

Mammalian Sperm Proteins Are Rapidly Evolving: Evidence of Positive Selection in Functionally Diverse Genes

Dara G. Torgerson, Rob J. Kulathinal,¹ and Rama S. Singh

Department of Biology, McMaster University, Hamilton, Ontario, Canada

A growing number of genes involved in sex and reproduction have been demonstrated to be rapidly evolving. Here, we show that genes expressed solely in spermatozoa represent a highly diverged subset among mouse and human tissue-specific orthologs. The average rate of nonsynonymous substitutions per site (K_a) is significantly higher in sperm proteins (mean $K_a = 0.18$; $N = 35$) than in proteins expressed specifically in all other tissues (mean $K_a = 0.074$; $N = 473$). No differences, however, are found in the synonymous substitution rate (K_s) between tissues, suggesting that selective forces, and not mutation rate, explain the high rate of replacement substitutions in sperm proteins. Four out of 19 sperm-specific genes with characterized function demonstrated evidence of strong positive Darwinian selection, including a protein involved in gene regulation, Protamine-1 (PRM1), a protein involved in glycolysis, GAPDS, and two egg-binding proteins, Adam-2 precursor (ADAM2) and sperm-adhesion molecule-1 (SAM1). These results demonstrate the rapid evolution of sperm-specific genes and highlight the molecular action of sexual selection on a variety of characters involved in mammalian sperm function.

Introduction

A fundamental question in evolutionary biology asks which processes generate the enormous diversity of species that are observed in nature. One approach to examine the molecular foundations of species diversity has been to identify the genes that show high divergence between species because these genes may have played an important role in the early stages of species formation. But the rate of nucleotide substitution that a particular gene may experience is contingent on a variety of factors such as its physical location in the genome (Wolfe, Sharp, and Li 1989; Casane et al. 1997; Matassi, Sharp, and Gautier 1999; Williams and Hurst 2000), base composition (Wolfe and Sharp 1993), and dispensability (Hirsh and Fraser 2001). A gene's spatial pattern of expression is also a major determinant of its evolutionary rate (Civetta and Singh 1995; Duret and Mouchiroud 2000) because genes of different functional classes may evolve under different selective pressures. For example, genes expressed in the brain and nervous system may possess significant functional constraints and may tend to evolve slowly (Kuma, Iwabe, and Miyata 1995; Duret and Mouchiroud 2000), whereas genes expressed in the immune system play a role in antigen recognition and coevolution and tend to evolve more rapidly (Hughes and Nei 1988; Kuma, Iwabe, and Miyata 1995; Hughes 1997; Duret and Mouchiroud 2000). Other evidence suggests that genes involved in sex and reproduction also are evolving rapidly (Singh and Kulathinal 2000; Swanson and Vacquier 2002), particularly proteins found in the male reproductive tract (Coulthart and Singh 1988; Wyckoff, Wang, and Wu 2000; Swanson et al. 2001a).

Key words: tissue-specific genes, sperm, rapid evolution, positive Darwinian selection.

Address for correspondence and reprints: Dara G. Torgerson, Department of Biology, McMaster University, 1280 Main Street West, Hamilton, Ontario, Canada L8S 4K1. E-mail: torgerdg@mcmaster.ca.

¹ Present address: Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, Massachusetts.

Mol. Biol. Evol. 19(11):1973–1980. 2002

© 2002 by the Society for Molecular Biology and Evolution. ISSN: 0737-4038

Recently, a number of studies have focused on the rapid evolution of seminal proteins in *Drosophila* and have reported the presence of extensive directional selection (Aguadé, Miyashita, and Langley 1992; Clark et al. 1995; Tsaur, Ting, and Wu 1998; Swanson et al. 2001a). One explanation is that sperm competition between different males results in the rapid divergence of sperm-associated proteins, increasing the probability of successful fertilization. In marine invertebrates, male- and female-expressed genes involved in sperm-egg recognition and binding also have been shown to evolve rapidly (Swanson and Vacquier 1995; Metz and Palumbi 1996; Swanson and Vacquier 1998; Swanson et al. 2001b; Swanson and Vacquier 2002), demonstrating that molecular coevolutionary processes may drive the evolution of sex- and reproduction-related traits. Similarly, positive Darwinian selection may often be the prevailing force acting on male reproductive genes in humans and primates, including several sperm genes (Rooney and Zhang 1999; Wyckoff, Wang, and Wu 2000). It is therefore expected that sperm-expressed genes, as a group, may show higher divergence because of intersexual selective pressures acting on these genes. Morphological evidence supports this hypothesis because sperm morphology reveals substantial diversity, even between closely related species (Eberhard 1985; Pitnick 1996). Yet, even with these examples, it has not been systematically tested whether sperm genes, as a general class, are rapidly evolving. In fact, an alternative explanation is that because sperm proteins perform a critical reproductive function, the vast majority of sperm-associated genes may possess significant selective constraints and thus may follow a distribution of divergence similar to other tissue-specific genes. Here, we report that proteins specifically found in mammalian sperm are rapidly evolving when compared with a large sample of tissue-specific orthologous genes from human and mouse lineages and further demonstrate that positive Darwinian selection has acted on a variety of sperm components.

