



# Genes Encoding Mammalian Oviductal Proteins Involved in Fertilization are Subjected to Gene Death and Positive Selection

Carla Moros-Nicolás<sup>1</sup> · Sophie Fouchécourt<sup>1</sup> · Ghylène Goudet<sup>1</sup> · Philippe Monget<sup>1</sup>

Received: 19 April 2018 / Accepted: 12 November 2018  
© The Author(s) 2018

## Abstract

Oviductal proteins play an important role in mammalian fertilization, as proteins from seminal fluid. However, in contrast with the latter, their phylogenetic evolution has been poorly studied. Our objective was to study in 16 mammals the evolution of 16 genes that encode oviductal proteins involved in at least one of the following steps: (1) sperm–oviduct interaction, (2) acrosome reaction, and/or (3) sperm–zona pellucida interaction. Most genes were present in all studied mammals. However, some genes were lost along the evolution of mammals and found as pseudogenes: annexin A5 (ANXA5) and deleted in malignant brain tumor 1 (DMBT1) in tarsier; oviductin (OVGP1) in megabat; and probably progesterone-associated endometrial protein (PAEP) in tarsier, mouse, rat, rabbit, dolphin, and megabat; prostaglandin D2 synthase (PTGDS) in microbat; and plasminogen (PLG) in megabat. Four genes [ANXA1, ANXA4, ANXA5, and heat shock 70 kDa protein 5 (HSPA5)] showed branch-site positive selection, whereas for seven genes [ANXA2, lactotransferrin (LTF), OVGP1, PLG, S100 calcium-binding protein A11 (S100A11), Sperm adhesion molecule 1 (SPAM1), and osteopontin (SPP1)] branch-site model and model-site positive selection were observed. These results strongly suggest that genes encoding oviductal proteins that are known to be important for gamete fertilization are subjected to positive selection during evolution, as numerous genes encoding proteins from mammalian seminal fluid. This suggests that such a rapid evolution may have as a consequence that two isolated populations become separate species more rapidly.

**Keywords** Protein evolution · Oviduct · Fertilization · Pseudogene · Positive selection

## Introduction

The oviduct, also called the Fallopian tube in human, is a tubular organ that plays an important role in mammalian fertilization. The oviduct is the location where fertilization takes place in amniotes. It is involved in gamete transportation and maturation, sperm capacitation, and it provides the appropriate microenvironment for the early embryo development (Coy et al. 2012a; Hunter 1998, 2012). The oviductal fluid (OF) is formed by the secretion of the epithelial cells and also results from a transudate of the blood plasma, and contains glycosaminoglycans, metabolites, amino acids, inorganic salts, and proteins (Aviles et al. 2010; Leese et al. 2008, 2001).

The proteome of the OF has been studied in several mammalian species, such as cow (Lamy et al. 2016), ewe (Soleilhavoup et al. 2016), mare (Smits et al. 2017), pig (Georgiou et al. 2005; Mondejar et al. 2012), and rabbit (Yu et al. 2016). It has been demonstrated that the oviductal proteome changes with the different stages of the oestrous cycle (Lamy et al. 2016; Soleilhavoup et al. 2016), the presence of gametes (Georgiou et al. 2005) or the presence of embryos (Smits et al. 2017). The proteome of the oviduct epithelial cells has also been studied in humans (Wang et al. 2016) and pigs (Seytanoglu et al. 2008).

We have chosen to study the evolution of 16 genes that have been demonstrated to be crucial for the biological function of oviduct. Oviductal proteins, like oviductin (OVGP1), annexins (ANXA), heat shock proteins (HSP), deleted in malignant brain tumor (DMBT1), or osteopontin (SPP1) among others have been extensively studied (Coy and Yanagimachi 2015). These proteins are involved in different steps that are crucial for fertilization. For instance, annexins and DMBT1 participate in the binding between

✉ Philippe Monget  
philippe.monget@inra.fr

<sup>1</sup> PRC, INRA, CNRS, IFCE, Université de Tours,  
37380 Nouzilly, France

the sperm and the oviduct (Ignatz et al. 2007; Teijeiro et al. 2011). Progesterone-associated endometrial protein (PAEP) also known as glycodelin (Gd), lactotransferrin (LTF), heat shock 70 kDa protein 5 (HSPA5), OVGPI, prostaglandin D2 synthase (PTGDS), plasminogen (PLG), and SPP1 participate in the zona pellucida–sperm interaction controlling polyspermy (Algarra et al. 2016; Chiu et al. 2007; Coy et al. 2008, 2012b; Goncalves et al. 2008a; Hao et al. 2006; Lachance et al. 2007; Marin-Briggiler et al. 2010; Mondejar et al. 2013; Zumoffen et al. 2013). Heat shock 70 kDa protein 8 (HSPA8) has a beneficial effect on sperm viability (Elliott et al. 2009; Lloyd et al. 2009; Moein-Vaziri et al. 2014). Natriuretic peptide A (NPPA) and glycodeilin-A (GdA) induce the acrosome reaction (Chiu et al. 2010; Zhang et al. 2006). Sperm adhesion molecule 1 (SPAM1) plays a role in the dispersion of the cumulus cells (Griffiths et al. 2008) and S100 calcium-binding protein A11 (S100A11) plays a role in sperm selection through its action on cumulus cells (Hanaue et al. 2011).

It is well known that genes related to reproduction evolve faster than genes expressed in most other tissues (Singh and Kulathinal 2000; Swanson and Vacquier 2002; Turner and Hoekstra 2008; Vacquier 1998); genes subjected to rapid evolution show a big percentage of amino acids substitution between species. For example, some genes involved in oocyte–sperm interaction were highlighted as ZP2, ZP3, or acrosine (Makalowski and Boguski 1998). Rapid evolution can also be associated to a loss of functionality, sometimes leading to pseudogenization (Bodmer and Ashburner 1984; Hellberg and Vacquier 1999; Li et al. 1981), or it can be related to an adaptive evolution induced by natural selection (Swanson and Vacquier 2002). This rapid evolution can be species-specific and it could even play a role in speciation, as it has been described for instance for CD9 in the oocyte and IZUMO1 in the sperm (Claw et al. 2014).

Several studies have been focused on the positive selection of the seminal plasma proteins in *Drosophila*, in rodents, and in primates (Meslin et al. 2015). However, to our knowledge, there are no studies focusing on the evolution and an eventual positive selection of the genes encoding oviductal proteins in any mammalian species. For this reason, the aim of this work was to study the evolution of these genes in mammals.

## Materials and Methods

### Phylogenetic and Syntenic Analyses

Sixteen oviductal proteins (ANXA1, ANXA2, ANXA4, ANXA5, DMBT1, PAEP, HSPA5, HSPA8, LTF, NPPA, OVGPI, PLG, PTGDS, SPAM1, SPP1, and S100A11), described in the literature as important for fertilization, were

chosen for the analysis. On this study, the genomes of 16 placental mammal species were sampled; different species were chosen due to the availability of their genome and to include species from all the phylogenetic tree of mammals (*Ailuropoda melanoleuca*, *Bos taurus*, *Canis lupus familiaris*, *Equus caballus*, *Felis catus*, *Homo sapiens*, *Myotis lucifugus*, *Mus musculus*, *Oryctolagus cuniculus*, *Ovis aries*, *Pongo abelii*, *Pteropus vampyrus*, *Rattus norvegicus*, *Sus scrofa*, *Tarsius syrichta*, and *Tursiops truncatus*). The version of Ensembl used to perform the analyses was Ensembl 87: Dec 2016, and the following versions of genomes were used: cat (*Felis\_catus\_6.2*), cow (UMD3.1), horse (EquCab2), human (GRCh38.p7), dog (CanFam3.1), dolphin (turTru1), megabat (pteVam1), microbat (Myoluc2.0), mouse (GRCm38.p5), orangutan (PPYG2), panda (ailMel1), pig (Sscrofa10.2), rabbit (OryCun2.0), rat (Rnor\_6.0), sheep (Oar\_v3.1), and tarsier (tarSyr1).

For all identified genes, we used the PhyleasProg web server v3.1 (<http://phyleasprog.inra.fr/>) (Busset et al. 2011) to study the evolution of the corresponding Ensembl protein ID (Table 1), retrieved from the Ensembl database. We carefully examined all reconstructed phylogenetic trees before the interpretation of the results, that we eventually corrected by synteny analysis as previously described (Tian et al. 2009), so that calculations were performed with correct orthologs.

### Identification of Pseudogenes—Inference of Positive Selection

We systematically performed the search for pseudogenes by using tBlastn in the studied genomes for genes for which one of the species of interest presented no ortholog. As previously described for seminal plasma genes, this methodology allowed us to test the hypothesis that evolution of oviductal genes in mammals is characterized by a gene loss pattern (Meslin et al. 2011, 2015; Tian et al. 2009). The pseudogene status was inferred in a genome if we found a premature stop codon or an indel in the sequence identified by the similarity search in the syntenic locus in comparison with the other species of interest. In the cases of absent genes without pseudogene observed in the syntenic locus, we could only make the hypothesis that the gene has been lost.

