DJ, Freeman BM. Physiology and biochemistry of the domestic fowl. Academic Press; 1984.

39. Saitou N, Nei M. The neighbor-joining method - a new method for reconstructing phylogenetic trees. Molecular Biology and Evolution 1987; 4:406-425.
40. Felsenstein J. Confidence-limits on phylogenies - an approach using the bootstrap Evolution 1985; 39:783-791.
41. Zuckerkandl E, Pauling L. Evolutionary divergence and convergence in proteins. 1965.
42. 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. Molecular Biology and Evolution 2011; 28:2731-2739.
43. Simpson RJ, Morgan FJ. Isolation and complete amino-acid-sequence of a basic low-molecular weight protein from black swan egg-white. International Journal of Peptide and Protein Research 1983; 22:476-481.

Figure 1: Evolutionary relationship of ovodefensin homologues. A distance matrix and hierarchical clustering of all known ovodefensin members based on spacing between cysteine residues (A) identifies six specific sub-families termed OvoDA-OvoDF. A phylogram indicating the evolutionary history of sub-families OvoDA, OvoDB, OvoDD and OvoDE was inferred using the Neighbor-Joining method [39] from the amino acid sequence of a core region from the conserved glycine residue until the fourth cysteine (B). This suggested that the spacing with the cysteine sequence motif of OvoDB evolved independently in avian and reptilian lineages and which are presented separately (C) and (D). Mature peptide sequences from sub-families OvoDA (E) and OvoDB (F) were analyzed individually in order to compare the full length of the molecule.

For B, E and F the percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) are shown next to the branches [40]. The branch lengths are proportional to the evolutionary distances which were computed using the Poisson correction method [41]. The units are the number of amino acid substitutions per site. All positions containing gaps and missing data were eliminated. Evolutionary analyses were conducted in MEGA5 [42]. A key for the abbreviations of the species Latin name can be found in supplementary data 3 and the sequences the alignments are based on in supplementary data 1.

Figure 2: Expression of ovodefensin mRNA in a range of adult chicken (n=4), turkey (n=4), duck (n=3) and zebra finch (n=5) tissues measured by RT-QPCR (mean ± sem). Expression was corrected for chicken, turkey, duck or zebra finch LBR expression to normalize

for any differences between tissues. Note to accommodate the large differences in expression the data are presented on the log scale.

Figure 3: Expression of *Gallus gallus OvoDB1* mRNA in magnum tissue at different stages of egg formation measured by RT-QPCR (n=8, mean ± sem). *Gallus gallus OvoDB1* expression was corrected for LBR expression. Pause represents a day when the hen did not ovulate so no egg is present. Magnum represents tissue taken when an egg is present in the magnum and early, mid and late describes the stage of shell formation in the shell gland and indicates the egg was not in the magnum when the sample was taken. ANOVA, P=0.269.

Figure 4: Expression of A) *Gallus gallus OvoDA1* and B) *Gallus gallus OvoDB1* in magnum tissue of laying (L) and out of lay (NL) birds measured by RT-QPCR (n=11, mean ± sem). Non-laying hens were those where the oviduct had regressed due to the withdrawal of gonadotrophic support with the onset of incubation behavior. Expression was corrected using LBR gene expression. Note the large difference in the Y-axis scale between A and B. Significance between laying state is indicated at P<0.001, (***).

Figure 5: Expression of *Gallus gallus OvoDA1* mRNA in the magnum of juvenile chicks treated with steroids measured by RT-QPCR (n=10, mean ± sem). *Gallus gallus OvoDA1* expression was corrected using LBR expression. Female chicks at 3 weeks of age were either primed with diethylstilbestrol or vehicle then subsequently treated with either progesterone, estrogen or vehicle. ANOVA indicated priming was significant at P<0.001; steroid treatment at

P<0.001. Significance between primed state within treatment is indicated, P<0.001(***) and significance between treatments regardless of primed state between the brackets, P<0.001(***).

Figure 6: Only the tubular gland cells of the magnum (A and C) region of the oviduct convincingly stained positive with the chicken anti-OvoDB1 antisera (108 CNK-1.3) whereas the surface epithelial cells (ciliated and non-ciliated) stained negative (A and C). The corresponding negative controls for A and C are shown in B and D.

Figure 7: The tubular gland cells and surface epithelium of the magnum (A), isthmus (C) and shell gland (E) regions of the oviduct stained positive with zebra finch anti-OvoDB1 antisera (113_KGE_2.1). The corresponding negative controls are shown in image B, D, and F. Breast muscle (G and H) was not reactive to the primary antibody.

Figure 8: *Gallus gallus* OvoDA1 was incubated for 3 hours at 37°C with *E. coli* DH5α, Avian Pathogenic *E. coli* χ7122 at 0.5-100μM (A) or *S. aureus* 8325-4 at 10-200μM (B) in PBS and the number of surviving bacteria were counted. Results are represented as a reduction in CFU/ml when compared to a 10% DMSO control.

Figure 9: The ED50 of *Gallus gallus* OvoDA1 did not significantly differ with a range of pH (P=0.42) (A) and NaCl concentrations (P=0.49) (B).

Supplementary data 1: All known and newly discovered ovodefensins

Supplementary data 2: ClustalW alignments of each of the 6 cysteine sequence motifs (OvoDA-OvoDF) and a global alignment of all ovodefensin peptides. A) CLUSTALW multiple sequence alignments of mature peptides of each cysteine sequence motif and B) a global alignment of ovodefensin peptides. The conserved cysteines of each ovodefensin peptide are shaded in grey and the conserved glycine is shaded in black. Complete conservation between molecules is indicated with **'*'**, strong similarity with **':'** and weak similarity with **'.'**.

