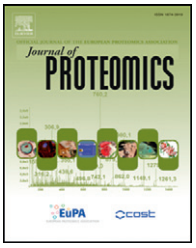
JOURNAL OF PROTEOMICS XX (2012) XXX–XXX

Available online at www.sciencedirect.com

SciVerse ScienceDirect

www.elsevier.com/locate/jprot



Rabbit zona pellucida composition: A molecular, proteomic and phylogenetic approach

I. Stetson[a, b, 1], M.J. Izquierdo-Rico[a, 1], C. Moros[a], P. Chevret[c], P.L. Lorenzo[d], J. Ballesta[a], P.G. Rebollar[f], R. Gutiérrez-Gallego[e], M. Avilés[a,*, 1]

[a]Department of Cell Biology and Histology, Faculty of Medicine and Faculty of Nursing, University of Murcia, 30100 Murcia, Spain
[b]Department of Biology, National University of Misiones (UNaM) of Posadas Misiones, Argentina
[c]Laboratoire de Biométrie et Biologie Evolutive, CNRS, UMR5558, Université de Lyon, Université Claude Bernard Lyon 1, Bat Mendel, 69622 Villeurbanne, France
[d]Department of Physiology, Complutense University, Madrid, Spain
[e]Bio-analysis Group, Neuroscience Research Programme, IMIM-Hospital del Mar, Department of Experimental and Health Sciences, University Pompeu Fabra (UPF), Barcelona Biomedical Research Park (PRBB), 08003 Barcelona, Spain
[f]Animal Production Department, Polytechnic University, Madrid, Spain

ARTICLE INFO

Article history:
Received 12 March 2012
Accepted 16 July 2012

Keywords:
Zona pellucida
Rabbit
Oocyte
Fertilization
Sperm–egg interaction

ABSTRACT

The zona pellucida (ZP) participates in sperm–egg interactions during the first steps of fertilization. Recent studies have shown that the ZP matrix of oocytes in several species is composed of four glycoproteins, designated as ZP1, ZP2, ZP3 and ZP4, rather than the three described in mouse, pig and cow. In this study, investigations were carried out to unveil a fourth glycoprotein in the rabbit (*Oryctolagus cuniculus*) ZP. Using total RNA isolated from rabbit ovaries, the complementary deoxyribonucleic acid (cDNA) encoding rabbit ZP1 was amplified by reverse transcribed polymerase chain reaction (RT-PCR). The ZP1 cDNA contains an open reading frame of 1825 nucleotides encoding a polypeptide of 608 amino acid residues. The deduced amino acid sequence of rabbit ZP1 showed high identity with other species: 70% identity with human and horse ZP1, and 67% identity with mouse and rat ZP1. At the proteomic level, peptides corresponding to the four proteins were detected by mass spectrometry. In addition, a molecular phylogenetic analysis of ZP1 showed that pseudogenization of this gene has occurred at least four times during the evolution of mammals. The data presented in this manuscript provide evidence, for the first time, that the rabbit ZP is composed of four glycoproteins.

1. Introduction

During the in vivo fertilization process, sperm interact with the extracellular coats that surround the mammalian oocyte. These coats are the cumulus oophorus and the zona pellucida (ZP). The ZP has been related with species-specific gamete recognition, the sperm acrosome reaction, the control of polyspermy and protection of the oviductal embryo [1–3].

* *Corresponding author at:* Department of Cell Biology and Histology, Faculty of Medicine and Faculty of Nursing, University of Murcia, Espinardo, 30100, Murcia, Spain. Tel.: +34 868884385.
E-mail address: maviles@um.es (M. Avilés).
[1] Both authors contributed equally to this work.

1874-3919/$ – see front matter © 2012 Published by Elsevier B.V.
http://dx.doi.org/10.1016/j.jprot.2012.07.027

It has been shown that the ZP genes that encode ZP proteins can be classified into six subfamilies: ZPA/ZP2, ZPB/ZP4, ZPC/ZP3, ZP1, ZPAX and ZPD [4]. However, not all these genes are present in all species. The ZP or equivalent extracellular coat in vertebrates is formed by several proteins ranging from 3 to 6 [5,4]. Phylogenetic studies and the finding of different pseudogenes suggest that the evolution of ZP genes is mainly produced by duplications and death of genes [6,4].

Early studies in mouse demonstrated that the ZP is formed of only three glycoproteins: ZP1, ZP2 and ZP3 [7]. Later, the presence of three glycoproteins was demonstrated in other species like pig [8] and cow [9]. However, in these species the proteins are ZP2, ZP3 and ZP4.

Moreover, analysis of the complete genome in different species suggests the existence of additional genes coding for ZP proteins and shows that mammalian ZP could be formed of four proteins. Some studies have reported the existence of four proteins in the ZP of species like human [10,11], rat [12] and hamster [13,14] and phylogenetic analysis has detected four genes in other species like chimpanzee and macaque [15].

Therefore, mammals could be classified into three categories according to their ZP composition. 1) Species with a ZP formed by ZP1, ZP2 and ZP3 (to date, includes only the mouse); 2) species showing three proteins, where ZP1 is not present (e.g. cow, dog and pig); and 3) species with four proteins (ZP1, ZP2, ZP3 and ZP4) as, for example, human, rat and hamster.

The confusing results obtained in different studies on the ZP composition in some species is mainly due to the scarce amount of ZP available and, especially, to the heterogeneous glycosylation of the ZP proteins, resulting in broad, partially overlapping, bands in SDS-PAGE [16,14]. These facts make the purification of these proteins very difficult and, subsequently, accurate analysis is also difficult. Moreover, the general acceptation of mouse zona pellucida model (with 3 proteins) made it difficult to take into consideration the analysis of the ZP composition in other species.

The development of mass spectrometric techniques has provided an important opportunity to identify the different proteins and glycoforms present in a complex mixture. Thus, proteomic analysis clarified ZP protein composition in human, rat, and hamster [11–13] and provided detailed information on the carbohydrate composition of the ZP proteins in some species [17–19].

In rabbit, characterization of the ZP by SDS-PAGE suggested the presence of three glycoproteins, ZP2 [20,21], ZP3 [22] and ZP4 [23,24], which migrate as one band with an apparent molecular mass of 85–95 kDa [23]. These proteins (ZP2 [20], ZP3 [22] and ZP4 [25]) were detected by molecular biology approaches and are deposited in the GenBank database (GenBank accession numbers: L12167 (ZP2), NM_001195720.1 (ZP3), NM_001082295 (ZP4)).

In addition, ZP1 cDNA sequence has been deposited after in silico analysis in GenBank with the accession number: XM_002709016.

The aim of this study was to demonstrate the presence of ZP1 mRNA in rabbit ovaries and ZP1 protein in the ovary and ZP from isolated oocytes.

2. Material and methods

2.1. Collection of rabbit (*Oryctolagus cuniculus*) ovaries

12 week-old New Zealand California white rabbits were used to obtain ovarian RNA (n=3). The rabbits were injected with 25 IU of pregnant mare serum gonadotropin (PMSG) to stimulate folliculogenesis and were sacrificed 48 h later by overdose of CO_2. Ovaries were obtained and frozen in liquid nitrogen and kept at −80 °C until use.

In addition, 20 female 8 week-old rabbits, obtained from a slaughterhouse (Conejos Susi, S.L., Alicante, Spain), were used for proteomic analysis.

2.2. Collection of rabbit (*O. cuniculus*) oocytes and obtention of the zona pellucida

Ovaries were obtained from 18 week-old animals (n=12) killed with an overdose of CO_2 and subjected to laparotomy. Cumulus–oocyte complexes (COCs) were obtained by aspiration with a 2 mL syringe and a 25 gauge needle from ovarian follicles, <1 mm in size, as previously described [26]. The COCs were placed in PBS 4-well dishes and the cumulus cells were removed by gentle pipetting using 2 mM hyaluronidase.

The oocyte ZPs were obtained after vigorous pipetting of each oocyte by using a narrow-bore micropipette, following four washes in PBS to eliminate the oocyte debris.

2.3. Purification of rabbit ovarian RNA, obtention of cDNA and amplification of the complete open reading frame of rabbit ZP1 gene

Total RNA was isolated using RNAqueous® kit (Ambion, Austin, Texas, USA) according to the manufacturer's instructions. The first-strand cDNA was synthesized from total RNA with the SuperScript First-Strand Synthesis System kit for RT-PCR (Invitrogen-Life Technologies, Carlsbad, CA, USA), according to the manufacturer's protocol.

Rabbit ZP1 was amplified using polymerase chain reaction (PCR) by means of specific primers (Table 1) designed according to the predicted cDNA sequence obtained from

Table 1 – Primers used in amplification of rabbit ZP1.

Primers	Forward (position in sequence)	Reverse	Amplified region (pb)
Fw1 and Rv1	at gac tgg ggg tcg cct ggt (1)	ctc ctg ggg cag atg gct acc tac	651
Fw2 and Rv2	ggt gga acg ctg gga agt gg (594)	gaa gat gga cgc ctg gat gg	522
Fw3 and Rv3	tct tca atg cca gcg act tc (1079)	ctca ggc cca caa aga cac ca	746
Fw4 and Rv4	aga ctt gct cat cta cgt gt (1571)	tta ttg agc ctg gtc ggt ga	314

Please cite this article as: Stetson I., et al, Rabbit zona pellucida composition: A molecular, proteomic and phylogenetic approach, J Prot (2012), http://dx.doi.org/10.1016/j.jprot.2012.07.027

genomic sequences in the ENSEMBL server (ensemble accession number: ENSOCUG00000015673).

PCR amplifications were performed using 3 μL of target cDNA, 0.5 μg of each primer, 200 μM of each dNTP and 1 IU of Advantage GC 2 Polymerase (Clontech Laboratories, CA, USA). PCR was carried out using an initial denaturation cycle of 2 min, and then 30 cycles of 1 min at 95 °C, 1 min at annealing temperature (depending on the primers) and then 1 min at 72 °C. The final extension time was 10 min at 72 °C. PCR products were analyzed by electrophoresis on 1.5% agarose gels. Four microliters of the PCR reaction mixture was mixed with loading buffer and separated for 90 min at 100 V before visualizing under UV light using ethidium bromide.

Amplicons were carefully excised from the agarose gels and purified with the QIAquick Gel Extraction Kit Protocol (Quiagen, Hilden, Germany) according to the manufacturer's protocol. After that, the amplicons were automatically sequenced. The sequences were analyzed to determine the degree of homology with other known sequences using the BLAST program (Basic Local Alignment Search Tool) (http://www.ncbi.nlm.nih.gov/blast/) [27]. Direct comparison between two sequences was made with the ALIGN program, and the multiple alignment of the ZP1 sequences of different species with the rabbit ZP1 sequence was carried out using Clustal W (http://www.ebi.ac.uk/clustalw/).

The amino acid sequences were analyzed with the following software packages: "SignalP" [28] to predict putative signal sequence and cleavage sites, and "NetOGlyc" [29] and "NetNglyc" [30] to predict potential O-linked and N-linked glycosylation sites, respectively.

In addition, amplicons corresponding to ZP2, ZP3 and ZP4 were amplified in the same conditions as ZP1, while the primers were designed based on cDNA sequences obtained from the GenBank database (ZP2: XM_002711834, ZP3: NM_001195720, ZP4: NM_001082295) (Supplementary material 1).

2.4. *Phylogenetic analysis of ZP1*

Sequences of ZP1 for different mammals were retrieved from GenBank (when mRNA sequences were available) and from ENSEMBL gene predictions (Supplementary material 2). All these predictions were checked manually to detect annotation errors especially those that are close to splicing sites. Similarity searches were performed using BLAST and BLAT against assembled genomes (http://ensembl.org), and TRACE data (deposited in the trace archive of GenBank) followed by manual compilation of data to predict further genes or exons missing from the ENSEMBL predictions. We also checked that the new sequences corresponded to a syntenic region of the corresponding chromosome or contig. Sequences were aligned using Muscle in SeaView [31] and the alignment was refined visually. Only the exonic portions were kept for the phylogenetic analysis. Phylogenetic trees were reconstructed using maximum likelihood with PhyML [32] and the robustness of the nodes was estimated with bootstrap percentage (n=1000). The appropriate model of evolution was determined using corrected Akaike information criterion (AICc) and Modeltest software [33].

3. Proteomic analysis

3.1. *Solubilization of rabbit ZP, SDS-PAGE and silver staining*

The rabbit ovaries (three different experiments: n=14; n=22 and n=37) were trimmed using small scissors and dissected to remove fat and connective tissue. Solubilized ZP was obtained according to the protocol previously described by our group [13,14].

