

Evolution of the Avian Eggshell Biomineralization Protein Toolkit – New Insights From Multi-Omics

Nathalie Le Roy, Lilian Stapane, Joël Gautron, Maxwell T Hincke

▶ **To cite this version:**

Nathalie Le Roy, Lilian Stapane, Joël Gautron, Maxwell T Hincke. Evolution of the Avian Eggshell Biomineralization Protein Toolkit – New Insights From Multi-Omics. Frontiers in Genetics, 2021, 12, 10.3389/fgene.2021.672433 . hal-03223848

HAL Id: hal-03223848
https://hal.inrae.fr/hal-03223848

Submitted on 11 May 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

frontiers in Genetics

REVIEW
published: 11 May 2021
doi: 10.3389/fgene.2021.672433

Evolution of the Avian Eggshell Biomineralization Protein Toolkit – New Insights From Multi-Omics

*Nathalie Le Roy[1], Lilian Stapane[1], Joël Gautron[1] and Maxwell T. Hincke[2,3]**

[1] INRAE, Université de Tours, Nouzilly, France, [2] Department of Innovation in Medical Education, University of Ottawa, Ottawa, ON, Canada, [3] Department of Cellular and Molecular Medicine, University of Ottawa, Ottawa, ON, Canada

The avian eggshell is a remarkable biomineral, which is essential for avian reproduction; its properties permit embryonic development in the desiccating terrestrial environment, and moreover, are critically important to preserve unfertilized egg quality for human consumption. This calcium carbonate ($CaCO_3$) bioceramic is made of 95% calcite and 3.5% organic matrix; it protects the egg contents against microbial penetration and mechanical damage, allows gaseous exchange, and provides calcium for development of the embryonic skeleton. In vertebrates, eggshell occurs in the Sauropsida and in a lesser extent in Mammalia taxa; avian eggshell calcification is one of the fastest known $CaCO_3$ biomineralization processes, and results in a material with excellent mechanical properties. Thus, its study has triggered a strong interest from the researcher community. The investigation of eggshell biomineralization in birds over the past decades has led to detailed characterization of its protein and mineral constituents. Recently, our understanding of this process has been significantly improved using high-throughput technologies (i.e., proteomics, transcriptomics, genomics, and bioinformatics). Presently, more or less complete eggshell proteomes are available for nine birds, and therefore, key proteins that comprise the eggshell biomineralization toolkit are beginning to be identified. In this article, we review current knowledge on organic matrix components from calcified eggshell. We use these data to analyze the evolution of selected matrix proteins and underline their role in the biological toolkit required for eggshell calcification in avian species. Amongst the panel of eggshell-associated proteins, key functional domains are present such as calcium-binding, vesicle-binding and protein-binding. These technical advances, combined with progress in mineral ultrastructure analyses, have opened the way for new hypotheses of mineral nucleation and crystal growth in formation of the avian eggshell, including transfer of amorphous $CaCO_3$ in vesicles from uterine cells to the eggshell mineralization site. The enrichment of multi-omics datasets for bird species is critical to understand the evolutionary context for development of $CaCO_3$ biomineralization in metazoans, leading to the acquisition of the robust eggshell in birds (and formerly dinosaurs).

Keywords: birds, eggshell, calcium carbonate biomineralization, organic matrix, multi-omics tools, evolution

OPEN ACCESS

Edited by:
Frederic Marin, Délégation Centre-Est, Center for the National Scientific Research, France

Reviewed by:
Paula Ramos-Silva, Naturalis Biodiversity Center, Netherlands
Marie Arul, Muséum National d'Histoire Naturelle, France

***Correspondence:**
Maxwell T. Hincke
mhincke@uottawa.ca

Specialty section:
This article was submitted to Evolutionary and Population Genetics, a section of the journal Frontiers in Genetics

Received: *25 February 2021*
Accepted: *08 April 2021*
Published: *11 May 2021*

Citation:
Le Roy N, Stapane L, Gautron J and Hincke MT (2021) Evolution of the Avian Eggshell Biomineralization Protein Toolkit – New Insights From Multi-Omics. Front. Genet. 12:672433. doi: 10.3389/fgene.2021.672433

INTRODUCTION

The avian eggshell is a calcitic biomineral that surrounds the telolecithal egg (i.e., possessing an uneven distribution of vitellus). The eggshell is essential to prevent desiccation during embryonic development and to regulate metabolic gas exchange. The shell is a remarkable physical barrier to protect the embryo against pathogens and mechanical shocks (Hincke et al., 2012; Gautron et al., 2021); moreover, the shell is a source of calcium for embryonic bone mineralization (Hincke et al., 2019). The egg is an autonomous source of all nutritive elements for embryo development, and therefore the unfertilized chicken egg is a high quality nutrient in the human diet. The study of the eggshell calcification process is of great importance to provide new insights into mechanisms of biomineralization, and to provide new tools to ensure the quality of the egg and its food safety for human consumption.

The oviduct is the organ of egg production in birds; it consists of six distinctly specialized segments that secrete the constituents of each egg compartment: infundibulum (vitelline membrane to enclose the egg yolk), magnum (secretion of egg white), white isthmus (elaboration of eggshell membranes), red isthmus (initiation of shell mineralization on mammillary cores), uterus (eggshell formation, cuticle deposition) and vagina (expulsion of mature egg). After ovulation of the ovocyte and its egg yolk, the egg white is secreted and deposited around the forming egg while it transits through the magnum segment. The eggshell membranes are deposited in the isthmus, and the calcitic eggshell is mineralized while it remains in the uterus (Nys et al., 2004). Eggshell mineralization takes place in an acellular uterine fluid secreted by uterine tissue, which contains mineral and organic precursors essential for shell mineralization (Gautron et al., 1997; Jonchère et al., 2012). *In vitro* crystallization tests using chicken uterine fluid have shown that this fluid strongly modifies the kinetics of calcite crystal formation and the resulting crystal morphology (Dominguez-Vera et al., 2000; Hernandez-Hernandez et al., 2008b). Finally, shell biomineralization is arrested with deposition of the phosphate-rich cuticle and the egg is laid. Eggshell mineralization follows five major stages: briefly, (1) amorphous calcium carbonate (ACC) is deposited on the entire surface of the outer eggshell membranes, which then (2) transforms into calcite at organic matrix clusters (mammillary cores); (3) calcite crystals nucleate from these sites and (4) grow rapidly with their *c*-axis becoming increasingly perpendicular to the eggshell surface. Two hours before egg expulsion, (5) mineralization is arrested and a thin layer of organic cuticle is deposited that covers the calcified layer and plugs the respiratory pores (Nys et al., 2004; Rodríguez-Navarro et al., 2015). For more detail regarding eggshell mineralization, the reader is referred to a new review by Gautron et al. (2021).

The avian eggshell is the result of an exceptional evolutionary strategy. Since the Late Devonian geological period (∼360 MYA – million years ago), the conquest of the terrestrial landscape challenged vertebrates to fulfill various vital functions such as breathing, locomotion and reproduction. Birds belong to the Sauropsida clade (that includes modern and extinct reptiles – turtles, lizards, snakes, crocodiles etc. – and birds) that appeared 315 MYA (Falcon-Lang et al., 2007). While amphibians have retained a need to lay their eggs in water, birds produce an impervious calcified barrier around the egg that they lay in terrestrial nests. Amongst laying-egg animals, birds possess the most solid eggshell; the soft-shelled eggs of turtles, lizards and snakes are less mineralized than bird eggshells, whereas crocodiles produce intermediate hard-shelled eggs (Choi et al., 2018). According to the fossil record, the first evidence of a calcified eggshell occurred at the Late Triassic/Early Jurassic, and belonged to a crocodilian (Bonaparte and Vince, 1979; Carpenter, 2000). In dinosaurs, the groups from which birds emerged, the oldest eggshells have been identified in the Early Jurassic, for which microstructural studies reveal a very thin shell (Garcia et al., 2006; Stein et al., 2019). According to numerous observations, the microstructure of dinosaur and bird eggshells is highly similar, with calcareous crystals forming an inner mammillary zone and outer palisade structure (Mikhailov, 1991). A recent study provided evidence for the independent evolution of calcified eggs in dinosaurs, with soft-shelled eggs as the ancestral character and the occurrence of at least three hard-shelled egg events (Norell et al., 2020).

Much knowledge of eggshell biomineralization has been obtained from studies utilizing the chicken egg. For decades, a combination of physical and biological approaches has led to increased understanding of this process. Microscopies (SEM and TEM), infrared and Raman spectroscopies, and X-ray diffraction have characterized the mineral phase (Rodríguez-Navarro et al., 2015; Pérez-Huerta and Dauphin, 2016; Choi et al., 2019). Protein purification, immunochemistry (Western blotting, colloidal gold immunocytochemistry) and proteomics were essential to identify occluded organic matrix constituents (Hincke et al., 2012; Gautron et al., 2021). Hundreds of proteins have been identified in eggshell proteomes from a small number of species (**Table 1**). Amongst this protein cortege, major functions have been assigned such as calcium-binding, matrix-organization, antimicrobial function, and so on (Marie et al., 2015a). The present review aims to describe six major proteins that have been identified as key actors in eggshell mineralization: Ovocalyxin-32 (OCX-32), Ovocalyxin-36 (OCX-36), Ovocleidin-116 (OC-116), Osteopontin (OPN), EGF (epidermal growth factor)-like repeats and discoidin domains 3 (EDIL3), and Ovocleidin-17 and its homologs (OC-17 and XCA). These six proteins possess essential functions (antimicrobial properties, regulation of $CaCO_3$ crystallization or vesicular transport of ACC) and are present in significant abundance to be considered as major members of the eggshell biomineralization toolkit (**Figure 1**); it is likely that they were recruited during evolutionary acquisition of eggshell formation.

AVIAN EGGSHELL: AN EXCEPTIONAL VERTEBRATE CALCIUM CARBONATE BIOMINERAL

The bird eggshell is a remarkable bioceramic that demonstrates exceptional mechanical properties; the eggshell appeared in the last common ancestor of amniotes around 326 MYA

TABLE 1 | List of eggshell proteomes and the presence of transcriptome/genome in bird and reptile species.

Species	Common name	Eggshell proteins	Uterine transcriptome/genome	Proteome references
Neognathae → Galloanserae				
Gallus gallus	Chicken	904	Yes/yes	Mann et al., 2006; Marie et al., 2015a
Meleagris gallopavo	Turkey	697	No/yes	Mann and Mann, 2013
Numida meleagris	Guinea fowl	154	Yes/no	Le Roy et al., 2019
Coturnix japonica	Japanese quail	622	Yes/yes	Mann and Mann, 2015
Anas platyrhynchos	Mallard duck	484	Yes/yes	Zhu et al., 2019
Neognathae → Neoaves				
Taeniopygia guttata	Zebra finch	475	No/yes	Mann, 2015
Palaeognathae				
Struthio camelus	Common ostrich	2	No/yes	Mann and Siedler, 2004
Dromaius novaehollandiae	Emu	2	No/yes	Mann and Siedler, 2006
Rhea americana	Greater rhea	2	No/yes	Mann and Siedler, 2006
Crocodilia				
Crocodylus siamensis	Siamese crocodile	58	No/no	Mikšík et al., 2018

Hundreds of proteins have been identified in eggshell proteomes in a small number of species. The presence of uterine transcriptome characterization and/or genome annotation in these species was checked in NCBI databases.

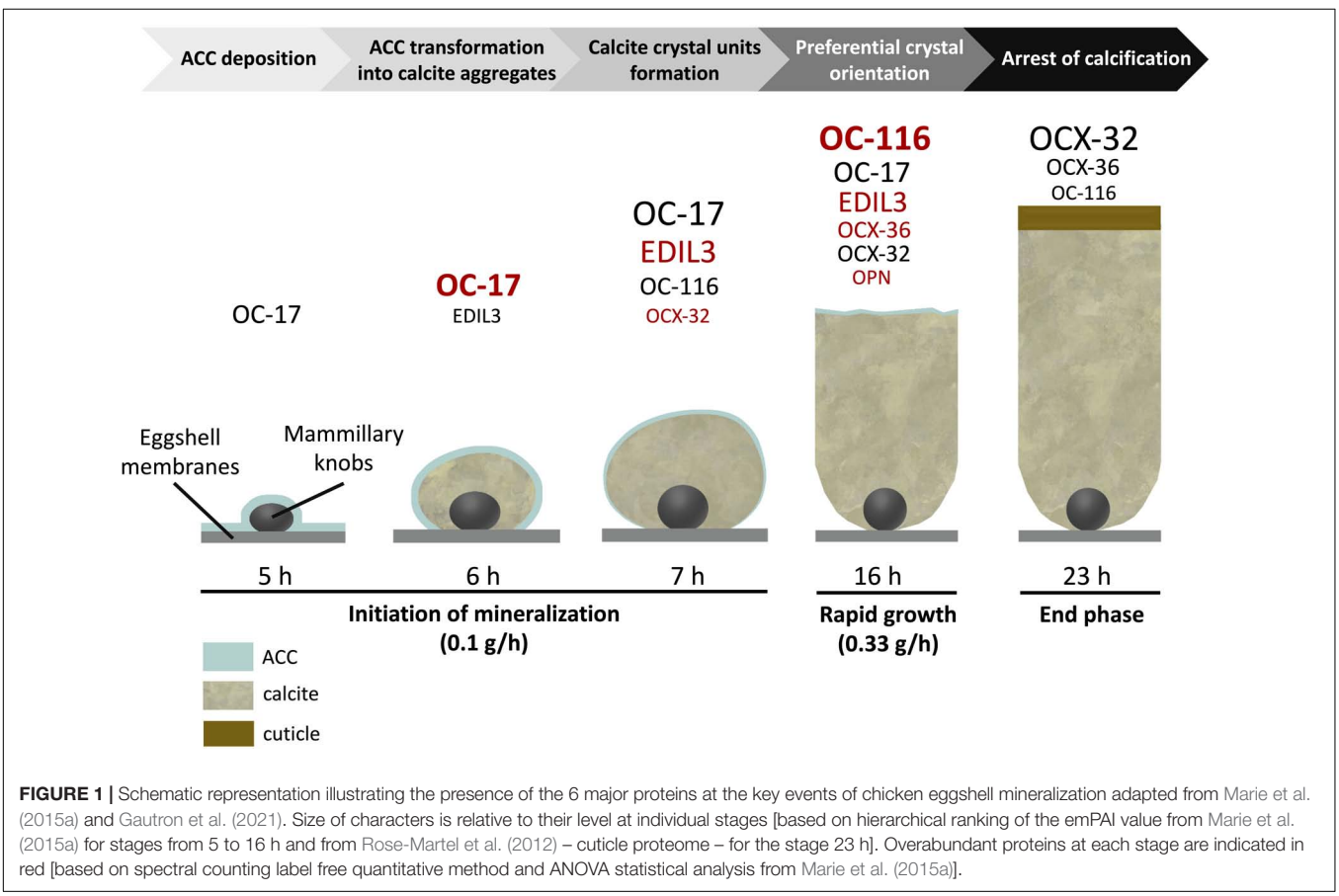


FIGURE 1 | Schematic representation illustrating the presence of the 6 major proteins at the key events of chicken eggshell mineralization adapted from Marie et al. (2015a) and Gautron et al. (2021). Size of characters is relative to their level at individual stages [based on hierarchical ranking of the emPAI value from Marie et al. (2015a) for stages from 5 to 16 h and from Rose-Martel et al. (2012) – cuticle proteome – for the stage 23 h]. Overabundant proteins at each stage are indicated in red [based on spectral counting label free quantitative method and ANOVA statistical analysis from Marie et al. (2015a)].