Materials and Methods

Human and mouse sperm-specific orthologs were retrieved from HomoloGene at the NCBI database,

Table 1
Sperm-Specific Human and Mouse Orthologous Genes

Gene	Human GenBank Acc. No.	Mouse GenBank Acc. No.	K_a	K_s^a	K_a/K_s^a
ACRV1.....	NM_001612	NM_007391	0.229	0.477	0.480
ADAM18.....	AJ133004	AF167405	0.223	0.502	0.445
ADAM2.....	NM_001464	U16242	0.281	0.546	0.515
SAM1.....	X84347	U33958	0.314	0.342	0.916
SPAG17.....	NM_017425	NM_011449	0.150	0.278	0.538
ZPBP.....	NM_007009	D17569	0.119	0.458	0.259
NASP.....	NM_002482	NM_016777	0.106	0.271	0.391
PRM1.....	Z46940	Z47352	0.327	0.515	0.634
PRM2.....	Z46940	Z47352	0.222	0.651	0.340
PRM3.....	Z46940	Z47352	0.139	0.345	0.404
TP1.....	M59924	X12521	0.0656	0.239	0.275
TP2.....	XM_028328	NM_013694	0.446	0.495	0.901
GAPDS.....	NM_014364	NM_008085	0.114	0.597	0.191
ODF1.....	Q14990	Q61999	0.0156	0.802	0.0195
ODF2.....	NM_002540	NM_013615	0.0088	0.478	0.018
PRKA1.....	NM_003488	NM_009648	0.162	0.406	0.402
PRKA3.....	U85715	AF093406	0.149	0.513	0.290
PRKA4.....	NM_003886	NM_009651	0.135	0.406	0.332
PRKAcAMP.....	AF239744	AF239743	0.0273	0.610	0.0448
FTHL17.....	AF285592	AF285569	0.355	0.431	0.823
LDH.....	J02938	L10389	0.156	0.494	0.316
MOV10L1.....	AF285604	AF285587	0.0954	0.619	0.154
MTL5.....	U86074	NM_010841	0.124	0.870	0.143
NR6A1.....	U64876	U14666	0.0096	0.341	0.0282
NXF.....	AF285596	AF285575	0.300	0.434	0.691
SPAG1.....	AF311312	AF181252	0.173	0.531	0.325
SPAG6.....	AF079363	AF173866	0.0181	0.610	0.0297
STK31.....	AF285599	AF285580	0.113	0.367	0.310
TAF2Q.....	AF285595	AF285574	0.271	0.324	0.837
TDRD1.....	AF285606	AF285591	0.167	0.663	0.252
TEX11.....	AF285594	AF285572	0.301	0.208	1.45
TEX12.....	AF285600	AF285582	0.083	0.287	0.288
TEX14.....	AF285601	AF285584	0.305	0.196	1.55
TEX15.....	AF285605	AF285589	0.274	0.311	0.884
USP26.....	AF285593	AF285570	0.492	0.189	2.61

^a Codons with doublet substitutions were removed.

which uses a reciprocal best hits criterion against two or more taxa to infer putative sequence orthology (<http://www.ncbi.nlm.nih.gov/HomoloGene/>). Sequence orthology was further confirmed by using BLASTp (Altschul et al. 1990) to retrieve human and mouse protein sequences, which were aligned using ClustalX, Version 1.81 (Thompson et al. 1997). Gene trees were drawn using the neighbor-joining algorithm (Saitou and Nei 1987) and were examined manually to confirm orthologs and exclude paralogs. Sperm specificity was determined through a critical review of the primary literature for each gene. DNA sequences coding for each sperm protein were then successfully aligned and each manually inspected using the amino acid alignment as a template. The expected numbers of substitutions per site at synonymous sites (K_s) and nonsynonymous sites (K_a) were calculated using Li's method (1993). For calculations of K_s , doublet substitutions were removed to avoid neighboring effects (substitutions between adjacent codon positions). Previous studies in mammals have found the correlation between K_s and K_a to be influenced by doublet substitutions (Duret and Mouchiroud 2000); thus, removing such substitutions reduces bias in the synonymous substitution rate because of mutations at nonsynonymous sites. A list of sperm-specific human-mouse

orthologous genes with accession numbers and values of K_a and K_s are found in table 1.

Estimates of K_a and K_s for other tissue-specific human-mouse orthologs were kindly provided by Laurent Duret at the Université Claude Bernard, Villeurbanne, France (also calculated using the method of Li [1993] with doublet substitutions removed). Only genes expressed exclusively in a single adult tissue were selected for analysis, and in the testis-specific data set, genes expressed in the sperm were excluded. Using a random sample of genes, calculated divergence estimates were equal to estimates of K_a and K_s from the original data set. A Student's *t*-test with a Bonferroni correction for multiple tests compared the estimates of K_a and K_s between sperm-specific and other tissue-specific proteins.

