

Adaptive Evolution of Gamete-Recognition Proteins in Birds

Sofia Berlin · Lujiang Qu · Hans Ellegren

Received: 5 June 2008 / Accepted: 3 September 2008 / Published online: 11 October 2008
© Springer Science+Business Media, LLC 2008

Abstract Gamete-recognition proteins have been shown to evolve by positive selection in diverse organism groups, such as marine invertebrates and mammals, although underlying evolutionary mechanisms driving this rapid divergence are poorly understood. However, several hypotheses have been put forward to explain the observed pattern, including different forms of sexual conflict and sperm competition. Because female gametes require more energy to produce than male gametes, female organisms suffer more when fertilisation goes wrong. One process that results in a failed mammalian fertilisation is polyspermy, when >1 sperm fertilises the egg. However in birds, there is no such sexual conflict because multiple sperm typically bind and fuse with the egg. If sexual conflict driven by polyspermy avoidance is important for the evolution of gamete-recognition proteins in vertebrates, we expect to find positive selection in the genes to be less pronounced in birds. We therefore sequenced six genes (*ZP1*, *ZP2*, *ZP4*, *ZPAX*, *CD9*, and *Acrosin*) encoding gamete-recognition proteins in several bird species to test for positive selection. For comparison, we also analysed

orthologous sequences in a set of mammalian species. We found no major differences in the occurrence of adaptive evolution and the strength of selection between bird and mammal orthologs. From this we conclude that polyspermy avoidance does not act as the main underlying evolutionary force shaping the rate of evolution in these genes. We discuss other possible processes that could explain positive selection of gamete-recognition proteins in birds and mammals, such as hybridisation avoidance, cryptic female choice, and postcopulatory sperm competition.

Keywords Adaptive evolution · Birds · Reproductive proteins · Sexual conflict · Sperm competition

Introduction

Numerous sequence comparisons between closely related taxa have shown that genes encoding gamete-recognition proteins (proteins that mediate sperm-egg interactions during fertilisation) are unusually diverse (Clark et al. 2006; Swanson and Vacquier 2002a; Swanson and Vacquier 2002b). It has been demonstrated that the rapid diversification of such proteins are often driven by positive selection in free-spawning marine species, such as sea urchins and abalones (Hellberg and Vacquier 1999; McCartney and Lessios 2004; Riginos and McDonald 2003; Springer and Crespi 2007; Vacquier et al. 1997; Yang et al. 2000), as well as in several vertebrate species (Calkins et al. 2007; Civetta 2003; Gasper and Swanson 2006; Swanson et al. 2003; Swanson et al. 2001; Turner and Hoekstra 2006). Although positive selection in gamete-recognition proteins thus appears to be a general phenomenon in diverse taxonomic groups, the underlying

Electronic supplementary material The online version of this article (doi:10.1007/s00239-008-9165-6) contains supplementary material, which is available to authorized users.

S. Berlin · L. Qu · H. Ellegren (✉)
Department of Evolutionary Biology, Evolutionary Biology
Centre, Uppsala University, Norbyvägen 18 D, 752 36 Uppsala,
Sweden
e-mail: Hans.Ellegren@ebc.uu.se

Present Address:

L. Qu
State Key Laboratory for AgroBiotechnology, College of
Animal Science and Technology, China Agricultural University,
Beijing 100094, China

evolutionary forces are still poorly understood. In marine invertebrates, which are characterised by external fertilization and no mating behaviour, it has been hypothesised that divergence between gamete-recognition proteins can establish prezygotic barriers to reproduction and hence play an important role in speciation (Metz and Palumbi 1996; Palumbi 1992; Vacquier et al. 1997). For example, it has been proposed that selection against hybridisation (reinforcement) drives the evolution of the bindin protein in *Echinometra* sea urchins (Geyer and Palumbi 2003). In these organisms, bindin shows extreme sequence divergence in sympatric populations, whereas this is not the case in allopatric populations.

In vertebrates, the presence of mating behaviour and mate choice add a behavioural pre-mating barrier to hybridisation, theoretically reducing the pressure on gametic recognition and binding proteins. A variety of alternative hypotheses have therefore been put forward to explain why gamete-recognition proteins evolve by positive selection in vertebrates, and these primarily involve sperm competition and sexual conflict as evolutionary driving forces. Sexual conflict can arise when conditions that are optimal in one sex simultaneously act to reduce fitness in the other, i.e., when the reproductive interests of the two sexes are not coincident. Sexual conflict over adaptive optima is thought to lead to a co-evolutionary chase between male and female characters (Gavrilets 2000; Rice and Holland 1997). This may apply to gamete-recognition proteins because there is a sexual conflict when sperm competition leads to fast rates of fertilisation (fertilisation being mediated by gamete-recognition proteins). Female organisms may benefit from a more moderate rate to prevent polyspermic fertilisation, i.e., when several sperm bind and fuse with the egg (Frank 2000). Polyspermy generally results in embryo mortality in most organisms (Gardner and Evans 2006), and elaborate mechanisms have evolved to avoid it. It is often assumed that the larger energy investment put into female gametes makes polyspermy more detrimental to female than male fitness, although it should be kept in mind that the cost of a failed fertilisation could either be a matter of egg contra sperm or egg contra entire ejaculate. If polyspermy avoidance is important for the rapid evolution of gamete-recognition proteins in vertebrates with internal fertilisation, then we can predict that the signatures of positive selection would be absent or weaker in species where polyspermy is norm.

