Journal of Heredity 2011:102(3):275–282
doi:10.1093/jhered/esq131
Advance Access publication January 27, 2011

© The American Genetic Association. 2011. All rights reserved.
For permissions, please email: journals.permissions@oup.com.

Evolution of 2 Reproductive Proteins, ZP3 and PKDREJ, in Cetaceans

Ana R. Amaral, Luciana M. Möller, Luciano B. Beheregaray, and M. Manuela Coelho

From the Centro de Biologia Ambiental, Faculdade de Ciências, Universidade de Lisboa, Campo Grande, 1749-016 Lisbon, Portugal (Amaral and Coelho); the Department of Biological Sciences, Macquarie University, Sydney, NSW 2109, Australia (Amaral and Beheregaray); and the School of Biological Sciences, Flinders University, Adelaide, SA 5001, Australia (Möller and Beheregaray).

Address correspondence to Ana R. Amaral, Departamento de Biologia Animal, Faculdade de Ciências, da Universidade de Lisboa, Edifício C2, Campo Grande, 1749-016 Lisbon, Portugal, or e-mail: aramaral@fc.ul.pt.

Abstract

The rapid evolution of proteins involved in reproduction has been documented in several animal taxa. This is thought to be the result of forces involved in sexual selection and is expected to be particularly strong in promiscuous mating systems. In this study, a range of cetacean species were used to analyze the patterns of evolution in 2 reproductive proteins involved in fertilization: the zona pellucida 3 (ZP3), present in the egg coat, and PKDREJ, localized in the sperm head. We targeted exons 6 and 7 of ZP3 and a part of the REJ domain in PKDREJ for a total of 958 bp in 18 species. We found very low levels of amino acid sequence divergence in both proteins, a very weak signal of positive selection in ZP3 and no signal in PKDREJ. These results were consistent with previous reports of a slow rate of molecular evolution in cetaceans but unexpected due to the existence of promiscuous mating systems in these species. The results raise questions about the evolution of reproductive isolation and species recognition in whales and dolphins.

Key words: *Cetacea, PKDREJ, sexual selection, ZP3*

Recent research on fertilization proteins, those mediating sperm–egg interactions, has revealed a pattern of rapid adaptive evolution in several animal groups, such as in marine invertebrates, birds, and mammals (Metz et al. 1998; Swanson et al. 2003; Turner and Hoekstra 2006; Calkins et al. 2007). This widespread phenomenon may have important consequences, like the establishment of barriers to fertilization that could lead to speciation (Swanson and Vacquier 2003). The selective forces of sperm competition, sexual selection, and sexual conflict have been suggested as drivers of the rapid evolution of these proteins (Swanson and Vacquier 2002). In mammals, the initial binding of sperm to the egg coat is thought to be the critical step of sperm–egg recognition (Wassarman and Litscher 2001). The egg coat comprises at least 3 glycoproteins with zona pellucida (ZP) domains: ZP1, ZP2, and ZP3, the latter being generally accepted to be the natural agonist that initiates the acrosome reaction on binding of sperm to egg (Wassarman et al. 2001). Moreover, ZP3 is one of the best-characterized mammalian fertilization proteins, containing a region described as the "sperm-combining" region (Chen et al. 1998; Wassarman et al. 2004).

PKDREJ is a protein localized in the plasma membrane of the acrosomal crescent region of the sperm head, whose expression has only been detected in the spermatogenic lineage (Butscheid et al. 2006). It has been recently shown that this protein controls acrosome exocytosis through the process of capacitation (Sutton et al. 2008), which represents a time delay between insemination and fertilization. In species where sperm competition exists as a form of postcopulatory sexual selection, genes that control the duration of capacitation could provide a selective paternal advantage and therefore could be targets of positive selection (Birkhead and Pizzari 2002). PKDREJ is therefore a candidate egg-binding sperm protein with a presumed role in cases of postcopulatory sperm competition (Sutton et al. 2008).

In mammals, initial studies of reproductive protein evolution used gene sequences from relatively distant species (Swanson et al. 2001). However, it has been suggested that an understanding of how amino acid changes affect fertilization, and consequently reproductive isolation, will only be possible by studying the patterns of evolution of these proteins in closely related species (Turner and Hoekstra 2008). Such studies have only been conducted in rodents for ZP3 (*Mus*, Jansa et al. 2003; *Peromyscus*, Turner and Hoekstra 2006; Australasian rodents, Swann et al. 2007) and primates for PKDREJ (Hamm et al. 2007). Patterns of positive selection were documented for both proteins in all

American Genetic Association
The Journal of Heredity
www.jhered.oxfordjournals.org
Downloaded from jhered.oxfordjournals.org at Flinders University of South Australia on July 7, 2011

these studies suggesting their key role in the egg–sperm binding process. Although not the aim of these studies, the authors have also found no relation between mating strategies, that is, different levels of sperm competition, in the studied species and the pattern of evolution of these proteins. Nevertheless, investigation of additional taxa is needed to confirm if this pattern of rapid evolution can be generalized across closely related and recently diverged species.

Here, we investigate the evolution of ZP3 and PKDREJ in cetaceans. These proteins were chosen based on their putative role in egg–sperm interaction as mentioned above.

