THE ROLE OF REPRODUCTIVE CONFLICTS IN GENETIC, PHENOTYPIC, AND SPECIES DIVERGENCE

by

Erica Athena Dianne Jeffery
B.Sc., University of Lethbridge, 2004

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

In the
Department of Biological Sciences

SIMON FRASER UNIVERSITY

Fall 2007

APPROVAL

Name: Erica Jeffery

Degree: Master of Science

Title of Thesis:

The role of reproductive conflicts in genetic, phenotypic, and species divergence

Examining Committee:

Chair: Dr. S. Bisgrove, Assistant Professor

Dr. B. Crespi, Professor, Senior Supervisor
Department of Biological Sciences, S.F.U.

Dr. F. Breden, Professor
Department of Biological Sciences, S.F.U.

Dr. A. Mooers, Associate Professor
Department of Biological Sciences, S.F.U.

Dr. M. Hart, Associate Professor
Department of Biological Sciences, S.F.U.
Public Examiner

---

30 November 2007
Date Approved

ABSTRACT

Connecting natural selection of phenotypes with molecular evolution is one of the central goals of evolutionary biology. Using phylogenetic methods, I tested the hypothesis that reproductive conflicts related to sperm competition drive the adaptive molecular evolution of primate reproductive proteins. To control for potential empirical or statistical biases in the data, I compared results from 22 'housekeeping' proteins to those of 28 reproductive proteins. Average correlation coefficients between sperm competition and adaptive molecular evolution were significantly greater for reproductive proteins than for control group proteins. Reproductive proteins implicated in seminal coagulation and sperm-egg interactions, including two female-expressed proteins, had particularly high correlation coefficients. These results suggest that inter- as well as intra-sexual reproductive conflicts generate adaptive divergence in reproductive proteins. The nature of molecular interactions may mean that reproductive conflicts between males and females at this level are particularly likely to lead to the reproductive isolation of allopatric populations.

Keywords: sexual conflict; sperm competition; reproductive proteins; primates; molecular evolution; reproductive isolation

Subject Terms: Sexual behaviour in animals; Agonistic behaviour in animals; Primates – Variation; Molecular evolution

ACKNOWLEDGEMENTS

Thank you to everyone who has helped me in my work. In particular, thank you to my supervisor, Dr. Bernard Crespi, and my committee members, Dr. Arne Mooers and Dr. Felix Breden, for the patience and guidance they have shown me through out my degree. Regarding the second chapter of this thesis, thank you to Kyle Summers for his help with PAML, and Andrew Meade for his help with *BayesTraits*, respectively. I am also very grateful to Nathaniel Wolfe for supplying sequences and alignments for several seminal proteins. Finally, I owe a debt of gratitude to my friends and colleagues in the FAB*-Lab and BERG research groups (SFU) for their encouragement and support. This work was supported by NSERC.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1: INTRODUCTION

Evolutionary biology attempts to connect changes at the species level with change occurring at the level of populations, and the individuals and genes that compose them. The methods used to undertake this goal are constantly undergoing development and paradigm shifts. An apt example is the study of the rapid evolution of reproductive proteins. Early demonstrations of rapid evolution in reproductive proteins relied on the use of polymorphism data, principally in the abalone gamete recognition protein, lysin (Lee and Vacquier, 1992), and in the accessory gland proteins found in *Drosophila* seminal fluid (Aguade et al., 1992). The introduction of methods that compare rates of nonsynonymous to synonymous nucleotide substitutions (ie., d_N versus d_S) was pivotal in that it allowed researchers to differentiate between rapid divergence due to neutral genetic drift, and rapid divergence due to selection (McDonald and Kreitman, 1991). The observation of patterns of sequence divergence in reproductive proteins that had previously only been observed in coevolved host-parasite proteins began a wave of research, first in marine invertebrates and *Drosophila* (Tsaur and Wu, 1997; Metz and Palumbi, 1996; Lee et al., 1995; Lee and Vacquier, 1992), then mammals (Swanson et al., 2003; Swanson et al., 2001; Rooney and Zhang, 1999), then plants, fungi, and prokaryotes (Clark et al., 2006; Swanson and Vacquier, 2002b).

Although hypotheses explaining this widespread phenomenon are plentiful, they have been, in general, difficult to test (Swanson and Vacquier, 2002a). The observation that primate seminal proteins semenogelin I and II appear to have undergone a greater number of selective sweeps in species with higher expected levels of sperm competition was thus opportune (Jensen-Seaman and Li, 2003; Kingan et al., 2003). Dorus et al. (2004) recognized that the development of branch-specific methods of estimating d_N/d_S ratios (Yang,

1998) allowed directed testing of the hypothesis that sperm competition drives the adaptive molecular evolution of these proteins. Since their finding that species-specific d_N/d_S estimates are positively correlated with female promiscuity in primates (Dorus et al., 2004), several authors have tested this hypothesis using other primate reproductive proteins (Hurle et al., 2007; Herlyn and Zischler, 2007; Hamm et al., 2007). My goal in the second chapter of this thesis was to build upon past work by exploring the generality of a positive correlation between female promiscuity and the adaptive evolution of reproductive proteins, relative to non-reproductive proteins. In doing so, I incorporated knowledge of the phylogenetic relationships between primate species, as well as protein function, to explain the patterns of evolution observed in these proteins.

Just as early observations of rapid protein evolution lead to hypotheses regarding the role of sexual selection, sexual conflict and species-recognition in adaptive molecular evolution, so too did they lead to suggestions of the importance of these phenomena in the evolution of reproductive isolation (Lee et al., 1995). If rapid evolution can be caused by reproductive competition and conflict, what implications might this have for the evolution of species? Thus, my goal in the third chapter of this thesis was to synthesize theoretical and empirical data regarding the influence of reproductive conflicts on gene flow between populations. By placing reproductive conflicts at different levels of biological organization within a common framework, I hope to encourage cross-fertilization of theory and methods between disciplines such that we are able to move together towards our common goal – a better understanding of the biological diversity that surrounds us.

CHAPTER 2: PHYLOGENETIC-COMPARATIVE ANALYSES LINKING THE ADAPTIVE MOLECULAR EVOLUTION OF PRIMATE REPRODUCTIVE PROTEINS TO SPERM COMPETITION

Abstract

Although many proteins involved in the insemination of females and the fertilization of their gametes appear to have undergone strong positive selection, very few studies have linked this adaptive molecular evolution with corresponding evidence of natural selection at the phenotypic level. In this study, I evaluate the hypothesis that sperm competition has had a widespread influence on the adaptive molecular evolution of primate reproductive proteins. To control for potential empirical or statistical biases in the data, I also analysed a group of highly-conserved, widely-expressed 'housekeeping' proteins. A total of 28 reproductive proteins and 22 control group proteins were included in this study. Using phylogenetic methods, I compared species-specific d_N/d_S to two measures of sperm competition: relative testes mass and female remating rate. After correcting for multiple comparisons, 9 reproductive proteins showed positive correlations between d_N/d_S and sperm competition measures. In contrast, there were no positive correlations among the control group proteins. Using meta-analytic methods, I standardized correlation coefficients, and weighted them as a function of sample size. Transformed correlation coefficients were significantly higher among reproductive protein comparisons than among control group comparisons. Reproductive proteins implicated in seminal coagulation and sperm-egg interactions, including two female-expressed proteins, had particularly high correlation coefficients. These results suggest that, despite the complexity of

evolutionary pressures acting upon this diverse group of proteins, elevated d_N/d_S in reproductive proteins are likely the mark of post-copulatory sexual selection and sexual conflict.

Introduction

Connecting natural selection of phenotypes with evolution at the molecular level is one of the central goals of evolutionary biology. An increasingly popular method for detecting evidence of past selection in protein-coding DNA is to calculate rates of nonsynonymous (d_N) versus synonymous (d_S) nucleotide change between orthologous coding sequences of closely related species. A d_N/d_S ratio > 1 indicates an excess of mutations altering the amino acid sequence, and is interpreted as selection for phenotypic change (i.e., positive selection). Positive selection has been purported via this method for many proteins involved in the insemination of females and fertilization of female gametes, collectively known as reproductive proteins (Swanson and Vacquier, 2002a). However, dN/dS alone is not a fully adequate indicator of positive selection without supporting evidence regarding selection on protein structure and or function (MacCullum and Hill, 2006).

Several mechanisms of positive selection have been proposed for reproductive proteins (Swanson and Vacquier, 2002a). Though none of the following mechanisms should be considered mutually exclusive, hypotheses can be divided into those that predict a correlation between female promiscuity and the adaptive evolution of reproductive proteins, and those that do not.

Mechanisms associated with female promiscuity

Sexual selection (sperm competition, cryptic female choice)

Post-mating, pre-fertilization competition between males over female gametes, i.e., sperm competition, may create strong selection for functional optimization of male reproductive proteins. Promiscuous and polyandrous

species are expected to have higher levels of sperm competition than polygynous and monogamous species, and may therefore have faster rates of reproductive protein evolution (Herlyn and Zischler, 2007; Dorus et al., 2004a).

Greater female promiscuity also increases female opportunity to exercise post-mating, i.e., cryptic, female choice. If female choice is constantly evolving, male phenotypes could undergo continuous positive selection. Addressing how and why female choice might change will thus strengthen this type of hypothesis. One possibility is that female-expressed reproductive proteins are more likely to be under relaxed selection than male-expressed reproductive proteins, especially under conditions of intense sperm competition (Swanson and Vacquier, 2002a). If female preferences are free to evolve in a relatively neutral manner, males may be subjected to constantly changing positive selection. However, the fact that many female-expressed proteins appear to have undergone positive selection themselves suggests that another mechanism is at work in these cases (Swanson et al., 2001).

Sexual conflict

Intersexual conflict over fertilization could cause selection on female choice to shift over time, also resulting in continuous female evolution. Sperm competition will select for traits that are beneficial to males, but potentially costly to females. Such intergenomic conflict could lead to ongoing antagonistic coevolution between female- and male-expressed reproductive proteins (Rice and Holland, 1997). Because increased sperm competition is expected to intensify postmating sexual conflict, this hypothesis also predicts that reproductive proteins will evolve more quickly if female promiscuity is high. However, unlike sperm competition, sexual conflict will influence the evolution of both male-expressed *and* female-expressed proteins.

Sexually transmitted pathogens

Pathogens that infect gametes and the reproductive tract could subject both male- and female-expressed proteins to positive selection. Antipathogenic adaptations may be particularly favored in female-expressed proteins that mediate fusion of sperm with the egg and travel along the female reproductive tract (Swanson and Vacquier, 2002a). Positive selection on these proteins would lead to corresponding positive selection on male-expressed proteins. More promiscuous primate species have higher white blood cells counts, perhaps due to a higher incidence of sexually transmitted infections (Nunn et al., 2000). Antagonistic coevolution with pathogens could therefore also contribute to a positive correlation between female promiscuity and reproductive protein evolution, particularly in reproductive proteins involved in host defense.

