Sexual selection drives weak positive selection in protamine genes and high promoter divergence, enhancing sperm competitiveness

Juan Martin-Coello¹, Hernán Dopazo², Leonardo Arbiza², Juan Ausío³, Eduardo R. S. Roldan¹,4,† and Montserrat Gomendio¹,5,*<sup>‡</sup>

1 Reproductive Ecology and Biology Group, Museo Nacional de Ciencias Naturales (CSIC), c/Jose´ Gutie´rrez Abascal 2, 28006 Madrid, Spain
2 Bioinformatics and Genomics Department, Centro de Investigación Principe Felipe (CIPF), Valencia 46013, Spain
3 Department of Biochemistry and Microbiology, University of Victoria, Victoria, British Columbia V8W 3P6, Canada
4 Department of Veterinary Basic Sciences, The Royal Veterinary College, London NW1 0TU, UK
5 Department of Zoology, University of Cambridge, Downing Street, Cambridge CB2 3EF, UK

Phenotypic adaptations may be the result of changes in gene structure or gene regulation, but little is known about the evolution of gene expression. In addition, it is unclear whether the same selective forces may operate at both levels simultaneously. Reproductive proteins evolve rapidly, but the underlying selective forces promoting such rapid changes are still a matter of debate. In particular, the role of sexual selection in driving positive selection among reproductive proteins remains controversial, whereas its potential influence on changes in promoter regions has not been explored. Protamines are responsible for maintaining DNA in a compacted form in chromosomes in sperm and the available evidence suggests that they evolve rapidly. Because protamines condense DNA within the sperm nucleus, they influence sperm head shape. Here, we examine the influence of sperm competition upon protamine 1 and protamine 2 genes and their promoters, by comparing closely related species of Mus that differ in relative testes size, a reliable indicator of levels of sperm competition. We find evidence of positive selection in the protamine 2 gene in the species with the highest inferred levels of sperm competition. In addition, sperm competition levels across all species are strongly associated with high divergence in protamine 2 promoters that, in turn, are associated with sperm swimming speed. We suggest that changes in protamine 2 promoters are likely to enhance sperm swimming speed by making sperm heads more hydrodynamic. Such phenotypic changes are adaptive because sperm swimming speed may be a major determinant of fertilization success under sperm competition. Thus, when species have diverged recently, few changes in gene-coding sequences are found, while high divergence in promoters seems to be associated with the intensity of sexual selection.

Keywords: protamine genes; gene expression; sexual selection; sperm competition; spermatozoa; speciation

1. INTRODUCTION
Phenotypic evolution may occur through changes in gene-coding sequences (CDSs) or regulatory regions, but the relative contribution of the two is a matter of debate. Some authors argue that regulatory mutations may have a predominant role in phenotypic evolution, mainly because the modular nature of regulatory regions largely frees them from deleterious pleiotropic effects, thus allowing selection to operate more efficiently by minimizing functional trade-offs (Carroll 2005; Wray 2007). The main argument against this view is that there is still little empirical evidence to support the role of regulatory mutations on phenotypic evolution, while there is ample evidence that structural mutations do play an important role (Hoekstra & Coyne 2007). However, this may be owing to the fact that fewer studies have examined the role of changes in regulatory regions, which is partly because, unlike coding regions, regulatory sequences are difficult to identify (Gilad et al. 2006). Several lines of evidence suggest that gene regulation may play an important role in evolutionary change, although this evidence is not conclusive. First, genomic studies have found interspecific divergence in gene expression, but they have not been able to link this divergence to phenotypic effects (Pollard et al. 2006). Second, single-locus studies have compared differences in phenotype among species with the pattern of expression of a single gene thought to influence that phenotype (Beldade et al. 2002; Reed & Serfas 2004; Abshanov et al. 2006). Because such comparisons commonly involve distantly related species, there are often differences in the coding regions as well as in the regulatory regions, making it difficult to rule out the effect of structural mutations. Third, there is evidence that

*Author and address for correspondence: Museo Nacional de Ciencias Naturales (CSIC), c/José Gutiérrez Abascal 2, 28006 Madrid, Spain (montseg@mn.cn.csic.es).
†These authors contributed equally to this work.