Cetaceans are thought to have diverged from *Hippopotamus* 53 million years ago (Ma; Arnason et al. 2004). Extant species have split into 2 main groups around 35 Ma: the Mysticeti (baleen whales) and the Odontoceti (toothed whales). The explosive radiation of delphinoids (especially the family Delphinidae) occurred 11–12 Ma, with some dolphin species having originated as recently as 1–3 Ma (Caballero et al. 2008; McGowen et al. 2009). This group of dolphins, referred to as the STDL species complex, includes the genera *Stenella–Sousa–Tursiops–Delphinus–Lagenodelphis* (Perrin and Reeves 2004) and provides an excellent case to test whether the rapid evolution of reproductive proteins is a phenomenon generalized across different closely related taxa. Several mating systems in both the mysticetes and the odontocetes have been reported to be promiscuous and thus characterized by sperm competition and sexual conflict (Berta and Sumich 1999). Nevertheless, the different life-history patterns of the 2 groups likely resulted in different mating strategies that could have influenced the evolution of reproductive proteins. It has also been reported that at least some cetaceans have a slow rate of molecular evolution (Martin and Palumbi 1993; Jackson et al. 2009), potentially limiting the adaptive potential of those genes. Our aim in this study was to test the hypothesis that positive Darwinian selection is acting on female and male reproductive proteins in cetaceans. Such result would lend support to models that propose sexual conflict and sperm competition as selective forces driving the divergence of these proteins and confirming their role in the sperm-binding process. For that, we studied patterns of evolution in 2 reproductive proteins, ZP3 and PKDREJ, and 2 nonreproductive proteins, MC1R and BMP4, in several cetacean species.

Materials and Methods

Genomic DNA was extracted from ethanol-preserved tissue using standard phenol–chloroform extraction. Species used in this study are specified in Supplementary Table S1. Comparisons of published DNA sequences from *Bos taurus*, *Sus scrofa*, *Ovis aries*, and *Mus musculus* were used to design primers for exons 6 and 7 of ZP3 (ZP3-F1, 5′-CTGCCACCTGAAGGTCACTC-3′ and ZP3-R1, 5′-GCGACTTCGGGGAACAGA-3′). These regions were chosen because they contain several sites identified as targets of selection in an analysis of divergent mammalian species, namely the sperm-binding region (Swanson et al. 2001). For PKDREJ, we used published primers (Demere et al. 2008). Primers for BMP4 exon 3 (BMP4-F3 5′-CCACCTTGTCATACTCATCCAG-3′; BMP4-R3 5′-AGAACATCCCAGGGACCAG-3′) were designed based on an alignment of published DNA sequences. For MC1R, sequences deposited in GenBank were used (accession numbers FJ773287–FJ773291, FJ773294, FJ773296, FJ773305, and FJ773313). Polymerase chain reactions (PCRs) were performed in 25 μl reactions containing 10–100 ng DNA, 0.2 mM each dNTP, 0.3 μM each primer, 1 U *Taq* Polymerase, and 1× Taq buffer. For ZP3 and BMP4, the thermocycle profile included one cycle of 95 °C for 2 min, followed by 20 cycles of 95 °C for 30 s and a touchdown from 65 to 55 °C for 1 min decreasing by 0.5 °C per cycle, and then 72 °C for 1.50 min. This was followed by 20 cycles of 95 °C for 30 s, 55 °C for 1 min, and 72 °C for 1.50 and a final extension step of 72 °C for 10 min. For PKDREJ, the thermocycle profile consisted of an initial denaturation step at 94 °C for 2 min followed by 35 cycles of 94 °C for 30 s, 60 °C for 30 s, and 72 °C for 45 s, followed by a final extension step at 72 °C for 7 min. PCR products were separated on 1.0% agarose gels, stained with ethidium bromide, and visualized with ultraviolet light. PCR products were cleaned with Exonuclease I and Shrimp Alkaline Phosphatase and directly sequenced in both directions. Sequences were aligned in Sequencher v. 4.2., with heterozygous nucleotide sites being coded as ambiguities (position 192 in exons 6 and 7 of ZP3 and positions 153, 219, 407, and 550 in PKDREJ, see Supplementary Figure S1). The identity of the sequenced fragments was confirmed by performing a basic alignment search tool (BLAST) of amino acid sequences obtained for all genes using the BLASTp algorithm (NCBI). Phylogenetic relationships for ZP3 and PKDREJ were constructed using the maximum likelihood (ML) method as implemented in PAUP* (v. 4b10; Swofford 2003). The best evolutionary model for each gene was determined using the Akaike Information Criterion in Modeltest (v. 3.7; Posada and Buckley 2004). Bayesian trees were constructed using MrBayes (Huelsenbeck and Ronquist 2001), and 10 000 000 generations of Monte Carlo Markov Chain (MCMC) were run using the program default priors as starting values for the model. Trees were sampled every 100 generations during the analysis. The first 600 000 generations were excluded as burn-in after examining the variation in log-likelihood scores over time.