One way to test this prediction is to study the molecular evolution of gamete-recognition proteins in birds because the principal feature of fertilisation in birds is physiologic polyspermy (Stepinska and Bakst 2007; Tarin and Cano 2000). Consequently, sexual conflict over fertilisation may be less pronounced in the avian system. In fact, it is unclear

whether a single sperm can activate an oocyte in birds, and the observation of a positive correlation between the number of sperm entering ova and ovum size could suggest that large ova may require more spermatozoa to ensure fertilization (Birkhead et al. 1994; Bramwell and Howarth 1992), making polyspermy adaptive. The vertebrate egg envelope is composed of a set of related proteins encoded by zona pellucida (ZP) genes (Hughes 2007; Lefievre et al. 2004; Wassarman 1988). These genes can be divided into five classes, ZP1, ZP2, ZP3, ZP4, and ZPAX. ZP genes have been identified in mammals, birds, amphibians, and fish (Litscher and Wassarman 2007). Although avian ZP proteins have not been functionally characterized in detail (compared with the situation in, for example, invertebrates and mice), the presence of ZP genes in the genomes of phylogenetically divergent eukaryotic lineages, coupled with the conserved nature of the fertilization process, suggest that their functional role in birds is similar to that in other organisms.

The molecular evolution of avian ZP3 has previously been studied by Berlin and Smith (2005) and Calkins et al. (2007), with some evidence for adaptive evolution provided by the latter study. In this study we sequenced and analysed avian ZP1, ZP2, ZP4, and ZPAX, along with two other genes known to be involved in gamete-recognition, CD9 (Miyado et al. 2000; Runge et al. 2007) and Acrosin (Baba et al. 1994). We investigated the role of positive selection in driving the evolution of these avian proteins and then compared this with the situation for mammalian orthologs.

Materials and Methods

Samples and Sequences

Tissues were collected from one female and one male mallard (*Anas platyrhynchos*), guinea fowl (*Numida meleagris*), pheasant (*Phasianus colchicus*), pigeon (*Columba livia*), quail (*Coturnix coturnix*), red grouse (*Lagopus lagopus scotica*), and turkey (*Meleagris gallopavo*) and were stored in RNAlater (Qiagen). Some avian and mammalian sequences were taken from GenBank as specified in Table 1 and Supplementary Table 1.

Laboratory Work

Total RNA was extracted from spleen, testes, and ovaries using TRIzol (Invitrogen). First-strand cDNA was synthesized from the total RNA using Oligo(dT)₂₀ primers (Invitrogen), and this cDNA was subsequently used as template for polymerase chain reaction (PCR). The PCR conditions were 95°C for 5 minutes, 35 cycles of 94°C

Table 1 Summary of species and GenBank accession numbers for mammalian gamete-recognition proteins analysed in this study

Species	<i>CD9</i>	<i>Acrosin</i>	<i>ZP1</i>	<i>ZP2</i>	<i>ZP4</i>
Human	NM_001769	NM_001097	NM_207341	BC096306	BC069521
Chimpanzee	XM_508947		XM_522022		XM_525105
Rhesus macaque	XM_001102751	XM_001113191	XM_001084628	XM_001091147	XM_001096846
Pig	NM_214006	NM_214033		NM_213848	NM_214045
Cow	BC147992	NM_173886		NM_173973	NM_173975
White-tufted-ear marmoset	EF215132			Y10767	
Domestic cat	D30786			U05776	NM_001009260
Norway rat	NM_053018	NM_012490	XM_001074922	NM_031150	NM_172330
House mouse	NM_007657	BC103577	U20448	NM_011775	
Harbour seal	EU085445				
Horse		XM_001490525	XM_001493722	XM_001494769	
Golden hamster			EU003563	AY876920	DQ838550
Dog		XM_843155		D45069	
Ermine				AY779765	AY779766
Rabbit		NM_001082336		L12167	NM_001082295
Guinea pig		Z12153			
Sheep		AJ278742			

for 1 minute, 55° to 58° C for 1 minute, and 72°C for 1 minute, and a final 10-minutes extension at 72°C. Primer sequences and combinations are listed in Supplementary Table 2. The PCR primers were also used as sequencing primers. All PCR products were cleaned before sequencing by adding 1 µl ExoSAP-IT (Amersham Biosciences) to every 3 µl PCR product. The reactions were incubated for 15 minutes at 37°C and for 15 minutes at 80°C. The samples were sequenced by Macrogen (Seoul, South Korea) on ABI 3730 instruments (Applied Biosystems). DNA chromatograms were edited and checked using Sequencher 4.2.2 (Gene Codes Corp., Ann Arbor).

Sequence Analyses

The sequences were aligned using CLUSTALW in the Alignment Explorer tool in MEGA 3.1 (Kumar et al.

Table 2 Descriptive data for the six gamete-recognition genes sequenced in birds

Gene	Putative function	N	L	S (M0)	ω (M0)
<i>CD9</i>	Sperm–egg fusion	9	669	1.62	0.29
<i>Acrosin</i>	Sperm–egg interaction	6	924	3.12	0.26
<i>ZP1</i>	Sperm–egg interaction	7	2814	1.78	0.28
<i>ZP2</i>	Sperm–egg interaction	7	2058	2.20	0.31
<i>ZP4</i>	Sperm–egg interaction	8	1629	1.47	0.36
<i>ZPAX</i>	Sperm–egg interaction	8	2469	2.02	0.39

N—number of taxa; L—length of gapped alignment (bp); S—number of substitutions per codon (M0 model); ω — d_N/d_S (M0 model)

