CSIRO PUBLISHING

Reproduction, Fertility and Development
https://doi.org/10.1071/RD16519

Composition of marsupial zona pellucida: a molecular and phylogenetic approach

Carla Moros-Nicolás[A,G], *Pascale Chevret*[B], *Ma José Izquierdo-Rico*[A], *William V. Holt*[C], *Daniela Esteban-Díaz*[A], *Manel López-Béjar*[D], *Eva Martínez-Nevado*[E], *Maria A. Nilsson*[F], *José Ballesta*[A] *and Manuel Avilés*[A,G]

[A]Department of Cell Biology and Histology, Faculty of Medicine, University of Murcia, Campus Mare Nostrum and IMIB, Murcia 30100, Spain.
[B]Laboratoire de Biométrie et Biologie Evolutive, UMR5558, CNRS, Université de Lyon, Université Claude Bernard Lyon 1, Villeurbanne 69622, France.
[C]Institute of Zoology, Regent's Park, London NW1 4RY, UK.
[D]Department of Animal Health and Anatomy, Universitat Autònoma de Barcelona, Barcelona 08193, Spain.
[E]Veterinary Department, Zoo-Aquarium Madrid, Casa de Campo s/n., Madrid 28011, Spain.
[F]Senckenberg Biodiversity and Climate Research Centre, Senckenberg Gesellschaft für Naturforschung, Senckenberganlage 25, Frankfurt am Main D-60325, Germany.
[G]Corresponding authors. Emails: carla.moros@um.es; maviles@um.es

Abstract. The zona pellucida (ZP) is an extracellular matrix that surrounds mammalian oocytes. In eutherians it is formed from three or four proteins (ZP1, ZP2, ZP3, ZP4). In the few marsupials that have been studied, however, only three of these have been characterised (ZP2, ZP3, ZP4). Nevertheless, the composition in marsupials may be more complex, since a duplication of the *ZP3* gene was recently described in one species. The aim of this work was to elucidate the ZP composition in marsupials and relate it to the evolution of the ZP gene family. For that, an *in silico* and molecular analysis was undertaken, focusing on two South American species (gray short-tailed opossum and common opossum) and five Australian species (brushtail possum, koala, Bennett's wallaby, Tammar wallaby and Tasmanian devil). This analysis identified the presence of *ZP1* mRNA and mRNA from two or three paralogues of *ZP3* in marsupials. Furthermore, evidence for *ZP1* and *ZP4* pseudogenes in the South American subfamily Didelphinae and for *ZP3* pseudogenes in two marsupials is provided. In conclusion, two different composition models are proposed for marsupials: a model with four proteins (ZP1, ZP2 and ZP3 (two copies)) for the South American species and a model with six proteins (ZP1, ZP2, ZP3 (three copies) and ZP4) for the Australasian species.

Additional keywords: evolution, fertilisation, oocyte, sperm–egg interaction.

Received 23 December 2016, accepted 20 September 2017, published online xx xxxxx xxxx

Introduction

During fertilisation, female and male gametes are involved in a series of complex events that allow the formation of a zygote. In marsupials, preimplantation embryos are enclosed by three acellular envelopes: the zona pellucida (ZP), a mucoid coat and an outer shell coat (Selwood 2000). Furthermore, in some marsupials there is a sub-zonal, extra-cellular matrix coat (Selwood 2000). In eutherian mammals, the structure and composition of the ZP has been studied in detail, but this is not the case in most marsupials. The ZP plays a key role in species-specific binding between gametes in most species, as well as in the induction of the acrosome reaction, polyspermy block and protection of the preimplantation embryo (Yanagimachi 1994; Dean 2004, 2007; Wassarman and Litscher 2009).

Eutherians can be classified into three categories according to their ZP composition: (1) species with a ZP formed from ZP1, ZP2 and ZP3, in which *ZP4* is a pseudogene (to date, only the house mouse; Bleil and Wassarman 1980; Lefièvre *et al.* 2004; Evsikov *et al.* 2006; Goudet *et al.* 2008), (2) species with three proteins, in which *ZP1* is a pseudogene (pig, cow, dog, common marmoset, dolphin, tarsier, Antarctic fur seal, Weddell seal and fox; Hedrick and Wardrip 1987; Noguchi *et al.* 1994; Goudet

et al. 2008; Stetson *et al.* 2012; Moros-Nicolás *et al.* 2017) and (3) species with four proteins (ZP1, ZP2, ZP3 and ZP4), such as human, rat, hamster, rabbit and cat (Hughes and Barratt 1999; Lefièvre *et al.* 2004; Hoodbhoy *et al.* 2005; Izquierdo-Rico *et al.* 2009; Jiménez- Movilla *et al.* 2009; Stetson *et al.* 2012, 2015). In 1994, Harris *et al.* (1994) proposed the terms ZPA, ZPB and ZPC for the different ZP proteins in eutherians; these same terms were later used by other authors in relation to marsupials (Haines *et al.* 1999; Voyle *et al.* 1999; Breed *et al.* 2002). However, identification of a fourth protein in some species of eutherians led some authors to use the nomenclature with numbers (Mate and McCartney 1998; McCartney and Mate 1999; Mate *et al.* 2003; McCartney *et al.* 2007; Au *et al.* 2008; Cui *et al.* 2010). Since this nomenclature is clearer, the same model has been followed in this manuscript.

The composition and functions of the marsupial ZP are less well characterised than those of eutherian mammals (Breed *et al.* 2002). It has conventionally been believed that the marsupial ZP is formed of three glycoproteins (McCartney and Mate 1999; Selwood 2000; Mate *et al.* 2003; Cui *et al.* 2010), which are similar to those of eutherian mammals (Selwood 2000). However, the genes or proteins have been characterised in very few species. Only in one species, the brushtail possum (*Trichosurus vulpecula*), has the open reading frame (ORF) of three cDNAs encoding ZP genes been cloned and sequenced: *ZP2*, *ZP3* and *ZP4* (Mate and McCartney 1998; Haines *et al.* 1999; McCartney and Mate 1999; Voyle *et al.* 1999; Mate *et al.* 2003). In other species, partial sequences of these genes have been described. For instance, *ZP2* has been partially amplified in fat-tailed dunnart (*Sminthopsis crassicaudata*; Voyle *et al.* 1999), Tammar wallaby (*Macropus eugenii*; McCartney *et al.* 2007) and in the stripe-faced dunnart (*Sminthopsis macroura*; Au *et al.* 2008); partial *ZP3* sequences have been described in Eastern grey kangaroo (*Macropus giganteus*), fat-tailed dunnart (*Sminthopsis crassicaudata*), gray short-tailed opossum (*Monodelphis domestica*), Bennett's wallaby (*Macropus rufogriseus*), swamp wallaby (*Wallabia bicolor*) and Tammar wallaby (*Macropus eugenii*; McCartney *et al.* 2007) and finally, a partial *ZP4* sequence has been described in Tammar wallaby (*Macropus eugenii*; McCartney *et al.* 2007).

The composition may be more complex than previously suspected, since recently a bioinformatic analysis identified two copies of the *ZP3* gene in one marsupial species, the gray short-tailed opossum (*Monodelphis domestica*; Meslin *et al.* 2012).

The aim of the present study was to focus in depth on the ZP composition in marsupials and to study the evolution of the ZP gene family. To this end, the ZP composition was analysed in five species from Australia and two species from South America. These were chosen as we had access to sequence data and/or biological material, and we specifically looked for pseudogenisation or duplication events that may have affected the ZP composition in this group of mammals.

Materials and methods

In the present study, two different types of analysis were performed: an *in silico* analysis and a molecular analysis using genomic DNA or RNA. The use of DNA or RNA depended on the availability of tissue samples from different marsupials (see Fig. S1, available as Supplementary Material to this paper).

Molecular analysis

DNA isolation

DNA was obtained from tissues of three marsupials: common opossum (*Didelphis marsupialis*), gray short-tailed opossum (*Monodelphis domestica*) and koala (*Phascolarctos cinereus*). DNA was extracted from ethanol-preserved tissues in common opossum and gray short-tailed opossum, which were obtained from the collection of Preserved Mammalian Tissues of the Institut des Sciences de l'Evolution of Montpellier (France), and from koala's blood preserved with ethylenediamine tetraacetic acid (EDTA) at 4°C, which was donated by the Zoo-Aquarium of Madrid (Spain). Total DNA was extracted using a QIAamp DNA Mini Kit (Qiagen) following the manufacturer's recommendations.