Received 14 February 2009
Accepted 11 March 2009

This journal is © 2009 The Royal Society

This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
changes in regulatory elements are linked to adaptive traits (Sucena & Stern 2000; Shapiro et al. 2004; Prud’homme et al. 2006), but individual mutations in regulatory elements have not been identified and all the cases involve trait loss.

Some studies have found that the rate of protein divergence between species is not correlated with the rate of expression divergence (Larracuente et al. 2008), although other studies have reached different conclusions (Khaitovich et al. 2005). This suggests that in some genes, positive selection may act at the level of protein sequence, but not at the level of gene expression, and vice versa in other genes. Genes with male-biased expression consistently show the greatest divergence between species at both sequence and expression levels (Ellegren & Porsch 2007).

Most studies have focused on differences between species in genes expressed in male reproductive tissues, showing evidence of positive selection when distantly related species are compared (Swanson & Vacquier 2002; Shapiro et al. 2004; Prud’homme et al. 2004; but see Hurle et al. 2007). The fact that these reproductive genes are expressed in the male germ line and accessory glands suggests that such rapid changes may have been favoured by sperm competition. Despite the popularity of this hypothesis, there is limited supporting evidence. An association between the rate of evolution of reproductive genes and levels of sperm competition has been reported for semenogelin II, a structural component of semen coagulum in primates (Dorus et al. 2004; but see Hurle et al. 2007), and Sos2, which encodes the major component of the rodent copulatory plug (Ramm et al. 2008). In addition, the rate of evolution of sperm zonadhesin has been found to be inversely related to the degree of sexual dimorphism in primates, which is consistent with the hypothesis that selection is related to levels of sperm competition (Herlyn & Zischler 2007). However, no study so far has shown how differences in gene and protein sequences lead to differences in ejaculate competitiveness upon which selection could plausibly operate. The evolution of expression of reproductive genes has received much less attention, and no study so far has examined whether sexual selection does play a role in the evolution of the expression of reproductive genes.

Among mammals, evidence of positive selection in reproductive proteins has often been found when comparing distantly related species, but studies on closely related species have generated conflicting results (Jansa et al. 2003; Turner & Hoekstra 2006). A study on muroid rodents has found evidence of positive selection in four out of seven reproductive genes examined, but evidence suggesting that sexual selection favours adaptive evolution has been found for only one reproductive gene (Ramm et al. 2008). Studies on groups of species where reproductive isolation has evolved recently are needed to identify the first changes promoting divergence between species, and to understand the adaptive significance of such early changes. Comparing divergence in CDSs and promoters during the incipient stages of speciation will allow us to find out whether sexual selection operates more efficiently at one level or the other. The fact that studies on closely related species either do not find evidence of positive selection in reproductive genes or fail to establish a link between evolutionary rates and levels of sexual Selection for most genes is suggested that, during the early stages of divergence, sperm competition may enhance sperm competitiveness largely by influencing gene expression.

Protamines have been found to be among the fastest-evolving reproductive proteins in mammals (Wyckoff et al. 2000; Torgerson et al. 2002; Ramm et al. 2008), but the selective forces driving this extremely rapid rate are a matter of debate. Wyckoff et al. (2000) found evidence of more diverged sequences among humans and chimpanzees than gorillas, and argued that sperm competition may play a role under the assumption that protamines influence sperm morphology and male fertilizing ability. However, this study has been criticized for two main reasons: it fails to provide a link between the processes that protamines regulate and sperm competition mechanisms, and it focuses on gene sequences whereas it is the levels of protamine expression that have been found to be related to male fertility (Clark & Civetta 2000). An alternative hypothesis suggests that protamines are under purifying selection (Rooney et al. 2000).