(Blair and Hedges, 2005; Ford and Benson, 2020). Across various bird species, eggshell strength is positively correlated with egg weight (Ar et al., 1979). Guinea fowl shows an elevated strength (**Figure 2**), which is related to its unique shell texture (Pérez-Huerta and Dauphin, 2016). In general, calcium carbonate crystal units of the bird eggshell palisade layer are parallel to each other and grow following the c-axis, i.e., perpendicular to the eggshell membranes and eventual eggshell

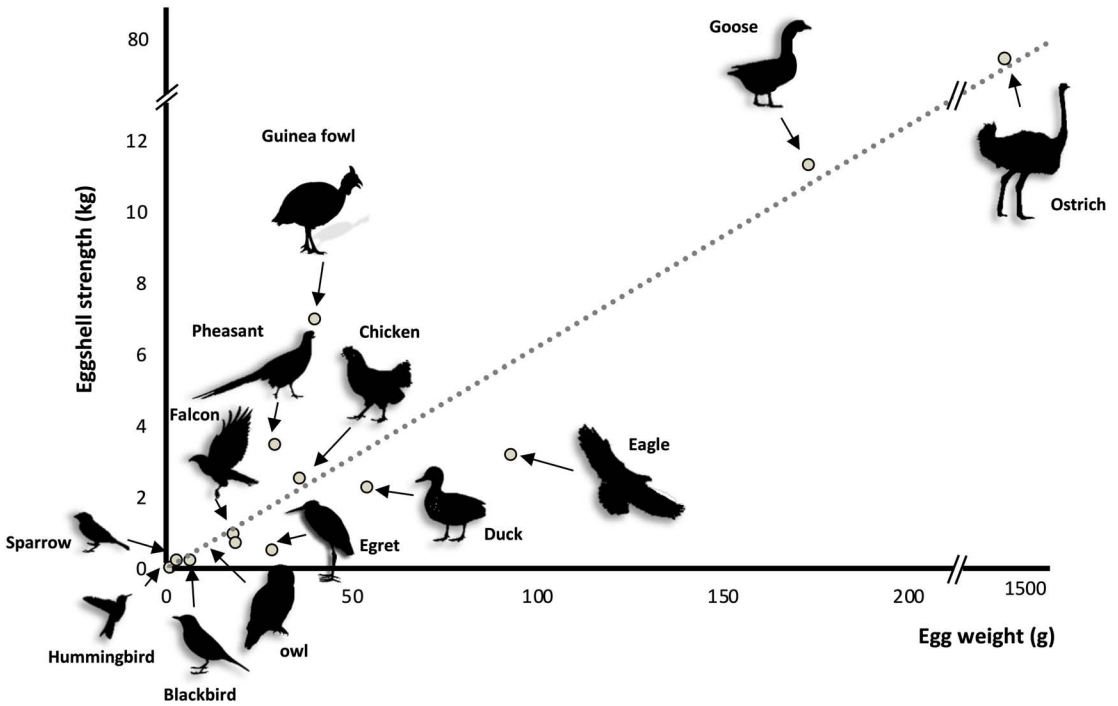


FIGURE 2 | Eggshell strength as function of egg weight in thirteen bird species [adapted from Ar et al. (1979)].

surface (Hernandez-Hernandez et al., 2008a,b; Rodríguez-Navarro et al., 2015). This organization is also observed in fossilized dinosaur eggshells (Mikhailov, 1991; Voris et al., 2018; Dawson et al., 2020). In Guinea fowl (*Numida meleagris*), the first third of deposited shell is similar in structure to that of other species, whereas the outer two-thirds are composed of smaller crystal units with varying crystallographic orientations that form an intricate interlacing pattern that greatly improves shell strength (Petersen and Tyler, 1966; Panheleux et al., 1999a; Song et al., 2000; Pérez-Huerta and Dauphin, 2016). These animals are endemic to Central Africa and lay their eggs on the ground. The elevated breaking strength of their eggshell is likely the result of a specific adaptation to their environment (e.g., predation). A recent microstructural study of other bird eggshells suggested that Guinea fowl is not the only avian species to demonstrate this peculiar feature, since Rhea eggshell also has a similar crystalline organization (Choi et al., 2019). Since both are ground-nesting species, it is tempting to correlate their vulnerable nest location to the interlaced crystalline organization of their strengthened eggshell, which could better protect the egg. However, in other ground-nesting species, such as ostrich, chicken, turkey, etc., the eggshell ultrastructure is columnar.

The ultrastructure, polymorph and nucleation/growth of calcium carbonate crystals are controlled by a specific macromolecular toolkit, the organic matrix (OM) (Chien et al., 2008; Hernandez-Hernandez et al., 2008a,b; Zhao et al., 2013; Gautron et al., 2019, 2021). The bird eggshell OM represents 3.5% of the total shell weight including shell membranes. In the calcified part, the OM represents 1.5–2% of the overall contents (Panheleux et al., 2000). This eggshell OM is composed of proteins and proteoglycans, and has been studied for several decades, especially in chicken, *Gallus gallus* (Leach, 1982; Hincke et al., 1995, 2012; Pines et al., 1995; Gautron et al., 1997, 2021; Nys et al., 1999; Panheleux et al., 1999b; Arias and Fernandez, 2001). Eggshell membranes are the physical support for initiation of shell formation and are composed of a collagen-rich network of fibers (Wong et al., 1984; Arias et al., 1991; Ahmed et al., 2017). As in other metazoan biomineralization systems (e.g., molluscan shell, coral exoskeleton, echinoderm skeleton, etc.), the proteins of the OM possess various functional domains that regulate the matrix organization and control mineral formation (Marie et al., 2010; Ramos-Silva et al., 2013; Gautron et al., 2021). In addition, antimicrobial activities have been identified in avian eggshell that reinforce protection of the embryo against pathogens (Wellman-Labadie et al., 2008a,b,c; Cordeiro et al., 2013; Marie et al., 2015a,b; Gautron et al., 2021). Multiple studies of the evolution of calcium carbonate biomineralization in metazoans have demonstrated that certain homologous proteins were independently recruited to support this process, such as carbonic anhydrases and C-type lectins (Blank et al., 2003; Matsubara et al., 2008; Moya et al., 2008; Le Roy et al., 2014; Voigt et al., 2014; Karakostis et al., 2016; Weber et al., 2016). In other cases, newly arisen genes became specific and essential in the biomineralization process; for instance, the *Enam* gene product (Enamelin) in mammals is involved in dental enamel mineralization. This gene either emerged in mammals after

the mammal/bird divergence or was lost in birds (Kawasaki et al., 2007). There are other examples of taxon-specific proteins involved in metazoan biomineralization, such as scleritin in the calcitic skeleton of octocorals (Conci et al., 2019; Le Roy et al., 2021), galaxin in the calcitic and aragonitic skeleton of corals (Conci et al., 2020), and pearlin in nacre of pearl oysters (Marie et al., 2011).

MULTI-OMICS: A SIGNIFICANT CONTRIBUTION TO THE IDENTIFICATION OF EGGSHELL ORGANIC MATRIX PROTEINS

In the past few years, a huge enrichment of the genomic and transcriptomic databases in Aves has widely contributed to the identification of eggshell OM proteins in diverse species. The current Ensembl database[1] contains genomes from 40 bird and 18 reptile species (**Figure 3A**), while NCBI lists genome assemblies from 507 bird (487 Neognathae and 20 Palaeognathae) and 64 reptile species. The 10,000 Genomes Project (B10K[2]) recently reported the genomes for 363 bird species including 267 new genomes, establishing a new pipeline to analyze the unprecedented scale of genomic data, and illustrating how these resources give improved resolution for genomic evolution analyses (Feng et al., 2020). The recent enrichment of genomic databases provides a critical tool for identification of bird eggshell proteome constituents by high-throughput mass spectrometry analysis (**Figure 3B**) (Mann et al., 2006, 2007a; Rose-Martel et al., 2012, 2015; Mann and Mann, 2013, 2015; Mann, 2015; Marie et al., 2015a; Gautron, 2019; Gautron et al., 2019, 2021; Le Roy et al., 2019; Zhu et al., 2019). Accurate gene annotation is critical to support proteomic approaches. For example, more than 1,300 chicken eggshell protein sequences with different identifiers were aligned to eliminate all redundancies; with this approach, 904 unique proteins were identified in the eggshell layers including membranes and cuticle (Gautron, 2019; Gautron et al., 2019). Another integrated analysis of chicken eggshell matrix enumerated a total of 676 eggshell matrix proteins in the mineralized shell (Yang et al., 2020). Additional bird eggshell proteomes have been studied quite extensively, identifying 697, 622, 475, and 484 proteins in the mineralized eggshell of turkey (*Meleagris gallopavo*), quail (*Coturnix japonica*), zebra finch (*Taeniopygia guttata*) and mallard (*Anas platyrhynchos*) eggshells, respectively (Mann and Mann, 2013, 2015; Mann, 2015;

[1]http://www.ensembl.org/index.html
[2]https://b10k.genomics.cn/index.html

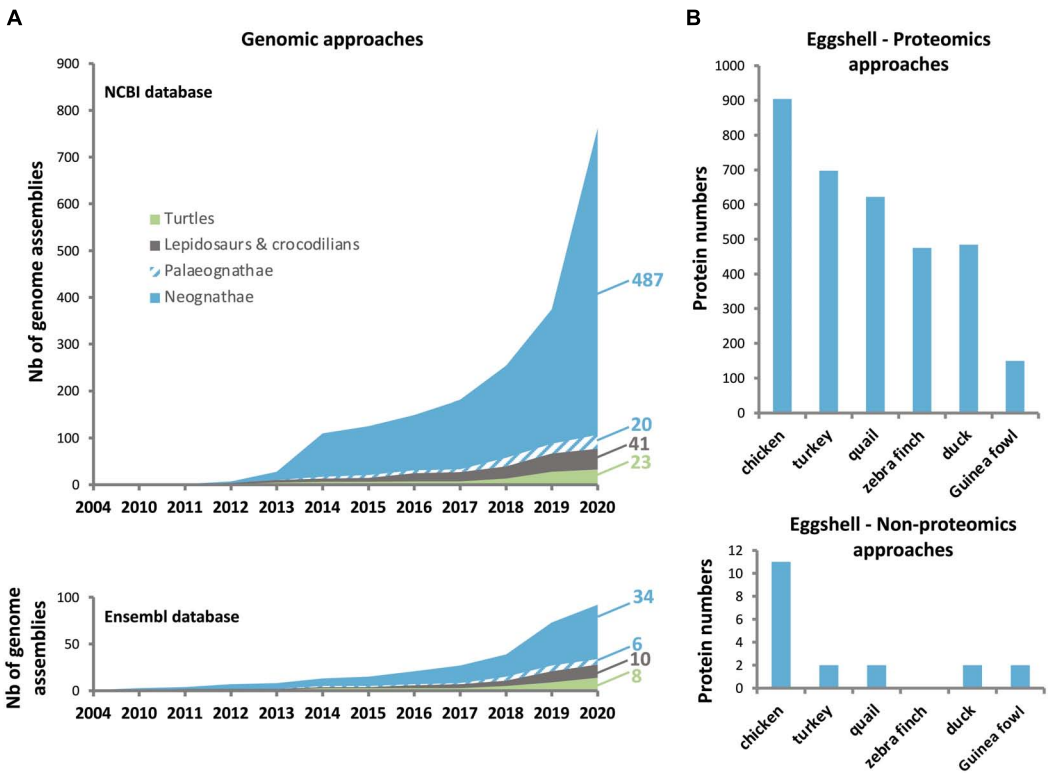


FIGURE 3 | Number of genomes sequenced in birds and reptiles, and the proteins identified in eggshell of different avian species. **(A)** Increase in number of bird and reptile genomes available in NCBI (upper: https://www.ncbi.nlm.nih.gov/), and Ensembl (lower: http://www.ensembl.org/index.html) databases since the chicken genome assembly in 2004. **(B)** Number of identified proteins using proteomics (upper) and non-proteomics (lower) approaches. Data were compiled from Hincke et al. (1995, 1999), Panheleux et al. (1999b), Gautron et al. (2001, 2007),Mann et al. (2006), Mann and Mann (2013, 2015), Mann (2015), Marie et al. (2015a), Zhu et al. (2019), and Le Roy et al. (2019).

Zhu et al., 2019), and 149 proteins in the entire eggshell of Guinea fowl (*Numida meleagris*) (Le Roy et al., 2019). This low number of identified proteins in Guinea fowl eggshell is possibly due to incomplete annotation of its genome (NumMel1.0.).

The next section will describe the evolutionary context for six major proteins that have been identified as key actors in chicken eggshell mineralization (OCX-36, OCX-32, OC-116, OPN, EDIL3, and OC-17/XCA).

VERTEBRATE PROTEINS RECRUITED IN THE EGGSHELL BIOMINERALIZATION PROCESS

Ovocalyxin-32: An Antimicrobial Protein That Influences Eggshell Quality

Ovocalyxin-32 (OCX-32) was originally identified in chicken as an eggshell-specific protein, and its gene is highly expressed in the uterus and the isthmus regions of the oviduct (Gautron et al., 2001). This last study localized OCX-32 in the outer shell (outer palisade layer, the vertical crystal layer and the cuticle). Proteomic analyses revealed abundant OCX-32 in the uterine fluid during the initial phase of mineralization, and its relative enrichment in the palisade region of the eggshell (Marie et al., 2015b,a). OCX-32 was also identified in the proteome of the insoluble fraction of the chicken eggshell organic matrix (Mikšík et al., 2007). OCX-32 possesses 32% identity with mammalian carboxypeptidase inhibitors, latexin and the retinoic acid receptor-responder 1 (RARRES1). Recombinant OCX-32 inhibited bovine carboxypeptidase and the growth of *Bacillus subtilis* (Xing et al., 2007), suggesting an antimicrobial role for OCX-32 in providing protection to the developing avian embryo. Proteomic analysis of the chicken eggshell cuticle demonstrated that OCX-32 is one of the most abundant constituents of this non-mineralized region, and could play a major role in the antimicrobial properties of the cuticle (Rose-Martel et al., 2012; Bain et al., 2013). Polymorphisms in the gene coding for OCX-32, *RARRES1* (gene synonym*: OCX32*), are significantly associated with egg production traits (Uemoto et al., 2009; Romé and Le Roy, 2016). In another study, the quantitative trait loci (QTLs) on chromosome 9 were investigated in an F2 generation that was an intercross between two chicken lines divergently selected for eggshell strength (Takahashi et al., 2010). *RARRES1/OCX32* was identified as a candidate gene influencing eggshell quality (e.g., egg weight, egg dimensions and eggshell weight), and *RARRES1*/*OCX32* SNPs (single-nucleotide polymorphisms) are associated with eggshell quality and mammillary knob layer thickness (Dunn et al., 2008). Trait association studies of non-synonymous SNPs also revealed a significant effect of OCX-32 on shell color in white egg lines and line-specific significant effects on albumen height, early egg weight, puncture score, and yolk weight (Fulton et al., 2012).

Recent proteomics analyses failed to identify OCX-32 in the turkey and quail eggshell (Mann and Mann, 2013, 2015). Nevertheless, in addition to the chicken eggshell organic matrix, this protein has been identified in the eggshell proteome of zebra finch, Guinea fowl and mallard duck (Mann, 2015; Le Roy et al., 2019; Zhu et al., 2019). Synteny analysis of the *RARRES1/OCX32* gene using NCBI database, demonstrates that it is homologous from fishes to mammals (**Figure 4**), suggesting a common ancestor in vertebrates. In addition, the chromosomal location of the *RARRES1/OCX32* gene is highly conserved in a syntenous gene locus from fishes to mammals (**Figure 4**). Therefore, while OCX-32 is highly conserved and may be an important member of the eggshell mineralization toolkit, it does not appear to be present in all eggshell proteomes, nor is it unique to calcium carbonate biomineralizing organisms.

The LBP/BPI/PLUNC Family Protein, Ovocalyxin-36

Ovocalyxin-36 (OCX-36) is a protein belonging to the bactericidal/permeability-increasing (BPI), lipopolysaccharide-binding proteins (LBP), and palate, lung and nasal epithelial clone (PLUNC) protein family (Chiang et al., 2011; Krasity et al., 2011; Baron et al., 2013). OCX-36 was first identified in the chicken eggshell; expression of its gene was detected in the uterus and to a lesser degree in red isthmus, which are located where eggshell mineralization occurs (Gautron et al., 2007). This protein is detected in uterine fluid and throughout the entire eggshell, especially at the inner part of the shell and at the mammillary layer (Gautron et al., 2007; Mikšík et al., 2007). Purification of OCX-36 revealed its antimicrobial activity against *Staphylococcus aureus,* and ability to bind lipopolysaccharide (LPS) from *Escherichia coli* and *S. aureus* lipoteichoic acid (LTA) (Cordeiro et al., 2013). These results support the proposed involvement of OCX-36 in the innate immune response, similar to other homologous members of the BPI/LBP/PLUNC family (Gautron et al., 2007, 2011). The OCX-36 protein sequence is composed of two lipid-binding domains BPI1 (BPI/LBP/CETP N-terminal domain) and BPI2 (BPI/LBP/CETP C-terminal domain) of about 200 amino acids each (**Supplementary Figure 1**) (Gautron et al., 2007). OCX-36 was initially thought to be eggshell-specific since this protein was first identified in chicken eggshell membranes and eggshell organic matrix. However, in addition to the distal oviduct, it is also expressed in the chicken intestine (Gautron et al., 2007; Tian et al., 2010; Chiang et al., 2011).

