The unique features of proteins depicting the chicken amniotic fluid

Mylène da Silva, Clara Dombre, Aurélien Brionne, Philippe Monget, Magali Chessé, Marion de Pauw, Maryse Mills, Lucie Combes-Soia, Valérie Labas, Nicolas Guyot, et al.

MCP Papers in Press. Published on February 14, 2018 as Manuscript RA117.000459

The unique features of proteins depicting the chicken amniotic fluid

Mylène Da Silva ‡, Clara Dombre §, Aurélien Brionne ‡, Philippe Monget §, Magali Chessé ‡, Marion De Pauw ‡, Maryse Mills ‡, Lucie Combes-Soia §¶, Valérie Labas §¶, Nicolas Guyot ‡, Yves Nys ‡, Sophie Réhault-Godbert ‡‖

‡ BOA, INRA, Université de Tours, 37380 Nouzilly, France.

§ PRC, INRA, CNRS, IFCE, Université de Tours, Nouzilly 37380, France.

¶ INRA, Plate-forme de Chirurgie et Imagerie pour la Recherche et l'Enseignement (CIRE), Pôle d'Analyse et d'Imagerie des Biomolécules (PAIB), F-37380 Nouzilly, France.

‖ Corresponding author: sophie.rehault-godbert@inra.fr; +33 (0)2 47 42 78 39; Fax : +33 (0)2 47 42 77 78

Running title: Proteomic specificities of the chicken amniotic fluid

SUMMARY

In many amniotes, the amniotic fluid is depicted as a dynamic milieu that participates in the protection of the embryo (cushioning, hydration, and immunity). However, in birds, the protein profile of the amniotic fluid remains unexplored, even though its proteomic signature is predicted to differ compared to human's. In fact, unlike humans, chicken amniotic fluid does not collect excretory products and its protein composition strikingly changes at mid-development due to the massive inflow of egg white proteins, which are thereafter swallowed by the embryo to support its growth. Using GeLC-MS/MS and shotgun strategies, we identified 91 non-redundant proteins delineating the chicken amniotic fluid proteome at day 11 of development, before egg white transfer. These proteins were essentially associated with the metabolism of nutrients, immune response and developmental processes. Forty-eight proteins were common to both chicken and human amniotic fluids, including serum albumin, apolipoprotein A1 and alpha-fetoprotein. We further investigated the effective role of the chicken amniotic fluid in innate defense and revealed that it exhibits significant antibacterial activity at day 11 of development. This antibacterial potential is drastically enhanced after egg white transfer, presumably due to lysozyme, avian beta-defensin 11, vitelline membrane outer layer protein 1, and beta-microseminoprotein-like as the most likely antibacterial candidates. Interestingly, several proteins recovered in the chicken amniotic fluid prior and after egg white transfer are uniquely found in birds (ovalbumin and related proteins X and Y, avian beta-defensin 11) or oviparous species (vitellogenins 1 and 2, riboflavin-binding protein). This study provides an integrative overview of the chicken amniotic fluid proteome and opens stimulating perspectives in deciphering the role of the avian egg-specific proteins in embryonic development, including innate immunity. These proteins may constitute

valuable biomarkers for poultry production to detect hazardous situations (stress, infection, etc.), that may negatively affect the development of the chicken embryo.

ABBREVIATIONS

AF, Amniotic fluid; ATCC, American Type Culture Collection; BMSP, Beta-microseminoprotein-like; ED, Day of incubation; emPAI: Exponentially modified Protein Abundance Index; EW, Egg white; HBP, Heparin-binding protein; *L.m.*, *Listeria monocytogenes*; MYA: Million Years Ago; NCBI, National Center for Biotechnology Information; *S*.E., *Salmonella enterica* Enteritidis; TSB, Trypticase soy broth.

INTRODUCTION

In oviparous species, the embryonic development depends on the various components, nutrients and structures composing the eggshell, the egg yolk, the egg white and the vitelline membrane (1). It also relies on the proper development of extra-embryonic structures, namely the yolk sac, the amniotic sac and the allantoic/chorioallantoic sac (1) (Fig. 1A). These structures develop at the very early stages of development and originate from embryonic tissues, but are not considered to be part of the embryo's body (2). They are discarded or resorbed at hatching. These living structures are partly preserved among amniotes species, but exhibit evolutionary particularities depending on the embryonic development mode (3). The yolk sac, which appears in the first stages of development, degenerates rapidly in mammals (Fig. 1B), while in some birds, it participates in digestive processes until the last stages of incubation prior to complete abdominal resorption at hatch. The yolk sac may have many other functions, which are temporally regulated all along the incubation: it resembles the liver in the synthesis of plasma proteins, the bone marrow in erythropoiesis, and the intestine, in digestion of nutrients and their transport to the embryo (4). Thus, the yolk sac plays different roles to support or replace the functions of several organs that have not yet reached their full functional capacity. The chorioallantoic sac is composed of the chorioallantoic membrane, which results from the fusion of the chorion and the allantois at ED5-ED6, and it includes the allantoic fluid. It is a highly vascularized structure that ensures many functions during chicken embryonic development: it collects nitrogenous and excretory products from the embryonic metabolism, it participates in respiratory exchanges, in calcium transport from the eggshell towards the embryo, in ion and water reabsorption from the allantoic fluid and thus in acid-base homeostasis (2, 5). In humans, the allantoic sac forms only a part of the umbilical cord. Concerning the amniotic sac, it was described in all amniotes as a structure, filled out with the

amniotic fluid (AF) protecting the embryo against mechanical shocks, dehydration or adhesion to the other extra-embryonic membranes. It also serves as a source of nutrients (6). It provides a favorable environment to accompany the development of the embryo: pH of about 7.1 to 7.3, stable temperature, and sensorial stimulation (taste, sense of smell and hearing) (7).

Human AF is a fluctuating milieu mainly composed of water (about 96.4%), minerals, trace elements, carbohydrates, hormones, glucose, lipids, urea, cells, free amino-acids, proteins and peptides (6, 8). Its biochemical composition changes with gestational age/developmental stage as a result of various physiological mechanisms including feto-maternal exchanges: AF swallowing and lung fluid production by the embryo, excretion of fetal urine and transfer of solutes and fluids across amniotic and uterine membranes (Fig.1B) and across skin especially before keratinization (6, 9). In humans, before skin keratinization (25 weeks of age), the composition of the AF is very similar to fetal blood plasma. As a dynamic milieu reflecting the physiological or pathological status of the embryo, the biochemical composition of the AF has been extensively characterized in humans since it contains molecular markers for the detection of embryonic abnormal development, inflammation, infection or pregnancy-related complications (10). Its role as a mechanical cushioning of the embryo, together with the fact that 25% of its total protein content is associated with the immune response or is related to defense (9), corroborate its primary role in the protection of the embryo. However, it seems that proteins of the human AF are not solely involved in the nutrition and protection of the embryo, but they display many other functions related to metabolism and development (11). In oviparous species, however, the functions of the AF are still poorly understood and there are several structural and biochemical particularities that suggest similar but also diverging functions. First, there are two major extra-embryonic fluids in birds, the amniotic and allantoic fluids, which are physically separated, to ensure different

functions. Similar to humans, the bird AF is contained in the amniotic sac and bathes the embryo, whereas the allantoic fluid is secreted in the chorioallantoic sac (Fig. 1A), which is an intestinal intussusception of the embryo that receives disposable wastes directly from the embryonic kidneys. These anatomical specificities deeply influence the biochemical composition of the AF, which thus in birds does not collect fetal urine. The second main difference, as compared with mammals, is that after ED12, egg white proteins are massively transferred into the amniotic sac (12), where they are absorbed orally by the embryo as a source of amino-acids to support its rapid growth (13-17) until hatching (ED21). Consequently, this process drastically impacts protein concentration of AF, which is barely measurable before the 11th day of incubation (ED11) (about 0.01 mg/mL) and reaches 200 mg/mL following egg white transfer (18). In fact, before ED11, the chicken AF is mainly composed of water and mineral elements, such as chloride, sodium, potassium, phosphorus, magnesium, calcium, iron, and sulfur, like human AF (19, 20). More recently, a study analyzing the major proteins of the chicken AF, revealed that it contains egg white proteins even before the massive egg white transfer at ED12 (18). Some of these proteins are egg yolk proteins while others may originate from the embryo (skin, feather) or its extra-embryonic membranes (amniotic, yolk, and chorioallantoic sacs). All these proteins have been associated with functions comparable to those described for human AF, including metabolism, immune system or tissue remodeling, but they also serve as a major source of amino-acids and energy for the embryo, especially during the second half of incubation. The avian specificity of some egg white and yolk proteins (21) recovered in the AF, together with the physical separation of amniotic sac from the embryonic urinary system (chorioallantoic sac) and the transfer of egg white proteins at mid-incubation, suggest that this fluid may have very specific biological functions related to birds or even oviparity.

In such a stimulating context, we explored the avian AF proteins and specificities using the chicken as a model of birds. Similar to human AF, with albumin, immunoglobulins, transferrin and haptoglobin (9, 11), chicken AF contains a few major proteins (ovalbumin, ovotransferrin, etc.) that complicate proteome profiling, masking the presence of proteins of lower abundance. Therefore, we designed a bottom-up proteomic approach combining two complementary strategies as GeLC-MS/MS (protein samples fractionated by SDS-PAGE are analyzed by nanoLC-MS/MS after in-gel digestion) and shotgun (protein samples are directly analyzed by nanoLC-MS/MS after in-solution digestion) to give an exhaustive view of the AF proteome at ED11, before the substantial transfer of egg white (which completely changes the global protein profile). We performed a functional annotation of AF proteome using Gene Ontology approaches: 1) on the 10 most abundant proteins representing 66% to 81% of the total protein contents for GeLC-MS/MS and shotgun analyses, respectively, and 2) on the complete list of proteins resulting from both analyses. This approach was compulsory because many of the major proteins composing egg structures do not have assigned functions in databases yet and consequently, the functional annotation of the entire list would enrich functions, which biological significance may be overestimated. Considering the importance of human amniotic fluid in innate immunity and to better appreciate the contribution of chicken amniotic fluid to embryo defense against microorganisms, we used an in-gel antibacterial assay combined to mass spectrometry, to identify antibacterial proteins and peptides contained in AF or in enriched fractions of AF, before and after egg white transfer. As the protein composition of the chicken AF is temporally regulated during incubation, AF functions are thoroughly modified following egg white transfer. Finally, a comparison with the published proteomes of the human AF (11) and a phylogenetic study on all

proteins identified in this study, allowed us to highlight chicken AF specificities, revealing a set of proteins, which are only present in birds and which physiological roles remain fully opened.

EXPERIMENTAL PROCEDURES

A diagram describing the experimental design is presented in Fig 2.

Samplings

Fluids were sampled as previously described by Da Silva *et al.* (18). Briefly, fertile eggs were incubated under standard conditions (45% relative humidity, 37.8°C, automatic turning every hour), after a three-day storage at 16°C and 85% relative humidity to ensure synchronization of developmental stages (UE1295, INRA, F-37380 Nouzilly, France). At ED11 and ED16, 40 eggs of comparable weight (62.4 ± 4.3 g) containing viable embryos (checked by candling) were tested with the Acoustic Egg Tester (KU Leuven, Belgium), and cracked eggs were discarded. By ED11, egg white was sampled with a syringe after drilling a hole in the eggshell. After removing the eggshell, the egg content was poured into a Petri dish and the AF was recovered with a syringe through the amniotic membrane. By ED16, the egg white was collected in the Petri dish with pliers due to its high viscosity. All samplings were performed under sterile conditions. Sex of individual embryos was determined using PCR (22).

Fluid characterization

After samplings, AF were centrifuged at 3,000 g (10 min, 4°C) to remove insoluble components. Volumes, pH (Microelectrode pH InLab 423, Fisher Scientific, Illkirch, France), osmolality (Fiske Mark 3 Osmometer, Advanced Instruments, Niederbronn-Les-Bains, France), and absorbance spectrum (Nanodrop, ND-100 Spectro, Wilmington, USA) were analyzed for each sample. The

total protein concentration for each sample was assessed using BioRad DC Protein Assay Kit II (BioRad, Marnes-la-Coquette, France). All samples (25 µL) containing loading buffer (5X loading buffer: 0.25 M Tris-HCl, 0.05% bromophenol blue, 50% glycerol, 5% SDS, pH 6.8) were independently loaded on a 12.5% SDS-PAGE (1 mm) using a Mini-Protean II electrophoresis cell (BioRad, Marnes-la-Coquette, France), and further stained with Coomassie Brilliant Blue G250 or silver nitrate. This overall characterization (pH, osmolality, protein concentration, absorbance and electrophoretic patterns, and embryo's sex) helped us to select homogenous samples that were stored at - 20°C for further analyses by mass spectrometry (Supplementary data S1).

In-Gel and in-Solution Tryptic Digestion

Twelve homogenous AF samples including six males and six females were pooled for protein identification. Proteins were either separated by a 4-20% SDS-PAGE followed by Coomassie Brilliant Blue G250 staining (GeLC-MS/MS analysis), or directly identified in solution (shotgun analysis) as described previously (18). The SDS-PAGE gel was cut in 20 sections (Fig. 3B) and each slice was rinsed separately in water and then acetonitrile. Proteins were then reduced with dithiothreitol, alkylated with iodoacetamide, and incubated overnight at 37°C in 25 mM NH_4HCO_3 with 12.5 ng/µL trypsin (Sequencing grade, Roche, Paris, France) as described by Shevchenko (23). Peptides were pooled and dried using a SPD1010 speedvac system. Peptide mixtures associated to each band and in-solution samples were analyzed using nanoLC-MS/MS.