Obtaining Bennett's wallaby ovarian RNA and cDNA synthesis

The use of animals was approved by the Chair of the Zoological Society of London Ethics Committee. Ovaries were obtained postmortem from five female adult wallabies from Whipsnade Zoo (Dunstable, Bedfordshire, UK) that had been culled during the course of a regular population control program. Immediately after collection, the ovaries ($n = 10$) were stored in RNAlater (Ambion) at −80°C until use. Total RNA was isolated using an RNAqueous kit (Ambion) according to the manufacturer's instructions. First-strand cDNA was synthesised using a SuperScript First-Strand Synthesis System kit for reverse transcription polymerase chain reaction (RT-PCR; Invitrogen-Life Technologies) according to the manufacturer's instructions.

Amplification of the open reading frame of Bennett's wallaby ZP genes

Bennett's wallaby ZP genes were completely (*ZP1*, *ZP2*, *ZP3-a*, *ZP3-b* and *ZP4*) or partially (*ZP3-c*) amplified using PCR and specific primers (see Table S1, available as Supplementary Material to this paper), which were designed according to the predicted cDNA sequences of the Tammar wallaby (*Macropus eugenii*), its closest relative in the Ensembl database (http://ensembl.org, verified 17 October 2017; accession numbers were: ENSMEUT00000000742 (*ZP1*), ENSMEUT00000013237 (*ZP2*), ENSMEUT00000000436 (*ZP3*) and ENSMEUT00000012954 (*ZP4*)). Furthermore, a Basic Local Alignment Search Tool (BLAST) search on the Tammar wallaby genome identified three copies of *ZP3*; additional sets of primers were designed to amplify these copies (*ZP3-a* and *ZP3-b*). Glyceraldehyde-3-phosphate dehydrogenase (*GAPDH*) primers designed from Tammar wallaby sequence (accession number: EF654515) were used as a positive control.

PCR amplifications were performed using 2 μL target cDNA, 0.5 μg each primer, 200 μM each dNTP and 1 IU Taq DNA polymerase (Fermentas). PCR was performed using an initial denaturation cycle of 3 min at 95°C, followed by 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 analysed by electrophoresis on 1.5% agarose gels. Amplicons were carefully excised from the agarose gels and purified with the QIAquick Gel Extraction Kit Protocol (Qiagen) according to the manufacturer's protocol or directly purified with the DNA Clean and Concentrator TM-5 (Zymo) according to the manufacturer's instructions. Amplicons were automatically sequenced using a 3500 Genetic Analyzer (Applied Biosystems).

Amplification of Exons 1 and 2 of the ZP4 *gene in common opossum, gray short-tailed opossum and koala*

DNA from three marsupials – two South American, common opossum (*Didelphis marsupialis*) and gray short-tailed opossum (*Monodelphis domestica*) and one Australasian, koala (*Phascolarctos cinereus*) – was used to amplify Exons 1 and 2 of the *ZP4* gene using the primers shown in Table S1. Opossum primers were designed using the *ZP4* gene from gray short-tailed opossum genome (version monDom5) for guidance, while koala primers were designed based on the predicted GenBank database cDNA sequence of the Tasmanian devil (*Sarcophilus harrisii*) *ZP4* (XM_003767794).

PCR amplifications were performed using 200–500 ng target DNA, 0.5 μg each primer, 200 μM each dNTP and 1 IU Taq DNA polymerase (Fermentas). PCR amplifications, sample purification and sequencing were performed as described above.

Bioinformatic analysis

New RNA sequences

Sequences were analysed to determine the degree of conservation with other known sequences using the BLAST program (www.ncbi.nlm.nih.gov/blast/, verified 17 October 2017). Direct comparison between two sequences was made with the ALIGN program and multiple sequence alignment was performed using Clustal Omega (www.ebi.ac.uk/Tools/msa/clustalo/, verified 17 October 2017). The amino acid sequences were analysed using the following software packages: SignalP (http://www.cbs.dtu.dk/services/SignalP/, verified 18 October 2017) and Smart Genome (www.smart.embl-heidelberg.de, verified 17 October 2017) to predict the signal peptide and different domains and NetNglyc (http://www.cbs.dtu.dk/services/NetNGlyc/, verified 18 October 2017) to predict potential N-linked glycosylation sites. The theoretical protein molecular weights and mature protein molecular weights were calculated using PeptideMass from ExPASy (www.web.expasy.org/peptide_mass/, verified 17 October 2017).

Phylogenetic analysis of marsupial ZP sequences

To complete the dataset available for marsupials, similarity searches were performed using BLAST or BLAST-like alignment tool (BLAT) in GenBank and in the genome data of gray short-tailed opossum (*Monodelphis domestica*), Tammar wallaby (*Macropus eugenii*) and Tasmanian devil (*Sarcophilus harrisii*) in Ensembl (Release 87; Table 1). Our RNA sequences of *ZP1*, *ZP2*, *ZP3-a*, *ZP3-b*, *ZP3-c* and *ZP4* of Bennett's wallaby (*Macropus rufogriseus*) were used as query sequences. The Ensembl gene predictions were checked manually to detect annotation errors, especially close to splicing sites. In some cases, further genes or exons missing from the Ensembl predictions were predicted. All the sequences were aligned using Muscle implemented in SeaView (www.doua.prabi.fr/software/seaview, verified 17 October 2017) and the alignment was refined visually to determine exon limits. The new sequences were also checked to ensure that they corresponded to a syntenic region of the corresponding chromosome. The sequences obtained for marsupials were completed with sequences retrieved from GenBank for other mammals, birds and amphibians (*ZP3* only). For the placentals, birds and amphibians, only RNA sequences were retrieved from GenBank (Table S2). For the platypus, for which no RNA sequences were available in GenBank, genomic predictions were retrieved from Ensembl and checked using the approach applied to the marsupial sequences (Table S2). Only the exonic portions were kept for the phylogenetic analysis. As *ZP3* sequences were very divergent from the other ZP sequences, the phylogenetic analysis was performed on two datasets. The first comprised *ZP1*, *ZP2* and *ZP4*, with *ZPAX* sequences used to root the phylogeny; the second comprised *ZP3* sequences of mammals and birds, with amphibian sequences serving as outgroups. Sequences were aligned using Muscle implemented in Seaview and the alignment was refined visually. The first dataset comprised 70 sequences and 2164 positions, the second 52 sequences and 1344 positions (Tables 1 and S2), after the removal of sites with too many indels (limit set at 2/3). Phylogenetic reconstructions were performed using maximum likelihood (ML) and Bayesian inference (BI) with PhyML (Guindon *et al.* 2010) and MrBayes Version 3.2 (Ronquist *et al.* 2012) respectively. Node robustness was estimated using 1000 bootstrap (BP) replicates in ML analyses, while posterior probabilities (PP) were used in BI analyses. The appropriate model of evolution (GTR+I+G) was determined using Akaike information criterion (AIC) and jModeltest software (Posada and Crandall 1998) for each dataset. For Bayesian inference, two Markov chain Monte Carlo (MCMC) analyses were run independently for 10 000 000 generations. Trees were sampled each 500 generations. The burn-in was graphically determined with Tracer Version 1.6 (Rambaut *et al.* 2014) and the average s.d. of split frequencies was also checked to ensure they remained <0.01 after the burn-in threshold. Ten percent of the trees were discarded as burn-in for each dataset.

Results

Evidence of the presence of six ZP transcripts in Bennett's wallaby ovaries

Transcript analysis of ZP1, ZP2 *and* ZP4

Full-length Bennett's wallaby *ZP1*, *ZP2* and *ZP4* were amplified and characterised. Sequences were submitted to GenBank (*ZP1*, KC954398; *ZP2*, KP057615 and *ZP4*, KC954399). The ORF of *ZP1* encodes a 616-amino acid polypeptide (Fig. S2) with a theoretical molecular mass of 69.16 kDa; *ZP2* encodes a 712-amino acid polypeptide, with a theoretical molecular mass of 79.28 kDa (Fig. S3) and *ZP4* encodes a 525-amino acid polypeptide (Fig. S4) with a theoretical molecular mass of 58.40 kDa.

These proteins possess the archetypal 'ZP module' (Jovine *et al.* 2005), a signature domain rich in cysteine residues.