In addition, solubilized ZP was also obtained from oocytes. For that, oocyte ZP (n=200) was solubilized at 65 °C for 30 min in PBS buffer; the sample was then centrifuged and the supernatant was recovered.

Partially purified ZP (ovary) and isolated ZP (oocytes) were dissolved in sample buffer in reducing conditions (5%-mercaptoethanol). After boiling for 5 min, samples were separated by 12% SDS-PAGE. In brief, 4% stacking and 12% separating gels were used with 25 mM Tris/0.2 M glycine buffer, pH 8.6, containing 0.1% SDS for 1.5 h at 150 V and room temperature. After electrophoresis, the gel was fixed in a 5% acetic acid/50% methanol solution for 30 min. The gel was then washed in a 50% methanol solution for 15 min followed by milliQ water for 15 min. Next, the gel was incubated in a 0.01% sodium thiosulfate solution for 1 min and, after two washes with milliQ water, the gel was incubated with 0.1% silver nitrate solution for 20 min at 4 °C. Finally, the gel was washed twice with milliQ water and incubated with 2% sodium carbonate solution with 250 μL of 35% formaldehyde solution, to visualize the protein bands. The proteins were immobilized by incubating for 5 min in a 5% acetic acid solution.

3.2. *HPLC–MS analysis*

HPLC–MS/MS analysis was used to identify the rabbit ZP proteins. The analysis was carried out on an HPLC–MS system consisting of an Agilent 1100 Series HPLC (Agilent Technologies, Santa Clara, CA) equipped with a μ-wellplate autosampler and a capillary pump, and connected to an Agilent Ion-Trap XCT Plus mass spectrometer (Agilent Technologies, Santa Clara, CA) equipped with an electrospray interface (ESI). Details of the mass LC–MS conditions are described below.

Samples were separated using SDS-PAGE and the bands were cut and washed twice with milliQ distilled water and then twice with 25 mM ammonium bicarbonate buffer pH 8.5 in 50% acetonitrile for 30 min at 37 °C. The bands were dried by vacuum evaporator, and then incubated with 50 μL of 25 mM ammonium bicarbonate buffer pH 8.5 with 50 mM tris (2-carboxyethyl) phosphine at 60 °C for 10 min. After removing the supernatant, samples were alkylated by adding 25 mM ammonium bicarbonate buffer pH 8.5 containing 100 mM iodoacetamide and allowed to stand for 1 h at room temperature in the dark. The supernatant was removed and the bands were washed with 25 mM ammonium bicarbonate buffer pH 8.5 and then with 25 mM ammonium bicarbonate buffer pH 8.5 in 50% acetonitrile for 15 min at 37 °C each time. After washing, the bands were dried using a vacuum

Please cite this article as: Stetson I., et al, Rabbit zona pellucida composition: A molecular, proteomic and phylogenetic approach, J Prot (2012), http://dx.doi.org/10.1016/j.jprot.2012.07.027

evaporator, and then incubated with 25 mM ammonium bicarbonate buffer pH 8.5 containing 0.3 μg of proteomic grade trypsin (Sigma-Aldrich) for 45 min at 4 °C and finally submitted to digestion for 16 h at 37 °C. The supernatant was collected in a new tube, and the bands were washed with 50 μL of a solution containing 50% acetonitrile and 0.5% TFA and then with 50 μL of acetonitrile for 30 min at 37 °C each time. These washes enhanced the extraction of digested fragments from the gel bands and, afterward, both supernatants were combined and dried using a vacuum evaporator.

In the case of soluble samples, these were diluted up to a final volume of 100 μL of 25 mM ammonium bicarbonate buffer pH 8.5 and then incubated with 50 mM tris (2-carboxyethyl) phosphine at 60 °C for 10 min. After that, samples were alkylated by adding 100 mM iodoacetamide and then left to stand for 1 h at room temperature in the dark. Finally, 0.3 μg of proteomic grade trypsin (Sigma-Aldrich) was added to each sample for a digestion time of 16 h at 37 °C. After this incubation, the tryptic digestion was stopped with 0.5% TFA and the samples were dried using a vacuum evaporator.

The tryptic digestions were separated and analyzed by HPLC–MS. Dried samples (both from solution digestion and in-gel digestion) were resuspended in 10 μL of buffer A, consisting of water/acetonitrile/formic acid (94.9:5:0.1). Samples were injected into a Zorbax SB-C18 HPLC column (5 μm, 150×0.5 mm, Agilent Technologies, Santa Clara, CA), thermostatted at 40 °C, at a flow rate of 10 μL/min. After injection, the column was washed with buffer A and the digested peptides were eluted using a linear gradient of 0–80% B (buffer B: water/acetonitrile/formic acid, 10:89.9:0.1) for 120 min.

The mass spectrometer was operated in the positive mode with a capillary spray voltage of 3500 V and a scan speed of 8100 (m/z)/s from 300 to 2200 m/z. The nebulizer gas (He) pressure was set at 15 psi, whereas the drying gas was set at a flow rate of 5 L/min at a temperature of 350 °C. MS/MS data were collected in an automated data-dependent mode. The most intense ions were sequentially fragmented using collision-induced dissociation (CID) with an isolation width of 2 Da and a relative collision energy of 35%. Data processing was performed with the Data Analysis program for LC/MSD Trap Version 3.2 (Bruker Daltonik, GmbH, Germany) and Spectrum Mill MS Proteomics Workbench (Agilent Technologies, Santa Clara, CA).

Data processing was performed with Data Analysis program for LC/MSD Trap Version 3.3 (Bruker Daltonik, GmbH, Germany) and Spectrum Mill MS Proteomics Workbench (Rev A.03.02.060B, Agilent Technologies, Santa Clara, CA, USA). Briefly, raw data were extracted under default conditions as follows: unmodified or carbamidomethylated cysteines; sequence tag length > 1; $[MH]^+$ 50–7000 m/z; maximum charge +7; minimum signal-to-noise (S/N) ratio 25; finding ^{12}C signals. The MS/MS search against mammalian sequences in the NCBInr database was performed with the following criteria: identity search mode; tryptic digestion with 2 maximum missed cleavages; carbamidomethylated cysteines; peptide charge +1, +2, +3; monoisotopic masses; peptide precursor mass tolerance 2.5 Da; product ion mass tolerance 0.7 amu; ESI ion trap instrument; minimum matched peak intensity 50%; oxidized methionine, N-terminal glutamine conversion to pyroglutamic acid and STY phosphorylation as variable modifications. Two or more validated peptides were considered to demonstrate the existence of the protein. Peptides were considered valid with a score threshold of 5, and a percentage-scored peak intensity of 60%. All database matches above the threshold score of 3 were reported and used for discussion purposes.

4. Results and discussion

4.1. *Analysis of rabbit ZP1 cDNA and amino acid sequences*

The open reading frame (ORF) of rabbit ZP1 was completely amplified and characterized for the first time in this work.

The ORFs corresponding to rabbit ZP2, ZP3 and ZP4 have been characterized in previous studies [20,22,25]. Moreover, amplifications of a fragment corresponding with each gene (ZP2, ZP3 and ZP4) were made to confirm the results (Fig. 1) demonstrating the presence of the four transcripts in the rabbit ovary.

Full-length rabbit ZP1 cDNA was obtained from the total RNA prepared from rabbit ovaries and the sequence was submitted to GenBank with the accession number HQ702467. The amplified sequence of ZP1 contains a single ORF of 1884 nucleotides (Fig. 2) and 100% similarity with the predicted rabbit ZP1 (XM_002709016). The ATG initiation codon which was predicted with the Pedersen and Nielsen algorithm, was found to be associated with vertebrate initiator codons. This sequence contains a stop codon (TAA) in positions 1882–1884.

The ORF of ZP1 encodes a polypeptide 627 amino acids long (Fig. 2) with a theoretical molecular weight of 68.7 kDa. A signal peptide of 20 amino acids with a cleavage between Gly20 and Gln21 was predicted with the Bendtsen et al. algorithm [34].

The ZP protein possesses the archetypal 'ZP domain', a signature domain comprised of 272 amino acid residues (^{279}Gln-Gly^{550}) rich in Cys residues (ten). Upstream of the ZP

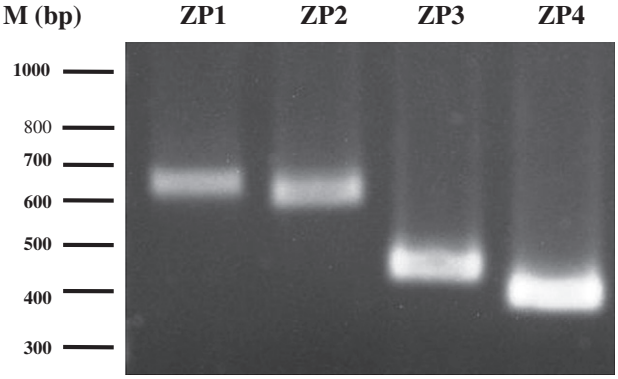


Fig. 1 – Analysis of ZP1, ZP2, ZP3 and ZP4 gene expression in rabbit ovary by RT-PCR. The primers used are in Table 1 (Fw1 and Rv1 for ZP1, ZP2, ZP3 and ZP4).

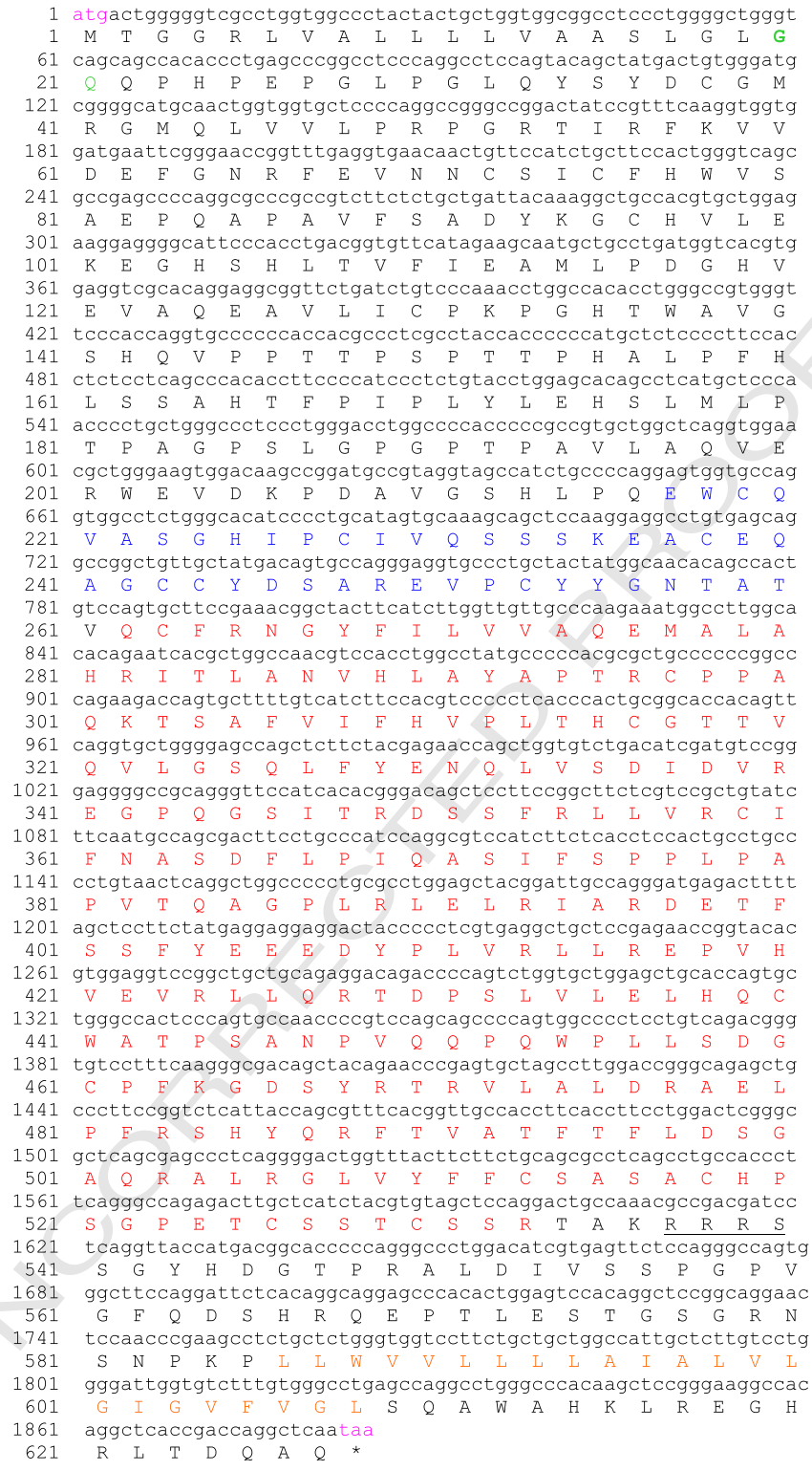


Fig. 2 – Nucleotide and deduced amino acid sequence of rabbit ZP1. The initial and final codons are in pink. The signal peptidase cleavage site is between Gly20 and Gln21 and is marked in green. The underlined amino acids indicate the C-terminal cleavage site. The zona domain is shown in red. The trefoil domain is shown in blue. The consensus furin cleavage-site is underlined. The transmembrane domain is in orange. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Please cite this article as: Stetson I., et al, Rabbit zona pellucida composition: A molecular, proteomic and phylogenetic approach, J Prot (2012), http://dx.doi.org/10.1016/j.jprot.2012.07.027

domain, a trefoil domain contains 45 residues (^{217}Glu-Thr^{260}). This domain is characteristic of ZP1 and ZP4 and is a region rich in Cys amino acids (six).