The BPI/LBP/PLUNC protein family belongs to the TULIP (tubular lipid-binding) superfamily, which split into two groups before the last eukaryote common ancestor: SMP-like proteins (synaptotagmin-like, mitochondrial and lipid-binding proteins) and BPI-like proteins (Alva and Lupas, 2016). The BPI/LBP/PLUNC family is only present in animals (Alva and Lupas, 2016). Indeed, members of this gene family are found in both vertebrate and invertebrate species (Solstad et al., 2007; Chiang et al., 2011; Krasity et al., 2011). In vertebrates, although 20–30% of amino acid identity was observed between chicken OCX-36 and other BPI family B proteins (also called LPLUNCs) (**Supplementary Table 1**), the similar organization of exons/introns in members of this gene family strongly suggests a common origin by multiple duplication events (Gautron et al., 2007; Tian et al., 2010). Synteny analysis of this gene family

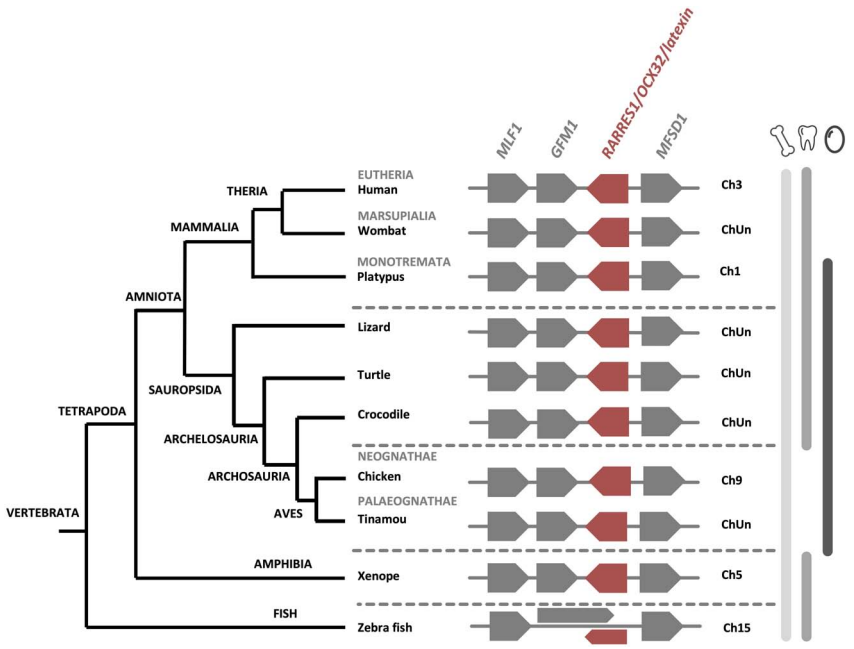


FIGURE 4 | Synteny of *RARRES1/OCX32* genes in vertebrates. The *RARRES1/OCX32* gene is represented by a red box and the flanking genes *MLF1* (myeloid leukemia factor 1), *GFM1* (G elongation factor mitochondrial 1) and *MFSD1* (major facilitator superfamily domain containing 1) are represented by gray boxes. ChUn, chromosome unknown. Gene IDs are listed in **Supplementary Table 2**. The left part of the figure depicts the phylogenetic relationship between vertebrate species (adapted from www.tolweb.org and Kapusta et al., 2017).

confirms a common ancestor for the genes encoding chicken OCX-36, *BPIFB3* (gene synonym: *OCX36*) and other *BPI family B* members (Tian et al., 2010; Gautron et al., 2011) (**Figure 5** and **Supplementary Table 1**). Since these previous analyses were performed, substantial new genomic and transcriptomic data has enriched this story. The synteny presented in **Figure 5** shows the presence of a *BPIFB3*/*OCX36* orthologous gene in reptiles (turtle and alligator), and other bird species from Palaeognathae (kiwi), Neoaves (zebra finch), and Galloanserae (duck). In addition, analysis of the platypus genome (*Ornithorhynchus anatinus*), an egg-laying mammal (Monotremata), reveals the presence of *BPIFB4-like* gene at the same location as *BPIFB3/OCX36* in birds and reptiles (**Figure 5**). Identity and similarity levels are higher between platypus BPIFB4-like and chicken OCX-36 than between chicken OCX-36 and other chicken BPIFB paralogs (**Supplementary Table 1**), suggesting that platypus BPIFB4-like is the ortholog of avian OCX-36. Phylogenetic analysis of OCX-36 and its relatives shows a first cluster containing chicken OCX-36, BPIFB4-like from other birds, OCX-36-like from reptiles and BPIFB4-like from platypus (**Figure 6**). This cluster, in addition to the results of synteny analysis, strongly suggests that all these genes are orthologs of chicken OCX-36. These new inputs expand the presence of *BPIFB3/OCX36* orthologous genes to Archelosauria (turtles, crocodiles, and birds) and Monotremata (egg-laying mammals) phyla. Finally, these new insights invalidate the previous hypothesis that *BPIFB3*/*OCX36* arose after the divergence of birds and mammals (Tian et al., 2010; Chiang et al., 2011; Gautron et al., 2011). The phylogenetic tree coupled with synteny strongly support that *BPIFB3*/*OCX36* appeared before the divergence of birds and mammals, which was likely lost in therian mammals (placentals and marsupials) (**Figures 5**, **6**). This phylogeny indicates that another member, *TENP* (transiently expressed in neural precursors), is the oldest gene in the BPI/LBP/LPLUNC family, and that the *BPIFB3/OCX36* gene is the result of three duplication events before tetrapod diversification and one event in amniotes.

New support for the specificity of OCX-36 protein orthologs to eggshell should be sought by investigation of diverse eggshell proteomes such as ratites, reptiles and monotremes. In the shell proteome of a crocodilian egg, an OCX-36 ortholog was not identified, but, as in birds, its paralog TENP-like is present in the shell organic matrix (Mann et al., 2006; Mikšík et al., 2018). In chicken, TENP is also found in egg white, vitelline membranes and egg yolk (Guérin-Dubiard et al., 2006; Mann, 2007; D'Ambrosio et al., 2008; Mann and Mann, 2008; Farinazzo et al., 2009). Apparently, TENP was recruited to the egg immune system in birds, while mammalian orthologs, BPIL1/BPIFB2/LPLUNC2, were recruited in ORL (olfactory epithelium, larynx, and tongue) tissue immunity (Andrault et al., 2003).

Co-option of SIBLING Bone Proteins, Osteopontin (OPN/SPP1) and Ovocleidin-116 (OC-116/MEPE) in Eggshell Biomineralization

OPN is a phosphoprotein (SPP1, secreted phosphoprotein 1, is the mammalian ortholog) found in both avian bone and

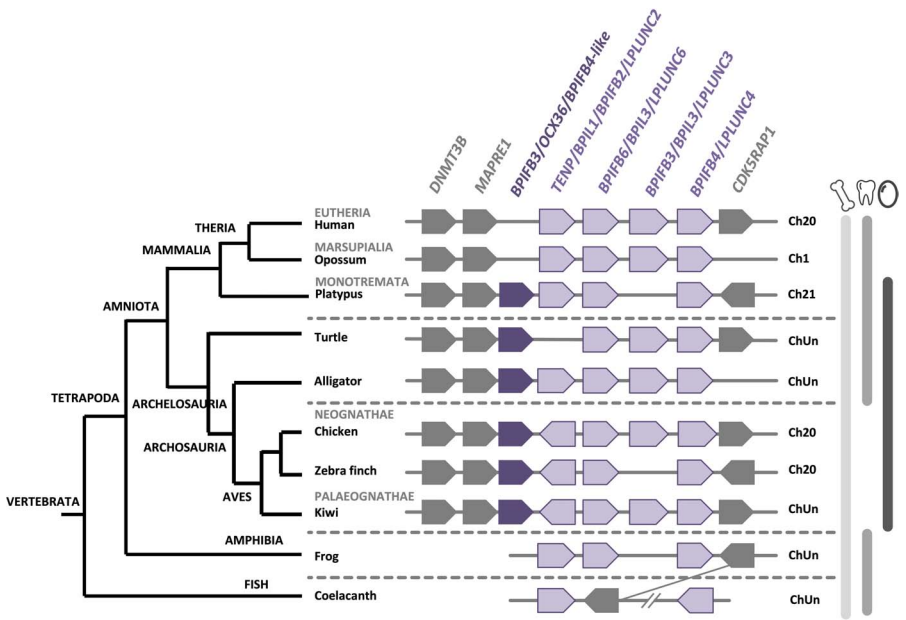


FIGURE 5 | Synteny of *BPIFB3/OCX36* genes in birds and its relatives in vertebrates. The *BPIFB3/OCX36* gene is represented by a dark purple box and its relatives are represented by light purple boxes. Flanking genes DNA (cytosine-5)-methyltransferase 3B (*DNMT3B*), Microtubule-associated protein RP/EB family member 1 (*MAPRE1*) and Mitochondrial tRNA methylthiotransferase (*CDK5RAP1*) are represented by gray boxes. ChUn, chromosome unknown. Gene IDs are listed in **Supplementary Table 2**. The left part of the figure depicts the phylogenetic relationship between vertebrate species (adapted from www.tolweb.org and Kapusta et al., 2017).

eggshell, as well as a variety of other tissue and cell types (Moore et al., 1991; Pines et al., 1995; Sodek et al., 2000; Fernandez et al., 2003; Chien et al., 2009; Hincke et al., 2012). In mice, OPN is strongly implicated in bone remodeling and fracture healing (McKee et al., 2011). In chicken, the oviduct expression of the OPN gene (*SPP1*) is entirely uterine-specific and is temporally associated with eggshell calcification through mechano-transcriptomic coupling of physical distension of the uterine wall to *SPP1* expression (Pines et al., 1995; Lavelin et al., 1998). Moreover, unusual patterns of uterine *SPP1* expression are associated with eggshell mineralization defects (Arazi et al., 2009). Localization by colloidal gold immunocytochemistry shows that OPN is concentrated in the palisade layer of the eggshell, where it is associated with parallel protein sheets of organic matrix, and more diffusely with the (104) crystallographic faces of eggshell calcite (Chien et al., 2008, 2009; Hincke et al., 2008). Specific OPN binding to the growing (104) crystal face during mineralization could modify the resistance of the shell to fracture along this plane. A functional interaction between OPN and the (104) eggshell calcite faces was supported by *in vitro* studies where synthetic calcite crystal growth at the (104) face was inhibited by added OPN (Chien et al., 2008). Nanoindentation and atomic force microscopy measurements suggest that OPN influences eggshell hardness and nanostructure, which in turn control the mechanical properties of the shell (Athanasiadou et al., 2018). SNPs in chicken *SPP1* are associated with eggshell fracture toughness (Dunn et al., 2009; Romé and Le Roy, 2016), which supports a regulatory role for OPN in mineralization. From fishes to mammals, OPN possesses a poly-aspartate motif, which is able to bind calcium and mediates binding to the mineral surface (**Supplementary Figure 2**; Sodek et al., 2000; Athanasiadou et al., 2018; Weber, 2018). However, in birds and reptiles, the OPN protein sequence exhibits a unique feature, a histidine-rich region that is suspected to originate from a microbial gene *via* a horizontal gene transfer event in early reptiles (**Supplementary Figure 2**; Weber, 2018). In mollusk shell, perlinhibin is a histidine-rich protein that inhibits calcium carbonate crystallization (Mann et al., 2007b), suggesting that this motif in eggshell OPN could play a similar calcite-specific role. In addition to the histidine-rich region, the C-terminal region is different between reptiles and non-reptilian vertebrates. In reptiles, the C-terminus is highly conserved, which supports a specialization of this protein with an important functional role in this vertebrate group (**Supplementary Figure 2**).

Ovocleidin-116 (OC-116; MEPE, matrix extracellular phosphoglycoprotein, is the mammalian ortholog) is a major component of the chicken uterine fluid, and is the most abundant matrix protein in the eggshell (Hincke et al., 1999; Mann et al., 2002, 2006; Marie et al., 2015a). It is an eggshell dermatan sulfate proteoglycan, which also possesses two *N*-glycosylated sites as well as *N*-glycan structures with fucosylated LacdiNAc (Nimtz et al., 2004). Immunostaining of the decalcified eggshell demonstrated the presence of OC-116 throughout the palisade layer and in the mammillary cone layer (Hincke et al., 1999). OC-116 is present in both soluble and insoluble fractions of the chicken eggshell matrix (Mann et al., 2002, 2006; Mikšík et al., 2007). Proteomics studies have identified OC-116 in the eggshell of chicken, turkey, quail, mallard duck and Guinea fowl, where

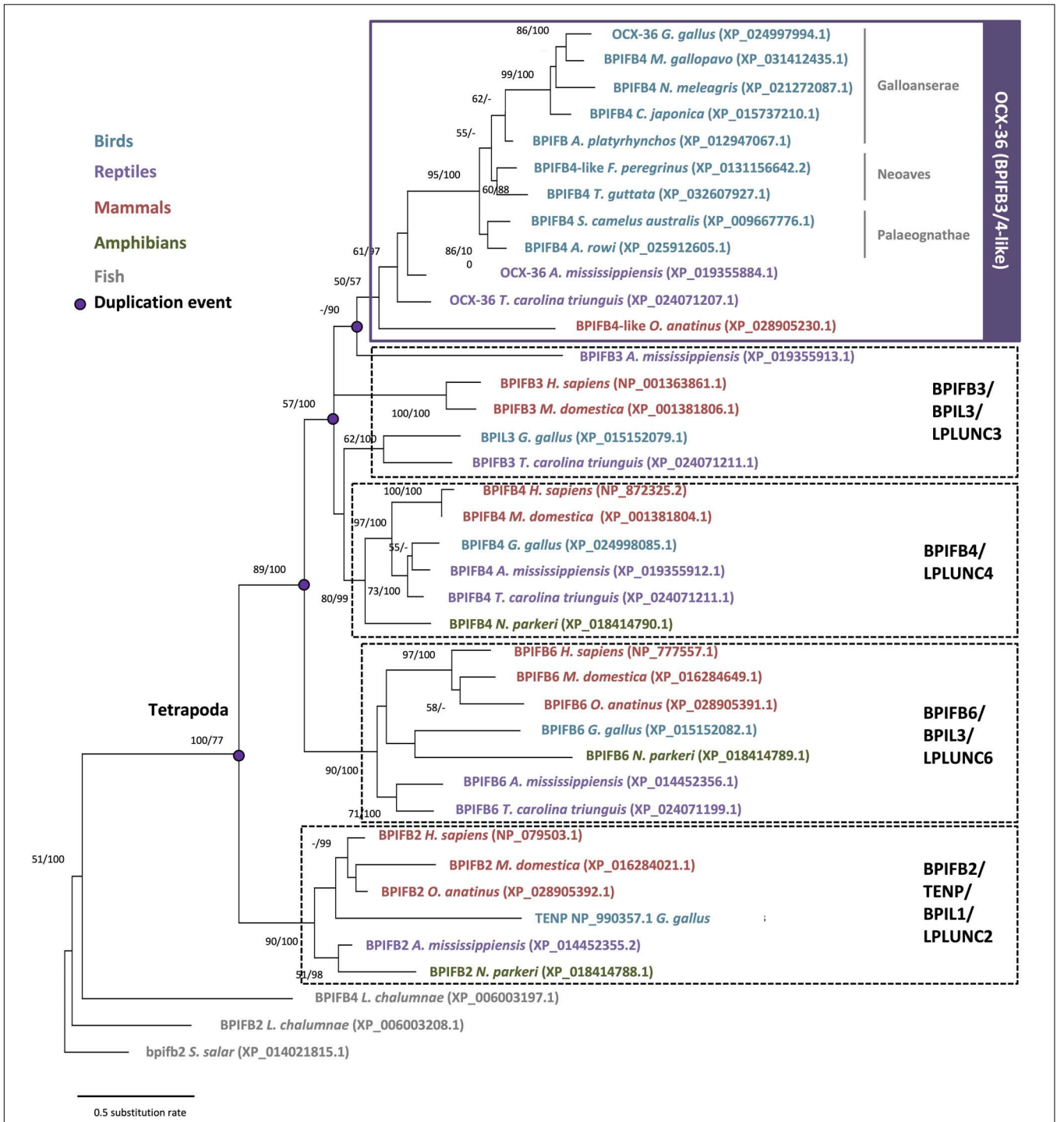


FIGURE 6 | Phylogenetic reconstruction of OCX-36 and its orthologs and paralogs in vertebrates. For phylogeny multiple alignment was performed using ClustalW and Gblocks (www.phylogeny.fr). Model of protein evolution used is JTT + G (MEGAX v10.1.8; https://www.megasoftware.net/). Topology of the tree corresponds to maximum likelihood method with 100 repetitions (MEGAX v10.1.8). Bootstrap values from maximum likelihood/Bayesian inference are indicated at each node when value is >50.

it is one of the most abundant eggshell constituent (Mann and Mann, 2013, 2015; Marie et al., 2015a; Le Roy et al., 2019; Zhu et al., 2019). In addition to the eggshell, OC-116 was reported in the chicken cortical and medullary bone (Horvat-Gordon et al., 2008). SNP analysis revealed that the *MEPE/OC116* gene is associated with shell thickness, elastic modulus and egg shape

(Dunn et al., 2009; Romé and Le Roy, 2016). The mammalian ortholog MEPE is involved in bone and teeth mineralization (Bardet et al., 2010a). Inactivation of the *Mepe* gene in mouse causes an increase in bone mass and mineralization due to an increase in osteoblast number and activity (Gowen et al., 2003). The role of OC-116/MEPE in mineralization is supported by the ASARM (acidic serine-aspartate rich MEPE-associated motif) sequence located at the C-terminus of the protein. When ASARM is phosphorylated, it inhibits mineralization by binding to hydroxyapatite crystals (Addison et al., 2008). This peptide is also involved in phosphatemia regulation (Rowe et al., 2004). Multiple sequence alignment of OC-116/MEPE proteins shows a high conservation of the ASARM peptide throughout tetrapods (**Figure 7**). The presence of numerous putative phosphorylation sites (7 in bird orthologs) suggests that its role in mineralization is also conserved in both bone and eggshell.