Supplementary data 3: Key for the abbreviations of each species Latin name.

Table 1 – Nomenclature of previously identified ovodefensins and peptides now recognized as ovodefensins.

New gene/ peptide name	Species	Previous name	First defined	Accession no.	Genome location
OvoDA1	*Gallus gallus*	Gallin_1	Mann [3]	GenBank CBE702(78-83)	Chr3; 106740601-106740723
OvoDA1_2	*Gallus gallus*	Gallin_2	Mann [3]	GenBank CBE702(78-83)	Chr 3; 106748086-106748208
OvoDA1_3	*Gallus gallus*	Gallin_3	Mann [3]	GenBank CBE702(78-83)	Chr 3; 106751375-106751497
OvoDA1	*Meleagris gallopavo*	Meleagrin	Odani et al [4]	UniProtKB/Swiss-Prot: P21376.2	Chr 2; 112906487-112907670
OvoDA1	*Cygnus atratus*	Cygnin	Simpson et al [43]	UniProtKB/Swiss-Prot: P02785.3	-
OvoDA1	*Anas platyrhynchos*	dBPS2	Naknukool et al [5]	UniProtKB/Swiss-Prot: P85123.1	Scaffold KB746044.1; 6812-6934
OvoDB1	*Anas platyrhynchos*	dBPS1	Naknukool et al [5]	UniProtKB/Swiss-Prot: P85124.1	Scaffold KB746044.1; 12527-12649
OvoDA1	*Taeniopygia guttata*	Taeniopygin 1	Gong et al [2]	EMBL: LN717251	Chr 3; 111190828-111191187
OvoDA1_2	*Taeniopygia guttata*	Taeniopygin 1_2	Gong et al [2]	EMBL: LN717252	Chr 3; 111176525-111176647
OvoDB1	*Taeniopygia guttata*	Taeniopygin 2	Gong et al [2]	EMBL: LN717253	Chr 3; 111193797-111194424
OvoDA1	*Columba livia*	Cygnin	Shapiro et al [27]	GenBank	-

				EMC87474.1	
OvoDB1	*Columba livia*	Small basic protein 1	Shapiro et al [27]	GenBank: EMC87475.1	-
OvoDB1	*Caretta caretta*	TEWP	Chattopadhyay et al [28]	UniProtKB/Swiss-Prot: P0CAP0.1	-

Table 2 – Names and sequences of primers used in the study.

Forward primer name	Forward primer sequence	Reverse primer name	Reverse primer sequence
Chicken OvoDB1 F/RTF	GTGCTCTTTGCTGTGCTCCT	Chicken OvoDB1 R/RTR	AGAGCTGCTCCTGCTCCAC
Turkey OvoDA1 F	TGGGCTAAGTTTCCCAGCTA	Turkey OvoDA1 R	AGTGACTGTGGCTGTCATCG
Turkey OvoDA1_2 F	AGCTCCTCTCCTCCAGCTCT	Turkey OvoDA1_2 R	TACTTTGCATGGGGTCAACA
Turkey OvoDB1 F/RTF	TGTGCTCCTGCTCTTTTCCT	Turkey OvoDB1 R/RTR	AGAGCCTCTCCTGCTCCACT
Duck OvoDA1 F	AGGGGGTGGCAGTTGAGTAG	Duck OvoDA1 R	GACGAAAGGCTGGAAAACAG
Duck OvoDB1 F	GAGATCCCCACCGTCCTC	Duck OvoDB1 R	TCAAAGCAACGAAACCGTCT
Zebra finch OvoDA1 F	AAATGGGGAAGAGCAATGG	Zebra finch OvoDA1 R	AATCGTGGGGACACCAGTAA
Zebra finch OvoDB1 F	GCTGTGGTCCTGGTGGATA	Zebra finch OvoDB1 R	TTGCTGCAAGCATCACTTTC
Chicken OvoDA1 RTF	CTCCAGCCTCGCTCACAC	Chicken OvoDA1 RTR	TTGAGAGGAGGGGATGACAC
Turkey OvoDA1/1_2 RTF	GCTGTCCTCCTGCTGGTCT	Turkey OvoDA1/1_2 RTR	CCTTGCAGTCGGAGGAGTAG
Duck OvoDA1 RTF	GCTACGGGTTTCCCCTTC	Duck OvoDA1 RTR	AGGTCGGGTACTGCTCCAG
Duck OvoDB1 RTF	GTGCTCCTGCTTTTCTCCAT	Duck OvoDB1 RTR	GCACGCAATGAACACAGC
Zebra finch OvoDA1 RTF	CCTTCCAGGCTATGGGAAGT	Zebra finch OvoDA1 RTR	GCAGTCGCCAGAGGTGTATT
Zebra finch OvoDB1 RTF	CGTTGTCTTTGCTGTTTTCCT	Zebra finch OvoDB1 RTR	GTGCTCCTCCCTCTCTCCTT
Chicken OvoDA1 SP5	ACACGTGCTCCAAGACACAG	Chicken OvoDA1 SP1	GCAGGCAGCAGTACATCTTG
		Chicken OvoDA1 SP2	GCCCAGATCTGTGTCTTGGA

Table 3 – Nomenclature and genome location of newly identified ovodefensins.