Mechanisms not associated with female promiscuity

Reinforcement and gene duplication

Selection for pre-zygotic reproductive isolation due to less fit hybrids, i.e., reinforcement, will favor divergence in proteins that mediate mating and fertilization (Dobzhansky, 1940). Reproductive proteins might thus evolve more quickly when closely related species are in sympatry than when they are in allopatry (Swanson and Vacquier, 2002a). However, once reproductive isolation is complete, selection for divergence will cease. Reinforcement alone may therefore not be adequate to explain the high frequency of positive selection that is observed in reproductive proteins. Similarly, duplication and subsequent specialization of reproductive genes could lead to a burst of adaptive evolution, but would likely be followed by purifying selection once protein function was optimized (Swanson and Vacquier, 2002a).

Testing hypotheses

A small number of studies have looked for associations between female promiscuity and positive selection, with varied results (Table 1). A total of four reproductive proteins have been analyzed: semenogelins I and II (SEMG1 and 2; Kingan et al., 2003; Dorus et al., 2004a, Herlyn and Zischler, 2007; Hurle et al., 2007), Zonadhesin (ZAN; Herlyn and Zischler, 2007), and PKDREJ (Hamm et al., 2007). With the exception of Hamm et al. (2007), these studies have relied on non-comparative methods. Such limitations call into question the robustness and generality of positive correlations between female promiscuity and adaptive evolution in reproductive proteins. I reanalyzed the above four proteins using phylogenetic-comparative method, and extended my analysis to include an additional 24 primate reproductive proteins, plus a control group of 22 widely-expressed cellular 'housekeeping' proteins. To my knowledge, this study is the first to address the role of female promiscuity in reproductive protein evolution on such a broad scale. Most importantly, the inclusion of a control group will increase the validity of my results by providing a standard against which to measure the effect of female promiscuity on reproductive protein evolution.

Methods

Reproductive proteins

A thorough search of the literature was made for any mention of seminal or gamete associated primate proteins. This search was updated regularly until July 2007 using the ISI Web of Knowledge – Web of Science online database (http://portal.isiknowledge.com.proxy.lib.sfu.ca/portal.cgi). Table 2 provides a complete list of the proteins included in this study, their expression, and known functions. Proteins were excluded from the analysis if there was evidence of substantial expression outside of the male or female reproductive tracts. Such proteins may have important functions unrelated to insemination or fertilization, and thus would be subject to selection pressures outside of the context of female remating. An exception to this criterion was diazepam binding inhibitor (DBI), due to the fact that it has an apparent function relating directly to sperm competition

(Kolmer et al., 1997). DBI is highly expressed in both late-spermatogenesis spermatids and mature spermatozoa, where it localizes to the mitochondria-enriched sperm midpiece. The volume of this midpiece is positively correlated with relative testes size in mammals, including primates (Anderson et al., 2005), which suggests that increased mitochondrial loading increases the competitive ability sperm. DBI fatty acid metabolism, as the primary energy source available to spermatozoa, may therefore also be an important factor in sperm motility (Kolmer et al., 1997). Additionally, DBI is an androgen-regulated, prostate-expressed protein that was highlighted in a recent analysis of proteins found in primate seminal fluid (Clark and Swanson, 2005).

Control group proteins

In order to test the hypothesis that post-copulatory sexual selection accelerates the evolution of reproductive proteins, it was important to compare my results to those from a control group of proteins. It is possible that adaptive molecular evolution in general is more rapid in promiscuous species for reasons unrelated to sperm competition. For example, in polygynous species small ratios of breeding males to breeding females may mean that less promiscuous species will tend to have lower effective population sizes than more promiscuous species. Although census adult sex ratios do not necessarily reflect breeding sex ratios, they do measure the number of reproductively mature individuals of each sex. Using available primate adult sex ratio data to calculate effective population size (van Noordwijk and van Schaik, 2004; Plavcan, 2004; van Schaik et al., 1999; Nunn, 1999; Dixson, 1998), there was a significant positive correlation between both effective population size and relative female remating rate (linear regression: $r^2 = 0.18$, $p < 0.0001$, $n = 108$) and effective population size and relative testes mass (linear regression: $r^2 = 0.078$, $p = 0.29$, $n = 61$). If effective population sizes do tend to be larger in more promiscuous species, genetic drift would occur at a lower rate in these species, allowing selection to operate more effectively.

Sampling error might also create a positive correlation between protein d_N/d_S estimates and sperm competition variables. For many primate proteins, there is a sequencing bias in favor of macaque and hominoid species commonly used in lab research. Repeated inclusion of the human-chimpanzee sister pair in my analyses was of particular concern: the common chimpanzee, *Pan troglodytes*, has one of the most promiscuous mating systems documented among primates (Dixson, 1998), whereas human males are expected to experience relatively low levels of sperm competition (Shackleford and Goetz, 2006). Because phylogenetic analyses stress the importance of differences between closely related species pairs, the contrast between rates of human and chimpanzee molecular evolution could exaggerate the apparent correlation between female promiscuity and rate of nucleotide substitution across primates (Harmon and Losos, 2005) – particularly if the number of substitutions in the chimpanzee coding sequence is inflated by publishing errors (Clark and Swanson, 2005). A control group comparison should indicate whether an observed effect is due to mating system-related selection pressures, or to other, confounding factors. My goal was to gauge the overall support for a correlation between female promiscuity and reproductive protein evolution, rather than to draw conclusions regarding mode of evolution for any individual protein.

The control group was drawn from a previously published list of mammalian housekeeping genes (Dorus et al., 2004b). Following the authors' criteria, these are widely-expressed genes with basic, conserved functions in cellular metabolism and protein synthesis. These genes have evolved at similar rates in both primate and rodent taxa (Dorus et al., 2004b), and there is no expectation that their evolution would be influenced by postcopulatory selection pressures.

Sequence analysis

All sequences were accessed online using GenBank at NCBI (www.ncbi.nlm.nih.gov/) – accession numbers are listed in Appendix 1. Coding sequences were aligned manually using Se-Al Sequence Alignment Editor

v2.0a11 (http://tree.bio.ed.ac.uk/; Rambaut, 2007). Any section of the sequence that appeared to vary due to an insertion or deletion mutation rather than a substitution mutation was omitted from the analysis. Portions of the sequence following premature stop codons, and regions for which more than one alignment was conceivable, were also removed.

Branch-specific d_N/d_S were estimated by the free-ratio maximum likelihood method using CODEML from the PAML package, v3.15 (http://abacus.gene.ucl.ac.uk/; Yang, 1998). Portions of the sequence that were missing for one or more of the species were excluded from the analysis in question (i.e., cleandata = 1). The equilibrium codon frequencies used in the codon substitution model were estimated from the average nucleotide frequencies at the three codon positions (CodonFreq = 2).

In several cases, PAML estimated exaggerated branch-specific dN/dS (373.7-∞) due to extremely low d_S estimates (0-0.0001). Sequences may have simply experienced too little time, and therefore too few mutations along these branches to allow for reliable dN/dS estimates (Dorus et al., 2004a). Branches with lower d_S estimates (0-0.0001) also had significantly lower d_N estimates than branches with higher d_S estimates (0.001-0.4), which suggests these branches are experiencing lower mutation rates at both nonsynonymous and synonymous sites ($t = -5.03$, $df = 481.7$, $p < 0.0001$). Rather than combining closely-related lineages to avoid the unreliable d_N/d_S estimates associated with short branches (Dorus et al., 2004a), I instead excluded terminal branches whose orginal d_S estimates had been less than 0.001. This threshold ($d_S = 0.001$) clearly differentiated the inflated d_N/d_S estimates from the more conservative d_N/d_S estimates (Figure 1). By excluding one of a pair of low d_S-sister species, a reasonable d_N/d_S could often be achieved for the remaining species. In this way I maximized the number of species that could be included in the analysis, and avoided averaging values across species.

I did not compare the likelihoods of branch-specific d_N/d_S models to single-ratio models to test for significant differences in d_N/d_S ratios between branches (Yang, 1998; Hamm et al., 2007). Demonstrating significant variation in d_N/d_S

estimates between branches could be an overly conservative criterion for linking protein evolution to specific selection pressures, given that several similar branches can mask variation between other branches. For example, if the majority of species included in the analysis have a high degree of female promiscuity, one might expect the majority of sequences to show similar high divergence rates. Although assigning branches to distinct d_N/d_S classes might improve this problem, internal branches must also be considered. Making such designations a priori would be difficult when little is known about ancestral phenotypes, such as is the case with mating systems. Furthermore, whether or not d_N/d_S estimates vary significantly between branches, significant correlations between d_N/d_S estimates and sperm competition variables suggests a close coupling between the two factors.

Comparison with mating system

I compared terminal-branch d_N/d_S with two indicators of sperm competition: 1) the number of sexual partners per peri-ovulatory period, i.e., female remating rate (Campbell, 2006; Singh et al., 2006; van Schaik et al., 1999; Dixson, 1998; Boinski, 1987); and 2) testes mass (Dixson and Anderson, 2004; Kappeler, 1997; Harcourt et al., 1995; Harcourt, 1991; Moller, 1988; Harcourt et al., 1981). In the case of several strepsirhine measurements (Kappeler, 1997), testis volume was converted to testes mass using the formula provided by Harcourt et al. (1995). Relative male body mass, a measure of sexual dimorphism, has been previously used to estimate sperm competition when data regarding female remating rates and/or testes mass are not available (Herlyn and Zischler, 2007). More promiscuous species will have reduced sexual dimorphism; however, so will more monogamous species. Although relative male body mass does tend to decrease as sperm competition increases in the dataset in question (Herlyn and Zischler, 2007), across the primate phylogeny there is no significant correlation between relative male body mass and either relative testes mass or female remating rate (linear regression: $r^2 = 0.0435$, $p = 0.0651$, $n = 79$;

$r^2 = 0.0176$, $p = 0.14$, $n = 125$). For this reason, relative male body mass was not used as an indicator of sperm competition in any of my analyses.

Both female remating rate and testes mass are positively correlated with adult male body mass in primates, although the correlation between male body mass and the former is relatively weak ($r^2 = 0.15$, $p < 0.0001$; $r^2 = 0.66$, $p < 0.0001$, respectively; calculated using natural logarithms). Residuals obtained by regression with male body mass for both variables were thus used in the following analyses. Testes mass measurements were regressed on paired adult male body mass measurements when possible (Dixson and Anderson, 2004; Moller, 1988; Harcourt et al., 1981). Sexual partner counts were regressed on the species average adult male body mass, using the largest sample size available (Thoren et al., 2006; Plavcan and van Schaik, 1997; Kappeler, 1991).