Evidence of positive selection in ZP3 and PKDREJ was tested using different ML methods as implemented in CODEML, as part of the PAML package (v. 4; Yang 2000). A likelihood ratio test (LRT) was used to examine the data for individual codons with d_N/d_S ratios (ω) significantly >1. This was done by comparing a null (neutral) model that does not allow $\omega > 1$, with an alternative model that does. The null models included a model with a d_N/d_S class between 0 and 1 and a class with $d_N/d_S = 1$ (M1a) and a model which assumes a beta distribution for d_N/d_S in the interval 0,1

Downloaded from jhered.oxfordjournals.org at Flinders University of South Australia on July 7, 2011

(M7). The alternative models include an additional class of sites with $d_N/d_S > 1$ estimated from the data set (M2a and M8). An additional test comparing results from M8 to a modified version of the model where the selection class has d_N/d_S set to 1 (model M8a) was performed. This test rules out the possibility that the neutral model M7 is rejected due to a poor fit of the beta distribution for neutral and negatively selected sites. The test statistic follows a 50:50 mix of a χ^2 distribution with one degrees of freedom (df) and a point mass of zero. If the LRT is significant, positive selection is inferred. A Bayesian analysis was used to calculate the posterior probability that each site is from a particular site class, and sites with high posterior probabilities coming from the class with $d_N/d_S > 1$ ($P >$ 95%) were considered to be under positive selection. This Bayes empirical Bayes (BEB) approach performs best when the data set is small and lacks information (Yang 2000; Yang et al. 2005).

We also used other methods that account for variation in both synonymous and nonsynonymous rates; the single-likelihood ancestor counting (SLAC) method, the fixed effects likelihood (FEL) method, and the random effects likelihood (REL) method (see Pond and Frost 2005b). These methods were implemented using the web interface DATAMONKEY (Pond and Frost 2005a). Additionally, we compared synonymous and nonsynonymous substitution rates of ZP3 and PKDREJ in rodents and primates with those of cetaceans in order to assess whether reproductive proteins are evolving slower or faster in these species. Such comparisons were also performed using the nonreproductive proteins, MC1R and BMP4. Sequences were retrieved from available databases (NCBI and Ensembl, accesion numbers in Supplementary Table S2) and truncated to correspond to the region amplified in cetaceans. Pairwise comparisons of d_N and d_S were obtained using the runmode = −2 option in CODEML, and mean estimations were then calculated over all species. Overall levels of nucleotide divergence among cetaceans were estimated using MEGA 4.0 (Tamura et al. 2007).

Results

A 355-bp fragment of ZP3 was sequenced, including exon 6 (92 bp), intron 6 (136 bp), and exon 7 (127 bp). Translation resulted in a fragment of 72 amino acids in total, corresponding to positions 279–354 of *M. musculus* (NP_035906, 48% identity), which includes the sperm-combining region (328–343) (Chen et al. 1998).

Alignment with *Mus* ZP3 revealed a 3 amino acid deletion. The existence of some conserved regions suggests that some domain structures predicted in *Mus* are likely retained in cetaceans. However, one (Ser-332) of the 2 serine residues identified to be essential for sperm receptor activity in mouse, rat, and human ZP3 (Chen et al. 1998) has been lost in cetaceans, whereas the other (Ser-334) has been retained in *Balaenoptera acutorostrata*, *B. musculus* (both Balaenopteridae), and *Phocoena phocoena* (Phocoenidae).

We found very low levels of amino acid sequence divergence in cetaceans (only 6.9% of sites differed among species) with all dolphin species of the STDL complex having identical amino acid sequences (Figure 1).

For PKDREJ, a 603-bp fragment was sequenced. Translation resulted in a fragment of 200 amino acids in total, corresponding to positions 217–419 of *Homo sapiens* (NP_006062.1, 53% identity), which falls in the REJ domain, a region predicted to be functionally important in the sperm–egg recognition process (Sutton et al. 2006). As in ZP3, amino acid sequence divergence was very low, with all dolphin species of the STDL complex having identical amino acid sequences (Figure 1). Nucleotide divergence was also very low (Table 2). For BMP4, a 771-bp fragment was sequenced. Translation resulted in a fragment of 257 amino acids in total, corresponding to positions 125–381 of *H. sapiens* (BAA06410.1).

Phylogenetic trees obtained for ZP3 and PKDREJ with ML and Bayesian methods were concordant (Figure 2). In all trees, phylogenetic relationships among STDL species were unresolved. Despite this low level of resolution, overall topology is in agreement with published cetacean phylogenies (McGowen et al. 2009). Moreover, it has been suggested that the detection of positive selection is largely unaffected by possible uncertainties in underlying phylogenies (Pie 2006).

The average d_N/d_S across all lineages and codon sites for both ZP3 and PKDREJ was calculated. When using the codon evolutionary model M0, which estimates a single d_N/d_S value across the whole tree, values were <1 in all analyses, suggesting that these genes are evolving under selective constraints (Table 1). However, these proteins could contain amino acid sites subjected to positive selection that would be masked by a higher proportion of sites under purifying selection with ω close to zero. We thus compared 3 pairs of models (M1a vs. M2a, M7 vs. M8, and M8 vs. M8a) using results from CODEML (Table 1). In both ZP3 and PKDREJ, the neutral models (M1a and M7) were not significantly different from the selection models (M2a and M8). However, in ZP3, the LRT test comparing the null model M8a with the selection model M8 was statistically significant at 5% if we consider the df for the χ^2 statistic to be the 50:50 mixture of point mass 0 and χ^2_1 with a critical value of 2.71. The BEB approach identified sites 14, 48, 55, and 59 as likely targets of positive selection with posterior probability values ranging from 0.54 to 0.78. In PKDREJ, even though none of the LRT tests was statistically significant, the BEB approach identified sites 67, 73, 77, and 109 as likely targets of selection with posterior probability values ranging from 0.53 to 0.80. The methods SLAC and FEL failed to identify any sites in ZP3 and PKDREJ under positive selection. Nevertheless, the REL method identified sites 14, 48, 55, and 59 in ZP3 as positively selected with posterior probabilities >0.95 (the same sites identified by the BEB approach) but failed to identify such sites in PKDREJ.