The PhyleasProg web server used the CODEML application from the PAML package version 4.4 to investigate selective pressure (Yang 2007). This allows the calculation of the ratio dN/dS which varied across codons and the estimation of the probability for each codon to be under positive selection. We used MUSCLE and PAL2NAL softwares to obtain the alignments (Thompson et al. 1994; Suyama et al. 2006). We systematically and carefully examined multiple alignments to avoid false-positive results. In particular, we did not consider amino acids that were at the boundary of

**Table 1** List of the 16 proteins studied. Human sequences of the proteins were used for analyses

Protein	Abbreviation	Protein accession number
Annexin 1	ANXA1	ENSP00000366109
Annexin 2	ANXA2	ENSP00000346032
Annexin 4	ANXA4	ENSP00000377833
Annexin 5	ANXA5	ENSP00000296511
Deleted in malignant brain tumor	DMBT1	ENSP00000357905
Heat shock 70 kDa protein 5	HSPA5	ENSP00000324173
Heat shock 70 kDa protein 8	HSPA8	ENSP00000432083
Lactotransferrin	LTF	ENSP00000231751
Natriuretic peptide A	NPPA	ENSP00000365663
Oviductin	OVGP1	ENSP00000358747
Progesterone-associated endometrial protein	PAEP	ENSP00000417898
Plasminogen	PLG	ENSP00000308938
Prostaglandin D2 synthase	PTGDS	ENSP00000360687
Sperm adhesion molecule 1	SPAM1	ENSP00000345849
Osteopontin	SPP1	ENSP00000237623
S100 calcium-binding protein A11	S100A11	ENSP00000271638

the alignments. We also systematically eliminated genes that presented sequence errors in Ensembl according to other databases such as RefSeq in NCBI, in particular for ANXA1 and PLG in rabbit; OVG1 in horse, megabat, and mouse; and SPAM1 in megabat.

Site Models implemented in PAML were used to evaluate if the intensity of selective pressure varies among sites in the sequences studied (Nielsen and Yang 1998). This allowed the estimation of  $\omega$  ratio to vary among sites (Nielsen and Yang 1998; Yang 2000). Five models and three comparisons are used in PhyleasProg: M1a ( $0 < \omega_0 < 1$  and  $\omega_1 = 1$ ) versus M2a ( $0 < \omega_0 < 1$ ,  $\omega_1 = 1$  and  $\omega_2 > 1$ ) (Wong et al. 2004; Yang et al. 2005), M7 ( $0 < \omega < 1$ ) versus M8 ( $0 < \omega < 1$  and  $\omega_S > 1$ ) (Yang 2000), and M8 versus M8a ( $0 < \omega < 1$  and  $\omega_S = 1$ ) (Swanson et al. 2003). We used likelihood ratio tests (LRTs) to compare the log likelihood values (Nielsen and Yang 1998). We used Bayes Empirical Bayes (BEB) method (Yang et al. 2005) implemented in PAML to estimate posterior probabilities of selection on each codon. Probabilities higher than 0.95 were considered significant.

Studying whether one or more species have undergone selection pressure on some of their genes during evolution has led us to use the Phyleasprog tool using branch-site models of PAML (Yang and Nielsen 2002; Zhang et al. 2005). These models can detect positive selection episodes on specific branches of the phylogenetic tree, designated by the user. Two types of branches are defined in these models, the so-called foreground branches for which positive selection will be allowed, (defined a priori by the user), and the so-called background branches, for which the  $dN/dS$  ratio is allowed to vary only between 0 and 1. Two models are used, one model allowing positive selection, called alternative model, and the other not allowing it, called null model.

In the latter model, the  $dN/dS$  ratio  $\omega_2$  is fixed to 1. For the alternative model, three classes are defined:  $\omega_0$ :  $dN/dS < 1$ ,  $\omega_1$ :  $dN/dS = 1$ , and  $\omega_2$ :  $dN/dS \geq 1$ . As for the site model, LRT (Nielsen and Yang 1998) and BEB (Yang et al. 2005) were used.

We tested simultaneously each branch of each phylogenetic tree for positive selection. The  $q$  value was used to control the statistical evidence associated with each branch tested because multiple-hypothesis tests were performed. Similar to the  $p$  value, we used the  $q$  value to measure the significance in terms of false discovery rate rather than false-positive rate. We used R package QVALUE to compute the  $q$  values (Storey and Tibshirani 2003). We considered positive selection on the foreground branch to be significant with a threshold of  $q < 10\%$  of false positives. After validation of the branch with the  $q$  value, we only considered sites with posterior probabilities of Bayes Empirical Bayes analyses superior to 95% or 99%. We discarded from subsequent analysis datasets with less than ten sequences, which is the minimum threshold required to obtain significant results, with excessively divergent sequences, or with sequences of genes for which annotations are not reliable.

## Results

### Identification of Pseudogenes and Paralogs

The present search for pseudogenes showed that ANXA5 and DMBT1 were lost in tarsier; OVG1 in megabat; and probably PAEP in tarsier, mouse, rat, rabbit, dolphin, and megabat; PTGDS in microbat; and PLG in megabat

(examples of traces of pseudogenes in Fig. 1; pseudogenes and species affected in Fig. 2).

Paralogues (*n*) were found for the next genes: OVGPI in sheep (1), cow (2), and pig (1); HSPA8 in rat (3) and dog (1); and SPAM1 in mouse (1), rat (1), horse (1), cow (1), and sheep (1) (Fig. 2).

### Inference of Positive Selection

Five genes did not show positive selection: DMBT1, HSPA8, NPPA, PAEP, and PTGDS. Seven genes showed site model positive selection: ANXA2, LTF, OVGPI, PLG, S100A11, SPAM1, and SPP1; and 11 genes showed branch model positive selection: ANXA1, ANXA2, ANXA4, ANXA5, HSPA5, LTF, OVGPI, PLG, S100A11, SPAM1, and SPP1.

Site model positive selection was found for OVGPI, PLG, S100A11, and SPP1 with the M8a versus M8 comparison, whereas LTF and SPAM1, LRTs for both comparisons (M1 vs. M2, M8a vs. M8) were significant ( $p < 0.001$ ) for the dataset. However, as these genes evolve particularly rapidly, the protein sequences are very divergent between species, and the identification of amino acids that are under positive selection was very difficult. So, we decided not to include this part of result because they are not confident.

The branch model showed positive selection for ANXA1, ANXA 2, ANXA4, ANXA5, HSPA5, LTF, OVGPI, PLG, S100A11, SPAM1, and SPP1 genes. The comparison between site models of PAML detects positive selection only if the  $\omega$  ratio averaged over all branches on the tree is greater than 1, but it can also be expected positive selection to affect only a few amino acid residues along particular lineages.

#### ANXA5 in tarsier

```
Query 156      LLQANRDPDAGIDEAQVEQDAQ 177
                LL ANRDPDAGI EAQVEQDAQ
Sbjct 1202681  LL*ANRDPDAGISEAQVEQDAQ 1202746
```

#### DMBT1 in tarsier

```
Query 1115      GSESSLALRLVNGGDRQCGRVEVLYRGSWGTVCDDYWDTNDANVVCRLGCGWAMSAPGN
1174
                G   LALRLVNG D+CQ  VEVLY  SWGTVCDD  DTNDANVVCRLGCG AM AP
Sbjct 161292      GKAYGLALRLVNGVDQCQSWVEVLY*DSWGTVCDDR*DTNDANVVCRLGCG*AMLAP-E
161468
```