In many cases, too little is known about reproductive protein function to compare sequence variation to relevant phenotypic variation. An exception is CatSper1 (Cation Sperm Channel 1), which directly influences the motility of sperm (Carlson et al., 2003). I compared species-specific d_N/d_S estimates for this protein to both percentage of motile sperm (Moller, 1988) and sperm midpiece volume (a likely indicator of individual sperm motility; Anderson et al., 2005). These were *a priori* comparisons testing a distinct hypothesis, and therefore were not pooled with the rest of the comparisons for the purposes of meta-analysis, or for the Bonferroni α-level corrections for multiple comparisons.

Non-phylogenetic comparative methods, in which species values are assumed to be statistically independent, can overestimate degrees of freedom and increase the likelihood of observing false positives in cases of phylogenetic correlation (Felsenstein, 1985). For this reason, I accounted for the phylogenetic structure in my data using the *Continuous* model from *BayesTraits* (www.evolution.rdg.ac.uk; Pagel, 1999). *Continuous* is a generalized least squares (GLS) model that uses a matrix of expected covariances among species to control for phylogenetic non-independence (Pagel, 1997). The analysis returns results equivalent to those of an independent contrasts analysis (Pagel, 1997; Felsenstein, 1985), and, as with an independent contrasts analysis, requires a

minimum of four species. Therefore, any protein with fewer than four species-specific $d_S > 0.001$ was necessarily excluded from the analysis. Twenty-eight of 42 candidate reproductive proteins, and 22 of 95 candidate control group proteins were ultimately included in this study. *BayesTraits-Continuous* returns the variances and covariance of the compared variables, which were used to calculate correlation coefficients in Microsoft Excel. The ln-likelihoods models assuming a correlation and assuming no correlation were compared using the Likelihood Ratio Test (df = 1) to attribute a *p*-value to the correlation (Pagel, 1999). All phylogenetic trees used in the above analyses were taken from an unpublished primate supertree (R. Vos, personal communication).

I also compared my variables using species-level regressions, calculated using JMP statistical software. Previous studies comparing terminal branch-specific dN/dS estimates and mating system traits have relied on species-level analyses; analyzing my data this way thus allows for more direct comparison with previous studies. Non-phylogenetic comparisons can provide statistically valid results, insofar as one of three following assumptions is met: 1) the species belong to a star phylogeny, such that they are all equally unrelated to one another; 2) the species values are solely the result of adaptive radiation, uninfluenced by Brownian motion (Harvey and Rambaut, 2000); or 3) rapid divergence between species erases similarities due to descent. Thus, if the phylogeny is reasonably diverse, and/or the correlation between the variables in question is reasonably strong, non-phylogenetic comparisons can provide useful tests of evolutionary hypotheses (Ricklefs and Stark, 1996).

Continuous also allows one to estimate the extent to which both female remating rate and testes mass were individually correlated with phylogeny, and then to compare the likelihood of those estimates to the likelihood of no phylogenetic correlation ($\lambda = 0$). Both sperm competition variables showed significant correlations with phylogeny (residual female remating rate: $\chi^2 = 18.21$, $p = 1.98 \times 10^{-5}$; residual testes mass: $\chi^2 = 21.79$, $p = 3.03 \times 10^{-6}$), supporting the use of phylogenetic-comparative methods.

Most of my analyses, especially those in the control group, included multiple hominoid primate species (i.e., *Pongo pygmaeus*, *Gorilla gorilla*, *Homo sapiens*, *Pan troglodytes*, and *Pan paniscus*), as well as two macaque species, *Macaca mulatta* and *Macaca fascicularis*. I retested for phylogenetic correlations in both sperm competition variables using only those species listed above. In this subset of primate species, neither female remating rate, nor testes mass were significantly correlated with phylogeny (residual female remating rate: $\chi^2 = 1.87$, $p = 0.17$; residual testes mass: $\chi^2 = 0.50$, $p = 0.48$). Thus, species-level analyses that are principally limited to hominoids and macaques may estimate correlations more accurately than those whose datasets extend to the rest of the primate phylogeny.

Multiple comparisons

Given that the control group proteins should not be affected by variation in post-copulatory sexual selection, I expected to find no evidence of a correlation between female promiscuity and d_N/d_S estimates in that group. However, if the null hypothesis is true, false positives are expected to occur at a rate corresponding to the chosen α-level (usually $\alpha = 0.05$). Resampling from the null distribution will increase the rate at which these false positives occur. Much discussion has centered around the best way of managing Type I error in these cases without becoming vulnerable to Type II error (Verhoeven et al., 2005). Although often criticized as unsuitable and overly conservative (Benjamini and Hochberg, 1995), the classic Bonferroni correction ($\alpha' = 0.05/m$, where m = number of comparisons) was considered an appropriate solution in my case for several reasons:

1) I was specifically concerned with group-wise (also known as family-wise) error rates – of the ~20 comparisons in the control group I would expect to see at least one significant result; I then questioned whether I would observe similar p-values at a similar frequency among the reproductive proteins

2) I was concerned primarily with Type I rather than Type II error

3) The average absolute effect size for my study, across groups, was $|r| = 0.62$ (Std dev = 0.30, SEM = 0.042, n = 50) – well above the small to moderate (r = 0.10-0.30) effect sizes at which loss of power is considered a concern (Nakagawa, 2004)

The reproductive proteins group and the control group were being compared to one another as separate populations, so the Bonferonni adjustment was conducted separately for each group. As relative testes mass and female remating rate are positively correlated with one another ($r^2 = 0.25$, $p < 0.0001$, n = 67), their comparisons with d_N/d_S estimates were considered redundant rather than independent comparisons. Similarly, the natural logarithms and untransformed d_N/d_S estimates could not be considered independent from one another. For this reason, only one comparison per protein was counted towards *m*. For the reproductive proteins, $\alpha' = 0.0018$ (*m* = 28); for the control group proteins, $\alpha' = 0.0023$ (*m* = 22).

To test for differences in the frequency of positive and negative correlations between the control group and reproductive proteins, I counted the number of protein comparisons that fell into the following groups: 1) no significant correlation; 2) positive correlation, $p < 0.05$; 3) positive correlation, $p < \alpha'$; 4) negative correlation, $p < 0.05$; 5) negative correlation, $p < \alpha'$. Each protein was counted only once: proteins with at least one correlation having a p-value less than α' were counted in categories 4) and 5); proteins with p-values ranging from 0.05-α' were counted in categories 2) and 3). A Pearson's ChiSquare test was used to determine whether or not the differences in counts between control group and reproductive proteins were significant.

Meta-analysis

Correlation coefficients were averaged by first using Fisher's Z-transformation to normalize the values (Z_r), weighting their average by the

inverse of the corresponding variances (wZ_r), and then back-converting to Pearson's r to give r_z (Corey et al., 1998). The p-value associated with r_z can be calculated from the average effect size and its standard error using a Z-test, which is on the same order of magnitude as the combined p-value calculated using the unweighted Z-method. The unweighted Z-method was appropriate for my data because the true level of replication was the number of proteins in each category, rather than the number of species in each comparison (Whitlock, 2005).

Average correlation coefficients (r_z) were calculated for both the control group and the reproductive protein group, as well as for each of five functional sub-groups within the reproductive protein group: sperm-egg interactions, seminal coagulation, sperm motility, spermatogenesis, and dissolution of seminal coagulum/host defense. Correlation coefficients in their weighted, standardized form (wZ_r) were compared between reproductive proteins and control group proteins using independent t-tests. Calculations were performed using a combination of Microsoft Excel and JMP statistical software.

Results

d_N/d_S estimates

Overall, d_N/d_S estimates were significantly higher among the reproductive proteins than among the control group proteins, as were both d_N estimates and d_S estimates individually (Figures 2 a,b,c). Figures 3a-d show the distribution of d_N/d_S estimates in relation to sperm competition-related traits.

Species-level analyses

In general, the species-level analyses offer weak, although suggestive, support in favor of a positive correlation between sperm competition variables and reproductive protein d_N/d_S estimates. Tables 3a-b show correlation coefficients and associated p-values from these analyses. None of the species-

level comparisons were significant at the Bonferroni-adjusted α-levels (reproductive proteins: $\alpha' = 0.0018$; control group proteins: $\alpha' = 0.0023$). At the 0.05 α-level, four reproductive proteins showed at least one positive significant correlation, whereas there were no significant negative correlations in this group. Among the control group proteins, there were two significant positive correlations, and two significant negative correlations. The difference in the frequency of positive and negative correlations between the two groups was not significant ($\chi^2 = 2.67$, df = 1, p = 0.10).

Averaged across reproductive protein comparisons, correlations between d_N/d_S estimates and female remating rate were positive, and significant (Table 5). Average correlations (r_z) between d_N/d_S estimates and relative testes mass were generally weaker, and nonsignificant. None of the averages calculated for control group comparisons were significant. Standardized, weighted correlation coefficients (wZ_r) were significantly larger among reproductive proteins than among control group proteins when d_N/d_S estimates were compared to female remating rate (Table 7b). The difference was not significant when d_N/d_S estimates were compared to relative testes mass.

Average r_z and associated significance were particularly high among proteins involved in sperm-egg interactions, and those involved in seminal coagulation. In the phylogenetic-comparative analyses detailed below, the sperm-egg interaction proteins OGP, PKDREJ, ZAN and ZP-4, and the seminal coagulation proteins SEMG1 and SEMG2 showed particularly strong correlations between d_N/d_S estimates and sperm competition variables. For this reason, I present species-level plots for these proteins in Figures 4 a-g. A species-level plot is also included for control group protein GSTM4; GSTM4 is of particular interest because results from the phylogenetic-comparative analyses suggest it may be a potential outlier.

Phylogenetic-comparative analyses

Results from my phylogenetic-comparative analyses strongly support a positive correlation between reproductive protein d_N/d_S ratios and sperm competition. Of the 28 reproductive proteins I analyzed, 12 showed positive correlations between d_N/d_S and sperm competition estimates, nine of which were significant at the Bonferroni-corrected α-level ($\alpha' = 0.0018$). Four reproductive proteins showed negative correlations between the dN/dS estimates and sperm competition variables, none of which were significant after the correction for multiple comparisons. In the case of the 22 control group proteins, four showed positive correlations and five showed negative correlations, although only two of the negative correlations were considered significant at the Bonferroni-corrected α-level ($\alpha' = 0.0023$) (Table 4a,b). The difference in the frequency of positive and negative correlations between the two groups was significant ($\chi^2 = 10.76$, df = 4, p = 0.029).

Standardized, weighted correlation coefficients (wZ_r) were significantly larger among reproductive proteins than among control group proteins in all cases (Table 7a, Figure 6a). Average correlations (r_z) between reproductive protein d_N/d_S estimates and sperm competition variables were positive and significant, whereas r_z from the control protein comparisons were weak and insignificant (Table 6). The average r_z of reproductive proteins involved in sperm-egg interactions, seminal coagulation, or sperm motility were generally significant, and positive. In contrast, the average r_z of reproductive proteins involved in spermatogenesis, the dissolution of seminal coagulum, or host defense were non-significant in all cases (Table 6). Despite this contrast, the wZ_r of the separate functional groups did not differ significantly from one another (one-way ANOVA, F = 0.99, p = 0.45).