NanoLC-MS/MS

The resulting peptide mixtures were analyzed using a LTQ Orbitrap Velos Mass Spectrometer (Thermo Fisher Scientific, Germany) coupled to an Ultimate® 3000 RSLC chromatographer

controlled by Chromeleon 6.8 software (Dionex, Amsterdam, The Netherlands). Five microliters of sample were desalted and pre-concentrated on a trap column (Acclaim PepMap 100 C18, 100 µm inner diameter x 2 cm long, 3 µm particles, 100 Å pores) for 10 min at 5 µl/min with 4% solvent B (0.1% formic acid, 15.9% water, 84% acetonitrile) in solvent A (0.1% formic acid, 97.9% water, 2% acetonitrile). Separation was conducted using a nanocolumn (Acclaim PepMap C18, 75 µm inner diameter × 50 cm long, 3 µm particles, 100 Å pores) at 300 nL/min by applying a gradient of 4 to 55% of solvent B during 90 min for GeLC-MS/MS analyses and during 120 min for shotgun analyses.

Data were acquired using Xcalibur 2.1 software (Thermo Fisher Scientific, San Jose, CA). The instrument was operated in positive mode in data-dependent mode. Survey full scan MS spectra (from 400 to 1800 m/z) were acquired with a resolution set at 60,000. The 20 most intense ions with charge states ≥ 2 were sequentially isolated (isolation width: 2 m/z; 1 microscan) and fragmented using CID mode (energy of 35% and wideband-activation enabled). Dynamic exclusion was active during 30 s with a repeat count of one. Polydimethylcyclosiloxane (m/z, 445.1200025) ions were used for internal calibration.

The mass spectrometry proteomics data have been submitted to the ProteomeXchange Consortium and are available *via* the PRIDE partner repository (24) with the dataset identifiers PXD008046 and 10.6019/PXD008046.

Protein Identification and Data Validation

The reliability between replicates was investigated by comparing chromatograms and spectra with Xcalibur 2.1 software. MS/MS ion searches were performed using Mascot search engine v 2.3.2 (Matrix Science, London, UK) *via* Proteome Discoverer 2.1 software (ThermoFisher Scientific)

against National Center for Biotechnology Information (NCBI) database with Chordata taxonomy (782,473 entries, downloaded in January 2017). Fragments and parents' tolerances were set at 0.80 Da and 5 ppm, respectively. The search parameters included trypsin as a protease with two allowed missed cleavages and carbamidomethylcysteine, methionine oxidation and acetylation of N-term protein as variable modifications. Mascot results obtained from the target and decoy databases searches were subjected to Scaffold 4.8.2 software (Proteome Software, Portland, USA), and displayed as clusters. Peptide and protein identification and validation were performed using the Peptide and Protein Prophet algorithms, respectively (95.0% probability, at least two different unique peptides), from three technical replicates. The abundance of identified proteins was estimated by calculating the emPAI using Scaffold Q+ software (version 4.4, Proteome Software, Portland, USA).

Keratins were not taken into consideration in the subsequent analysis as they might be either contaminant from human skin and/or chicken skin. The list of keratins identified in the analysis is however available (Supplementary data S2).

Functional Annotation Using Gene Ontology

Gene ontology terms annotations for biological processes and cellular components category provided by the GO consortium (http://www.geneontology.org/) were investigated using protein database with UniprotKB (http://www.uniprot.org), and genomic databases with Ensembl (http://www.ensembl.org) and NCBI (https://www.ncbi.nlm.nih.gov).

Comparison between Human and Chicken AF proteomes

IPI numbers and gene symbols corresponding to the proteins identified in the human AF were retrieved from the publication of Cho *et al.* (11), which integrated all previous studies on the human AF proteome. Earlier versions of the human protein sequences were recovered at the European Bioinformatics Institute website (ftp://ftp.ebi.ac.uk/pub/databases/IPI/) and updated according to the Ensembl database. Human and chicken gene symbols were compared and for associated genes, protein sequences were aligned using BLAST (https://blast.ncbi.nlm.nih.gov/Blast.cgi) to identify similarity and identity percentages between both species. Human orthologues were systematically checked using Ensembl compara (http://www.ensembl.org) to confirm previous alignments.

Phylogenetic analysis

Phylogenetic branches of emergence were defined in order to determine the moment of appearance of each gene in the tree of life. Eight possible branches for gene birth were considered: *Opisthokontas* (animals and fungi ~ 1215 million years ago (MYA)), *Bilateria* (bilateral animals ~ 937 MYA), *Chordata* (chordates ~ 722 MYA), *Vertebrata* (vertebrates ~ 535 MYA), *Tetrapoda* (tetrapods ~ 371 MYA), *Amniota* (amniotes ~ 296 MYA), *Sauropsida* (reptiles and birds ~ 276 MYA), and *Mammalia* (mammals ~ 220 MYA). For all identified proteins, the corresponding Ensembl protein ID was retrieved from the Ensembl database and the related phylogenetic trees were analyzed to highlight the specificity of proteins evolution (http://www.ensembl.org) (25). Eighty-nine trees were studied. In order to complement the phylogenetic trees, the conservation of synteny was systematically observed using Genomicus (http://www.genomicus.biologie.ens.fr) and Mapviewer (https://www.ncbi.nlm.nih.gov/mapview) (21, 26). The gene was defined as conserved if both surrounding/adjacent genes were the same.

Purification of heparin binding proteins

Heparin-Sepharose chromatography was performed according to manufacturer's instructions. Briefly, 2 mL of beads (Heparin sepharose 6 Flast Flow, GE Healthcare, Sweden) were loaded onto a polypropylene column (QIAGEN, Courtaboeuf, France) and washed with 10 volumes of water and 10 volumes of washing buffer (50 mM Tris-HCl, 150 mM NaCl, pH 7.4). After loading the crude sample (pool of ten individuals, five males and five females), the beads were washed extensively with the washing buffer until the absorbance at 220 nm reached zero, as previously reported (27). Elution of bound proteins was achieved with 50 mM Tris-HCl, 2 M NaCl, pH 7.4, until the absorbance at 220 nm reached zero. Eluted fractions were desalted and concentrated by ultrafiltration (Ultracel-3K, Merck Millipore, Molsheim, France), and analyzed by 12.5% SDS-PAGE under non-reducing and non-boiling conditions to preserve protein integrity, followed by Coomassie Brilliant Blue G250 staining.

Antibacterial assays

Antibacterial tests were conducted by direct detection of antibacterial activities after SDS-PAGE, a method adapted from Bhunia *et al.* (28). Pathogenic bacterial strains, *Salmonella enterica* serovar Enteritidis ATCC 13076 (*S.*E.) and *Listeria monocytogenes* EGD strain (*L.m.*) were provided by the International Centre for Microbial Resources (CIRM, https://www6.inra.fr/cirm_eng/Pathogenic-Bacteria) from the French National Institute for Agricultural Research (INRA, France). Pre-cultures of *S.*E. and *L.m.* were performed overnight in Trypticase Soy Broth (TSB, BD Biosciences) and in Brain Heart Infusion broth (BHI, BD Difco), respectively. This pre-culture was then used to inoculate a new culture broth (TSB or BHI) so that the mid-exponential phase was obtained after 3 or 4 hours of incubation depending on strains, with

shaking at 37 °C. Bacteria were centrifuged at 2000 *g* for 10 min at 4°C, washed twice with cold 10 mM sodium phosphate buffer (pH 7.4), and re-suspended in cold sodium phosphate buffer. Bacteria ($7.5x10^6$ Colony Forming Unit) were introduced in 25 mL of autoclaved nutrient-poor agar (10 mM phosphate buffer containing 0.03% TSB medium, 1% low-endosmosis agarose (wt/vol) (Sigma-Aldrich, Saint-Quentin-Fallavier, France), and 0.02% Tween 20.

In parallel, maximum of 20 µg of protein pools (see above) were loaded on two identical gels and further separated by 15% SDS-PAGE under non-reducing and non-boiling conditions, while maintaining the electrophoresis system at 4°C to avoid protein degradation. The first gel was stained with Coomassie Brilliant Blue G250 for further identification of antibacterial proteins by mass spectrometry. The second gel was dedicated to the antibacterial assay and was washed separately with 2.5% Triton X-100 and MilliQ water (4 x 15 min, 4°C), to eliminate SDS and renature proteins. Thereafter, the washed gel was covered with the nutrient-poor agar containing bacteria and incubated for 3 h at 37°C to allow protein diffusion in the agar. A second nutrient-rich agar (10 mM phosphate buffer containing 6% TSB medium, 1% low-endosmosis agarose (wt/vol)) was poured on the nutrient-poor agar to allow bacterial growth. After an overnight incubation at 37°C, clear zones were defined as inhibition zones. These inhibition zones were superimposed on the Coomassie-stained gel to locate bands containing antibacterial proteins. These bands were cut from the Coomassie-stained gel and further processed as described above, for protein identification by mass spectrometry.

Beta-microseminoprotein-like (BMSP) and avian beta-defensin 11 (AvBD11), two recently characterized antibacterial proteins from egg white were used as positive external controls (2 µg of proteins/well, data not shown). They were obtained as previously described (27).

Experimental Design and Statistical Rationale

To address our scientific questions, we decided to use a conventional chicken laying strain (ISA Brown, Hendrix Genetics, St Brieuc, France) at 38 weeks of age, that were raised at the UE1295 Pôle d'Expérimentation Avicole de Tours (INRA, F-37380 Nouzilly, France). Eighty fertilized eggs were incubated under standard conditions after a three-day storage at 16°C and 85% relative humidity to favor synchronization of developmental stages - knowing that conditions and duration of storage prior to incubation negatively impact embryo survival and development. The developmental stages were confirmed using the atlas of chicken developmental stages (29). Egg weight and eggshell quality were both checked and protein profiles of each individual sample were analyzed by SDS-PAGE (Supplementary data S1, sheets #1 and #2). In parallel, the sex of the embryo was determined for all corresponding samples as described in Experimental Procedures to generate a pool of combined male and females AF that could encompass most proteins composing AF, regardless of the sex. The combination of all these parameters allowed us to define 12 samples with an equilibrated sex ratio that were all homogenous in terms of stages of development, biochemical parameters (pH, osmolality, absorbance spectrum, etc.), and SDS-PAGE protein profiles. All experiments were conducted in compliance with the European legislation on the "Protection of Animals Used for Experimental and Other Scientific Purposes" (2010/63/UE) and under the supervision of an authorized scientist (S. Réhault-Godbert, Authorization no. 37-144). Two complementary bottom-up proteomic strategies (GeLC-MS/MS and shotgun analyses) were combined to give a representative overview of protein diversity in chicken AF samples. Protein sequences were retrieved from Scaffold and blasted to check their GenBank accession numbers (NCBI), and to find corresponding Ensembl and Uniprot IDs. Chicken proteins identified in other species were also searched by BLAST alignments. Functional annotation and search for

antibacterial candidates were performed by combining Ensembl, Uniprot and NCBI data on chicken proteins and their orthologues. The identification of antibacterial candidates was completed by investigating protein and peptide antibacterial domains using specific sequence analyzing tools such as InterPro (https://www.ebi.ac.uk/interpro/). For the production of enriched fractions, we defined optimal conditions (volume of samples, volume of heparin-sepharose beads) depending on the initial concentration of proteins in AF and egg white samples from ED0, ED8, ED11, ED14 or ED16 stages. All samples (raw, flow-through and eluted fractions) were analyzed by SDS-PAGE. Protein profiles of egg white at ED11 and ED16 were compared to previously published data for experimental validation (27, 30). For antibacterial assays, at least three independent assays were performed on *Listeria monocytogenes* and *Salmonella enterica* Enteritidis strains, using two positive controls (one peptide and one protein) purified from chicken egg white. The phylogenetic analysis and corresponding figures were conducted combining data available in Ensembl databases and in literature.

RESULTS

Proteomic profiling of AF – The global protein profile of AF assessed by SDS-PAGE is completely different when comparing the first half (ED8-11) to the second half of development (ED12-16), due to the massive transfer of egg white into the AF from ED12 onwards (Fig. 3A). In this context, to better appreciate the intrinsic protein composition of chicken AF, a bottom-up proteomic study combining two different strategies was conducted on AF samples collected at ED11, the last day before egg white transfer. AF of 40 eggs corresponding to 17 females and 22 males (plus one embryo for which tissue sample was discarded) were collected as detailed in Experimental

Procedures. Individual SDS-PAGE profiles (using both Coomassie Brilliant Blue and silver stainings, Supplementary data S1, sheet #2), protein concentrations, pH, osmolality values, and spectra of absorbance were determined for each AF sample (Supplementary data S1, sheet #1). Since no significant differences could be detected between individuals and regardless of the sex after integration of all parameters, we constituted a pool representing 12 homogenous samples with equilibrated sex ratio for further analyses by high-resolution mass spectrometry. We performed both in-gel and in-solution independent tryptic digestions followed by nanoLC-MS/MS analysis for protein identification and quantification. For in-gel digestion, a 4-20% SDS-PAGE gel was cut (Fig. 3B) and each slice was further processed as mentioned in Experimental Procedures. We filtered data for keratins and other proteins that could reflect sample contamination by experimenters (Supplementary data S2). Ninety-one non-redundant proteins were identified, including 49 for shotgun analysis and 70 for GeLC-MS/MS analysis (Fig. 4). Twenty-eight proteins were common to both approaches (Fig. 4). Some of the identified proteins sometimes refer to other bird species than *Gallus gallus,* such as *Coturnix japonica* (japanese quail) or fish *Tachysurus fulvidraco* (yellow catfish) (Supplementary data S3), either because the *Gallus gallus* counterpart is still missing in the genome annotation, or because these specific variants have not yet been identified in the chicken species. Raw files, proteins, peptides scores, information and quantitative values are compiled in Supplementary data S2 and S3.