Gene/peptide name	Species	Accession no.	Genome build/ location
OvoDB1	*Gallus gallus*	EMBL: LN717248	Nov'11; Chr 3; 106744444-106744572
OvoDB1	*Meleagris gallopavo*	EMBL: LN717250	Dec'09; Chr 2; 112960926-112961582
OvoDB1	*Melopsittacus undulatus*	-	melUnd1; Scaffold JH556573; 2958896-2959015
OvoDB1	*Pelodiscus sinensis*	-	PelSin_1.0; Scaffold JH212527; 3638-3745
OvoDB1_2	*Pelodiscus sinensis*	-	PelSin_1.0; Scaffold JH210373; 9969-10076
OvoDB1	*Chrysemys picta bellii*	-	Dec'11; Scaffold JH584573; 2758224-2758331
OvoDC1	*Chrysemys picta bellii*	-	Dec'11; Scaffold JH584573; 2800433-2800549
OvoDC1_2	*Chrysemys picta bellii*	-	Dec'11; Scaffold JH584573; 2828206-2828322
OvoDC2	*Chrysemys picta bellii*	-	Dec'11; Scaffold JH584573; 2778988-2779104
OvoDE1	*Anolis carolinensis*	-	May'10; Scaffold Un_GL343551; 118450-118557
OvoDE2	*Anolis carolinensis*	-	May'10; Scaffold Un_GL343551; 106991-107098
OvoDE2_2	*Anolis carolinensis*	-	May'10; Scaffold Un_GL343551; 131862-131969
OvoDE2_3	*Anolis carolinensis*	-	May'10; Scaffold Un_GL343551;

			142606-142713
OvoDE3	*Anolis carolinensis*	-	May'10; Scaffold Un_GL343551; 53724-53831
OvoDF1	*Anolis carolinensis*	-	May'10; Scaffold Un_GL343551; 64537-64644
OvoDF2	*Anolis carolinensis*	-	May'10; Scaffold Un_GL343553; 4868-4966
OvoDF3	*Anolis carolinensis*	-	May'10; Scaffold Un_GL343551; 60162-60269
OvoDD1	*Alligator mississippiensis*	-	Aug'12; Scaffold JH736037; 19230-19340
OvoDD2	*Alligator mississippiensis*	-	Aug'12; Scaffold JH736037; 4956-5063
OvoDA1	*Ficedula albicol*	-	FicAlb_1.4; Scaffold JH603338; 1433524-1433405
OvoDB1	*Ficedula albicol*	-	FicAlb_1.4; Scaffold JH603338; 1441614-1441721
OvoDA1	*Geospiza fortis*	-	Apr'12; Scaffold JH760839; 55-174
OvoDA1_2	*Geospiza fortis*	-	Apr'12; Scaffold JH739967; 2617679-2617798
OvoDA1_3	*Geospiza fortis*	-	Apr'12; Scaffold JH739967; 2632098-2632217

Table 4 – Summary of antimicrobial activity of *Gallus gallus* OvoDA1, *Gallus gallus* OvoDB1 and *Anas platyrhynchos* OvoDA1.

Strain[#]	***Gallus gallus* OvoDA1**		***Gallus gallus* OvoDB1**		***Anas platyrhynchos* OvoDA1**	
	% reduction[*]	ED50[**]	% reduction[*]	ED50[**]	% reduction[*]	ED50[**]
DH5α	98.6	3.0 (±0.77)	35.0	-	86.8	53.1(±55.0)
APEC	42.3	-	-	-	-	-
S.aureus	99.44	16.4 (±24.9)	-	-	-	-

[*]Percent reduction in CFU/ml at 100μM is shown for each peptide when activity is observed. [**]The ED50 (μM, ± SEM) value was calculated where a fall in viability greater than 50% was observed. [#]Bacterial strains used were *E. coli* K-12 (DH5α), avian pathogenic *E. coli* O78:H9 (χ7122) and *Staphylococcus aureus* (8325-4).