I predicted that the reproductive and control protein groups would differ in the frequency and magnitude of positive correlations, but not in their distributions of negative correlations. To test this prediction, I compared first positive wZ_r, and then negative wZ_r between the two groups. In order to summarize wZ_r across all

four possible comparisons (i.e., d_N/d_S compared to both sperm competition variables, and $\ln(d_N/d_S)$ compared to both variables), I used results from only one comparison per protein – the comparison associated with the lowest p-value, whether or not $p < 0.05$. When positive wZ_r were compared, values were significantly higher among reproductive protein comparisons than among control group comparisons (Figure 6b). In contrast, when negative wZ_r were compared, there was no difference between the two groups (Figure 6c).

The null hypothesis for the Likelihood Ratio Test-statistic (LRT-statistic) follows an approximate Chi-square distribution, with degrees of freedom equal to the difference in number of parameters between the compared models. Thus, if results from either the reproductive proteins or the control group were truly representative of the null hypothesis, the LRT-statistics for that group would be expected to follow to a Chi-square distribution with df = 1. A Cramer-von Mises W Test for goodness of fit found that both the control group and reproductive protein group LRT-statistic distributions were significantly different from the expected null distribution ($W^2 = 0.53$, $p = 0.036$; $W^2 = 2.13$, $p = 0.0010$).

Outliers are commonly defined as points that lie more than 1.5 times the interquartile distance beyond either the upper or lower quartiles (Frigge et al., 1989). The control group distribution had a single point, GSTM4, that lay beyond this outlier threshold (i.e., >18.17; Figure 3a). Excluding this potential outlier, the difference between the control group distribution and the null became non-significant ($W^2 = 0.42$, $p = 0.066$).

False positives

Phylogenetic-comparative analyses can exhibit inflated Type I error rates if the compared traits vary more within species than between (Harmon and Losos, 2005). Although relative testes mass and female promiscuity data were used based on the largest reported sample size, in some cases only one measurement was available. In general, intraspecific variation in primate testes mass will be lower than interspecific variation (Harcourt, 1997). However, certain pairs of species could show more within than between-species variation in

female promiscuity. Mating systems with dissimilar behavioural dynamics are capable of resulting in similar average female mating rates. For example, the white-handed gibbon, *Hylobates lar*, and the hamadryas baboon, *Papio hamadryas*, both have average female mating rates of about 1.5 mates per peri-ovulatory period (van Schaik et al., 1999). Whereas gibbons are generally considered a monogamous genus, most baboons species are highly promiscuous. Hamadryas baboons are unique in that larger groups are composed of many smaller harems, each guarded by a dominant reproductive male, although sperm competition between dominant and subdominant males is not uncommon (Zinner et al., 2006). White-handed gibbons live in small groups of 3-5 individuals. Groups were previously assumed to comprise a heterosexual pair and their juvenile offspring, but in actuality vary frequently from this structure (Fuentes, 2000). Sampling error could potentially inflate the difference in female mating rates between these two species, further exaggerating the phylogenetic contrast, despite both species having similar levels of sperm competition. The exaggerated contrast might then generate a correlation where there is none (Harmon and Loso, 2005). Although measurement error is a concern for some species measurements, this bias should have affected both reproductive protein and control group analyses equally.

Loci under positive selection are expected to show higher rates of sequence variation between species than within; negative selection, on the other hand, should decrease inter-specific divergence relative to intra-specific polymorphisms (Bustamante et al., 2005; Sawyer and Hartl, 1992). Inflation of false positive rates due to sequence sampling error may therefore be more of an issue in my control group than in the case of my reproductive protein analyses. This may explain why the control group, despite generally agreeing with the null hypothesis, generated two significant correlations between sequence evolution and reproductive traits. Harmon and Losos (2005) simulated the effect of measurement error on Type I error rates in phylogenetic analyses, under different ratios of inter:intra-specific variation in traits. Two false positives out of 22 control group comparisons, or a Type I error rate of 9.1%, is comparable to

the 10% upper average error rate they reported (Harmon and Losos, 2005). Type I error rates became particularly inflated when pectinate (i.e., highly asymmetrical) trees, such as the one used for my analysis of GSTM4, were used in phylogenetic analysis. If interspecific variation is high relative to within species variation, measurement error has little effect on the results – even if only one sample is used per species (Harmon and Losos, 2005). In general, reproductive proteins would be expected to have higher between than within-species variation. Possible exceptions are BOULE and DAZL, which appear to have undergone stabilizing selection (Tung et al., 2006).

It is surprising that BOULE, a reproductive protein with no evidence of positive selection, would show strong positive correlations between d_N/d_S estimates and sperm competition variables (Table 4a), especially when DAZL, a homologous protein with a similar function, does not. Even more surprisingly, these correlations carry over into the results from the species-level analyses (Table 3a). It is possible that branch-specific analyses lack the power to detect inter-specific variation in positive selection when it is restricted to a small number of nucleotide sites, the locations of which may also vary between species. Alternatively, the positive relationship between female promiscuity and molecular divergence in BOULE may indicate stronger purifying selection in less promiscuous species and relaxed selection in more promiscuous species, although it not obvious why this would be the case.

False negatives

One could argue that my small sample sizes and resulting low statistical power have prevented us from detecting existing correlations, positive or otherwise, among the control group proteins. Larger sample sizes would, of course, be ideal – unfortunately, my study was limited to existing, publicly available sequences. Control group sample sizes ranged from 4-6 sequences per protein (Table 4b). Although low, five of the 9 most significant reproductive protein comparisons also had sample sizes within this range (Table 4a). I should be able to detect the same effect size in the control group with sample sizes of 4-

5 sequences per protein, committing Type II errors only 5% of the time (Faul et al., 2007). To test for the influence of sample size on my results, I compared standardized, weighted correlation coefficients, wZ_r, between control group and reproductive proteins, excluding reproductive proteins for which more than 5 species had been included in the analysis. wZ_r continued to be significantly larger among reproductive proteins, although only marginally so ($t = 1.79$, $df = 18.32$, one-tailed $p = 0.045$). Nonetheless, this result suggests that it is not discrepancy in sample sizes driving the differences between the two groups. Sampling from different sections of the primate phylogeny in the control group analyses versus the reproductive protein analyses could also have influenced the results of my analysis. However, although reproductive protein analyses occasionally included Strepsirrhine primates and New World monkeys, sequences for both protein groups were drawn principally from Old World monkeys and hominoids, with substantial overlap.

CatSper1 analyses – sperm midpiece volume and sperm motility

CatSper1 dN/dS estimates were only weakly correlated with female remating rate using phylogenetic-comparative methods ($r^2 = 0.28$, $p = 0.032$), and were not significant at the Bonferroni-corrected α-level. However, d_N/d_S estimates were positively correlated with both sperm midpiece volume ($r^2 = 0.58$, $p = 0.0052$) and percentage of motile sperm ($r^2 = 0.56$, $p = 0.011$).

Results from the corresponding species-level analyses also show a lack of definite correlation between d_N/d_S estimates and sperm competition variables (Figures 5a,b), but increasing trends when d_N/d_S estimates are compared to variables linked to sperm motility: sperm midpiece volume ($r^2 = 0.18$, $p = 0.22$; Figure 5c) and percent motile sperm ($r^2 = 0.63$, $p = 0.018$; Figure 5d).

Discussion

My main objective in this study was to test the prediction that proteins involved in the insemination of females and fertilization of female gametes will

show positive correlations between d_N/d_S estimates and sperm competition-related traits more often and of a greater magnitude than proteins with basic cellular functions not relating directly to sperm competition. Using phylogenetic-comparative methods, I observed significant differences in the frequency of positive and negative correlations between my sample of reproductive proteins and my sample of control group proteins. Whereas nine of the 28 reproductive proteins showed at least one positive correlation between d_N/d_S estimates and sperm competition variables (significant at the Bonferroni-corrected α-level), none of the 22 control group proteins showed positive correlations of equivalent significance (Table 4a,b).

Furthermore, correlations between d_N/d_S estimates and sperm competition variables were significantly stronger among reproductive proteins: standardized, weighted correlation coefficients (wZ_r) were significantly larger (i.e., more positive) when testes mass and female remating rate were compared to d_N/d_S estimates from reproductive proteins, than when they were compared to d_N/d_S estimates from control group proteins (Figure 6a). This difference was specifically due to differences in the magnitude of positive correlations (Figure 6b) – when only negative correlations were compared, the two groups did not differ (Figure 6c). These results clearly imply that sperm competition-correlated selection pressures are driving the adaptive evolution of many reproductive proteins.

The fact that the results from my species-level analyses were generally weaker than the results from my phylogenetic-comparative analyses may seem surprising. Often, phylogenetic dependency is expected to make variables appear more correlated rather than less correlated (Hurle et al., 2007). However, this is not always the case. Because accounting for the effects of phylogeny decreases the standard error of estimated regression coefficients, phylogenetic-comparative methods will tend to have higher power, and lower Type I error rates than non-phylogenetic-comparative methods (Rohlf, 2006). This will sometimes result in higher estimates of regression coefficients, as is the case with my analysis.

Given that female remating rate and testes mass are both strongly correlated with phylogeny, comparisons with these variables using phylogenetic methods are expected to give more accurate results than non-phylogenetic methods. The results of the species-level analyses are still informative in that they indicate the extent of *observable*, rather than *mechanistic*, correlations between variables. For example, although the results of my phylogenetic-comparative analyses suggest a true correlation between ZAN d_N/d_S estimates and sperm competition variables (Table 4a), one should not necessarily expect to observe a higher d_N/d_S estimate for this protein along the lineage of a more promiscuous species (Figure 4c).

d_N/d_S estimates

The rapid evolution of reproductive proteins is well established (Swanson and Vacquier, 2002b; Clark et al., 2006). Although I actively avoided restricting my data set to proteins that have shown evidence of positive selection, all but two of the reproductive proteins I analyzed appear to be rapidly evolving (Table 8). Given that positive selection has been demonstrated for the majority of these proteins, it is not surprising that branch-specific d_N/d_S estimates were significantly higher among the reproductive proteins than among the control group proteins (Figure 2a). Whereas d_N/d_S estimates in the control group never exceeded one, 21% of the d_N/d_S estimates in the reproductive protein group were greater than one. These higher d_N/d_S estimates were not due to decreased d_S estimates; rather d_S was significantly higher among the reproductive proteins than among the control group proteins (Figure 2c).