Sperm proteins with more than two orthologs and with characterized function were further tested for evidence of positive selection. A maximum likelihood approach was implemented using the program codeml (Yang 1997), which uses a codon substitution model of evolutionary change. This method detects positive selection at the level of the codon (Yang 1994) and uses a likelihood ratio test to test various models of selection against a neutral model. Orthologous sperm-specific sequences from different mammalian species were extract-

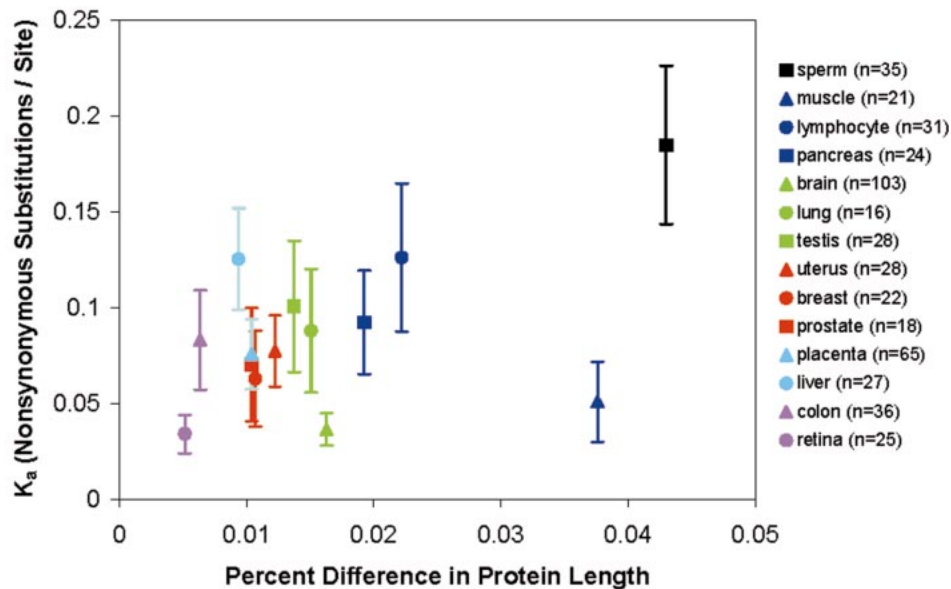


FIG. 1.—Comparison of mean protein divergence among various tissue-specific orthologs from mouse and human. Mean K_a , nonsynonymous substitution rate, and the mean length difference between orthologous coding regions were calculated. Error bars represent twice the standard error of mean. Sperm-specific genes have a significantly higher K_a than do genes expressed in most other tissue types, excluding the liver and lymphocyte, where lower values of K_a are found but are not significantly lower. Sperm-specific proteins also have the highest protein-coding length differences, indicating the occurrence of a larger number of insertion and deletion events during divergence than with other tissue types.

ed from GenBank. DNA sequences producing significant matches ($E < 10^{-6}$) were then aligned with ClustalX, Version 1.81 (Thompson et al. 1997) and neighbor-joining trees (Saitou and Nei 1987) drawn. Sequences were bootstrapped 1,000 times, and a consensus tree was drawn and examined manually to select orthologous genes and exclude paralogs. A bootstrap level of 90% was used as a cut-off value for a significant node. Bayes theorem was also used to calculate the posterior probability that each codon belongs to a certain class of ω (where $\omega = d_N/d_S$). Codons that identify with classes of $\omega > 1$ are purported to be under positive selection. Because of the high divergence of many

genes, we confirmed tests of positive selection by using a more conservative DNA sequence alignment by removing the immediate regions flanking indels until all sequences revealed identical nucleotides at one codon.

Results and Discussion

Sperm-specific genes show a significantly higher nonsynonymous substitution rate (K_a) than do other tissue-specific genes combined [mean K_a (sperm) = 0.18 ($N = 35$) versus mean K_a (nonsperm) = 0.073 ($N = 473$); $P < 0.001$]. Similarly, if we examine each tissue type separately, we find a significantly higher K_a in sperm-specific genes than in genes expressed in most of the other tissue types (excluding the liver and lymphocyte, where lower values of K_a are found but are not significantly lower; fig. 1). Using a second axis of divergence, we find that sperm-specific proteins have evolved the largest change in protein size, indicating the occurrence of a larger number of indel events compared with genes from other tissues (fig. 1). But the correlation of higher rates of amino acid substitutions to greater differences in coding region size is weak and may not represent a common pattern among mammalian genes ($R^2 = 0.23$, $P = 0.082$). K_a has been correlated previously to K_s (Wolfe and Sharp 1993; Mouchiroud, Gautier, and Bernardi 1995; Makalowski and Boguski 1998; Duret and Mouchiroud 2000), but a higher synonymous substitution rate in sperm-expressed genes is not correlated to the observed higher rate of amino acid substitution because there is no difference in K_s between sperm- and other tissue-specific genes [mean K_s (sperm) = 0.45 ($N = 35$) versus mean K_s (nonsperm) = 0.41 ($N = 473$); $P = 0.123$]. We also find that values of K_s do not vary between tissue types (fig. 2), suggesting that

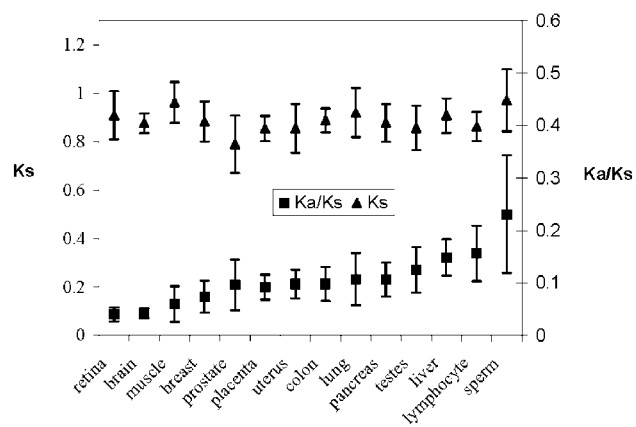


FIG. 2.—Mean synonymous substitution rate (K_s) and mean ratio of nonsynonymous to synonymous substitution rates (K_a/K_s) for various tissue-specific orthologs in mouse and human. Error bars represent twice the standard error of mean. Mean values of K_s do not significantly vary between tissue types, but sperm-specific genes have a significantly higher mean ratio of K_a/K_s than do proteins from five other tissues.