Table 1. List of accession numbers (from GenBank and/or ENSEMBL) of the marsupial sequences used in the phylogenetic analysis

Common names are indicated within brackets

Species	*ZP1*	*ZP2*	*ZP4*	*ZP3-a*	*ZP3-b*	*ZP3-c*
Macropus rufogriseus (Bennett's wallaby)	This study: KC954398	This study: KP057615	This study: KC954399	This study: KP057617	This study: KP057616	This study: KP057618
Macropus eugenii (Tammar wallaby)	ENSMEUT00000000742/ GeneScaffold_2489	ENSMEUT00000013237/ GeneScaffold_1912	ENSMEUT00000012954/ GeneScaffold_3106	Scaffold34246	Scaffold281138 + Scaffold51844	ENSMEUT00000000436/ Scaffold47341
Monodelphis domestica (gray short-tailed opossum)	XM_007497604–5-6–7/ ENSMOD-T00000035038/ Chromosome 5	XM_001377066/ ENSMOD-T00000009148/ Chromosome6			XM_007504383/ ENSMOD-T00000017774/ Chromosome 8	XM_007485985/ ENSMOD-T00000017391/ ENSMOD-T00000041947/ Chromosome 2
	ENSMODG00000011553 (pseudogene)		Chromosome 2 (pseudogene)	Chromosome 1 (pseudogene)	Chromosome 2 (pseudogene)	Chromosome 5 (two pseudogenes)
Sarcophilus harrisii (Tasmanian devil)	XM_003774026/ ENSSHA-T00000010914/ GL864905	XM_003763031/ ENSSHA-T00000007293/ GL842571	XM_003767794/ ENSSHA-T00000010737/ GL856801	XM_003759001/ ENSSHA-T00000004209/ GL835290	XM_003771645/ ENSSHA-T00000020027/ GL861656 GL841557, GL841204 (pseudogenes)	XM_012549947/ ENSSHA-T00000013822/ GL857013
Sminthopsis crassicaudata (fat-tailed dunnart)		AF263015				
Sminthopsis macroura (stripe-faced dunnart)		EF093503				
Trichosurus vulpecula (common brushtail possum)		AF079525	AF263013	AF079524		

Upstream of the ZP module, a trefoil domain, also rich in cysteine amino acids, is present in ZP4. A basic amino acid may serve as a consensus furin cleavage site, which is followed by a transmembrane domain (TMD) and a cytoplasmic tail. The molecular masses of the putative mature proteins were estimated to be 54.03 kDa, 67.26 kDa and 49.29 kDa for ZP1, ZP2 and ZP4 respectively. Three, nine and five potential N-glycosylation sites (Asn-X-S/T) were predicted in mature ZP1, ZP2 and ZP4 proteins respectively (Figs S2, S3 and S4).

Transcript analysis of ZP3

Three different *ZP3* copies were amplified in Bennett's wallaby. We propose to name these copies *ZP3-a*, *ZP3-b* and *ZP3-c*. This nomenclature has been previously used by other authors for a similar case (Bhat *et al.* 2014). The sequences were submitted to GenBank with the accession numbers (*ZP3-a*, KP057617; *ZP3-b*, KP057616 and *ZP3*-c, KP057618). Full-length *ZP3-a* and *ZP3-b* were amplified, whereas *ZP3-c* mRNA was partially amplified.

Each sequence has the archetypal ZP module, being a region rich in cysteine amino acids. A dibasic motif – part of, but distinct from, the proprotein convertase (furin) cleavage site – was present in ZP3-a and ZP3-c; however, these four amino acids were not identified in ZP3-b, where only two different amino acids at 366–367 position were identified (Fig. 1). A hydrophobic patch was also partially conserved in the three different Bennett's wallaby ZP3 proteins (Fig. 1). The potential N-glycosylation sites varied among the three copies, being three in ZP3-a, four in ZP3-b and five in ZP3-c (Fig. 1). The identity among the three copies varied between 50 and 52%.

Thus, full-length Bennett's wallaby *ZP1*, *ZP2* and *ZP4* mRNA were obtained from the total RNA prepared from wallaby ovaries, as well as sequences of three different copies of *ZP3* (*ZP3-a*, *ZP3-b* and *ZP3-c*). Therefore, the presence of six transcripts in the ovary of this species was demonstrated.

ZP4 *pseudogenisation in marsupials*

In the South American marsupial *Monodelphis domestica*, the BLAST search revealed a sequence similar to *ZP4* on Chromosome 2. This sequence presented several indels and stop codons along its sequence, moreover the start codon (ATG) was replaced by an AAG. This strongly indicates that it is a pseudogene (Figs 2 and S5). By contrast, in the Australasian marsupials analysed, *Trichosurus vulpecula*, *Macropus eugenii* and *Sarcophilus harrisii*, the *ZP4* sequences included a start codon (ATG) but no stop codons within their sequences (Fig. 2).

In *Didelphis marsupialis*, the DNA study showed an ATT instead of an initiation codon, a large indel and five stop codons. In *Monodelphis domestica*, the molecular analysis showed the presence of one stop codon and a large indel, as seen in the *in silico* analysis (Fig. 2). In *Phascolarctos cinereus*, no stop codons were detected in its sequence (Fig. 2).

In summary, these analyses indicate that pseudogenisation seems restricted to the South American subfamily Didelphinae, whereas all the Australian marsupials examined so far probably express *ZP4* mRNA (Fig. 2).

Phylogenetic analysis of ZP sequences of marsupials

Bayesian phylogenies of two different datasets are presented in Fig. 3 (for *ZP1*, *ZP2* and *ZP4*) and Fig. 4 (for *ZP3*) with sequences retrieved from GenBank and Ensembl databases (Tables 1 and S2).

For the first dataset (*ZP1*, *ZP2* and *ZP4*), the monophyly of each of the ZP genes is robustly (*ZP2* and *ZP1*, PP = 1, BP ≥ 95) or moderately (*ZP4*, PP = 0.85, BP = 66) supported. *ZP2* is the most ancient one, whereas *ZP1* and *ZP4* are the result of a younger duplication (Fig. 3). The three ZP sequences seem to be functional in Tammar wallaby (*Macropus eugenii*), Bennett's wallaby (*Macropus rufogriseus*) and Tasmanian devil (*Sarcophilus harrisii*), but only *ZP1* and *ZP2* seem to be functional in gray short-tailed opossum (*Monodelphis domestica*). One pseudogene of *ZP1* and one of *ZP4* were also found in *Monodelphis domestica*. The pseudogenisation is recent for *ZP1*, where this pseudogene is closely related to the putative functional copy of *ZP1*, whereas for *ZP4* no evidence of another functional copy was found.

The phylogenetic tree reconstructed with all the *ZP3* sequences found in the databases and with the new ones described in this work (*ZP3-a* and *ZP3-b*) indicates that there are clearly three distinct copies of the *ZP3* gene in marsupials and several pseudogenes (Fig. 4*a*). The three copies seem to be functional in the genome of the Australasian marsupials, Tasmanian devil (*Sarcophilus harrisii*), Tammar wallaby (*Macropus eugenii*) and Bennett's wallaby (*Macropus rufogriseus*), whereas in gray short-tailed opossum (*Monodelphis domestica*) we found only functional copies of *ZP3-b* and *ZP3-c* and a non-functional copy of *ZP3-a*, characterised by the presence of four stop codons in its sequence (Fig. S6). *ZP3-b* is the most ancient copy and *ZP3-a* and *ZP3-c* are the result of a younger duplication (PP = 0.99, BP = 81). *ZP3-a* and *ZP3-b* were also found in the platypus genome, *ZP3-a* in birds and *ZP3-c* in placentals. *ZP3-a* and *ZP3-b* of marsupials and platypus form two robustly supported groups (PP = 1, BP = 99 for *ZP3-a*, PP = 1, BP = 100 for *ZP3-b*). *ZP3-a* and *ZP3-c* are robustly associated with bird *ZP3* (PP = 1, BP = 93) and placentals *ZP3* (PP = 1, BP = 100) respectively (Fig. 4*a*). Pseudogenisation of *ZP3* copies appeared several times in two marsupial genomes: in *Sarcophilus harrissi* two pseudogenes for *ZP3-b* were found, whereas in *Monodelphis domestica* one pseudogene for *ZP3-a* and *ZP3-b* and two pseudogenes for *ZP3-c* were found. Based on this phylogeny, it could be hypothesised that at least two duplications took place during the evolution of *ZP3* in amniotes (Fig. 4*b*), the first leading to *ZP3-b* and the second to *ZP3-a* and *ZP3-c*. *ZP3-a* and *ZP3-b* were probably lost in placental mammals and *ZP3-b* and *ZP3-c* in birds.

Our analysis indicates that the evolution of *ZP3* is more complex than the evolution of the other ZP genes. Pseudogenisation is more frequent in *Monodelphis domestica* than in the other marsupials analysed, especially in the case of *ZP3*, whereas no pseudogenisation was found for *ZP2* in marsupials.