2004). We used the codeml program in the PAML package version 4 (Yang 1997; Yang 2007) to perform likelihood ratio tests of positive selection for each gene. For these analyses we considered models of codon evolution which allow for variation in ω , which is the ratio of nonsynonymous to synonymous substitutions per nonsynonymous and synonymous sites (d_N/d_S or K_A/K_S), among codons but assume the same distribution in all lineages. We performed three likelihood ratio tests (LRT), which are thought to provide reliable tests of positive selection (Swanson et al. 2003; Wong et al. 2004), according to the following:

1. M1a-M2a LRT: The M1a model (one ω class between 0 and 1, and one class of $\omega = 1$) is compared with the M2a model (same as M1a model plus an extra class of >1).
2. M7-M8 LRT: The M7 model (a discretised beta distribution for ω between 0 and 1 with 10 equal class proportions) is compared with the M8 model (same as the M7 model plus an extra class of $\omega \geq 1$).
3. M8a-M8 LRT: The M8a model (same as M7 plus an extra class of $\omega = 1$) is compared with the M8 model).

For all LRTs, equilibrium codon frequencies were obtained using the average base composition at the three codon positions (CodonFreq = 2), and the transition–transversion rate ratio was estimated from the data. The sequences were analysed with gaps included. This type of analysis requires an unrooted phylogeny, and we used the following topologies of species trees depending on from the number of taxa from which sequence data were obtained. For the nine bird species analysed for *CD9*, the

following topology was used: ((((((red grouse, turkey), pheasant), quail), chicken), guineafowl), duck), pigeon, zebrafinch) (Kaiser et al. 2007). For *ZP4* and *ZPAX*, for which eight species were analysed, the following topology was used: ((((((turkey, pheasant), quail), chicken), guineafowl), duck), pigeon, zebrafinch). The following seven species were analysed for *ZP1* and *ZP2*: ((((((turkey, pheasant), quail), chicken), guineafowl), duck), zebrafinch). Finally, six species were analysed for *Acrosin* using the following tree: (((((pheasant, quail), chicken), guineafowl), duck), zebrafinch). For analyses of the mammalian orthologues (Table 1), the species trees that were used were based on Murphy et al. (2001) and are presented in the Supplementary Material.

Results

Birds

We sequenced six gamete-recognition genes (*CD9*, *Acrosin*, *ZP1*, *ZP2*, *ZP4*, and *ZPAX*) across a suite of bird species (Table 2 and Supplementary Tables 1 and 3) and analysed the pattern of molecular evolution in these genes. Three pairs of selection-neutral models were compared for each gene using results from codeml: M1a and M2a, M7 and M8, and M8a and M8 (Table 3). For two of the genes (*ZP4* and *ZPAX*), the selection models (M2a and M8) were not significantly different from the neutral models (M1a, M7, and M8a), and we therefore find no evidence for a strong role of positive selection affecting the evolution of these genes in birds. For the other four genes (*CD9*, *Acrosin*, *ZP1*, and *ZP2*), the selection model M8 fit the data significantly better than the neutral model M7 (*CD9*: $2\Delta\ln L = 13.5$, $df = 2$, $p = 0.001$; *Acrosin*: $-2\Delta\ln L = 8.2$, $p = 0.02$; *ZP1*: $-2\Delta\ln L = 10.3$, $p = 0.006$; and *ZP2*: $-2\Delta\ln L = 9.7$, $p = 0.008$). We obtained similar results for the M8 versus M8a comparisons, where the M8 model was a significantly better fit to the data than the M8a model for all four genes (*CD9*: $-2\Delta\ln L = 9.8$, $df = 1$, $p = 0.002$; *Acrosin*: $-2\Delta\ln L = 3.9$, $p = 0.05$; *ZP1*: $-2\Delta\ln L = 5.3$, $p = 0.02$; and *ZP2*: $-2\Delta\ln L = 5.0$, $p = 0.02$). For the more conservative M1a versus M2a comparison, the selection model was significantly different than the neutral model for the *CD9* gene ($-2\Delta\ln L = 8.3$, $df = 2$, $p = 0.02$) but not for the other three genes. These results suggest that *CD9*, *Acrosin*, *ZP1*, and *ZP2* have evolved under the influence of positive selection in birds. Detailed results from all genes and all models, including putative selected sites, as determined through the Bayes empiric Bayes analyses for models M2 and M8, are listed in Table 3. The number of selected sites varied between 1 and 7 at the $p = 0.10$ significance level for the four different genes.

Mammals

The same models were compared for the mammalian orthologs of the investigated bird genes using available sequence data (Tables 4 and 5 [*ZPAX* absent in mammals]). The selection model M8 fit the data significantly better than the neutral M7 model for all genes except for *ZP1* (*CD9*: $-2\Delta\ln L = 10.7$, $df = 2$, $p = 0.005$; *Acrosin*: $-2\Delta\ln L = 45.5$, $p < 0.001$; *ZP2*: $-2\Delta\ln L = 26.1$, $p < 0.001$; and *ZP4*: $-2\Delta\ln L = 9.8$, $p = 0.008$). Similarly, the selection model M8 fit the data significantly better than the neutral model M8a for the same four genes (*CD9*: $-2\Delta\ln L = 4.2$, $df = 1$, $p = 0.04$; *Acrosin*: $-2\Delta\ln L = 28.2$, $p < 0.001$; *ZP2*: $-2\Delta\ln L = 15.4$, $p < 0.001$; and *ZP4*: $-2\Delta\ln L = 8.5$, $p = 0.004$). However, in the more conservative test, the selection model M2a fit the data significantly better than the neutral model M1a for *Acrosin* ($-2\Delta\ln L = 32.0$, $df = 2$, $p < 0.001$) and *ZP2* ($-2\Delta\ln L = 12.5$, $p = 0.002$) but not for the other genes. To summarize, four of five mammalian gamete-recognition genes showed evidence for adaptive evolution, confirming previous observations (Swanson et al. 2001).