Protamines are the most abundant sperm nuclear basic proteins in many species and are responsible for maintaining DNA in a compacted form in chromosomes in sperm (Oliva 2006). In mammals, there are two types of protamine: protamine 1 (Prm 1) and protamine 2 (Prm 2). Prm 1 is present in all species of vertebrates studied, but Prm 2 is present only in some species, including rodents, ungulates and primates (Corzett et al. 2002), which could indicate a more basic and conserved function for Prm 1 and an accessory function for Prm 2 in some species.

Here, we test whether sexual selection drives rapid evolution in Prm 1 and Prm 2 genes and their promoters by comparing 10 closely related species of Mus that differ in the levels of sperm competition. Previous work has shown that sperm competition has had a profound influence on the evolution of reproductive traits in this group of species, influencing not only the rate of sperm production but also the proportion of sperm ready to fertilize and their sensitivity to ovum signals (Gomendio et al. 2006). In addition, we have been able to show recently that, in this group of species, sperm competition plays an important role in favouring rapid changes both in sperm (increased sperm competitiveness) and ovum (increased ovum defensiveness), which lead to incipient reproductive barriers (Martin-Coello et al. 2009).

2. MATERIAL AND METHODS

(a) Collection of testes

The study includes males from 10 species of Mus: Mus cookii, Mus fascinus, Mus macedonicus, Mus musculus bactrianus, Mus musculus castaneus, Mus musculus domesticus, Mus musculus musculus, Mus pahari, Mus spicilegus and Mus spretus. Apodemus sylvaticus was used as an outgroup. Previous studies have shown that, despite being closely related, interspecific differences in levels of sperm competition are associated with increases in sperm competitiveness and ovum defensiveness, which lead to asymmetric reproductive barriers (Gomendio et al. 2006; Martin-Coello et al. 2009).

Males of these Mus species were purchased from the Institut des Sciences de l’Evolution, CNRS-Université Montpellier 2, France. Apodemus sylvaticus males were caught in the wild (Sierra de Guadarrama, Madrid, Spain). Animals were kept and bred at the MNCN animal facility under 14 L: 10 D, and 22°C. All the males were kept in individual cages after weaning and allowed free access to food and water.
Males (n = 5 for each species) were sacrificed by cervical dislocation and weighed. Testes were removed, weighed, frozen by direct immersion in liquid nitrogen and stored at −70°C. Relative testes size was calculated according to the rodent regression equation of Kenagy & Trombulak (1986), i.e. as observed mass of both testes/expected mass, where expected mass = 0.031 × body mass0.77. For each species, relative testes size was used as a reliable indicator of sperm competition levels.

(b) Sperm parameters
Spermatozoa were collected from epididymides and vasa deferentia by allowing sperm to swim out for 7 min into a HEPES-buffered modified Tyrode’s medium (mT-H; Shi & Roldan 1995).

Sperm morphology was assessed by staining sperm smears with Giemsa (Watson 1975; Tamuli & Watson 1994). Smears were air-dried and fixed in 4 per cent formaldehyde in tetratus–phosphate buffer (TPB). The smears were washed with water for 10 min, then submerged for 60 min in a Giemsa solution (4.5 ml Giemsa stock solution, 3 ml buffer TPB and 32.5 ml distilled water). The stained smears were washed with distilled water, dried and mounted using DePeX (BDH, Madrid, Spain). Sperm abnormalities were classified depending on whether they affected the head, midpiece or rest of the flagellum.

Sperm dimensions were obtained in Giemsa-stained cells. Images were captured using a microscope (Nikon Labophot 2) with a 40 × objective under bright field and a monochrome charge-coupled device video camera (Sony SSC-M370CE). Images were digitized and analysed in an IBM-compatible computer using VISILOG software (VISILOG v. 4.1.3 Rev 6, Noesis, Vélizy, France).