Discussion

Marsupial ZP may be formed of four to six proteins, depending on the species

Until now it has been believed that marsupial ZP consists of three proteins (McCartney and Mate 1999; Selwood 2000;

Fig. 1. Comparison of amino acid sequences of Bennett's wallaby ZP3-a, ZP3-b and ZP3-c with mouse and rat ZP3. The deduced amino acid sequences of wallaby ZP3 copies were aligned with the ZP3 sequences of the other species using ClustalW. The accession numbers of the sequences used are as follows: Bennett's wallaby ZP3-a (KP057617), ZP3-b (KP057616) and ZP3-c (KP057618), mouse ZP3 (NP_035906) and rat ZP3 (NP_446214). Identical amino acids are marked by an asterisk (*). Colon (:) represents conserved residues and dot (.) represents semi-conserved residues. The ZP module is shown in red. The consensus furin cleavage-site is underlined. The transmembrane domain is in orange. The cysteine residues are in green. The potential N-glycosylation sites are in purple. Colour version of this figure is available online.

Mate *et al.* 2003; Cui *et al.* 2010). The brushtail possum (*Trichosurus vulpecula*) is the only species where three mRNAs encoding ZP proteins have been reported (*ZP2*, *ZP3* and *ZP4*; Mate and McCartney 1998, 2003; Haines *et al.* 1999; McCartney and Mate 1999; Voyle *et al.* 1999), suggesting that its ZP is formed of three proteins with a composition similar to that of pig, cow, dog and fox (Hedrick and Wardrip 1987; Noguchi *et al.* 1994; Goudet *et al.* 2008; Moros-Nicolás *et al.* 2017). Nevertheless, using bioinformatics, Meslin *et al.* (2012) identified the presence of a *ZP1* gene and a duplication of the *ZP3* gene in one marsupial species, the gray short-tailed opossum (*Monodelphis domestica*). It is surprising that *ZP1* mRNA or protein has not

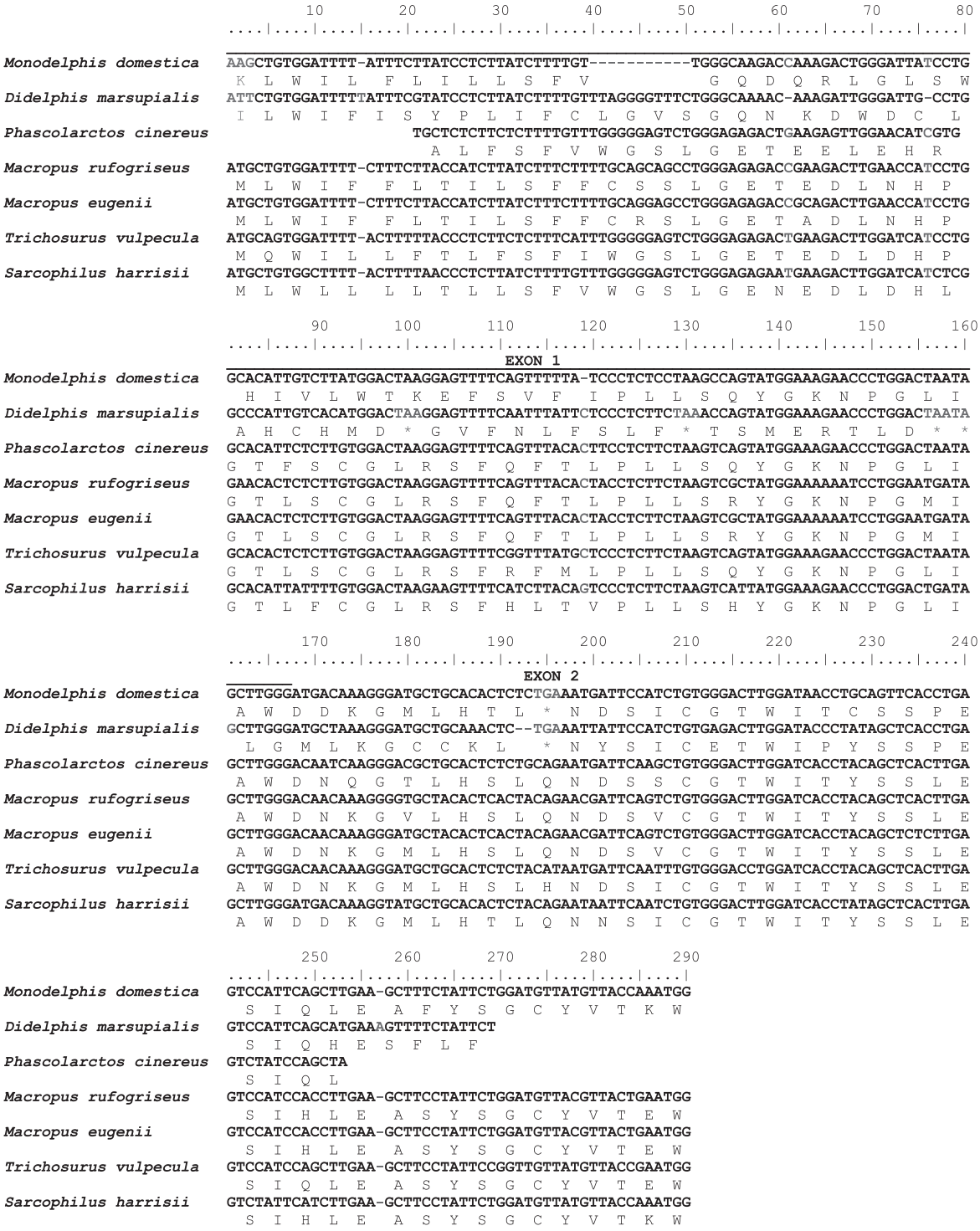

Fig. 2. Alignment and translation of Exons 1 and 2 of ZP4 of seven marsupials. The stop codons and sites with indels are indicated in red and the modified first codon observed in the sequence of *Monodelphis domestica* and *Didelphis marsupialis* are in blue. For the five other marsupials, there is no evidence of ZP4 pseudogenisation.