The incidence of putatively selected sites in mammalian orthologs (Table 5) was approximately as high as in the corresponding bird genes. For the three genes positively selected in both birds and mammals, the frequency of selected sites were 9.5% (116 of 1217 sites) and 9.0% (118 of 1314), respectively ($\chi^2 = 0.21$ [not significant]). For all genes analysed, the frequencies were in both cases 1.8%. We finally sought to determine if the same sites of orthologs had been subject to positive selection in both mammals and birds. Figure 1 shows an amino-acid alignment of human and chicken *CD9*. There is a tendency for clustering selected sites in a homologous region of the two species toward the 3' end of the protein, but hardly any individual sites identified as positively selected in both species. For *acrosin* and *ZP2*, the overlap in adaptively evolving regions is less clear, with the exception of three adjacent codons of *acrosin* evolving under positive selection (Supplementary Figs. 1 and 2).

Discussion

We found that four (*CD9*, *Acrosin*, *ZP1*, and *ZP2*) of six (also including *ZP4* and *ZPAX*) gamete-recognition genes showed evidence for adaptive evolution in birds. For the orthologous genes in mammals, *CD9*, *Acrosin*, *ZP2*, and *ZP4* evolved by positive selection, whereas this was not seen for *ZP1*. There were no significant differences in the number of codons per gene that were positively selected in mammals and birds, suggesting that the pattern and strength of selection is similar in these vertebrate groups for the investigated

Table 3 Tests of positive selection among avian gamete-recognition genes

Gene	M1a-M2a $-2\Delta\ln L(df = 2)$	Parameter estimates under M2a	M7-M8 $-2\Delta\ln L$ ($df = 2$)	Parameter estimates under M8	M8a-M8 $-2\Delta\ln L$ ($df = 1$)	Positively selected sites ^a
<i>CD9</i>	8.26	$p1: 0.78; \omega1: 0.08$	13.48	$p0 = 0.94$	9.81	38, 91, 119, 151, 153, 154, 156, 158, 168, 173, 175, 178, 185
	$p = 0.02$	$p2: 0.18; \omega2: 1.00$	$p = 0.001$	$q = 1.21$	$p = 0.002$	
		$p3: 0.04; \omega3: 3.89$		$(p1 = 0.06) \omega = 2.88$		
<i>Acrosin</i>	3.29	$p1: 0.62; \omega1: 0.04$	8.18	$p0 = 0.84$	3.93	11, 12, 21, 23, 24, 25, 26, 27, 29, 32, 39, 54, 55, 57, 58, 67, 69, 70, 75, 76, 80, 81,
	NS	$p2: 0.32; \omega2: 1.00$	$p = 0.02$	$q = 0.75$	$p = 0.05$	83, 85, 90, 91, 92, 93, 94, 105, 108, 124, 126, 129, 134, 138, 139, 140, 141, 144,
		$p3: 0.06; \omega3: 2.58$		$(p1 = 0.16) \omega = 1.83$		145, 146, 149, 150, 151, 152, 160, 171, 174, 179, 182, 207, 208, 215, 228, 229,
<i>ZPI</i>	4.97	$p1: 0.75; \omega1: 0.07$	10.30	$p0 = 0.82$	5.28	236, 240, 241, 245, 257, 264, 265, 266, 286, 295, 302, 304
	$p = 0.08$	$p2: 0.17; \omega2: 1.00$	$p = 0.006$	$q = 4.16$	$p = 0.02$	4, 16, 21, 26, 28, 41, 43, 79, 84, 96, 99, 112, 113, 134, 147, 148, 155, 156, 163, 164,
		$p3: 0.08; \omega3: 1.96$		$(p1 = 0.18) \omega = 1.60$		166, 170, 174, 179, 184, 186, 188, 204, 208, 211, 318, 325, 349, 366, 390, 396,
<i>ZP2</i>	3.51	$p1: 0.70; \omega1: 0.10$	9.67	$p0 = 0.95$	5.04	397, 402, 413, 479, 514, 527, 556, 567, 569, 589, 616, 665, 668, 706, 707, 729,
	NS	$p2: 0.28; \omega2: 1.00$	$p = 0.008$	$q = 0.88$	$p = 0.02$	754, 770, 771, 772, 812, 814, 829, 844, 866, 868, 879, 888, 910, 913, 917, 924,
		$p3: 0.01; \omega3: 3.87$		$(p1 = 0.05) \omega = 2.43$		926, 932, 935, 938
<i>ZP4</i>	0.39	$p1: 0.73; \omega1: 0.14$	2.43	$p0 = 0.90$	1.04	21, 23, 53, 96, 114, 119, 120, 153, 160, 164, 180, 192, 244, 247, 286, 287, 292, 301,
	NS	$p2: 0.24; \omega2: 1.00$	$p = NS$	$q = 1.67$	$p = NS$	303, 342, 378, 421, 424, 517, 546, 607, 631, 643, 646, 649, 654, 655, 666, 670,
		$p3: 0.03; \omega3: 1.73$		$(p1 = 0.10) \omega = 1.61$		672
<i>ZPAX</i>	0.98	$p1: 0.62; \omega1: 0.08$	2.77	$p0 = 0.94$	1.70	
	NS	$p2: 0.37; \omega2: 1.00$	$p = NS$	$q = 1.05$	$p = NS$	
		$p3: 0.01; \omega3: 3.11$		$(p1 = 0.06) \omega = 3.38$		