The sequence showed high hydrophobicity in the N-terminal (signal peptide) and C-terminal regions, the latter corresponding to the transmembrane domain (TMD) between Leu586 and Leu608, which is followed by a cytoplasmic tail. A basic amino acid domain (^{537}Arg-Arg-Arg-Ser^{540}) upstream of the TMD may serve as a consensus furin cleavage site [35–37].

Taking into account the presence of a signal peptide and that the putative cutting site is Arg538, as observed for mouse and rat ZP1 [36,37], the molecular mass of the putative mature protein is estimated to be 57.3 kDa.

The amino acid sequence showed high similarity with the ZP1 of other mammals (Fig. 3). The protein sequence of rabbit ZP1 is 70% identical to human and horse ZP1, and 67% identical to mouse ZP1. The similarity with mammalian ZP1, the presence of the same domains and the conservation of the

```
HUMAN       ------------------------------------------------------------
CHIMPANZEE  ------------------------------------------------------------
HORSE       MGSSAFCPRLRGEISAACWSLSSPVALILLQALGAGAVSLQAPKAPPEPEPSPSGGRWGL 60
RABBIT      ------------------------------------------------------------
MOUSE       ------------------------------------------------------------
MACAQUE     ------------------------------------------------------------

HUMAN       --------MAGGSATTWGYPVALLLLVAT-LGLGRWLQPDPGLPGLRHSYDCGIKGMQLL 51
CHIMPANZEE  --------MAGGSATTWGYRVALLLLVAT-LGLGRRLQPNPGLPGLRHSYDCGIKGMQLL 51
HORSE       LWGLWVYLVAGTSAMVWGLSVA-LLLVAT-LGLGQQPRLEPGLLGLRHTYDCGMKGLQLL 118
RABBIT      --------------MTGGRLVALLLLVAASLGLGQQPHPEPGLPGLQYSYDCGMRGMQLV 46
MOUSE       --------------MAWGCFVVLLLLAAAPLRLGQRLHLEP---GFEYSYDCGVRGMQLL 43
MACAQUE     ----------------MRLLQCVLLCVSLSLVLSGQHKPET--PGYSSVLHCGLWSFQFA 42
                             ** .:  * *.   : :.   *     .**: .:*:

HUMAN       VFPRPGQTLR--FKVVDEFGNRFDVNNCSICYHWVTSRPQEPAVFSADYRGCHVLEKDGR 109
CHIMPANZEE  VFPRPGQTLR--FKVVDEFGNRFDVNNCSICYHWVTSRPQEPAVFSADYRGCHVLEKDGR 109
HORSE       VFPPPGWTVR--FKVVDEFGNTFEVNNCSICYHWVTAKPKGPVVFSADYKGCHVLEKDGR 176
RABBIT      VLPRPGRTIR--FKVVDEFGNRFEVNNCSICFHWVSAEPQAPAVFSADYKGCHVLEKEGH 104
MOUSE       VFPRPNQTVQ--FKVLDEFGNRFEVNNCSICYHWVTSEAQEHTVFSADYKGCHVLEKDGR 101
MACAQUE     VNLSQEATSPPVLITWDNQGLLHKLQNDSDCGTWIRKRPGSSVVLEATYSSCYVTEWDSH 102
            *      *    : . *: *  ..::* * *  *:  ..   .*:.* * .*:* * :.:

HUMAN       FHLRVFMEAVLPNGRVDVAQDATLICPKPDPSRTLDSQLAPPAMFSVSTPQTLSFLPTSG 169
CHIMPANZEE  FHLRVFMEAVLPNGRVDVAQDATLICPKPDPSRTPDSQLAPPAMFSVSTPQTLSFLPTSG 169
HORSE       SQLRVLIEAVLPEGRVDVARDVTLICPKPDHTWPLDSYLVPPSAP--STPHTRPLRPT-- 232
RABBIT      SHLTVFIEAMLPDGHVEVAQEAVLICPKPGHTWAVGSHQVPPTTPSPTTPHALPFHLSSA 164
MOUSE       FHLRVFIQAVLPNGRVDIAQDVTLICPKPDHTVTPDPYLAPPTTPEPFTPHAFALHPIPD 161
MACAQUE     YIMPVGVEGVGVAEHKMVPERKLLKCP---------------------------------- 129
              : * ::.:    :  :..   * **

HUMAN       HTSQGSGHAFPSPLDPGHSSVHPTPALPSPGPGPTLATLAQPHWGTLEHWDVNKRDYIGT 229
CHIMPANZEE  HTSQGSGHAFPSPLDPGHSSVHPTPALPSPGPGPTLATLAQPHWGTLEHWDVNKRDYIGT 229
HORSE       ---------------PEHSFVHPTPALPSLGPGPTHPTLAQPQWGTLGPWEVDKPASIGT 277
RABBIT      HT-------FPIPLYLEHSLMLPTPAGPSLGPGPTPAVLAQ-----VERWEVDKPDAVGS 212
MOUSE       HTLAGSGHTGLTTLYPEQSFIHPTPAPPSLGPGPAGSTVPHSQWGTLEPWELTELDSVGT 221
MACAQUE     ------------------------------------------------MDLLARDAPDT 140
                                                             ::      .:

HUMAN       HLSQEQCQVASGHLPCIVRRTSKEACQQAGCCYDNTREVPCYYGNTATVQCFRDGYFVLV 289
CHIMPANZEE  HLSQERCQVASGHLPCIVRRTSKEACQQAGCCYDNTREVPCYYGNTATVQCFRDGYFVLV 289
HORSE       HLAQEQCQVASGRIPCRVSGSSREACQQAGCCYDNTREVPCYYGNTATVQCFRNGHFILV 337
RABBIT      HLPQEWCQVASGHIPCIVQSSSKEACEQAGCCYDSAREVPCYYGNTATVQCFRNGYFILV 272
MOUSE       HLPQERCQVASGHIPCMVNGSSKETCQQAGCCYDSTKEEPCYYGNTVTLQCFKSGYFTLV 281
MACAQUE     DWCDS--IPARDRLPCAPSPISRGDCEGLGCCYSS--ENSCYYGNTVTLRCTREGHFSIA 196
            .  :.    * .::**     *:  *:  ****..  * .******.*::* :.*:* :.
```

Fig. 3 – Comparison of amino acid sequences of ZP1 from human, chimpanzee, horse, rabbit, mouse and macaque. The deduced amino acid sequence of rabbit ZP1 was aligned with the ZP1 sequences of the other species using the Clustal W program. The accession numbers of the sequences used are as follows: horse ZP1 (XP_001493772), human ZP1 (NP_997224), mouse ZP1 (NP_033606) and rabbit ZP1. Identical amino acids are marked by an asterisk. The colon (:) represents conserved residues and the dot (.) represents semi-conserved residues. The potential signal peptidase cleavage is between Gly20 and Gln21. The zona domain is shown in red. The trefoil domain is shown in blue. The consensus furin cleavage-site is underlined. The transmembrane domain is in orange. The cystein residues are in green. The potential N-glycosylation sites are in pink. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Please cite this article as: Stetson I., et al, Rabbit zona pellucida composition: A molecular, proteomic and phylogenetic approach, J Prot (2012), http://dx.doi.org/10.1016/j.jprot.2012.07.027

```
HUMAN          VSQEMALTHRITLANIHLAYAP-TSCSPTQHTEAFVVFYFPLTHCGTTMQVAGDQLIYEN 348
CHIMPANZEE     VSQEMALTHRITLANIHLAYAP-TSCSPTQHMEAFVVFYFPLTHCGTTMQVAGDQLIYEN 348
HORSE          VSQETASAHSFTLANVRLAYAP-TGCSPTQETGSFALFQFPLTHCGTTVQVVGNRLIYEN 396
RABBIT         VAQEMALAHRITLANVHLAYAP-TRCPPAQKTSAFVIFHVPLTHCGTTVQVLGSQLFYEN 331
MOUSE          MSQETALTHGVLLDNVHLAYAP-NGCPPTQKTSAFVVFHVPLTLCGTAIQVVGEQLIYEN 340
MACAQUE        VSRNVVSPP-LLLDSVRLALRNDSACNPVMATQAFVLFHFPFTSCGTTRRITGDRAVYEN 255
               :::: . .  . * .::**    . * *.    :*.:* .*:* ***: :: *.: .***

HUMAN          WLVSGIHIQKGPQGSITRDSTFQLHVRCVFNAS-DFLPIQASIFPPPSPAPMTQPGPLRL 407
CHIMPANZEE     WLVSGIHIQKGPQGSITRDSTFQLHVRCVFNAS-DFLPIQASIFPPPSPAPMTQPGPLRL 407
HORSE          QLVSDMDVRRGPQGSITREGTFRLHMHCIFNAS-DFLPLQASIFPPPSPAAVTQSGPLRL 455
RABBIT         QLVSDIDVREGPQGSITRDSSFRLLVRCIFNAS-DFLPIQASIFSPPLPAPVTQAGPLRL 390
MOUSE          QLVSDIDVQKGPQGSITRDSAFRLHVRCIFNAS-DFLPIQASIFSPQPPAPVTQSGPLRL 399
MACAQUE        ELVATRDVKNGSRGSVTRDSIFRLHVSCSYSVSSNSLPIKVQVFTLPPPFPETQPGPLTL 315
                **:  .::.*.:**:**:. *:* : * :..* : **::..:*.   * . **.*** *

HUMAN          ELRIAKDETFSSYYGEDDYPIVRLLREPVHVEVRLLQRTDPNLVLLLHQCWGAPSANPFQ 467
CHIMPANZEE     ELRIAKDETFSSYYGEDDYPIVRLLREPVHVEVRLLQRTDPNLVLLLHQCWGAPSANPFQ 467
HORSE          ELRIAKDVTYGSYYGEGDYPIVRLLREPVPVEVRLLQRTDPSLVLVLHQCWATPSANPVQ 515
RABBIT         ELRIARDETFSSFYEEEDYPLVRLLREPVHVEVRLLQRTDPSLVLELHQCWATPSANPVQ 450
MOUSE          ELRIATDKTFSSYYQGSDYPLVRLLREPVYVEVRLLQRTDPSLVLVLHQCWATPTTSPFE 459
MACAQUE        ELQIAKDKNYGSYYGVGDYPVVKLLRDPIYVEVSILHRTDPSLGLLLHQCWATPSTDPLS 375
               **:** * .:.*:*   ***:*:***:*: *** :*:****.* * *****.:*::.*..

HUMAN          QPQWPILSDGCPFKGDSYRTQMVALDGAT--PFQSHYQRFTVATFALLD-SGSQRALRGL 524
CHIMPANZEE     QPQWPILSDGCPFKGDSYRTQMVALDGAT--PFQSHYQRFTVATFALLD-SGSQRALRGL 524
HORSE          QPQWPILWDGCPFDGDSYRTRLVALDGAEL-PFPSHYQRFTVATFVLLD-SGSQRALRGP 573

RABBIT         QPQWPLLSDGCPFKGDSYRTRVLALDRAEL-PFRSHYQRFTVATFTFLD-SGAQRALRGL 508
MOUSE          QPQWPILSDGCPFKGDNYRTQVVAADREAL-PFWSHYQRFTITTFMLLD-SSSQNALRGQ 517
MACAQUE        QPQWPILVKGCPYIGDNYQTQLIPVQKALDLPFPSHYQRFSIFTFSFVDPTVEKQALRGP 435
               *****:* .***: **.*:*:::. :     ** ******:: ** ::* :  :.****

HUMAN          VYLFCSTSACHTSGLETCSTACSTGTTRQRRSSGHRNDTARPQDIVSSPGPVGFEDSYGQ 584
CHIMPANZEE     VYLFCSTSACHTSGLETCSTACSTGTTRQRRSSGHRNDTARPQDIVSSPGPVGFEDSYGQ 584
HORSE          VYFFCSASACAPSGLETCATACSSRTARQRRSQSHRSDTAEPQNIVSSPGPVHFEGTHGQ 633
RABBIT         VYFFCSASACHPSGPETCSSTCSSRTAKRRRSSGYHDGTPRALDIVSSPGPVGFQDSHRQ 568
MOUSE          VYFFCSASACHPLGSDTCSTTCDSGIARRRRSSGHHNITLRALDIVSSPGAVGFEDAAKL 577
MACAQUE        VHLHCSVSVCQPAETPSSVRTCPDLSRRRKFSTIFQNTTAS----VSSKGPMILLQATKD 491
               *::.**.*.* .    :.  :*     ::: *  .:. *      *** *.: :  :

HUMAN          EPTLGPTDSNGNSSLRPLLWAVLLLPAVALVLGFGVFVGLSQTWAQKLWESNRQ----- 638
CHIMPANZEE     EPTLGPTDSNGNSSLRPLLWVVLSLPAVALVLGFGVFVGLRQTWAQKLWESNRQ----- 638
HORSE          EPTLRPTGSTRNSKPRPLLWMVLLLVAIALVLGVGIFVGLRQARAQKLQEGNRG----- 687
RABBIT         EPTLESTGSGRNSNPKPLLWVVLLLLAIALVLGIGVFVGLSQAWAHKLREGHRLTDQAQ 627
MOUSE          EP----SGSSRNSSSR----MLLLLLAITLALAAGIFVGLIWAWAQKLWEGIRY----- 623
MACAQUE        PP--EKLRAPVDS-------KVLWVAGLSGTLILGGLVVSYLAIKQLNCPDQTCQ---- 537
                *      :  :*        :* : .:: .*  * :*    :  :    .
```