The low number of amino acid substitutions in ZP3 and PKDREJ found among cetacean species contrasts to what

Downloaded from jhered.oxfordjournals.org at Flinders University of South Australia on July 7, 2011

has been described for rodents and primates. This observation is supported when d_N and d_S values are averaged across these groups, with cetaceans showing much lower synonymous and nonsynonymous substitution rates than rodents or primates (Table 2). However, when d_N/d_S is compared instead, values are actually higher in cetaceans for ZP3. In PKDREJ, rodents present the highest ratio, followed by cetaceans and then primates (Table 2). Within cetaceans, if mysticetes and odontocetes are considered separately, d_N values are slightly higher than d_S values for mysticetes in ZP3 sequences, resulting in $\omega > 1$. When rates of protein evolution given by the d_N/d_S values are compared for the nonreproductive proteins, BMP4 and MC1R, cetaceans show lower or similar values to the ones obtained for primates and rodents, which is in contrast to the pattern seen for the reproductive proteins, as described above (Table 2).

a)

	292	*308*	*327*	*334*	*339*
	1	3	**4**	**5**	**5**
	4	0	**8**	**5**	**9**
Consensus	**R**	N	**S**	**Y**	**R**
B. acutorostrata	.	S	**H**	**S**	**T**
B. musculus	.	.	**R**	**S**	**T**
B. physalus	.	.	.	.	.
K. breviceps	.	.	.	.	.
P. phocoena	**Q**	.	.	**S**	.
G. melas	.	.	.	.	.
S. fluviatilis	.	.	.	.	.
S. chinensis	.	.	.	.	.
S. coeruleoalba	.	.	.	.	.
S. attenuata	.	.	.	.	.
S. frontalis	.	.	.	.	.
S. longirostris	.	.	.	.	.
L. hosei	.	.	.	.	.
D. delphis	.	.	.	.	.
D. capensis	.	.	.	.	.
D. c. tropicalis	.	.	.	.	.
T. truncatus	.	.	.	.	.
T. aduncus	.	.	.	.	.

b)

	2	3	3	3	3	4	6	**6**	**7**	7	**7**	9	9	1 0	**1 0**	1	1	1	1	1	1	1	1
	9	0	3	5	7	4	4	**7**	**3**	4	**7**	6	9	5	**9**	2	5	9	7	9	6	2	6
Consensus	A	H	N	R	F	Y	F	**E**	**I**	R	**T**	Y	T	D	**I**	I	K	Q	E	G	H	E	R
B. acutorostrata	.	R	I	S	.	F	.	**G**	.	H	**A**	.	K	.	**M**	.	R	E	D	E	.	D	.
B. musculus	.	R	I	S	.	F	.	**G**	.	H	**A**	.	K	.	.	.	R	E	D	E	.	D	.
B. physalus	.	R	I	S	.	F	.	**G**	.	H	**A**	.	K	.	.	.	R	E	D	E	.	D	.
K. breviceps	.	R	I	.	L	.	.	.	**V**	H	**A**	.	K	G	**M**	S	R	.	D	E	.	.	G
P. phocoena	V	.	I	.	.	.	.	.	.	.	.	H	K	.	.	.	.	.	D	E	Q	.	.
G. melas	.	.	.	.	.	.	L	**G**	.	.	**A**	.	.	.	.	.	.	.	.	.	.	.	.
S. fluviatilis	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
S. chinensis	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
S. coeruleoalba	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
S. attenuata	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
S. frontalis	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
S. longirostris	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
L. hosei	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
D. delphis	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
D. capensis	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
D. c. tropicalis	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. truncatus	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
T. aduncus	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.

Figure 1. Amino acid (aa) sequence alignment of variable amino acid sites in (**a**) exons 6–7 (aa 1–72, corresponding to aa 279–354 of *Mus musculus*) of ZP3 and (**b**) PKDREJ (aa 1–200). Sites identified by ML methods as being likely subjected to positive selection are in bold.