NS—not significant

^a Predicted by Bayes Empiric Bayes analysis (Yang et al. 2005); $p > 95\%$ for sites in bold text; $p > 90\%$ for sites in italic text; and $p > 70\%$ for underlined sites

Table 4 Descriptive data for the five gamete-recognition genes in mammals

Gene	N	L	S (M0)	ω (M0)
<i>CD9</i>	10	681	1.82	0.15
<i>Acrosin</i>	11	981 ^a	3.73	0.29
<i>ZP1</i>	7	2154	1.89	0.37
<i>ZP2</i>	13	2280	4.59	0.58
<i>ZP4</i>	10	1788	4.50	0.42

N—number of taxa; L—length of gapped alignment (bp); S—number of substitutions per codon (M0 model); ω — d_N/d_S (M0 model); N§not significant

^a 3' end removed because there were repeat regions, which were difficult to align properly

genes. The results were obtained using a similar number of sequences with similar sequence divergences for the two organismal groups, giving comparable power in the detection of adaptive evolution. We therefore conclude that the molecular evolution of gamete-recognition genes among birds and mammals is not dramatically different despite the fact that there are pronounced differences in reproductive biology between these groups.

The initial stage of avian fertilisation involves the penetration of the inner perivitelline layer (IPVL; analogous to mammalian zona pellucida) by multiple sperm. Multiple sperm then enter into the cytoplasm of a germinal disc (Bakst and Howarth 1977; Birkhead et al. 1994; Okamura and Nishiyama 1978a; Okamura and Nishiyama 1978b; Tarin and Cano 2000). In chicken, >100 sperm can enter an ovum without resulting in decreased fertility (Bramwell et al. 1995). After the sperm have penetrated the IPVL, the outer perivitelline layer is formed, which serves to block pathologic polyspermy (Stepinska and Bakst 2007). Spermatozoal nuclei that have entered the egg cytoplasm decondense and transform into male pronuclei (Okamura and Nishiyama 1978b; Waddington et al. 1998). However, only a single sperm pronucleus fuses with the egg pronucleus, whereas the other supernumary male pronuclei migrate to the periphery and are degraded by DNases present in the ovum (Stepinska and Bakst 2007). Overall, the IPVL does not seem to form such a critical barrier to fertilization as the zona pellucida does in mammals.

It is important to note that the observation of similar rates of molecular evolution of avian and mammalian gamete-recognition proteins does not necessarily prove that adaptive evolution of these proteins in birds is not driven by the same mechanism as in mammals. Although there are arguments in favour of polyspermy avoidance being the fundamental driver in mammals, firm evidence for this conclusion is lacking. We cannot therefore exclude that other forces are acting in both birds and mammals and that the impact of these forces is what gives rise to the similar

patterns in the two lineages. However, if one assumes that polyspermy avoidance is a major driver in mammals, then our data suggest that the mechanism is different in birds. This is in line with the study of Calkins et al. (2007) who, based on the observation of adaptive evolution of avian *ZP3*, concluded that “polyspermy avoidance is not sufficient to explain positive Darwinian selection in reproductive proteins across taxonomic groups.”

If polyspermy avoidance does not explain the rapid evolution of avian gamete-recognition proteins, we must seek alternative explanations. One possibility relates to the fact that heterospecific fertilization (hybridization) causes embryo mortality or results in a hybrid offspring with low fitness. Gamete-recognition proteins may thus evolve under positive selection to form prezygotic copulatory barriers to heterospecific sperm. However, whereas in mammals sperm interaction with the zona pellucida is thought of as a species-specific event (Wassarman et al. 2001), the role for gamete-recognition proteins as reproductive barriers in birds is questionable (Birkhead and Brillard 2007) because it has been demonstrated that chicken sperm can bind to the IPVL of both closely and distantly related species (Stewart et al. 2004). Nevertheless, this suggests that effective postcopulatory barriers have an important role in avian speciation because bird hybrids are rare (Birkhead and Brillard 2007).

An alternative possibility is that the evolution of gamete-recognition proteins in birds is related to postcopulatory sexual selection because sperm competition is a widespread phenomenon among birds, and cryptic female choice is thought to be common (Birkhead et al. 2004; Birkhead and Pizzari 2002). It is not known how specific sperm are selected and at which stage between insemination and fertilisation sperm selection takes place, although Birkhead and Brillard (2007) list five stages where it could happen: (1) when sperm traverse the vagina; (2) when sperm enter or exit the sperm-storage tubules (SSTs); (3) when sperm are transported from the sperm-storage tubules to the infundibulum; (4) when sperm penetrate the IPVL; and (5) when sperm locate and or fuse with the female pronucleus. Potentially, gamete-recognition proteins may act as the barrier between sperm and egg in one of the later stages.