Table 2
Likelihood Ratio Test of Positive Selection in Sperm Proteins with Different Functions

FUNCTIONAL CLASS	PROTEIN	N	L _c	d _N /d _S	2Δℓ		PARAMETER ESTIMATES UNDER M8: P ₁ (ω > 1), ω
					M3 versus M0	M8 versus M7	
A. Egg binding							
ADAM2	4	728	0.34	128.6**	15.9**	P ₁ = 0.04, ω = 6.85
ADAM18	3	718	0.34	15.3**	3.7	P ₁ = 0.02, ω = 7.97
ACRV1	4	251	0.44	27.2**	0.2	P ₁ = 0.04, ω = 2.20
SPAG17	5	147	0.59	13.0*	5.7	P ₁ = 0.10, ω = 3.69
SAM1	4	508	0.53	98.9**	24.3**	P ₁ = 0.10, ω = 4.20
ZBPB	3	349	0.23	20.6**	0.04	P ₁ = 0.31, ω = 0.76
B. Transcription and translation							
NASP	3	666	0.38	68.0**	5.7	P ₁ = 0.40, ω = 1.29
PRM1	7	50	1.14	16.6**	12.7**	P ₁ = 0.28, ω = 6.50
PRM2	5	94	0.25	7.7	1.0	P ₁ = 0.29, ω = 1.07
PRM3	3	97	0.14	1.3	0.7	P ₁ = 0.05, ω = 1.03
TP1	5	55	0.18	6.4	1.1	P ₁ = 0.34, ω = 0.58
TP2	8	108	0.31	21.1**	5.0	P ₁ = 0.21, ω = 1.59
C. Motility							
GAPDS	3	408	0.16	79.4**	31.6**	P ₁ = 0.08, ω = 4.04
ODF1	4	591	0.01	9.4	5.6	P ₁ = 0.10, ω = 0.15
ODF2	4	591	0.02	2.9	1.6	P ₁ = 0.16, ω = 0.17
PRKA1	3	852	0.36	52.4**	5.1	P ₁ = 0.18, ω = 1.77
PRKA3	3	831	0.21	107.6**	3.1	P ₁ = 0.45, ω = 0.59
PRKA4	4	834	0.22	36.8**	4.7	P ₁ = 0.02, ω = 2.63
PRKAcAMP	3	143	0.03	1.8	0.7	P ₁ = 0.22, ω = 0.17

NOTE.—N, number of sequences; L_c, number of codons after alignment gaps were removed; d_N/d_S, average ratio calculated under M0; P₁, proportion of sites under positive selection; *, significant at 5% level; **, significant at 1% level.

different selective forces and not differential mutation rates explain the differences in K_a between tissues.

In the immune system, high rates of nonsynonymous nucleotide substitution have been found in the major histocompatibility complex (Hughes, Ota, and Nei 1990) and immunoglobulins (Tanaka and Nei 1989; Hughes 1997), both of which are directly involved in the identification of foreign antigens. The selective forces promoting amino acid diversity in the immune system may be coevolving with foreign or toxic substances. To maintain their ability to recognize antigens, lymphocyte proteins may also experience a high rate of amino acid substitution. Similar mechanisms have also been proposed to explain the rapid evolution of sperm proteins (Swanson and Vacquier 1998; Swanson, Aquadro, and Vacquier 2001; Swanson et al. 2001b; Swanson and Vacquier 2002). Sperms must recognize and interact with proteins on the oocyte surface for successful fertilization, and such proteins in the female may also be subjected to rapid evolutionary change. For example, in the marine invertebrate, abalone, sperm lysin is believed to evolve rapidly by selective pressure caused by the rapid concerted evolution of the egg receptor (Swanson and Vacquier 1998). In mating systems such as those found in many aquatic organisms, where sperms are released freely to seek out female oocytes, a selective drive toward specificity may be expected because of the sperm's increased chance of encountering foreign material or unsuitable oocytes. But in mammalian systems, complex behavioral and ecological premating barriers exist, so we might not expect sperm-egg coevolution to drive nonsynonymous substitutions to the extent seen in

free-spawning fertilization systems. But positive selection is often a primary force acting on male reproductive genes in humans and primates (Wyckoff, Wang, and Wu 2000) and appears to be driving the evolution of several mammalian egg surface proteins (Swanson et al. 2001b). Because sperm genes, in general, have a higher proportion of replacement substitutions than do proteins from other tissues, positive selection may be acting on a wider array of mammalian sperm genes other than those involved in primary sperm-egg interactions. We therefore tested for signatures of positive selection against sperm genes with characterized function to examine precise mechanisms of rapid evolution and to determine the extent of positive selection on sperm genes.