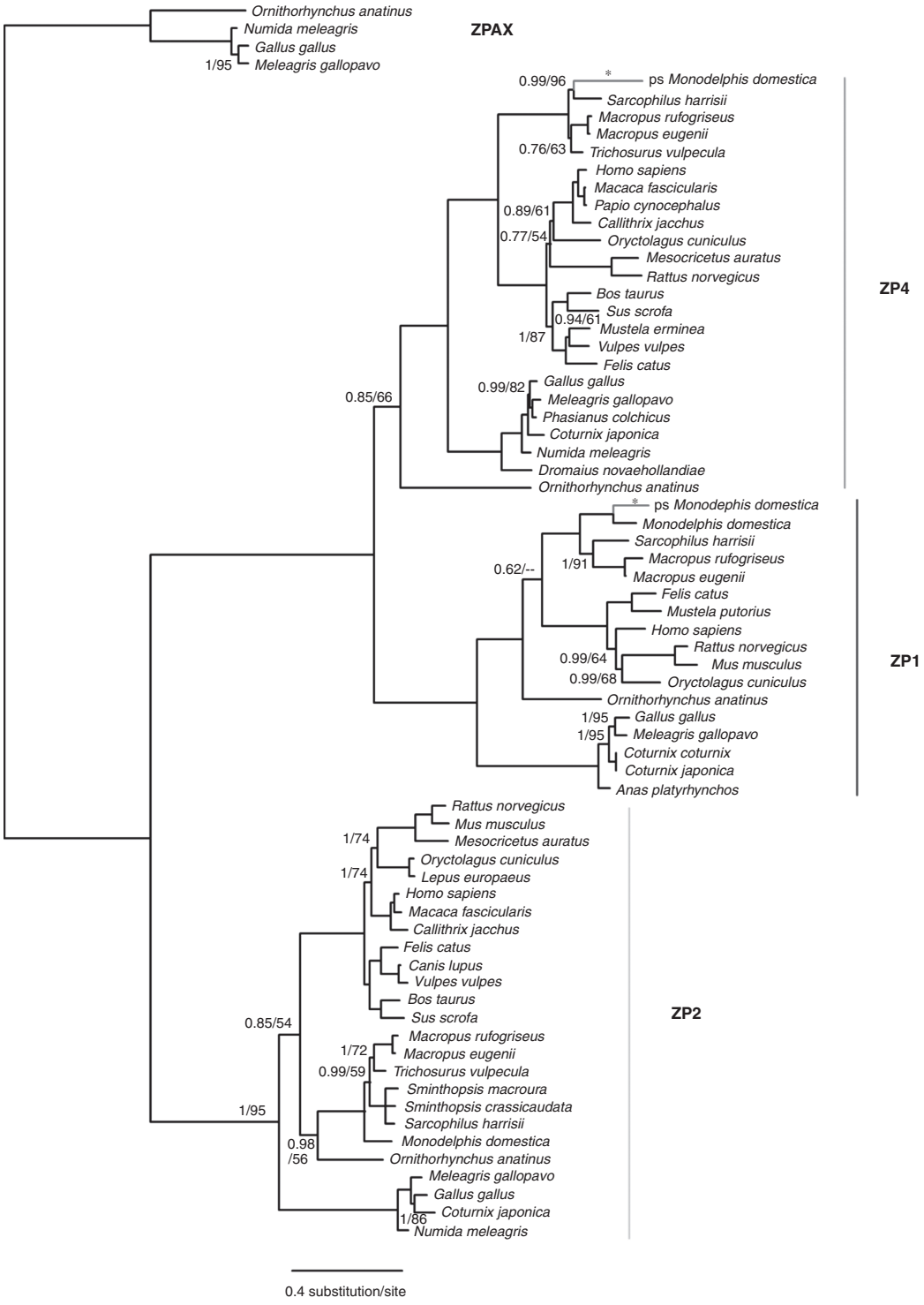

Fig. 3. Phylogenetic tree of *ZP1*, *ZP2*, *ZP4* and *ZPAX*. Asterisk (*) and red branch indicate pseudogenisation; the pseudogenes are indicated by ps before the species name. *ZP1* and marsupial species are shown in blue, *ZP2* and marsupial species are shown in orange and *ZP4* and marsupial species are shown in green. Posterior probability and bootstrap percentage are indicated for each node as follows: PP/BP, except for nodes where PP = 1 and BP > 95.

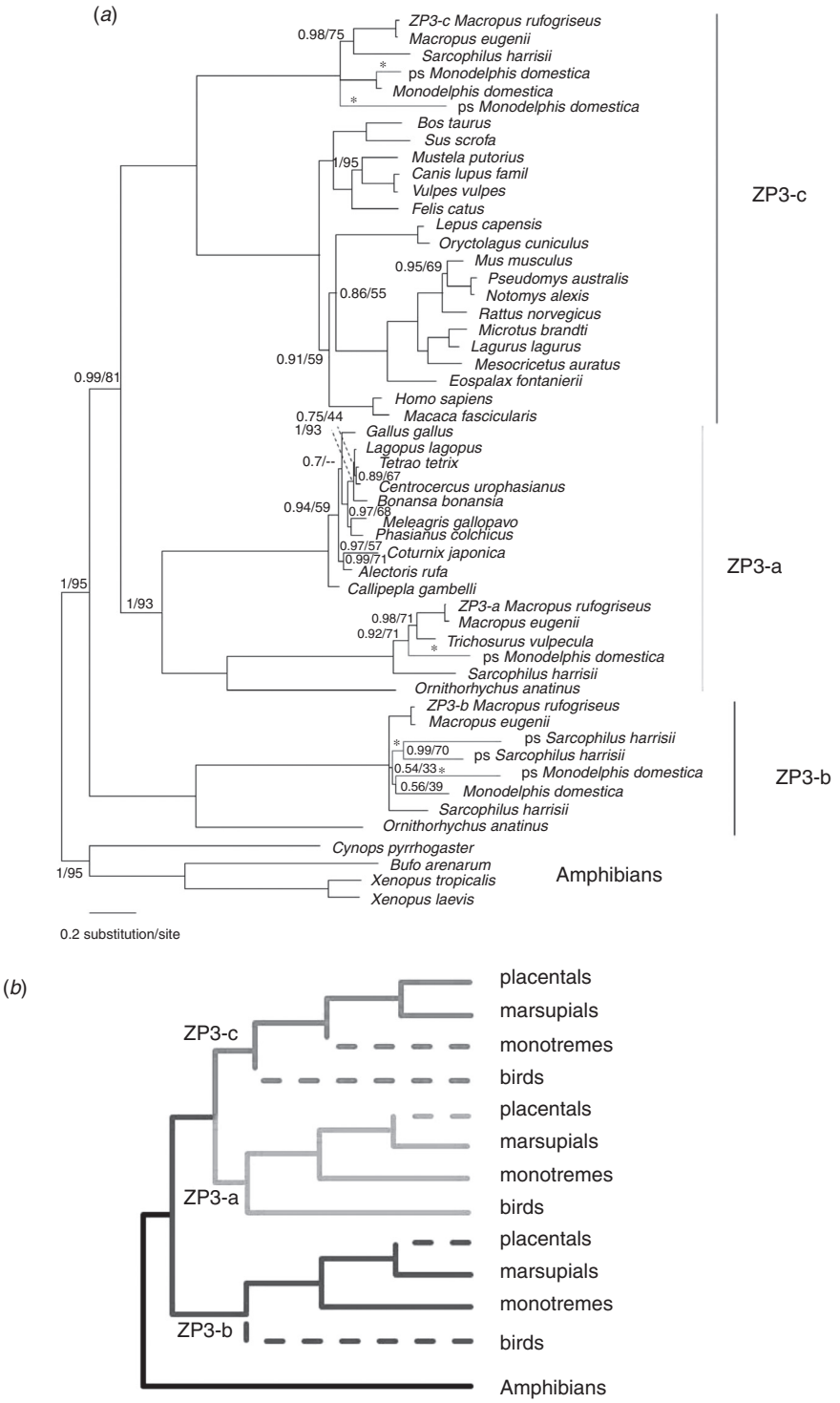

Fig. 4. (*a*) Phylogenetic tree of *ZP3*. Asterisk (*) and red branch indicate pseudogenisation; the pseudogenes are indicated by ps before the species name. *ZP3-a* and marsupial species are shown in orange, *ZP3-b* and marsupial species are shown in blue and *ZP3-c* and marsupial species are shown in green. Posterior probability and bootstrap percentage are indicated for each node as follows: PP/BP, except for nodes where PP = 1 and BP > 95. (*b*) Hypothesis for the evolution of *ZP3* in amniotes. The colours of *ZP3-a*, *ZP3-b* and *ZP3-c* are the same as in (*a*), dashed lines and names in black indicate that a gene became extinct in that lineage.

been identified previously in marsupials. In the present work, the presence of *ZP1* mRNA has been detected, for the first time, in one marsupial (Bennett's wallaby), confirming the data obtained by genomic sequencing of a close relative, the Tammar wallaby. Furthermore, the study demonstrates the presence of transcripts for *ZP2*, *ZP3-a*, *ZP3-b*, *ZP3-c* and *ZP4* in Bennett's wallaby ovaries, suggesting that its ZP could be formed of six proteins. In contrast, the ZP of the gray short-tailed opossum seems to be formed of four proteins (ZP1, ZP2, ZP3-b and ZP3-c), *ZP3-a* and *ZP4* being pseudogenes. The identification for the first time of three different *ZP3* genes helps to clear up previous phylogenetic data (Breed *et al.* 2002; Spargo and Hope 2003).

There is no information on the presence of the different ZP3 proteins in marsupial ZP; however, analysis of the protein sequence (Fig. 1) strongly suggests the secretion of ZP3-a and ZP3-c, due to the presence of a consensus furin cleavage site. Moreover, the hydrophobic patch previously described in mouse ZP3 (Zhao *et al.* 2003) is well conserved in ZP3-c and, to a lesser extent, in ZP3-a, compared with mouse and rat ZP3. In the case of ZP3-b, only two basic amino acids are present at position 366–367 and the hydrophobic patch is not as well conserved as in ZP3-a and ZP3-c, suggesting that this protein is not secreted and remains attached to the oocyte plasma membrane. However, it has previously been reported that the presence of the proprotein convertase (furin) cleavage site is not necessary for ZP4 secretion in cat and human (Zhao *et al.* 2003). Additionally, mutation of the proprotein convertase cleavage site does not prevent the secretion or incorporation of ZP3 protein in different systems (Kiefer and Saling 2002; Qi *et al.* 2002; Zhao *et al.* 2002) or in transgenic mouse lines (Zhao *et al.* 2002). Proteomic analyses have shown that cleavage of the ZP protein occurs immediately upstream of the two basic residues of the furin site (Boja *et al.* 2003, 2005). Consequently, an alternative cleavage site is possible. Further proteomic or heterologous protein expression studies are necessary to clarify this point.