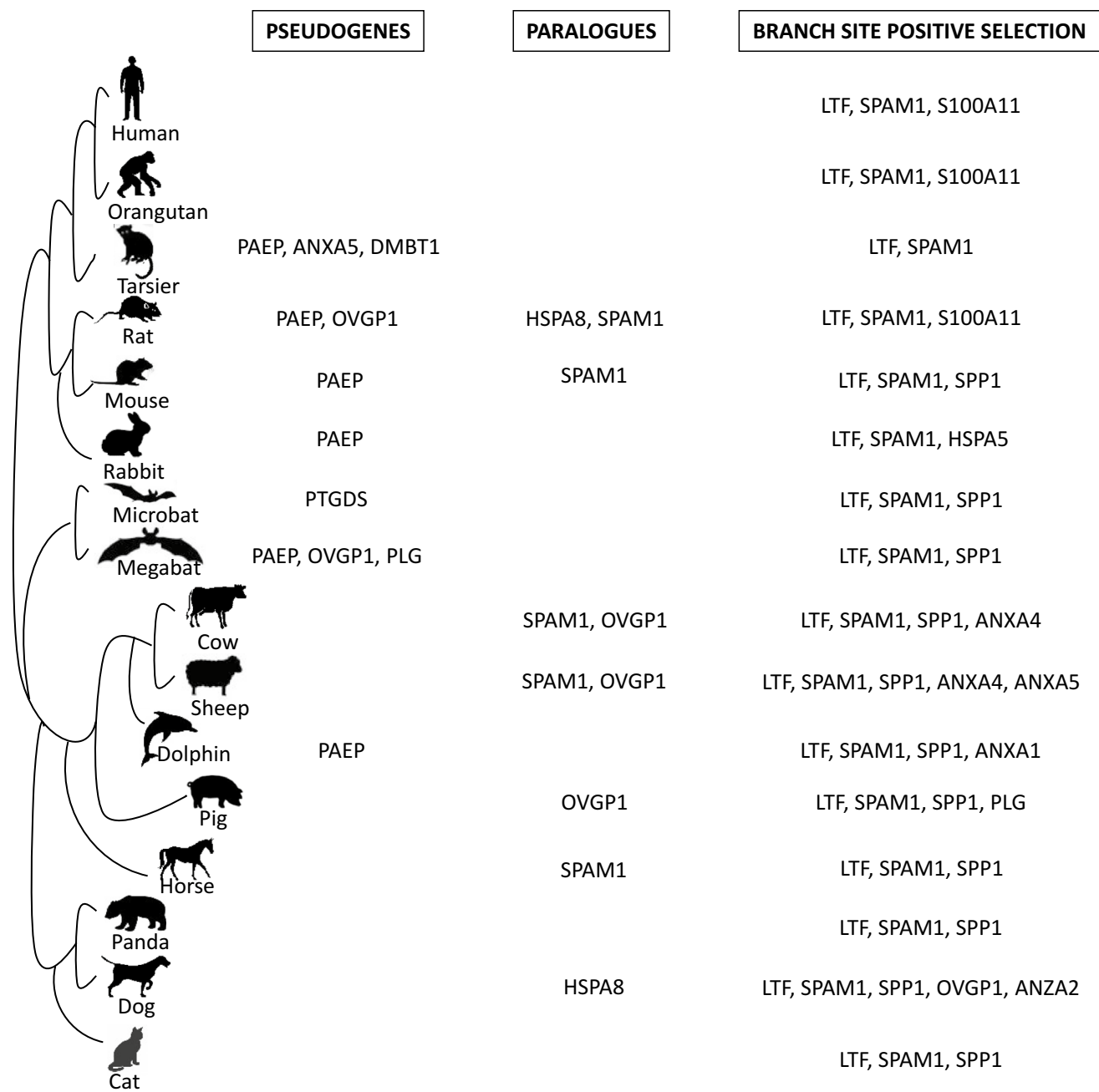
```
Query 1175      ARFGQSGPIVLDDVRCSGHESYLWSCPHNGWLSHNCGHEDAGVICASASQSQP 1228
                FGQSGPI+LD++ C  HE+YLWSC H+ WLSHNCG  EDAGV+CS  ++P
Sbjct 161469      IPFGQSGPILLDNMCCFRHETYLWSCSHHSWLSHNCGRSEDAGVLCSGGPTRP 161630
```

#### OVGPI in megabat

```
Query 240      AYAMNYWRKLGAPSEKLIMGIPTYGRTFRLKASKNGLQARAIGPASPGKYTKQEGFLAY
299
                A+A++YW KLGA SE L+MG PTYG  FRLK SKNGLQA AIGPASPGKY KQ GFLAY
Sbjct 7214429    AHAIYW*KLGALSE*LLMGFPTYGYMFRLLKTSKNGLQAEIIGPASPGKYIKQSGFLAY
7214608
```

**Fig. 1** Identification of marks of pseudogenes. A tBlastn analysis allowed the identification of exons presenting STOP codons (\*) within tarsier ANXA5, tarsier DMBT1, and megabat OVGPI. For

each alignment, the upper number corresponds to the amino acid position and the lower number to the genomic position of nucleotides on the corresponding chromosome



**Fig. 2** Proteins present in oviductal fluids and phylogenetic results showing gene loss, gene duplication, and positive selection. The list of oviductal proteins involved in the different processes of sperm stor-

age and gamete fertilization in different mammals was established in the “Material and Methods” section

So, the branch-site models were used (Yang and Nielsen 2002), which are dedicated to detect such signals of local episodic positive selection, to determine whether different species have undergone selective pressure. Sixteen species were tested as foreground branch with branch-site models of PAML, branches that were tested are indicated in Fig. 2.

Significant LRTs (with at least  $p < 0.05$ ) were found for ANXA1 in dolphin; ANXA 2 and OVGP1 in dog; ANXA4

in ruminants (cow and sheep); ANXA5 in sheep; HSPA5 in rabbit; PLG in pig; S100A11 in human, orangutan, and rat; SPP1 in microbat, megabat, horse, dog, panda, cat, pig, dolphin, cow, sheep, and mouse; and LTF and SPAM1 in all the studied species suggesting that orthologs of these genes might have been subjected to positive selection in these species (Table 2).

**Table 2** Parameter estimates and likelihood scores for branch-site evolutionary models

Species	Model	$\ell^a$	Estimates of parameters <sup>b</sup>	$2\Delta\ell^c$	Positively selected sites (BEB) <sup>d</sup>
ANXA1—parameter estimates and likelihood scores for branch-site models for species					
Dolphin	Null	-2590.515664	$\rho_0=0.557, (\rho_1=0.123), \omega_0=0.046, (\omega_1=1)$	3.915*	Not allowed
	Alternative	-2588.557995	$\rho_0=0.750, \rho_1=0.160, (\rho_2=0.087), \omega_0=0.049, (\omega_1=1), \omega_2=7.275$		1 Site $p > 95\%$ : 5N
ANXA2—parameter estimates and likelihood scores for branch-site models for species					
Dog	Null	-3141.701658	$\rho_0=0.312, (\rho_1=0.007), \omega_0=0.071, (\omega_1=1)$	6.423*	Not allowed
	Alternative	-3138.489876	$\rho_0=0.535, \rho_1=0.011, (\rho_2=0.072), \omega_0=0.072, (\omega_1=1), \omega_2=2.726$		8 Sites $p > 99\%$ : 32A, 108A109E, 110S, 111L, 172I, 249H, 281Q. 16 sites $p > 95\%$ : 80T, 143P, 159T, 162P, 184W, 190W, 209T, 235G, 242G, 246Q, 282I
ANXA4—parameter estimates and likelihood scores for branch-site models for species					
Ruminants (cow, sheep)	Null	-2537.951438	$\rho_0=0.886, (\rho_1=0.074), \omega_0=0.034, (\omega_1=1)$	4.451*	Not allowed
	Alternative	-2535.725639	$\rho_0=0.915, \rho_1=0.074, (\rho_2=0.009), \omega_0=0.034, (\omega_1=1), \omega_2=18.872$		Cow: 1 site $p > 95\%$ : 216A Sheep: 1 site $p > 95\%$ : 217T
ANXA5—parameter estimates and likelihood scores for branch-site models for species					
Sheep	Null	-2580.297972	$\rho_0=0.734, (\rho_1=0.063), \omega_0=0.018, (\omega_1=1)$	3.937*	Not allowed
	Alternative	-2578.329334	$\rho_0=0.903, \rho_1=0.077, (\rho_2=0.018), \omega_0=0.018, (\omega_1=1), \omega_2=29.298$		1 Site $p > 95\%$ : 309R
HSPA5—parameter estimates and likelihood scores for branch-site models for species					
Rabbit	Null	-5737.704417	$\rho_0=0.962, (\rho_1=0.017), \omega_0=0.001, (\omega_1=1)$	10.735**	Not allowed
	Alternative	-5732.336651	$\rho_0=0.975, \rho_1=0.016, (\rho_2=0.007), \omega_0=0.001, (\omega_1=1), \omega_2=998.999$		1 Site $p > 95\%$ : 329S
LTF—parameter estimates and likelihood scores for branch-site models for species					
Mammals	Null	-12709.754534	$\rho_0=0.632, (\rho_1=0.169), \omega_0=0.108, (\omega_1=1)$	90.879***	Not allowed
	Alternative	-12664.314683	$\rho_0=0.605, \rho_1=0.357, (\rho_2=0.036), \omega_0=0.106, (\omega_1=1), \omega_2=3.836$		Rabbit: 4 sites $p > 99\%$ : 32P, 39S, 62L, 642A; 3 sites $p > 95\%$ : 36A, 61A, 183Q Mouse: 4 sites $p > 99\%$ : 31N, 38L, 61R, 641Q; 3 sites $p > 95\%$ : 35E, 60T, 182D Rat: 4 sites $p > 99\%$ : 32R, 39F, 62P, 641C; 3 sites $p > 95\%$ : 36Q, 61M, 182S Tarsier: 4 sites $p > 99\%$ : 32N, 39Y, 62R, 534R; 3 sites $p > 95\%$ : 36T, 61V, 183R Orangutan: 4 sites $p > 99\%$ : 34Q, 41F, 64T, 646S; 3 sites $p > 95\%$ : 38T, 63P, 185G Human: 4 sites $p > 99\%$ : 32Q, 39F, 62I, 644S; 3 sites $p > 95\%$ : 36T, 61P, 183G Dog: 4 sites $p > 99\%$ : 32K, 39S, 62Q, 642T; 3 sites $p > 95\%$ : 36K, 61R, 183K Panda: 4 sites $p > 99\%$ : 32K, 39S, 62Q, 640A; 3 sites $p > 95\%$ : 36A, 61H, 183R Cat: 4 sites $p > 99\%$ : 32Q, 39T, 62Q, 645D; 3 sites $p > 95\%$ : 36T, 61R, 183K Microbat: 4 sites $p > 99\%$ : 32P, 39F, 62T, 644R; 3 sites $p > 95\%$ : 36T, 61Y, 183R Megabat: 4 sites $p > 99\%$ : 32K, 39S, 62K, 641S; 3 sites $p > 95\%$ : 36A, 61H, 183T Horse: 4 sites $p > 99\%$ : 32P, 39A, 62F, 642P; 3 sites $p > 95\%$ : 36A, 61S, 183K Pig: 4 sites $p > 99\%$ : 32T, 39R, 60T, 638K; 3 sites $p > 95\%$ : 36S, 59P, 178N Dolphin: 4 sites $p > 99\%$ : 32L, 39Y, 62F, 627K; 3 sites $p > 95\%$ : 36L, 61R, 183K Cow: 4 sites $p > 99\%$ : 32Q, 39R, 62L, 642K; 3 sites $p > 95\%$ : 36F, 61A, 183Q Sheep: 4 sites $p > 99\%$ : 32P, 39Y, 62L, 642K; 3 sites $p > 95\%$ : 36S, 61A, 183K