The comparison of the rate of nonsynonymous substitutions with the rate of synonymous substitutions has commonly been used to measure the selective pressures on a gene (i.e., K_a/K_s). If the rate of amino acid replacement is higher than the rate of synonymous changes, the gene is typically thought to have evolved under positive selection. But there is no clear cutoff between neutral evolution, K_a/K_s = 1, and positive selection, K_a/K_s > 1. For example, a ratio of K_a/K_s = 1.1 may be interpreted as either positive or neutral selection. In our human and mouse comparison of orthologous genes, most of the values of K_a/K_s are less than 1, providing no indication of positive selection driving sperm protein evolution but rather a relaxed selective constraint. Yet on average, sperm-specific genes have a significantly higher ratio of K_a/K_s in sperm proteins than in genes from five other tissues (fig. 2), and there is a twofold increase in the average K_a/K_s of sperm-specific genes

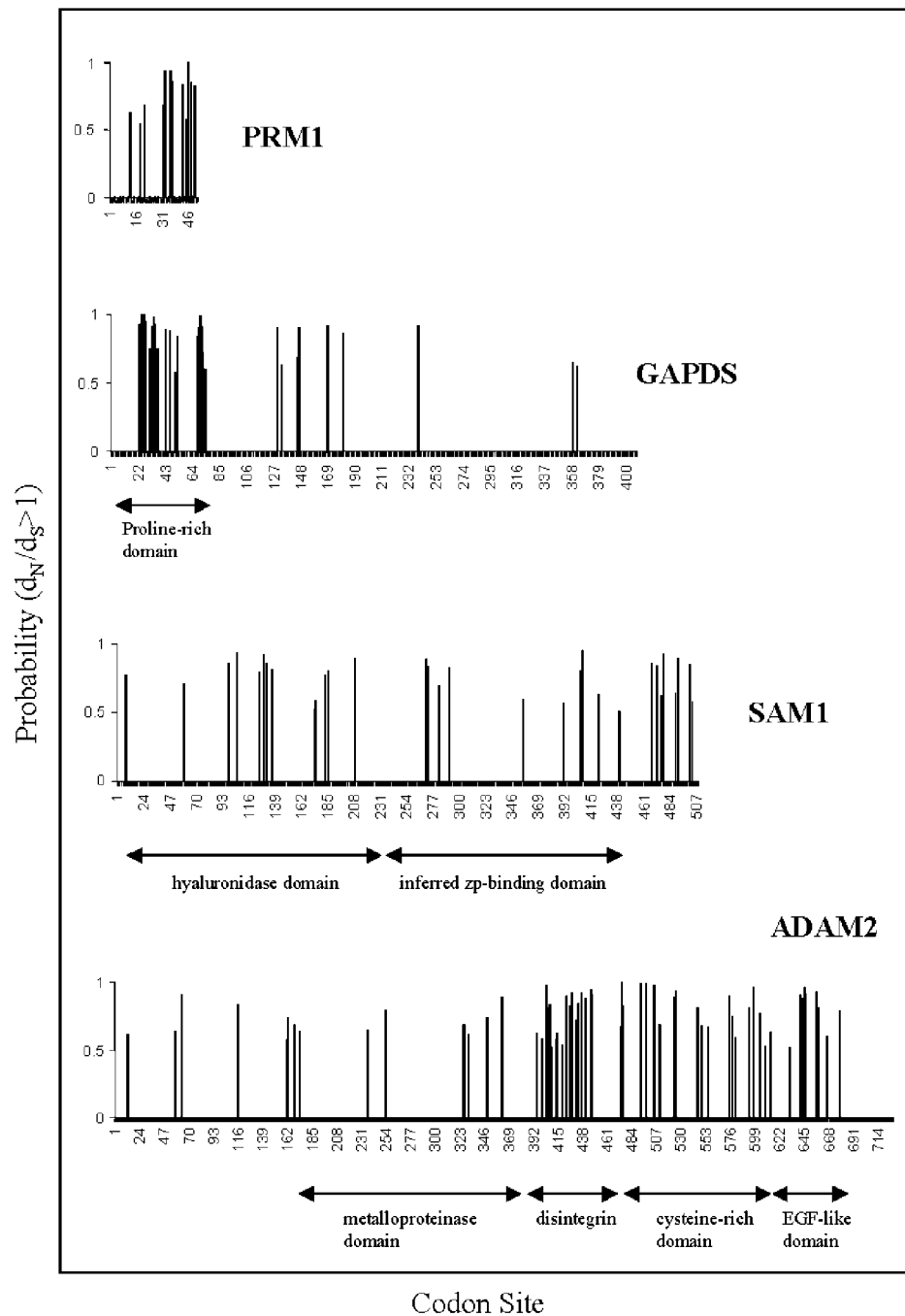


FIG. 3.—Location of codons that are likely to be under positive selection in four sperm-specific proteins in mammals. Codon numbering is based on the aligned human and mouse sequences with indels removed, and the alignments are available by request from D.G.T. In SAM1 and ADAM2, positively selected sites coincide with some regions hypothesized to interact with egg proteins. In GAPDS the first 72 amino acids are unique to the sperm-specific isoform where the majority of codons under positive selection are located.

when compared with all tissue-specific genes [mean K_a/K_s (sperm) = 0.50 ($N = 35$) versus mean K_a/K_s (non-sperm) = 0.19 ($N = 473$); $P = 0.001$]. Although the data may suggest prematurely that sperm genes have lower selective constraints than do other tissue-specific genes, ratios of K_a/K_s are generally not powerful predictors of positive selection because they represent an average value over all codons and cannot identify positively selected amino acid sites. A maximum likelihood

approach has been more successful in detecting positively selected sites within a gene, even when the average K_a/K_s over the entire gene is less than 1 (Yang 1994; Swanson et al. 2001b). This method is preferred over average measures of K_a/K_s in detecting positive selection because it can test whether similar selective forces are acting over the entire length of a gene.