To conclude, we have shown that gamete-recognition proteins evolve by positive selection in birds, similar to what is the case in mammals (Swanson et al. 2001). The evolutionary forces driving the rapid divergence of these genes are likely to be different forms of sexual conflict between female and male gametes, a conflict created by way of the differential cost of failed or suboptimal fertilisation between male and female organisms. If polyspermy avoidance is not the main mechanism for this in birds, as has been postulated for other organisms, postcopulatory sexual selection, such as cryptic female choice, could create the sexual conflict

Table 5 Tests of positive selection among mammalian gamete-recognition genes

Gene	M1a–M2a –2ΔlnL (df = 2)	Parameter estimates under M2a	M7–M8 –2ΔlnL (df = 2)	Parameter estimates under M8	M8a–M8 –2ΔlnL (df = 1)	Positively selected sites ^a
<i>CD9</i>	1.73	<i>p</i> 1: 0.86, <i>ω</i> 1: 0.04	10.73	<i>p</i> 0 = 0.92 <i>p</i> = 0.24	4.20	<u>52</u> , <u>133</u> , <u>147</u> , <u>157</u> , <u>170</u> , <u>171</u> , <u>173</u> , <u>174</u> , <u>175</u> , <u>176</u> , <u>177</u> , <u>185</u> , <u>189</u>
	NS	<i>p</i> 2: 0.10, <i>ω</i> 2: 1.00	<i>p</i> = 0.005	<i>q</i> = 3.03	<i>p</i> = 0.04	
		<i>p</i> 3: 0.04, <i>ω</i> 3: 1.92		(<i>p</i> 1 = 0.09) <i>ω</i> = 1.67		
<i>Acrosin</i>	31.96	<i>p</i> 1: 0.70, <i>ω</i> 1: 0.06	45.54	<i>p</i> 0 = 0.91 <i>p</i> = 0.24	28.16	<u>65</u> , <u>69</u> , <u>76</u> , <u>78</u> , <u>79</u> , <u>120</u> , <u>126</u> , <u>164</u> , <u>165</u> , <u>166</u> , <u>167</u> , <u>169</u> , <u>224</u> , <u>249</u> , <u>259</u>
	<i>p</i> < 0.001	<i>p</i> 2: 0.25, <i>ω</i> 2: 1.00	<i>p</i> < 0.001	<i>q</i> = 0.89	<i>p</i> < 0.001	
		<i>p</i> 3: 0.05, <i>ω</i> 3: 3.34		(<i>p</i> 1 = 0.09) <i>ω</i> = 2.28		
<i>ZPI</i>	0	<i>p</i> 1: 0.63, <i>ω</i> 1: 0.11	2.89	<i>p</i> 0 = 0.63 <i>p</i> = 12.85	5.54 × 10 ⁻⁴	
	NS	<i>p</i> 2: 0.12, <i>ω</i> 2: 1.00	<i>p</i> = NS	<i>q</i> = 99.00	<i>p</i> = NS	
		<i>p</i> 3: 0.25, <i>ω</i> 3: 1.00		(<i>p</i> 1 = 0.37) <i>ω</i> = 1.00		
<i>ZP2</i>	12.52	<i>p</i> 1: 0.46, <i>ω</i> 1: 0.15	26.05	<i>p</i> 0 = 0.87 <i>p</i> = 0.54	15.35	<u>2</u> , <u>3</u> , <u>7</u> , <u>15</u> , <u>16</u> , <u>23</u> , <u>32</u> , <u>40</u> , <u>43</u> , <u>55</u> , <u>62</u> , <u>66</u> , <u>68</u> , <u>82</u> , <u>84</u> , <u>87</u> , <u>92</u> , <u>112</u> , <u>118</u> , <u>122</u> , <u>128</u> , <u>130</u> , <u>131</u> , <u>132</u> , <u>134</u> , <u>142</u> , <u>144</u> , <u>148</u> , <u>150</u> , <u>157</u> , <u>166</u> , <u>167</u> , <u>168</u> , <u>175</u> , <u>176</u> , <u>177</u> , <u>178</u> , <u>182</u> , <u>183</u> , <u>184</u> , <u>185</u> , <u>190</u> , <u>191</u> , <u>196</u> , <u>204</u> , <u>208</u> , <u>213</u> , <u>219</u> , <u>224</u> , <u>226</u> , <u>250</u> , <u>252</u> , <u>258</u> , <u>261</u> , <u>263</u> , <u>267</u> , <u>291</u> , <u>298</u> , <u>308</u> , <u>312</u> , <u>315</u> , <u>331</u> , <u>337</u> , <u>350</u> , <u>351</u> , <u>352</u> , <u>353</u> , <u>365</u> , <u>421</u> , <u>452</u> , <u>472</u> , <u>509</u> , <u>517</u> , <u>557</u> , <u>562</u> , <u>564</u> , <u>568</u> , <u>575</u> , <u>611</u> , <u>612</u> , <u>628</u> , <u>629</u> , <u>646</u> , <u>652</u> , <u>710</u> , <u>715</u> , <u>716</u> , <u>718</u> , <u>748</u> , <u>751</u> , <u>775</u> , <u>758</u>
	<i>p</i> = 0.002	<i>p</i> 2: 0.50, <i>ω</i> 2: 1.00	<i>p</i> < 0.001	<i>q</i> = 0.58	<i>p</i> < 0.001	
		<i>p</i> 3: 0.04, <i>ω</i> 3: 2.36		(<i>p</i> 1 = 0.13) <i>ω</i> = 1.68		
<i>ZP4</i>	0	<i>p</i> 1: 0.52, <i>ω</i> 1: 0.10	9.78	<i>p</i> 0 = 0.96 <i>p</i> = 0.33	8.50	<u>13</u> , <u>14</u> , <u>30</u> , <u>33</u> , <u>45</u> , <u>87</u> , <u>137</u> , <u>149</u> , <u>150</u> , <u>196</u> , <u>286</u> , <u>376</u> , <u>382</u> , <u>409</u> , <u>415</u> , <u>427</u> , <u>456</u> , <u>478</u> , <u>485</u> , <u>506</u> , <u>507</u> , <u>555</u> , <u>560</u> , <u>587</u>
	NS	<i>p</i> 2: 0.29, <i>ω</i> 2: 1.00	<i>p</i> = 0.008	<i>q</i> = 0.43	<i>p</i> = 0.004	
		<i>p</i> 3: 0.19, <i>ω</i> 3: 1.00		(<i>p</i> 1 = 0.04) <i>ω</i> = 2.19		