Figure 1

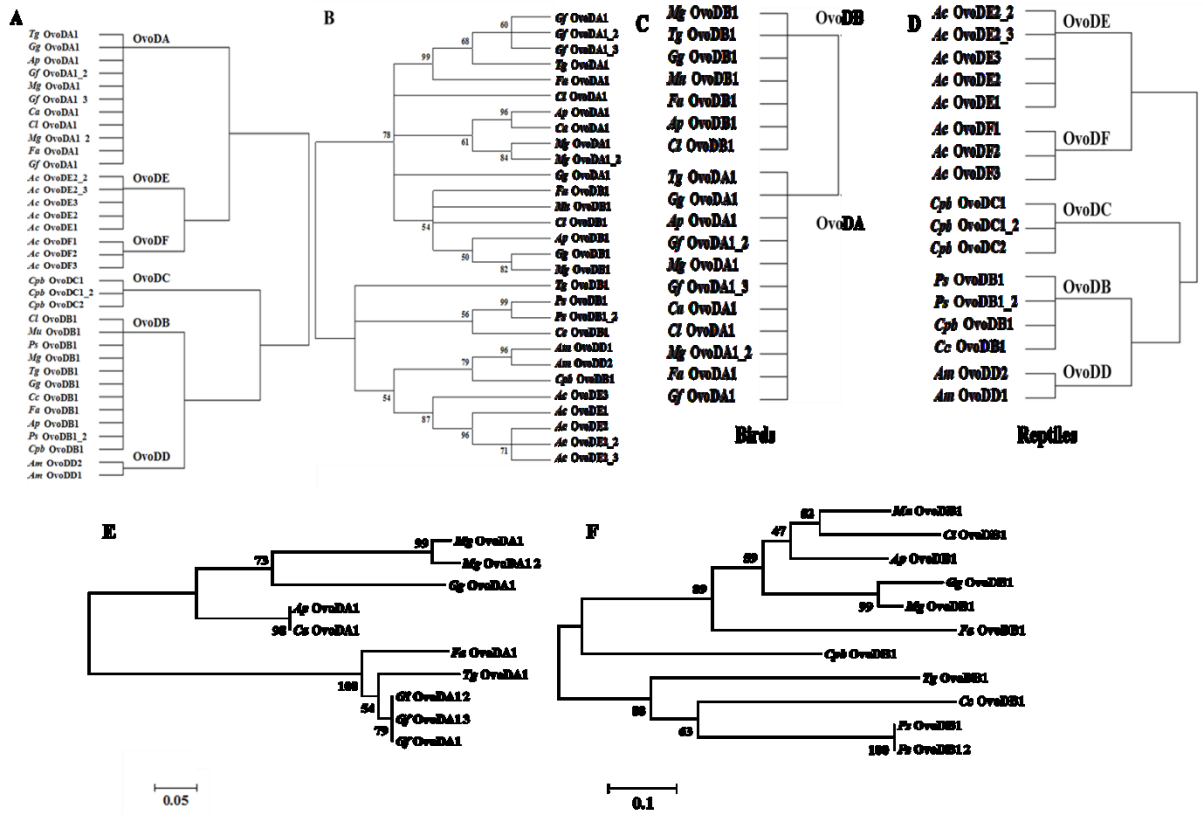

A
B
C
D
OvoDA
OvoDE
OvoDF
OvoDC
OvoDB
OvoDD
Birds
Reptiles
E
F
0.05
0.1

Figure 2

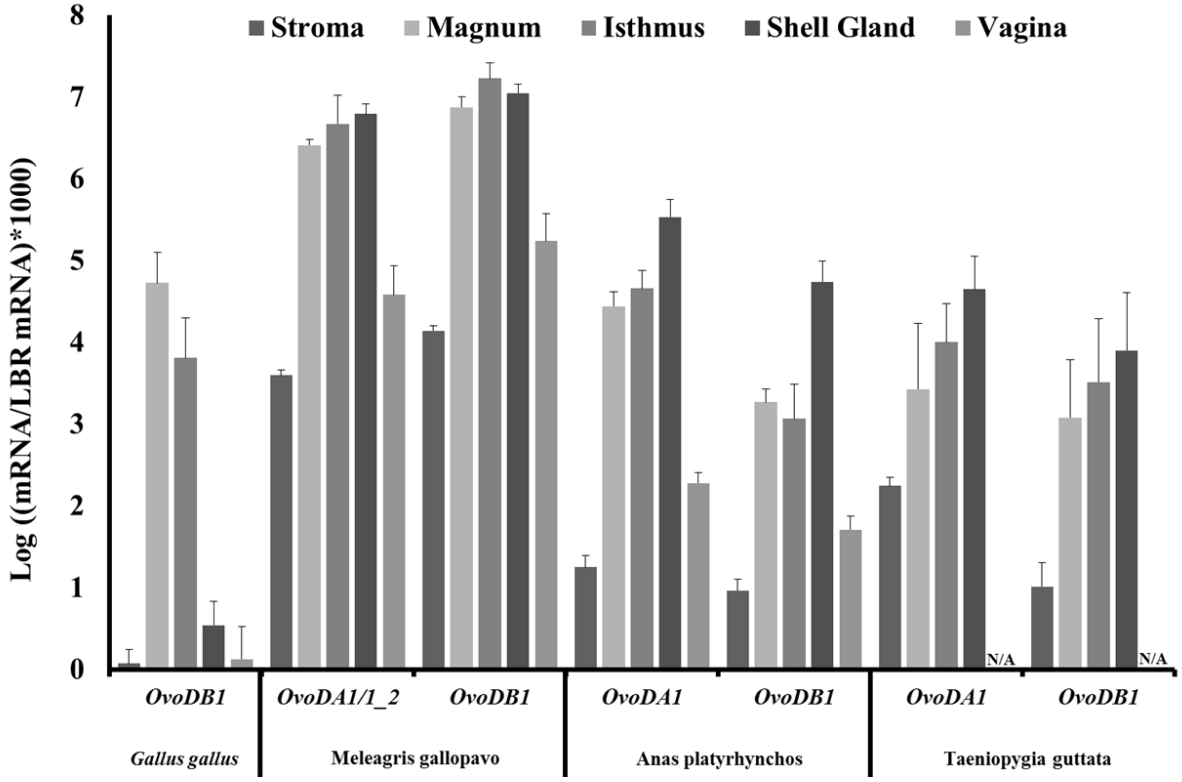


Figure 3

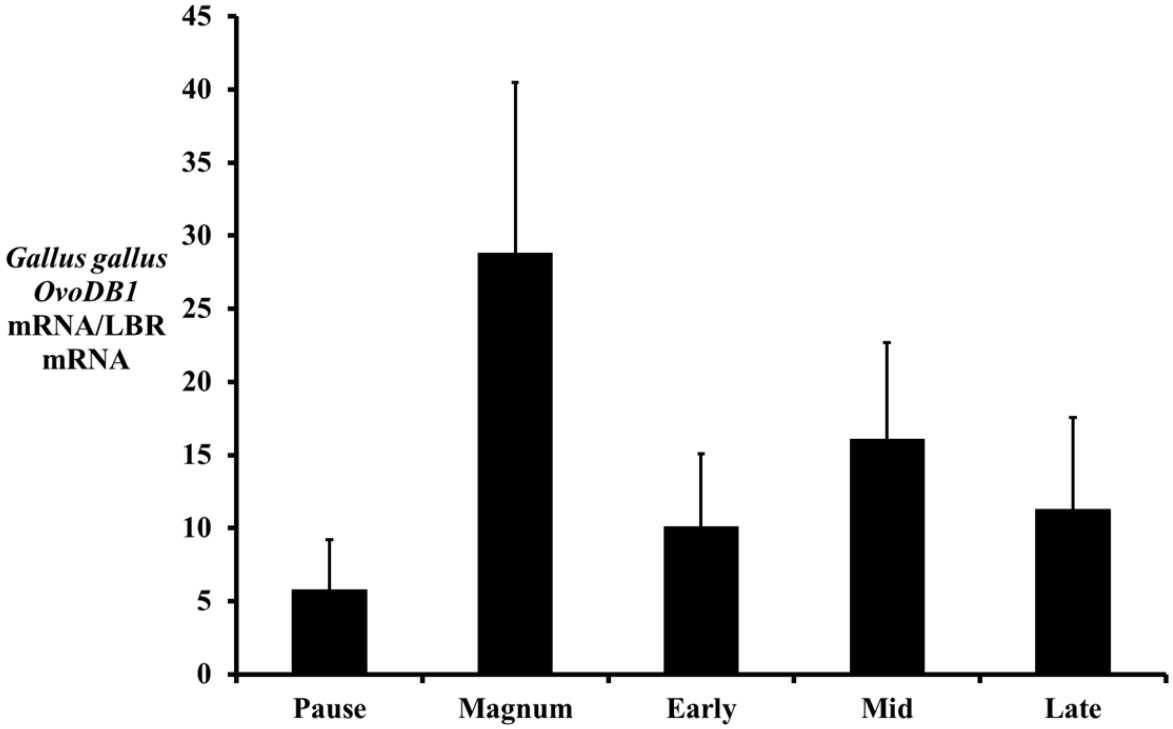


Figure 4