Control group results

The low d_N/d_S estimates of control group proteins suggest that, whereas negative correlations among the reproductive proteins would indicate a *decrease* in positive selection as female promiscuity increases, the same results in the control protein group would indicate an *increase* in negative, or purifying

selection, with increasing female promiscuity. The control group proteins were chosen for their highly conserved roles in basic cellular regulation. These proteins are expressed in a wide variety of cellular tissues, such that selection acting on a specific tissue type, i.e., reproductive tissues, should not unduly influence molecular evolution. However, it is possible that if males engage in intense sperm competition, maintaining proper cellular regulation in reproductive tissues may be particularly important in order to ensure male fertility. Therefore, if anything, one might expect to see negative correlations between d_N/d_S estimates and sperm competition among these proteins more often than positive correlations. Such a trend would exaggerate the apparent effect female promiscuity has on reproductive protein evolution relative to the control.

There was a general trend in the control group toward negative correlations between the rate of molecular evolution and female promiscuity. Two proteins in particular, GSTM4 and KNSL6, show strong negative correlations between d_N/d_S estimates and mating system variables (Table 4b). For GSTM4, the strength of the correlation is extreme enough for the protein to be considered an outlier (Figure 7a). A biological explanation for the observed negative correlations is more plausible in the case of KNSL6. Kinesin-Like 6 protein (KNSL6) is a microtubule depolymerase that corrects kinetochore-microtubule attachment errors (Huang et al., 2007). Although expressed in several tissues that contain rapidly dividing cells types, KNSL6 is particularly highly-expressed in the thymus and testes (Kim et al., 1997). It is possible that KNSL6 regulation of spindle assembly and chromosome segregation during mitosis and meiosis is under higher stabilizing selection in species with increased sperm production – such as in species with high levels of sperm competition.

Overall, however, the support for a negative correlation between d_N/d_S estimates and mating system in control group proteins is not strong. Comparisons among control group proteins consistently had weak average r_z of low significance, whose upper confidence intervals were above zero (Tables 5 and 6). Additionally, with the exclusion of GSMT4, the distribution of LRT-

statistics from this group was not significantly different from the expected null hypothesis distribution (Figure 7a).

Reproductive protein results - comparing hypotheses

As a group, the reproductive proteins deviated significantly from the expectations of the null hypothesis (Table 6, Figure 7b). This result was due in particular to proteins with strong positive correlations (i.e., those with p-values lower than the Bonferroni-corrected α-level, α') between d_N/d_S estimates and sperm competition variables. Surprisingly, four of the eight putative sperm-egg interaction proteins I analysed fell into this category: PKDREJ, ZAN, OGP, and ZP-4 (Table 4a). This pattern was striking; average r_z were consistently more significant for proteins involved in sperm-egg interactions than for any other functional category (Table 6). The same trend can be observed in the results from the species-level analyses, although the correlations are generally weaker (Table 5). Although there is general evidence for a positive relationship between sperm competition and reproductive protein evolution across functional categories, the strength of the correlation in the case of sperm-egg interaction proteins particularly supports intersexual conflict as a mechanism of adaptive molecular evolution.

Sperm-egg interaction proteins are unique in that male and female proteins interact directly with one another. Male-expressed and female-expressed proteins pairs are therefore expected to coevolve with one another. Sperm competition will make the fitness costs of not coevolving greater for males than for females (Swanson and Vacquier, 2002a) – that is, unless the male phenotypes favored by sperm competition are directly costly to females. For example, sperm competition increases the risk of polyspermy, i.e., ovum fertilization by multiple sperm. Male gametes will be more competitive if they bore through the egg matrix as quickly as possible; unfortunately, the race to fertilize may lead to multiple sperm fusing with the egg before blocks to polyspermy are implemented, increasing the chances of both pathogen infection and lethal polyploidy (Rice and Holland, 1997). The risk of polyspermy may be particularly

important in mammalian species, whose eggs have only slow blocks to polyspermy (Swanson and Vacquier, 2002a). In this situation, sequence changes that *decrease* sperm-egg binding efficiency will be favored in female-expressed egg-coat proteins, whereas changes that *increase* sperm-egg binding efficiency will be favored in the male-expressed counterparts. Because of direct interaction with one another, such protein pairs are more likely to undergo ongoing antagonistic coevolution than other male- and female-expressed reproductive proteins (Swanson and Vacquier, 2002a). The importance of intersexual conflict in accelerating the adaptive evolution of reproductive proteins may explain why strong positive correlations between d_N/d_S estimates and sperm competition were more rare in proteins from other functional categories (Table 4a).

Although this is not the first study to observe a positive correlation between sperm competition and molecular evolution in a putative male-expressed sperm-egg interaction protein (Herlyn and Zischler, 2007), this is the first to find evidence of such correlations in female-expressed reproductive proteins (OGP and ZP-4). Despite this evidence being suggestive of intersexual conflict, it is not conclusive – in the case of OGP, only five species were included in the analysis, whereas in the analysis of ZP-4 there were only four species (Table 4a). Furthermore, of the three zona pellucida proteins I analysed, ZP-4 was the only protein to show any evidence of a correlation, in spite of the fact that both ZP-2 and ZP-3 play roles in the induction of the acrosome reaction (Gahlay et al., 2002). Rigorous testing of the male-female antagonistic coevolution hypothesis for reproductive proteins requires including results from more sperm-egg interaction proteins across more species, as well as comparing the d_N/d_S estimates of interacting male- and female- expressed proteins. Other forces influencing the evolution of female-expressed reproductive proteins, such as genetic drift or selection for anti-microbial adaptations, might sometimes, but not always, affect sperm-receptor binding. These forces should thus tend to decouple the adaptive evolution of interacting male- and female-expressed proteins, whereas intersexual conflict should lead to positive correlations between the d_N/d_S estimates of interacting proteins. Such a comparison is

theoretically possible for putative zona pellucida-binding proteins PKDREJ and ZAN (Hamm et al., 2007; Lea et al., 2001). Unfortunately, there was not enough overlap between datasets to compare either PKDREJ or ZAN dN/dS estimates with dN/dS estimates from any of the zona pellucida glycoproteins in the current study.

CatSper1 results

My results are meant to be evaluated in conjunction with one another, rather than taken individually. Generalized across functional groups, and in comparison with my control group, I found strong support for the influence of sperm competition in the evolution of reproductive proteins. However, drawing conclusions regarding the mechanism of evolution for any particular protein requires specifying how molecular divergence relates to phenotypic divergence, and how phenotypic divergence relates to fitness. For example, after using species-level, linear regression to show a positive correlation between SEMG2 d_N/d_S estimates and female remating rate, Dorus et al. (2004a) were able to show an increasing trend between the same d_N/d_S estimates and seminal coagulation rankings (Dixson and Anderson, 2002). Unfortunately, for many reproductive proteins, the relationship between molecular variation and phenotypic variation is poorly characterized.

The sperm cation channel, CatSper1, is an exception. The protein is exclusively expressed in the membrane of developing spermatids, where it localizes to the sperm tail midpiece (Li et al., 2006). CatSper1 is necessary for proper sperm motility (Carlson et al., 2003), to the point that male mice lacking it are infertile (Ren et al., 2001). Using phylogenetic-comparative analyses, primate CatSper1 d_N/d_S estimates were positively correlated with both the percentage of sperm in male ejaculate that are motile, and with average sperm midpiece volume (μm^3). Results from species-level analyses showed similar patterns (Figures 5c,d). Because CatSper1 localizes to the sperm midpiece, this correlation may be influenced by physical constraints regarding protein conformation and orientation. However, sperm midpiece volume has also been

connected to flagellar function, and thus sperm motility. Mitochondria are confined to the midpiece in sperm, such that larger midpieces can carry more mitochondria, and result in more competitive sperm (Anderson et al., 2005). Sperm midpiece volume is positively correlated with both mating system and relative testes size in mammals (Anderson et al., 2005), suggesting that larger midpieces are in fact important in sperm competition.

Although both sperm motility and sperm midpiece volume are positively correlated with expected sperm competition (Moller, 1988; Anderson et al., 2005), correlations between CatSper1 d_N/d_S estimates and sperm competition variables were weaker than correlations between d_N/d_S estimates and sperm motility variables (Table 4a, Figures 5a,b). Imperfect correlations between sperm competition and phenotypic evolution, and phenotypic evolution and sequence evolution may have also obscured the relationship between d_N/d_S and sperm competition variables for other reproductive proteins in my analysis.

Comparison with previous studies

SEMG1, SEMG2 – the seminal coagulation proteins

Dorus et al. (2004a) were the first to show a positive correlation between terminal branch-specific d_N/d_S estimates and female promiscuity in their study of primate SEMG2. More recently, Hurle et al. (2007) extended the data set to include New World monkeys and strepsirrhines, in addition to hominoids and Old World monkeys, and failed to find a correlation. When the same authors analysed SEMG1, they found only a nonsignificant increasing trend between dN/dS estimates and female remating rate (Hurle et al., 2007).

In my analysis of SEMG2, I was able to include a wider range of Old World monkeys than Dorus and colleagues (2004a), although only one New World monkey and no strepsirrhine primates were included. The species-level analysis found no correlation between SEMG2 d_N/d_S estimates and sperm competition variables (Table 3a). However, the phylogenetic-comparative analysis found a strong positive correlation with female remating rate, and a

lesser correlation with relative testes mass (Table 4a). Similarly, I included two New World monkeys, and three Old World monkeys in my analysis of SEMG1, and found only a weak positive correlation with female remating rate using species-level, linear regressions (Table 3a), but strong correlations with both female remating rate and testes mass using phylogenetic-comparative methods (Table 4a). Whereas other branches of the primate phylogeny show a relative conservation of mating systems, divergence of mating systems between closely-related species is pronounced in the hominoids (Dixson, 1998). Although both female remating rate and testes mass are correlated with phylogeny in primates, the correlations are not significant when limited to hominoid species. Therefore, in analyses that extend to Old World monkeys, New World monkeys, and strepsirrhines, phylogenetic-comparative analyses are expected to give more accurate results (Felsenstein, 1985).

ZAN, PKDREJ – sperm-egg receptor candidate proteins

Herlyn and Zischler (2007) pointed out that analyses of primate reproductive protein evolution are usually restricted to catarrhines (Old World monkeys, gibbons, great apes), and that this may limit my ability to generalize results to other species. Unfortunately, testes mass and female remating data is less readily available for platyrrhines and strepsirrhine species. In order to include a wider range of species in their analysis of zonadhesin (ZAN), the authors used sexual body mass dimorphism to approximate sperm competition – species with large male:female body mass ratios should tend to have greater pre-copulatory male competition, and therefore less sperm competition (Herlyn and Zischler, 2007). Across the 16 species they included in their analysis, there was a significant negative correlation between ZAN d_N/d_S estimates and relative male body mass (Table 1).