Fig. 3 (*continued*).

Cys (Fig. 3) strongly suggest that the amplified ORF corresponds to rabbit ZP1.

A total of 12 potential O-glycosylation sites were predicted in the mature protein and two potential N-glycosylation sites (Asn-X-S/T) are present in mature rabbit ZP1 in position Asn71 and Asn362. These equivalent positions are conserved in horse, human, mouse and rat ZP1; however, an additional N-glycosylation site, Asn49, is present in the mouse and rat species [36,37] and is lacking in horse, human and rabbit ZP1 (Fig. 4 and Supplementary material 3).

4.2. *Mass spectrometry of rabbit ZP glycoproteins*

Following amplification of ZP1 ORF, the next step was to confirm the expression of the four proteins in rabbit ZP. For this purpose, the rabbit ZP extracted from ovaries or oocytes, as described in the Material and methods section, was analyzed directly by MS/MS or after separation of the proteins by SDS-PAGE electrophoresis followed by silver staining of the gel. In this last situation, gel segments were reduced and alkylated, trypsinized and analyzed by LC–ESI-MS–MS. A summary of the peptides identified is included in Table 2 and Fig. 4.

Several peptides corresponding to the immature version of the proteins were detected, indicating that ZP expression is continuous and elevated or, alternatively, that signal peptides and the carboxyl terminal region are not efficiently removed. These regions of the proteins could probably be identified because the protocol used for ZP isolation differs from those of previous studies performed in other species.

Taking into account that HPLC–MS analysis can be considered as a semiquantitative technique, the fact that the coverage of ZP2 and ZP3 was similar (≈50%) might indicate the

Please cite this article as: Stetson I., et al, Rabbit zona pellucida composition: A molecular, proteomic and phylogenetic approach, J Prot (2012), http://dx.doi.org/10.1016/j.jprot.2012.07.027

Rabbit ZP1

```
  1 MTGGRLVALL LLVAASLGLG QQPHPEPGLP GLQYSYDCGM RGMQLVVLPR PGRTIRFKVV
 61 DEFGNRFEVN NCSICFHWVS AEPQAPAVFS ADYKGCHVLE KEGHSHLTVF IEAMLPDGHV
121 EVAQEAVLIC PKPGHTWAVG SHQVPPTTPS PTTPHALPFH LSSAHTFPIP LYLEHSLMLP
181 TPAGPSLGPG PTPAVLAQVE RWEVDKPDAV GSHLPQEWCQ VASGHIPCIV QSSSKEACEQ
241 AGCCYDSARE VPCYYGNTAT VQCFRNGYFI LVVAQEMALA HRITLANVHL AYAPTRCPPA
301 QKTSAFVIFH VPLTHCGTTV QVLGSQLFYE NQLVSDIDVR EGPQGSITRD SSFRLLVRCI
361 FNASDFLPIQ ASIFSPPLPA PVTQAGPLRL ELRIARDETF SSFYEEEDYP LVRLLREPVH
421 VEVRLLQRTD PSLVLELHQC WATPSANPVQ QPQWPLLSDG CPFKGDSYRT RVLALDRAEL
481 PFRSHYQRFT VATFTFLDSG AQRALRGLVY FFCSASACHP SGPETCSSTC SSRTAKRRRS
541 SGYHDGTPRA LDIVSSPGPV GFQDSHRQEP TLESTGSGRN SNPKPLLWVV LLLLAIALVL
601 GIGVFVGLSQ AWAHKLREGH RLTDQAQ
```

Rabbit ZP2

```
  1 MQVSNSGSRG KRLPWPSLTK FTFPYLSPPS SSSACTWLFL FFTLVTSVNS IYFLQLSDPA
 61 FPGTVTCNEN EIMVEFPSYV GTKTLHASVV DPLGVEMLNC TYILDPEKLT LRVPYKACTR
121 AVHGGHQMSI RVMNNSAAL RHTDVEYQFFC PVEQTLEFSK SAACTKDFMS LSFPRIPTGL
181 GDSTMVNESQ MGWMVQAGHG PGAQTLSLEE AKGQGFGVLI DDNKMTLSVL LNATGVTHYV
241 EGTSHLHTMF LKLSLVSPGQ KMTFPSRAIC LSGPVTCNAT HMTLTIPEFP GKLESVSIEN
301 RNITVSQLHD QGIDVEAING LRLHFSKTVL KTKFSEKCLH DQLYISSLKL TFNLELDTVS
361 TVINPECPCD SPASIVSGEL CTQDGFMDFE VYTHQTKPAL NLDTLRVGSS SCQPVFKAQS
421 QGLVRFRIPL NGCGTRHKFE DEKVIYENEV HALWENLPPS KISRDSEFRM TVQCYYTRDD
481 MLLNANIKSL PPPVASVKPG PLALSLQTYP DESYQQPYRV NEYPIVKYLR QPIYMEVRVL
541 NRNDPNIKLA LDDCWATSSM DPASLPKWSI VMDGCEYSLD NYQTNFHPVG SSVTYPEHYQ
601 RFDVKTFAFV SEAQARSSLV YFHCSALICN QHYPDSPLCS VTCPGSSRHR RATGNTEEER
661 VTASLPGPIL LLPNGSSFRG VGDSKEHGMA GDVTSKTMAA VAAVAGVVAT LGFISYLCKK
721 RTMMLSH
```

Rabbit ZP3

```
  1 MGLSYGLFVC LLLWGGSELC CPQPLWFWQG GTRQPAPSVT PVVVECLEAR LVVTVSRDLF
 61 GTGKLIQEAD LSLGPEGCEP QASTDAVVRF EVGLHECGNS VQVTDDSLVY SSFLAGNLSI
121 LRTNRAEVPI ECRYPRQGNV SSRAILPTWV PFWTTVLSEE RLVFSLRLME ENWSREKMSP
181 TFHLGDTAHL QAEVRTGSHP PLLLFVDRCV ATPTRDQSGSP YHTIVDLHG CLVDGLSDGA
241 SKFKAPRPKP DVLQFMVAVF HFANDSRHTV YITCHLRVIPA QQAPDRLNK ACSFNQSSSS
301 WAPVEGSADI CECCGNGDCD LIAGSPMNQN HAARSSLRSRR HVTEEADVT VGPLIFLGKA
361 GDPAGTEGLA SAAQATLVLG LRMATIVFLA VAAVVLGLTRG RHAASHPRS ASQ
```

Rabbit ZP4

```
  1 MAPGSTMWLL GYIFLCFPVS PKPFALIKQE TPTDPGVLHC RPWNFKFTIN FQNQETGSSP
 61 VLVTWDNQGR LHRLQNDTDC GTRVGEGPGP SVVLEANYSS CYVTESEPYY VMLVGVEEVD
121 AAGQNLVTKQ QLLKCPMHLP APDAGLCDSV PVQDRLPCAT APISQEDCEE LGCCHSSEEV
181 NACYYGNTVT SHCTQEGHFS IAVSRNVSSP PLHLDSVHLV FGNDSECQPV VATRAFVLFL
241 FPFTACGTTR QITGDRAIYE NELLATREVR TWSRGSITRD SIFRLRVSCS YSISSSALPV
301 DMHVLTLPPP LPETQPGPLT VVLQIAKDKD YHSYYTMDDY PVVKLLRDPI YVDVSILYRT
361 DPYLGLRLHQ CWATPRTNPL YQPQWPILVK GCPYTGDNYQ TQLIPVQEAF DLPFPSHHQR
421 FSISTFSFLD SSVAKEALKG PIYLHCSVSV CQPTGTQSCT VTCPIDSRRR NSDINFQNST
481 ANISSKGPMI LLQATEDPSE KLHKHSGVPV HPGALWVAGL SGIFIIGALL VSYVAIRTRR
```

Please cite this article as: Stetson I., et al, Rabbit zona pellucida composition: A molecular, proteomic and phylogenetic approach, J Prot (2012), http://dx.doi.org/10.1016/j.jprot.2012.07.027

similar abundance of these two proteins. However, ZP4 seems less abundant (≈43%) and ZP1 may be the least abundant protein in the ZP matrix (≈33%). Taking into account these data, it seems that the levels of ZP2 and ZP3 in the ZP matrix may be higher than the levels of ZP1 and ZP4, even though one should realize that this is indicative only given the sequence differences and ionization efficiency differences of the different peptides. Nevertheless, a similar situation is observed in mouse, in which the levels of ZP1 mRNA are four times lower than those of ZP2 and ZP3 [38] and so only 56% of the ZP1 polypeptide chain can be identified by direct MS/MS, compared with the 96 and 100% of ZP2 and ZP3 respectively [36]. Other species show a similar pattern to human or rat, in both of which proteomic analysis has revealed a coverage percentage that is similar between ZP2 and ZP3, with ZP1 being the least abundant protein [11,12]. Future quantitative proteomic analysis will be performed to ascertain the ZP glycoprotein stoichiometry of the mature ZP.

4.2.1. ZP1

A total of 21 different peptides were identified in the different analyses yielding a sequence coverage of 33.9%. None of the identified peptides contained an N-glycosylation site, suggesting that both the described consensus sequences may be occupied in the mature glycoprotein. An N-glycosylation site present in the N-terminal region of the mouse and rat ZP1 protein was not conserved in the rabbit and human ZP1 [36,37] (Supplementary material 3). On the other hand, 35 out of the predicted O-glycosylation sites were contained in the identified peptides, from which it can be deduced that these residues are either not glycosylated at all or are, at most, partially glycosylated. A similar result was previously reported for ZP1 in mouse and rat where proteomic analysis did not detect O-glycosylation sites [36,37].

4.2.2. ZP2

Forty peptides corresponding to rabbit ZP2 were detected. These peptides correspond to 55.7% of the protein. 65 potential sites of O-glycosylation and five sites of N-glycosylation (Asn99, Asn134, Asn278 and Asn302) can be localized in the detected peptides (Fig. 4). These residues might not be glycosylated at all or be partially glycosylated. In contrast, the conserved Asn corresponding to Asn99 and Asn278 was detected as glycosylated peptides in mouse and rat [36,37]. Other differences in the degree of N-glycosylation in this protein between cow, human, pig and rabbit can also be detected (Supplementary material 3).