ZP4 *pseudogenisation: a molecular synapomorphy of Didelphinae?*

The ZP is an extracellular matrix that surrounds mammalian oocytes. It has been extensively studied in eutherian mammals, where it has been seen to have several functions in fertilisation and protection of the preimplantion embryo (Yanagimachi 1994; Dean 2004, 2007; Wassarman and Litscher 2009). It has been reported that it is formed of three or four proteins, depending on the species. Moreover, this composition is not dependent on the mammalian order or family. Thus, three proteins have been observed in some rodents, such as *Mus musculus*, where *ZP4* is pseudogenised (Bleil and Wassarman 1980; Lefièvre *et al.* 2004; Evsikov *et al.* 2006; Goudet *et al.* 2008); however, in the rat and the hamster, it is formed of four proteins (Hoodbhoy *et al.* 2005; Izquierdo-Rico *et al.* 2009; Jiménez- Movilla *et al.* 2009) despite the close evolutionary relationship. Similar results are observed in other species, such as the pig, cow, dog, dolphin, Philippine tarsier, common marmoset, Antarctic fur seal, Weddell seal and fox (Hedrick and Wardrip 1987; Noguchi *et al.* 1994; Goudet *et al.* 2008; Stetson *et al.* 2012; Moros-Nicolás *et al.* 2017), indicating that the pseudogenisation events of *ZP1* or *ZP4* occurred several times during mammalian evolution, as previously reported (Lefièvre *et al.* 2004; Evsikov *et al.* 2006; Goudet *et al.* 2008; Stetson *et al.* 2012; Moros-Nicolás *et al.* 2017). In contrast, ZP2 and ZP3 proteins are present in all species described to date.

Information about the composition of marsupial ZP and its evolution is very scarce. Eutherian mammals and marsupials shared a common ancestor around 143–178 million years ago (Ma) (Warren *et al.* 2008; Luo *et al.* 2011) and the divergence between Australasian and South American marsupials is dated around 80 Ma (Beck 2008; Meredith *et al.* 2008; Mitchell *et al.* 2014). The pseudogenisation of *ZP1* or *ZP4* genes during the long period of marsupial evolution may have affected the ZP composition of marsupials, a hypothesis explored in this study.

Marsupials can be classified into two different cohorts: Ameridelphia and Australidelphia, with two (Didelphimorphia and Paucituberculata) and five (Dasyuromorphia, Diprotodontia, Peramelemorphia, Microbiotheria and Notoryctemorphia) orders respectively (Szalay 1982; Aplin and Archer 1987); however, recent molecular studies clearly indicate that the Ameridelphia are not monophyletic (see, for example, Meredith *et al.* 2008; Nilsson *et al.* 2010). For decades, there has been controversy concerning the South American order Microbiotheria in the marsupial tree (see references in Nilsson *et al.* 2010), but recently a study using nuclear genes and retroposons identified the position of this order as the sister group of all Australian marsupials, leading to the suggestion of the name Euaustralidelphia (Nilsson *et al.* 2010). However, as mentioned by these authors, the use of additional markers might be interesting for delving more deeply into the phylogeny of marsupials and to confirm the data obtained to date. Furthermore, a recent study including 193 marsupial species clearly indicates the position of the different orders in the marsupial tree, greatly contributing to its classification (Mitchell *et al.* 2014).

In this study, the *ZP4* gene of different marsupials was analysed and two species (gray short-tailed opossum and common opossum) from the South American order Didelphimorphia were found to have no functional *ZP4* gene. The pseudogenisation process includes several nucleotide mutations and indels in the sequence leading to the presence of stop codons and the absence of the initiation codon coding for methionine (Fig. 2). Based on our phylogeny, we concluded that the pseudogenisation of *ZP4* occurred after the split between the South American and Australasian marsupials, ~80 Ma and before the divergence of *Didelphis marsupialis* and *Monodelphis domestica*, between 20 and 30 Ma (Meredith *et al.* 2008; Jansa *et al.* 2014; Fig. 5). The question remains concerning the Paucituberculata, Microbiotheria and Notoryctemorphia orders, for which no samples were available for us to study. The *ZP3* and *ZP4* pseudogenisation events could be used as additional molecular markers for the phylogenetic classification of marsupials, as suggested previously by some authors (Nilsson *et al.* 2010).

Conclusions

In conclusion, this study suggests that marsupial ZP could be formed by four to six glycoproteins, depending on the species. In South American marsupials, more specifically in the subfamily

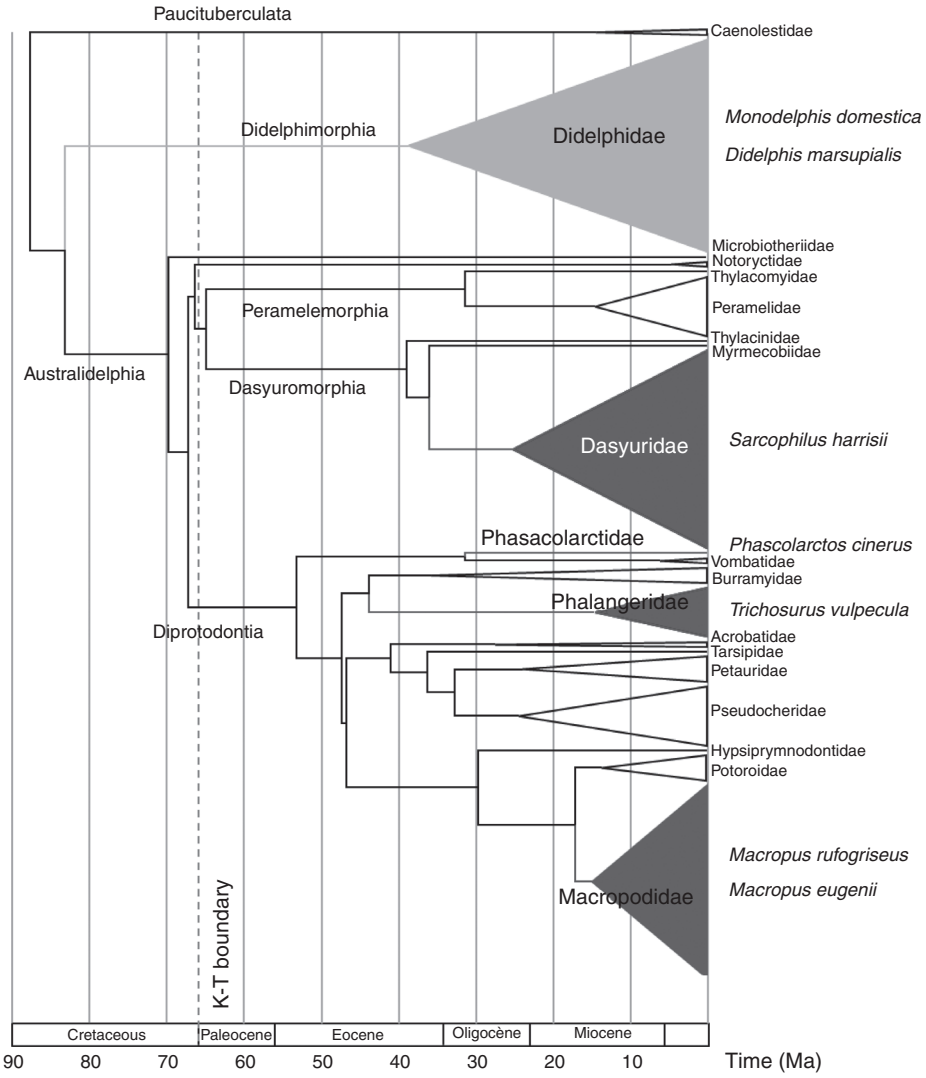

Fig. 5. Marsupial phylogenetic tree and *ZP4* evolution. The species with a functional *ZP4* are indicated in green and the species with a pseudogene are indicated in red (modified from Mitchell *et al.* 2014).

Didelphinae, the ZP is possibly formed by four glycoproteins (ZP1, ZP2, ZP3-b and ZP3-c) and *ZP3-a* and *ZP4* are probably pseudogenes, whereas in Australasian marsupials the ZP is possibly formed by six proteins (ZP1, ZP2, ZP3-a, ZP3-b, ZP3-c and ZP4). The identification of different ZP compositions in several marsupials and the first description of additional ZP proteins could have implications for the cellular understanding of female gamete formation, sperm interaction with the ZP and fertilisation. Moreover, the present findings have potential relevance from a biological, ecological and economic point of view, related to a development of new contraceptive vaccines for the population control of some species, mainly in Australia and New Zealand. However, future analyses are necessary to estimate the relative abundance of these proteins, their glycosylation pattern and their immunogenicity. The biological significance of the differing composition of marsupial ZP remains to be clarified.

Conflicts of interest

The authors declare no conflicts of interest.