In support of a relationship between ZAN d_N/d_S estimates and sperm competition, I found d_N/d_S estimates were positively correlated with both female remating rate and testes mass, using phylogenetic analysis (Table 4a). However, the results using species-level analysis were non-significant in all cases (Table

3a). Whereas Herlyn and Zischler included 16 species in their correlation, I was only able to include 13 (Table 3a). This decreased power to detect an effect may in part explain the contrast between my results and their own (Herlyn and Zischler, 2007).

Hamm et al. (2007) used maximum likelihood-based analysis to test for a correlation between mating system and adaptive molecular evolution in PKDREJ, another sperm-egg receptor candidate. The authors compared a model in which lineages were assigned to d_N/d_S classes on the basis of expected sperm competition to one in which a single d_N/d_S ratios is estimated for all branches; the first was not significantly more likely than the second. In contrast, I found strong evidence of a correlation between PKDREJ d_N/d_S estimates and female remating rate (Table 4a). Because of variation within stereotypical mating system classes (Dixson, 1999), analyses that use discrete categories to measure sperm competition may have less power to detect correlations than analyses using continuous variables. A positive correlation between PKDREJ d_N/d_S estimates and sperm competition variables agrees with the relatively strong correlations I observed in sperm-egg interaction proteins in general.

Limitations

Hurle et al. (2007) offer several good reasons why one should not expect to find significant correlations between lineage-specific d_N/d_S estimates and female promiscuity, even if there is such a causal relationship. Firstly, it is unlikely that mating systems remain fixed throughout time. The longer the branch, the more likely it is that selection pressures other than the ones observed today have shaped its evolution. Secondly, assuming that species can be characterized by one mating system or another is likely an oversimplification of the facts. There is increasing evidence supporting intra-specific variation in primate mating behaviour, particularly in species with dispersed or pair-living social systems (Goossens et al., 2006; Fuentes, 2000). Thus, the mating behaviour that is thought to be typical of a species may in fact not be. Finally, estimates of sperm competition are generally based on behavioural observations

rather than methods that offer more precision and accuracy, such as genetic analysis (de Ruiter, 2004). Even using relative testes mass as an indication of sperm competition intensity may be misleading in some cases (Schülke et al., 2004).

Potential correlations might also have been masked due to imprecision or biases in my estimates of molecular evolution. My d_N/d_S estimates reflected the rates of molecular evolution averaged across nucleotide sites, such that neutral or negative selection at some sites may obscure the pattern of evolution at positively selected sites. Analyzing different regions of the coding sequence could thus give stronger or weaker correlations between substitution rates and sperm competition. However, if different sites are the targets of selection in different species, it may be misleading to focus only on sites that show high across-species divergence when making comparisons. The site and branch-site models provided by PAML allow one to test predictions regarding both these scenarios (Yang and Swanson, 2002; Zhang et al., 2005).

Sequence divergence other than that due to nucleotide point substitutions should also be considered. Indel (insertion/deletion) substitutions that are three nucleotides long (and therefore do not disrupt the reading frame) appear to be positively selected in primates (Podlaha and Zhang, 2003). Longer indels are particularly favored by selection, perhaps because a longer N-terminus could cause the ion channel to remain activated for longer. SEMG1 and 2 also show variation in sequence length between species, which relates directly to variation in the viscosity of the seminal coagulum (Jensen-Seaman and Li, 2003). If positive selection often takes the form of changes in sequence length or organization, d_N/d_S estimates represent only a portion of reproductive protein divergence. Furthering our understanding of how changes in reproductive genes relate to protein structure and function, and how protein function relates to male and female fitness, is necessary to explain diversity in these proteins.

For all these reasons, my results should be simultaneously treated as conservative estimates of the strength of the relationship between sperm competition-related selection and the adaptive divergence of reproductive

proteins, and as tenuous examples of correlations between sperm competition variables and d_N/d_S estimates for any particular protein.

Conclusion

Wyckoff et al. (2000) first suggested a possible correlation between adaptive molecular evolution and mating system in their analysis of the primate protamine gene cluster. Although I did not find positive correlations between female promiscuity and sequence divergence in either protamines 1 or 2, I did find persuasive evidence that female promiscuity is positively correlated with reproductive protein divergence more often than expected. Despite the imprecision of comparing sequence-wide d_N/d_S estimates to proximate measures of historical sperm competition, my results suggest that such comparisons are useful in elucidating the causes of molecular divergence. Extending similar techniques to contexts outside of reproductive protein evolution will improve our understanding of molecular evolution, and of the relationship between selection and divergence in general.

CHAPTER 3: LINKING REPRODUCTIVE CONFLICT TO ANTAGONISTIC COEVOLUTION, DIVERGENCE, AND REPRODUCTIVE ISOLATION

Abstract

Conflict-driven coevolution is expected to be characterized by rapid divergence. It has been proposed that this rapid divergence could lead to speciation. I suggest that conflicts associated with various stages of reproduction are particularly likely to result in reproductive isolation (i.e., speciation), because 1) they will specifically cause divergence in reproductive traits characters, and 2) the outcome of reproductive conflicts will often directly influence the extent of gene flow between two populations. I present a common conceptual framework within which to discuss the outcomes of intragenomic, intraspecific, and interspecific conflicts. Furthermore, I make predictions regarding the influence of conflict-interactions on gene flow between populations. In contrast to current hypotheses, I suggest that the intensity of reproductive conflict does not directly influence the evolution of reproductive isolation. Instead, relative arms level (which is determined by both the level of conflict and the cost of further counter-adaptation) is an important factor in determining between-population interactions. I find that speciation will more be likely if 1) counter-adaptations are qualitative (i.e., arms level is *not* correlated with the magnitude of the trait) rather than quantitative (i.e., arms level *is* correlated with the magnitude of the trait), 2) the reproductively-parasitized party counter-adapts passively (by *decreasing* the specificity of conflict-interactions) rather than *actively* (by increasing the specifity of conflict-interactions), and 3) there are substantial fitness costs to interactions between individuals with very different arms levels. I present examples of reproductive conflicts that illustrate each of these scenarios.

Based on these predictions, I suggest that conflicts mediated by signal-receptor mechanisms, and those in which counter-adaptations by the losing party supply an 'antidote' to the winning party's 'toxicity,' are the most likely to lead to the evolution of reproductive isolation. This conclusion is supported by examples from the reproductive conflict and speciation literature. Finally, I suggest possibilities for further research that will help to bridge the gap between our understanding of conflict and coevolution, and speciation.

Introduction

Evolutionary conflicts

The phrase 'evolutionary conflict of interest' has been used to describe a wide range of biotic interactions, including host-parasite interactions, parent-offspring conflicts, intersexual conflicts, and intragenomic conflicts. Although definitions of evolutionary conflict are often context-specific, all share the following characteristic: selection behaves in an antagonistic manner with respect to two parties such that adaptive evolution by one party decreases the fitness of the other, and vice versa (Arnqvist and Rowe, 2005; Hurst and Werren, 2001). This antagonism has been described as the two parties having mutually exclusive optima for some shared trait of interest (Rowe and Day, 2006; Lessells, 2006; Hardling et al., 2001). The term 'shared trait' is interpreted broadly to mean any event (either the expression of a trait or an event occurring due to the expression of a trait) that 1) directly influences the fitness of both parties, and 2) is determined to some extent by the phenotype of each of the two parties. For example, a shared trait of interest in parent-offspring conflicts may be the rate of nutrient transfer between the parent and offspring, or whether or not offspring cannibalism occurs in a particular situation.

This definition allows us to specify what sorts of conflict-interactions are considered evolutionary conflicts of interest, and what sorts are not. In general, competition between conspecifics acting within the same behavioural or ecological niche will not be considered evolutionary conflict. When two

individuals compete with one another within the same niche, although, their evolutionary interests may be mutually exclusive, they are *competing* with one another rather than *conflicting* with one another. For example, in the context of male-male competition, males are 'in conflict' with one another – each male would prefer to fertilize as many females as possible, at the expense of his competitors. However, if males adapt such that they are able to fertilize more females (i.e., average mating rate goes up), selection will be operating directionally, rather than antagonistically. In contrast, if an alternative male reproductive tactic exists, then an evolutionary conflict of interest may develop between males in each of the alternative roles. For example, males may guard territories in order to procure copulations when they are large, and alternatively sneak copulations when they are small. Adaptive evolution that increases the fertilization rate of males when they are in the role of sneakers would simultaneously decrease the fitness of males in the role of guarders – sneakers and guarders would be in conflict over the optimum rate of fertilization by sneakers.

Recent sexual conflict literature has emphasized the role of direct fitness costs in the creation of evolutionary conflicts between males and females (Tregenza et al., 2006; Chapman et al., 2003; Gavrilets et al., 2001). Evolutionary conflicts over direct benefits, such as nuptial feeding and biparental care, are also well-established cases of sexual conflict (Kondoh, 2001; Westneat and Sargent, 1996; Dawkins, 1976). However, it is worth pointing out that the potential for evolutionary conflicts between the sexes due to indirect (i.e., genetic) costs and benefits has long been recognized (Parker, 2006; Gage et al., 2002; Parker, 1979; Trivers, 1972; Fisher, 1930). One example is the use of dishonest signals of genetic quality by males (van Dorn and Weissing, 2006; Hill, 1994; Johnstone and Grafen, 1993). Another example is that of asymmetry between the sexes in the costs of hybrid disadvantage (Parker and Partridge, 1998; Dawkins, 1976). Females generally invest more in and have fewer offspring than males, such that copulations that result in hybrid offspring with reduced fitness will be relatively costly. For males, on the other hand, the

benefits of having many mates, both heterospecific and conspecific, may overwhelm the costs of occasionally having hybrid offspring with reduced fitness.

Evolutionary conflicts can occur between genes within a genome, as well as between both conspecific and heterospecific genomes – because sexual recombination breaks up temporary coalitions of genes, it is possible for genes to spread through a population independently of one another, even if it is at each other's expense (Rice and Holland, 1997; Hurst et al., 1996; Dawkins, 1976). For the purposes of this paper, I will designate evolutionary conflicts as being either intragenomic, intraspecific, or interspecific (the last two categories both falling under intergenomic conflicts). Although in some cases these categories overlap, they provide a practical hierarchy for discussing examples of evolutionary conflict. Additionally, theoretical developments in each of these areas have been largely independent of one another, with few exceptions (Summers et al., 2003; Hardling et al., 2001; Higashi and Yamamura, 1994; Slatkin and Maynard Smith, 1979). By drawing parallels between evolutionary conflicts at different levels of biological organization, I hope to integrate developments from various disciplines into a common theoretical framework. Among other things, this will allow us to make general, testable predictions regarding the outcomes of evolutionary conflicts, and their role in gene flow between species.