4.2.3. ZP3

A total of eighteen peptides from rabbit ZP3 could be convincingly identified in the different experiments. This corresponds to 50.3% of the sequence (Fig. 4). Two sites of N-glycosylation (Asn139 and Asn264) were detected in the analysis (Fig. 4). These two N-glycosylation sites may not be glycosylated in the native protein or, at most, be partially glycosylated. In contrast, these conserved sites are glycosylated in human, mouse and rat [39,36,37] (Supplementary material 3). Different studies have suggested that carbohydrates play a key role during the sperm–ZP interaction in different species. N-glycans of human ZP3 have important roles in the induction of the AR [40]. However, more attention has been paid to O-glycans. In particular, O-linked oligosaccharides in mice ZP3 have been analyzed in depth, although there is a controversy about the exact role played by O-glycans in this process [41–44]. Thirty-four potential O-glycosylation sites were detected in this analysis, suggesting that they are not occupied or only partially occupied. In recombinant human, mouse and rat ZP3, two clusters of O-glycans have been reported, some of them are similar among species; however, they are not identical and this probably contributes to the species specificity of the gamete interaction. The first cluster corresponds to the amino acid residues 156–173 [39]. The Thr155 and Thr156 in mouse and human, respectively, are glycosylated and have been suggested to play a role in sperm binding [42]; however, the corresponding amino acid in rabbit ZP3, Thr154, is at least partially unoccupied, indicating that differences exist between these species. A similar result was observed with the second cluster observed in the human ZP3, which corresponds to Thr260, Thr264 and Thr281. Identification in this study of the peptide (Aa 251–296) in rabbit ZP3 indicates that there is no a clustering because the first two Thrs are not conserved. The Thr279 is conserved but is not totally glycosylated. The role of the different oligosaccharide chains in fertilization in rabbit remains unresolved thus far.

4.2.4. ZP4

Twenty-three peptides were detected in ZP4 (43.5% coverage). Forty-one O-glycosylation and two N-glycosylation (Asn478, Asn482) sites were observed, which might mean that these sites are not glycosylated in mature protein. Asn75 and Asn206 are not detected in the peptides probably because they are occupied. The corresponding amino acid in the rat [12] and pig are also glycosylated (Supplementary material 3).

Different O-glycosylation sites were detected in the pig and rat ZP4 [12]. In rat ZP4, the precise amino acid involved was not determined; however, in the pig ZP4 the amino acid residues (Ser293 and Thr303) are glycosylated and are conserved in the rabbit ZP4. A peptide including this region is not detected suggesting that probably it is glycosylated as observed in pig ZP4. Future glycomic studies are necessary to obtain more precise information about the ZP4 glycosylation.

Thus, although the mouse ZP composition might originally have supported the hypothesis that mammalian ZP has three proteins, the mouse model has been revealed to be an exception within mammals [4]. This study demonstrates that the rabbit ZP is formed by four proteins, as is the human ZP, making this species a good animal model for understanding the role played

Fig. 4 – Rabbit ZP1, ZP2 (XP_002711880), ZP3 (NP_001182649), and ZP4 (NP_001075764) amino acid sequences. Underlined sequences are the tryptic peptides obtained by MS/MS. The detected putative N-glycosylation sites are in red and the O-glycosylation sites are in blue. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Please cite this article as: Stetson I., et al, Rabbit zona pellucida composition: A molecular, proteomic and phylogenetic approach, J Prot (2012), http://dx.doi.org/10.1016/j.jprot.2012.07.027

t2.1 **Table 2 – Peptides identified by proteomic analysis. Peptides with a score higher than 5, and percentage-scored peak intensity of 60%, which are the threshold criteria for a positive identification, are shown in red. The * indicates that the peptide has been also detected in ZP from oocytes.**

Q2

t2.2

Peptides	Theoretical $[M+H]^+$	Sequence	z	m/z	Number of detections	Score	SPI
ZP1							
GMQLVVLPRPGRTIR*	1693.0059	42–56	3	565.16	2	3.36	75
GMQLVVLPRPGRTIRFKVVDEFGNR	2884.6096	42–66	3	967.62	7	5.71	53.2
EVPCYYGNTATVQCFR	1850.8206	250–265	2	983.42	2	10.57	83.8
ITLANVHLAYAPTRCPPAQK*	2164.1701	283–302	3	741.80	1	3.86	62.6
EGPQGSITRDSSFR	1536.7406	341–354	3	565.89	1	3.91	51.7
EGPQGSITRDSSFRLLVR	2018.0783	341–358	3	674.09	1	5.21	60.5
IARDETFSSFYEEEDYPLVRLLR	2848.4157	394–416	3	950.07	1	3.06	57.9
LLREPVHVEVR*	1346.7908	414–424	2	674.17	1	4.10	87.7
GDSYRTRVLALDRAELPFR	2235.1998	465–483	2	1028.34	1	3.61	54.8
GDSYRTRVLALDRAELPFRSHYQR	2906.5138	465–488	3	996.28	1	4.21	61.4
TRVLALDR	943.5689	470–477	1	943.04	1	5.20	54.3
SHYQRFTVATFTFLDSGAQRALR*	2672.3697	484–506	3	918.66	9	5.34	86.4
FTVATFTFLDSGAQR	1660.8335	489–503	2	831.27	1	6.60	77.3
ALRGLVYFFCSASACHPSGPETCSSTCSSR	3124.3749	504–533	3	1095.04	1	5.51	71.5
ALRGLVYFFCSASACHPSGPETCSSTCSSRTAK	3424.5546	504–536	3	1188.22	1	6.31	54.5
GLVYFFCSASACHPSGPETCSSTCSSR	2784.1526	507–533	3	982.78	8	5.63	59.6
GLVYFFCSASACHPSGPETCSSTCSSRTAKR	3240.4335	507–537	3	1137.15	1	5.11	59.6
GLVYFFCSASACHPSGPETCSSTCSSRTAKRR	3396.5346	507–538	3	1133.40	1	4.58	55.6
RSSGYHDGTPRALDIVSSPGPVGFQDSHR	3095.5047	539–567	3	1059.79	3	6.22	79.0
SSGYHDGTPR*	1076.4761	540–549	2	578.23	3	3.78	61.8
SSGYHDGTPRALDIVSSPGPVGFQDSHR*	2399.4036	540–567	3	980.36	1	3.59	86.9
ZP2							
GKRLPWPSLTK	1282.7636	10–20	2	641.80	1	5.20	57.6
TLHASVVDPLGVEMLNCTYILDPEK*	2814.4058	84–108	3	973.15	17	10.24	63.2
TLHASVVDPLGVEMLNCTYILDPEKLTLR	2814.4058	84–112	3	1134.69	1	3.45	50.6
LTLRVPYKACTR	1420.8098	109–120	3	520.18	8	5.95	59.8
LTLRVPYKACTRAVHGGHQMSIR*	2594.3924	109–131	3	924.72	7	5.62	74.1
VPYKACTRAVHGGHQMSIR	2194.1014	113–131	3	750.36	1	5.62	54.8
VPYKACTRAVHGGHQMSIRVMNNSAALR	3067.5617	113–141	3	1074.28	2	5.13	54.6
ACTRAVHGGHQMSIR*	1623.7960	117–131	2	821.41	2	5.58	53.8
ACTRAVHGGHQMSIRVMNNSAALR*	2580.2822	117–141	3	919.09	8	4.97	54.7
AVHGGHQMSLRVMNNSAALR	2149.0871	121–141	3	743.70	2	5.11	65.2
HTDVEYQFFCPVEQTLEFSK	2504.1444	142–160	2	1252.16	1	16.57	94.0
SAACTKDFMSLSFPR	1660.7827	161–175	2	899.35	1	4.09	50.8
LSLVSPGQKMTFPSR	1647.8892	253–267	2	872.86	1	3.08	71.0
MTFPSRAICLSGPVTCNATHMTLTIPEFPGK	3320.6304	262–292	3	1145.56	1	5.13	53.5
AICLSGPVTCNATHMTLTIPEFPGK*	2601.8879	268–292	3	918.12	14	8.95	95.6
NITVSQLHDQGIDVEAINGLRLHFSK	2904.5332	302–327	3	995.05	1	3.90	52.5
LHFSKTVLKTKFSEK*	1793.0325	323–337	2	897.05	1	3.75	59.9
TVLKTKFSEKCLHDQLYISSLK	2581.4064	328–349	3	887.71	4	4.71	64.9
FSEKCLHDQLYISSLK	1910.9686	334–349	3	637.09	1	3.95	60.4
VGSSSCQPVFKAQSQGLVRFR	2281.1875	407–427	3	814.33	1	4.27	61.6
VGSSSCQPVFKAQSQGLVRFRIPLNGCGTRHK	3457.8061	407–438	6	599.62	1	4.40	54.8
AQSQGLVRFRIPLNGCGtRHKFEDEK	2986.5433	418–443	3	1022.77	1	3.16	57.3
FRIPLNGCGTRHKFEDEK	2147.0820	426–443	3	762.16	1	5.39	67.9
DSEFRMTVQCYYTR	1798.7892	465–478	3	654.00	1	3.51	70.0
DSEFRMTVQCYYTRDDMLLNANIK	2926.3538	465–488	3	999.96	1	4.97	67.7
MTVQCYYTRDDMLLNANIK	2292.0826	470–488	3	820.47	1	5.87	51.8
VNEYPIVK	961.5358	520–527	1	961.32	1	3.97	81.0
VNEYPIVKYLRQPIYMEVR	2410.2957	520–538	3	836.52	1	4.48	80.0
YLRQPIYMEVRVLNRNDPNIK	2631.4193	528–548	3	930.48	2	3.27	69.3
QPIYMEVRVLNRNDPNIK*	2199.1708	531–548	3	733.72	2	5.83	57.5
NDPNIKLALDDCWATSSMDPASLPK*	2702.2806	543–567	3	933.83	13	6.56	52.4
LALDDCWATSSMDPASLPK	2020.9360	549–567	2	1039.91	1	4.28	51.2
SSLVYFHCSALICNQHYPDSPLCSVTCPGSSRHRR	3920.8205	617–651	6	677.37	1	4.78	60.7
VTASLPGPILLLPNGSSFRGVGDSK	2482.3669	661–685	3	827.92	1	3.55	75.8
GVGDSKEHGMAGDVTSK*	1674.7757	680–696	2	917.15	3	6.01	71.5
EHGMAGDVTSKTMAAVAAVAGVVATLGFISYLCKK	3496.8006	686–720	3	1199.02	3	5.12	56.5
EHGMAGDVTSKTMAAVAAVAGVVATLGFISYLCKKRTMMLSH	4353.2054	686–727	5	906.90	4	6.47	60.7
EHGMAGDVTSK*	1131.5104	686–696	2	606.24	1	4.19	63.3

Please cite this article as: Stetson I., et al, Rabbit zona pellucida composition: A molecular, proteomic and phylogenetic approach, J Prot (2012), http://dx.doi.org/10.1016/j.jprot.2012.07.027

Table 2 (*continued*)