Objective measures of sperm motility were recorded in spermatozoa suspended in mT-H medium, using a computer-aided sperm analyser (Hobson Sperm Tracker, Sheffield, UK). Assessments were made within 5 min of sperm collection for all species. A total of five descriptors of sperm motility were scored by analysing a minimum of 100 tracks per sample: (i) curvilinear velocity (VCL), velocity over total distance moved, including all deviations of sperm head movement; (ii) straight-line velocity (VSL), velocity calculated using the straight-line distance between the beginning and end of the sperm track; (iii) average path velocity (VAP), velocity over a calculated, smoothed path, which is a shorter distance than that used for calculating VCL; (iv) amplitude of lateral head displacement, the mean value of the extreme side-to-side movement of the sperm head in each beat cycle; and (v) linearity (LIN), the ratio (as a percentage) of the distances of straight-line track length/actual track length (this value is 100% for a completely linear track).

(c) DNA isolation (phenol–chloroform extraction)
DNA extraction was performed by using a modified phenol–chloroform–isoamylalcohol method (Sambrook et al. 1989).

Testes were ground in a mortar with liquid nitrogen until a fine powder was obtained. The powder was thawed in 10 ml of extraction buffer (10 mM Tris–HCl (pH 8.0), 0.1 M EDTA, 20 mg RNAse, 0.6% sodium dodecyl sulphate). The sample was mixed by vortexing and incubating for 60 min at 37°C, then 50 μl of stock solution of 20 mg proteinase K ml−1 were added. The sample was mixed by gentle swirling and incubated at 50°C for 3 hours with gentle swirling. The sample was then cooled to room temperature and an equal volume of phenol–chloroform (25 : 24) was added. The sample was then incubated overnight at 4°C.

The following day the sample was centrifuged at 5000 r.p.m. in a microfuge for 10 min at 4°C in order to separate the aqueous and organic phases. The supernatant was transferred to a clean tube to which an equal volume of phenol–chloroform (25 : 24) was added. The sample was incubated for 10 min at 4°C, and then centrifuged at 5000 r.p.m. for 10 min at 4°C to separate the aqueous and organic phases. This step was repeated twice. After the last centrifugation, the supernatant was transferred to another tube and an equal volume of phenol–chloroform–isoamylalcohol (25 : 24 : 1) was added, followed by incubation for 10 min at 4°C. The sample was centrifuged again for 10 min at 4°C, the supernatant transferred to another tube, and 0.2 volumes of 10 M NH4OAc and two volumes of 95 per cent ethanol were added.

A bent Pasteur pipette was used to allow the DNA to clump onto the pipette, which was then dipped twice in a tube with 5 ml of 70 per cent ethanol. Collected DNA was transferred to a tube with 2 ml of TE (10 mM Tris–HCl (pH 8.0), 0.1 mM EDTA) and stored at 4°C.

(d) PCR and cloning of PCR products
Genomic DNA extracted from testes was used as template in polymerase chain reactions (PCRs).

Primers used for protamine 1 (Prm 1) were

— forward: 5'-CTCCCGGCAAGCCAGCACC-3',
— reverse: 5'-GGACTTGCTATTCTGTGCAT-3'.

Primers used for protamine 2 (Prm 2) were

— forward: 5'-CTCCTCCTATCTCCGGACAC-3',
— reverse: 5'-ATGGACAGGCTGGGGAGGC-3'.

Primers used for protamine 1 promoter (Prm 1 promoter) were

— forward: 5'-CTGGCCGACAGCATCGGTATCT-3',
— reverse: 5'-TCCTCAGGACATGGTGGGCC-3'.

Primers used for protamine 2 promoter (Prm 2 promoter) were

— forward: 5'-ATTCGGTGATTGGAACCATGCT-3',
— reverse: 5'-AAAGTTTGCTTTGTCATGCT-3'.

Primers were designed using the M. m. musculus DNA sequences for Prm 1 (GenBank accession number NM 013637) and Prm 2 (GenBank accession number NM 008933).

PCRs contained PCR buffer (50 mM KCl, 10 mM Tris–HCl (pH 8.3), 1.5 mM MgCl2), 200 μM each of the four deoxynucleoside triphosphates, 2 U of Taq DNA polymerase, 0.5 mM of each primer and 100 ng of DNA.