Figure 4 illustrates the results using log10 of the Exponentially Modified Protein Abundance Index (emPAI), and log10 of the percentage of the protein sequence coverage (Supplementary data S3). According to this graph, a set of ten proteins emerged to be very abundant compared with others. Regardless of the methods, apolipoprotein A1 (APOA1), ovotransferrin (TF) and alpha-fetoprotein (AFP) were three of the most abundant proteins found in AF at ED11 (Fig. 4). If we consider the

ten proteins with the highest emPAI for both shotgun and GeLC-MS/MS approaches, seven were common to both approaches: APOA1, TF, AFP, TTR (transthyretin), OVAL (ovalbumin), LYZ (lysozyme) and ALB (serum albumin). APOC3 (apolipoprotein C-III), SPINK7 (ovomucoid) and MDK (midkine) were found in the top ten resulting from in-solution method whereas GC (vitamin-D protein), FSTL1 (follistatin-related protein A) and ACTB (actin, cytoplasmic 1) were identified in the top ten resulting from GeLC-MS/MS method (Supplementary data S3). The ten major proteins of AF after integration of quantitative values from both approaches are compiled in Table 1. These proteins represent about 66% to 81% of the total protein content for GeLC-MS/MS and shotgun analyses, respectively. A primary analysis of the Gene Ontology terms of these major proteins revealed that six of them are associated with lipid, vitamin or hormone metabolisms (APOA1, AFP, ALB, GC, TTR, APOC3), while two additional proteins participate in embryo's defense against pathogens (TF and LYZ) (Table 1). On the other hand, OVAL, which is specific to birds (31), is assumed to have a role in embryo's nutrition as a source of amino acids while SPINK7 is a very potent protease inhibitor containing three inhibitory kazal-like domains. Despite the lack of information related to its physiological targets, SPINK7 presumably regulates endogenous proteolytic activities, preventing uncontrolled protein degradation and preserving overall protein integrity.

Except APOC3, all these proteins were described to be secreted (Table 1).

Gene Ontology analysis of the chicken AF proteome – An analysis of the occurrence of the Gene Ontology terms related to the biological processes and molecular functions, revealed that the most representative function of AF proteins is associated with cellular adhesion and migration, followed by functions associated with metabolism and transport of lipids, vitamins and carbohydrates (Fig.

5). We defined six categories of general functions knowing that many proteins have multiple functions and appear in various groups (Supplementary data S3, sheet #2). Three processes seem to be exacerbated when comparing the first approach on the ten major proteins (Table 1) and the second performed on the whole list of proteins (Fig. 5): cellular migration and adhesion, metabolism (and transport) of lipids and vitamins, and immunity (defense). Twenty-one genes are classified in "immune response" (Fig. 5). This category includes effectors/antibacterial molecules such as TF and LYZ, but also immunoglobulins (CAO79236.1 and LOC107051274), mucins (MUC5ACL, MUC5B), protein TENP (BPIFB2), ovoinhibitor (SPINK5) or ovalbumin-related protein X (OVALX). It is noteworthy that ten proteins could not be assigned to any known function, including two of the major proteins identified in this work: OVAL and SPINK7 (Table 1). Interestingly, among these 10 proteins, seven are potent protease inhibitors whose biological functions in the egg are still uncertain.

Comparison with the human AF proteome – Chicken AF proteome at ED11 was compared with the 936 proteins identified in the human AF proteome (11) to shed light on chicken AF specificities. Out of the 936 proteins identified by Cho *et al.* (11), after removal of redundancies and considering the updated human genome assembly (GRCh38.p10, GCA_000001405.25, Dec 2013), we ended with 842 proteins, plus 81 proteins retrieved from the other studies (11) (Supplementary data S3, sheet #4). This data integration led to 923 proteins that are identified so far in the human AF. When considering gene symbols, 48 proteins were found to be common to both chicken and human AF (Fig. 6). We confirmed their orthology relationship using Ensembl database. Although they do not share the same gene symbol, LOC107051274 and CAO79236.1

(immunoglobulins) can be likened to human immunoglobulins G, which are omnipresent in human AF (11).

Assessment of the antibacterial potential of the chicken AF and comparison to egg white – To further assess the biological significance of AF as a source of antibacterial proteins, we adapted a method from Bhunia *et al.* (28) that allows the direct detection of antibacterial activities after a SDS-PAGE. Briefly, AF proteins and peptides were separated by SDS-PAGE under non-boiling and non-reducing conditions. After removing SDS from the gel, a nutrient-poor agar containing bacteria was poured on the SDS-PAGE gel and left on the bench for three hours to allow for proteins diffusion. Thereafter, a nutrient-rich agar was added to favor bacterial growth. Regions were bacteria could not grow, because of the presence of antibacterial proteins and peptides, were detected as translucent areas. Antibacterial activity was tested against *Listeria monocytogenes* (*L.m.*; Gram-positive) and *Salmonella enterica* Enteritidis (*S*.E.; Gram-negative), two bacteria against which egg proteins have already demonstrated their activity under similar experimental conditions (27). Because AF protein composition at ED11 is thoroughly modified from ED12 onwards, due to the presence of egg white proteins masking the intrinsic protein composition of AF, we explored the antibacterial potential of AF at both ED11 and ED16 (Fig. 7A and 7B). Assays were performed on both pools of total AF, but also on fractions enriched in antibacterial molecules obtained after heparin-affinity chromatography (HBP, heparin-binding proteins), as previously described (27). At ED11, results show only weak antibacterial activity against Gram-positive *L.m.* for crude AF, and also for flow-through and eluted (HBP) fractions collected from heparin-chromatography (Fig. 7A). Overall, the antibacterial activity against both *L.m.* and *S*.E. seems to increase with the egg white transfer, with new inhibition zones appearing for both bacterial strains

at ED16, particularly for the eluted fraction: one to five anti-*L.m.* zones from ED11 to ED16, *versus* only one anti-*S.*E. zone at ED16 (Fig. 7A and 7B).

The HBP fraction of egg white has been previously reported to contain antibacterial proteins and activities (27, 30). In order to explore the regulation of egg white antibacterial activities all along incubation, we performed antibacterial assays on egg white HBP, sampled and purified at ED0, 8 and 14 (Fig. 8). These stages of development were chosen knowing that egg white transfer into the AF is nearly complete at ED16 (almost no residual egg white at this stage). The antibacterial activity of the egg white HBP against both *L.m.* and *S.*E. changes during embryonic development. Indeed, the strong activity detected at ED0 at 19 kDa (Fig. 8, ED0) against *L.m.*, progressively disappears to the benefit of anti-*L.m.* activities in the higher molecular masses (> 250 kDa) or lower molecular masses (about 14 kDa) (Fig. 8, ED8 and 14). In addition, three bands exhibiting anti-*S.*E. activities at ED0 (19 kDa and 14 kDa), were barely detectable at later stages (ED8 and ED14, Fig. 8). Antibacterial activity of egg white HBP at ED14 is in agreement with the antibacterial activities of the HBP in the AF at ED16: one inhibition zone in the higher masses (> 250 kDa), and three bands in the lower masses (< 15 kDa) for *L.m.*, as opposed to one inhibition zone at 15 kDa for *S.*E. An anti-*L.m.* zone between 25 and 37 kDa for the HBP at ED16 solely appears in AF, which might be associated with specific activation or expression of antibacterial proteins in this fluid at this stage.

Identification of candidate proteins responsible for antibacterial activity in AF at ED11 and ED16, by mass spectrometry - In parallel to gels prepared for antibacterial assays, a SDS-PAGE realized under the same conditions was stained with Coomassie Brilliant Blue. The superimposition of both the stained SDS-PAGE gel and the gel exhibiting antibacterial activities allowed us to select

SDS-PAGE bands (labeled A to H, Fig. 7A and 7B), which were susceptible to contain antibacterial proteins or peptides. These bands were analyzed by mass spectrometry at ED11 (band A, Fig. 7A) and at ED16 (bands B to H, Fig. 7B), which led to the identification of 29 proteins (Table 2) (Supplementary data S4). Among the most likely candidates possessing antibacterial activity, we could only find LYZ for AF at ED11. For sample at ED16, the most probable molecules are LYZ, but also AvBD11, vitelline membrane outer layer protein 1 (VMO1), avidin (AVD), beta-microseminoprotein-like (LOC101750704 *G. gallus*/LOC104403769 *N. notabilis*), OVALX, TF, BPIFB2, ovocleidin-17 (OC-17), midkine (MDK) (27, 30, 32-34). The antibacterial activity of SPINK5 was previously demonstrated against *Bacillus* but not against the two strains tested in the present study (35).

Phylogenetic analysis – To further explore specificities of the chicken model, we analyzed the evolutionary history of the 91 proteins identified in chicken AF at ED11 (Fig. 9), combined to the 18 new genes identified from the antibacterial assays at ED16 (Table 2). Among the 109 proteins analyzed, the corresponding genes of 100 of them were found in Ensembl database, but the evolution of only 90 of them could be successfully investigated. The 19 genes without an Ensembl accession number may be located in the non-sequenced or non-annotated region of the chicken genome or lack a related phylogenetic tree (AFP; APOA2; APOC3; ASTL; COL1A1; DCHS1; DMBTIL3; DSG4; LOC104403769 / LOC101750704; LOC107051274; LOC107323914; MUC6; OC-17; OVSTL; PSAP; SLC3A2; SPINT4; immunoglobulin heavy chain variable region CAO79236.1 (no GeneID); transferrin ADK35120.1 (no GeneID)). Studying the moment of appearance in the tree of life of each gene coding for these proteins, revealed that most of them (88%) appeared before 371 MYA (tetrapods) with a peak between 937 MYA and 722 MYA

(bilateria) for almost 50% of them (Fig. 9A). Orthologues of most of these genes (80 genes) are present in primates and other mammals. Nine proteins lacking orthologues in mammals were identified. Four of these genes, AvDB11, OVAL, OVALY and OVALX are bird-specific and are also found in duck or turkey, but not in lizard or turtle (31, 36) (Fig. 9B). These genes appeared 104 MYA after divergence between birds and reptiles. APOV1, present in bird and reptile genomes, emerged 276 MYA. AVD and RBP appeared in vertebrates and chordates, respectively (21, 37, 38). These genes are present in reptile and bird genomes but also in fishes or frogs, suggesting that they have been likely lost in the mammalian branch. Concerning RBP, the death of the gene in mammalian genome could not be confirmed because we could not find any pseudogenes. Genes encoding VTG1 and VTG2, which appeared in bilateria, are found in all egg-laying species including platypus (VTG), which is the only oviparous mammal. Some authors have shown that these VTG1- and VTG2-encoding genes progressively lost their functions and became pseudogenes during mammalian evolution (38) concomitantly with the appearance of lactation and placentation. Several authors have described the process of pseudogenization of VTG genes during mammalian evolution (21, 38). OC-116 gene has been described as orthologous to MEPE gene in other species, although OC-116 has likely acquired some tissue specificities (eggshell biomineralization) resulting from divergent and adaptive evolution (21). Regarding SPINK7 (ovomucoid) and SPINK5 (ovoinhibitor) genes (respectively ENSGALG00000003512 and ENSGALG00000031496), which appeared between 296 MYA and 276 MYA, the story is quite tricky as the data available in Ensembl were not sufficient to better define their evolutionary fate in various amniotes (Fig. 9). Both chicken proteins share 50.24% sequence identity. In *Gallus gallus*, they are located right next to each other on chromosome 13 within a 33-kb region, closed to other SPINK genes, which have all probably arisen by local duplication from a common ancestor

gene. All these proteins are characterized by one to 14 Kazal-like domains, which are known as evolutionary conserved domains, commonly occurring in tandem arrays, and usually associated with inhibitory activities against serine proteases. This SPINK gene family seems to evolve very quickly, and unfortunately, synteny analysis of these genes does not help in strengthening the corresponding phylogenetic tree, because all these genes are located within the same locus. Nevertheless, both chicken genes seem to have specifically duplicated in *Sauropsida* before the speciation of birds and reptiles. SPINK7 (ovomucoid) and SPINK5 (ovoinhibitor) genes encode proteins of 210 amino acids and 472 amino acids, respectively, the latter being characterized by a C-terminal extension of 252 amino acids containing four additional Kazal-like domains. A low percentage of sequence identity of chicken SPINK5 is also observed with mammalian SPINK5 proteins (less than 30%), associated with a difference in the length of protein sequences and the number of corresponding Kazal-like domains (seven for chicken SPINK5/ENSGALG00000031496 and 14 for human SPINK5). Ultimately, the phylogenetic tree strongly suggests that both genes have been lost in several mammals, in particular in primates, hinting at a pseudogenization in mammals, in parallel to a speciation after duplication in birds.

DISCUSSION

In amniotes, AF plays a crucial role in maintaining a stable and protecting milieu. In humans, the exhaustive characterization of its biochemical components (including proteins) at various stages of gestation in normal and pathological situations constitutes an extremely valuable approach to help in the diagnosis of human fetal disorders and infections. In avian species, there is evidence that AF also participates in homeostasis around the embryo (18) and in its protection against physical constraints (19). Nevertheless, little is known about the physiological role of AF proteins

in birds, and there is to date only few proteomic data available on this fluid for these species. Yet, we suspect that the development of the embryo in an egg, with no mother-connected tissues, requires some inherent specificities and structural particularities that may influence AF protein profile and composition in birds.