NS—not significant

^a Predicted by Bayes Empiric Bayes analysis (Yang et al. 2005); *p* > 95% for sites in bold text; *p* > 90% for sites in italic text; and *p* > 70% for underlined sites

Chicken PVKGGTKCIKYLIFGFNFVFWLAGTAVLAIGLWLRFDVQTKSIFELESNN
 Human PVKGGTKCIKYLIFGFNFIFWLAGI AVLAI GLWLRFD SQTKSIFEQETNN

Chicken --MTFYTGVIYILIGAGALMMLVGLGCCGALQESQCMLGLFFVFLVFVIFA
 Human NNSSFYTGVIYILIGAGALMMLVGLGCCGAVQESQCMLGLFFGFLLVIFA

Chicken LEIATAIWFANKEKIVEELQDFYRETYE--KRSQPAAREILKAFHFALN
 Human IETAAAIWGYSHKDEVIKEVQEFYKDTYNKLT KDEPQRETLKATHYALN

Chicken CCGITGGLEQQLMESCPKKTLPESFTTTPCPKAIDDDVFN SKLNIVIGAVGL
 Human CCGLAGGVEQFISDIPC KKDVL ETTFTV KSCPDAI KEVFDNKPHIIGAVGI

Chicken GIAVIMIFGMIFSMVLCCAIRRNREMV
 Human GIAVVMIFGMIFSMILCCAIRRNREMV

Fig. 1 Amino-acid alignment of chicken and human CD9. Positively selected sites are indicated in bold text

underlying the rapid evolution of gamete-recognition proteins. It will be interesting to study the molecular evolution of gamete-recognition proteins in other organisms in which physiologic polyspermy is norm, such as in some reptiles, amphibians, newts, and salamanders.

Acknowledgments The authors thank Stuart Piertney for the red grouse tissue and Niclas Backström, Niclas Kolm, and Judith Mank for useful discussions and comments on the manuscript. We also thank Jiao Xiang for help with laboratory work. Financial support was obtained from the Swedish Research Council.