Downloaded from jhered.oxfordjournals.org at Flinders University of South Australia on July 7, 2011

Table 1 Tests of adaptive evolution for (a) ZP3 and (b) PKDREJ using site models

Model code	*p*	d_N/d_S	*l*	Parameters	Positively selected sites: BEB	Positively selected sites: LRT
(a) ZP3						
M0	1	0.497		$\omega = 0.497$		
M1a	2	0.288	−340.476	$p_0 = 0.712$ ($p_1 = 0.288$) $w0 = 0.000$ ($\omega 1 = 1.000$)		M1a-M2a 2.768
M2a	4	0.542	−339.092	$p_0 = 0.847$ $p_1 = 0.000$ $p_2 = 0.152$ $\omega 0 = 0.000$ ($\omega 1 = 1.000$) $\omega 2 = 3.554$	48H (0.602); 55S (0.665); 59T (0.594)	
M7	2	0.200	−340.479	β (0.012, 0.005)		M7-M8 2.774
M8	4	0.621	−339.092	$p_0 = 0.847$ ($p_1 = 0.152$) β (0.005, 76.667) $\omega = 3.555$	14R (0.545); 48H (0.729); 55S (0.783); 59T (0.722)	
M8a	5	0.288	−340.475	$p_0 = 0.711$ ($p_1 = 0.288$) $\omega = 1.000$ β (0.007, 1.486)		M8-M8a 2.766*
(b) PKDREJ						
M0	1	0.604		$\omega = 0.604$		
M1a	2	0.578	−1123.576	$p_0 = 0.422$ ($p_1 = 0.578$) $\omega 0 = 1.000$ ($\omega 1 = 1.000$)		M1a-M2a 0.576
M2a	4	0.623	1123.288	$p_0 = 0.897$ $p_1 = 0.000$ $p_2 = 0.103$ $\omega 0 = 0.384$ ($w1 = 1.000$) $\omega 2 = 2.710$	67G (0.706); 77A (0.539)	
M7	2	0.600	−1123.588	β (0.007, 0.005)		M7-M8 0.6
M8	4	0.623	−1123.288	$p_0 = 0.898$ ($p_1 = 0.102$) β (62.042, 99.000) $\omega = 2.714$	67G (0.804); 73I (0.530); 77A (0.648); 109I (0.587)	
M8a	5	0.577	−1123.576	$p_0 = 0.422$ ($p_1 = 0.578$) β (0.007, 1.988) $\omega = 1.000$		M8-M8a 0.576

p, number of parameters relating to variation in d_N/d_S for each model; d_N/d_S, ratio averaged across all sites and lineages; $\omega = d_N/d_S$; *p* = proportion of codon sites in each class of ω; β = parameters for beta distribution; *l*, log likelihood of the model.

*$P < 0.05$.

Discussion

In cetaceans, the occurrence of female promiscuity, leading to the existence of sperm competition and sexual conflict, would suggest that the evolution of reproductive proteins would be rapid, driven by positive selection, because an increased mating rate escalates sexual conflict (Birkhead and Pizzari 2002; Swanson and Vacquier 2002). However, in this study, we found very low levels of amino acid divergence in ZP3 and PKDREJ between species, particularly among Delphinines. This lack of polymorphism resulted in the failure to reject the null model in favor of the positive selection model for PKDREJ, despite the fact that one of the codons (67) identified as likely target of positive selection, is very close to the codons indentified as under positive selection in the human REJ domain (codon 285 in the human PKDREJ, Hamm et al. 2007). For ZP3, only one

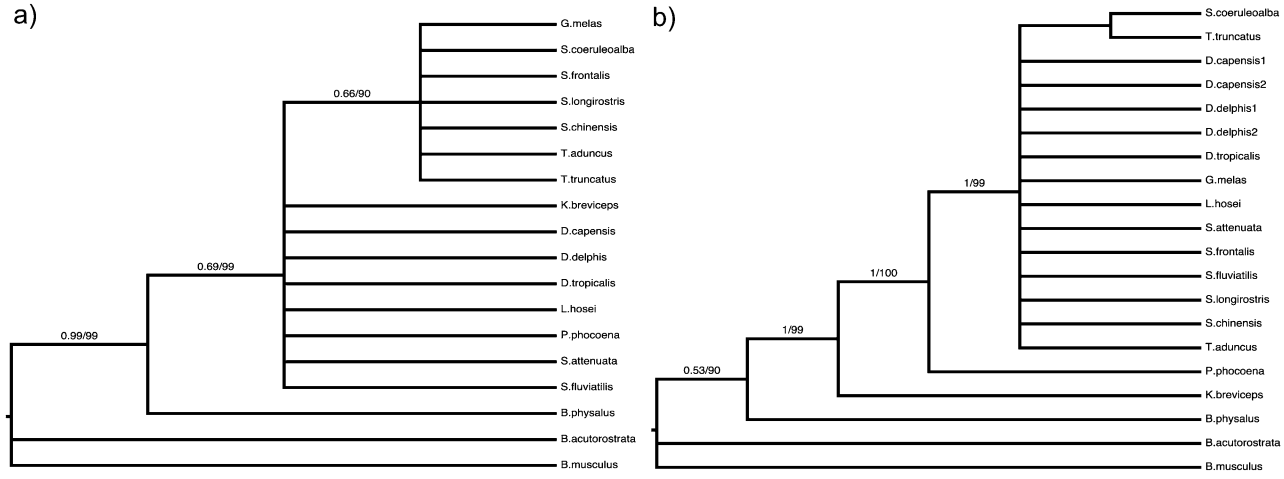

Figure 2. Phylogenetic trees of cetaceans based on nucleotide sequences of (**a**) ZP3 and (**b**) PKDREJ genes. Individuals are identified by species name. Values above branches represent Bayesian posterior probabilities and Maximum Parsimony bootstrap support values, respectively.