Both OPN/SPP1 and OC-116/MEPE belong to the SIBLING (small integrin-binding ligand N-linked glycoprotein) family, with three other protein members: integrin-binding sialoprotein (IBSP), dentin sialophosphoprotein (DSPP) and dentin matrix protein 1 (DMP1). Genes coding for these five proteins are clustered together through tetrapods (**Figure 8A**), and they all have a role in biomineralization (Rowe, 2012); however, none of them appears to be specific to calcium carbonate (eggshell) or calcium phosphate (bone, teeth) mineralization. They possess similar molecular properties such as integrin-binding and calcium-binding (Bardet et al., 2010b). Among the five SIBLING members, OPN and OC-116 have been widely studied in the chicken eggshell, as described above. Moreover, DMP1 and IBSP were detected in the eggshell matrix (by proteomics and Western blotting), and their genes are expressed in uterine tissue (Mann et al., 2006; Horvat-Gordon et al., 2008); however, *DSPP* gene, involved in dentin formation, was secondarily lost in ancestors of birds during late Cretaceous when they become toothless (**Figure 8A**) (Kawasaki and Weiss, 2008; Sire et al., 2008; Mcknight and Fisher, 2009; Kawasaki, 2011; Sire and Kawasaki, 2012). It has been suggested that the entire SIBLING gene family, including OC-116/MEPE and OPN/SPP1, which were initially involved in bone formation, was co-opted for the eggshell calcium carbonate mineralization process in birds (Sire and Kawasaki, 2012). In the phylogenetic reconstruction of OC-116 and OPN, the distribution of both proteins follows the accepted phylogenetic relationships in tetrapods, with a clear split between mammals and reptiles including birds (**Figures 8B,C**). In the Aves clade, Palaeognathae is the basal group; while in Neognathae, Neoaves sequences are well separated from Galloanserae for OPN, and from Galliformae for OC-116. In the recent crocodilian eggshell proteome, OC-116 was identified but not OPN (Mikšík et al., 2018), suggesting the recruitment of OC-116, at least for eggshell mineralization, in the Archosauria linage (Aves and Crocodylia). In seven turtle species and platypus (*O. anatinus*, the only monotreme genome in NCBI database), although gene coding for SPP1 is present, synteny analysis reveals the absence of *MEPE/OC116* from the SIBLING locus (**Figure 8A**), which implies a loss of this gene in these two lineages.

The Glycoprotein EDIL3: A Novel Candidate for Calcium Carbonate Delivery in Eggshell Mineralization

The glycoprotein EDIL3 (EGF-like repeats and discoidin domains 3) was identified in the chicken eggshell by proteomics analysis (Mann et al., 2006; Marie et al., 2015a). The EDIL3 sequence contains three EGF-like domains and two F5/8C (discoidin) domains; it was first identified as an extracellular matrix protein involved in embryonic vascular development in mouse (Hidai et al., 1998). The three EGF-like domains are present in EDIL3 orthologs in all vertebrates except in fishes (Stapane et al., 2019), and the third domain possesses a calcium-binding site, which suggests a potential role of EDIL3 in calcium carbonate crystallization (Marie et al., 2015a). In addition, an RGD (arginine, glycine, and aspartate) motif is present in the second EGF-like domain, through which it can bind integrins (Stapane et al., 2019). Integrins are transmembrane proteins involved in cell-cell and cell-extracellular matrix interactions, and vesicular trafficking (Théry et al., 1999; Hynes, 2002; Gatti et al., 2005; Bridgewater et al., 2012). The second F5/8C domain exhibits a phospholipid-binding site, which give to the protein the ability to complex vesicle and/or cell membranes (**Supplementary Figure 3A**; Stapane et al., 2019). In the EDIL3 sequence of some bird species (*G. gallus*, *A. platyrhynchos*, *Aquila chrysaetos*, *Dromaius novaehollandiae*, etc.), the first EGF-like domain also contains an RGD motif suggesting an even higher affinity of the protein for integrins (**Supplementary Figure 3B**). Proteomics demonstrate the presence of EDIL3 in bird and crocodile eggshell (Marie et al., 2015a; Mikšík et al., 2018; Le Roy et al., 2019). However, in the eggshell proteome of *Crocodylus siamensis*, EDIL3 protein was identified with six peptides matching with EDIL3 from *Alligator mississippiensis* (NCBI accession KYO21076.1) (Mikšík et al., 2018). Surprisingly, amongst the six peptides, only one matched with a domain present in EDIL3 proteins (first F5/8C domain); however, the five other peptides matched with the IG-like (immunoglobulin-like) and LINK_2 domain (hyaluronan-binding region) that is found in HAPLN1 (hyaluronan and proteoglycan link protein 1). Moreover, the annotated *A. mississippiensis* EDIL3 protein is 843 aa in length instead of about 480 aa for the other EDIL3 proteins, and seems to be composed of both EDIL3 and HAPLN1 protein features (**Supplementary Figure 3C**). Indeed, the N-terminal part of alligator EDIL3 shows 53.8% of identity with chicken EDIL3 and the C-terminal part of alligator EDIL3 possesses 54.3% of identity with chicken HAPLN1. These contradictions indicate that the potential identification of the EDIL3 ortholog in crocodilian eggshell requires confirmation (**Supplementary Figure 3**). In chicken, EDIL3 is not an eggshell-specific protein, although it exhibits a high relative abundance in the shell OM (Marie et al., 2015a). According to the emPAI (exponentially modified protein abundance index) values of proteins from chicken eggshell proteome at four calcification stages, EDIL3 is the fifth most abundant protein at the early stages of biomineralization, corresponding to the transformation of ACC into calcite crystals (Marie et al., 2015a).

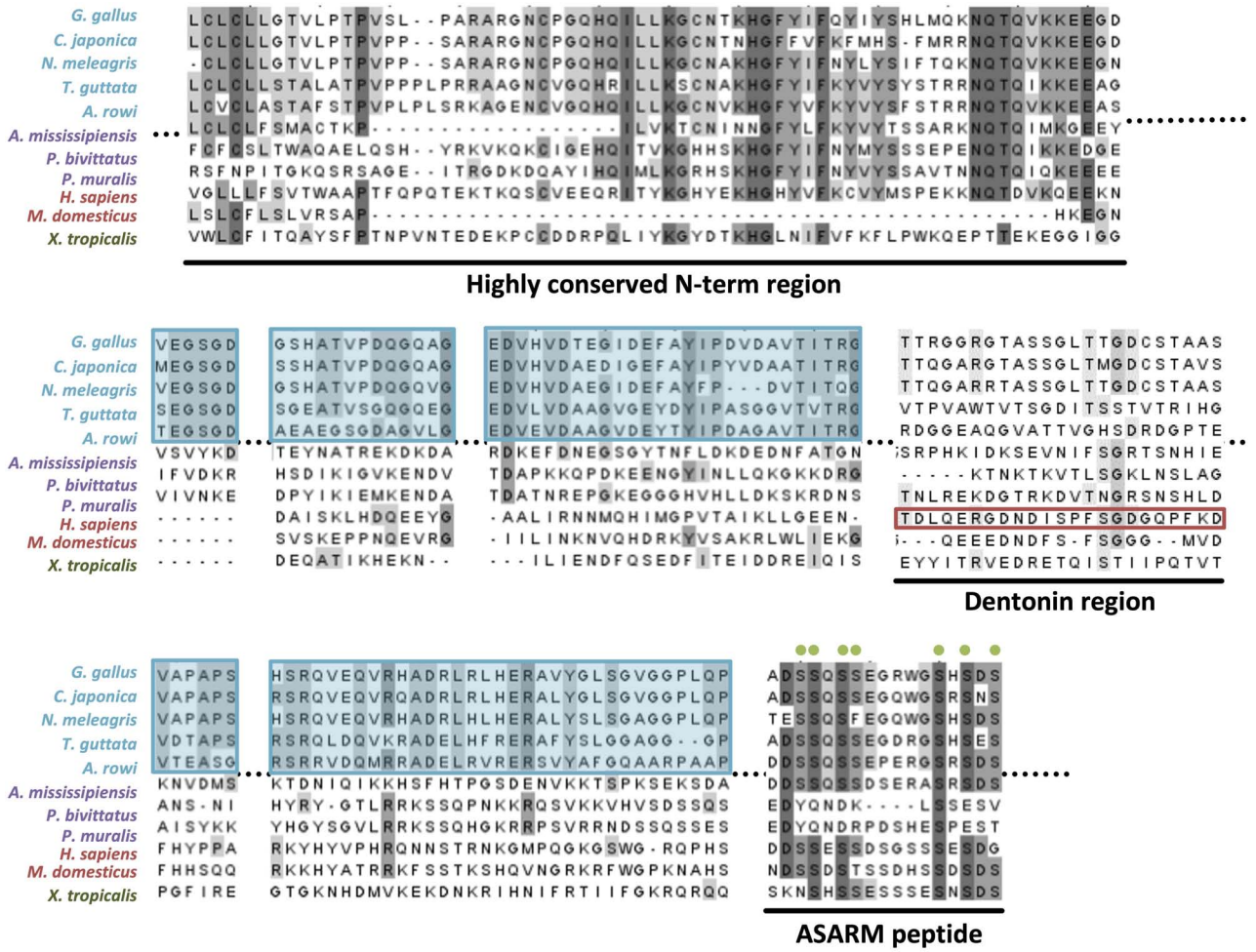


FIGURE 7 | Multiple alignment of OC-116/MEPE proteins in tetrapods. The alignment was built with ClustalW (www.phylogeny.fr) and edited with Jalview v2.11.1.4 (https://www.jalview.org/). Gray colors of letters correspond to percentage of identity (dark gray > 80% identity, medium gray > 60% identity, and light gray > 40% identity). The dentonin region identified in human MEPE, and regions with amino acid consensus in birds, are, respectively, framed in red and blue. Putative phosphorylated sites of the ASARM (acidic, serine, and aspartic acid-rich motif) peptide are indicated by green dots above the alignment. Sequences aligned are from *Gallus gallus* (AAF00982.3), *Coturnix japonica* (XP_015716951.1), *Numida meleagris* (XP_021251779.1), *Taeniopygia guttata* (XP_030127798.1), *Apteryx rowi* (XP_025920573.1), *Alligator mississippiensis* (XP_019343676.1), *Python bivittatus* (XP_025029286.1), *Podarcis muralis* (XP_028598766.1), *Homo sapiens* (XP_006714341.1), *Monodelphis domestica* (XP_007495981.1), and *Xenopus tropicalis* (XP_002938672.2).

In chicken, *EDIL3* gene expression is up-regulated in isthmus and uterus compared with bone, duodenum, kidney, liver and magnum, and is significantly higher in the oviduct segments at early stages (6 and 7 h post-ovulation, initiation of mineralization) than at 16 h post-ovulation (mid-calcification) (Stapane et al., 2019, 2020). Immunohistochemistry in uterine cross-sections confirms the presence of high levels of EDIL3 at the early stages of mineralization in tubular gland cells (5 and 6 h post-ovulation) (Stapane et al., 2020). Moreover, proteomics and Western blot analyses revealed the presence of EDIL3 in extracellular vesicles isolated from chicken uterine fluid (Stapane et al., 2019, 2020). These vesicles are proposed to mediate the transportation of ACC to the mineralization site. Indeed, vesicles have been demonstrated to play roles in ACC stabilization in invertebrate biomineralization models such as sea urchins, molluscan shell and coral skeleton (Levi-Kalisman et al., 2002;

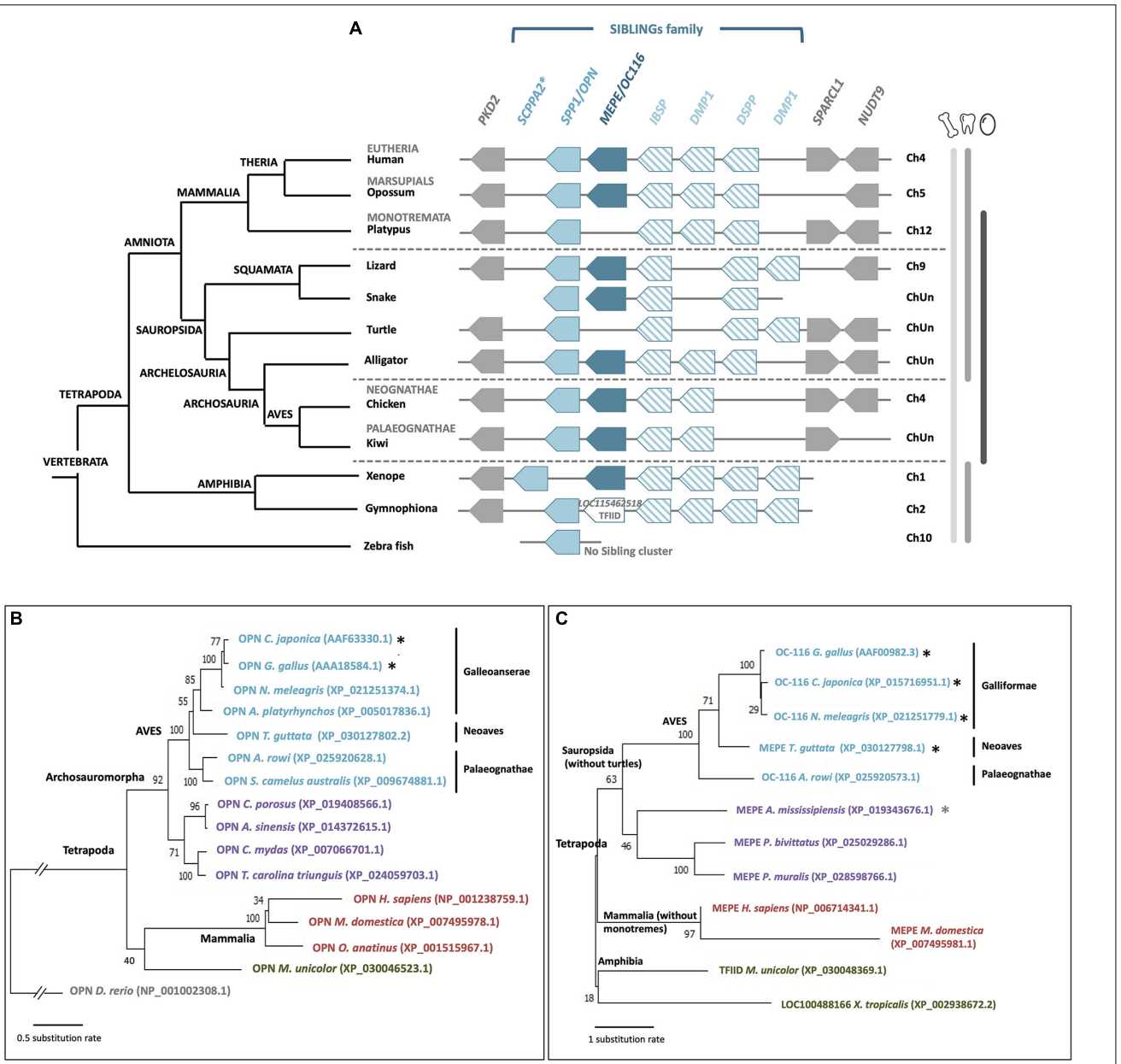


FIGURE 8 | Synteny and phylogeny of *OPN/SPP1* and *MEPE/OC116* genes and corresponding proteins in vertebrates. **(A)** *OPN/SPP1* is represented by a light blue box and *MEPE/OC116* is represented by a dark blue box. The other SIBLING genes are integrin-binding sialoprotein (*IBSP*), dentin sialophosphoprotein (*DSPP*) and dentin matrix protein 1 (*DMP1*), which are represented by empty boxes with oblique blue lines. In Gymnophiona (amphibian), instead of *MEPE* we observed the presence of the transcription initiation factor TFIID subunit 1-like gene (*LOC115462518*, XP_030048369.1). Flanking genes are polycystin 2 (*PKD2*), SPARC-like 1 (*SPARCL1*) and nudix hydrolase 9 (*NUDT9*), which are represented by gray boxes. ChUn: Chromosome Unknown. Blue asterisk indicates that *SCPPA2* is a hypothetical ortholog of *spp1* (Kawasaki, 2011). Gene IDs are listed in **Supplementary Table 2**. In the left part, the phylogenetic tree for vertebrate species is represented (adapted from www.tolweb.org and Kapusta et al., 2017). **(B)** Phylogeny of OPN/SPP1 in vertebrates, which was reconstructed using multiple alignment performed with ClustalW and Gblocks (www.phylogeny.fr) and JTT + G model of protein evolution (MEGAX v10.1.8; https://www.megasoftware.net/). Topology of the tree corresponds to maximum likelihood method (MEGAX v10.1.8). Bootstrap values from maximum likelihood are indicated at each node. The symbol//indicates a gap of 0.8 substitution rate to add to the basal branch. Black asterisks indicate species where the protein was identified in the eggshell. **(C)** Phylogenetic reconstruction of OC-116/MEPE in vertebrates was performed using multiple alignment constructed with ClustalW and Gblocks (www.phylogeny.fr) and JTT + G + F model of protein evolution (MEGAX v10.1.8). Topology of the tree corresponds to maximum likelihood method (MEGAX v10.1.8). Bootstrap values from maximum likelihood are indicated at each node. Black asterisks indicate species where the protein was identified in the eggshell and the gray asterisk indicates a related species (*C. siamensis*) where the protein was identified in the eggshell.