Evolutionary conflicts over reproduction, and reproductive isolation

Numerous authors have suggested that the reciprocal, antagonistic selection that characterizes evolutionary conflicts may lead to the rapid divergence of allopatric populations, and subsequent reproductive isolation (Summers et al., 2003; Gavrilets, 2003; Orr and Presgraves, 2000; Hurst and Schilthuizen, 1998; Haldane, 1992). Evolutionary conflicts of interest may be important sources of speciation, as they are expected to drive divergence even in the absence of prominent ecological selection pressures (West-Eberhard, 1983). I suggest that evolutionary conflicts concerning reproduction, whether they be over the production, fertilization or development of gametes into embryos and offspring, will be particularly likely to result in speciation. The reason for this is

twofold: Firstly, when reproductive conflicts *do* result in allopatric divergence, that divergence will be specifically in traits involved in reproduction. If two populations are sufficiently divergent, this will lead specifically to reproductive incompatibilities between individuals from the two populations. Furthermore, if antagonistic coevolution occurs as a result of antagonistic selection, over time it will act to increase the complexity and redundancy of reproductive interactions within populations (Malik and Henikoff, 2002). The more convoluted the processes leading up to successful reproduction, the greater the potential for malfunction when the system is perturbed (Summers et al., 2002). If traits mediating reproduction are counter-adapted to a specific coevolutionary partner, those traits will likely be maladaptive in the context of a hybrid genome (Haldane, 1949).

Secondly, the outcome of conflict-interactions will often directly determine the extent of gene flow. This effect is obvious in conflicts of interest over mating and fertilization – in general, optimal mating rates are higher for males than for females, such that outcomes in favour of males will tend to increase mating rates (and therefore gene flow) between populations, whereas outcomes in favour of females will tend to decrease gene flow between populations (Parker, 2006; Parker and Partridge, 1998). Although less self-evident, this prediction can also be generalized to reproductive conflicts at other levels of biotic interaction. Let us consider the following example of intragenomic conflict:

In a process termed meiotic drive, or segregation distortion, certain gene sequences (i.e., driving elements) are disproportionately over-represented among the gametes produced by meiosis (Hurst and Werren, 2001). One method of accomplishing this is by biasing the outcome of oogenesis in their favor. Unlike spermatogenesis, oogenesis discards one chromosome of each homologous pair into the first polar body (Cummings, 1988). Thus, any chromosome that increases its chances of being included in the final gamete, rather than being discarded in a polar body, will be favored by selection (Henikoff and Malik, 2002). Although meiotic drive directly benefits the driving element, it can also compromise the fitness of other genes in the genome. The driver may be linked

to a deleterious allele, or centromeric misalignments during spermatogenesis may lead to male sterility. Selection should therefore favor the evolution of a suppressor gene that restores chromosomal parity during meiosis (Hurst and Schilthuizen, 1998).

The antagonistic selection between driving elements and their suppressors can be described as an evolutionary conflict over the rate at which drivers are transmitted to gametes during gametogenesis. If drivers 'win' the conflict-interaction, driver transmission will be high, and migrant drivers will spread quickly through novel populations. If, on the other hand, suppressors 'win' the conflict-interaction, they will successfully limit the rate at which migrant drivers are inherited. In general, the evolutionary interests of the reproductively exploitative party will tend to promote gene flow between two populations, whereas the interests of the reproductively exploited party will tend to limit it. Predicting the outcomes of conflict-interactions will thus be important when predicting the role of evolutionary conflicts in reproductive isolation.

Conflict dynamics

Discussions of the expected evolutionary outcomes of reproductive conflicts vary between disciplines. Predictions regarding the outcome of intragenomic conflicts tend to be made on a case-by-case basis (Kondoh and Higashi, 2000; Haig, 1993; Hurst et al., 1996), perhaps because the phenotypic and selective mechanisms associated with the suppression of selfish genetic elements vary greatly from one system to another. Discussions of intergenomic (both inter- and intraspecific) conflicts have mostly focused on battleground models that specify the conditions under which conflicts of interest are likely to occur, but pay little attention to the outcome of the conflict itself (Kolliker et al., 2005; Arnqvist, 2004; Gomulkiewicz et al., 2003; Godfray, 1995; Higashi and Yamamura, 1994).

Outcome-oriented theory is best developed in the field of sexual conflict, where interests have typically centered around the potential for ongoing sexually antagonistic coevolution (Chapman et al., 2003; Gavrilets et al., 2001; Hill, 1994).

Recent discussions, however, suggest that so-called 'arms races' may be less likely than previously thought (Parker et al., 2006; Rowe et al., 2005; Hardling et al., 2001). The existence of a variety of possible evolutionary outcomes, dependent upon the specific constraints associated with the conflict, agrees with suggestions by authors in other disciplines (Summers et al., 2003; Gomulkiewicz et al., 2003; Hurst et al., 1996; Godfray, 1995). Ongoing antagonistic coevolution will serve to escalate the conflict between two parties, such that each party will become increasingly invested in costly counter-adaptations. In many cases, resolution, de-escalation, or transformation of the conflict by transferring antagonistic selection to another shared trait may be less costly alternative for both parties (Rowe et al., 2005; Gomulkiewicz et al., 2003; Hurst, 1996). Table 9 compares findings from theoretical studies regarding the expected outcomes of evolutionary conflicts over reproduction. Despite differences in terminology, I argue that these examples may all be interpreted according to a generalized set of basic outcomes: antagonistic resolution, mutualistic resolution, stalemate, and arms race (ongoing or cycling).

In the following section, I discuss a verbal model that places reproductive conflicts within a common conceptual framework, and allows several general predictions to be made regarding the affect of reproductive conflicts on gene flow between populations. Evolutionary constraints associated with particular conflicts will determine the outcome of within-population conflict-interactions, which will in turn determine the outcome of between-population conflict-interactions. I then apply these predictions to specific examples of reproductive conflicts. By identifying the biological mechanisms that characterize various conflicts, I discuss whether or not some reproductive conflicts may be more likely to result in certain outcomes, and therefore whether or not some forms of conflict may be more likely to result in speciation than others.

Reproductive conflicts: evolutionary outcomes and between-population gene flow

Existing predictions

Two lines of argument have been used to connect reproductive conflicts with allopatric speciation: 1) antagonistic coevolution within populations may lead to rapid divergence between populations (Summers et al., 2003); and 2) in the context of sexual conflict over mating and/or fertilization, the outcome of the conflict may influence how likely mating and/or fertilization are to occur between individuals from different populations (Parker and Partridge, 1998). Parker (2006) describes the first hypothesis as the "engine of speciation" hypothesis, and the second as an "outcome moderated" mechanism of sexual conflict-driven speciation. Whereas the "engine of speciation" hypothesis predicts that speciation rates will be higher when reproductive conflicts are more intense, the "outcome moderated" hypothesis predicts that speciation rates will depend on which party wins the conflict. If males win, then females will be more likely to mate with males, including males from other populations, and the resulting gene flow will counteract reproductive isolation between the populations. If females win, the gene flow between populations will be relatively restricted, and speciation will be more likely.

It is possible to reconcile these two hypotheses with one another by recognizing that the predictions they make are about different aspects of gene flow between populations. Higher mating rates could increase the rate at which hybridization occurs, but, if divergence between the populations is high enough, those hybridizations may have a very low success rate. Which of the two mechanisms has a greater influence on overall gene flow will depend upon the nature and extent of the divergence between populations, which in turn will be determined by the outcomes of conflict interactions within populations.

Whereas the "engine of speciation" hypothesis has been used to explain how conflict-driven diversification in general could lead to speciation (Summers, et al., 2003), "outcome moderated" speciation would seem to be limited to

situations in which the probability of mating and fertilization determines the extent of reproductive isolation. However, I suggest that the outcomes of conflicts over reproduction will often have implications for gene flow between populations. In general, conflicts over reproduction occur when one party, designated 'i' (for example, a male), attempts to exploit the reproductive potential of a second party, designated 'j' (for example, a female), in such a way that i increases its own reproductive fitness at the expense of its 'partner.' Successful exploitation of j by i will increase the extent to which descendents of i advance to future generations in parallel with descendents of j. If i and j are from different populations (e.g., a male from one population, and a female from a different population), this reproductive association between individuals from different populations will serve to increase gene flow between the two populations.

I suggest that the outcome of within-population reproductive conflicts will often influence gene flow between populations even if i and j represent different species (i.e., if the reproductive conflict is interspecific). For example, a reproductive conflict may exist between a pollinator (i) and its host plant (j) over seed parasitism. If pollinators are able to successfully parasitize host plants from a population other than their own, this will encourage pollinator migration from one population to another. All else equal, increased pollinator migration will serve to increase gene flow between pollinator populations. If instead host plants limit seed parasitism by migrant pollinators, migration and therefore gene flow will be limited between pollinator populations. Because parasitism by migrants increases the likelihood of pollination by migrants, by decreasing migration between pollinator populations the host plants will also be decreasing gene flow between their own populations. Due to the fact that pollinators and host plants are reproductively co-dependent (i.e., reproduction by one requires reproduction by the other, and vice versa), the outcome of between population conflict-interactions will influence the gene flow of both parties.

If, on the other hand, the reproductive dependency between i and j is asymmetric (i.e., i is dependent upon j, but j is not dependent upon i), the influence of between-population conflict-interactions on gene flow will also be

one-sided. For example, pollinator 'cheaters' in a pollinator-host plant system may parasitize seeds without pollinating their hosts. Although successful parasitism by migrants will still encourage gene flow between parasite populations, it will not directly influence the extent of gene flow between host plant populations. In general, the more reproductively co-dependent two parties are, the greater the potential for outcome moderated speciation.

Out-comes of within-population conflict-interactions

A verbal model

Parker (1979) suggested that the long-term outcome of an arms race between males (i) and females (j) would be determined by the value to each sex of winning the conflict (V_i and V_j), versus the cost (c_i and c_j) to each sex of achieving and maintaining the arms level (a_i and a_j) necessary to win the conflict. The model assumes that any increase in arms level that allows an individual to win a conflict-interaction with a net fitness benefit will spread through the population, with the end result that all individuals of the same sex will have the same arms level. At any point in time, whichever sex has the higher arms level will always win the conflict interaction. Each sex will continue to increase their arms level in an alternating fashion until, for one of the two sexes, the current cost of upgrading an arms level past that of the opponent exceeds the future payoff of winning the conflict interaction (Parker, 2006). V_i/c_i and V_j/c_j determining the maximum allowable arms levels of i and j, respectively. If $V_i/c_i > V_j/c_j$, i will eventually win the conflict; if $V_i/c_i < V_j/c_j$, j will be the eventual winner (Parker, 1979).

In order to generalize Parker's model to conflicts other than those between the sexes, I will designate i as the reproductively parasitic party, and j as the reproductively parasitized party. For i, winning the conflict means successfully exploiting j's reproductive potential, whereas for j winning the conflict means avoiding exploitation. Because of this dissimilarity, the value of winning may often be greater for the parasitic party (i.e., $V_i > V_j$; Parker, 2006), although the reverse is also plausible. If the value of winning is lower for j than for i, V_j will determine

the level of conflict between the two parties – if V_j is close to zero, the cost to j of losing the conflict is also relatively small, so the overall level of conflict is low; if V_j is very high, the overall level of conflict will be high.