**Table 2** (continued)

Species	Model	$\ell^a$	Estimates of parameters <sup>b</sup>	$2\Delta I^c$	Positively selected sites (BEB) <sup>d</sup>
Artiodactyls (pig, dolphin, cow, sheep)	Null	-12706.881324	$\rho_0=0.608, (\rho_1=0.337), \omega_0=0.104, (\omega_1=1)$	10.855**	Not allowed
	Alternative	-12701.453814	$\rho_0=0.622, \rho_1=0.354, (\rho_2=0.022), \omega_0=0.108, (\omega_1=1), \omega_2=6.298$		Dolphin: 1 site $p > 95\%$ : 583V Sheep: 1 site $p > 95\%$ : 597T
Pig	Null	-12707.737211	$\rho_0=0.587, (\rho_1=0.333), \omega_0=0.106, (\omega_1=1)$	18.188***	Not allowed
	Alternative	-12698.643012	$\rho_0=0.621, \rho_1=0.356, (\rho_2=), \omega_0=0.022, \omega_1=1, \omega_2=999.000$		1 Site $p > 95\%$ : 49N
Dolphin	Null	-12709.096501	$\rho_0=0.594, (\rho_1=0.338), \omega_0=0.108, (\omega_1=1)$	7.501**	Not allowed
	Alternative	-12705.345589	$\rho_0=0.629, \rho_1=0.353, (\rho_2=0.015), \omega_0=0.109, (\omega_1=1), \omega_2=16.085$		1 Site $p > 95\%$ : 507D
Rodents (mouse, rat)	Null	-12694.848599	$\rho_0=0.543, (\rho_1=0.306), \omega_0=0.091, (\omega_1=1)$	4.198*	Not allowed
	Alternative	-12692.749587	$\rho_0=0.586, \rho_1=0.327, (\rho_2=0.085), \omega_0=0.097, (\omega_1=1), \omega_2=2.516$		Mouse: 1 site $p > 99\%$ : 180Q Rat: 1 site $p > 99\%$ : 180Q
Rat	Null	-12694.360640	$\rho_0=0.481, (\rho_1=0.260), \omega_0=0.100, (\omega_1=1)$	21.139***	Not allowed
	Alternative	-12683.791136	$\rho_0=0.561, \rho_1=0.300, (\rho_2=0.136), \omega_0=0.101, (\omega_1=1), \omega_2=5.662$		1 Site $p > 99\%$ : 79S; 5 sites $p > 95\%$ : 86F, 145A, 186R, 514F, 696I
Chiroptera (microbat, megabat)	Null	-12709.210769	$\rho_0=0.614, (\rho_1=0.350), \omega_0=0.106, (\omega_1=1)$	7.019**	Not allowed
	Alternative	-12705.700835	$\rho_0=0.630, \rho_1=0.355, (\rho_2=0.012), \omega_0=0.109, (\omega_1=1), \omega_2=11.651$		Microbat: 1 site $p > 95\%$ : 12L Megabat: 1 site $p > 95\%$ : 12W
OVGP1—parameter estimates and likelihood scores for branch-site models for species					
Dog	Null	-4378.052266	$\rho_0=0.519, (\rho_1=0.239), \omega_0=0.074, (\omega_1=1)$	3.908*	Not allowed
	Alternative	-4376.097858	$\rho_0=0.621, \rho_1=0.279, (\rho_2=0.098), \omega_0=0.076, (\omega_1=1), \omega_2=4.454$		1 Site $p > 95\%$ : 70T
PLG—parameter estimates and likelihood scores for branch-site models for species					
Pig	Null	-13182.998657	$\rho_0=0.775, (\rho_1=0.197), \omega_0=0.072, (\omega_1=1)$	6.215*	Not allowed
	Alternative	-13179.890971	$\rho_0=0.792, \rho_1=0.199, (\rho_2=0.006), \omega_0=0.073, (\omega_1=1), \omega_2=21.805$		1 Site $p > 95\%$ : 620K
S100A11—parameter estimates and likelihood scores for branch-site models for species					
Primates (human, orangutan)	Null	-1468.549296	$\rho_0=0.795, (\rho_1=0.153), \omega_0=0.044, (\omega_1=1)$	4.798*	Not allowed
	Alternative	-1466.149841	$\rho_0=0.825, \rho_1=0.158, (\rho_2=0.015), \omega_0=0.044, (\omega_1=1), \omega_2=19.752$		Human: 1 site $p > 95\%$ : 78S Orangutan: 1 site $p > 95\%$ : 78S
Rat	Null	-1464.292877	$\rho_0=0.697, (\rho_1=0.128), \omega_0=0.038, (\omega_1=1)$	6.934**	Not allowed
	Alternative	-1460.825580	$\rho_0=0.811, \rho_1=0.150, (\rho_2=0.036), \omega_0=0.041, (\omega_1=1), \omega_2=45.087$		1 Site $p > 95\%$ : 8L