In the present work, we identified 91 non-redundant proteins with high confidence, in the chicken AF at ED11, before egg white transfer, which is higher than the 47 proteins identified in a preliminary study (18). This number is however far less than the number of proteins identified in two surveys performed simultaneously on the human AF (219 (9) and 923 (11)). This difference can be partly explained by the several exchanges with maternal tissues existing during gestation in human, which results in some specific AF protein profile, and the fact that in humans, metabolic wastes originating from the embryo are indeed recovered in the AF, which is actually not the case in the chicken model (6).

All combined results indicate that more than half of the proteins identified in the chicken AF at ED11 have orthologues that have been identified as components of the human AF, which highlights that the overall protein composition of the AF before egg white transfer exhibits high similarities with the human AF. This is strengthened by the observation that, five of the top-ten high abundant proteins identified in the chicken AF at ED11 (Table 1 - ALB, TF, AFP, GC and APOA1) are also listed in the 15 high abundant proteins recovered in human AF (9, 11). It is to note that the TF identified in the chicken AF corresponds to ovotransferrin, which shares 51.76% and 51.73% sequence identity with human serotransferrin (P02787) and human lactotransferrin (P02788), respectively. This moderate sequence identity between ovotransferrin and its human homologues, together with some specific structural features, such as glycosylation and number of disulfide bridges (39), suggest similar functions with regard to their iron-binding capacity and

storage, but also possibly diverging functions. As an example, both lactotransferrin and ovotransferrin, but not serotransferrin, are depicted as antibacterial proteins (40, 41). Interestingly, two of the top-ten abundant proteins have no assigned functions: OVAL and SPINK7. These two proteins are major proteins of egg white (42) and appear belatedly in the timeline of evolution, with OVAL and very likely SPINK7, being specific to birds (Fig. 9).

Because of these similitudes in protein composition, most functions allocated to the most abundant proteins composing the human AF resemble that of the chicken AF: 1) metabolism and transport of vitamins, lipids and hormones, 2) immune response, and 3) hemostasis and homeostasis. The analysis of the list of chicken AF proteins using automatic tools dedicated to Gene Ontology annotation, emphasized functions associated with many aspects of developmental biology such as morphogenesis and organogenesis, which include cell proliferation, adhesion and migration processes (Fig. 5). These functions were also significantly highlighted in the analysis of human AF proteins (9, 11). However, the biological significance of these functions in the chicken and human AF are likely to be overestimated since most proteins allocated to this function are very-low-abundant proteins and may reflect some embryonic or extraembryonic tissues desquamation (Supplementary data S2 and S3, sheet #2).

Overall, it is noteworthy that on the ten most abundant proteins in the chicken AF, four of them are major egg white proteins (OVAL, SPINK7, LYZ, TF) (42) whereas ALB, APOA1, GC, TTR, APOC3 are egg yolk abundant proteins (43). The gene expression of OVAL, SPINK7, LYZ and TF, which together account for about 80% of the egg white proteins, is under hormonal control in the oviduct and will not be expressed until the onset of sexual maturity of hens (44). OVAL, SPINK7 and TF are thus likely to flow from egg white into the amniotic sac, rather than being expressed by embryonic tissues. As for LYZ, the ongoing hypothesis is that it also originates from

egg white although it is also expressed by a wide range of tissues (45). On the other hand, APOA1, ALB, GC, TTR and APOC3 represent highly-abundant proteins of egg yolk (43), which also suggests that they are somehow transferred to the chicken AF from the egg yolk, or from the yolk sac where they are highly expressed (4). The molecular mechanisms by which egg white and egg yolk proteins enter the AF during the first half of incubation are still not completely solved (16, 18). In contrast, the embryonic origin of AFP is unequivocal since it is a fetus-specific protein present in a range of (extra)embryonic tissues, highly expressed in the yolk sac and the embryo's liver (46). Although all these proteins are the most abundant proteins identified in the chicken AF at ED11, their protein concentration in the AF at this stage (0.01 mg/mL) is not comparable to those recovered in the egg yolk (> 10 mg/mL, 22% of dry matter (47)) and egg white (380 mg/mL (30)). The low protein concentration of chicken AF at ED11 suggests that the presence of proteins in this fluid may reflect some passive transfer from these egg compartments towards AF, rather than an active mechanism. This concentration is also lower as compared with that of human AF: it is about 4 mg/mL at 15 weeks of gestation (10), it reaches its maximum (6 mg/mL) at 22-27 weeks of gestation and starts to decrease to about 2 mg/mL at 37-40 weeks of gestation (48). It also highlights that chicken AF cannot meet embryo's requirements in energy up to ED11, suggesting that during the first half of incubation, the egg yolk is the major source of energy/lipoproteins for the embryo. Some proteolytic activities (18) have been however detected in AF at ED11, but, because of the low abundance of the corresponding proteins, the biological significance of these activities in AF remains questionable as compared to the egg yolk, which concentrates hydrolytic enzymes (43, 49). The same observation can be made for the faint antibacterial activities detected at ED11 in the chicken AF, compared with the egg white, which concentrates high amounts of antibacterial proteins such as active lysozyme (30, 42).

These considerations along with the detection of antibacterial activities, corroborates the role of chicken AF in protecting the embryo against a range of potential physical, physicochemical and antibacterial pressures. Anyway, it also emphasizes some particularities such as the presence in this fluid of bird-specific proteins involved in embryo nutrition, eggshell biomineralization and egg defense (OVAL, OVALY, OVALX, OC-116, AvBD11) and likely SPINK7/ovomucoid and SPINK5/ovoinhibitor, recently renamed SPIK7 and SPIK5 in NCBI Gene database (May 2017) and which functions are still speculative. We also identified some oviparous-specific proteins (VTG1, VTG2, AVD, APOV1) (Fig. 9B), that appears concomitantly with the appearance of the egg-laying type of reproduction. Additionally, it revealed the presence of 43 proteins that have been uniquely found in the chicken AF as compared with the human AF (Fig. 6). Nevertheless, the high number of proteins reported for human AF as compared with chicken AF, merits further studies. In particular, a comprehensive study of the proteome profiling of the allantoic fluid would probably answer some of these questions, as we might find allantoic fluid-specific proteins whose orthologues have been actually identified in human AF. Future studies of this specific fluid would confirm or not whether proteins from both chicken amniotic and allantoic fluids actually encompass most proteins identified in the human AF.

To conclude, the story of the chicken AF remains quite simple up to ED11, with a protein composition that is comparable to that of human AF or at least with similar general functions. However, the protein composition of AF is deeply revised after egg white transfer, which occurs between ED11 and ED12 (18). AF protein profile then completely merges with egg white protein profile (Fig. 3A) (16, 18, 50), which ten major proteins consist of OVAL, LYZ, TF, OVALY, SPINK7, VMO1, SPINK5, AVD, OVALX, and CST3 (42). Despite the presence of AF hydrolytic enzymes (18), no major proteolytic degradation of egg white proteins was detected after its inflow

into the amniotic cavity, up to ED19. These data suggest that abundant egg white antiproteases, namely SPINK7, SPINK5, cystatin (CST3) and ovostatin (OVST), remain active during the whole duration of incubation. It is rather interesting to note that the inflow of egg white actually reinforces some proteins that were already present as major proteins in ED11-AF (OVAL, LYZ, TF and SPINK7, Table 1). More than half of these major egg white proteins are related to innate defense (LYZ, TF, VMO1, SPINK5, AVD, OVALX) while three of them are protease inhibitors (SPINK5, SPINK7, CST3). Therefore, we suspected that chicken AF enriched in egg white proteins from ED12 onwards, acquires an increased antibacterial potential. An in-gel antibacterial assay was developed to better appreciate the relevance and effective activity of these so-called antibacterial proteins in the AF-egg white mixture. We showed that ED16-AF displays higher antibacterial potential against the two bacterial strains tested, namely *Listeria monocytogenes* (Gram-positive strain) and *Salmonella enterica* Enteritidis (Gram-negative strain) compared with ED11-AF. A total of 29 proteins were identified in the corresponding zones lacking bacterial growth (Table 2), including LYZ, AvBD11, VMO1, AVD, LOC101750704 (BMSP), OVALX, TF, OC-17, BPIFB2 and MDK that were all previously reported to exhibit antibacterial activities (27, 32-34, 41, 51-56). These results demonstrate that the antibacterial potential of egg white proteins is not impaired by the incubation temperature, nor by proteolytic degradation (thanks to egg white antiproteases), and that its antimicrobial potential remains effective after transfer into the AF. However, subtle differences are detected in AF at ED16 after completion of egg white transfer, but also in remnant egg white at ED8-14. Indeed, additional antibacterial zones were visualized on anti-*Listeria monocytogenes* assays performed with ED16-AF and ED8-14-EW (Fig. 7 and 8, respectively). Unexpectedly, they correspond to proteins or protein complexes/association of very high molecular masses (>250 kDa). These high molecular mass complexes were identified as a mixture

of egg white proteins: clusterin (CLU, 62 kDa), SPINK5 (52 kDa), OVALX (44 kDa), AVD (tetramer of 17 kDa units), OVAL (43 kDa), TF (78 kDa), ovomucin (234 kDa), LYZ (14 kDa), BPIFB2 (47 kDa), etc. The progressive appearance of these complexes suggests some specific regulation occurring during incubation. Indeed, the re-distribution of the egg white water towards extraembryonic sacs during the first half of the development, concentrates egg white proteins, thus promoting proteins interactions between antibacterial proteins/peptides of low molecular masses (LYZ, LOC101750704 (BMSP), AvBD11, etc.) with proteins of higher molecular masses (ovomucin). These protein complexes do not seem to affect the antibacterial protein/peptide activities, which are still effective even before oral absorption by the embryo (Fig. 8), thus forming another barrier against bacteria around the embryo's body. As the embryo moves and develops in the AF until hatching, it can be hypothesized that this mixture of antimicrobial molecules and mucins may also constitute a protective biofilm deposited on embryo feathers and skin to protect the embryo and the newborn chick. This is partly supported by several articles that reported the presence of lysozyme and defensin-like molecules in human *vernix caseoa* (57, 58), a protective lipid-rich substance that covers the skin of the human fetus during the last trimester of gestation and the newborn. Whether this enhanced antimicrobial potential of AF prior to the swallowing has some relevant biological significance with respect to the protection of the embryonic digestive tract should also retain major attention. Indeed, the fate of AF-egg white proteins mixture that is ingested by the embryo is still poorly understood. These proteins remain very stable in egg white during incubation but also in AF (16, 18) and all along the digestive tract of the embryo, up to ED19 (50). Similarly, some egg white typical profiles can be visualized in the yolk sac content at ED20 just before emergence of the chick, with OVAL, LYZ and TF as major proteins (13, 50). These data together with the fact that OVAL was recovered as a native form in central nervous

system and other embryonic organs (16), suggest that this protein and possibly related proteins OVALX and OVALY may have other functions than nutrition (31). From these data, we infer that intrinsic egg yolk proteases exhibit limited proteolytic activities on egg white content at this stage. This can be partly explained by the gradual increase of the yolk pH, which impacts the activity of aspartic proteases (59), such as chicken cathepsin D, a key enzyme in yolk processing (49). It seems that everything converges to protect egg white-AF protein content from uncontrolled proteolytic and thermal degradation (presence of numerous proteases inhibitors together with the progressive increase in the thermostable S-ovalbumin during incubation (60, 61)) to be utilized by the embryo and/or the chick at a very precise moment of its development/growth.

Besides its major interests for comparative biology approaches and for deciphering the functions and specificities of avian extraembryonic compartments, these results (protein profile and composition) constitute a reference start point for analyses of pathological conditions due to infection or impaired development occurring during the chicken embryonic development. As examples, in humans, differences in concentration of AFP, defensins, vitamin-binding proteins, APOA1, and TTR in AF (also identified in chicken AF) have been associated with many metabolic and developmental disorders such as trisomy, developmental delays, preterm-labor, inflammation or infections (62-64). These data combined to other high throughput methods such as metabolomics or epigenetics, and general egg quality traits, may constitute useful tools for poultry production. Indeed, these approaches may help to further investigate the impact of chicken lines, housing systems and nutrition of hens (65, 66) or of more subtle egg manipulations including conditions of egg storage prior to the incubation (67) or thermal manipulation of eggs (68), on many aspects of chicken embryo's development and health. We believe that the identification of a set of molecular markers for abnormal embryonic development in chicken extraembryonic fluids,

will contribute to the development of tools to improve poultry management and, to increase the robustness of chicks and chickens to help them face contrasted and changing environments.