Downloaded from jhered.oxfordjournals.org at Flinders University of South Australia on July 7, 2011

Table 2 Nonsynonymous (d_N) and synonymous (d_S) substitution rates estimated based on pairwise comparisons of (a) ZP3, (b) PKDREJ, (c) MC1R, and (d) BMP4 sequences from cetaceans, rodents, and primates

	d_N/d_S	d_N	d_S	d
(a) ZP3				
Primates	0.2860	0.1407	0.1025	
Rodents	0.3070	0.2029	0.6605	
Cetaceans	0.4250	0.0068	0.0161	0.0100
Mysticetes	1.1120	0.0175	0.0157	0.0160
Odontocetes	0.4096	0.0066	0.0160	0.0050
(b) PKDREJ				
Primates	0.2640	0.0205	0.0777	
Rodents	0.8750	0.1089	0.1245	
Cetaceans	0.6230	0.0162	0.0259	0.0050
Mysticetes	0.2950	0.0030	0.0102	0.0060
Odontocetes	0.6260	0.0164	0.0263	0.0090
(c) BMP4				
Primates	0.0660	0.0017	0.0260	
Rodents	0.0690	0.0297	0.4330	
Cetaceans	0.0490	0.0011	0.0220	
(d) MC1R				
Primates	0.0902	0.0375	0.4153	
Rodents	0.1191	0.0626	0.5252	
Cetaceans	0.1088	0.0100	0.0915	

Nucleotide divergence (d) estimated for cetacean species.

test (M8-M8a) of 5 was statistically significant, identifying a few codons as likely targets of selection. One of such codons (55) corresponds to a serine residue (Ser-334) identified to be essential for sperm receptor activity in mouse, rat, and human ZP3 (Chen et al. 1998). This residue, however, has only been retained in 3 species: *B. acutorostrata*, *B. musculus*, and *P. phocoena*. Nevertheless, we would expect a strong signal of selection to be present in these proteins for the reasons described above. In Delphinines, a group of species recently diverged, it is possible that fertilization specificities are evolving slowly as a result of greater functional constraint on these reproductive proteins or reduced selective pressures for species recognition. In fact, hybridization among dolphin species seems to occur, both in captivity (Zornetzer and Duffield 2003) and in the wild (Bérubé 2002). It would, however, be expected in closely related species, where lineages have not yet sorted out completely, that differences in the genome would be found in the so-called speciation genes, those affecting target phenotypic traits or those involved in species recognition, such as ZP3 and PKDREJ (Wu 2001).

Other factors such as long generation times and intrinsic demographic features could also be dictating the slower evolution of proteins across the genome in cetaceans and therefore affecting reproductive proteins as well. This is supported by our results, where both reproductive and nonreproductive proteins show overall low number of synonymous and nonsynonymous substitutions in cetaceans. It has been generally accepted that long-lived larger mammals experience a slower mutation rate than small-bodied mammals due to several genomic features relating substitution rates with generation times and life-history traits (e.g., Martin and Palumbi 1993; Bromham 2009; Jackson et al. 2009). Our estimates of the rates of amino acid evolution appear to support this theory because rodents present a consistently higher rate of amino acid evolution when compared with primates and cetaceans for all proteins, with the exception of ZP3. Although cetaceans present a marginally higher d_N/d_S ratio for this protein, this likely results from the overall low number of both nonsynonymous and synonymous substitutions. When the mysticetes are analyzed separately, the ratio is even higher due to higher number of nonsynonymous changes present in this lineage when compared with the odontocetes (see Figure 1). This is likely explained by noise at low levels of divergence and not necessarily by the signal left by positive selection.

Studies in primates have shown a positive correlation between the intensity of sperm competition and degree of polyandry and the strength of positive selection in genes encoding for structural components of semen coagulum (semenogelin I, Kingan et al. 2003 and semenogelin II, Dorus et al. 2004), therefore supporting this theory. However, such association was not found between adaptive evolution of ZP3 and PKDREJ and the potential for sperm competition in rodents and primates (Turner and Hoekstra 2006; Hamm et al. 2007; Swann et al. 2007). The same appears to apply for cetaceans. Although varying degrees of sperm competition have been suggested across taxa (e.g., Connor et al. 2000), we did not find a strong overall pattern of positive selection in ZP3 and PKDREJ. However, it should be noted that within mysticetes, Balaenids appear to show higher levels of sperm competition than Balaenopterids. Female promiscuity has in fact been well documented in some species (e.g., minke whales, Skaug et al. 2008; humpback whales, Clapham and Palsboll 1997; gray whales, Swartz et al. 2006; and right whales, Frasier et al. 2007). Unfortunately, it was not possible to include all these species in this study, and its inclusion could have changed the results presented here, namely in the signal for positive selection in ZP3. Within odontocetes, our sampling of Delphinid species, for which more information on levels of sperm competition is available, would have enabled us to make such comparisons, if it was not for the complete lack of amino acid substitutions we observed.

It should be mentioned that the reduced statistical power to detect positive selection due to low polymorphism among our study species may have influenced our results, particularly because the use of χ^2 distribution makes the LRTs very conservative for short closely related sequences (Anisimova et al. 2001). Additionally, we cannot rule out that for PKDREJ, other regions than the one analyzed here could be targets of positive selection in cetaceans. Nonetheless, the results obtained in this study were surprising and should initiate a discussion on the evolutionary forces driving the evolution of reproductive proteins in cetaceans and processes that may be dictating the establishment of reproductive isolation and species recognition. Future studies should focus on the study of additional species and reproductive proteins.