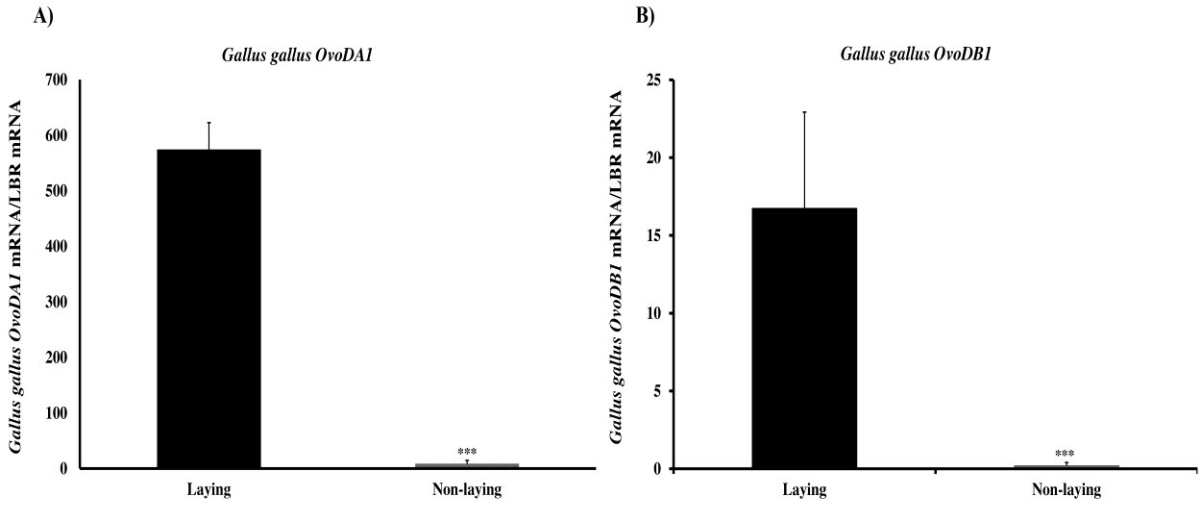


Figure 5

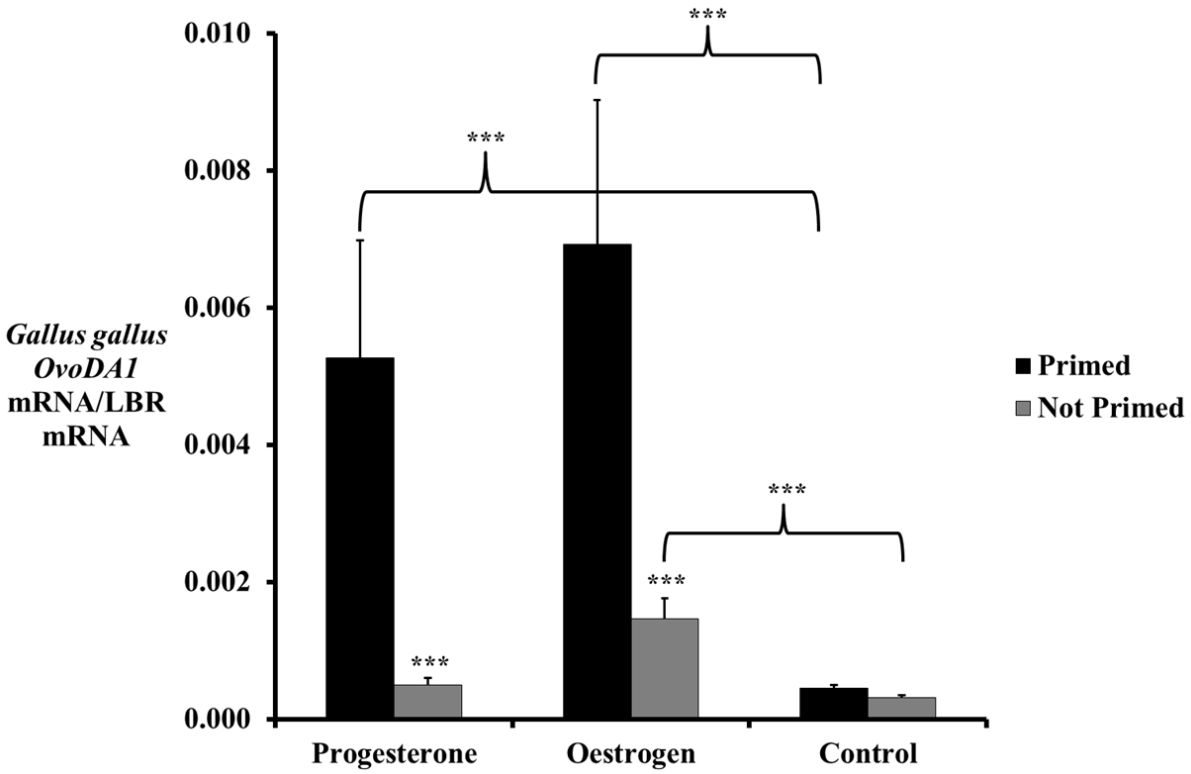

0.010
0.008
0.006
0.004
0.002
0.000
Gallus gallus OvoDA1 mRNA/LBR mRNA
***
***
***
***
***
Progesterone
Oestrogen
Control
Primed
Not Primed