**Table 2** (continued)

Species	Model	$\ell^a$	Estimates of parameters <sup>b</sup>	$2\Delta I^c$	Positively selected sites (BEB) <sup>d</sup>
SPAM1—parameter estimates and likelihood scores for branch-site models for species					
Mammals	Null	-9332.838271	$\rho_0=0.486, (\rho_1=0.222), \omega_0=0.088, (\omega_1=1)$	55.457***	Not allowed
	Alternative	-9305.109656	$\rho_0=0.461, \rho_1=0.470, (\rho_2=0.067), \omega_0=0.086, (\omega_1=1), \omega_2=2.833$		Rabbit: 3 sites $p > 99\%$ : 121Q, 131L, 403G Mouse: 3 sites $p > 99\%$ : 121S, 131L, 401Q Rat: 3 sites $p > 99\%$ : 121S, 131H, 403R Tarsier: 3 sites $p > 99\%$ : 122K, 132S, 403I Orangutan: 3 sites $p > 99\%$ : 122D, 132L, 403D Human: 3 sites $p > 99\%$ : 122D, 132T, 403D Microbat: 3 sites $p > 99\%$ : 121A, 131E, 403K Dog: 3 sites $p > 99\%$ : 122K, 132S, 404V Panda: 3 sites $p > 99\%$ : 122K, 132A, 404D Cat: 3 sites $p > 99\%$ : 122K, 132S, 404V Pig: 3 sites $p > 99\%$ : 122R, 132L, 403A Dolphin: 3 sites $p > 99\%$ : 122E, 132S, 404Q Cow: 3 sites $p > 99\%$ : 123S, 133A, 405M Sheep: 3 sites $p > 99\%$ : 122N, 132A, 404M Horse: 3 sites $p > 99\%$ : 118A, 128S, 399K
Rodents (mouse, rat, rab- bit) and primates (human, orangutan, tarsier)	Null	-9332.291017	$\rho_0=0.481, (\rho_1=0.494), \omega_0=0.085, (\omega_1=1)$	20.661***	Not allowed
	Alternative	-9321.960143	$\rho_0=0.474, \rho_1=0.485, (\rho_2=0.039), \omega_0=0.087, (\omega_1=1), \omega_2=4.121$		Rabbit: 1 site $p > 95\%$ : 131L Mouse: 1 site $p > 95\%$ : 131L Rat: 1 site $p > 95\%$ : 131H Tarsier: 1 site $p > 95\%$ : 132S Orangutan: 1 site $p > 95\%$ : 132L Human: 1 site $p > 95\%$ : 132T
Microbat	Null	-9328.404724	$\rho_0=0.415, (\rho_1=0.432), \omega_0=0.077, (\omega_1=1)$	21.585***	Not allowed
	Alternative	-9317.611891	$\rho_0=0.443, \rho_1=0.456, (\rho_2=0.100), \omega_0=0.081, (\omega_1=1), \omega_2=11.322$		1 site $p > 99\%$ : 374L ; 2 sites $p > 95\%$ : 392A, 423M
Pig	Null	-9328.072812	$\rho_0=0.400, (\rho_1=0.422), \omega_0=0.078, (\omega_1=1)$	10.202**	Not allowed
	Alternative	-9322.971449	$\rho_0=0.468, \rho_1=0.491, (\rho_2=0.039), \omega_0=0.081, (\omega_1=1), \omega_2=13.407$		1 site $p > 95\%$ : 77T ; 1 site $p > 99\%$ : 413N
SPP1—parameter estimates and likelihood scores for branch-site models for species					
Microbat, megabat, horse, dog, panda, cat, pig, dol- phin, cow, sheep	Null	-3405.427792	$\rho_0=0.521, (\rho_1=0.158), \omega_0=0.180, (\omega_1=1)$	4.387*	Not allowed
	Alternative	-3403.234241	$\rho_0=0.611, \rho_1=0.142, (\rho_2=0.245), \omega_0=0.226, (\omega_1=1), \omega_2=1.603$		Microbat: 3 sites $p > 95\%$ : 291H, 299P Megabat: 3 sites $p > 95\%$ : 108S, 254R, 262P Horse: 3 sites $p > 95\%$ : 107S, 280S, 288R Dog: 3 sites $p > 95\%$ : 104L, 244H, 252S Panda: 3 sites $p > 95\%$ : 104G, 266H, 274S Cat: 3 sites $p > 95\%$ : 104T, 266N, 274P Pig: 3 sites $p > 95\%$ : 107A, 271H, 279S Dolphin: 3 sites $p > 95\%$ : 87A, 256R, 264L Cow: 3 sites $p > 95\%$ : 103S, 244L, 252H Sheep: 3 sites $p > 95\%$ : 105S, 247L, 255H
Dog, panda, cat, pig, dolphin, cow, sheep	Null	-3408.017517	$\rho_0=0.559, (\rho_1=0.230), \omega_0=0.201, (\omega_1=1)$	6.525*	Not allowed
	Alternative	-3404.754608	$\rho_0=0.602, \rho_1=0.234, (\rho_2=0.162), \omega_0=0.221, (\omega_1=1), \omega_2=2.002$		Dog: 1 site $p > 95\%$ : 104L Panda: 1 site $p > 95\%$ : 104G Cat: 1 site $p > 95\%$ : 104T Pig: 1 site $p > 95\%$ : 107A Dolphin: 1 site $p > 95\%$ : 87A Cow: 1 site $p > 95\%$ : 103S Sheep: 1 site $p > 95\%$ : 105S
Dog, panda, cat	Null	-3404.586199	$\rho_0=0.443, (\rho_1=0.204), \omega_0=0.197, (\omega_1=1)$	18.383***	Not allowed
	Alternative	-3395.394529	$\rho_0=0.556, \rho_1=0.245, (\rho_2=0.197), \omega_0=0.207, (\omega_1=1), \omega_2=4.297$		Dog: 2 sites $p > 99\%$ : 104L, 240L; 2 sites $p > 95\%$ : 75S, 182S Panda: 2 sites $p > 99\%$ : 104G, 262I; 2 sites $p > 95\%$ : 75S, 182S Cat: 2 sites $p > 99\%$ : 104T, 262T; 2 sites $p > 95\%$ : 75A, 182S



**Table 2** (continued)

Species	Model	$\ell^a$	Estimates of parameters <sup>b</sup>	$2\Delta I^c$	Positively selected sites (BEB) <sup>d</sup>
Dog, panda	Null	-3410.889147	$\rho_0=0.600, (\rho_1=0.399), \omega_0=0.208, (\omega_1=1)$	10.162**	Not allowed
	Alternative	-3405.808026	$\rho_0=0.586, \rho_1=0.393, (\rho_2=0.018), \omega_0=0.208, (\omega_1=1), \omega_2=33.914$		Dog: 1 site $p > 95\%$ : 104L Panda: 1 site $p > 95\%$ : 104G
Pig, dolphin, cow, sheep	Null	-3410.383104	$\rho_0=0.584, (\rho_1=0.368), \omega_0=0.204, (\omega_1=1)$	8.998**	Not allowed
	Alternative	-3405.883945	$\rho_0=0.569, \rho_1=0.372, (\rho_2=0.057), \omega_0=0.206, (\omega_1=1), \omega_2=4.776$		Pig: 1 site $p > 95\%$ : 176Q Dolphin: 1 site $p > 95\%$ : 156Y Cow: 1 site $p > 95\%$ : 172S Sheep: 1 site $p > 95\%$ : 175S
Mouse	Null	-3408.601339	$\rho_0=0.350, (\rho_1=0.254), \omega_0=0.186, (\omega_1=1)$	7.352**	Not allowed
	Alternative	-3404.925140	$\rho_0=0.565, \rho_1=0.394, (\rho_2=0.039), \omega_0=0.195, (\omega_1=1), \omega_2=30.713$		1 Site $p > 95\%$ : 35L

<sup>a</sup>Log-likelihood values

<sup>b</sup> $\rho_0, \rho_1,$  and  $\rho_2$  are the proportions of codons subject to purifying, neutral, and positive selection, respectively.  $\omega_0, \omega_1,$  and  $\omega_2$  represented dN/dS for each class (purifying, neutral, and positive selection, respectively)

<sup>c</sup>\*Significant at  $p < 0.05$  threshold; \*\*significant at  $p < 0.01$  threshold; \*\*\*significant at  $p < 0.001$  threshold

<sup>d</sup>Bayes Empirical Bayes inference of amino acid sites under positive selection

## Discussion

Previous studies have identified numerous proteins in the OF from several mammalian species (Georgiou et al. 2005; Lamy et al. 2016; Mondejar et al. 2012; Smits et al. 2017; Soleilhavoup et al. 2016; Yu et al. 2016). Their expression changes with stages of the oestrous cycle (Lamy et al. 2016; Soleilhavoup et al. 2016), presence of gametes (Georgiou et al. 2005), or presence of embryos (Smits et al. 2017). Some of these proteins are involved in crucial steps for fertilization, such as binding between sperm and oviduct (Ignotz et al. 2007; Teijeiro et al. 2011), zona pellucida–sperm interaction (Algarra et al. 2016; Coy et al. 2012a; Goncalves et al. 2008a; Hao et al. 2006; Lachance et al. 2007; Mondejar et al. 2013; Zumoffen et al. 2013), sperm viability (Elliott et al. 2009; Lloyd et al. 2009; Moein-Vaziri et al. 2014), acrosome reaction (Chiu et al. 2010; Zhang et al. 2006), dispersion of cumulus cells (Griffiths et al. 2008), and sperm selection (Hanaue et al. 2011). The present work suggests that the high diversity of proteins present in the OF of mammals is associated with a species-specific evolutionary pattern of the corresponding genes by fairly frequent pseudogenization. Gene death is usually caused by introducing any changes that lead to a premature stop codon. Now, with the availability of complete animal genome sequences, it is possible to test hypotheses concerning gene death by searching for the trace of a pseudogene.

Pseudogenization has been previously demonstrated for OVGPI in the rat (Tian et al. 2009). Our study shows that OVGPI is lost in megabat. OVGPI, also known as oviductin, has been shown to play a role in the fertilization process in goat (Pradeep et al. 2011), pig (McCauley et al. 2003),

cow (Martus et al. 1998), buffalo (Choudhary et al. 2017), hamster (Saccary et al. 2013; Schmidt et al. 1997; Yang et al. 2015), and human (O'Day-Bowman et al. 1996; Zhao et al. 2016): it attaches to the zona pellucida and the spermatozoa, and improves in vitro fertilization and embryo development. Interestingly, targeted invalidation of Ovgp1 in the mouse is without any consequence in female reproduction (Araki et al. 2003), implying that at least in murine species, Ovgp1 is not essential for fertility. Gene loss by pseudogenization has already been reported for genes encoding seminal fluid proteins, such as for prostate-specific transglutaminase (TGM4) genes, and semenogelins (SEM1 and SEM2) in gorillas (Carnahan and Jensen-Seaman 2008), as well as TGM4 in pigs (Meslin et al. 2015).