REFERENCES

1. Moran, E. T. Jr. (2007) Nutrition of the developing embryo and hatchling. Poult. Sci. 86, 1043-1049

2. Bellairs, R., and Osmond, M. (2014) The Atlas of Chick Development, Third edition Ed., Academic Press, Oxford, UK

3. Sheng, G., and Foley, A. C. (2012) Diversification and conservation of the extraembryonic tissues in mediating nutrient uptake during amniote development. Ann. N. Y. Acad. Sci. 1271, 97-103

4. Yadgary, L., Wong, E. A., and Uni, Z. (2014) Temporal transcriptome analysis of the chicken embryo yolk sac. BMC Genomics 15, 690

5. Gabrielli, M. G., and Accili, D. (2010) The chick chorioallantoic membrane: a model of molecular, structural, and functional adaptation to transepithelial ion transport and barrier function during embryonic development. J. Biomed. Biotechnol. 2010, 940741

6. Underwood, M. A., Gilbert, W. M., and Sherman, M. P. (2005) Amniotic fluid: not just fetal urine anymore. J. Perinatol. 25, 341-348

7. Bakalar, N. (2012) Sensory science: partners in flavour. Nature 486, S4-5

8. Orczyk-Pawilowicz, M., Jawien, E., Deja, S., Hirnle, L., Zabek, A., and Mlynarz, P. (2016) Metabolomics of Human Amniotic Fluid and Maternal Plasma during Normal Pregnancy. PloS one 11, e0152740

9. Michaels, J. E. A., Dasari, S., Pereira, L., Reddy, A. P., Lapidus, J. A., Lu, X. F., Jacob, T., Thomas, A., Rodland, M., Roberts, C. T., Gravett, M. G., and Nagalla, S. R. (2007) Comprehensive

proteomic analysis of the human amniotic fluid proteome: Gestational age-dependent changes. J. Proteome Res. 6, 1277-1285

10. Romero, R., Kusanovic, J. P., Gotsch, F., Erez, O., Vaisbuch, E., Mazaki-Tovi, S., Moser, A., Tam, S., Leszyk, J., Master, S. R., Juhasz, P., Pacora, P., Ogge, G., Gomez, R., Yoon, B. H., Yeo, L., Hassan, S. S., and Rogers, W. T. (2010) Isobaric labeling and tandem mass spectrometry: A novel approach for profiling and quantifying proteins differentially expressed in amniotic fluid in preterm labor with and without intra-amniotic infection/inflammation. J. Matern. Fetal Neonatal. Med. 23, 261-280

11. Cho, C. K. J., Shan, S. J., Winsor, E. J., and Diamandis, E. P. (2007) Proteomics analysis of human amniotic fluid. Mol. Cell. Proteomics 6, 1406-1415

12. Yoshizaki, N., Ito, Y., Hori, H., Saito, H., and Iwasawa, A. (2002) Absorption, transportation and digestion of egg white in quail embryos. Dev. Growth Differ. 44, 11-22

13. Geelhoed, S. E., and Conklin, J. L. (1966) An electrophoretic study of proteins in chick embryonic fluids. J. Exp. Zool. 162, 257-261

14. Baintner, K., and Fehér, G. (1974) Fate of egg white trypsin inhibitor and start of proteolysis in developing chick embryo and newly hatched chick. Dev. Biol. 36, 272-278

15. Cirkvenčič, N., Narat, M., Dovč, P., and Benčina, D. (2012) Distribution of chicken cathepsins B and L, cystatin and ovalbumin in extra-embryonic fluids during embryogenesis. Br. Poult. Sci. 53, 623-630

16. Sugimoto, Y., Sanuki, S., Ohsako, S., Higashimoto, Y., Kondo, M., Kurawaki, J., Ibrahim, H. R., Aoki, T., Kusakabe, T., and Koga, K. (1999) Ovalbumin in developing chicken eggs migrates from egg white to embryonic organs while changing its conformation and thermal stability. J. Biol. Chem. 274, 11030-11037

17. Muramatsu, T., Hiramoto, K., Koshi, N., Okumura, J., Miyoshi, S., and Mitsumoto, T. (1990) Importance of albumen content in whole-body protein synthesis of the chicken embryo during incubation. Br. Poult. Sci. 31, 101-106

18. Da Silva, M., Labas, V., Nys, Y., and Rehault-Godbert, S. (2017) Investigating proteins and proteases composing amniotic and allantoic fluids during chicken embryonic development. Poult. Sci. 96, 2931-2941

19. Romanoff, A. L. (1960) The Avian Embryo. Structural and functional development, The Macmillan Compagny, New York, USA

20. Romanoff, A. L., and Romanoff, A. J. (1967) Biochemistry of the avian embryo: A Quantitative Analysis of Prenatal Development, Interscience Publishers, a division of John Wiley & Sons, Inc., New York (USA), London (UK), Sydney (Australia)

21. 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

22. Clinton, M., Haines, L., Belloir, B., and McBride, D. (2001) Sexing chick embryos: a rapid and simple protocol. Br. Poult. Sci. 42, 134-138

23. Shevchenko, A., Wilm, M., Vorm, O., and Mann, M. (1996) Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. Anal. Chem. 68, 850-858

24. Vizcaino, J. A., Deutsch, E. W., Wang, R., Csordas, A., Reisinger, F., Rios, D., Dianes, J. A., Sun, Z., Farrah, T., Bandeira, N., Binz, P. A., Xenarios, I., Eisenacher, M., Mayer, G., Gatto, L., Campos, A., Chalkley, R. J., Kraus, H. J., Albar, J. P., Martinez-Bartolome, S., Apweiler, R., Omenn, G. S., Martens, L., Jones, A. R., and Hermjakob, H. (2014) ProteomeXchange provides globally coordinated proteomics data submission and dissemination. Nat. Biotechnol. 32, 223-226

25. Vilella, A. J., Severin, J., Ureta-Vidal, A., Heng, L., Durbin, R., and Birney, E. (2009) EnsemblCompara GeneTrees: Complete, duplication-aware phylogenetic trees in vertebrates. Genome Res. 19, 327-335

26. Dufourny, L., Levasseur, A., Migaud, M., Callebaut, I., Pontarotti, P., Malpaux, B., and Monget, P. (2008) GPR50 is the mammalian ortholog of Mel1c: evidence of rapid evolution in mammals. BMC Evol. Biol. 8, 105

27. Guyot, N., Labas, V., Harichaux, G., Chesse, M., Poirier, J. C., Nys, Y., and Rehault-Godbert, S. (2016) Proteomic analysis of egg white heparin-binding proteins: towards the identification of natural antibacterial molecules. Sci. Rep. 6, 27974

28. Bhunia, A. K., Johnson, M. C., and Ray, B. (1987) Direct Detection of an antimicrobial peptide of Pediococcus-Acidilactici in Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis J. Ind. Microbiol. 2, 319-322

29. Hamburger, V., and Hamilton, H. L. (1951) A series of normal stages in the development of the chick embryo. J. Morphol. 88, 49-92

30. Guyot, N., Rehault-Godbert, S., Slugocki, C., Harichaux, G., Labas, V., Helloin, E., and Nys, Y. (2016) Characterization of egg white antibacterial properties during the first half of incubation: A comparative study between embryonated and unfertilized eggs. Poult. Sci. 95, 2956-2970

31. Da Silva, M., Beauclercq, S., Harichaux, G., Labas, V., Guyot, N., Gautron, J., Nys, Y., and Rehault-Godbert, S. (2015) The Family Secrets of Avian Egg-Specific Ovalbumin and Its Related Proteins Y and X. Biol. Reprod. 93, 71

32. Wellman-Labadie, O., Lakshminarayanan, R., and Hincke, M. T. (2008) Antimicrobial properties of avian eggshell-specific C-type lectin-like proteins. FEBS letters 582, 699-704

33. Rehault-Godbert, S., Labas, V., Helloin, E., Herve-Grepinet, V., Slugocki, C., Berges, M., Bourin, M.-C., Brionne, A., Poirier, J.-C., Gautron, J., Coste, F., and Nys, Y. (2013) Ovalbumin-related Protein X Is a Heparin-binding Ov-Serpin Exhibiting Antimicrobial Activities. J. Biol. Chem. 288, 17285-17295

34. Svensson, S. L., Pasupuleti, M., Walse, B., Malmsten, M., Morgelin, M., Sjogren, C., Olin, A. I., Collin, M., Schmidtchen, A., Palmer, R., and Egesten, A. (2010) Midkine and pleiotrophin have bactericidal properties: preserved antibacterial activity in a family of heparin-binding growth factors during evolution. J. Biol. Chem. 285, 16105-16115

35. Bourin, M., Gautron, J., Berges, M., Attucci, S., Le Blay, G., Labas, V., Nys, Y., and Rehault-Godbert, S. (2011) Antimicrobial Potential of Egg Yolk Ovoinhibitor, a Multidomain Kazal-like Inhibitor of Chicken Egg. J. Agric. Food Chem. 59, 12368-12374

36. Dombre, C., Guyot, N., Moreau, T., Monget, P., Da Silva, M., Gautron, J., and Rehault-Godbert, S. (2017) Egg serpins: The chicken and/or the egg dilemma. Semin. Cell. Dev. Biol. 62, 120-132

37. Ahlroth, M. K., Grapputo, A., Laitinen, O. H., and Kulomaa, M. S. (2001) Sequence features and evolutionary mechanisms in the chicken avidin gene family. Biochem. Biophys. Res. Commun. 285, 734-741

38. Brawand, D., Wahli, W., and Kaessmann, H. (2008) Loss of egg yolk genes in mammals and the origin of lactation and placentation. PLoS Biol. 6, e63

39. Metz-Boutigue, M. H., Jolles, J., Mazurier, J., Schoentgen, F., Legrand, D., Spik, G., Montreuil, J., and Jolles, P. (1984) Human lactotransferrin: amino acid sequence and structural comparisons with other transferrins. Eur. J. Biochem. 145, 659-676

40. Bellamy, W., Takase, M., Yamauchi, K., Wakabayashi, H., Kawase, K., and Tomita, M. (1992) Identification of the bactericidal domain of lactoferrin. Biochim. Biophys. Acta 1121, 130-136

41. Baron, F., Jan, S., Gonnet, F., Pasco, M., Jardin, J., Giudici, B., Gautier, M., Guerin-Dubiard, C., and Nau, F. (2014) Ovotransferrin plays a major role in the strong bactericidal effect of egg white against the Bacillus cereus group. J. Food Prot. 77, 955-962

42. Mann, K. (2007) The chicken egg white proteome. Proteomics 7, 3558-3568

43. Mann, K., and Mann, M. (2008) The chicken egg yolk plasma and granule proteomes. Proteomics 8, 178-191

44. Palmiter, R. D. (1972) Regulation of protein synthesis in chick oviduct. I. Independent regulation of ovalbumin, conalbumin, ovomucoid, and lysozyme induction. J. Biol. Chem. 247, 6450-6461

45. Nakano, T., and Graf, T. (1991) Goose-type lysozyme gene of the chicken: sequence, genomic organization and expression reveals major differences to chicken-type lysozyme gene. Biochim. Biophys. Acta 1090, 273-276

46. Slade, B., and Milne, J. (1977) Localization and synthesis of alpha-fetoprotein in the chicken. Cell Tissue Res. 180, 411-419

47. Anton, M. (2013) Egg yolk: structures, functionalities and processes. Journal of the science of food and agriculture 93, 2871-2880

48. Malik, G. K., Sinha, S. M., Saksena, P. N., Kapoor, A. K., Mehra, P., Bagchi, M., Agarwal, D. K., and Tuteja, N. (1981) Amniotic fluid proteins in relation to fetal maturity. Indian J. Pediatr. 48, 149-152

49. Retzek, H., Steyrer, E., Sanders, E. J., Nimpf, J., and Schneider, W. J. (1992) Molecular cloning and functional characterization of chicken cathepsin D, a key enzyme for yolk formation. DNA Cell Biol. 11, 661-672

50. Nelson, T. C., Groth, K. D., and Sotherland, P. R. (2010) Maternal investment and nutrient use affect phenotype of American alligator and domestic chicken hatchlings. Comp. Biochem. Physiol. A-Mol. Integr. Physiol. 157, 19-27

51. Maehashi, K., Ueda, M., Matano, M., Takeuchi, J., Uchino, M., Kashiwagi, Y., and Watanabe, T. (2014) Biochemical and functional characterization of transiently expressed in neural precursor (TENP) protein in emu egg white. J. Agric. Food Chem. 62, 5156-5162

52. Elo, H. A., and Korpela, J. (1984) The occurrence and production of avidin: a new conception of the high-affinity biotin-binding protein. Comp. Biochem. Physiol. B 78, 15-20

53. Matulova, M., Rajova, J., Vlasatikova, L., Volf, J., Stepanova, H., Havlickova, H., Sisak, F., and Rychlik, I. (2012) Characterization of chicken spleen transcriptome after infection with Salmonella enterica serovar Enteritidis. PloS one 7, e48101

54. Nakimbugwe, D., Massehalck, B., Atanassova, M., Zewdie-Bosuner, A., and Michiels, C. W. (2006) Comparison of bactericidal activity of six lysozymes at atmospheric pressure and under high hydrostatic pressure. Int. J. Food Microbiol. 108, 355-363

55. Sellier, N., Vidal, M. L., Baron, F., Michel, J., Gautron, J., Protais, M., Beaumont, C., Gautier, M., and Nys, Y. (2007) Estimations of repeatability and heritability of egg albumen antimicrobial activity and of lysozyme and ovotransferrin concentrations. Br. Poult. Sci. 48, 559-566

56. Herve-Grepinet, V., Rehault-Godbert, S., Labas, V., Magallon, T., Derache, C., Lavergne, M., Gautron, J., Lalmanach, A. C., and Nys, Y. (2010) Purification and Characterization of Avian

beta-Defensin 11, an Antimicrobial Peptide of the Hen Egg. Antimicrob. Agents Chemother. 54, 4401-4408

57. Tollin, M., Bergsson, G., Kai-Larsen, Y., Lengqvist, J., Sjovall, J., Griffiths, W., Skuladottir, G. V., Haraldsson, A., Jornvall, H., Gudmundsson, G. H., and Agerberth, B. (2005) Vernix caseosa as a multi-component defence system based on polypeptides, lipids and their interactions. Cell. Mol. Life Sci. 62, 2390-2399

58. Yoshio, H., Tollin, M., Gudmundsson, G. H., Lagercrantz, H., Jornvall, H., Marchini, G., and Agerberth, B. (2003) Antimicrobial polypeptides of human vernix caseosa and amniotic fluid: implications for newborn innate defense. Pediatr. Res. 53, 211-216

59. Cunningham, M., and Tang, J. (1976) Purification and properties of cathepsin D from porcine spleen. J. Biol. Chem. 251, 4528-4536