Acknowledgements

The authors thank François Catzeflis for the ethanol-preserved tissue from the tissue collection of the Institut des Sciences de l'Evolution of Montpellier and Dr Axel Janke (LOEWE-Biodiversity and Climate Research Centre, Frankfurt, Germany) for providing DNA samples. Phylogenetic analyses were performed using the computing facilities of the CC LBBE/PRABI. We acknowledge the authors that sequenced the genomes of the following species: gray short-tailed opossum (Mikkelsen *et al.* 2007), Tammar wallaby (Renfree *et al.* 2011) and Tasmanian devil (Murchison *et al.* 2012); recognition is also due to the platypus genome study (Warren *et al.* 2008). This work was supported by Ministerio de Ciencia e Innovación Español (AGL2012–40180-C03–02 and AGL2015-7159), The European Commission, The European Regional Development Fund and Fundación Séneca (0452/GERM/06).

References

Aplin, K. P., and Archer, M. (1987). Recent advances in marsupial systematics with a new syncretic classification. In 'Possums and Opossums: Studies in Evolution'. (Ed. M. Archer.) pp. 15–72. (Surrey Beatty and Sons and the Royal Zoological Society of New South Wales: Sydney.)

Au, P. C. K., Whitley, J., Vaux, D., Selwood, L., and Familari, M. (2008). Identification of novel and known ovary-specific genes including ZP2, in a marsupial, the stripe-faced dunnart. *Mol. Reprod. Dev.* **75**, 318–325. doi:10.1002/MRD.20708

Beck, R. M. D. (2008). A dated phylogeny of marsupials using a molecular supermatrix and multiple fossil constraints. *J. Mammal.* **89**, 175–189. doi:10.1644/06-MAMM-A-437.1

Bhat, R., Chakraborty, M., Mian, I. S., and Newman, S. A. (2014). Structural divergence in vertebrate phylogeny of a duplicated prototype galectin. *Genome Biol. Evol.* **6**, 2721–2730. doi:10.1093/GBE/EVU215

Bleil, J. D., and Wassarman, P. M. (1980). Structure and function of the zona pellucida: identification and characterization of the proteins of the mouse oocyte's zona pellucida. *Dev. Biol.* **76**, 185–202. doi:10.1016/0012-1606(80)90371-1

Boja, E. S., Hoodbhoy, T., Fales, H. M., and Dean, J. (2003). Structural characterization of native mouse zona pellucida proteins using mass spectrometry. *J. Biol. Chem.* **278**, 34189–34202. doi:10.1074/JBC.M304026200

Boja, E. S., Hoodbhoy, T., Garfield, M., and Fales, H. M. (2005). Structural conservation of mouse and rat zona pellucida glycoproteins. Probing the native rat zona pellucida proteome by mass spectrometry. *Biochemistry* **44**, 16445–16460. doi:10.1021/BI051883F

Breed, W. G., Hope, R. M., Wiebkin, O. W., Spargo, S. C., and Chapman, J. A. (2002). Structural organization and evolution of the marsupial zona pellucida. *Reproduction* **123**, 13–21. doi:10.1530/REP.0.1230013

Cui, X., Duckworth, J. A., Molinia, F. C., and Cowan, P. E. (2010). Identification and evaluation of an infertility-associated ZP3 epitope from the marsupial brushtail possum (*Trichosurus vulpecula*). *Vaccine* **28**, 1499–1505. doi:10.1016/J.VACCINE.2009.11.052

Dean, J. (2004). Reassessing the molecular biology of sperm–egg recognition with mouse genetics. *BioEssays* **26**, 29–38. doi:10.1002/BIES.10412

Dean, J. (2007). The enigma of sperm–egg recognition in mice. *Soc. Reprod. Fertil. Suppl.* **63**, 359–365.

Evsikov, A. V., Graber, J. H., Brockman, J. M., Hampl, A., Holbrook, A. E., Singh, P., Eppig, J. J., Solter, D., and Knowles, B. B. (2006). Cracking the egg: molecular dynamics and evolutionary aspects of the transition from the fully grown oocyte to embryo. *Genes Dev.* **20**, 2713–2727. doi:10.1101/GAD.1471006

Goudet, G., Mugnier, S., Callebaut, I., and Monget, P. (2008). Phylogenetic analysis and identification of pseudogenes reveal a progressive loss of zona pellucida genes during evolution of vertebrates. *Biol. Reprod.* **78**, 796–806. doi:10.1095/BIOLREPROD.107.064568

Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* **59**, 307–321. doi:10.1093/SYSBIO/SYQ010

Haines, B. P., Rathjen, P. D., Hope, R. M., Whyatt, L. M., Holland, M. K., and Breed, W. G. (1999). Isolation and characterisation of a cDNA encoding a zona pellucida protein (ZPB) from the marsupial *Trichosurus vulpecula* (brushtail possum). *Mol. Reprod. Dev.* **52**, 174–182. doi:10.1002/(SICI)1098-2795(199902)52:2<174::AID-MRD8>3.0.CO;2-7

Harris, J. D., Hibler, D. W., Fontenot, G. K., Hsu, K. T., Yurewicz, E. C., and Sacco, A. G. (1994). Cloning and characterisation of the zona pellucida gene and cDNAs from a variety of mammalian species: the ZPA, ZPB and ZPC gene families. *DNA Seq.* **4**, 361–393. doi:10.3109/10425179409010186

Hedrick, J. L., and Wardrip, N. J. (1987). On the macromolecular composition of the zona pellucida from porcine oocytes. *Dev. Biol.* **121**, 478–488. doi:10.1016/0012-1606(87)90184-9

Hoodbhoy, T., Joshi, S., Boja, E. S., Williams, S. A., Stanley, P., and Dean, J. (2005). Human sperm do not bind to rat zonae pellucidae despite the presence of four homologous glycoproteins. *J. Biol. Chem.* **280**, 12721–12731. doi:10.1074/JBC.M413569200

Hughes, D. C., and Barratt, C. L. (1999). Identification of the true human orthologue of the mouse Zp1 gene: evidence for greater complexity in the mammalian zona pellucida. *Biochim. Biophys. Acta* **1447**, 303–306. doi:10.1016/S0167-4781(99)00181-5

Izquierdo-Rico, M. J., Jiménez-Movilla, M., Llop, E., Pérez-Oliva, A. B., Ballesta, J., Gutierrez-Gallego, R., Jimenez-Cervantes, C., and Avilés, M. (2009). Hamster zona pellucida is formed by four glycoproteins: ZP1, ZP2, ZP3 and ZP4. *J. Proteome Res.* **8**, 926–941. doi:10.1021/PR800568X

Jansa, S. A., Barker, F. K., and Voss, R. S. (2014). The early diversification history of didelphid marsupials: a window into South America's "Splendid Isolation". *Evolution* **68**, 684–695. doi:10.1111/EVO.12290

Jiménez-Movilla, M., Martínez-Alonso, E., Castells, M. T., Izquierdo-Rico, M. J., Saavedra, M. D., Gutiérrez-Gallego, R., Fayrer-Hosken, R., Ballesta, J., and Avilés, M. (2009). Cytochemical and biochemical evidences for a complex tridimensional structure of the hamster zona pellucida. *Histol. Histopathol.* **24**, 599–609.

Jovine, L., Darie, C. C., Litscher, E. S., and Wassarman, P. M. (2005). Zona pellucida domain proteins. *Annu. Rev. Biochem.* **74**, 83–114. doi:10.1146/ANNUREV.BIOCHEM.74.082803.133039

Kiefer, S. M., and Saling, P. (2002). Proteolytic processing of human zona pellucida proteins. *Biol. Reprod.* **66**, 407–414. doi:10.1095/BIOLREPROD66.2.407

Lefièvre, L., Conner, S. J., Salpekar, A., Olufowobi, O., Ashton, P., Pavlovic, B., Lenton, W., Afnan, M., Brewis, I. A., Monk, M., Hughes, D. C., and Barratt, C. L. (2004). Four zona pellucida glycoproteins are expressed in the human. *Hum. Reprod.* **19**, 1580–1586. doi:10.1093/HUMREP/DEH301

Luo, Z. X., Yuan, C. X., Meng, Q. J., and Ji, Q. (2011). A Jurassic eutherian mammal and divergence of marsupials and placentals. *Nature* **476**, 442–445. doi:10.1038/NATURE10291

Mate, K. E., and McCartney, C. A. (1998). Sequence and analysis of zona pellucida 2 cDNA (ZP2) from a marsupial, the brushtail possum, *Trichosurus vulpecula. Mol. Reprod. Dev.* **51**, 322–329. doi:10.1002/(SICI)1098-2795(199811)51:3<322::AID-MRD12>3.0.CO;2-4