References

- Baba T, Azuma S, Kashiwabara S, Toyoda Y (1994) Sperm from mice carrying a targeted mutation of the acrosin gene can penetrate the oocyte zona pellucida and effect fertilization. *J Biol Chem* 269:31845–31849
- Bakst MR, Howarth B Jr (1977) Hydrolysis of the hen's perivitelline layer by cock sperm in vitro. *Biol Reprod* 17:370–379
- Berlin S, Smith NG (2005) Testing for adaptive evolution of the female reproductive protein ZPC in mammals, birds and fishes reveals problems with the M7–M8 likelihood ratio test. *BMC Evol Biol* 5:65
- Birkhead TR, Brillard JP (2007) Reproductive isolation in birds: postcopulatory prezygotic barriers. *Trends Ecol Evol* 22:266–272
- Birkhead TR, Chaline N, Biggins JD, Burke T, Pizzari T (2004) Nontransitivity of paternity in a bird. *Evolution* 58:416–420
- Birkhead TR, Pizzari T (2002) Postcopulatory sexual selection. *Nat Rev Genet* 3:262–273
- Birkhead TR, Sheldon BC, Fletcher F (1994) A comparative-study of sperm-egg interactions in birds. *J Reprod Fertil* 101:353–361
- Bramwell RK, Howarth B Jr (1992) Preferential attachment of cock spermatozoa to the perivitelline layer directly over the germinal disc of the hen's ovum. *Biol Reprod* 47:1113–1117
- Bramwell RK, Marks HL, Howarth B (1995) Quantitative-determination of spermatozoa penetration of the perivitelline layer of the hen ovum as assessed on oviposited eggs. *Poult Sci* 74:1875–1883
- Calkins JD, El-Hinn D, Swanson WJ (2007) Adaptive evolution in an avian reproductive protein: ZP3. *J Mol Evol* 65:555–563
- Civetta A (2003) Positive selection within sperm-egg adhesion domains of fertilin: an ADAM gene with a potential role in fertilization. *Mol Biol Evol* 20:21–29
- Clark NL, Aagaard JE, Swanson WJ (2006) Evolution of reproductive proteins from animals and plants. *Reproduction* 131:11–22
- Frank SA (2000) Sperm competition and female avoidance of polyspermy mediated by sperm-egg biochemistry. *Evol Ecol Res* 2:613–625
- Gardner AJ, Evans JP (2006) Mammalian membrane block to polyspermy: new insights into how mammalian eggs prevent fertilisation by multiple sperm. *Reprod Fertil Dev* 18:53–61
- Gasper J, Swanson WJ (2006) Molecular population genetics of the gene encoding the human fertilization protein zonadhesin reveals rapid adaptive evolution. *Am J Hum Genet* 79:820–830
- Gavrilets S (2000) Rapid evolution of reproductive barriers driven by sexual conflict. *Nature* 403:886–889
- Geyer LB, Palumbi SR (2003) Reproductive character displacement and the genetics of gamete recognition in tropical sea urchins. *Evolution* 57:1049–1060
- Hellberg ME, Vacquier VD (1999) Rapid evolution of fertilization selectivity and lysin cDNA sequences in teguline gastropods. *Mol Biol Evol* 16:839–848
- Hughes DC (2007) ZP genes in avian species illustrate the dynamic evolution of the vertebrate egg envelope. *Cytogenet Genome Res* 117:86–91
- Kaiser VB, van Tuinen M, Ellegren H (2007) Insertion events of CR1 retrotransposable elements elucidate the phylogenetic branching order in galliform birds. *Mol Biol Evol* 24:338–347
- Kumar S, Tamura K, Nei M (2004) MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. *Brief Bioinform* 5:150–163
- Lefevre L, Conner SJ, Salpekar A, Olufowobi O, Ashton P, Pavlovic B, Lenton W, Afnan M, Brewis IA, Monk M et al (2004) Four zona pellucida glycoproteins are expressed in the human. *Hum Reprod* 19:1580–1586
- Litscher ES, Wassarman PM (2007) Egg extracellular coat proteins: from fish to mammals. *Histol Histopathol* 22:337–347
- McCartney MA, Lessios HA (2004) Adaptive evolution of sperm binding tracks egg incompatibility in neotropical sea urchins of the genus *echinometra*. *Mol Biol Evol* 21:732–745
- Metz EC, Palumbi SR (1996) Positive selection and sequence rearrangements generate extensive polymorphism in the gamete recognition protein binding. *Mol Biol Evol* 13:397–406
- Miyado K, Yamada G, Yamada S, Hasuwa H, Nakamura Y, Ryu F, Suzuki K, Kosai K, Inoue K, Ogura A et al (2000) Requirement of CD9 on the egg plasma membrane for fertilization. *Science* 287:321–324
- Murphy WJ, Eizirik E, O'Brien SJ, Madsen O, Scally M, Douady CJ, Teeling E, Ryder OA, Stanhope MJ, de Jong WW et al (2001) Resolution of the early placental mammal radiation using Bayesian phylogenetics. *Science* 294:2348–2351
- Okamura F, Nishiyama H (1978a) The passage of spermatozoa through the vitelline membrane in the domestic fowl, *Gallus gallus*. *Cell Tissue Res* 188:497–508
- Okamura F, Nishiyama H (1978b) Penetration of spermatozoon into the ovum and transformation of the sperm nucleus into the male pronucleus in the domestic fowl, *Gallus gallus*. *Cell Tissue Res* 190:89–98
- Palumbi SR (1992) Marine speciation on a small planet. *Trends Ecol Evol* 7:114–118
- Rice WR, Holland B (1997) The enemies within: intergenomic conflict, interlocus contest evolution (ICE), and the intraspecific Red Queen. *Behav Ecol Sociobiol* 41:1–10
- Riginos C, McDonald JH (2003) Positive selection on an acrosomal sperm protein, M7 lysin, in three species of the mussel genus *Mytilus*. *Mol Biol Evol* 20:200–207
- Runge KE, Evans JE, He ZY, Gupta S, McDonald KL, Stahlberg H, Primakoff P, Myles DG (2007) Oocyte CD9 is enriched on the microvillar membrane and required for normal microvillar shape and distribution. *Dev Biol* 304:317–325

- Springer SA, Crespi BJ (2007) Adaptive gamete-recognition divergence in a hybridizing *Mytilus* population. *Evolution* 61:772–783
- Stepinska U, Bakst MR (2007) Fertilization. In: Jamieson BGM (ed) *Reproductive biology and phylogeny of birds*. Science Publishers, Enfield, NH, pp 553–587
- Stewart SG, Bausek N, Wohlrab F, Schneider WJ, Janet Horrocks A, Wishart GJ (2004) Species specificity in avian sperm: perivitelline interaction. *Comp Biochem Physiol A* 137:657–663
- Swanson WJ, Nielsen R, Yang Q (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. *Annu Rev Ecol Evol Syst* 33:161–179
- Swanson WJ, Yang Z, Wolfner MF, Aquadro CF (2001) Positive Darwinian selection drives the evolution of several female reproductive proteins in mammals. *Proc Natl Acad Sci USA* 98:2509–2514
- Tarin JJ, Cano A (2000) *Fertilization in protozoa and metazoan animals*. Springer, Berlin, Germany
- 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
- Vacquier VD, Swanson WJ, Lee YH (1997) Positive Darwinian selection on two homologous fertilization proteins: what is the selective pressure driving their divergence? *J Mol Evol* 44(suppl 1): S15–S22
- Waddington D, Gribbin C, Sterling RJ, Sang HM, Perry MM (1998) Chronology of events in the first cell cycle of the polyspermic egg of the domestic fowl (*Gallus domesticus*). *Int J Dev Biol* 42:625–628
- Wassarman PM (1988) Zona pellucida glycoproteins. *Annu Rev Biochem* 57:415–442
- Wassarman PM, Jovine L, Litscher ES (2001) A profile of fertilization in mammals. *Nat Cell Biol* 3:E59–E64
- 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* 168:1041–1051
- Yang Z (1997) PAML: a program package for phylogenetic analysis by maximum likelihood. *Comput Appl Biosci* 13:555–556
- Yang Z (2007) PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol* 24:1586–1591
- Yang Z, Swanson WJ, Vacquier VD (2000) Maximum-likelihood analysis of molecular adaptation in abalone sperm lysin reveals variable selective pressures among lineages and sites. *Mol Biol Evol* 17:1446–1455