Downloaded from jhered.oxfordjournals.org at Flinders University of South Australia on July 7, 2011

Supplementary Material

Supplementary material can be found at http://www.jhered.oxfordjournals.org/.

Funding

A.R.A. is supported by Macquarie University and a PhD grant (SFR/BD/27245/2006) funded by Fundação para Ciência e Tecnologia (Portugal).

Acknowledgments

The authors wish to thank S. L. Kosakovsky Pond for guidance with data analyses in Datamonkey; M. C. Silva and J. R. Powell for helpful discussions; and 2 anonymous reviewers for helpful comments on previous versions of the manuscript. Kelly Robertson (SWFSC) and Marina Sequeira (ICN) provided tissue samples.

References

Anisimova M, Bielawski JP, Yang ZH. 2001. Accuracy and power of the likelihood ratio test in detecting adaptive molecular evolution. Mol Biol Evol. 18:1585–1592.

Arnason U, Gullberg A, Janke A. 2004. Mitogenomic analyses provide new insights into cetacean origin and evolution. Gene. 333:27–34.

Berta A, Sumich JL. 1999. Marine mammals: evolutionary biology. San Diego (CA): Academic Press.

Bérubé M. 2002. Hybridism. In: Perrin WF, Wursig B, Theweissen JGM, editors. Encyclopedia of marine mammals. San Diego (CA): Academic Press Inc. p. 596–600.

Birkhead TR, Pizzari T. 2002. Postcopulatory sexual selection. Nat Rev Genet. 3:262–273.

Bromham L. 2009. Why do species vary in their rate of molecular evolution? Biol Lett. 5:401–404.

Butscheid Y, Chubanov V, Steger K, Meyer D, Dietrich A, Gudermann T. 2006. Polycystic kidney disease and receptor for egg jelly is a plasma membrane protein of mouse sperm head. Mol Reprod Dev. 73:350–360.

Caballero S, Jackson J, Mignucci-Giannoni AA, Barrios-Garrido H, Beltran-Pedreros S, Montiel-Villalobos MG, Robertson KM, Baker CS. 2008. Molecular systematics of South American dolphins Sotalia: sister taxa determination and phylogenetic relationships, with insights into a multi-locus phylogeny of the Delphinidae. Mol Phylogenet Evol. 46:252–268.

Calkins JD, El-Hinn D, Swanson WJ. 2007. Adaptive evolution in an avian reproductive protein: ZP3. J Mol Evol. 65:555–563.

Clapham PJ, Palsboll PJ. 1997. Molecular analysis of paternity shows promiscuos mating in female humpback whales (*Megaptera novaengliae*, Borowski). Proc R Soc Lond Ser B Biol Sci. 264:95–98.

Chen J, Litscher ES, Wassarman PM. 1998. Inactivation of the mouse sperm receptor, mZP3, by site-directed mutagenesis of individual serine residues located at the combining site for sperm. Proc Natl Acad Sci U S A. 95:6193–6197.

Connor RC, Read Aj, Wrangham R. 2000. Male Reproductive Strategies and Social Bonds. In: Mann J, Connor RC, Tyack PL, Whitehead H, editors. Cetacean Societies: field studies of dolphins and whales. The chicago university press. p. 247–270.

Demere TA, McGowen MR, Berta A, Gatesy J. 2008. Morphological and molecular evidence for a stepwise evolutionary transition from teeth to baleen in mysticete whales. Syst Biol. 57:15–37.

Dorus S, Evans PD, Wyckoff GJ, Choi SS, Lahn BT. 2004. Rate of molecular evolution of the seminal protein gene SEMG2 correlates with levels of female promiscuity. Nat Genet. 36:1326–1329.

Frasier TR, Hamilton PK, Brown MW, Conger LA, Knowlton AR, Maryx MK, Slay CK, Kraus SD, White BN. 2007. Patterns of male reproductive success in a highly promiscuous whale species: the endangered North Atlantic right whale. Mol Ecol. 16:5277–5293.

Hamm D, Mautz BS, Wolfner MF, Aquadro CF, Swanson WJ. 2007. Evidence of amino acid diversity-enhancing selection within humans and among primates at the candidate sperm-receptor gene PKDREJ. Am J Hum Genet. 81:44–52.

Huelsenbeck JP, Ronquist F. 2001. MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics. 17:754–755.

Jackson JA, Baker CS, Vant M, Steel DJ, Medrano-Gonzalez L, Palumbi SR. 2009. Big and slow: phylogenetic estimates of molecular evolution in baleen whales (Suborder Mysticeti). Mol Biol Evol. 26:2427–2440.

Jansa SA, Lundrigan BL, Tucker PK. 2003. Tests for positive selection on immune and reproductive genes in closely related species of the murine genus Mus. J Mol Evol. 56:294–307.

Kingan SB, Tatar M, Rand DM. 2003. Reduced polymorphism in the chimpanzee semen coagulating protein, semenogelin I. J Mol Evol. 57:159–169.