Our study shows that PAEP is lost in tarsier, rat, mouse, rabbit, megabat, and dolphin. PAEP, also named glycodelin-A, interacts with fucosyltransferase on human sperm plasma membrane thus inhibiting spermatozoa–zona pellucida binding (Chiu et al. 2007). Moreover, pre-treatment of spermatozoa with glycodelin-A enhances zona pellucida-induced calcium influx and zona pellucida-induced acrosome reaction (Chiu et al. 2010). Thus, PAEP is involved in sperm selection and preparation for fertilization. In tarsier, rat, mouse, rabbit, megabat, and dolphin, this role may not be essential, or may be assumed by other proteins.

Two other proteins have been lost in tarsier: ANXA5 and DMBT1. ANXA5 is a candidate for the bovine oviductal epithelium sperm receptors, it binds with high affinity to heparin and related glycosaminoglycans and antibodies to this protein block sperm–oviduct binding (Talevi and Gualtieri 2010). However, since other annexins, such as ANXA1, ANXA2, and ANXA4 are also candidates for the oviductal

epithelium sperm receptors, the loss of ANXA5 may be replaced by other annexins. DMBT1 has been shown to be involved in the mechanism of fertilization in equine and porcine species (Ambruosi et al. 2013): pre-incubation of oocytes with DMBT1 induces an increase in *in vitro* fertilization rate and an interaction between DMBT1 and spermatozoa has been shown using surface plasmon resonance studies. Moreover, DMBT1 has been proposed to be implicated in sperm selection in the oviduct through acrosome alteration and suppression of motility (Teijeiro and Marini 2012). As previously mentioned for ANXA5, in tarsier, other candidates for sperm binding and selection are probably involved.

Finally, our study shows that PTGDS and PLG are lost in microbat and megabat, respectively. In bovine, prostaglandin D2 synthase (PTGDS) is involved in sperm binding to the ZP, *in vitro* fertilization and embryonic development, reaction with both oocytes and spermatozoa (Goncalves et al. 2008a, b). In swine, plasminogen (PLG) regulates sperm entry into the oocyte: sperm binding to oocytes triggers the releasing of plasminogen activators and the generated plasmin causes supernumerary spermatozoa detachment from the zona pellucida (Coy et al. 2012b). Since these proteins are lost in microbat and megabat, some other proteins are probably involved in this regulation.

Overall, this study revealed that the evolution of mammals was accompanied by the progressive loss of genes coding for oviductal proteins involved in the process of fertilization. The significance of such loss remains to be further investigated. It may be related, at least for bats, to sperm storage in the oviduct. Prolonged sperm storage in the female genital tract over winter is a common feature of reproduction in some bats, so that they can mate in the autumn but postpone fertilization until the spring (Holt 2011). Thus, oviductal proteins involved in sperm attachment to the oviduct or sperm selection within the oviduct may have evolved in a different way in these species.

In our study, gene duplication has been observed in several species: orangutan, rat, mouse, cow, sheep, pig, horse, cat, and dog. The duplication could be a preservation mechanism through which some daughter copies keep ancient functions, whereas others evolve toward new biological functions.

OVGP1 is lost in rat and megabat. Cow, sheep, and pig have two copies of the gene. Porcine OVGP1 has been shown to have an effect on the number of sperm bound to the zona pellucida, on sperm penetration rate, and polyspermy rate (Kouba et al. 2000). Sperm adhesion molecule 1 (SPAM1) is duplicated in rat, mouse, cow, sheep, and horse. SPAM1 is a hyaluronidase which can bind to sperm during capacitation and increase cumulus dispersal efficiency and ability of sperm to penetrate the cumulus of oocytes (Griffiths et al. 2008; Martin-DeLeon 2006). Prothmann et al. (2012) found that the rate of evolution of the gene encoding SPAM1 could

be correlated with relative testis weight in monkeys as well as the uni-male versus multi-male breeding system. This suggests that different levels of sperm competition might account for species-specific sequence evolution of SPAM1 in these species. The evolution of gene encoding the male zonadhesin gene also may be related to the uni-male breeding system in other primates (Herlyn and Zischler 2007, 2008). In the present work, we find a positive selection of this gene in all the studied species. However, it is too speculative here to make a hypothesis concerning breeding system because we did the work on a greater diversity of mammals, with wild and domesticated animals, and very different breeding systems. However, such a question would be very interesting to investigate with a smaller and especially more homogenous group of species.

S100 calcium-binding protein A11 (S100A11) is duplicated in rat, pig, and cat. This protein plays a role in sperm selection through its action on cumulus cells (Hanaue et al. 2011). Annexin A2 (ANXA2) is duplicated in sheep and dog. ANXA2 has been shown to be present at the apical surface of the oviductal epithelial cells in sow and it is the main sperm binding isoform in pig (Teijeiro et al. 2009). Other annexins, such as ANXA1, ANXA4, and ANXA5 are other candidates for the oviductal epithelium sperm receptors (Talevi and Gualtieri 2010). Finally, HSPA8 is duplicated in orangutan, rat, sheep, and pig. HSPA8 is involved in maintenance of bovine, ovine, and porcine sperm survival in the oviduct (Elliott et al. 2009; Lloyd et al. 2009). Overall, all the proteins that are duplicated are involved in sperm selection and competence to reach the oocyte.

In our study, we have observed branch-site positive selection in all the species that we have studied.

SPAM1 showed positive selection in all the species analyzed: human, orangutan, tarsier, rat, mouse, rabbit, microbat, megabat, cow, sheep, dolphin, pig, horse, cat, dog, and panda. Moreover, SPAM1 has not been lost in the sampled species, but was duplicated in rat, mouse, cow, sheep, and horse. Since it is involved in the ability of sperm to penetrate the cumulus of oocytes (Griffiths et al. 2008; Martin-DeLeon 2006), this could highlight its species-specific role in gamete recognition.

Lactotransferrin (LTF) also showed positive selection in all the species analyzed. Moreover, LTF has neither been lost nor duplicated in the sampled species. LTF has been shown to interact with spermatozoa and oocytes and to modulate gamete interaction (Zumoffen et al. 2013). This rapid evolution may contribute to the diversity of mating systems and may explain part of the loss in interspecific fecundity after speciation (Meslin et al. 2015).

Osteopontin (SPP1) showed positive selection in many species: mouse, microbat, megabat, cow, sheep, dolphin, pig, horse, cat, dog, and panda. Moreover, SPP1 has neither been lost nor duplicated in the sampled species. SPP1 has been

shown to improve bovine sperm capacitation (Monaco et al. 2009) and in vitro development of porcine embryos (Hao et al. 2008). Moreover, osteopontin plays an important role in the regulation of pig polyspermic fertilization; it decreases polyspermy and increases fertilization efficiency during IVF (Hao et al. 2006).

Annexins showed positive selection in cow (ANXA4), sheep (ANXA4 and ANXA5), dolphin (ANXA1), and dog (ANXA2). Several annexins, such as ANXA1, ANXA2, ANXA4, and ANXA5 are involved in oviductal epithelium sperm receptors. Their rapid evolution may also contribute to the loss in interspecific fecundity.

S100A11, also involved in the interaction between sperm and cumulus cells, showed positive selection only in human, orangutan and rat, and gene duplication in rat, pig and cat.

Heat shock 70 kDa protein 5 (HSPA5) showed positive selection only in the rabbit. It has neither been lost nor duplicated in the sampled species, contrary to HSPA8, which was duplicated in orangutan, rat, sheep, and pig. Heat shock proteins A5 and A8 play a role in the modulation of boar sperm function in the oviductal reservoir (Yeste et al. 2014).

Plasminogen (PLG) showed positive selection only in the pig. In porcine, PLG regulates sperm entry into the oocyte (Coy et al. 2012b). Its rapid evolution could highlight the species-specific gamete recognition.

OVGP1 showed positive selection only in the dog. It is lost in rat and megabat, and duplicated in cow, sheep, and pig. It has been shown to have an effect on the number of sperm bound to the zona pellucida, on sperm penetration rate, and polyspermy rate (Kouba et al. 2000).

## Conclusion

Overall, the present work highlights the evolution of genes encoding proteins involved in gamete transportation and maturation, sperm capacitation, and early embryo development. These genes are subjected to a particularly rapid evolution by duplication as well as species-specific positive selection, suggesting a diversification of their functions and of their role in the different mammalian species, with possible impact on the process of speciation, as proposed for the evolution of genes encoding proteins of seminal fluid of different species (Meslin et al. 2015). Moreover, several previous works have suggested that genes encoding proteins involved in sperm–egg interactions might undergo co-evolution and will also have correlated evolutionary rates due to compensatory changes on both the sperm and egg (Claw et al. 2014). In the present paper and one previous one (Meslin et al. 2015), one could hypothesize such a kind of co-evolution for proteins of seminal fluid and proteins of oviductal fluid.

**Funding** This work was supported by INRA and grant from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No. 677353. Carla Moros-Nicolás was supported by a post-doctoral fellowship from “Consejería de Educación y Universidades de la CARM, Fundación Séneca-Agencia de Ciencia y Tecnología de la Región de Murcia.”