Using this approach, sperm-specific genes with known functions were tested for the presence of posi-

tively selected codon sites (table 2). The first test compares an evolutionary model that estimates a single class of the parameter, $\omega = d_N/d_S$, with the constraint, $0 < \omega < 1$ (M0), to a model that allows for three site classes, including one that estimates ω greater than unity (M3). A total of 13 out of 19 sperm-specific genes show a significantly better fit to model M3 (table 2), suggesting heterogeneity in ω across most of the sperm-specific genes. A second, more conservative test compared models that assume a beta distribution on the parameter ω : one model, where $0 < \omega < 1$ (M7), is compared with a model that contains the additional site class, $\omega > 1$ (M8). Four out of 19 sperm-specific genes show a significantly better fit to model M8 over M7 with relatively large values of ω (table 2), indicating that these genes contain positively selected amino acids. Surprisingly, these four sperm genes represent diverse functional classes, including a protein involved in chromatin condensation, protamine-1 (PRM1), a protein involved in glycolysis, sperm-specific glyceraldehyde-3-phosphate dehydrogenase (GAPDS), and two proteins involved in sperm-egg binding, Adam 2 precursor (ADAM2) and sperm adhesion molecule 1 (SAM1).

Further analyses of these four genes reveal many interesting aspects of the selective process in mammalian sperm evolution. Codons were identified in each gene that are likely, under a Bayesian framework, to be under positive selection (fig. 3). Many of the positively selected sites in the two egg-binding proteins, ADAM2 and SAM1, lie in the regions of the protein believed to interact with the female egg. In ADAM2, sperm disintegrins are thought to interact with egg integrins allowing for secondary binding, suggesting that coevolution with egg integrins may drive positive selection in this region. In SAM1 most of the protein may interact with the egg, coinciding with a more even dispersal of positively selected sites: the hyaluronidase domain digests the hyaluronic acid of the egg, whereas the inferred zona pellucida-binding domain may be involved in egg recognition. Our findings also demonstrate that positive selection resulting from intersexual forces does not always present itself in the most obvious way, such as in sperm-egg binding. PRM1 replaces transitional protein-2 (TP2) in the sperm head for condensation of the chromatin during spermatogenesis and, at first glance, does not appear to directly interact with egg proteins. Once fertilization takes place, PRM1 interacts with an acidic amino acid motif of the β subunit of casein kinase II (Ohtsuki et al. 1996), an important regulatory protein found in the fertilized egg responsible for cellular metabolic alteration. Also, chromatin condensation may be important in determining the shape of the sperm head (Curry and Watson 1995), which may affect the ability of the spermatid to fertilize the egg. On the other hand, positive selection in sperm proteins may not solely be male-female coevolutionary processes. The enzyme GAPDS is not known to interact directly with female proteins but is likely to have an essential role in regulating energy production for motility (Welch et al. 2000). A distinct clustering of positively selected sites in GAPDS occurs in the proline-rich 72 amino acid segment that

does not have a homologous region in the somatic GAPD, suggesting that this region may have a unique and adaptive function in sperms. Proline-rich regions may mediate protein-protein interactions (Williamson 1994), and in sperms this region is hypothesized to anchor GAPDS to other glycolytic enzymes or to the fibrous sheath for efficient ATP diffusion to motor proteins of the flagellum (Welch et al. 2000). Positive selection in GAPDS may therefore be driven by sperm competition for energy production in fertilization or by coevolution with other sperm proteins rather than with female proteins as suggested by ADAM2, SAM1, and PRM1. Overall, the identification of positive selection in such a wide array of spermatozoan functional classes demonstrates the highly cryptic nature of sexual selection and emphasizes the far-reaching effects of sexual selection on a variety of molecular mechanisms.

Selection on sexual traits may be an important and ubiquitous driver of evolutionary change (Carson 1997; Singh and Kulathinal 2000). Ever since Darwin (1871) first attempted to explain the often-extreme sexual dimorphism between males and females, sexual selection has been proposed to play a major role in the generation of phenotypic diversity by speciation (Lande 1981; Carson 1997). Recently, the focus of sexually selected targets has broadened from male secondary sexual characters to a whole range of traits involved in sex and reproduction. In particular, sperm morphology has become an excellent example of a highly diverged male structure (Eberhard 1985; Pitnick 1996), and sexual selection has been suggested to drive its evolution through such processes as sperm competition (Karr and Pitnick 1999), male-female coevolution (Swanson et al. 2001*b*), and sexual conflict (Rice 1996). Sexual selection has been recently implicated at the molecular level on *Drosophila* accessory gland proteins used to assist sperm in successful fertilization (Cirera and Aguadé 1997; Tsaur, Ting, and Wu 1998) and *de novo* genes expressed exclusively in the *Drosophila* testis (Nurminsky et al. 1998). In mammals, female egg proteins used to bind sperm have been recently demonstrated to evolve through positive selection (Swanson et al. 2001*b*), suggesting a molecular coevolutionary arms race. Our study demonstrates positive Darwinian selection on various mammalian sperm components and suggests that sexual selective mechanisms in the form of male-female coevolution, sperm competition, and sexual conflict may have played a critical role in the generation of phenotypic diversity in mammalian lineages.