ZP3							
QPAPSVTPVVVECLEARLVVTVSR	2549.4125	34–57	3	877.52	1	3.44	61.3
QPAPSVTPVVVECLEARLVVTVSRDLFGTGK	3267.7775	34–64	3	1104.02	1	3.94	94.1
LVVTVSR	773.4885	51–57	1	773.55	5	4.90	53.7
DLFGTGK*	737.3834	58–64	1	737.60	12	9.53	74.9
TNRAEVPIECRYPRQGNVSSR*	2432.2217	123–143	3	856.27	1	3.16	50.5
QGNVSSRAILPTWVPFWTTVLSEER*	2837.4950	137–161	3	984.42	3	4.62	60.5
AILPTWVPFWTTVLSEERLVFSLR	2860.5765	144–167	3	980.71	1	6.18	78.2
MSPTFHLGDTAHLQAEVRTGSHPPLLLFVDR	3442.7694	178–208	3	1175.04	1	4.69	50.3
DQSGSPYHTIVDLHGCLVDGLSDGASK*	2771.2946	216–242	3	943.90	18	13.38	74.1
FKAPRPKPDVLQFMVAVFHFANDSR*	2917.5299	243–267	3	973.10	1	3.01	59.4
APRPKPDVLQFMVAVFHFANDSR*	2642.3666	245–267	3	881.35	6	5.21	55.5
HTVYITCHLR*	1242.6417	268–277	2	650.33	1	4.11	72.7
HTVYITCHLRVIPAQQAPDRLNK*	2673.4411	268–290	3	918.57	1	6.11	75.1
SSLRSRRHVTEEADVTVGPLIFLGK*	2767.5219	335–359	3	923.12	4	5.42	68.5
SRRHVTEEADVTVGPLIFLGK	2324.2726	339–359	3	827.93	1	3.29	57.9
RHVTEEADVTVGPLIFLGK*	2081.1395	340–359	3	721.25	1	4.03	71.6
MATIVFLAVAAVVLGLTR*	1845.1036	383–399	2	970.90	2	4.80	68.5
MATIVFLAVAAVVLGLTRGRHAASHPR*	2814.6041	383–408	3	970.87	4	5.51	57.1
ZP4							
FTINFQNQETGSSPVLVTWDNQGR*	2738.3174	47–70	3	966.51	2	5.59	58.1
QQLLKCPMHLPAPDAGLCDSVPVQDR	2831.4006	130–155	3	984.75	5	5.31	84.5
CPMHLPAPDAGLCDSVPVQDR	2221.0204	135–155	3	778.71	1	4.65	73.9
AFVLFLFPFTACGTTR	1847.9518	235–250	2	924.46	4	16.01	91.8
QITGDRAIYENELLATR*	1963.0249	251–264	3	648.94	2	6.95	65.7
QITGDRAIYENELLATREVR	2347.2370	251–270	3	836.85	3	4.86	75.9
QITGDRAIYENELLATREVRTWSR*	2877.4971	251–274	3	987.25	2	3.49	61.2
AIYENELLATR	1292.6850	257–267	2	648.04	6	18.30	97.5
AIYENELLATREVRTWSRGSITR*	2721.4436	257–279	3	960.92	3	4.00	55.6
EVRTWSRGSITRDSIFR*	2066.0895	268–284	3	741.91	4	4.96	78.3
TWSRGSITRDSIFR*	1681.8774	271–284	2	921.26	1	3.49	67.7
DSIFR	637.3309	280–284	1	637.93	5	6.85	89.2
DKDYHSYYTMDDYPVVK	2138.9381	328–344	3	739.56	2	4.85	87.0
DKDYHSYYTMDDYPVVKLLR*	2521.2073	328–347	3	841.06	1	3.30	51.6
DYHSYYTMDDYPVVK	1895.8162	330–344	3	632.41	1	3.68	79.7
LLRDPIYVDVSILYR*	1835.0431	345–359	2	918.96	3	14.09	90.6
LLRDPIYVDVSILYRTDPYLGLR*	2750.5245	345–367	3	970.47	5	6.12	56.7
TNPLYQPQWPILVK	1696.9427	377–390	2	848.13	2	11.44	74.5
FSISTFSFLDSSVAK	1635.8270	421–435	2	818.34	3	12.47	79.6
EALKGPIYLHCSVSVCQPTGTQSCTVTCPIDSR	3493.6588	436–468	3	1191.29	1	3.82	84.7
RNSDINFQNSTANISSK	1895.9211	470–486	2	1028.34	1	3.40	58.5
NSDINFQNSTANISSK*	1739.8200	471–486	2	950.03	2	4.33	85.6
GPMILLQATEDPSEK*	1628.8205	487–501	2	895.71	1	3.68	93.7

by the ZP for several reasons. First, the similarity between rabbit and human ZP proteins is generally greater than in other species (except primates and horse, Supplementary material 4). Second, ZP biogenesis in rabbit is similar to primates including human [45–50]. In human and rabbit ovarian follicles, oocytes and granulosa cells contribute to the formation of the ZP [45,46]; however, in hamster, mice and rat the ZP is only formed by the oocytes [51–57]. Third, these animals are used for human consumption in numerous countries, and so large amounts of rabbit ovaries are available in the slaughterhouses and could contribute to reducing the number of animal sacrifices necessary for research purposes.

4.2.5. *ZP1 phylogeny*

It was previously reported using in silico approaches that ZP genes have been gradually lost during the evolution of vertebrates [4]. Thus, in several mammals the ZP1, ZPB/ZP4, ZPD, and/or ZPAX gene is lacking. These data suggested the presence of pseudogenes in the genome of these species. In particular, the same authors did not find ZP1 in rabbit and other species [4]. Our results demonstrate, however, that rabbit ZP is composed of four glycoproteins, including ZP1. The difference observed between both studies is probably due to the incompleteness of the rabbit genome sequence available at that time.

Additionally, four ZPs sequences are annotated in the genome of pika (*Ochotona princeps*), which belongs to the same order as the rabbit (the Lagomorpha), suggesting that this species also have four proteins. Future proteomic analyses are necessary to confirm these in silico findings. Our similarity search also found sequences corresponding to ZP1 in the genome of *Canis* (chromosome 18) and *Bos* (chromosome 29),

as already mentioned by Goudet et al.; and new ones in the genome of *Sus* (chromosome 2), *Callithrix* (chromosome 11), and *Spermophilus* (scaffold_129656) (Supplementary material 2). Our analysis of the sequences indicates that at least 4 sequences are probably pseudogenes due to the presence of deletions and/or stop codons in addition to *Canis* and *Bos* [4]: *Callithrix*, *Tarsius* and two other cetartiodactyls: *Sus* and *Tursiops* (Fig. 5).

The phylogenetic tree reconstructed with PhyML with the GTR+I+G model of sequence evolution is presented in Fig. 6. The topology is congruent with the classic phylogeny of mammals. Pseudogenization of the ZP1 occurred at least four times during the evolution of mammals (indicate by red branches in Fig. 6).

ZP1 and ZP4, two paralogous genes from the ZPB subfamily, were formed by gene duplication. Previous phylogenetic studies indicated that they share a common ancestral gene [4,5,58].

Gene duplication is possible in three situations: a) the ancestral function is partitioned and shared by the two members of the duplicated pair (subfunctionalization), b) one duplicate acquires a new function while the other retains the original function (neofunctionalization) and c) one gene degenerates to a pseudogene by accumulation of mutations and the other maintains the original gene function. The last situation corresponds with the species with three glycoproteins in which ZP1 or ZP4 is lost.

Therefore, the common origin of ZP1 and ZP4 is suggested by the observation that both are involved in identical molecular mechanisms. Studies have shown that ZP3-induced acrosome reaction (AR) involves activation of the G(i)-coupled receptor pathway, whereas ZP1- and ZP4-mediated ARs are independent of this pathway. The ZP3-induced AR involves the activation of T-type voltage-operated calcium channels (VOCCs), whereas ZP1- and ZP4-induced ARs involve both T- and L-type VOCCs [59].

Thus, in species with four glycoproteins, the fact that ZP1 and ZP4 participate in AR through similar pathways may indicate a possible subfunctionalization.

The function played by each ZP protein is not totally clear and differs in the species. A structural function was initially attributed to the ZP1 based on the mouse model [60,61]. In human, ZP1 has been related with the induction of the AR [62,63]. However, ZP4 seems to be implicated more directly in the fertilization process in numerous mammal species [64–67]. Thus, in human, it was reported that ZP4 is involved

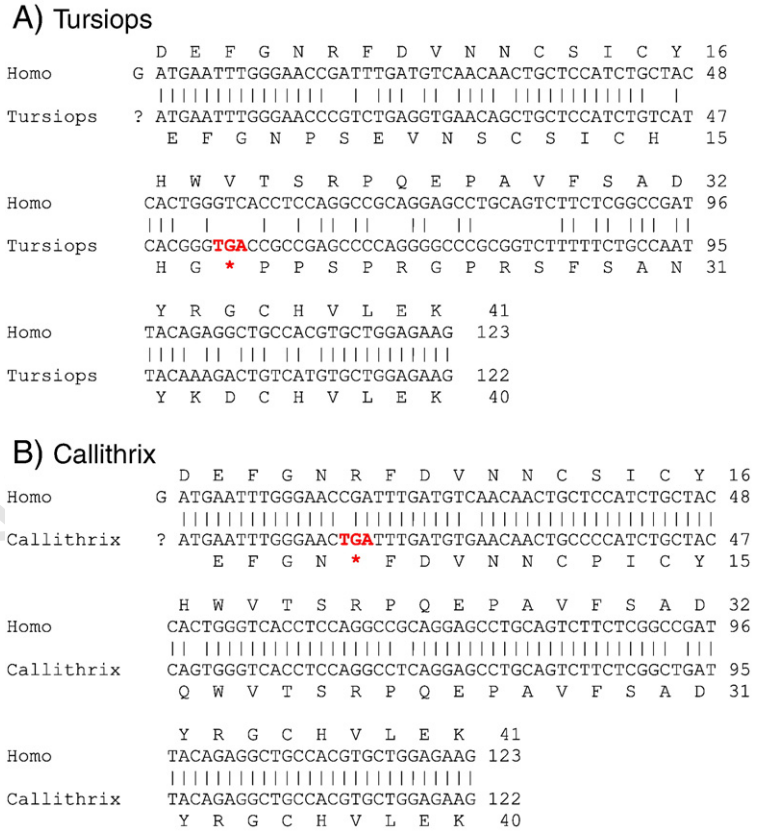


Fig. 5 – Illustration of the presence of stop codon in the different sequences of the putative pseudogenes. All the sequences are aligned to the ZP1 sequence of *Homo*. A) Exon 2 of *Tursiops*. B) Exon 2 of *Callithrix*. Exon 1 is missing in these two species, but the last nucleotide of the first exon of *Homo* was added for the translation. C) Exon 3 and exon 4 (underlined) of *Tarsius*. D) First 180 bp of the exon 3 of *Sus* (the sequence of *Sus* is highly incomplete with numerous indels).

Please cite this article as: Stetson I., et al, Rabbit zona pellucida composition: A molecular, proteomic and phylogenetic approach, J Prot (2012), http://dx.doi.org/10.1016/j.jprot.2012.07.027

C) Tarsius

```
          D  G  R  F  H  L  R  V  F  M  E  A  V  L  P  N  G  R  V  D  20
Homo      GATGGGCGTTTCCACCTGAGGGTGTTCATGGAGGCTGTGCTGCCCAATGGTCGTGTGGAT 60
          |  ||||||||||||||||||||  |||||  |||||     |||  ||||  |    ||  ||||||
Tarsius   GGTGGGCGTTTCCACCTGAGGATGTTCGTGGAGAGCATGCAGCCCGACCATCAGGTGGAT 60
          G  G  R  F  H  L  R  M  F  V  E  S  M  Q  P  D  H  Q  V  D  20

          V  A  Q  D  A  T  L  I  C  P  K  P  D  P  S  R  T  L  D  S  40
Homo      CAGGCACAAGACGCTACTCTGATCTGTCCCAAACCTGACCCCTCCCGGACTCTGGACTCC 120
          |||||||  |||  |  ||||||  ||  ||||||||||||||||  ||  ||    |  ||||||||
Tarsius   TAGGCACAGGACACCGCTCTGACCTATCCCAAACCTGACCCCACCTGGGTCCCGGACTCC 120
          V  A  Q  D  T  A  L  T  Y  P  K  P  D  P  T  W  V  P  D  S  40

          Q  L  A  P  P  A  M  F  S  V  S  T  P  Q  T  L  S  F  L  P  60
Homo      CAGCTGGCACCACCCGCCATGTTCTCTGTCTCAACCCCACAAACCCTTTCCTTCCTCCCC 180
          |  |  ||||||||||  ||  ||||||  ||||  ||||  |  |  |    ||  |||  ||||
Tarsius   TACCCGGCACCACCCACCGAGTTCTCACTCTCTGCCCCTAATAGC---TCTTTCTCCCCC 177
          Y  P  A  P  P  T  E  F  S  L  S  A  P  N  S     S  F  S  P  59

          T  S  G  H  T  S  Q  G  S  G  H  A  F  P  S  P  L  D  P  G  80
Homo      ACCTCTGGCCATACCTCCCAAGGCTCTGGCCATGCCTTTCCCAGCCCACTGGACCCAGGG 240
          ||||  |||||  ||  ||    ||  |  ||||  |||  |  |  ||||||  ||||||  ||  |
Tarsius   TCCTCCGGCCACGCCCCCG--GGGCCCAGCCACGCCCTGCTCAGCCCTCTGGACAAAGAG 235
          S  S  G  H  A  P     G  A  Q  P  R  P  A  Q  P  S  G  Q  R    78

          H  S  S  V  H  P  T  P  A  L  P  S  P  G  P  G  P  T  L  A  100
Homo      CACAGCTCTGTCCACCCAACCCCTGCTTTACCATCCCCTGGACCTGGACCTACCCTCGCC 300
          |||||||||  |||||||||  |  ||  |  |    |  |||  |   |||  |||  |||  |||  |  ||
Tarsius   CACAGCTCTATCCACCCAAGCGCTTCCTCTTCCTCCTCCAGACTTGGGCCTGCCCACCCC 295
          A  Q  L  Y  P  P  K  R  F  L  F  L  L  Q  T  W  A  C  P  P    98

          T  L  A  Q  P  H  W  G  T  L  E  H  W  D  V  N  K  R  D  Y  120
Homo      ACCCTGGCTCAACCCCACTGGGGCACCTTGGAACACTGGGATGTGAACAAACGAGATTAC 360
          ||||||||||||||  ||  |  ||||||||||||  ||  ||||||  |||  ||  |||  ||  ||
Tarsius   ACCCTGGCTCAACTCCTCGGGGGCACCTTGGGACCCTGGGAAGTGGACGAACCAGGTTCT 355
          H  P  G  S  T  P  R  G  H  L  G  T  L  G  S  G  R  T  R  F  118

          I  G  T  H  L  S  Q  E  Q  C  Q  V  A  S  G  H  L  P  C  I  140
Homo      ATAGGTACCCACCTGAGCCAGGAGCAGTGCCAGGTGGCCTCAGGGCACCTCCCCTGCATC 420
          ||||||||||  ||||  ||||||  |||||||  |||||||||  ||||  |  ||||||||||
Tarsius   TTAGGTACCCATCTGACCCAGGAACAGTGCCGGGTGGCCTCCGGGCCCATCCCCTGCATC 375
          F  R  Y  P  S  D  P  G  T  V  P  G  G  L  R  A  H  P  L  H   138

          V  R  R  T  S  K  E  A  C  Q  Q  A  G  C  C  Y  D  N  T  R  160
Homo      GTGAGAAGAACTTCAAAAGAAGCCTGTCAGCAGGCTGGCTGCTGCTATGACAACACCAGA 480
          ||||  |    ||||  ||  ||||||||||||||  ||||||||||||||||  ||||
Tarsius   ATGAGTGG---CCCAAAGGAGTCCTGTCAGCAGGCTGACTGCTGCTATGACAACATCAGA 432
          H  E  W     P  K  G  V  L  S  A  G  *  L  L  L  *  Q  H  Q   157

          E  V  P  C  Y  Y  G  N  T   169
Homo      GAGGTTCCCTGTTACTATGGCAACACAG 508
          ||||||||||  ||  ||||||||||||||
Tarsius   GAGGTTCCCTGCTATTATGGCAACACAG 460
          R  G  S  L  L  L  W  Q  H   166
```