Resulting PCR products were purified using Wizard PCR Prep DNA Purification System (Promega, Alcobendas, Spain). Purified PCR products were then inserted into pCR 2.1-TOPO vector (Invitrogen, Barcelona, Spain) following the manufacturer’s instructions and transformed into TOP10 competent cells (Invitrogen). Transformed cells were plated on Luria–Bertani (LB)-agar medium containing ampicillin and 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal) for the selection of positive clones. Positive colonies

Proc. R. Soc. B (2009)
were picked and used to inoculate in tubes containing LB medium plus ampicillin, which were incubated overnight in humidified containers at 37°C with shaking.

The following day plasmid was purified using Wizard Plus Minipreps DNA Purification System (Promega), followed by restriction digestion with ECO RI (Roche Diagnostics, Barcelona, Spain) in order to confirm positive clones; positive clones were finally sequenced.

Sequence alignments were performed using the distance-based program ClustalX v. 1.83 (Thompson et al. 1997) with default parameters.

(e) Phylogenetic analyses
A phylogenetic tree was constructed using sequences available in GenBank for the species used in this study (table S1 in the electronic supplementary material). CDSs were aligned using translated protein sequences as templates by means of Muscle (Edgar 2004) and default parameters. Phylogenetic testing of the best nucleotide substitution model was done using Modeltest program (Posada & Crandall 1998). The un ungapped number of characters ranged from 11 038 to 13 073 for M. m. domesticus, M. cookii, M. macedonicus, M. m. castaneus, M. m. musculus, M. spretus and M. spicilegus, from 4894 to 8397 for M. famulus, M. pahari and A. sylvaticus, and were 1645 for M. m. bactrianus. Maximum-likelihood and Bayesian analyses were run in PhyML (Guindon & Gascuel 2003) and MrBayes (Ronquist & Huelsenbeck 2003) programs according to the best-fit model of DNA. Convergence of the four Markov chains was obtained in 1 000 000 generations, and 500 samples out of 1000 were used to summarize the posterior probability of all the trees. Note that the species with the lower number of un ungapped characters produces a polytomy with the lowest posterior probability value (0.79) observed in the tree (figure 16). Comparative analyses by independent contrasts (CAIC, v. 2.6.9; Purvis & Rambout 1995) were used to control for phylogenetic effects.

(f) Positive Darwinian selection
Evolutionary rates at the codon level were computed using the CODEML program from the PAML (v. 3.15) package (Yang 2007). Maximum-likelihood site models were fitted to protamine sequences using the PhyML tree topology deduced above. Likelihood ratio tests were computed between results of alternative models in order to infer events of positive selection. Rate of variation among sites was modelled using four categories, codon frequencies were estimated under the F3 × 4 model and the kappa (ts/tv) parameter was optimized in all of the ML site models. The site models computed were M1, M2, M7 and M8 (Yang et al. 2000).

(g) Evolutionary parameters of promoters
Promoter sequences for both protamine 1 and protamine 2, as well as the more common transcription activation elements (TATA box, Y-box, Tet1, CAAT box, CRE and half CRE sequence; Mayr & Montminy 2001; Aoki & Carrell 2003), were also analysed in order to find out whether differences in the sequence of promoters and the position of transcription activation elements could influence protamine expression.

The evolutionary distances of Prm 1 and Prm 2 promoters were compared against the concatenated alignments of introns (INT) available for the genes SmcX Salivary

Androgen binding protein gene, zp2-3, zp3-3, SmcY and tcpI. INT sequences were downloaded from GenBank. Orthologous promoter regions of protamine genes and concatenated sequences of INT were aligned using Muscle. After Modeltest optimization, promoters of protamine 1, protamine 2 and INT sequences fitted HKY + G, K80 + G and TN93 DNA models, respectively. In order to compare evolutionary rates between promoters and INT sequences, we generated a distribution of branch lengths for each group of sequences based on bootstrap replications. PhyML program was run according to the best-fit model on each alignment using 10 000 replicates, assuming a constrained topology corresponding to the protein tree deduced in this study. Apodemus sylvaticus was excluded since no INT information was obtained from the GenBank database. Therefore, divergence was estimated as the branch length between the internal node joining M. m. musculus and M. famulus, and each of the descendant species of this node.
The mean branch length distribution values for Prm 1 promoter, Prm 2 promoter and the INT were compared and used to study the relationship between promoter evolution, relative testes size and sperm swimming velocity. Linear regressions were performed using the R statistics package (Ihaka & Gentleman 1996).