Addadi et al., 2003; Weiner and Dove, 2003; Mass et al., 2017). In chicken uterus samples examined by transmission electron microscopy, extracellular vesicles are observed in uterine cells and fluid, and vesicles are seen budding from cells into the uterine lumen. Energy-dispersive electron spectroscopy and selected area electron diffraction revealed the presence of ACC in extracellular vesicles purified from the uterine fluid (Stapane et al., 2020). Based on these and other results, extracellular vesicles are proposed to play a role in ACC-mediated calcification of the eggshell. EDIL3 is proposed to bind vesicle membrane (phospholipid-binding site/integrin-binding site) and to guide these vesicles from uterine cell cytosol to the mineralization site (calcium-binding site/integrin-binding site) in the uterine fluid of chicken (Stapane et al., 2019, 2020).

The gene *EDIL3* is highly conserved in vertebrates (**Supplementary Figure 4**) (Stapane et al., 2019). The flanking genes in this locus are versican (*VCAN*) and hyaluronan and proteoglycan link protein 1 (*HAPLN1*). Phylogenetic analysis of EDIL3 and its paralog MFGE8 in animals demonstrates the appearance of both proteins after a duplication event in vertebrates 480 MYA (Stapane et al., 2019). EDIL3 was subsequently recruited to the eggshell mineralization process, at least in the Aves phylum. Interestingly, its paralog MFGE8 is also detected in the eggshell proteome of birds; however, its abundance is much lower than EDIL3, and *MFGE8* expression is not specific to tissues responsible for eggshell mineralization or to the initial stages of mineralization (Stapane et al., 2019).

A PROTEIN SPECIFIC TO THE EGGSHELL BIOMINERALIZATION PROCESS

C-Type Lectin Proteins in Eggshell Organic Matrix: Ovocleidin-17 Homologs

The C-type lectin protein Ovocleidin-17 (OC-17) is an eggshell-specific protein, which was first purified and partially sequenced from the chicken eggshell (Hincke et al., 1995). The mRNA sequence was determined only recently by *de novo* transcriptomic assembly (Zhang et al., 2014). OC-17 contains a C-type lectin (CTL) domain and possesses two phosphorylated serine residues (Mann, 1999; Mann and Siedler, 1999). The CTL proteins are a huge family of proteins including at least seven subgroups such as hyalectans, asialoglycoprotein receptors, collectins, selectins, natural killer group transmembrane receptors, macrophage mannose receptors and simple lectins (Zelensky and Gready, 2005). OC-17 and its homologs correspond to a simple lectin, with a short amino acid sequence (about 150 aa) and only one CTL domain. Proteomics analysis demonstrated that OC-17 is a highly abundant protein in the eggshell matrix in chicken and Guinea fowl (Marie et al., 2015a; Le Roy et al., 2019). Moreover, CTL proteins that are homologs of OC-17 have been identified in eggshells of many bird species, including ostrich, emu, and rhea (Mann and Siedler, 2004, 2006). In each of these ratites, two homologous CTL eggshell proteins were identified and named according to the bird species: Struthiocalcin-1 and 2 (SCA-1 and -2) for ostrich, Dromaiocalcin-1 and -2 (DCA-1 and -2) for emu and Rheacalcin-1 and -2 (RCA-1 and -2) for rhea. For easier reading in the present review, we have termed these proteins XCA-1 and XCA-2. In contrast, only one CTL protein (OC-17) is present in chicken eggshell, which aligns with the XCA-2 group of other bird species.

C-type lectin proteins have been identified in the biomineralization process of invertebrates. For instance, in the sea urchin *Strongylocentrotus purpuratus*, SM50 is a protein containing a C-type lectin domain in addition to glycine-rich and proline-rich regions. The study of this C-type lectin domain revealed that it influences the biomineralization of $CaCO_3$ (Rao et al., 2013). In the same manner, in the freshwater pearl mussel, a C-type lectin protein called perlucin, already identified in the shell proteomes of mollusks, is involved in nacre formation (Lin et al., 2013). Purified OC-17 modifies calcite crystallization *in vitro* (Reyes-Grajeda et al., 2004). *In silico* molecular dynamics simulations suggest three protein configurations of OC-17, which is able to bind calcium carbonate surfaces through its positively charged guanidino group of specific arginine residues (Freeman et al., 2010, 2011). Thus, CTL proteins could play a role in eggshell formation by binding to specific calcite crystal faces (Wallace and Schiffbauer, 2016). In addition, chicken OC-17 and its goose ortholog (ansocalcin) exhibit an antimicrobial activity, and could play a potential role in innate immunity of the avian embryo (Wellman-Labadie et al., 2008a). Although the presence of one versus two OC-17 paralogs in the eggshell of Palaeognathae birds has been proposed to correlate with eggshell thickness (Mann and Mann, 2015), there is currently no experimental evidence to support this hypothesis.

Synteny, protein multiple alignment and phylogenetic analysis of XCA-1, XCA-2 and OC-17-like confirm that XCA-2 is ortholog to OC-17-like and XCA-1 is paralog to OC-17-like/XCA-2 (**Figures 9**, **10**, and **Supplementary Figure 5**). OC-17-like and XCAs are also similar to other vertebrate C-type lectin (CTL) proteins, such as REG4 (Regenerating Islet-derived protein 4) and Lithostathine (also known as REG1A and REG1B) in humans. REG1 is a pancreatic CTL protein involved in the inhibition of $CaCO_3$ precipitation in the bicarbonate-rich pancreatic juice (Bernard et al., 1992). The gene encoding REG1/Lithostathine is not found in the genome of reptiles and birds, whereas *REG4* is present in numerous bird and crocodilian species. The OC-17-like/XCAs gene symbol in reptiles and birds is different for each species (i.e., *LOC numbers*; **Supplementary Table 3**); in order to simplify our discussion, we use the arbitrary nomenclature *OC17-like/XCA2* and *XCA1*. Synteny analysis shows that the *REG4* gene is located on a different chromosome than the *OC17-like/XCA2* and *XCA1* gene locus (e.g., in tufted duck *REG4* is located on chromosome 8 and *OC17-like/XCA2* and *XCA1* are located on chromosome 1; **Figure 9**). *REG4* is flanked by the same genes in crocodiles, birds and in mammals, but it is absent in turtles and lepidosaurs (lizards, snakes, etc.) at the same locus. On the other hand, *OC17-like/XCA2* and/or *XCA1* genes are clustered together and flanked by the same genes in birds and reptiles; however, they are absent from the same locus in mammalian and amphibian genomes. These data support the hypothesis that OC-17-like/XCA-2 and XCA-1 are eggshell specific proteins and that

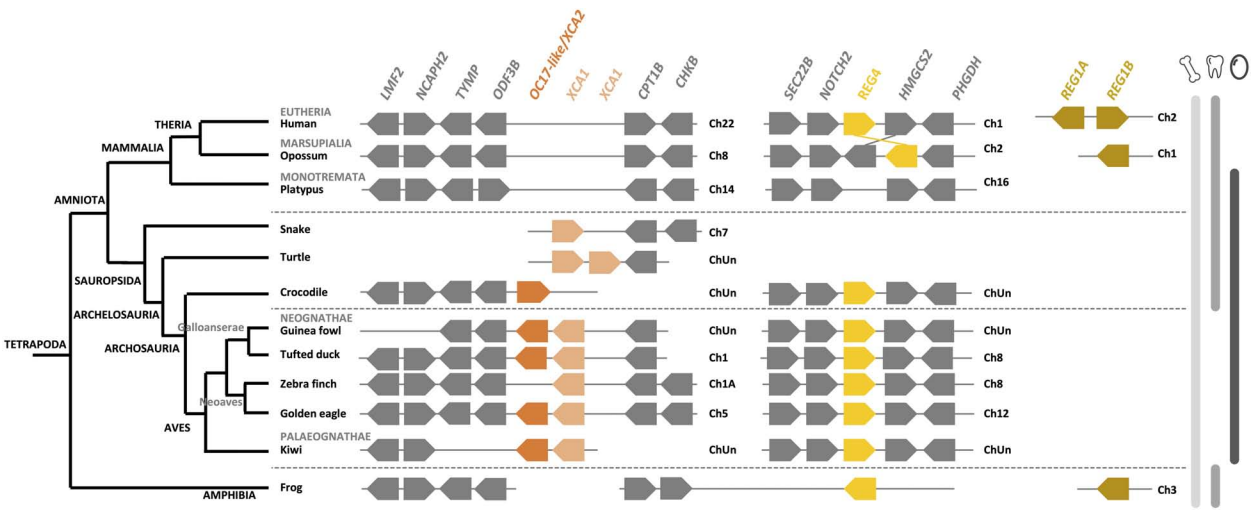


FIGURE 9 | Synteny of *OC17-like*, *XCA* and *REG* orthologs and paralogs in birds and its relatives in vertebrates. *OC17-like/XCA2* is represented by a dark orange box and *XCA1* by a light orange box. *REG4* (regenerating Islet-derived protein 4) is represented by a yellow box and *REG1 A* and *B* (Lithostathine) are represented by a brown box. Flanking genes are represented by gray boxes: lipase maturation factor 2 (*LMF2*), condensing-2 complex subunit H2 (*NCAPH2*), thymidine phosphorylase (*TYMP*), outer dense fiber protein 3B (*ODF3B*), carnitine *O*-palmitoyltransferase 1 (*CPT1B*), choline/ethanolamine kinase (*CHKB*), vesicle-trafficking protein (*SCE22B*), neurogenic locus notch homolog protein 2 (*NOTCH2*), hydroxymethylglutaryl-CoA synthase (*HMGCS2*) and D-3-phosphoglycerate dehydrogenase (*PHGDH*). Gene IDs are listed in **Supplementary Table 2**. The left part of the figure depicts the phylogenetic relationship between vertebrate species (adapted from www.tolweb.org and Kapusta et al., 2017).

these genes are specific to vertebrates that produce a calcitic shell; however, no ortholog has yet been identified in the crocodilian eggshell proteome (Mikšík et al., 2018). Deeper investigation needs to be done in crocodilian eggshells, but also in other reptile eggshells (e.g., snakes, lizards, and turtles), in order to determine if OC-17-like/XCA-2 and/or XCA-1 are widespread in reptile and bird eggshells or if they strictly correspond to the bird eggshell biomineralization process.

The pairwise alignment of chicken REG4 and OC-17 amino acid sequences exhibits 29.2% identity (58.4% similarity), which supports a common origin of both proteins (**Supplementary Table 4**). Phylogenetic reconstruction shows that REG4 and OC-17-like/XCAs are divided into two distinct groups (**Figure 10**). Regarding these observations, the phylogeny indicates that OC-17-like/XCAs arose from a duplication event in Sauropsida. In this clade, squamates (lizards and snakes) are in basal position with only one form of XCA. Then, three groups split: bird OC-17-like/XCA-2 (including crocodilian XCA-2), turtle XCA-1 and bird XCA-1 (**Figure 10**). The synteny of *OC17-like/XCA2* and *XCA1* shows that duplication of the ancestral gene occurred on the same chromosome and the phylogeny suggests that *XCA1* is closer to the ancestral form of the duplicated gene. These observations might indicate that *OC17-like/XCA2* is the result of a duplication event in archosaurs with a loss of *XCA1* in crocodilians. In turtles, two *XCA1* paralogs are also present but they are clustered together suggesting an independent duplication event in the turtle phylum.

In birds, XCA paralogs are present in both Neognathae (Neoaves and Galloanserae), and Palaeognathae (ratites) phyla. Nevertheless, in Neognathae, each bird species does not possess the two paralogs inside sub groups. In Neoaves we notice that only XCA-1 is present in common starling (*Sturnus vulgaris*), falcons (*Falco cherrug* and *Falco rusticolis*) and zebra finch (*Taeniopygia guttata*), whereas both paralogs are present in golden eagle (*A. chrysaetos chrysaetos*) (**Supplementary Figure 6**). In Galloanserae, some species exhibit one paralog such as chicken (*G. gallus*) and pheasant (*Phasianus colchicus*) (Mann et al., 2006; Marie et al., 2015a), for which the genome position is unknown. Conversely, other Galloanserae species possess both paralogs such as Guinea fowl, black swan (*Cygnus atratus*) and tufted duck (*Athya fuligula*) (**Supplementary Figure 6**) (Le Roy et al., 2019). The mallard duck eggshell proteome demonstrated the presence of an ortholog to chicken OC-17 protein (Zhu et al., 2019), which may correspond to mallard SCA-2-like protein translated from the newly submitted (December 2020) mallard genome in NCBI (Accession XP_038024161.1; Gene ID: 119713911). In this genome, the gene coding for SCA-2-like is located next to RCA-1-like (Gene ID: 119713283) in the same gene cluster containing *OC17-like/XCA2* and *XCA1*, as observed in other birds (**Figure 9** and **Supplementary Figure 6**). In Palaeognathae, both paralogs are present in emu, ostrich and rhea eggshells, but in two *Apteryx* species (kiwis), for which eggshell proteomes are not available, one species exhibits two adjacent paralogs and the other species has only one paralog in its genome (**Supplementary Figure 6**). Nevertheless, the lack of genome data (gene sequencing and scaffolding genome assembly) is possibly the reason for the absence of the second paralog in all these species of birds, crocodilian etc. Indeed, the chicken OC-17 transcript has a very high GC content (72.17%), which could account for the observed difficulty to sequence this gene in the chicken genome

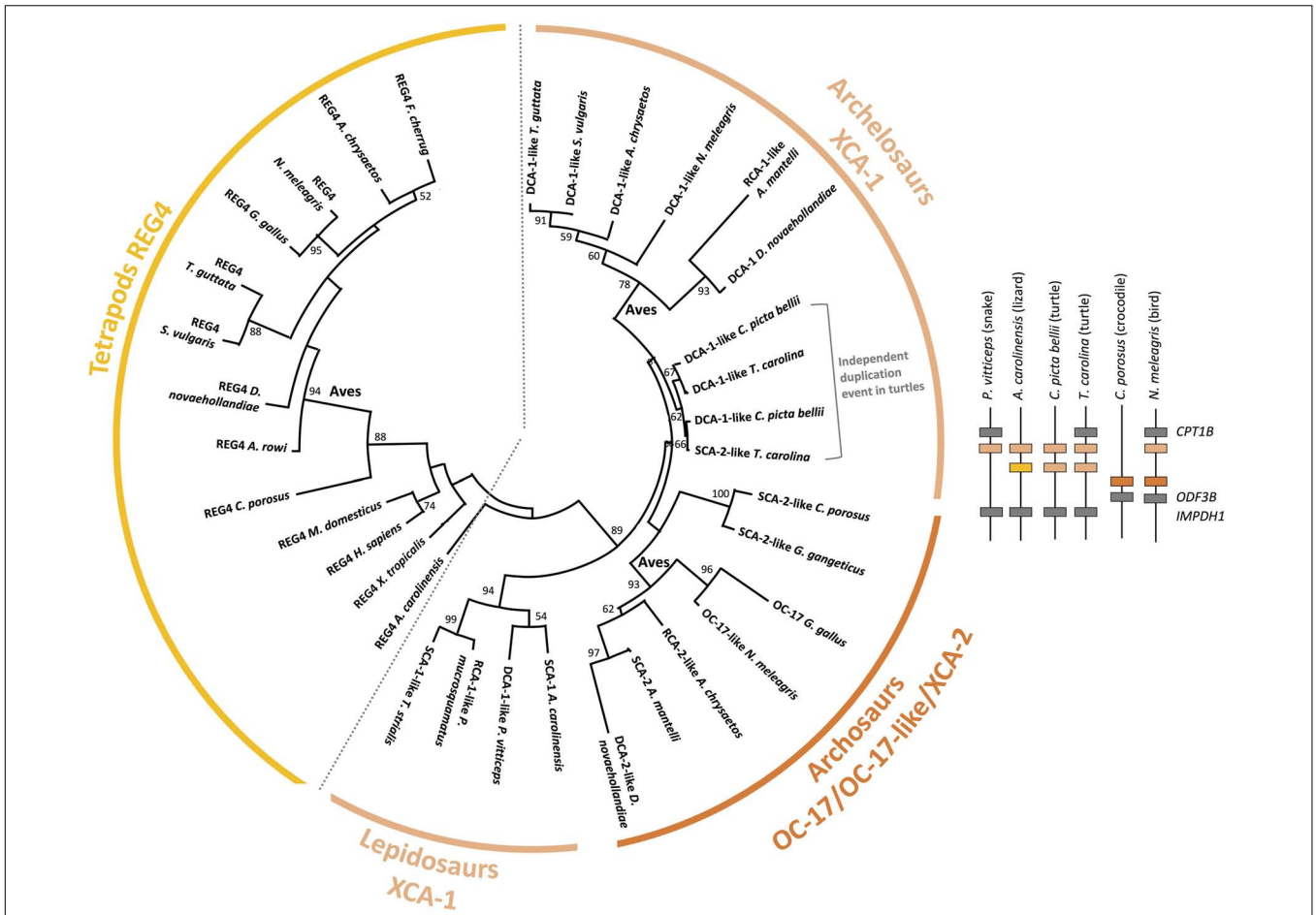


FIGURE 10 | Phylogenetic reconstruction of OC-17 and its orthologs and paralogs in vertebrates. Phylogeny was performed using the maximum likelihood method with 100 repetitions (MEGAX v10.1.8; https://www.megasoftware.net/) using ClustalW multiple alignment and Gblocks (www.phylogeny.fr), and the WAG + G model of protein evolution. Bootstrap values from maximum likelihood are indicated at each node when value is >50. Synteny of *OC-17-like/XCA-2* and *XCA-1* genes is represented for five reptilian species (*Pogona vitticeps*, *Anolis carolinensis*, *Chrysemys picta bellii*, *Terrapene carolina triunguis,* and *Crocodylus porosus*) and one bird (*Numida meleagris*).

and in genome projects of other bird species. Hence, this lack of *OC17-like/XCAs* annotation in bird genomes reduces the possibility to identify orthologous OC-17 peptides using proteomics approaches. This is the case in turkey and quail eggshell proteomes (Mann and Mann, 2013; Mann, 2015). In the Guinea fowl, OC-17-like (71.56% GC) and DCA-1-like (72.01% GC) were detected in its eggshell proteome, likely because the genome assembly of *Numida meleagris* that is available in the NCBI database (NumMel1.0) was built using the *G. gallus* genome.