Acknowledgments

We thank G. B. Golding and R. A. Morton for their invaluable suggestions and an anonymous reviewer for helpful comments on the manuscript. This work was supported by the Natural Sciences and Engineering Research Council of Canada through a graduate scholarship to D.G.T and a research grant to R.S.S.

LITERATURE CITED

- AGUADÉ, M., N. MIYASHITA, and C. H. LANGLEY. 1992. Polymorphism and divergence in the Mst26A male accessory gland gene region in *Drosophila*. *Genetics* **130**:755–770.
- ALTSCHUL, S. F., W. GISH, W. MILLER, E. W. MYERS, and D. J. LIPMAN. 1990. Basic local alignment search tool. *J. Mol. Biol.* **215**:403–410.
- CARSON, H. L. 1997. Sexual selection: a driver of genetic change in Hawaiian *Drosophila*. *J. Hered.* **88**:343–352.
- CASANE, D., S. BOISSINOT, B. H. J. CHANG, L. C. SHIMMIN, and W.-H. LI. 1997. Mutation pattern variation among regions of the primate genome. *J. Mol. Evol.* **45**:216–226.
- CIRERA, S., and M. AGUADÉ. 1997. Evolutionary history of the sex-peptide (Acp70A) gene region in *Drosophila melanogaster*. *Genetics* **147**:189–197.
- CIVETTA, A., and R. S. SINGH. 1995. High divergence of reproductive tract proteins and their association with postzygotic reproductive isolation in *Drosophila melanogaster* and *Drosophila virilis* group species. *J. Mol. Evol.* **41**:1085–1095.
- CLARK, A. G., M. AGUADE, T. PROUT, L. G. HARSHMAN, and C. H. LANGLEY. 1995. Variation in sperm displacement and its association with accessory gland protein loci in *Drosophila melanogaster*. *Genetics* **139**:189–201.
- COULTHART, M. B., and R. S. SINGH. 1988. Differing amounts of genetic polymorphism in testes and male accessory glands of *Drosophila melanogaster* and *Drosophila simulans*. *Mol. Biol. Evol.* **5**:182–191.
- CURRY, M. R., and P. F. WATSON. 1995. Sperm structure and function. Pp. 45–69 in J. G. GRUDZINSKAS and J. L. YOVOVICH, eds. *Gametes—the spermatozoon*. Cambridge University Press, Cambridge.
- DARWIN, C. 1871. *The descent of man, and selection in relation to sex*. John Murray, London.
- DURET, L., and D. MOUCHIROUD. 2000. Determinants of substitution rates in mammalian genes: expression pattern affects selection intensity but not mutation rate. *Mol. Biol. Evol.* **17**:68–74.
- EBERHARD, W. G. 1985. *Sexual selection and animal genitalia*. Harvard University Press, Cambridge.
- HIRSH, A. E., and H. B. FRASER. 2001. Protein dispensability and rate of evolution. *Nature* **411**:1046–1049.
- HUGHES, A. L. 1997. Rapid evolution of immunoglobulin superfamily C2 domains expressed in immune system cells. *Mol. Biol. Evol.* **14**:1–5.
- HUGHES, A. L., and M. NEI. 1988. Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. *Nature* **335**:167–170.
- HUGHES, A. L., T. OTA, and M. NEI. 1990. Evolutionary relationships of class II major-histocompatibility-complex genes in mammals. *Mol. Biol. Evol.* **7**:491–514.
- KARR, T. L., and S. PITNICK. 1999. Sperm competition: defining the rules of engagement. *Curr. Biol.* **9**:R787–R790.
- KUMA, K., N. IWABE, and T. MIYATA. 1995. Functional constraints against variations on molecules from the tissue level: slowly evolving brain-specific genes demonstrated by protein kinase and immunoglobulin supergene families. *Mol. Biol. Evol.* **12**:123–130.
- LANDE, R. 1981. Models of speciation by sexual selection on polygenic traits. *Proc. Natl. Acad. Sci. USA* **78**:3721–3725.
- LI, W. 1993. Unbiased estimation of the rates of synonymous and nonsynonymous substitution. *J. Mol. Evol.* **26**:96–99.
- MAKALOWSKI, W., and M. S. BOGUSKI. 1998. Evolutionary parameters of the transcribed mammalian genome: an analysis of 2,820 orthologous rodent and human sequences. *Proc. Natl. Acad. Sci. USA* **95**:9407–9412.
- MATASSI, G., P. M. SHARP, and C. GAUTIER. 1999. Chromosomal location effects on gene sequence evolution in mammals. *Curr. Biol.* **9**:786–791.
- METZ, E. C., and S. R. PALUMBI. 1996. Positive selection and sequence rearrangements generate extensive polymorphism in the gamete recognition protein binding. *Mol. Biol. Evol.* **13**:397–406.
- MOUCHIROUD, D., C. GAUTIER, and G. BERNARDI. 1995. Frequencies of synonymous substitutions in mammals are gene-specific and correlated with frequencies of nonsynonymous substitutions. *J. Mol. Evol.* **40**:107–113.
- NURMINSKY, D. I., M. V. NURMINSKAYA, D. DE AGUIAR, and D. L. HARTL. 1998. Selective sweep of a newly evolved sperm-specific gene in *Drosophila*. *Nature* **396**:572–575.
- OHTSUKI, K., Y. NISHIKAWA, H. SAITO, H. MUNAKATA, and T. KATO. 1996. DNA-binding sperm proteins with oligo-arginine clusters function as potent activators for egg CK-II. *FEBS Lett.* **378**:115–120.
- PITNICK, S. 1996. Investment in testes and the cost of making long sperm in *Drosophila*. *Am. Nat.* **148**:57–80.
- RICE, W. R. 1996. Sexually antagonistic male adaptation triggered by experimental arrest of female evolution. *Nature* **381**:232–234.
- ROONEY, A. P., and J. ZHANG. 1999. Rapid evolution of a primate sperm protein: relaxation of functional constraint or positive Darwinian selection? *Mol. Biol. Evol.* **16**:706–710.
- SAITOU, N., and M. NEI. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**:406–425.
- SINGH, R. S., and R. J. KULATHINAL. 2000. Sex gene pool evolution and speciation: a new paradigm. *Genes Genet. Syst.* **75**:119–130.
- SWANSON, W. J., C. F. AQUADRO, and V. D. VACQUIER. 2001. Polymorphism in abalone fertilization proteins is consistent with the neutral evolution of the egg's receptor for lysin (VERL) and positive Darwinian selection of sperm lysin. *Mol. Biol. Evol.* **18**:376–383.
- SWANSON, W. J., A. G. CLARK, H. M. WALDRIP-DAIL, M. F. WOLFNER, and C. F. AQUADRO. 2001a. Evolutionary EST analysis identifies rapidly evolving male reproductive proteins in *Drosophila*. *Proc. Natl. Acad. Sci. USA* **98**:7375–7379.
- SWANSON, W. J., and V. D. VACQUIER. 1995. Extraordinary divergence and positive Darwinian selection in a fusogenic protein coating the acrosomal process of abalone spermatozoa. *Proc. Natl. Acad. Sci. USA* **92**:4957–4961.
- . 1998. Concerted evolution in an egg receptor for a rapidly evolving abalone sperm protein. *Science* **281**:710–712.
- . 2002. The rapid evolution of reproductive proteins. *Nat. Rev. Genet.* **3**:137–144.
- SWANSON, W. J., Z. YANG, M. F. WOLFNER, and C. F. AQUADRO. 2001b. Positive Darwinian selection drives the evolution of several female reproductive proteins in mammals. *Proc. Natl. Acad. Sci. USA* **98**:2509–2514.
- TANAKA, T., and M. NEI. 1989. Positive Darwinian selection observed at the variable-region genes of immunoglobulins. *Mol. Biol. Evol.* **6**:447–459.
- THOMPSON, J. D., T. J. GIBSON, F. PLEWNIAK, F. JEANMOUGIN, and D. G. HIGGINS. 1997. The CLUSTALX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res.* **24**:4876–4883.
- TSUR, S., C. TING, and C. WU. 1998. Positive selection driving the evolution of a gene of male reproduction, Acp26Aa, of *Drosophila*: II. Divergence versus polymorphism. *Mol. Biol. Evol.* **15**:1040–1046.