Promoter and gene sequences for both protamines of all species included in this study have been deposited in GenBank (accession numbers: FJ411373–FJ411394).

3. RESULTS AND DISCUSSION

(a) Sequence analysis and phylogeny

We have examined a group of 10 species of Mus that, despite being closely related, differ in the levels of sperm competition. As shown in figure 1a, these species show clear differences in relative testes size, which is a reliable indicator of levels of sperm competition in most taxa. Thus, the 10 species cover the whole range from high to low sperm competition levels.

To perform comparative analyses, we used the available sequences for a variety of genes (table S1 in the electronic supplementary material) and constructed a phylogenetic tree derived from a gapped alignment of 15 415 characters for the 10 Mus taxa, using A. syrichticus as an outgroup (figure 1b). MODELTEST selected the TN93 + G + I, with 70 per cent of invariant sites and rate heterogeneity parameter value of 0.725. No differences in topologies were observed between the results of ML and Bayesian approaches (figure 1b). All the clades found the maximum posterior probability with the exception of the group clustering M. m. musculus, M. m. bactrianus and M. m. castaneus ($p = 0.79$). In this case, we follow the analysis assuming the polytomy.

Protamine sequences were obtained and amino acid sequences deduced (figs S1 and S2 in the electronic supplementary material). Genetic divergence among sequences was low (see branch lengths of figure 1b, and figs S1 and S2 in the electronic supplementary material). Protamine amino acid sequences are identical for all species of Mus, with two exceptions: M. famulus and M. pahari revealed two amino acid substitutions (fig. S1B in the electronic supplementary material). The amino acid sequences of protamine 2 reveal changes on four residues: 22, 80, 98 and 106. Ingroup and outgroup species differentiate at a single residue, 98. Mus cooki and M. pahari share the basal state condition at residue 80, while M. spicilegus shows a derived state at residue 22, and M. sputes, M. famulus and M. spicilegus show a common derived state at position 106 (fig. S2B in the electronic supplementary material).

(b) Positive selection in protamine genes

Using the phylogenetic tree of figure 1, we tested adaptive evolution on protamine genes by maximum-likelihood site test methods. Protamine 1 did not show evidence of adaptive evolution when using the site methods (table 1). By contrast, in protamine 2, positive selection was detected at residue 106 ($p < 0.01$) when likelihood ratio tests were applied to compare the alternative selection and nearly neutral models (table 1). In both cases, the likelihood ratio tests found statistically significant differences between them when considering the three possible topologies to arrange the dichotomy of figure 1b. Thus, in contrast to the results obtained for Prm 1, the statistical analyses performed on Prm 2 suggest low values of gene divergence mainly modelled by a Darwinian process of adaptive evolution (figure 2).