EVOLUTION OF ORGANIC MATRIX PROTEINS IN EGGSHELL BIOMINERALIZATION

During amniote evolution, reproduction was freed from reliance on the aquatic environment with the emergence of two possible reproductive strategies: egg-laying vs. placentation. In sauropsids and some mammals (Monotremata), the egg-laying strategy was based on a soft or hard-shelled egg to protect the embryo. This adaptation has reached its most advanced development in birds that emerged 102 MYA. The avian eggshell proteome exhibits both co-opted proteins and eggshell-specific proteins. Over the last several decades, the development of high-throughput technologies has helped to characterize and enumerate the complexity of the organic matrices in multiple avian species. The evolution of the eggshell reproductive strategy in sauropsids was accompanied by (1) the recruitment of existing genes for eggshell biomineralization (OCX-32, EDIL3, OC-116/MEPE, and OPN), and (2) the birth of new genes from duplication events, which are highly specialized for this process (XCA-1 and OCX-36; **Figure 11**). It is intriguing that avian OPN orthologs appear to have acquired a His-rich domain which may be related to calcitic biomineralization. Therefore, insight into the common eggshell toolkit which is responsible for eggshell mineralization in birds is emerging.

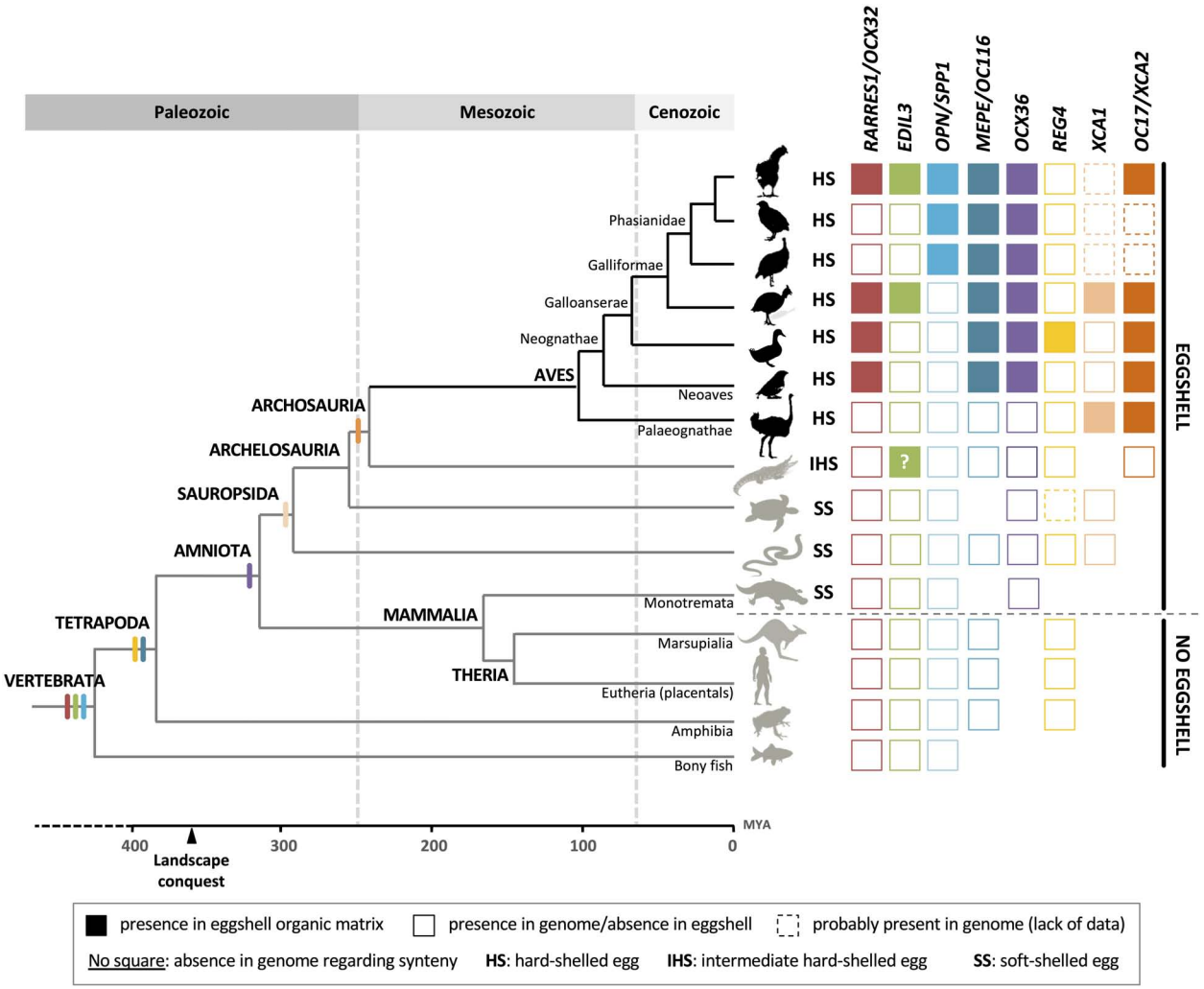


FIGURE 11 | Schematic summary of the presence of genes encoding investigated proteins in vertebrate genomes and the identification of their related proteins in Sauropsida eggshells. Filled squares indicate the presence of the protein in eggshell organic matrix whereas empty squares indicate the presence in genome but the absence in eggshell. Squares with a dotted line indicate a putative presence in genome (lack of data). The absence of square in the figure illustrates an absence in genome with respect to synteny. The "?" in the crocodilian EDIL3 square indicates the uncertainty of true identification of this protein in the crocodilian eggshell proteome (see **Supplementary Figure 3**). Genomic data can come from different species inside the same apical group. The divergent times and topology of the tree are from (www.tolweb.org; Warren et al., 2008; Phillips et al., 2009; Shen et al., 2011; Jarvis et al., 2014; Kapusta et al., 2017).

The huge outpouring of genomic data for bird species from the Bird 10,000 Genomes Project (2015–2020) is of vital importance to better understand the evolution of genes coding for eggshell proteins inside Aves (Palaeognathae vs. Neognathae; Neoaves vs. Galloanserae). In addition, this genomic database enrichment is crucial for future eggshell proteome studies in other bird groups (especially in Palaeognathae, for which only 2 proteins – the paralogs XCA-1 and XCA-2 – have yet been identified in the eggshell organic matrix). Finally, in order to enrich the evolutionary perspective, expanded analysis of eggshell proteomes should be performed in several reptilian species for which genomic/transcriptomic dataset are available, and in monotreme species such as platypus. Although the only crocodilian eggshell proteome exhibits common proteins with bird eggshell proteomes, it suffers from a very low number of identified proteins (58 against 904 for the chicken eggshell proteome), likely due to the absence of genomic/transcriptomic data for this species and incomplete annotation of genomic datasets of other crocodilians. And finally, the eggshell matrix databases must be enriched with post-translational modification (PTM) information, especially phosphorylation and glycosylation, which will permit cross-species comparisons for further insight. Currently, this information only exists in a comprehensive manner for the chicken eggshell phospho-proteome (Mann et al., 2007a).

This review identifies challenges and proposes new strategies to better understand the evolution of eggshell biomineralization, such as the multiplication of eggshell proteomics analyses in

basal birds (ratites), sister groups of birds (crocodiles, turtles, squamates), and in more distant groups (monotremes).

AUTHOR CONTRIBUTIONS

MH and NL coordinated writing of the manuscript and edited the review. JG and LS contributed to the manuscript, wrote part of the review, and approved the final version. All authors contributed to the article and approved the submitted version.

FUNDING

MH acknowledges funding from NSERC (RGPIN-2016-04410) and is grateful to Le STUDIUM for support during the preparation of this manuscript. He is a Le STUDIUM Research Fellow, Loire Valley Institute for Advanced Studies, Orleans-Tours, and BOA, INRAE, Centre Val de Loire, Nouzilly, France. JG thanks the Agence Nationale de la Recherche for funding (Impact project ANR-13-BSV-0007-01). The postdoctoral position of NL was funded by an Agreenskills fellowship. LS acknowledges the University of Tours and the "Region Centre" for financial support during his doctoral thesis.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2021.672433/full#supplementary-material

REFERENCES

Addadi, L., Raz, S., and Weiner, S. (2003). Taking advantage of disorder: amorphous calcium carbonate and its roles in biomineralization. *Adv. Mater.* 15, 959–970. doi: 10.1002/adma.200300381

Addison, W. N., Nakano, Y., Loisel, T., Crine, P., and McKee, M. D. (2008). MEPE-ASARM peptides control extracellular matrix mineralization by binding to hydroxyapatite: an inhibition regulated by PHEX cleavage of ASARM. *J. Bone Miner. Res.* 23, 1638–1949. doi: 10.1359/jbmr.080601

Ahmed, T. A. E., Suso, H. P., and Hincke, M. T. (2017). In-depth comparative analysis of the chicken eggshell membrane proteome. *J. Proteomics* 155, 49–62. doi: 10.1016/j.jprot.2017.01.002

Alva, V., and Lupas, A. N. (2016). The TULIP superfamily of eukaryotic lipid-binding proteins as a mediator of lipid sensing and transport. *Biochim. Biophys. Acta* 1861, 913–923. doi: 10.1016/j.bbalip.2016.01.016

Andrault, J. B., Gaillard, I., Giorgi, D., and Rouquier, S. (2003). Expansion of the BPI family by duplication on human chromosome 20: characterization of the RY gene cluster in 20q11.21 encoding olfactory transporters/antimicrobial-like peptides. *Genomics* 82, 172–184. doi: 10.1016/S0888-7543(03)00102-2

Ar, A., Rahn, H., and Paganelli, C. V. (1979). The avian egg: mass and strength. *Condor* 81, 331–337. doi: 10.2307/1366955

Arazi, H., Yoselewitz, I., Malka, Y., Kelner, Y., Genin, O., and Pines, M. (2009). Osteopontin and calbindin gene expression in the eggshell gland as related to eggshell abnormalities. *Poult. Sci.* 88, 647–653. doi: 10.3382/ps.2008-00387

Arias, J. L., and Fernandez, M. S. (2001). Role of extracellular matrix molecules in shell formation and structure. *Worlds. Poult. Sci. J.* 57, 353–357. doi: 10.1079/wps20010024

Arias, J. L., Fernandez, M. S., Dennis, J. E., and Caplan, A. I. (1991). The fabrication and collagenous substructure of the eggshell membrane in the isthmus of the hen oviduct. *Matrix* 11, 313–320. doi: 10.1016/S0934-8832(11)80202-7

Athanasiadou, D., Jiang, W., Goldbaum, D., Saleem, A., Basu, K., Pacella, M. S., et al. (2018). Nanostructure, osteopontin, and mechanical properties of calcitic avian eggshell. *Sci. Adv.* 4:eaar3219. doi: 10.1126/sciadv.aar3219

Bain, M. M., McDade, K., Burchmore, R., Law, A., Wilson, P. W., Schmutz, M., et al. (2013). Enhancing the egg's natural defence against bacterial penetration by increasing cuticle deposition. *Anim. Genet.* 44, 661–668. doi: 10.1111/age.12071

Bardet, C., Delgado, S., and Sire, J. Y. (2010a). MEPE evolution in mammals reveals regions and residues of prime functional importance. *Cell. Mol. Life Sci.* 67, 305–320. doi: 10.1007/s00018-009-0185-1

Bardet, C., Vincent, C., Lajarille, M. C., Jaffredo, T., and Sire, J. Y. (2010b). OC-116, the chicken ortholog of mammalian MEPE found in eggshell, is also expressed in bone cells. *J. Exp. Zool. B Mol. Dev. Evol.* 314, 653–662. doi: 10.1002/jez.b.21366

Baron, O. L., van West, P., Industri, B., Ponchet, M., Dubreuil, G., Gourbal, B., et al. (2013). Parental transfer of the antimicrobial protein LBP/BPI protects Biomphalaria glabrata eggs against oomycete infections. *PLoS Pathog.* 9:e1003792. doi: 10.1371/journal.ppat.1003792

Bernard, J. P., Adrich, Z., Montalto, G., De Caro, A., De Reggi, M., Sarles, H., et al. (1992). Inhibition of nucleation and crystal growth of calcium carbonate by human lithostathine. *Gastroenterology* 103, 1277–1284. doi: 10.1016/0016-5085(92)91516-7

Blair, J. E., and Hedges, S. B. (2005). Molecular phylogeny and divergence times of deuterostome animals. *Mol. Biol. Evol.* 22, 2275–2284. doi: 10.1093/molbev/msi225

Blank, S., Arnoldi, M., Khoshnavaz, S., Treccani, L., Kuntz, M., Mann, K., et al. (2003). The nacre protein perlucin nucleates growth of calcium carbonate crystals. *J. Microsc.* 212, 280–291. doi: 10.1111/j.1365-2818.2003.01263.x

Bonaparte, J. F., and Vince, M. (1979). El hallazgo del primer nido de dinosaurios triásicos, (*Saurischia*, *Prosauropoda*), Triásico Superior de Patagonia, Argentina. *Ameghiniana* 16, 173–182.

Bridgewater, R. E., Norman, J. C., and Caswell, P. T. (2012). Integrin trafficking at a glance. *J. Cell Sci.* 125, 3695–3701. doi: 10.1242/jcs.095810

Carpenter, K. (2000). *Eggs, Nests, and Baby Dinosaurs: A Look at Dinosaur Reproduction*. Bloomington, IN: Indiana University Press, doi: 10.5860/choice.38-4

Chiang, S. C., Veldhuizen, E. J. A., Barnes, F. A., Craven, C. J., Haagsman, H. P., and Bingle, C. D. (2011). Identification and characterisation of the BPI/LBP/PLUNC-like gene repertoire in chickens reveals the absence of a LBP gene. *Dev. Comp. Immunol.* 35, 285–295. doi: 10.1016/j.dci.2010.09.013

Chien, Y. C., Hincke, M. T., and McKee, M. D. (2009). Avian eggshell structure and osteopontin. *Cells Tissues Organs.* 189, 38–43. doi: 10.1159/000151374

Chien, Y. C., Hincke, M. T., Vali, H., and McKee, M. D. (2008). Ultrastructural matrix-mineral relationships in avian eggshell, and effects of osteopontin on calcite growth in vitro. *J. Struct. Biol.* 163, 84–99. doi: 10.1016/j.jsb.2008.04.008

Choi, S., Han, S., Kim, N. H., and Lee, Y. N. (2018). A comparative study of eggshells of gekkota with morphological, chemical compositional and crystallographic approaches and its evolutionary implications. *PLoS One* 13:e0199496. doi: 10.1371/journal.pone.0199496

Choi, S., Han, S., and Lee, Y. N. (2019). Electron backscatter diffraction (EBSD) analysis of maniraptoran eggshells with important implications for microstructural and taphonomic interpretations. *Palaeontology* 62, 777–803. doi: 10.1111/pala.12427

Conci, N., Lehmann, M., Vargas, S., and Wörheide, G. (2020). Comparative proteomics of octocoral ans cleractinian skeletomes and the evolution of coral calcification. *Genome Biol. Evol.* 12, 1623–1635.