60. Castellano, A. C., Barteri, M., Bianconi, A., Bruni, F., Della Longa, S., and Paolinelli, C. (1996) Conformational changes involved in the switch from ovalbumin to S-ovalbumin. Z Naturforsch C. 51, 379-385

61. Yamasaki, M., Takahashi, N., and Hirose, M. (2003) Crystal structure of S-ovalbumin as a non-loop-inserted thermostabilized serpin form. J. Biol. Chem. 278, 35524-35530

62. Spencer, K., Muller, F., and Aitken, D. A. (1997) Biochemical markers of trisomy 21 in amniotic fluid. Prenat. Diagn. 17, 31-37

63. Liu, Y., Liu, Y., Du, C., Zhang, R., Feng, Z., and Zhang, J. (2016) Diagnostic value of amniotic fluid inflammatory biomarkers for subclinical chorioamnionitis. Int. J. Gynaecol. Obstet. 134, 160-164

64. Hsu, T. Y., Lin, H., Hung, H. N., Yang, K. D., Ou, C. Y., Tsai, C. C., Cheng, H. H., Chung, S. H., Cheng, B. H., Wong, Y. H., Chou, A. K., and Hsiao, C. C. (2016) Two-Dimensional

Differential Gel Electrophoresis to Identify Protein Biomarkers in Amniotic Fluid of Edwards Syndrome (Trisomy 18) Pregnancies. PloS one 11, e0145908

65. Wolanski, N. J., Renema, R. A., Robinson, F. E., Carney, V. L., and Fancher, B. I. (2007) Relationships among egg characteristics, chick measurements, and early growth traits in ten broiler breeder strains. Poult. Sci. 86, 1784-1792

66. Karcher, D. M., Jones, D. R., Abdo, Z., Zhao, Y., Shepherd, T. A., and Xin, H. (2015) Impact of commercial housing systems and nutrient and energy intake on laying hen performance and egg quality parameters. Poult. Sci. 94, 485-501

67. Bakst, M. R., Welch, G. R., Fetterer, R., and Miska, K. (2016) Impact of broiler egg storage on the relative expression of selected blastoderm genes associated with apoptosis, oxidative stress, and fatty acid metabolism. Poult. Sci. 95, 1411-1417

68. Loyau, T., Hennequet-Antier, C., Coustham, V., Berri, C., Leduc, M., Crochet, S., Sannier, M., Duclos, M. J., Mignon-Grasteau, S., Tesseraud, S., Brionne, A., Metayer-Coustard, S., Moroldo, M., Lecardonnel, J., Martin, P., Lagarrigue, S., Yahav, S., and Collin, A. (2016) Thermal manipulation of the chicken embryo triggers differential gene expression in response to a later heat challenge. BMC Genomics 17, 329

ACKNOWLEDGEMENTS

The authors wish to thank Joël Delaveau and Christophe Rat (INRA, UE PEAT 609, F-37380 Nouzilly, France) for providing fertilized eggs; Angelina Trotereau and Nathalie Winter (UMR UR1282 Infectiologie Animale et Santé Publique, F-37380 Nouzilly, France) who gave us access to L2 laboratories, which was required to perform antimicrobial assays using bacterial pathogenic strains. This work was supported by MUSE project (Région Centre-Val de Loire, 2014-00094512) and SAPhyR-11 project (Région Centre-Val de Loire, 2017-119983). The high-resolution mass

spectrometer (LTQ Velos Orbitrap) was financed (SMHART project n°3069) by the European Regional Development Fund (ERDF), the Conseil Régional du Centre, the French National Institute for Agricultural Research (INRA) and the French National Institute of Health and Medical Research (INSERM). The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication. The authors are very grateful to Région Centre Val de Loire for financing Mylène Da Silva's PhD.

DATA AVAILABILITY

Data are available *via* the PRIDE partner repository with the dataset identifiers PXD008046 and 10.6019/PXD008046.

AUTHORS CONTRIBUTION

MDS and SRG designed experiments, collected samples, analyzed and integrated data and wrote the manuscript. CD conducted the phylogenic analysis, realized the related figures and was in charge of writing the corresponding experimental procedures, results and discussion of the manuscript. AB performed Gene Ontology annotation of AF proteins and contributed to the critical analysis of proteomic data. MC and MDP adapted and conducted the antimicrobial experiments and MDP also carried out heparin-sepharose chromatography of AF and SDS-PAGE analyses of protein fractions. PM used his expertise to analyze phylogenetic results and contributed to the overall discussion of the data. LCS and VL were involved in the initial design of the proteomic experiments and conducted proteome profiling. NG contributed to the overall discussion of the data, and brought his specific expertise on egg white antimicrobial proteins and peptides, as well as on antimicrobial assays. YN participated in the conception of the project, scientific analyses of

the data and discussions. All authors read the manuscript and contributed to the critical revision of the paper.

FIGURE LEGENDS

Figure 1. Schematic representation of the extraembryonic structures during the chicken (A) and human embryonic development (B); chicken embryo at mid-incubation (11 days) and human embryo mid-gestation (21 weeks), respectively.

Figure 2. Diagram describing the experimental design. A., preparation of biological samples, B., proteomic analyses using Shotgun and GeLC-MS/MS analyses, C., antibacterial potential of egg samples and identification of antibacterial candidates. AF, amniotic fluid; Cb, Coomassie Brilliant Blue staining; ED, Day of incubation; EW, egg white; HBP, heparin-binding proteins.

Figure 3. SDS-PAGE analysis of the chicken amniotic fluid (AF) during incubation. A. 12.5% SDS-PAGE analysis of the AF from days 8-11 (ED8-11) to days 12-16 (ED12-16) of the incubation (2 µg and 10 µg, respectively) followed by Coomassie Brilliant Blue staining. B. 4-20% SDS-PAGE profile of AF at ED11 for GeLC-MS/MS analysis. Horizontal lanes and numbers indicate the position of gel slices (1-20) prepared for in-gel digestion by trypsin. Molecular masses are expressed in kDa.

Figure 4. Quantitative distribution of the 91 proteins identified in the chicken amniotic fluid at ED11, for shotgun and GeLC-MS/MS mass spectrometry analyses. Proteins are represented according to log10 of their Exponentially Modified Protein Abundance Index (emPAI) and log10 of the percentage of their protein sequence coverage. The most abundant proteins identified after compilation of data resulting from both approaches are illustrated in a dotted square. They were further analyzed for functional annotation as shown in Table 1.

Figure 5. Functions of the proteins identified in the chicken amniotic fluid at 11th day of incubation (ED11). Using Gene Ontology terms, as described in Experimental Procedures, 80 genes could be assigned to five major functions whereas the biological function of 10 proteins are still not documented in this database (this latter group includes ovalbumin (OVAL) and ovomucoid (SPINK7), two proteins of high abundance (Table 1, Supplementary data S3)).

Figure 6. Venn diagrams representing the specific and overlapping proteins identified in the human (during the whole gestation) and chicken (11th day of incubation) amniotic fluid (AF). Human AF data were obtained from the work of Cho *et al.* (11), which combines results from nine publications (Supplementary data 3, sheet #3).

Figure 7. **Antibacterial activity of the chicken amniotic fluid (AF) before and after the egg white transfer into the amniotic sac at ED11 (A) and ED16 (B)**. AF samples were analyzed by Coomassie Brilliant Blue staining (Cb) and their antibacterial activities were assessed against *Listeria monocytogenes* (*L.m.)* and *Salmonella enterica* Enteritidis (*S.*E.). Flow-through and eluted fractions were obtained after affinity chromatography on heparin-sepharose as described in Experimental Procedures. Black arrows indicate inhibition zones corresponding to an impaired bacterial growth (due to bacteriostatic or bactericidal activities of AF proteins or peptides). Grey letters A to H correspond to bands that were cut from the Cb gel, prior analysis by mass spectrometry for protein identification (Table 2).

Figure 8. **Antibacterial activity of the heparin-bound proteins of egg white at ED0 (start of incubation), at ED8 (8th day of incubation) and ED14 (14th day of incubation).** Proteins were analyzed by Coomassie Brilliant Blue staining (Cb) and their antibacterial activities were assessed against *Listeria monocytogenes* (*L.m.)* and *Salmonella enterica* Enteritidis (*S.*E.), as described in Experimental Procedures. Black arrows indicate inhibition zones corresponding to an impaired bacterial growth (due to bacteriostatic or bactericidal activities of egg white proteins or peptides).

Figure 9. Evolutionary scenario of the chicken amniotic fluid (AF) proteins. (A) Evolutionary frieze of 76 of the 91 proteins identified in the AF on the 11th day of incubation (ED11), and of 14 of the 29 potential antibacterial candidates identified in the AF at ED16 (in grey). The 10 genes underlined in yellow are specific to birds and/or reptiles. (B) Life tree of the bird and/or reptile specific 10 genes (underlined in yellow) that display bird and/or reptile specificities. The presence of a gene is depicted by a solid symbol, while the loss of a gene is depicted by a hollow symbol and marked with "Ψ." The divergent time of lineages refer to Tian *et al.* (21) and phylogenetic analysis using Ensembl (http://www.ensembl.org/index.html). For SPINK5 and SPINK7 genes, which are marked with a question mark, we cannot state on the distribution of these genes within amniotes, based on the data currently available in Ensembl databases and despite the synteny analysis of these genes. This issue is discussed in the Results section.

Table 1. Top-ten abundant proteins found in the chicken amniotic fluid at ED11 (11th day of incubation) using shotgun and GeLC-MS/MS analysis. Quantitative values are expressed as the percentage of the Exponentially Modified Protein Abundance Index (emPAI). Words in italics refer to hypothetical/unknown functions. Proteins that are also listed in the top 15 high abundance proteins from the human amniotic fluid are indicated with an asterisk (9, 11). Proteins that were not detected in one or the other approach are indicated by a dash in the emPAI column.

Gene symbol	Gene Description	GeLC-MS/MS % emPAI	Shotgun % emPAI	Biological Process	Cellular component
APOA1*	Apolipoprotein A-I	27.8	39.1	Lipid metabolism	Secreted
OVAL	Ovalbumin	7.8	2.0	*Nutrition*	Secreted
TF*	Ovotransferrin	7.8	17.6	Defense response; Iron metabolism	Secreted
AFP*	Alpha-fetoprotein	7.5	7.7	Blood pressure homeostasis; Development (folliculogenesis)	Secreted
LYZ	Lysozyme C	3.4	1.7	Defense response	Secreted
ALB*	Serum albumin	3.3	1.6	Blood pressure homeostasis; Fatty acid, DNA, ion metabolism	Secreted
GC*	Vitamin D-binding protein	3.2	0.4	Vitamin metabolism	Secreted
TTR	Transthyretin isoform 1	2.8	3.5	Vitamin and hormone metabolisms	Secreted
SPINK7	Ovomucoid isoform 1	-	2.4	*Regulation of proteolytic processes*	Secreted
APOC3	Apolipoprotein C-III	-	7.1	Lipid metabolism	Extracellular region

Table 2. Potential antibacterial candidates identified in the heparin-binding fraction of the chicken amniotic fluid at ED11 and ED16. Proteins contained in the SDS-PAGE bands corresponding to bacterial inhibition zones (*Listeria monocytogenes* and *Salmonella* Enteritidis, Fig. 7) were analyzed using mass spectrometry. Bands where spectral counts were the highest are indicated in bold.

Identified Proteins (29)	Accession Number	Gene symbol	Molecular Mass (kDa)	Corresponding SDS-PAGE bands (Fig.7)
Mutant cysteine-rich FGF receptor [*G. gallus*]	AAB39211.1	GLG1	122	**G**
Avidin [*G. gallus*]	CAC34569.1	AVD	17	B, **C**, D, E, F, G, H
OvoglobulinG2 type BB [*G. gallus*]	BAM13273.1	BPIFB2	47	B, **C**, D, E, F
Ovomacroglobulin, ovostatin [*G. gallus*]	CAA55385.1	OVST	164	**C**
Sodium/potassium-transporting ATPase subunit alpha-1, precursor [*G. gallus*]	NP_990852.1	ATP1A1	112	**A**
Peptidyl-prolyl cis-trans isomerase B precursor [*G. gallus*]	NP_990792.1	PPIB	19	**E**
Mucin-6, partial [*G. gallus*]	XP_015142236.1	MUC6	98	B, **C**
Metalloproteinase inhibitor 2 precursor [*G. gallus*]	NP_989629.1	TIMP2	22	**E**
Metalloproteinase inhibitor 3 precursor [*G. gallus*]	NP_990818.1	TIMP3	20	**E**
Gallinacin-11 precursor [*G. gallus*]	NP_001001779.1	AvBD11	12	**C**
Ovoinhibitor precursor [*G. gallus*]	NP_001025783.2	SPINK5	52	B, **C**
Ovalbumin-related protein Y [*G. gallus*]	NP_001026172.1	OVALY	44	B, C, **D**
Elongation factor 1-alpha 1 [*G. gallus*]	NP_001308445.1	EEF1A1	50	**G**
Vitelline membrane outer layer protein 1 precursor [*G. gallus*]	NP_001161233.1	VMO1	20	C, E, **F**, G
Astacin-like metalloendopeptidase precursor [*G. gallus*]	NP_001292019.1	ASTL	46	**D**
Lysozyme C [*G. gallus*]	P00698.1	LYZ	16	A, B, C, D, E, F, G, **H**
Ovalbumin [*G. gallus*]	P01012.2	OVAL	43	A, B, **C**, D, E, F, G, H
Ovotransferrin [*G. gallus*]	P02789.2	TF	78	B, **C**
Ovocleidin-17 [*G. gallus*]	Q9PRS8.2	OC-17	15	**G**
Ovomucin [*G. gallus*]	XP_003641415.1	MUC5B	234	B, **C**
Tumor necrosis factor receptor superfamily member 6B [*G. gallus*]	XP_004947191.1	TNFRSF6B	38	**D**
Beta-actin [*G. gallus*]	CAA25004.1	ACTB	46	**A**, G
Beta-microseminoprotein-like [N. notabilis]	*XP_010010315.1*	*LOC104403769*	*15*	*C, **E**, F, G*
Ovalbumin-related protein X [*G. gallus*]	XP_015137660.1	OVALX	45	B, C, **D**, E, F, G
Clusterin [*G. gallus*]	XP_015140573.1	CLU	62	**C**
Midkine [*G. gallus*]	XP_015142525.1	MDK	16	**G**
Alpha-2-macroglobulin-like protein 1 [*G. gallus*]	XP_015148230.1	A2ML1	160	B, **C**
ATP synthase subunit alpha, mitochondrial [*G. gallus*]	NP_989617.1	ATP5A1Z	56	**A**
Deleted in malignant brain tumors 1 protein-like [*G. gallus*]	XP_015156102.1	DMBT1L3	100	B, **C**