(c) Promoter evolution

We also tested the possibility that sexual selection may enhance sperm competitiveness by promoting changes in gene regulation. We compared the degree of divergence in protamine 1 and protamine 2 promoters (Prm 1 and Prm 2, respectively) using two different models for the distribution of $\omega$ among sites ($2\Delta_{M1&M2} = 12.60; 2\Delta_{M7&M8} = 12.62 > \chi^2_{0.001,d.f. = 1} = 10.83$). According to the Bayes empirical Bayes (BEB) analysis, residue 106 has a mean value of $\omega$ between 0.728 and 0.824 with a posterior probability of being positively selected between 0.994 and 0.998 in models M2 or M8, respectively. Residue 22 changing from glutamic acid (E) to glycine (G) in M. spicilegus does not reach the significance cut-off value ($P_{\omega > 1} = 53.0\%$). The $\omega$ ratio is taken as the average over all sites in the alignments. PSS is the number of positively selected sites, inferred above a 50% posterior probability cut-off.)
promoter and Prm 2 promoter; figs S3 and S4 in the electronic supplementary material). Deletion and mutational analyses have revealed that changes in the sequence of Prm 2 promoter may enhance Prm 2 transcription by more than fivefold (Yiu & Hecht 1997). In addition, differences in the levels of expression of Prm 2 between species have been associated with the differences in the efficiency of the promoters (Bunick et al. 1990). Both Prm 1 and Prm 2 promoters showed higher divergence rates than INT. In order to understand whether the genetic divergence of promoters of both protamine genes may be related to the intensity of sexual selection, the fit of genetic divergence of promoters of both protamine genes with relative testes size. In general, Prm 2 promoter is more highly significant than INT. In order to understand whether the deletion and mutational analyses have revealed that changes in the sequence of Prm 2 promoter may enhance Prm 2 transcription by more than fivefold (Yiu & Hecht 1997). In addition, differences in the levels of expression of Prm 2 between species have been associated with the differences in the efficiency of the promoters (Bunick et al. 1990). Both Prm 1 and Prm 2 promoters showed higher divergence rates than INT. In order to understand whether the genetic divergence of promoters of both protamine genes may be related to the intensity of sexual selection, the fit of a linear regression model of the mean evolutionary rates is related to sperm viability, DNA integrity and fertilizing ability (Aoki et al. 2006). In mice, haploinsufficiency of protamines increases sperm morphological abnormalities and damage in sperm DNA, and decreases sperm motility (Cho et al. 2001). More specifically, Prm 2-deficient male mice have increased DNA damage, inefficient chromatin packaging, modified sperm heads and changes in the acrosome (Cho et al. 2003).

Thus, we tested the hypothesis that the rapid divergence observed in Prm 2 promoters is associated with changes in sperm head dimensions and sperm swimming velocity. We analysed sperm morphology and sperm swimming velocity in these 10 Mus taxa. All analyses were corrected for phylogenetic effects using the phylogeny in figure 1a. After controlling for phylogenetic effects, absolute testes size was found to be associated with head width (n=7 contrasts, r^2=0.764, p=0.004), the proportion of sperm with head abnormalities (n=7 contrasts, r^2=0.728, p=0.007) and sperm swimming velocity (VAP: n=7 contrasts, r^2=0.688, p=0.01; VCL: n=7 contrasts, r^2=0.548, p=0.03; VSL: n=7 contrasts, r^2=0.645, p=0.01; LIN, n=7 contrasts, r^2=0.631, p=0.01). Thus, species of rodents with inferred high levels of sperm competition have (i) a lower proportion of sperm with head abnormalities, and (ii) sperm with narrower sperm heads, which (iii) swim faster and follow straighter trajectories.

Furthermore, we were able to establish a direct link between divergence in Prm 2 promoters and sperm swimming velocity. Thus, the mean values of divergence of the regulatory region controlling the expression of Prm 2 show a significant association with sperm swimming velocity (VSL, r^2=0.659, p=0.02; figure 4),
and a marginally non-significant association with LIN ($r^2 = 0.4448$, $p = 0.061$). Previous studies by our research group were able to show that faster-swimming sperm have more elongated and narrower heads (Malo et al. 2006).

Thus, one possibility that deserves further study is that the association between sperm competition, divergence in Prm 2 promoters and sperm swimming velocity is mediated by changes in the degree of condensation of DNA within the sperm head, which in turn may influence sperm head shape.

By contrast, no relationships were found between divergence in Prm 1 promoters and sperm head morphology or dimensions, or sperm swimming velocity.

Our findings show that intense sexual selection is associated with rapid divergence in Prm 2 promoters, which, in turn, correlates with faster sperm swimming velocity.

**Table 2. Promoter evolution and relative testes size.** (Statistical values of linear regression models between the genetic divergence of the regulatory elements of protamines and relative testes size. Note that, irrespective of the way in which the relative divergence of Prm 2 promoter in relation to Prm 1 promoter and/or INT sequences is analysed, all the variables considering the promoter region of Prm 2 fit a linear regression model with statistical confidence ($p < 0.05$). Genetic divergence of the Prm 1 promoter, Prm 2 promoter and INT sequences was estimated as the mean distribution value of branch lengths computed after a 10 000 bootstrap-based analysis (see fig. S3 in the electronic supplementary material). n.s., non-significant at 95% of statistical confidence. *Significant at 95% of statistical confidence. **Significant at 99% of statistical confidence. ***Significant at 99.9% of statistical confidence.)