Figure 6

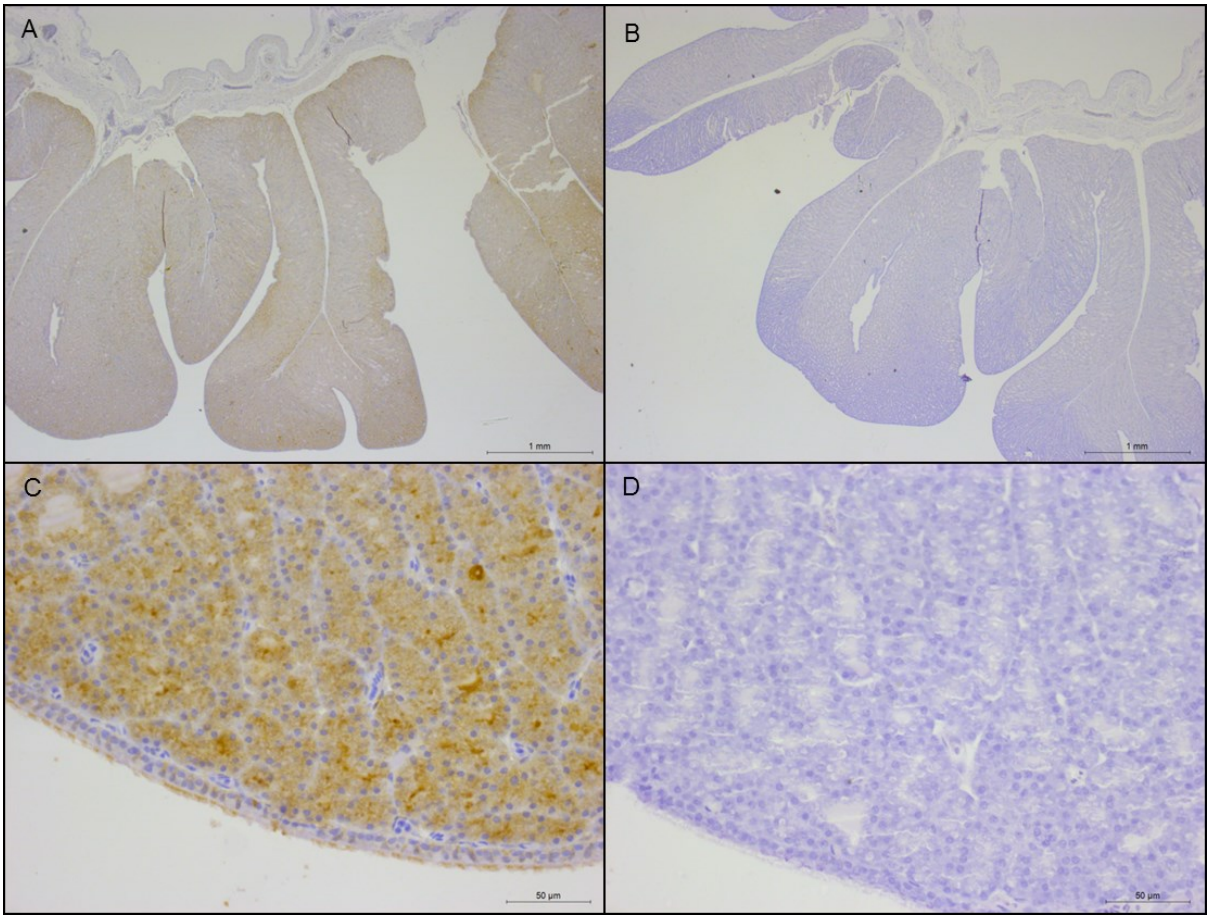


Figure 7

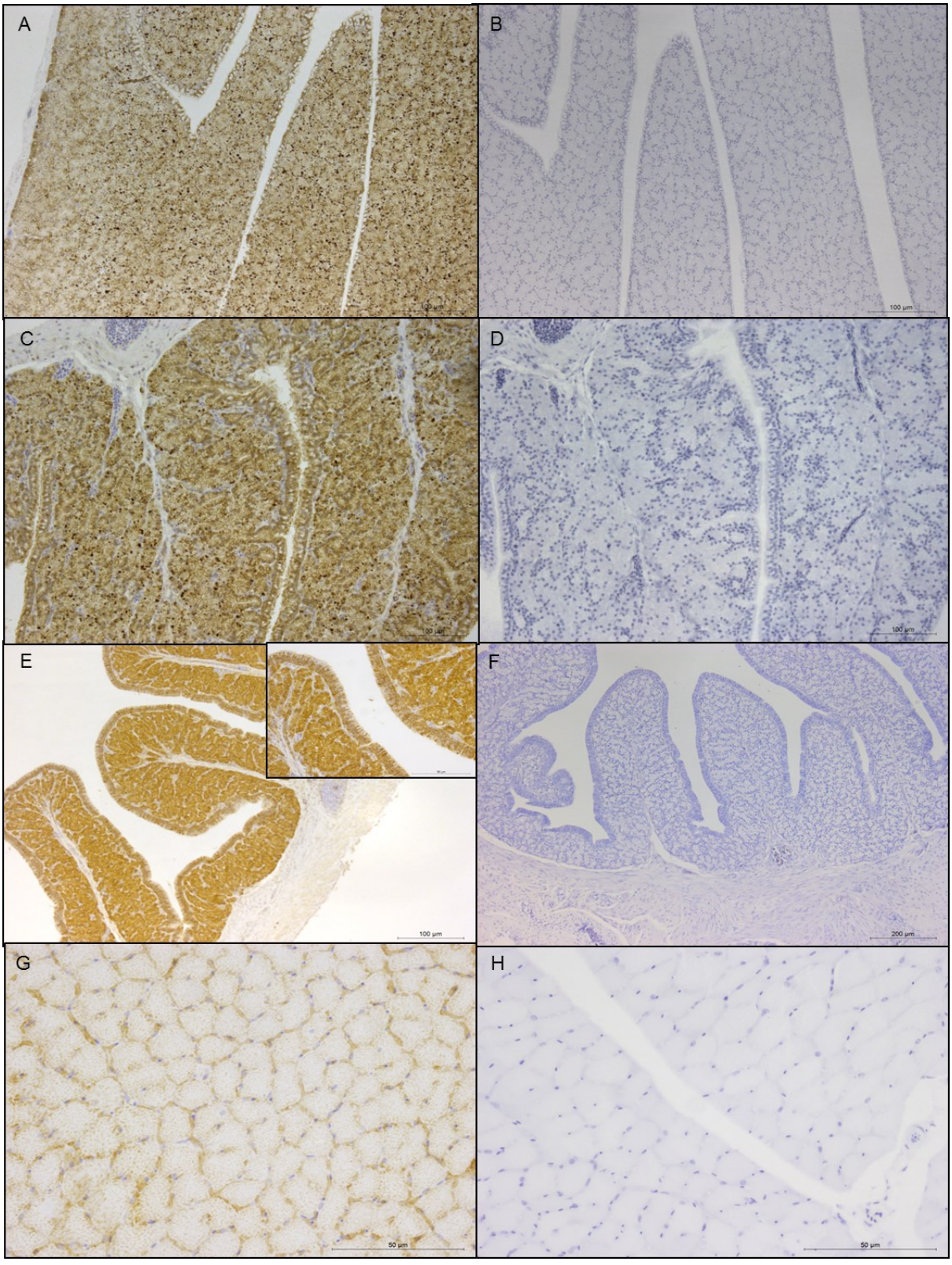


Figure 8

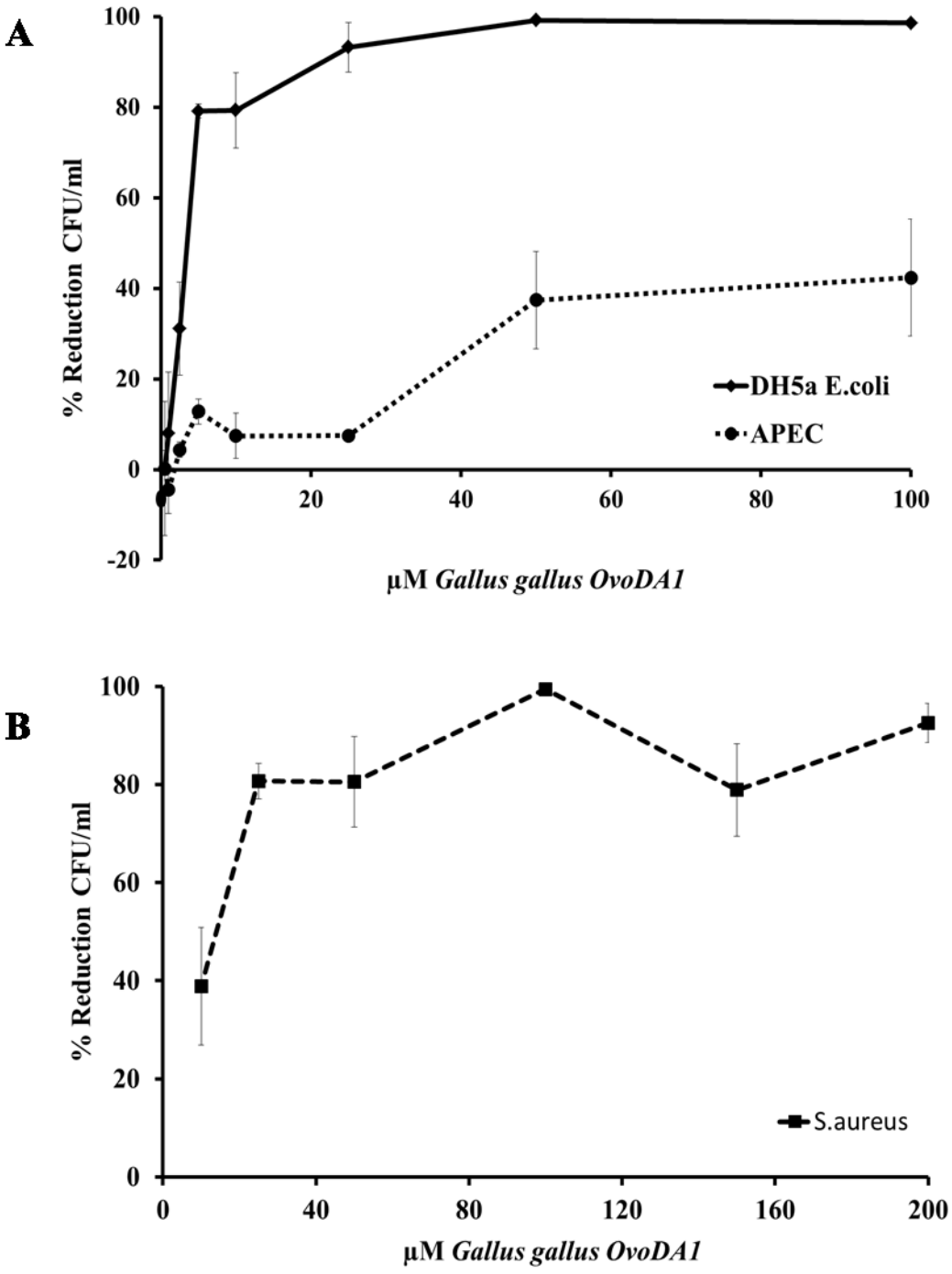
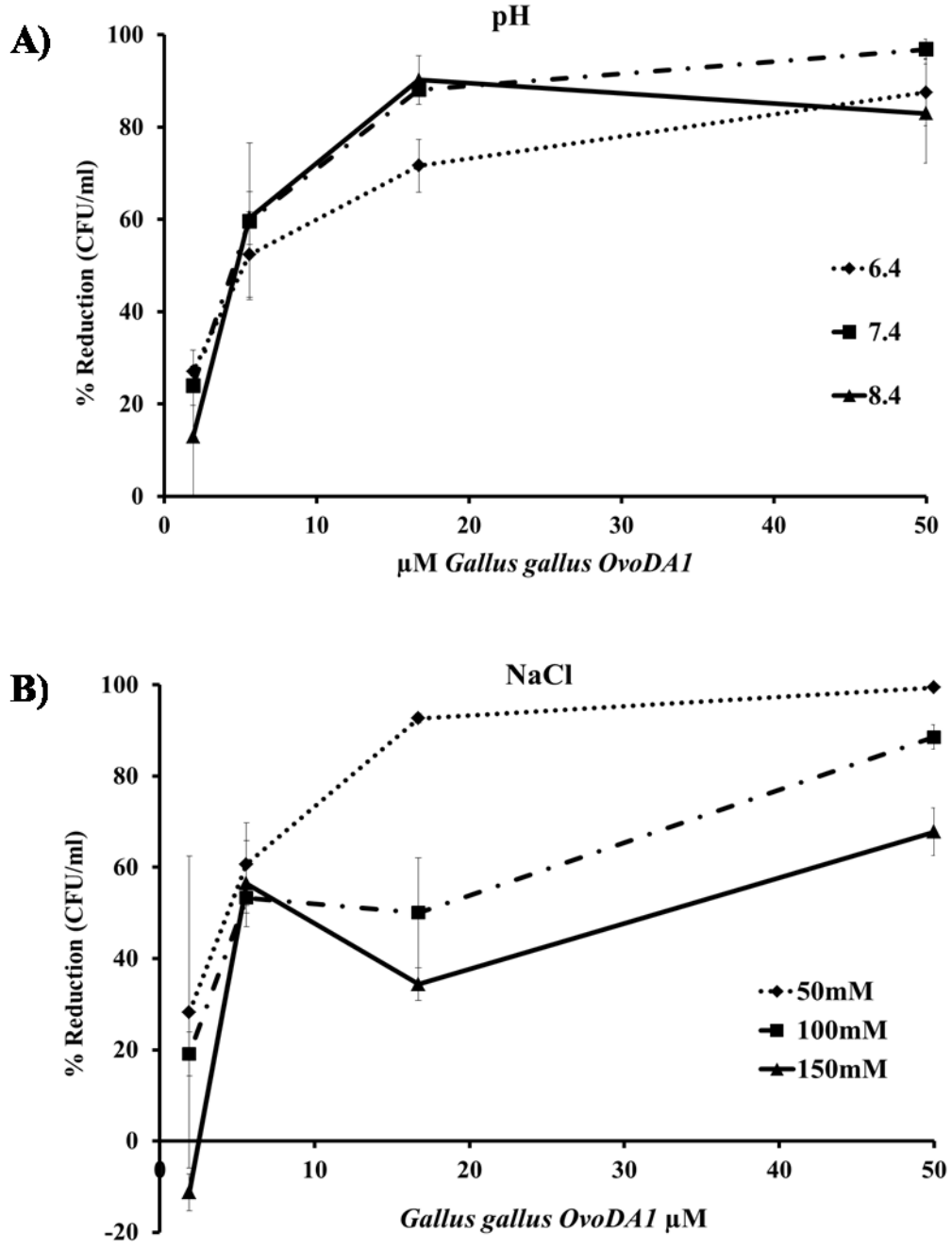

A
100
80
60
40
20
0
-20
% Reduction CFU/ml
20
40
60
80
100
DH5a E.coli
APEC
µM Gallus gallus OvoDA1
B
100
80
60
40
20
0
% Reduction CFU/ml
0
40
80
120
160
200
S.aureus
µM Gallus gallus OvoDA1

Figure 9


A)
pH
% Reduction (CFU/ml)
µM Gallus gallus OvoDA1
6.4
7.4
8.4
B)
NaCl
% Reduction (CFU/ml)
Gallus gallus OvoDA1 µM
50mM
100mM
150mM