D) Sus

```
          D  G  R  F  H  L  R  V  F  M  E  A  V  L  P  N  G  R  V  D    20
Homo      GATGGGCGTTTCCACCTGAGGGTGTTCATGGAGGCTGTGCTGCCCAATGGTCGTGTGGAT    60
          ||  ||        ||  |  |     ||  |||  |   |||  |     ||  |  |  |
Sus       AATCAGCAGCCGAGCCAGCAGAAAGCCA-GGAAG---TGGAGTTCC-TGTT-GCGGCTCA    55
          N  Q  Q  P  S  Q  Q  K  A   R  K     W  S  S   C   C  G  S    18

          V  A  Q  D  A  T  L  I  C  P  K  P  D  P  S  R  T  L  D  S    40
Homo      GTGGCACAAGACGCTACTCTGATCTGTCCCAAACCTGACCCCTCCCGGACTCTGGACTCC   120
          |      ||||  ||     ||||||  ||||||||||  |  ||  ||||     |||  |
Sus       GCAGATGAAGAACCT-----CATCTGTGGCAAACCTGACCACACCTGGACCTGACACTAC   110
          A  D  E  E  P       H  L  W  Q  T  *  P  H  L  D  L  T  L     36

          Q  L  A  P  P  A  M  F  S  V  S  T  P  Q  T  L  S  F  L  P    60
Homo      CAGCTGGCACCACCCGCCATGTTCTCTGTCTCAACCCCACAAACCCTTTCCTTCCTCCCC   180
          |  |||               |||||  |  |  ||||  ||  |    |  ||  |||  ||||
Sus       CCACTG-------------CATTCTCACTTCCTGCCCCGCAGCCTTGTCCCCTCCACCCC   157
          P  T               A  F  S  L  P  A  P  Q  P  C  P  L  H  P    52
```

Fig. 5 (*continued*).

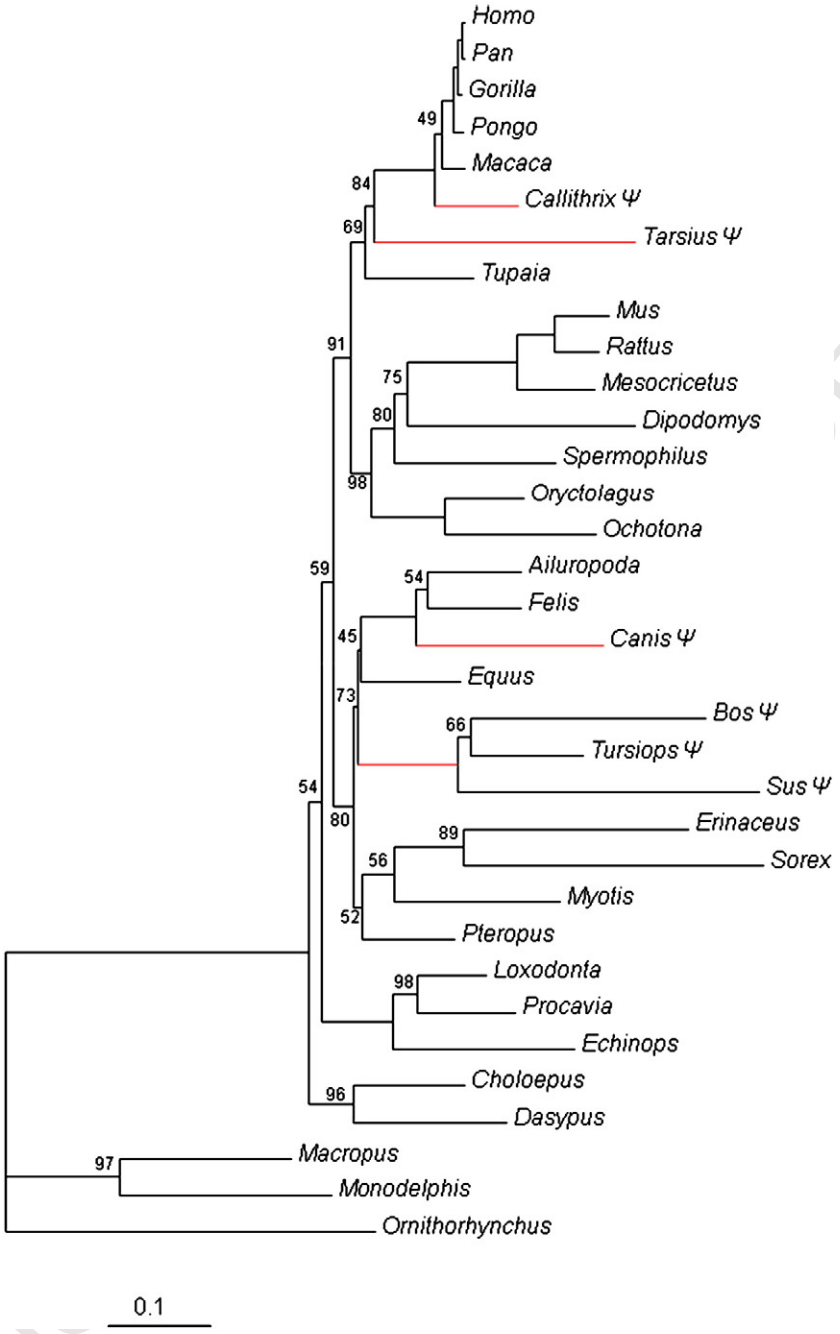


Fig. 6 – Phylogenetic relationship of ZP1 proteins reconstructed with PhyML. The numbers above each node are the bootstrap supports (only the values<99 are shown). The symbol ψ indicates possible pseudogenes. The branches in red indicate the position of the probable pseudogenization event. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

in sperm binding and the induction of the AR [68]. In other species like pig and cow, it was reported that both ZP3 and ZP4 act as receptors of the spermatozoa [69,70].

Evidence suggests that when a gene suffers duplication, the functional divergence of gene copies is a major factor promoting their retention in the genome. So, species with four ZP glycoproteins, like rabbit, the two copies might play a different role and the two genes are necessary for the ZP to play its role correctly.

5. Concluding remarks

In summary, in this study, the cDNA encoding ZP1 has been identified in rabbit (*O. cuniculus*) ovaries. The nucleotide sequence shows a high similarity with the ZP1 of other mammals. Mass spectrometric analysis confirmed the presence of ZP1, ZP2, ZP3 and ZP4 proteins in rabbit ovaries and oocytes. Phylogenetic analysis indicates that

Please cite this article as: Stetson I., et al, Rabbit zona pellucida composition: A molecular, proteomic and phylogenetic approach, J Prot (2012), http://dx.doi.org/10.1016/j.jprot.2012.07.027

the pseudogenization of ZP1 has occurred at least four times during the evolution of mammals. Finally, due to the similar composition and expression pattern, rabbit ZP could be proposed as a suitable experimental model for studying the human ZP and its role during fertilization.

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.jprot.2012.07.027.

Acknowledgments

This work was supported by the Fundación Seneca (04542/GERM/06), MICINN (AGL2009-12512-C02-01-02) and MICINN-AGL08-2283. We thank Dr. Alejandro Torrecillas Sánchez of the Proteomic Service of University of Murcia.