Martin AP, Palumbi SR. 1993. Body size, metabolic-rate, generation time, and the molecular clock. Proc Natl Acad Sci U S A. 90:4087–4091.

McGowen MR, Spaulding M, Gatesy J. 2009. Divergence date estimation and a comprehensive molecular tree of extant cetaceans. Mol Phylogenet Evol. 53:891–906.

Metz EC, Robles-Sikisaka R, Vacquier VD. 1998. Nonsynonymous substitution in abalone sperm fertilization genes exceeds substitution in introns and mitochondrial DNA. Proc Natl Acad Sci U S A. 95:10676–10681.

Perrin WF, Reeves RR. 2004. Appendix 5: report of the working group on species- and subspecies-level taxonomy. In: Reeves RR, Perrin WF, Taylor BL, Baker CS, Mesnick SL, editors. Report of the Workshop on Shortcomings of Cetacean Taxonomy in Relation to Needs of Conservation and Management. La Jolla (CA): NOAA Technical Memorandum. p. 25–60.

Pie MR. 2006. The influence of phylogenetic uncertainty on the detection of positive darwinian selection. Mol Biol Evol. 23:2274–2278.

Pond SLK, Frost SDW. 2005a. Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. Bioinformatics. 21:2531–2533.

Pond SLK, Frost SDW. 2005b. Not so different after all: a comparison of methods for detecting amino acid sites under selection,. Mol Biol Evol. 22:1208–1222.

Posada D, Buckley T. 2004. Model selection and model averaging in phylogenetics: advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests. Syst Biol. 53:793–808.

Skaug HJ, Bérubé M, Rew MB, Palsboll PJ. 2008. Genetic analyses reveal promiscuous mating in female common minke whales, *Balaenoptera acutorostrata*. J Cetacean Res Manage. 9:249–251.

Sutton KA, Jungnickel MK, Florman HM. 2008. A polycystin-1 controls postcopulatory reproductive selection in mice. Proc Natl Acad Sci U S A. 105:8661–8666.

Sutton KA, Jungnickel MK, Ward CJ, Harris PC, Florman HM. 2006. Functional characterization of PKDREJ, a male germ cell-restricted polycystin. J Cell Physiol. 209:493–500.

Swann CA, Cooper SJB, Breed WG. 2007. Molecular evolution of the carboxy terminal region of the zona pellucida 3 glycoprotein in murine rodents. Reproduction. 133:697–708.

Downloaded from jhered.oxfordjournals.org at Flinders University of South Australia on July 7, 2011

Swanson WJ, Nielsen R, Yang QF. 2003. Pervasive adaptive evolution in mammalian fertilization proteins. Mol Biol Evol. 20:18–20.

Swanson WJ, Vacquier VD. 2002a. The rapid evolution of reproductive proteins. Nat Rev Genet. 3:137–144.

Swanson WJ, Vacquier VD. 2002b. Reproductive protein evolution. Ann Rev Ecol Syst. 33:161–179.

Swanson WJ, Zhang ZH, Wolfner MF, Aquadro CF. 2001. Positive Darwinian selection drives the evolution of several female reproductive proteins in mammals. Proc Natl Acad Sci U S A. 98:2509–2514.

Swartz SL, Taylor BL, Rugh DJ. 2006. Gray whale *Eschrichtius robustus* population and stock identity. Mamm Rev. 36:66–84.

Swofford DL. 2003. PAUP*. Phylogenetic analysis using parsimony (*and other methods). Version 4.0. Sunderland (MA): Sinauer Associates.

Tamura K, Dudley J, Nei M, Kumar S. 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol. 24:1596–1599.

Turner LM, Hoekstra HE. 2006. Adaptive evolution of fertilization proteins within a genus: variation in ZP2 and ZP3 in deer mice (Peromyscus). Mol Biol Evol. 23:1656–1669.

Turner LM, Hoekstra HE. 2008. Causes and consequences of the evolution of reproductive proteins. Int J Dev Biol. 52:769–780.

Wassarman PM, Jovine L, Litscher ES. 2001. A profile of fertilization in mammals. Nat Cell Biol. 3:E59–E64.

Wassarman PM, Jovine L, Litscher ES. 2004. Mouse zona pellucida genes and glycoproteins. Cytogenet Genome Res. 105:228–234.

Wassarman PM, Litscher ES. 2001. Towards the molecular basis of sperm and egg interaction during mammalian fertilization. Cells Tissues Organs. 168:36–45.

Wu CI. 2001. The genic view of the process of speciation. J Evol Biol. 14:851–865.

Yang Z. 2000. Phylogenetic analysis by maximum likelihood (PAML). London: University College.

Yang ZH, Wong WSW, Nielsen R. 2005. Bayes empirical Bayes inference of amino acid sites under positive selection. Mol Biol Evol. 22: 1107–1118.

Zornetzer HR, Duffield DA. 2003. Captive-born bottlenose dolphin x common dolphin (*Tursiops truncatus* × *Delphinus capensis)* intergeneric hybrids. Can J Zool. 81:1755–1762.

Received November 24, 2010; Revised December 14, 2010; Accepted December 14, 2010

Corresponding Editor: C Scott Baker

Downloaded from jhered.oxfordjournals.org at Flinders University of South Australia on July 7, 2011