## Compliance with Ethical Standards

**Conflict of interest** The authors declare that they have no conflict of interest.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Algarra B, Han L, Soriano-Ubeda C, Aviles M, Coy P, Jovine L, Jimenez-Movilla M (2016) The C-terminal region of OVGP1 remodels the zona pellucida and modifies fertility parameters. *Sci Rep* 6:32556
- Ambrosi B, Accogli G, Douet C, Canepa S, Pascal G, Monget P, Moros C, Holmskov U, Mollenhauer J, Robbe-Masselot C, Vidal O, Desantis S, Goudet G (2013) Deleted in malignant brain tumor 1 is secreted in the oviduct and involved in the mechanism of fertilization in equine and porcine species. *Reproduction* 146:119
- Araki Y, Nohara M, Yoshida-Komiya H, Kuramochi T, Ito M, Hoshi H, Shinkai Y, Sendai Y (2003) Effect of a null mutation of the oviduct-specific glycoprotein gene on mouse fertilization. *Biochem J* 374:551
- Aviles M, Gutierrez-Adan A, Coy P (2010) Oviductal secretions: will they be key factors for the future ARTs? *Mol Hum Reprod* 16:896
- Bodmer M, Ashburner M (1984) Conservation and change in the DNA sequences coding for alcohol dehydrogenase in sibling species of *Drosophila*. *Nature* 309:425
- Busset J, Cabau C, Meslin C, Pascal G (2011) PhyleasProg: a user-oriented web server for wide evolutionary analyses. *Nucleic Acids Res* 39:W479
- Carnahan SJ, Jensen-Seaman MI (2008) Hominoid seminal protein evolution and ancestral mating behavior. *Am J Primatol* 70:939
- Chiu PC, Chung MK, Koistinen R, Koistinen H, Seppala M, Ho PC, Ng EH, Lee KF, Yeung WS (2007) Glycodelin—a interacts with fucosyltransferase on human sperm plasma membrane to inhibit spermatozoa-zona pellucida binding. *J Cell Sci* 120:33
- Chiu PC, Wong BS, Lee CL, Lam KK, Chung MK, Lee KF, Koistinen R, Koistinen H, Gupta SK, Seppala M, Yeung WS (2010) Zona pellucida-induced acrosome reaction in human spermatozoa is potentiated by glycodelin—a via down-regulation of extracellular signal-regulated kinases and up-regulation of zona pellucida-induced calcium influx. *Hum Reprod* 25:2721
- Choudhary S, Kumaresan A, Kumar M, Chhillar S, Malik H, Kumar S, Kaushik JK, Datta TK, Mohanty AK (2017) Effect of recombinant and native buffalo OVGP1 on sperm functions and in vitro embryo development: a comparative study. *J Anim Sci Biotechnol* 8:69
- Claw KG, George RD, Swanson WJ (2014) Detecting coevolution in mammalian sperm-egg fusion proteins. *Mol Reprod Dev* 81:531

- Coy P, Yanagimachi R (2015) The common and species-specific roles of oviductal proteins in mammalian fertilization and embryo development. *Bioscience* 65:973
- Coy P, Canovas S, Mondejar I, Saavedra MD, Romar R, Grullon L, Matas C, Aviles M (2008) Oviduct-specific glycoprotein and heparin modulate sperm-zona pellucida interaction during fertilization and contribute to the control of polyspermy. *Proc Natl Acad Sci USA* 105:15809
- Coy P, Garcia-Vazquez FA, Visconti PE, Aviles M (2012a) Roles of the oviduct in mammalian fertilization. *Reproduction* 144:649
- Coy P, Jimenez-Movilla M, Garcia-Vazquez FA, Mondejar I, Grullon L, Romar R (2012b) Oocytes use the plasminogen-plasmin system to remove supernumerary spermatozoa. *Hum Reprod* 27:1985
- Elliott RM, Lloyd RE, Fazeli A, Sostaric E, Georgiou AS, Satake N, Watson PF, Holt WV (2009) Effects of HSPA8, an evolutionarily conserved oviductal protein, on boar and bull spermatozoa. *Reproduction* 137:191
- Georgiou AS, Sostaric E, Wong CH, Snijders AP, Wright PC, Moore HD, Fazeli A (2005) Gametes alter the oviductal secretory proteome. *Mol Cell Proteomics* 4:1785
- Goncalves RF, Barnabe VH, Killian GJ (2008a) Pre-treatment of cattle sperm and/or oocyte with antibody to lipocalin type prostaglandin D synthase inhibits in vitro fertilization and increases sperm-oocyte binding. *Anim Reprod Sci* 106:188
- Goncalves RF, Staros AL, Killian GJ (2008b) Oviductal fluid proteins associated with the bovine zona pellucida and the effect on in vitro sperm-egg binding, fertilization and embryo development. *Reprod Domest Anim* 43:720
- Griffiths GS, Miller KA, Galileo DS, Martin-DeLeon PA (2008) Murine SPAM1 is secreted by the estrous uterus and oviduct in a form that can bind to sperm during capacitation: acquisition enhances hyaluronic acid-binding ability and cumulus dispersal efficiency. *Reproduction* 135:293
- Hanaue M, Miwa N, Uebi T, Fukuda Y, Katagiri Y, Takamatsu K (2011) Characterization of S100A11, a suppressive factor of fertilization, in the mouse female reproductive tract. *Mol Reprod Dev* 78:91
- Hao Y, Mathialagan N, Walters E, Mao J, Lai L, Becker D, Li W, Critser J, Prather RS (2006) Osteopontin reduces polyspermy during in vitro fertilization of porcine oocytes. *Biol Reprod* 75:726
- Hao Y, Murphy CN, Spate L, Wax D, Zhong Z, Samuel M, Mathialagan N, Schatten H, Prather RS (2008) Osteopontin improves in vitro development of porcine embryos and decreases apoptosis. *Mol Reprod Dev* 75:291
- Hellberg ME, Vacquier VD (1999) Rapid evolution of fertilization selectivity and lysin cDNA sequences in teguline gastropods. *Mol Biol Evol* 16:839
- Herlyn H, Zischler H (2007) Sequence evolution of the sperm ligand zonadhesin correlates negatively with body weight dimorphism in primates. *Evolution* 61:289
- Herlyn H, Zischler H (2008) The molecular evolution of sperm zonadhesin. *Int J Dev Biol* 52:781
- Holt WV (2011) Mechanisms of sperm storage in the female reproductive tract: an interspecies comparison. *Reprod Domest Anim* 46(Suppl 2):68
- Hunter RH (1998) Have the Fallopian tubes a vital role in promoting fertility? *Acta Obstet Gynecol Scand* 77:475
- Hunter RH (2012) Components of oviduct physiology in eutherian mammals. *Biol Rev Camb Philos Soc* 87:244
- Ignatz GG, Cho MY, Suarez SS (2007) Annexins are candidate oviductal receptors for bovine sperm surface proteins and thus may serve to hold bovine sperm in the oviductal reservoir. *Biol Reprod* 77:906
- Kouba AJ, Abeydeera LR, Alvarez IM, Day BN, Buhi WC (2000) Effects of the porcine oviduct-specific glycoprotein on fertilization, polyspermy, and embryonic development in vitro. *Biol Reprod* 63:242
- Lachance C, Bailey JL, Leclerc P (2007) Expression of Hsp60 and Grp78 in the human endometrium and oviduct, and their effect on sperm functions. *Hum Reprod* 22:2606
- Lamy J, Labas V, Harichaux G, Tsikis G, Mermillod P, Saint-Dizier M (2016) Regulation of the bovine oviductal fluid proteome. *Reproduction* 152:629
- Leese HJ, Tay JI, Reischl J, Downing SJ (2001) Formation of Fallopian tubal fluid: role of a neglected epithelium. *Reproduction* 121:339
- Leese HJ, Hugentobler SA, Gray SM, Morris DG, Sturmey RG, Whitear SL, Sreenan JM (2008) Female reproductive tract fluids: composition, mechanism of formation and potential role in the developmental origins of health and disease. *Reprod Fertil Dev* 20:1
- Li WH, Gojobori T, Nei M (1981) Pseudogenes as a paradigm of neutral evolution. *Nature* 292:237
- Lloyd RE, Elliott RM, Fazeli A, Watson PF, Holt WV (2009) Effects of oviductal proteins, including heat shock 70 kDa protein 8, on survival of ram spermatozoa over 48 h in vitro. *Reprod Fertil Dev* 21:408
- Makalowski W, Boguski MS (1998) Evolutionary parameters of the transcribed mammalian genome: an analysis of 2820 orthologous rodent and human sequences. *Proc Natl Acad Sci USA* 95:9407
- Marin-Briggiler CI, Gonzalez-Echeverria MF, Munuce MJ, Ghersevich S, Caille AM, Hellman U, Corrigan VM, Vazquez-Levin MH (2010) Glucose-regulated protein 78 (Grp78/BiP) is secreted by human oviduct epithelial cells and the recombinant protein modulates sperm-zona pellucida binding. *Fertil Steril* 93:1574
- Martin-DeLeon PA (2006) Epididymal SPAM1 and its impact on sperm function. *Mol Cell Endocrinol* 250:114
- Martus NS, Verhage HG, Mavrogianis PA, Thibodeaux JK (1998) Enhancement of bovine oocyte fertilization in vitro with a bovine oviductal specific glycoprotein. *J Reprod Fertil* 113:323
- McCauley TC, Buhi WC, Wu GM, Mao J, Caamano JN, Didion BA, Day BN (2003) Oviduct-specific glycoprotein modulates sperm-zona binding and improves efficiency of porcine fertilization in vitro. *Biol Reprod* 69:828
- Meslin C, Brimau F, Nagnan-Le Meillour P, Callebaut I, Pascal G, Monget P (2011) The evolutionary history of the SAL1 gene family in eutherian mammals. *BMC Evol Biol* 11:148
- Meslin C, Laurin M, Callebaut I, Druart X, Monget P (2015) Evolution of species-specific major seminal fluid proteins in placental mammals by gene death and positive selection. *Contrib Zoo* 84:217
- Moein-Vaziri N, Phillips I, Smith S, Alminana C, Maside C, Gil MA, Roca J, Martinez EA, Holt WV, Pockley AG, Fazeli A (2014) Heat-shock protein A8 restores sperm membrane integrity by increasing plasma membrane fluidity. *Reproduction* 147:719
- Monaco E, Gasparrini B, Boccia L, De Rosa A, Attanasio L, Zicarelli L, Killian G (2009) Effect of osteopontin (OPN) on in vitro embryo development in cattle. *Theriogenology* 71:450
- Mondejar I, Acuna OS, Izquierdo-Rico MJ, Coy P, Aviles M (2012) The oviduct: functional genomic and proteomic approach. *Reprod Domest Anim* 47(Suppl 3):22
- Mondejar I, Martinez-Martinez I, Aviles M, Coy P (2013) Identification of potential oviductal factors responsible for zona pellucida hardening and monospermy during fertilization in mammals. *Biol Reprod* 89:67
- Nielsen R, Yang Z (1998) Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. *Genetics* 148:929
- O'Day-Bowman MB, Mavrogianis PA, Reuter LM, Johnson DE, Fazleabas AT, Verhage HG (1996) Association of oviduct-specific glycoproteins with human and baboon (*Papio anubis*) ovarian oocytes and enhancement of human sperm binding to human hemizonae following in vitro incubation. *Biol Reprod* 54:60