<table>
<thead>
<tr>
<th>variable</th>
<th>$R^2$</th>
<th>$p$-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>INT</td>
<td>-0.0957</td>
<td>0.5662n.s.</td>
</tr>
<tr>
<td>Prm 1 promoter</td>
<td>-0.0434</td>
<td>0.4322n.s.</td>
</tr>
<tr>
<td>Prm 1 promoter/INT</td>
<td>-0.1631</td>
<td>0.8960n.s.</td>
</tr>
<tr>
<td>Prm 2 promoter</td>
<td>0.7065</td>
<td>0.0055**</td>
</tr>
<tr>
<td>Prm 2 promoter/INT</td>
<td>0.5826</td>
<td>0.0168*</td>
</tr>
<tr>
<td>Prm 2 promoter – Prm 1 promoter</td>
<td>0.8091</td>
<td>0.0014**</td>
</tr>
<tr>
<td>(Prm 2 promoter – Prm 1 promoter)/INT</td>
<td>0.8599</td>
<td>0.0005***</td>
</tr>
<tr>
<td>Prm 2 promoter/Prm 1 promoter</td>
<td>0.8233</td>
<td>0.0011**</td>
</tr>
<tr>
<td>(Prm 2 promoter – INT)/(Prm 1 promoter – INT)</td>
<td>0.5619</td>
<td>0.0195**</td>
</tr>
</tbody>
</table>
sperm competition intensity, which modify male reproductive traits (including sperm size) in just a few generations (LaMunyon & Ward 2002). These findings support the idea that, for short evolutionary time scales, CDSs would be expected to evolve more slowly than promoter regions.

By integrating several levels of analysis, we have been able to show how changes in behaviour (female sexual promiscuity) create new selective pressures (sperm competition), which favour minor changes in gene sequence (molecular adaptive evolution) and major changes in its regulation (promoter genetic divergence), which are associated with changes in both sperm design and sperm performance. We conclude that, in the incipient stages of speciation, sexual selection may favour more rapid divergence in the regulation of reproductive genes than in their structure, resulting in different evolutionary dynamics for coding and promoter sequences.

All animal manipulations were performed in accordance with the Spanish Animal Protection Regulation, RD1201/2005, which conforms to European Union Regulation 2003/65/CE.

We are very grateful to François Bonhomme and Annie Orth from the Institut des Sciences de l’Evolution, CNRS-Université Montpellier 2, France, for rearing colonies of these wild mice and for facilitating our purchases of animals. We thank Concepcion Magaña for doing some initial ground work and W. V. Holt for allowing us to analyse mouse sperm kinetics in his laboratory. This work was funded by the Spanish Ministry of Education and Science (grants to M.G., E.R.S.R. and H.D.) and the Natural Science and Engineering Research Council (NSERC) of Canada (grant to J.A.). E.R.S.R. is the recipient of a Royal Society Wolfson Research Merit Award. J.M.-C. enjoyed a PhD studentship from the Spanish Ministry of Education and Science.

REFERENCES


Figure 4. Relationship between mean values of divergence of the regulatory region controlling the expression of *Prm 2* and sperm swimming velocity (straight-line velocity: VSL, $R^2=0.659$, $p=0.02$). The divergence unit is the mean number of nucleotide substitution per site.


Hurle, B., Swanson, W., NISC Comparative Sequencing Program & Green, E. D. 2007 Comparative sequence analyses reveal rapid and divergent evolutionary changes of the WFDC locus in the primate lineage. *Genome Res.* 17, 276–286. (doi:10.1101/gr.600407)


*Sexual selection and protamine genes* J. Martin-Coello et al. 2435

Proc. R. Soc. B (2009)