Conci, N., Wörheide, G., and Vargas, S. (2019). New non-bilaterian transcriptomes provide novel insights into the evolution of coral skeletomes. *Genome Biol. Evol.* 11, 3068–3081. doi: 10.1093/gbe/evz199

Cordeiro, C. M. M., Esmaili, H., Ansah, G., and Hincke, M. T. (2013). Ovocalyxin-36 is a pattern recognition protein in chicken eggshell membranes. *PLoS One* 8:e84112. doi: 10.1371/journal.pone.0084112

D'Ambrosio, C., Arena, S., Scaloni, A., Guerrier, L., Boschetti, E., Mendieta, M. E., et al. (2008). Exploring the chicken egg white proteome with combinatorial peptide ligand libraries. *J. Proteome Res.* 7, 3461–3474. doi: 10.1021/pr800193y

Dawson, R. R., Field, D. J., Hull, P. M., Zelenitsky, D. K., Therrien, F., and Affek, H. P. (2020). Eggshell geochemistry reveals ancestral metabolic thermoregulation in Dinosauria. *Sci. Adv.* 6, 9361–9375. doi: 10.1126/sciadv.aax9361

Dominguez-Vera, J. M., Gautron, J., Garcia-Ruiz, J. M., and Nys, Y. (2000). The effect of avian uterine fluid on the growth behavior of calcite crystals. *Poult. Sci.* 79, 901–907. doi: 10.1093/ps/79.6.901

Dunn, I. C., Joseph, N. T., Bain, M., Edmond, A., Wilson, P. W., Milona, P., et al. (2008). Polymorphisms in eggshell organic matrix genes are associated with eggshell quality measurements in pedigree Rhode Island Red hens. *Anim. Genet.* 40, 110–114. doi: 10.1111/j.1365-2052.2008.01794.x

Dunn, I. C., Wilson, P. W., Lu, Z., Bain, M. M., Crossan, C. L., Talbot, R. T., et al. (2009). New hypotheses on the function of the avian shell gland derived from microarray analysis comparing tissue from juvenile and sexually mature hens. *Gen. Comp. Endocrinol.* 163, 225–232. doi: 10.1016/j.ygcen.2009.03.006

Falcon-Lang, H. J., Benton, M. J., and Stimson, M. (2007). Ecology of earliest reptiles inferred from basal Pennsylvanian trackways. *J. Geol. Soc.* 164, 1113–1118. doi: 10.1144/0016-76492007-015

Farinazzo, A., Restuccia, U., Bachi, A., Guerrier, L., Fortis, F., Boschetti, E., et al. (2009). Chicken egg yolk cytoplasmic proteome, mined via combinatorial peptide ligand libraries. *J. Chromatogr. A.* 1216, 1241–1252. doi: 10.1016/j.chroma.2008.11.051

Feng, S., Stiller, J., Deng, Y., Armstrong, J., Fang, Q., Reeve, A. H., et al. (2020). Dense sampling of bird diversity increases power of comparative genomics. *Nature* 587, 252–257. doi: 10.1038/s41586-020-2873-9

Fernandez, M. S., Escobar, C., Lavelin, I., Pines, M., and Arias, J. L. (2003). Localization of osteopontin in oviduct tissue and eggshell during different stages of the avian egg laying cycle. *J. Struct. Biol.* 143, 171–180. doi: 10.1016/j.jsb.2003.08.007

Ford, D. P., and Benson, R. B. J. (2020). The phylogeny of early amniotes and the affinities of *Parareptilia* and *Varanopidae*. *Nat. Ecol. Evol.* 4, 57–65. doi: 10.1038/s41559-019-1047-3

Freeman, C. L., Harding, J. H., Quigley, D., and Rodger, P. M. (2010). Structural control of crystal nuclei by an eggshell protein. *Angew. Chem. Int. Ed. Engl.* 49, 5135–5137. doi: 10.1002/anie.201000679

Freeman, C. L., Harding, J. H., Quigley, D., and Rodger, P. M. (2011). Simulations of ovocleidin-17 binding to calcite surfaces and its implications for eggshell formation. *J. Phys. Chem. C.* 115, 8175–8183. doi: 10.1021/jp200145m

Fulton, J. E., Soller, M., Lund, A. R., Arango, J., and Lipkin, E. (2012). Variation in the ovocalyxin-32 gene in commercial egg-laying chickens and its relationship with egg production and egg quality traits. *Anim. Genet.* 43, 102–113. doi: 10.1111/j.1365-2052.2012.02384.x

Garcia, G., Marivaux, L., Pélissié, T., and Vianey-Liaud, M. (2006). Earliest laurasian sauropod eggshells. *Acta Palaeontol. Pol.* 51, 99–104.

Gatti, J.-L., Métayer, S., Belghazi, M., Dacheux, F., and Dacheux, J.-L. (2005). Identification, proteomic profiling, and origin of ram epididymal fluid exosome-like vesicles. *Biol. Reprod.* 72, 1452–1465. doi: 10.1095/biolreprod.104.036426

Gautron, J. (2019). Proteomics analysis of avian eggshell matrix proteins: toward new advances on biomineralization. *Proteomics* 19, 2. doi: 10.1002/pmic.201900120

Gautron, J., Guyot, N., Brionne, A., and Réhault-Godbert, S. (2019). "Bioactive minor egg components," in *Eggs as Functional Foods and Nutraceuticals for Human Health*, ed. J. Wu (London: Royal Society of Chemistry), 259–284.

Gautron, J., Hincke, M. T., Mann, K., Panheleux, M., Bain, M., McKee, M. D., et al. (2001). Ovocalyxin-32, a novel chicken eggshell matrix protein. isolation, amino acid sequencing, cloning, and immunocytochemical localization. *J. Biol. Chem.* 276, 39243–39252. doi: 10.1074/jbc.M104543200

Gautron, J., Hincke, M. T., and Nys, Y. (1997). Precursor matrix proteins in the uterine fluid change with stages of eggshell formation in hens. *Connect. Tissue Res.* 36, 195–210. doi: 10.3109/03008209709160220

Gautron, J., Murayama, E., Vignal, A., Morisson, M., McKee, M. D., Réhault, S., et al. (2007). Cloning of ovocalyxin-36, a novel chicken eggshell protein related to lipopolysaccharide-binding proteins, bactericidal permeability-increasing proteins, and plunc family proteins. *J. Biol. Chem.* 282, 5273–5286. doi: 10.1074/jbc.M610294200

Gautron, J., Réhault-Godbert, S., Pascal, G., Nys, Y., and Hincke, M. T. (2011). Ovocalyxin-36 and other LBP/BPI/PLUNC-like proteins as molecular actors of the mechanisms of the avian egg natural defences. *Biochem. Soc. Trans.* 39, 971–976. doi: 10.1042/BST0390971

Gautron, J., Stapane, L., Le Roy, N., Nys, Y., Rodriguez-Navarro, A. J., and Hincke, M. T. (2021). Avian eggshell biomineralization: an update on its structure, mineralogy and protein tool kit. *BMC Mol. Cell Biol.* 22:11. doi: 10.1186/s12860-021-00350-0

Gowen, L. C., Petersen, D. N., Mansolf, A. L., Qi, H., Stock, J. L., Tkalcevic, G. T., et al. (2003). Targeted disruption of the osteoblast/osteocyte factor 45 gene (Of45) results in increased bone formation and bone mass. *J. Biol. Chem.* 278, 1998–2007. doi: 10.1074/jbc.M203250200

Guérin-Dubiard, C., Pasco, M., Mollé, D., Désert, C., Croguennec, T., and Nau, F. (2006). Proteomic analysis of hen egg white. *J. Agric. Food Chem.* 54, 3901–3910. doi: 10.1021/jf0529969

Hernandez-Hernandez, A., Gomez-Morales, J., Rodriguez-Navarro, A. B., Gautron, J., Nys, Y., and Garcia Ruiz, J. M. (2008a). Identification of some proteins in the process of hen eggshell formation. *Cryst. Growth Des.* 8, 4330–4339. doi: 10.1021/cg800786s

Hernandez-Hernandez, A., Vidal, M. L., Gomez-Morales, J., Rodriguez-Navarro, A. B., Labas, V., Gautron, J., et al. (2008b). Influence of eggshell matrix proteins on the precipitation of calcium carbonate (CaCO3). *J. Cryst. Growth* 310, 1754–1759. doi: 10.1016/j.jcrysgro.2007.11.170

Hidai, C., Zupancic, T., Penta, K., Mikhail, A., Kawana, M., Quertermous, E. E., et al. (1998). Cloning and characterization of developmental endothelial locus-1: an embryonic endothelial cell protein that binds the αvβ3 integrin receptor. *Genes Dev.* 12, 21–33. doi: 10.1101/gad.12.1.21

Hincke, M. T., Chien, Y.-C., Gerstenfeld, L. C., and McKee, M. D. (2008). Colloidal-gold immunocytochemical localization of osteopontin in avian eggshell gland and eggshell. *J. Histochem. Cytochem.* 56, 467–476. doi: 10.1369/jhc.2008.950576

Hincke, M. T., Da Silva, M., Guyot, N., Gautron, J., McKee, M. D., Guabiraba-Brito, R., et al. (2019). Dynamics of structural barriers and innate immune components during incubation of the avian egg: critical interplay between autonomous embryonic development and maternal anticipation. *J. Innate Immun.* 11, 111–124. doi: 10.1159/000493719

Hincke, M. T., Gautron, J., Tsang, C. P. W., McKee, M. D., and Nys, Y. (1999). Molecular cloning and ultrastructural localization of the core protein of an eggshell matrix proteoglycan, ovocleidin-116. *J. Biol. Chem.* 274, 32915–32923. doi: 10.1074/jbc.274.46.32915

Hincke, M. T., Nys, Y., Gautron, J., Mann, K., Rodriguez-Navarro, B., and McKee, M. D. (2012). The eggshell: structure, composition and mineralization. *Front. Biosci.* 17:1266–1280. doi: 10.2741/3985

Hincke, M. T., Tsang, C. P. W., Courtney, M., Hill, V., and Narbaitz, R. (1995). Purification and immunochemistry of a soluble matrix protein of the chicken eggshell (ovocleidin 17). *Calcif. Tissue Int.* 56, 578–583. doi: 10.1007/BF00298593

Horvat-Gordon, M., Yu, F., Burns, D., and Leach, R. M. (2008). Ovocleidin (OC 116) is present in avian skeletal tissues. *Poult. Sci.* 87, 1618–1623. doi: 10.3382/ps.2008-1

Hynes, R. O. (2002). Integrins: bidirectional, allosteric signaling machines. *Cell* 110, 673–687. doi: 10.1016/s0092-8674(02)00971-6

Jarvis, E. D., Mirarab, S., Aberer, A. J., Li, B., Houde, P., Li, C., et al. (2014). Whole-genome analyses resolve early branches in the tree of life of modern birds. *Science* 346, 1320–1331. doi: 10.1126/science.1253451

Jonchère, V., Brionne, A., Gautron, J., and Nys, Y. (2012). Identification of uterine ion transporters for mineralisation precursors of the avian eggshell. *BMC Physiol.* 12:10. doi: 10.1186/1472-6793-12-10

Kapusta, A., Suh, A., and Fescotte, C. (2017). Dynamics of genome size evolution in birds and animals. *Proc. Natl. Acad. Sci. U. S. A.* 114, E1460–E1469. doi: 10.1073/pnas.1616702114

Karakostis, K., Costa, C., Zito, F., Brümmer, F., and Matranga, V. (2016). Characterization of an alpha type carbonic anhydrase from Paracentrotus lividus sea urchin embryos. *Mar. Biotechnol.* 18, 384–395. doi: 10.1007/s10126-016-9701-0

Kawasaki, K. (2011). The SCPP gene family and the complexity of hard tissues in vertebrates. *Cells Tissues Organs.* 194, 108–112. doi: 10.1159/000324225

Kawasaki, K., Buchanan, A. V., and Weiss, K. M. (2007). Gene duplication and the evolution of vertebrate skeletal mineralization. *Cells Tissues Organs.* 186, 7–24. doi: 10.1159/000102678

Kawasaki, K., and Weiss, K. M. (2008). SCPP gene evolution and the dental mineralization continuum. *J. Dent. Res.* 87, 520–531. doi: 10.1177/154405910808700608

Krasity, B. C., Troll, J. V., Weiss, J. P., and McFall-Ngai, M. J. (2011). LBP/BPI proteins and their relatives: conservation over evolution and roles in mutualism. *Biochem. Soc. Trans.* 39, 1039–1044. doi: 10.1042/BST0391039

Lavelin, I., Yarden, N., Ben-Bassat, S., Bar, A., and Pines, M. (1998). Regulation of osteopontin gene expression during egg shell formation in the laying hen by mechanical strain. *Matrix Biol.* 17, 615–623. doi: 10.1016/s0945-053x(98)90112-3

Le Roy, N., Combes-Soia, L., Brionne, A., Labas, V., Rodriguez-Navarro, A. B., Hincke, M. T., et al. (2019). Guinea fowl eggshell quantitative proteomics yield new findings related to its unique structural characteristics and superior mechanical properties. *J. Proteomics* 209:103511. doi: 10.1016/j.jprot.2019.103511

Le Roy, N., Ganot, P., Aranda, M., Allemand, D., and Tambutté, S. (2021). The skeletome of the red coral *Corallium* rubrum indicates an independent evolution of biomineralization process in octocorals. *BMC Evol. Biol.* 21:1. doi: 10.1186/s12862-020-01734-0

Le Roy, N., Jackson, D. J., Marie, B., Ramos-Silva, P., and Marin, F. (2014). The evolution of metazoan α-carbonic anhydrases and their roles in calcium carbonate biomineralization. *Front. Zool.* 11:75. doi: 10.1186/s12983-014-0075-78

Leach, R. M. (1982). Biochemistry of the organic matrix of the eggshell. *Poult. Sci.* 61, 2040–2047. doi: 10.3382/ps.0612040

Levi-Kalisman, Y., Raz, S., Weiner, S., Addadi, L., and Sagi, I. (2002). Structural differences between biogenic amorphous calcium carbonate phases using X-ray absorption spectroscopy. *Adv. Funct. Mater.* 12, 43–48.

Lin, J. Y., Ma, K. Y., Bai, Z. Y., and Le Li, J. (2013). Molecular cloning and characterization of perlucin from the freshwater pearl mussel, *Hyriopsis cumingii*. *Gene* 526, 210–216. doi: 10.1016/j.gene.2013.05.029

Mann, K. (1999). Isolation of a glycosylated form of the chicken eggshell protein ovocleidin and determination of the glycosylation site. Alternative glycosylation/phosphorylation at an N-glycosylation sequon. *FEBS Lett.* 463, 12–14. doi: 10.1016/S0014-5793(99)01586-0

Mann, K. (2007). The chicken egg white proteome. *Proteomics* 7, 3558–3568. doi: 10.1002/pmic.200700397

Mann, K. (2015). The calcified eggshell matrix proteome of a songbird, the zebra finch (*Taeniopygia guttata*). *Proteome Sci.* 13, 29. doi: 10.1186/s12953-015-0086-1

Mann, K., Hincke, M. T., and Nys, Y. (2002). Isolation of ovocleidin-116 from chicken eggshells, correction of its amino acid sequence and identification of disulfide bonds and glycosylated Asn. *Matrix Biol.* 21, 383–387. doi: 10.1016/S0945-053X(02)00031-8

Mann, K., Maèek, B., and Olsen, J. V. (2006). Proteomic analysis of the acid-soluble organic matrix of the chicken calcified eggshell layer. *Proteomics* 6, 3801–3810. doi: 10.1002/pmic.200600120

Mann, K., and Mann, M. (2008). The chicken egg yolk plasma and granule proteomes. *Proteomics* 8, 178–191. doi: 10.1002/pmic.200700790

Mann, K., and Mann, M. (2013). The proteome of the calcified layer organic matrix of turkey (*Meleagris gallopavo*) eggshell. *Proteome Sci.* 11:40. doi: 10.1186/1477-5956-11-40

Mann, K., and Mann, M. (2015). Proteomic analysis of quail calcified eggshell matrix: a comparison to chicken and turkey eggshell proteomes. *Proteome Sci.* 13:22. doi: 10.1186/s12953-015-0078-1

Mann, K., Olsen, J. V., Macek, B., Gnad, F., and Mann, M. (2007a). Phosphoproteins of the chicken eggshell calcified layer. *Proteomics* 7, 106–115. doi: 10.1002/pmic.200600635

Mann, K., Siedler, F., Treccani, L., Heinemann, F., and Fritz, M. (2007b). Perlinhibin, a cysteine-, histidine-, and arginine-rich miniprotein from abalone (*Haliotis laevigata*) nacre, inhibits in vitro calcium carbonate crystallization. *Biophys. J.* 93, 1246–1254. doi: 10.1529/biophysj.106.100636

Mann, K., and Siedler, F. (1999). The amino acid sequence of ovocleidin-17, a major protein of the avian eggshell calcified layer. *Biochem. Mol. Biol. Int.* 47, 997–1007. doi: 10.1080/15216549900202123

Mann, K., and Siedler, F. (2004). Ostrich (*Struthio camelus*) eggshell matrix contains two different C-type lectin-like proteins. Isolation, amino acid sequence, and posttranslational modifications. *Biochim. Biophys. Acta.* 1696, 41–50. doi: 10.1016/j.bbapap.2003.09.006

Mann, K., and Siedler, F. (2006). Amino acid sequences and phosphorylation sites of emu and rhea eggshell C-type lectin-like proteins. *Comp. Biochem. Physiol. - B Biochem. Mol. Biol.* 143, 160–170. doi: 10.1016/j.cbpb.2005.11.003

Marie, B., Le Roy, N., Zanella-Cléon, I., Becchi, M., and Marin, F. (2011). Molecular evolution of mollusc shell proteins: insights from proteomic analysis of the eddible mussel Mytilus. *J. Mol. Evol.* 72, 531–546. doi: 10.1007/s00239-011-9451-6

Marie, B., Marie, A., Jackson, D. J., Dubost, L., Degnan, B. M., Milet, C., et al. (2010). Proteomic analysis of the organic matrix of the abalone *Haliotis* asinina calcified shell. *Proteome Sci.* 8:54. doi: 10.1186/1477-5956-8-54

Marie, P., Labas, V., Brionne, A., Harichaux, G., Hennequet-Antier, C., Rodríguez-Navarro, A. B., et al. (2015a). Quantitative proteomics provides new insights into chicken eggshell matrix protein functions during the primary events of mineralisation and the active calcification phase. *J. Proteomics* 126, 140–154. doi: 10.1016/j.jprot.2015.05.034

Marie, P., Labas, V., Brionne, A., Harichaux, G., Hennequet-Antier, C., Nys, Y., et al. (2015b). Quantitative proteomics and bioinformatic analysis provide new insight into protein function during avian eggshell biomineralization. *J. Proteomics* 113, 178–193. doi: 10.1016/j.jprot.2014.09.024

Mass, T., Giuffre, A. J., Sun, C. Y., Stifler, C. A., Frazier, M. J., Neder, M., et al. (2017). Amorphous calcium carbonate particles form coral skeletons. *Proc. Natl. Acad. Sci. U. S. A.* 114, E7670–E7678. doi: 10.1073/pnas.1707890114

Matsubara, H., Hayashi, T., Ogawa, T., Muramoto, K., Jimbo, M., and Kamiya, H. (2008). Modulating effect of acorn barnacle C-type lectins on the crystallization of calcium carbonate. *Fish. Sci.* 74, 418–424. doi: 10.1111/j.1444-2906.2008.01539.x

McKee, M. D., Pedraza, C. E., and Kaartinen, M. T. (2011). Osteopontin and wound healing in bone. *Cells Tissues Organs.* 194, 313–319. doi: 10.1159/000324244

Mcknight, D. A., and Fisher, L. W. (2009). Molecular evolution of dentin phosphoprotein among toothed and toothless animals. *BMC Evol. Biol.* 9:299. doi: 10.1186/1471-2148-9-299

Mikhailov, K. (1991). Classification of fossil eggshells of amniotic vertebrates. *Acta Palaeontol. Pol.* 36, 193–238.