- Pradeep MA, Jagadeesh J, De AK, Kaushik JK, Malakar D, Kumar S, Dang AK, Das SK, Mohanty AK (2011) Purification, sequence characterization and effect of goat oviduct-specific glycoprotein on in vitro embryo development. *Theriogenology* 75:1005
- Prothmann A, Laube I, Dietz J, Roos C, Mengel K, Zischler H, Herlyn H (2012) Sexual size dimorphism predicts rates of sequence evolution of SPERM adhesion molecule 1 (SPAM1, also PH-20) in monkeys, but not in hominoids (apes including humans). *Mol Phylogenet Evol* 63:52
- Saccary L, She YM, Oko R, Kan FW (2013) Hamster oviductin regulates tyrosine phosphorylation of sperm proteins during in vitro capacitation. *Biol Reprod* 89:38
- Schmidt A, Mavrogianis PA, O'Day-Bowman MB, Verhage HG (1997) Species-specific effect of oviductal glycoproteins on hamster sperm binding to hamster oocytes. *Mol Reprod Dev* 46:201
- Seytanoglu A, Georgiou AS, Sostaric E, Watson PF, Holt WV, Fazeli A (2008) Oviductal cell proteome alterations during the reproductive cycle in pigs. *J Proteome Res* 7:2825
- Singh RS, Kulathinal RJ (2000) Sex gene pool evolution and speciation: a new paradigm. *Genes Genet Syst* 75:119
- Smits K, Nelis H, Van Steendam K, Govaere J, Roels K, Ververs C, Leemans B, Wydooghe E, Deforce D, Van Soom A (2017) Proteome of equine oviductal fluid: effects of ovulation and pregnancy. *Reprod Fertil Dev* 29:1085
- Soleilhavoup C, Riou C, Tsikis G, Labas V, Harichaux G, Kohnke P, Reynaud K, de Graaf SP, Gerard N, Druart X (2016) Proteomes of the female genital tract during the oestrous cycle. *Mol Cell Proteomics* 15:93
- Storey JD, Tibshirani R (2003) Statistical significance for genomewide studies. *Proc Natl Acad Sci USA* 100:9440
- Suyama M, Torrents D, Bork P (2006) PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Res* 34:W609
- Swanson WJ, Vacquier VD (2002) The rapid evolution of reproductive proteins. *Nat Rev Genet* 3:137
- Swanson WJ, Nielsen R, Yang Q (2003) Pervasive adaptive evolution in mammalian fertilization proteins. *Mol Biol Evol* 20:18
- Talevi R, Gualtieri R (2010) Molecules involved in sperm-oviduct adhesion and release. *Theriogenology* 73:796
- Teijeiro JM, Marini PE (2012) The effect of oviductal deleted in malignant brain tumor 1 over porcine sperm is mediated by a signal transduction pathway that involves pro-AKAP4 phosphorylation. *Reproduction* 143:773
- Teijeiro JM, Ignatz GG, Marini PE (2009) Annexin A2 is involved in pig (*Sus scrofa*) sperm-oviduct interaction. *Mol Reprod Dev* 76:334
- Teijeiro JM, Dapino DG, Marini PE (2011) Porcine oviduct sperm binding glycoprotein and its deleterious effect on sperm: a mechanism for negative selection of sperm? *Biol Res* 44:329
- Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22:4673
- Tian X, Pascal G, Fouchecourt S, Pontarotti P, Monget P (2009) Gene birth, death, and divergence: the different scenarios of reproduction-related gene evolution. *Biol Reprod* 80:616
- Turner LM, Hoekstra HE (2008) Causes and consequences of the evolution of reproductive proteins. *Int J Dev Biol* 52:769
- Vacquier VD (1998) Evolution of gamete recognition proteins. *Science* 281:1995
- Wang C, Liu Y, Chang C, Wu S, Gao J, Zhang Y, Chen Y, Zhong F, Deng G (2016) Human fallopian tube proteome shows high coverage of mesenchymal stem cells associated proteins. *Biosci Rep* 36:e00297
- Wong WS, Yang Z, Goldman N, Nielsen R (2004) Accuracy and power of statistical methods for detecting adaptive evolution in protein coding sequences and for identifying positively selected sites. *Genetics* 168:1041
- Yang Z (2000) Maximum likelihood estimation on large phylogenies and analysis of adaptive evolution in human influenza virus A. *J Mol Evol* 51:423
- Yang Z (2007) PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol* 24:1586
- Yang Z, Nielsen R (2002) Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages. *Mol Biol Evol* 19:908
- Yang Z, Wong WS, Nielsen R (2005) Bayes empirical bayes inference of amino acid sites under positive selection. *Mol Biol Evol* 22:1107
- Yang X, Zhao Y, Yang X, Kan FW (2015) Recombinant hamster oviductin is biologically active and exerts positive effects on sperm functions and sperm-oocyte binding. *PLoS ONE* 10:e0123003
- Yeste M, Holt WV, Bonet S, Rodriguez-Gil JE, Lloyd RE (2014) Viable and morphologically normal boar spermatozoa alter the expression of heat-shock protein genes in oviductal epithelial cells during co-culture in vitro. *Mol Reprod Dev* 81:805
- Yu H, Reiser J, Besenfelder U, Razzazi-Fazeli E, Bergquist J, Brem G, Artemenko K, Mayrhofer C (2016) Exploring the oviductal fluid proteome by a lectin-based affinity approach. *Proteomics* 16:2962
- Zhang J, Nielsen R, Yang Z (2005) Evaluation of an improved branch-site likelihood method for detecting positive selection at the molecular level. *Mol Biol Evol* 22:2472
- Zhang M, Hong H, Zhou B, Jin S, Wang C, Fu M, Wang S, Xia G (2006) The expression of atrial natriuretic peptide in the oviduct and its functions in pig spermatozoa. *J Endocrinol* 189:493
- Zhao Y, Yang X, Jia Z, Reid RL, Leclerc P, Kan FW (2016) Recombinant human oviductin regulates protein tyrosine phosphorylation and acrosome reaction. *Reproduction* 152:561
- Zumoffen CM, Gil R, Caille AM, Morente C, Muncu MJ, Ghersevich SA (2013) A protein isolated from human oviductal tissue in vitro secretion, identified as human lactoferrin, interacts with spermatozoa and oocytes and modulates gamete interaction. *Hum Reprod* 28:1297