Mate, K. E., Buist, J. M., and Duckworth, J. A. (2003). Expression in *Escherichia coli* and immunological characterization of three zona pellucida proteins (ZP1, ZP2 and ZP3) from a marsupial, the brushtail possum (*Trichosurus vulpecula*). *Mol. Reprod. Dev.* **64**, 136–143. doi:10.1002/MRD.10240

McCartney, C. A., and Mate, K. E. (1999). Cloning and characterization of a zona pellucida 3 cDNA from a marsupial, the brushtail possum *Trichosurus vulpecula. Zygote* **7**, 1–9. doi:10.1017/S0967199499000337

McCartney, C. A., Harris, M. S., Rodger, J. C., and Mate, K. E. (2007). Towards a ZP-based contraceptive for marsupials: comparative analysis and developmental expression of marsupial ZP genes. *Mol. Reprod. Dev.* **74**, 1581–1589. doi:10.1002/MRD.20707

Meredith, R. W., Westerman, M., Case, J. A., and Springer, M. S. (2008). A phylogeny and timescale for marsupial evolution based on sequences for five nuclear genes. *J. Mamm. Evol.* **15**, 1–36. doi:10.1007/S10914-007-9062-6

Meslin, C., Mugnier, S., Callebaut, I., Laurin, M., Pascal, G., Poupon, A., Goudet, G., and Monget, P. (2012). Evolution of genes involved in gamete interaction: evidence for positive selection, duplications and losses in vertebrates. *PLoS One* **7**, e44548. doi:10.1371/JOURNAL.PONE.0044548

Mikkelsen, T. S., Wakefield, M. J., Aken, B., Amemiya, C. T., Chang, J. L., Duke, S., Garber, M., Gentles, A. J., Goodstadt, L., Heger, A., *et al.* (2007). Genome of the marsupial *Monodelphis domestica* reveals innovation in non-coding sequences. *Nature* **447**, 167–177. doi:10.1038/NATURE05805

Mitchell, K. J., Pratt, R. C., Watson, L. N., Gibb, G. C., Llamas, B., Kasper, M., Edson, J., Hopwood, B., Male, D., Armstrong, K. N., Meyer, M., Hofreiter, M., Austin, J., Donnellan, S. C., Lee, M. S., Phillips, M. J., and Cooper, A. (2014). Molecular phylogeny, biogeography, and habitat preference evolution of marsupials. *Mol. Biol. Evol.* **31**, 2322–2330. doi:10.1093/MOLBEV/MSU176

Moros-Nicolás, C., Leza, A., Chevret, P., Guillén-Martínez, A., González-Brusi, L., Boué, F., Lopez-Bejar, M., Ballesta, J., Avilés, M., and Izquierdo-Rico, M. J. (2017). Analysis of ZP1 gene reveals differences in zona pellucida composition in carnivores. *Reprod. Fertil. Dev*. doi:10.1071/RD17022

Murchison, E. P., Schulz-Trieglaff, O. B., Ning, Z., Alexandrov, L. B., Bauer, M. J., Fu, B., Hims, M., Ding, Z., Ivakhno, S., Stewart, C., *et al.* (2012). Genome sequencing and analysis of the Tasmanian devil and its transmissible cancer. *Cell* **148**, 780–791. doi:10.1016/J.CELL.2011.11.065

Nilsson, M. A., Churakov, G., Sommer, M., Van Tran, N., Zemann, A., Brosius, J., and Schmitz, J. (2010). Tracking marsupial evolution using archaic genomic retroposon insertions. *PLoS Biol.* **8**, e1000436. doi:10.1371/JOURNAL.PBIO.1000436

Noguchi, S., Yonezawa, N., Katsumata, T., Hashizume, K., Kuwayama, M., Hamano, S., Watanabe, S., and Nakano, M. (1994). Characterization of the zona pellucida glycoproteins from bovine ovarian and fertilized eggs. *Biochim. Biophys. Acta* **1201**, 7–14. doi:10.1016/0304-4165(94)90143-0

Posada, D., and Crandall, K. A. (1998). MODELTEST: testing the model of DNA substitution. *Bioinformatics* **14**, 817–818. doi:10.1093/BIOINFORMATICS/14.9.817

Qi, H., Williams, Z., and Wassarman, P. M. (2002). Secretion and assembly of zona pellucida glycoproteins by growing mouse oocytes microinjected with epitope-tagged cDNAs for mZP2 and mZP3. *Mol. Biol. Cell* **13**, 530–541. doi:10.1091/MBC.01-09-0440

Rambaut, A., Suchard, M. A., Xie, D., and Drummond, A. J. (2014). Tracer v1.6. Available at: http://tree.bio.ed.ac.uk/software/tracer/.

Renfree, M. B., Papenfuss, A. T., Deakin, J. E., Lindsay, J., Heider, T., Belov, K., Rens, W., Waters, P. D., Pharo, E. A., Shaw, G., *et al.* (2011). Genome sequence of an Australian kangaroo, *Macropus eugenii*, provides insight into the evolution of mammalian reproduction and development. *Genome Biol.* **12**(8), R81. doi:10.1186/GB-2011-12-8-R81

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D. L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M. A., and Huelsenbeck, J. P. (2012). MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* **61**, 539–542. doi:10.1093/SYSBIO/SYS029

Selwood, L. (2000). Marsupial egg and embryo coats. *Cell Tissues Organs* **166**, 208–219. doi:10.1159/000016733

Spargo, S. C., and Hope, R. M. (2003). Evolution and nomenclature of the zona pellucida gene family. *Biol. Reprod.* **68**, 358–362. doi:10.1095/BIOLREPROD.102.008086

Stetson, I., Izquierdo-Rico, M. J., Moros, C., Chevret, P., Lorenzo, P. L., Ballesta, J., Rebollar, P. G., Gutiérrez-Gallego, R., and Avilés, M. (2012). Rabbit zona pellucida composition: a molecular, proteomic and phylogenetic approach. *J. Proteomics* **75**, 5920–5935. doi:10.1016/J.JPROT.2012.07.027

Stetson, I., Avilés, M., Moros, C., García-Vázquez, F. A., Gimeno, L., Torrecillas, A., Aliaga, C., Bernardo-Pisa, M. V., Ballesta, J., and Izquierdo-Rico, M. J. (2015). Four glycoproteins are expressed in the cat zona pellucida. *Theriogenology* **83**, 1162–1173. doi:10.1016/J.THERIOGENOLOGY.2014.12.019

Szalay, F. S. (1982). A new appraisal of marsupial phylogeny and classification. In 'Carnivorous Marsupials'. (Ed. M. Archer.) p. 621–640. (Royal Zoological Society of New South Wales: Sydney.)

Voyle, R. B., Haines, B. P., Loffler, K. A., Hope, R. M., Rathjen, P. D., and Breed, W. G. (1999). Isolation and characterisation of zona pellucida A (ZPA) cDNAs from two species of marsupial: regulated oocyte-specific expression of ZPA transcripts. *Zygote* **7**, 239–248. doi:10.1017/S0967199499000623

Warren, W. C., Hillier, L. W., Marshall Graves, J. A., Birney, E., Ponting, C. P., Grützner, F., Belov, K., Miller, W., Clarke, L., Chinwalla, A. T., *et al.* (2008). Genome analysis of the platypus reveals unique signatures of evolution. *Nature* **453**, 175–183. doi:10.1038/NATURE06936

Wassarman, P. M., and Litscher, E. S. (2009). The multifunctional zona pellucida and mammalian fertilization. *J. Reprod. Immunol.* **83**, 45–49. doi:10.1016/J.JRI.2009.06.259

Yanagimachi, R. (1994). Mammalian fertilization. In 'Physiology of Reproduction'. (Eds E. Knobil and J. D. Neill.) pp. 189–317. (Raven Press: New York.)

Zhao, M., Gold, L., Ginsberg, A. M., Liang, L. F., and Dean, J. (2002). Conserved furin cleavage site not essential for secretion and integration of ZP3 into the extracellular egg coat of transgenic mice. *Mol. Cell. Biol.* **22**, 3111–3120. doi:10.1128/MCB.22.9.3111-3120.2002

Zhao, M., Gold, L., Dorward, H., Liang, L. F., Hoodbhoy, T., Boja, E., Fales, H. M., and Dean, J. (2003). Mutation of a conserved hydrophobic patch prevents incorporation of ZP3 into the zona pellucida surrounding mouse eggs. *Mol. Cell. Biol.* **23**, 8982–8991. doi:10.1128/MCB.23.24.8982-8991.2003