Mikšík, I., Eckhardt, A., Sedláková, P., and Mikulikova, K. (2007). Proteins of insoluble matrix of avian (*Gallus Gallus*) eggshell. *Connect. Tissue Res.* 48, 1–8. doi: 10.1080/03008200601003116

Mikšík, I., Paradis, S., Eckhardt, A., and Sedmera, D. (2018). Analysis of siamese crocodile (*Crocodylus siamensis*) eggshell proteome. *Protein J.* 37, 21–37. doi: 10.1007/s10930-017-9750-x

Moore, M. A., Gotoh, Y., Rafidi, K., and Gerstenfeld, L. C. (1991). Characterization of a cDNA for chicken osteopontin: expression during bone development, osteoblast differentiation, and tissue distribution. *Biochemistry* 30, 2501–2508. doi: 10.1021/bi00223a029

Moya, A., Tambutté, S., Bertucci, A., Tambutté, E., Lotto, S., Vullo, D., et al. (2008). Carbonic anhydrase in the scleractinian coral Stylophora pistillata: characterization, localization, and role in biomineralization. *J. Biol. Chem.* 283, 25475–25484. doi: 10.1074/jbc.M804726200

Nimtz, M., Conradt, H. S., and Mann, K. (2004). LacdiNAc (GalNAcβ1-4GlcNAc) is a major motif in N-glycan structures of the chicken eggshell protein ovocleidin-116. *Biochim. Biophys. Acta.* 1675, 71–80. doi: 10.1016/j.bbagen.2004.08.007

Norell, M. A., Wiemann, J., Fabbri, M., Yu, C., Marsicano, C. A., Moore-Nall, A., et al. (2020). The first dinosaur egg was soft. *Nature* 583:7816. doi: 10.1038/s41586-020-2412-8

Nys, Y., Gautron, J., Garcia-Ruiz, J. M., and Hincke, M. T. (2004). Avian eggshell mineralization: biochemical and functional characterization of matrix proteins. *CR Palevol.* 3, 549–562. doi: 10.1016/j.crpv.2004.08.002

Nys, Y., Hincke, M. T., Arias, J. L., Garcia-Ruiz, J. M., and Solomon, S. E. (1999). Avian eggshell mineralization. *Poult. Avian Biol. Rev.* 10, 143–166.

Panheleux, M., Kälin, O., Gautron, J., and Nys, Y. (1999a). Features of eggshell formation in guinea fowl: kinetics of shell deposition, uterine protein secretion and uterine histology. *Br. Poult. Sci.* 40, 632–643. doi: 10.1080/00071669987025

Panheleux, M., Bain, M., Fernandez, M. S., Morales, I., Gautron, J., Arias, J. L., et al. (1999b). Organic matrix composition and ultrastructure of eggshell: a comparative study. *Br. Poult. Sci.* 40, 240–252. doi: 10.1080/00071669987665

Panheleux, M., Nys, Y., Williams, J., Gautron, J., Boldicke, T., and Hincke, M. T. (2000). Extraction and quantification by ELISA of eggshell organic matrix proteins (ovocleidin-17, ovalbumin, ovotransferrin) in shell from young and old hens. *Poult. Sci.* 79, 580–588. doi: 10.1093/ps/79.4.580

Pérez-Huerta, A., and Dauphin, Y. (2016). Comparison of the structure, crystallography and composition of eggshells of the guinea fowl and graylag goose. *Zoology* 119, 52–63. doi: 10.1016/j.zool.2015.11.002

Petersen, J., and Tyler, C. (1966). The strength of guinea fowl (*Numida meleagris*) egg shells. *Br. Poult. Sci.* 7, 291–296. doi: 10.1080/00071666608415635

Phillips, M. J., Bennett, T. H., and Lee, M. S. (2009). Molecules, morphology, and ecology indicate a recent, amphibious ancestry for echidnas. *Proc. Natl. Acad. Sci. U. S. A.* 106, 17089–17094. doi: 10.1073/pnas.0904649106

Pines, M., Knopov, V., and Bar, A. (1995). Involvement of osteopontin in egg shell formation in the laying chicken. *Matrix Biol.* 14, 765–771. doi: 10.1016/S0945-053X(05)80019-8

Ramos-Silva, P., Kaandorp, J., Huisman, L., Marie, B., Zanella-Cléon, I., Guichard, N., et al. (2013). The skeletal proteome of the coral *Acropora* millepora: the evolution of calcification by co-option and domain shuffling. *Mol. Biol. Evol.* 30, 2099–2112. doi: 10.1093/molbev/mst109

Rao, A., Seto, J., Berg, J. K., Kreft, S. G., Scheffner, M., and Cölfen, H. (2013). Roles of larval sea urchin spicule SM50 domains in organic matrix self-assembly and calcium carbonate mineralization. *J. Struct. Biol.* 183, 205–215. doi: 10.1016/j.jsb.2013.06.004

Reyes-Grajeda, J. P., Moreno, A., and Romero, A. (2004). Crystal structure of ovocleidin-17, a major protein of the calcified Gallus gallus eggshell: implications in the calcite mineral growth pattern. *J. Biol. Chem.* 279, 40876–40881. doi: 10.1074/jbc.M406033200

Rodríguez-Navarro, A. B., Marie, P., Nys, Y., Hincke, M. T., and Gautron, J. (2015). Amorphous calcium carbonate controls avian eggshell mineralization: a new paradigm for understanding rapid eggshell calcification. *J. Struct. Biol.* 190, 291–303. doi: 10.1016/j.jsb.2015.04.014

Romé, H., and Le Roy, P. (2016). Régions chromosomiques influençant les caractères de production et de qualité des oeufs de poule TT - chromosomal regions influencing egg production and egg quality traits in hens. *INRA Prod. Anim.* 29, 117–128. doi: 10.20870/productions-animales.2016.29.1.2521

Rose-Martel, M., Du, J., and Hincke, M. T. (2012). Proteomic analysis provides new insight into the chicken eggshell cuticle. *J. Proteomics* 75, 2697–2706. doi: 10.1016/j.jprot.2012.03.019

Rose-Martel, M., Smiley, S., and Hincke, M. T. (2015). Novel identification of matrix proteins involved in calcitic biomineralization. *J. Proteomics* 116, 81–96. doi: 10.1016/j.jprot.2015.01.002

Rowe, P. S. N. (2012). The chicken or the egg: PHEX, FGF23 and SIBLINGs unscrambled. *Cell Biochem. Funct.* 30, 355–375. doi: 10.1002/cbf.2841

Rowe, P. S. N., Kumagai, Y., Gutierrez, G., Garrett, I. R., Blacher, R., Rosen, D., et al. (2004). MEPE has the properties of an osteoblastic phosphatonin and minhibin. *Bone* 34, 303–319. doi: 10.1016/j.bone.2003.10.005

Shen, X.-X., Liang, D., Wen, J.-Z., and Zhang, P. (2011). Multiple genome alignments facilitate development of NPCL markers: a case study of tetrapod phylogeny focusing on the position of turtles. *Mol. Biol. Evol.* 28, 3237–3252. doi: 10.1093/molbev/msr148

Sire, J. Y., Delgado, S. C., and Girondot, M. (2008). Hen's teeth with enamel cap: from dream to impossibility. *BMC Evol. Biol.* 8:246. doi: 10.1186/1471-2148-8-246

Sire, J.-Y., and Kawasaki, K. (2012). "Origin and evolution of bone and dentin, and of their phosphorylated, acid-rich matrix proteins," in *Phosphorylated Extracellular Matrix Proteins of Bone and Dentin*, ed. M. Goldberg (Sharjah: Bentham Science), 3–60.

Sodek, J., Ganss, B., and McKee, M. D. (2000). Osteopontin. *Crit. Rev. Oral Biol. Med.* 11, 279–303. doi: 10.1177/10454411000110030101

Solstad, T., Stenvik, J., and Jørgensen, T. (2007). mRNA expression patterns of the BPI/LBP molecule in the Atlantic cod (Gadus morhua L.). *Fish Shellfish Immunol.* 23, 260–271. doi: 10.1016/j.fsi.2006.10.002

Song, K. T., Choi, S. H., and Oh, H. R. (2000). A comparison of egg quality of pheasant, chukar, quail and guinea fowl. *Asian-Australas. J. Anim. Sci.* 13, 986–990. doi: 10.5713/ajas.2000.986

Stapane, L., Le Roy, N., Ezagal, J., Rodriguez-Navarro, A. B., Labas, V., Combes-Soia, L., et al. (2020). Avian eggshell formation reveals a new paradigm for vertebrate mineralization via vesicular amorphous calcium carbonate. *J. Biol. Chem.* 295, 15853–15869. doi: 10.1074/jbc.ra120.014542

Stapane, L., Le Roy, N., Hincke, M. T., and Gautron, J. (2019). The glycoproteins EDIL3 and MFGE8 regulate vesiclemediated eggshell calcification in a new model for avian biomineralization. *J. Biol. Chem.* 294, 14526–14545. doi: 10.1074/jbc.RA119.009799

Stein, K., Prondvai, E., Huang, T., Baele, J. M., Sander, P. M., and Reisz, R. (2019). Structure and evolutionary implications of the earliest (Sinemurian, Early Jurassic) dinosaur eggs and eggshells. *Sci. Rep.* 9:4424. doi: 10.1038/s41598-019-40604-8

Takahashi, H., Sasaki, O., Nirasawa, K., and Furukawa, T. (2010). Association between ovocalyxin-32 gene haplotypes and eggshell quality traits in an F2 intercross between two chicken lines divergently selected for eggshell strength. *Anim. Genet.* 41, 541–544. doi: 10.1111/j.1365-2052.2010.02034.x

Théry, C., Regnault, A., Garin, J., Wolfers, J., Zitvogel, L., Ricciardi-Castagnoli, P., et al. (1999). Molecular characterization of dendritic cell-derived exosomes: selective accumulation of the heat shock protein hsc73. *J. Cell Biol.* 147, 599–610. doi: 10.1083/jcb.147.3.599

Tian, X., Gautron, J., Monget, P., and Pascal, G. (2010). What makes an egg unique? Clues from evolutionary scenarios of egg-specific genes. *Biol. Reprod.* 83, 893–900. doi: 10.1095/biolreprod.110.085019

Uemoto, Y., Suzuki, C., Sato, S., Sato, S., Ohtake, T., Sasaki, O., et al. (2009). Polymorphism of the ovocalyxin-32 gene and its association with egg production traits in the chicken. *Poult. Sci.* 88, 2512–2517. doi: 10.3382/ps.2009-00331

Voigt, O., Adamski, M., Sluzek, K., and Adamska, M. (2014). Calcareous sponge genomes reveal complex evolution of α-carbonic anhydrases and two key biomineralization enzymes. *BMC Evol. Biol.* 14:230. doi: 10.1186/s12862-014-0230-z

Voris, J. T., Zelenitsky, D. K., Therrien, F., and Tanaka, K. (2018). Dinosaur eggshells from the lower Maastrichtian St. Mary River Formation of southern Alberta, Canada. *Can. J. Earth Sci.* 55, 272–282. doi: 10.1139/cjes-2017-0195

Wallace, A., and Schiffbauer, J. (2016). Proteins from the past. *Elife* 5:e20877. doi: 10.7554/eLife.17092

Warren, W. C., Hillier, L. W., Graves, J. A. M., Birney, E., Ponting, C. P., Grützner, F., et al. (2008). Genome analysis of the platypus reveals unique signatures of evolution. *Nature* 453, 175–183. doi: 10.1038/nature06936

Weber, E., Weiss, I. M., Cölfen, H., and Kellermeier, M. (2016). Recombinant perlucin derivatives influence the nucleation of calcium carbonate. *CrystEngComm* 18, 8439–8444. doi: 10.1039/c6ce01878e

Weber, G. F. (2018). The phylogeny of osteopontin - analysis of the protein sequence. *Int. J. Mol. Sci.* 19:2557. doi: 10.3390/ijms19092557

Weiner, S., and Dove, P. M. (2003). An overview of biomineralization processes and the problem of the vital effect. *Rev. Mineral. Geochem.* 54, 1–29. doi: 10.2113/0540001

Wellman-Labadie, O., Lakshminarayanan, R., and Hincke, M. T. (2008a). Antimicrobial properties of avian eggshell-specific C-type lectin-like proteins. *FEBS Lett.* 582, 699–704. doi: 10.1016/j.febslet.2008.01.043

Wellman-Labadie, O., Picman, J., and Hincke, M. T. (2008b). Antimicrobial activity of cuticle and outer eggshell protein extracts from three species of domestic birds. *Br. Poult. Sci.* 49, 133–143. doi: 10.1080/00071660802001722

Wellman-Labadie, O., Picman, J., and Hincke, M. T. (2008c). Antimicrobial activity of the Anseriform outer eggshell and cuticle. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* 149, 640–649. doi: 10.1016/j.cbpb.2008.01.001

Wong, M., Hendrix, M. J. C., Von der Mark, K., Little, C., and Stern, R. (1984). Collagen in the egg shell membranes of the hen. *Dev. Biol.* 104, 28–36. doi: 10.1016/0012-1606(84)900332

Xing, J., Wellman-Labadie, O., Gautron, J., and Hincke, M. T. (2007). Recombinant eggshell ovocalyxin-32: expression, purification and biological activity of the glutathione S-transferase fusion protein. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* 147, 172–177. doi: 10.1016/j.cbpb.2007.01.015

Yang, R., Geng, F., Huang, X., Qiu, N., Li, S., Teng, H., et al. (2020). Integrated proteomic, phosphoproteomic and N-glycoproteomic analyses of chicken eggshell matrix. *Food Chem.* 330:127167. doi: 10.1016/j.foodchem.2020.127167

Zelensky, A. N., and Gready, J. E. (2005). The C-type lectin-like domain superfamily. *FEBS J.* 272, 6179–6217. doi: 10.1111/j.1742-4658.2005.05031.x

Zhang, Q., Liu, L., Zhu, F., Ning, Z., Hincke, M. T., Yang, N., et al. (2014). Integrating de novo transcriptome assembly and cloning to obtain chicken ovocleidin-17 full-length cDNA. *PLoS One* 9:e93452. doi: 10.1371/journal.pone.0093452

Zhao, K., Wang, M., Wang, X., Wu, C., Xu, H., and Lu, J. R. (2013). Crystal growth of calcite mediated by ovalbumin and lysozyme: atomic force microscopy study. *Cryst. Growth Des.* 13, 1583–1589. doi: 10.1021/cg301820w

Zhu, F., Zhang, F., Hincke, M., Yin, Z. T., Chen, S. R., Yang, N., et al. (2019). iTRAQ-based quantitative proteomic analysis of duck eggshell during biomineralization. *Proteomics* 19:e1900011. doi: 10.1002/pmic.201900011

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Le Roy, Stapane, Gautron and Hincke. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.