In italics: LOC104103769 from *N. Notabilis* has a *G. Gallus* homologue (LOC101750704), which has been erroneously removed from Pubmed database following the release of Gallus_gallus-5.0 assembly in 2015. This chicken homologue has been unambiguously identified in egg white (25) and shares 88.7% sequence identity with LOC104403769 from *N. notabilis.*

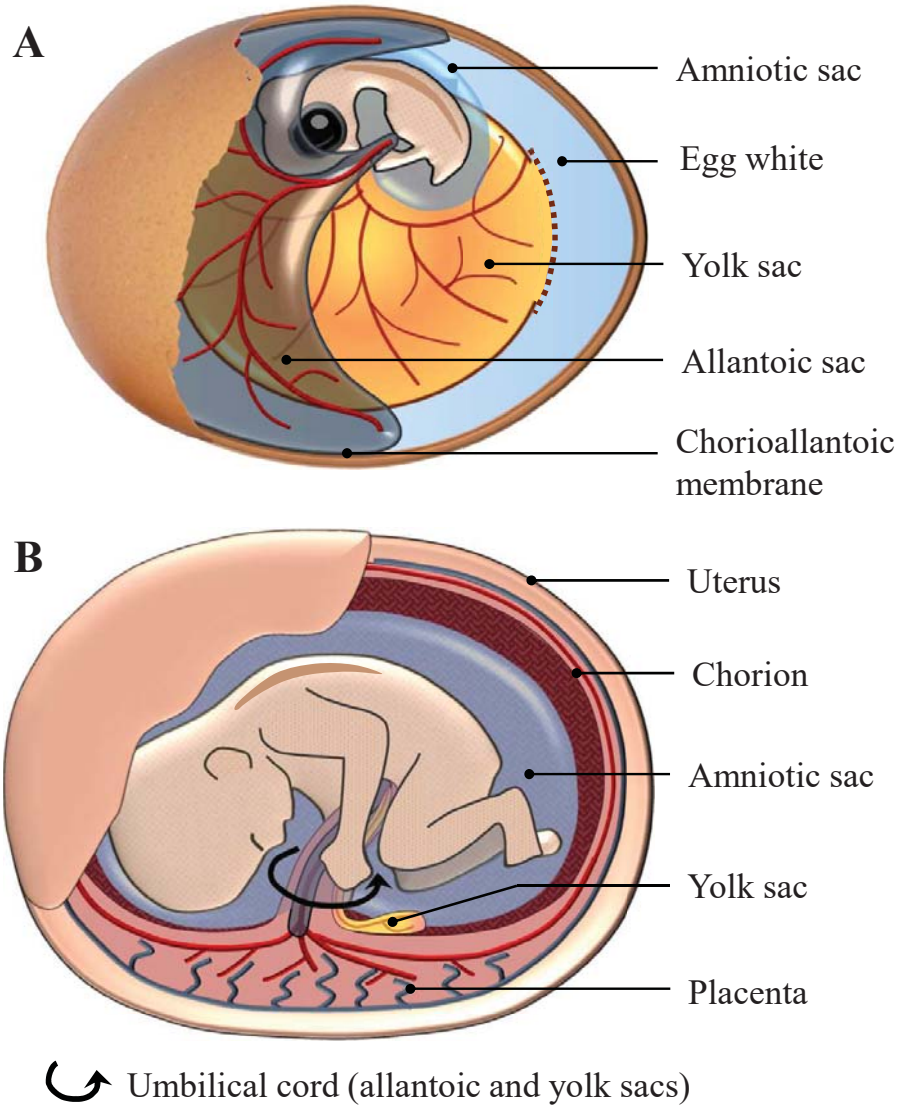