REFERENCES

[1] Yanagimachi R. Mammalian fertilization. In: Knobil E, Neill JD, editors. Physiology of reproduction. New York: Raven Press; 1994. p. 189-317.
[2] Dean J. The enigma of sperm–egg recognition in mice. Soc Reprod Fertil Suppl 2007;63:359-65.
[3] Wassarman PM, Litscher ES. The multifunctional zona pellucida and mammalian fertilization. J Reprod Immunol 2009;83:45-9.
[4] Goudet G, Mugnier S, Callebaut I, Monget P. Phylogenetic analysis and identification of pseudogenes reveal a progressive loss of zona pellucida genes during evolution of vertebrates. Biol Reprod 2008;78:796-806.
[5] Bausek N, Waclawek M, Wolfgang JS, Wohlrab F. The major chicken egg envelope protein ZP1 is different from ZPB and is synthesized in the liver. J Biol Chem 2000;275:28866-72.
[6] Spargo SC, Hope RM. Evolution and nomenclature of zona pellucida gene family. Biol Reprod 2003;68:358-62.
[7] Bleil JD, Wassarman PM. Structure and function of the zona pellucida: identification and characterization of the proteins of the mouse oocyte's zona pellucida. Dev Biol 1980;76:185-202.
[8] Hedrick JL, Wardrip NJ. On the macromolecular composition of the zona pellucida from porcine oocytes. Dev Biol 1987;121:478-88.
[9] Noguchi S, Yonezawa N, Katsumata T, Hashizume K, Kuwayama M, Hamano S, et al. Characterization of the zona pellucida glycoproteins from bovine ovarian and fertilized eggs. Biochim Biophys Acta 1994;1201:7–14.
[10] Hughes DC, Barratt CL. Identification of the true human orthologue of the mouse Zp1 gene: evidence for greater complexity in the mammalian zona pellucida. Biochim Biophys Acta 1999;1447:303-6.
[11] Lefièvre L, Conner SJ, Salpekar A, Olufowobi O, Ashton P, Pavlovic B, et al. Four zona pellucida glycoproteins are expressed in the human. Hum Reprod 2004;19:580-6.
[12] Hoodbhoy T, Joshi S, Boja ES, Williams SA, Stanley P, Dean J. Human sperm do not bind to rat zonae pellucidae despite the presence of four homologous glycoproteins. J Biol Chem 2005;280:12721-31.
[13] Izquierdo-Rico MJ, Jimenez-Movilla M, Llop E, Pérez-Oliva AB, Ballesta J, Gutierrez-Gallego R, et al. Hamster zona pellucida is formed by four glycoproteins. J Proteome Res 2009;8:926-41.
[14] Jiménez-Movilla M, Martínez-Alonso E, Castells MT, Izquierdo-Rico MJ, Saavedra MD, Gutiérrez-Gallego R, et al. Cytochemical and biochemical evidences for a complex tridimensional structure of the hamster zona pellucida. Histol Histopathol 2009;24:599-609.
[15] Goudet G, Mugnier S, Callebaut I, Monget P. Phylogenetic analysis and identification of pseudogenes reveal a progressive loss of zona pellucida genes during evolution of vertebrates. Biol Reprod 2008;78:796-806.
[16] Dunbar BS, Timmons TM, Skinner SM, Prasad SV. Molecular analysis of a carbohydrate antigen involved in the structure and function of zona pellucida glycoproteins. Biol Reprod 2001;65:951-60.
[17] Easton RL, Patankar MS, Lattanzio FA, Leaven TH, Morris HR, Clark GF, et al. Structural analysis of murine zona pellucida glycans. Evidence for the expression of core 2-type O-glycans and the Sd(a) antigen. J Biol Chem 2000;275:7731-42.
[18] von Witzendorff D, Ekhlasi-Hundrieser M, Dostalova Z, Resch M, Rath D, Michelmann HW, et al. Analysis of N-linked glycans of porcine zona pellucida glycoprotein ZPA by MALDI-TOF MS: a contribution to understanding zona pellucida structure. Glycobiology 2005;15:475-88.
[19] von Witzendorff D, Maass K, Pich A, Ebeling S, Kölle S, Kochel C, et al. Characterization of the acidic N-linked glycans of the zona pellucida of prepuberal pigs by a mass spectrometric approach. Carbohydr Res 2009;344:1541-9.
[20] Lee VH, Schwoebel E, Prasad S, Cheung P, Timmons TM, Cook R, et al. Identification and structural characterization of the 75-kDa rabbit zona pellucida protein. J Biol Chem 1993;268:12412-7.
[21] Schwoebel ED, Vandevoort CA, Lee VH, Lo YK, Dunbar BS. Molecular analysis of the antigenicity and immunogenicity of recombinant zona pellucida antigens in a primate model. Biol Reprod 1992;47:857-65.
[22] Harris JD, Hibler DW, Fontenot GK, Hsu KT, Yurewicz EC, Sacco AG. Cloning and characterization of zona pellucida genes and cDNAs from a variety of mammalian species: the ZPA, ZPB and ZPC gene families. DNA Seq 1994;4:361-93.
[23] Prasad SV, Mujtaba S, Lee VH, Dunbar BS. Immunogenicity enhancement of recombinant rabbit 55-kilodalton zona pellucida protein expressed using the baculovirus expression system. Biol Reprod 1995;52:1167-78.
[24] Lee VH. Expression of rabbit zona pellucida-1 messenger ribonucleic acid during early follicular development. Biol Reprod 2000;63:401-8.
[25] Schwoebel E, Prasad S, Timmons TM, Cook R, Kimura H, Niu EM, et al. Isolation and characterization of a full-length cDNA encoding the 55-kDa rabbit zona pellucida protein. J Biol Chem 1991;266:7214-9.
[26] Arias-Álvarez M, García-García RM, Torres-Rovira L, González-Bulnes A, Rebollar PG, Lorenzo PL. Influence of leptin on in vitro maturation and steroidogenic secretion of cumulus–oocyte complexes through JAK2/STAT3 and MEK 1/2 pathways in the rabbit model. Reproduction 2010;139:523-32.
[27] Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic Local Alignment Search Tool. J Mol Biol 1990;215:403-10.
[28] Nielsen H, Engelbrecht J, Brunak S, von Heijne G. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Eng 1997;10:1-6.
[29] Hansen JE, Lund O, Tolstrup N, Gooley AA, Williams KL, Brunak S. NetOglyc: prediction of mucin type O-glycosylation sites based on sequence context and surface accessibility. Glycoconj J 1998;15:115-30.
[30] Gupta R, Brunak S. Prediction of glycosilation across the human proteome and the correlation to protein function. Pac Symp Biocomput 2002:310-22.
[31] Gouy M, Guindon S, Gascuel O. SeaView version 4: a multiplatform graphical user interface for sequence

Please cite this article as: Stetson I., et al, Rabbit zona pellucida composition: A molecular, proteomic and phylogenetic approach, J Prot (2012), http://dx.doi.org/10.1016/j.jprot.2012.07.027

alignment and phylogenetic tree building. Mol Biol Evol 2010;27:221-4.
[32] Guindon S, Gascuel O. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst Biol 2003;52:696-704.
[33] Posada D, Crandall KA. MODELTEST: testing the model of DNA substitution. Bioinformatics 1998;14:817-8.
[34] Bendtsen JD, Nielsen H, von Heijne G, Brunak S. Improved prediction of signal peptides: SignalP 3.0. J Mol Biol 2004;340:783-95.
[35] Ducker P, Brunak S, Blom N. Prediction of proprotein convertase cleavage sites. Protein Eng Des Sel 2004;17:107-12.
[36] Boja ES, Hoodbhoy T, Fales HM, Dean J. Structural characterization of native mouse zona pellucida proteins using mass spectrometry. J Biol Chem 2003;278:34189-202.
[37] Boja ES, Hoodbhoy T, Garfield M, Fales HM. Structural conservation of mouse and rat zona pellucida glycoproteins. Probing the native rat zona pellucida proteome by mass spectrometry. Biochemistry 2005;44:16445-60.
Q17 [38] Epifano O, Liang LF, Familiari M, Moos MC, Dean J. Coordinate expression of the three zona pellucida genes during mouse oogenesis. Development 1995;121:1947-56.
[39] Zhao M, Boja ES, Hoodbhoy T, Nawrocki J, Kaufman JB, Kresge N, et al. Mass spectrometry analysis of recombinant human ZP3 expressed in glycosylation-deficient CHO cells. Biochemistry 2004;43:12090-104.
[40] Chakravarty S, Kadunganattil S, Bansal P, Sharma RK, Gupta SK. Relevance of glycosylation of human zona pellucida glycoproteins for their binding to capacitated human spermatozoa and subsequent induction of acrosomal exocytosis. Mol Reprod Dev 2008;75:75-88.
[41] Dean J. The enigma of sperm–egg recognition in mice. Soc Reprod Fertil Suppl 2007;63:359-65.
[42] Visconti PE, Florman HM. Mechanisms of sperm–egg interactions: between sugars and broken bonds. Sci Signal 2010;3:35.
[43] Clark GF. Molecular models for mouse sperm–oocyte binding. Glycobiology 2011;21:3-5.
[44] Gahlay G, Gauthier L, Baibakov B, Epifano O, Dean J. Gamete recognition in mice depends on the cleavage status of an egg's zona pellucida protein. Science 2010;329:216-9.
[45] Lee VH, Dunbar BS. Developmental expression of the rabbit 55-kDa zona pellucida protein and messenger RNA in ovarian follicles. Dev Biol 1993;155:371-82.
[46] Dunbar BS, Avery S, Lee V, Prasad S, Schwahn D, Schwoebel E, et al. The mammalian zona pellucida: its biochemistry, immunochemistry, molecular biology, and developmental expression. Reprod Fertil Dev 1994;6:331-47.
[47] Hinsch KD, Hinsch E, Meinecke B, Töpfer-Petersen E, Pfisterer S, Schill WB. Identification of mouse ZP3 protein in mammalian oocytes with antisera against synthetic ZP3 peptides. Biol Reprod 1994;51:193-204.
[48] Grootenhuis AJ, Philipsen HL, de Breet-Grijsbach JT, van Duin M. Immunocytochemical localization of ZP3 in primordial follicles of rabbit, marmoset, rhesus monkey and human ovaries using antibodies against human ZP3. J Reprod Fertil Suppl 1996;50:43-54.
[49] Martínez ML, Fontenot GK, Harris JD. The expression and localization of zona pellucida glycoproteins and mRNA in cynomolgus monkeys (*Macaca fascicularis*). J Reprod Fertil Suppl 1996;50:35-41.
[50] Gook DA, Edgar DH, Borg J, Martic M. Detection of zona pellucida proteins during human folliculogenesis. Hum Reprod 2008;23:394-402.
[51] Bleil JD, Wassarman PM. Synthesis of zona pellucida proteins by denuded and follicle-enclosed mouse oocytes during culture in vitro. Proc Natl Acad Sci U S A 1980;77:1029-33.
[52] Fléchon JE, Pavlok A, Kopecný V. Dynamics of zona pellucida formation by the mouse oocyte. An autoradiographic study. Biol Cell 1984;51:403-6.
[53] Epifano O, Liang LF, Familiari M, Moos MC, Dean J. Coordinate expression of the three zona pellucida genes during mouse oogenesis. Development 1995;121:1947-56.
[54] Sinowatz F, Kölle S, Töpfer-Petersen E. Biosynthesis and expression of zona pellucida glycoproteins in mammals. Cells Tissues Organs 2001;168:24-35.
[55] El Mestrah M, Castle PE, Borrosa G, Kan FW. Subcellular distribution of ZP1, ZP2, and ZP3 glycoproteins during folliculogenesis and demonstration of their topographical disposition within the zona matrix of mouse ovarian oocytes. Biol Reprod 2002;66:866-76.
[56] Scobie GA, Kerr LE, MacDuff P, Aitken RJ. Cloning, sequencing and site of origin of the rat sperm receptor protein, ZP3. Zygote 1999;7:27-35.
[57] Izquierdo-Rico MJ, Gimeno L, Jiménez-Cervantes C, Ballesta J, Avilés M. Biosynthesis of hamster zona pellucida is restricted to the oocyte. Theriogenology 2011;75:463-72.
[58] Conner SJ, Hughes DC. Analysis of fish ZP1/ZPB homologous genes—evidence for both genome duplication and species-specific amplification models of evolution. Reproduction 2003;126:347-52.
[59] Gupta SK, Bhandari B. Acrosome reaction: relevance of zona pellucida glycoproteins. Asian J Androl 2011;13:97–105.
[60] Wassarman PM. Zona pellucida glycoproteins. Annu Rev Biochem 1988;57:415-42.
[61] Rankin T, Talbot P, Lee E, Dean J. Abnormal zonae pellucidae in mice lacking ZP1 result in early embryonic loss. Development 1999;126:3847-55.
[62] Ganguly A, Bukovsky A, Sharma RK, Bansal P, Bhandari B, Gupta SK. In humans, zona pellucida glycoprotein-1 binds to spermatozoa and induces acrosomal exocytosis. Hum Reprod 2010;25:1643-56.
[63] Ganguly A, Bansal P, Gupta T, Gupta SK. 'ZP domain' of human zona pellucida glycoprotein-1 binds to human spermatozoa and induces acrosomal exocytosis. Reprod Biol Endocrinol 2010;8:110 [11].
[64] Topper EK, Kruijt L, Calvete J, Mann K, Töpfer-Petersen E, Woelders H. Identification of bovine zona pellucida glycoproteins. Mol Reprod Dev 1997;46:344-50.
[65] Govind CK, Hasegawa A, Koyama K, Gupta SK. Delineation of a conserved B cell epitope on bonnet monkey (*Macaca radiata*) and human zona pellucida glycoprotein-B by monoclonal antibodies demonstrating inhibition of sperm–egg binding. Biol Reprod 2000;62:67-75.
[66] Chakravarty S, Suraj K, Gupta SK. Baculovirus-expressed recombinant human zona pellucida glycoprotein-B induces acrosomal exocytosis in capacitated spermatozoa in addition to zona pellucida glycoprotein-C. Mol Hum Reprod 2005;11:365-72.
[67] Prasad SV, Wilkins B, Skinner SM, Dunbar BS. Evaluating zona pellucida structure and function using antibodies to rabbit 55 kDa ZP protein expressed in baculovirus expression system. Mol Reprod Dev 1996;43:519-29.
[68] Chiu PC, Wong BS, Chung MK, Lam KK, Pang RT, Lee KF, et al. Effects of native human zona pellucida glycoproteins 3 and 4 on acrosome reaction and zona pellucida binding of human spermatozoa. Biol Reprod 2008;79:869-77.
[69] Yurewicz EC, Sacco AG, Gupta SK, Xu N, Gage DA. Hetero-oligomerization-dependent binding of pig oocyte zona pellucida glycoproteins ZPB and ZPC to boar sperm membrane vesicles. J Biol Chem 1998;273:7488-94.
[70] Kanai S, Yonezawa N, Ishii Y, Tanakura M, Nakano M. Recombinant bovine zona pellucida glycoproteins ZP3 and ZP4 coexpressed in Sf9 cells form a sperm-binding active hetero-complex. FEBS J 2007;274:5390-405.

Please cite this article as: Stetson I., et al, Rabbit zona pellucida composition: A molecular, proteomic and phylogenetic approach, J Prot (2012), http://dx.doi.org/10.1016/j.jprot.2012.07.027