- WELCH, J. E., P. L. BROWN, D. A. O'BRIEN, P. L. MAGYAR, D. O. BUNCH, C. MORI, and E. M. EDDY. 2000. Human glyceraldehyde 3-phosphate dehydrogenase-2 gene is expressed specifically in spermatogenic cells. *J. Androl.* **21**:328–338.
- WILLIAMS, E. J. B., and L. D. HURST. 2000. The proteins of linked genes evolve at similar rates. *Nature* **407**:900–903.
- WILLIAMSON, M. P. 1994. The structure and function of proline-rich regions in proteins. *Biochem. J.* **297**:249–260.
- WOLFE, K. H., and P. M. SHARP. 1993. Mammalian gene evolution; nucleotide sequence divergence between mouse and rat. *J. Mol. Evol.* **37**:441–456.
- WOLFE, K. H., P. M. SHARP, and W.-H. LI. 1989. Mutation rates differ among regions of the mammalian genome. *Nature* **337**:283–285.
- WYCKOFF, G. J., W. WANG, and C. WU. 2000. Rapid evolution of male reproductive genes in the descent of man. *Nature* **503**:304–309.
- YANG, Z. 1994. Estimating the pattern of nucleotide substitution. *J. Mol. Evol.* **39**:105–111.
- . 1997. PAML: a program package for phylogenetic analysis by maximum likelihood. *Comput. Appl. Biosci.* **13**:555–556.

ANTONY DEAN, reviewing editor

Accepted July 17, 2002