Fig. 1

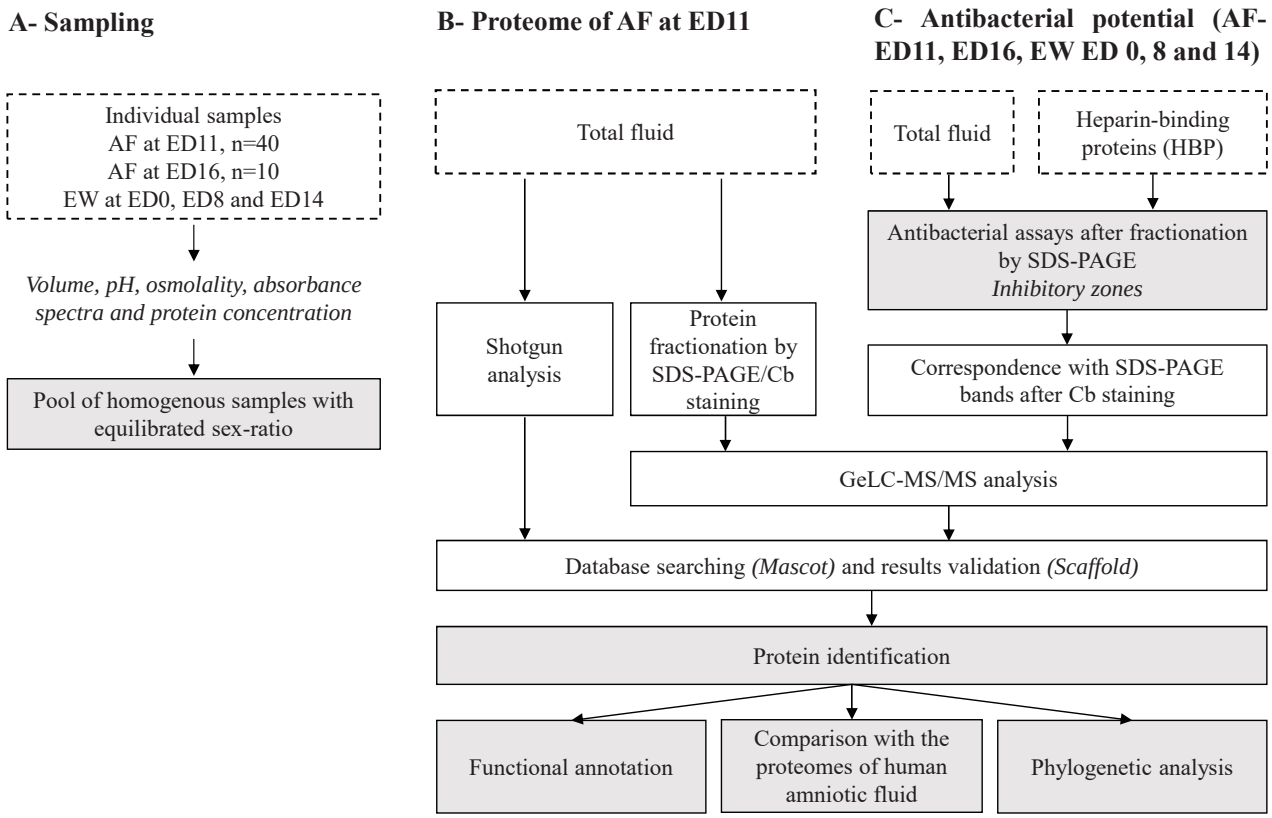

Fig. 2

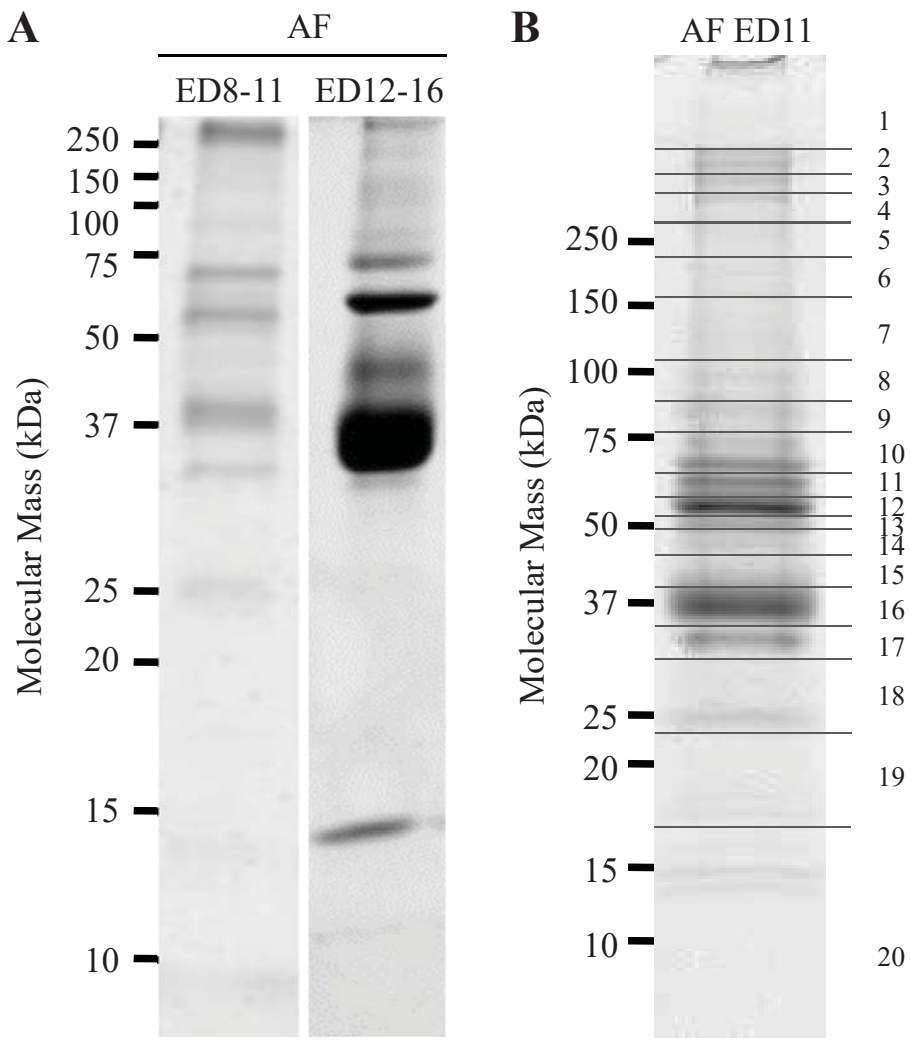

Fig. 3

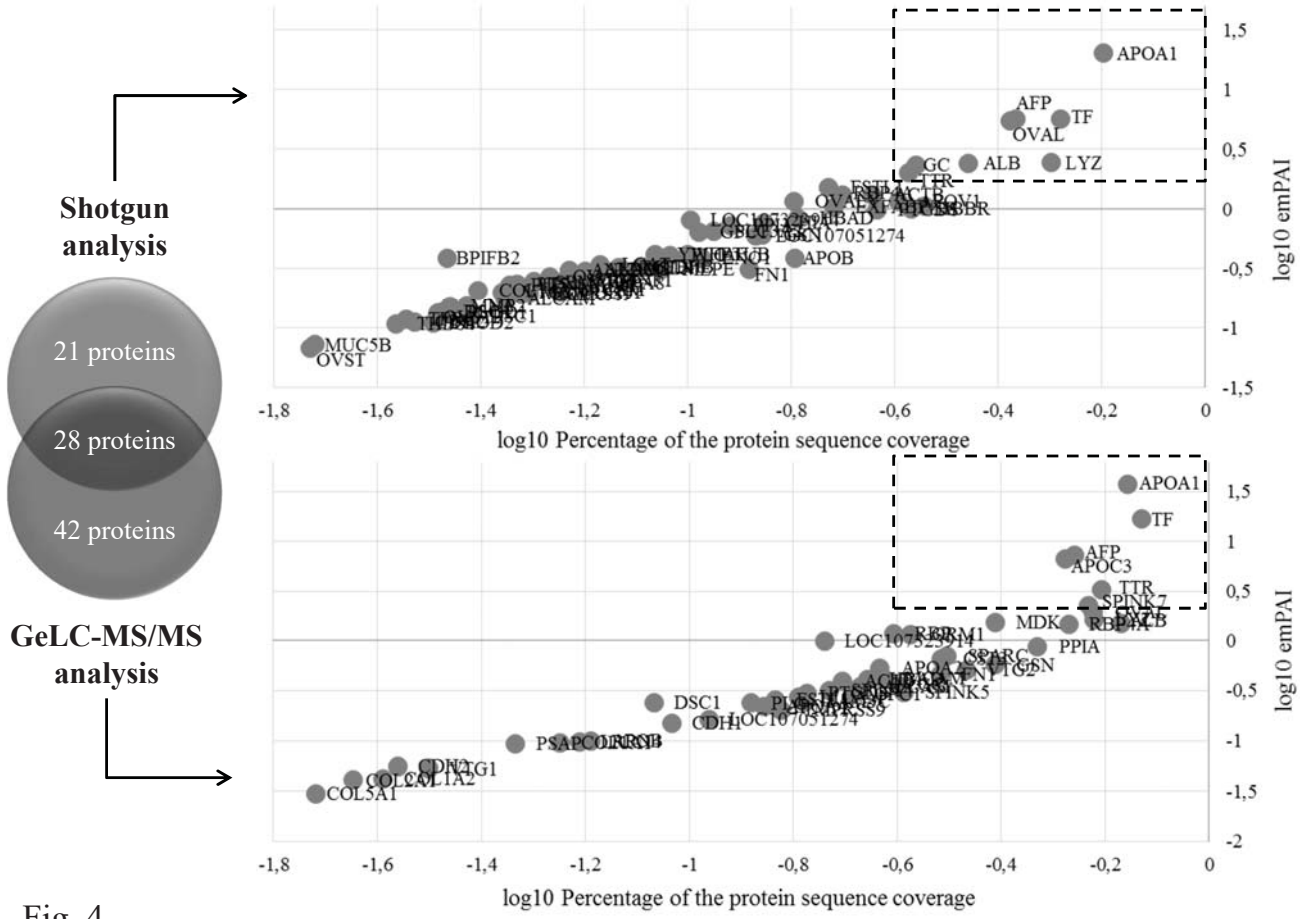

Fig. 4

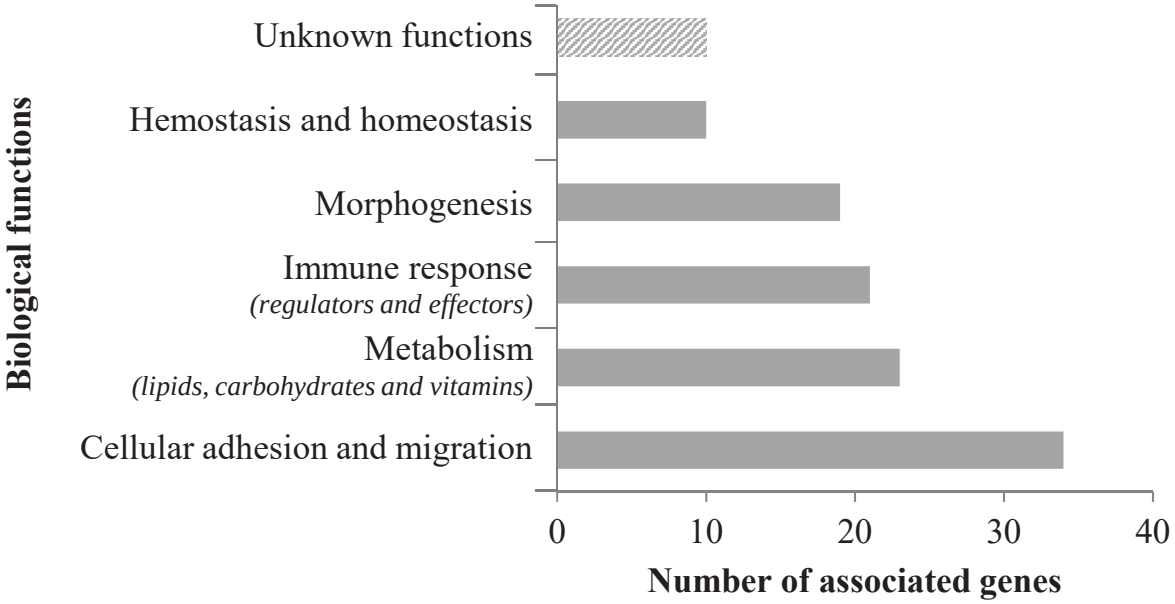

Fig. 5

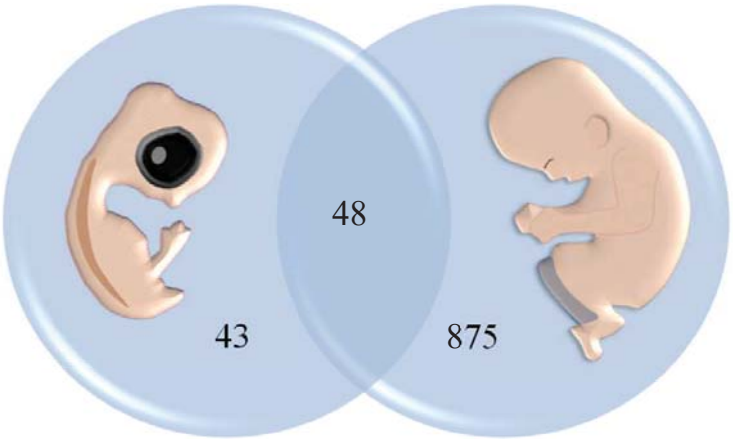

Fig. 6

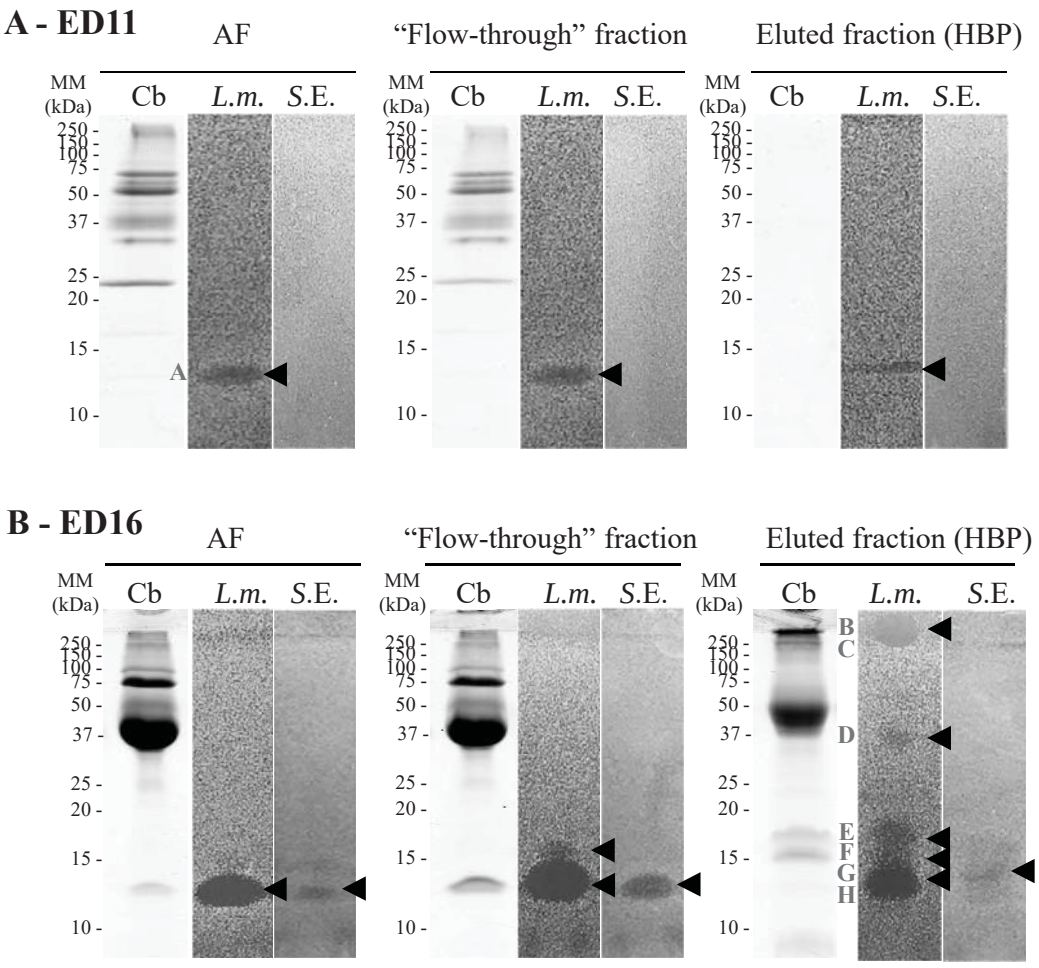

Fig. 7

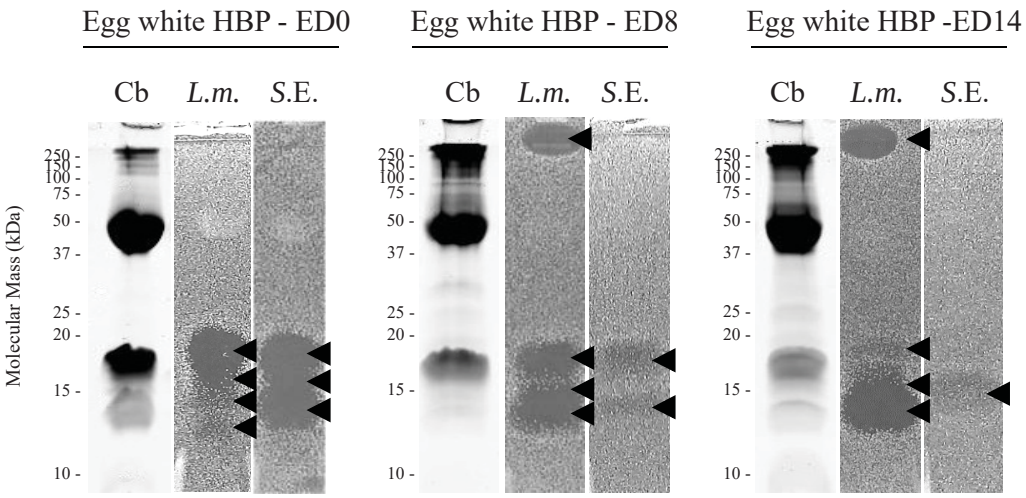

Fig. 8

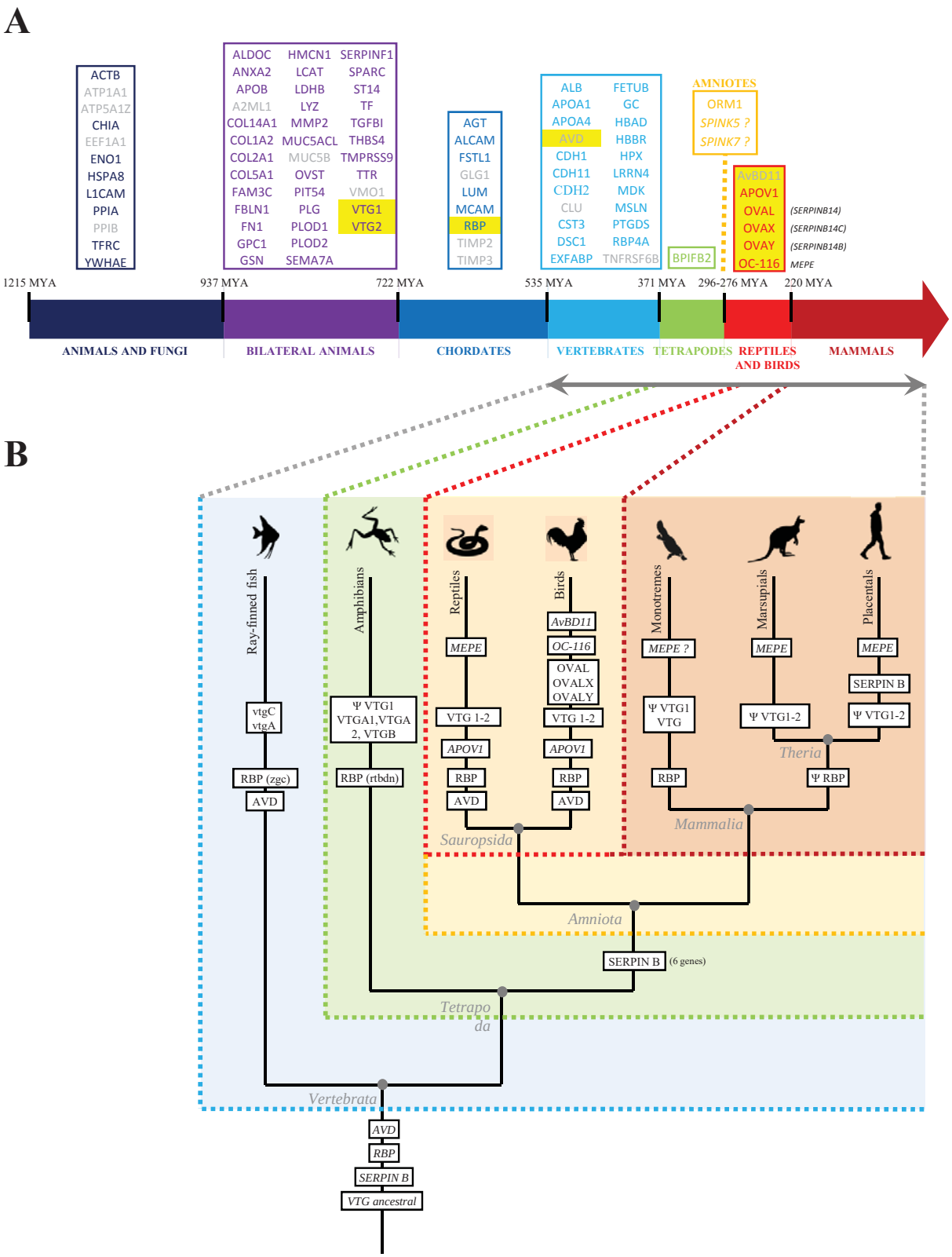

Fig. 9