Just as the nature of the conflict will influence the relative value of winning for each party, the nature of the armaments will influence their relative cost. I suggest that counter-adaptations that are specific solutions to overcoming the other party's armament will usually be more costly than counter-adaptations that represent general solutions. For example, a counter-adaptation that consists of blocking the other party with a specific structure will tend to be more costly than a counter-adaptation that consists of making random changes to a signal in order to evade the other party's receptor. The reasons for this are two-fold. Firstly, general solutions will be easier to find – there are more possible ways of degrading a signal than improving it. Secondly, because general solutions are more plentiful, it will be easier to find one with a relatively low cost – specific solutions, on the other hand, will be restricted to a finite number of potentially high-cost counter-adaptations. I will call armaments that *increase* the specificity of the interaction between i and j 'active' counter-adaptations, and armaments that *decrease* the specificity of the interaction 'passive' counter-adaptations. Both i and j may employ either active or passive counter-adaptations. An increase in the length of female water strider abdominal spines in order to avoid costly mating attempts by males would be an example of active counter-adaptation by females (Arnqvist and Rowe, 2002). Random changes in egg coat proteins to avoid binding with multiple sperm would be an example of passive counter-adaptation (Swanson and Vacquier, 2002). Because different sorts of counter-adaptations are associated with different evolutionary trajectories as well as different costs, considering the nature of a counter-adaptation will influence the predictions I make regarding both within and between-population outcomes.

Predictions

Table 10 and Figures 8.1-8.5 together summarize possible evolutionary outcomes of conflict-interactions, based on the verbal model outlined above.

Populations may reach stable equilibria (i.e., the conflict is resolved – either antagonisitically, mutualistically, or by stalemate), or alternatively may reach unstable equilibria (i.e., cycling arms race), or no equilibrium at all (ongoing arms race). These outcomes correspond to the basic outcomes predicted by theoretical and empirical work (Table 10).

In considering these predicted outcomes, I was able to generate three novel predictions relating to the nature of counter-adaptations. Firstly, because passive counter-adapations potentially allow a party to increase its arms level at a very low cost, situations in which j employs passive, rather than active counter-adaptations may be more likely to result in ongoing arms races (Table 10, Scenario 3; Figure 9.1b). This effect could explain the rapid evolution of sperm-egg receptors in a variety of taxa – random changes in egg coat proteins may allow females to easily avoid costly polyspermy by decreasing sperm-egg binding efficiency (Swanson and Vacquier, 2002).

Additionally, the nature of the counter-adaptations may influence the result of unstable equilibria. Parker (2006) explains that, once an outcome has been decided in favor of the winning party, there may be little benefit to the losing party in maintaining costly arms levels. If this is the case, individuals with lower arms levels may actually be more fit because they do not pay the costs of their heavily armed peers. Once the losing party has decreased their arms level, there will be little benefit to the winning party to maintain a costly arms level, i.e., the conflict will de-escalate (Table 10, Scenario 3; Parker, 2006). I suggest that two factors will influence this outcome:

1) It may be difficult to decrease passive counter-adaptations to previous arms levels, because doing so would require moving from a less specific to a more specific state; in such a case selection may favor a mutualistic resolution of the conflict (Table 10, Scenario 5). If costly passive counter-adaptations are rare, mutualistic resolution may also be rare.
2) Increasing the disparity in arms levels between i and j may be inherently costly; for example, if i produces a toxin, and j the costly

antidote, j may be forced to maintain high levels of the antidote, despite losing the conflict. This factor may explain why many conflicts do reach stable equilibria (Härdling and Smith, 2005).

Outcomes of between-population conflict-interactions

Extension of the verbal model

Just as comparison of maximum arms levels allowed us to generate predictions regarding the evolutionary outcome of within-population conflict-interactions between i and j, comparisons of current arms levels between i and j from different populations will allow us to predict the outcomes of between-population interactions. Because counter-adaptations in general are costly, and because direct costs may be associated with exceeding an opponents arms level by too much, the arms level of the winning party should never exceed the arms level of the losing party by any more than is minimally necessarily to win the conflict with certainty. Although within-population arms levels are expected to track one another closely, average arms levels could differ between populations for a variety of environmental and genetic reasons. I therefore base comparisons between populations on the assumption that within-population variation in arms level will generally be lower than between-population variation in arms level.

In addition to varying relative arms level between populations, I also considered the effect of counter-adaptations being quantitative in nature (arms level is correlated with a unit of magnitude that can be used to describe the counter-adaptation) versus qualitative (arms level is not correlated with magnitude). Summers et al., (2003) suggest that quantitative counter-adaptations will be associated with low within-population diversity because of directional selection on armaments, whereas qualitative counter-adaptations will tend to be associated with higher within-population polymorphism due to negative frequency-dependent selection on corresponding counter-adaptations. I argue that quantitative counter-adaptations should also lead to lower between-population divergence. Immediately after a vicariance event, two sister populations will share armaments that are similar in cost and function. If the

armaments are quantitative, arms levels will increase in each population in a relatively predictable, similar way (i.e., more is better). As a result, counter-adaptations from one population will continue to be biologically relevant to conflicts in the second population. In contrast, if armaments are qualitative, even very low levels of divergence will create counter-adaptations that are population-specific. If counter-adaptations by j are active, their effectiveness will be population-specific (i.e., j will be locally adapted), and they will therefore be maladaptive in conflict-interactions with i from other populations. If instead counter-adaptations by j are passive, they will effectively disrupt a variety of counter-adaptations by opponents (i), and their functionality will be able to be generalized to other populations. In such a situation, i will be more locally adapted, and the outcome of conflict-interactions between i and j from different populations will be in j's favor.

I consider three different aspects of gene flow in an attempt to characterize the influence of conflict-interactions between individuals from different populations on the evolution of reproductive isolation:

1) Contest outcomes –

   If the arms level of i exceeds that of j (i.e., $a_i > a_j$,) i will succeed in reproductively exploiting j. All other things being equal, this will increase the association between descendents of i and descendents of j, promoting gene flow between the two populations. If j successfully prevents reproductive exploitation by i, descendents of i and j will be less likely to be associated with one another, and gene flow between the two populations will be reduced.

2) Migrant versus resident contest success –

   If immigrants are able to out-compete their peers in reproductive conflict-interactions, immigrants will have higher fitness than residents, and migrant alleles will spread through the population. In contrast, if immigrants have lower fitness than residents, gene flow between the two populations will be discouraged. For the purposes

of these comparisons, I assumed that whereas the costs of maintaining and using an armament should be at least partially determined by factors specific to the immigrant's population of origin, the value of winning will be determined to a greater extent by the context within which the conflict is played out – that is, the population receiving the migrant.

3) Migrant versus resident hybridization success –

If a disparity in arms levels itself entails costs to one or both parties, hybridization between two highly diverged populations will be costly. Because reproductive parasitism results in partially overlapping reproductive interests between the two parties, I will assume that any such divergence cost (k) that is imposed on i will also be imposed on j. k_i and k_j denote divergence costs resulting from the arms level of i exceeding that of j, versus the arms level exceeding that of i, respectively.

When predicting the overall extent and direction of gene flow between the two populations, I assumed that outcomes at each of these three levels would have an equal impact on gene flow. Although this assumption was made for the sake of simplicity, in actuality it is unrealistic. Nonetheless, I was able to generate useful predictions regarding how within-population conflict-interactions are expected to scale up to between-population conflict-interactions. All other things being equal, my predictions reflect *how* the outcomes of specific conflict-interactions will influence reproductive isolation, if not *to what extent* (Table 11).

Predictions

Qualitative counter-adaptations

If average arms levels do not vary significantly between the two populations, the extent of gene flow will be determined by whether i or j is the winner of conflict-interactions in each population. In this situation, my predictions correspond to those of the outcome moderated-speciation hypothesis (Parker and Partridge, 1998) – outcomes in favor of i will tend to increase gene flow

between populations, whereas outcomes in favor of j will tend to decrease gene flow, although not necessarily to the point of reproductive isolation (Table 11, Scenarios 1a,b).

When populations do differ significantly in average arms level, individuals from the high arms level population will have relatively high fitness in the low arms level population, and individuals from the low arms level population will have relatively low fitness in the high arms level population. The result will be that gene flow is strongly biased – although gene flow from the high arms level to the low arms level population will increase, gene flow from the low arms level population to the high arms level population will be blocked, resulting in a slight net decrease in gene flow (Table 11, Scenarios 2a-d). This will be the case independent of whether within-population conflicts have been antagonistically resolved in favor of i or j.

However, if, instead of antagonistic resolution, within-population conflicts remain unresolved (Table 11, Scenario 3b) or end in a stalemate (Table 11, Scenario 3c) then gene flow is expected to decrease slightly. In both ongoing arms races and stalemates, the outcome of interactions between i and j will be less certain – sometimes i will win, and sometimes j will win. Especially in the case of stalemates, the balance between the two parties will be so fragile that any immigrant with a slightly lower arms level will be selected against.

In general, populations that share similar quantitative armaments should be less likely to evolve complete reproductive isolation than those with population-specific qualitative armaments. However, slight decreases in gene flow between populations when coupled with stronger barriers may restrict gene flow enough to result in speciation. For example, when divergence costs are factored in, such that interactions between individuals with disparate arms levels are unsuccessful with respect to hybridization, reproductive isolation becomes much more likely.

Qualitative counter-adaptations

When counter-adaptations are qualitative, the outcome of between-population conflict-interactions will be determined by two factors: 1) whether

counter-adaptations by the two parties are active or passive, and 2) whether or not divergence costs (k) influence the success of hybridization. If at least one of the populations has a relatively high arms level (i.e., relatively developed counter-adaptations), this divergence between the two populations may be enough to result in postzygotic reproductive isolation (Table 11, Scenarios 5-7c,d).

Independent of potential divergence costs, the nature of the counter-adaptations will mediate the influence of within-population outcomes on the extent of overall gene flow between populations. If both i and j employ active counter-adaptations, the outcome of encounters between i and j from different populations will not be easily decided because both parties will be equally maladapted to one another. In such cases, the party that wins conflict-interactions in its own population will do relatively poorly in other populations, resulting in a slight overall decrease in gene flow (Table 11, Scenario 5).

If, on the other hand, counter-adaptations by i are passive, i will succeed in interactions with foreign j. Because of this, if resident j win conflict-interactions, migrant j will be relatively less successful, reducing gene flow between populations (Table 11, Scenario 6b,d). If instead j employs passive counter-adaptations, the outcomes will be reversed: j will succeed in interactions with foreign i, and migrant i will have lower fitness than residents when party i wins within-population conflicts (Table 11, Scenario 7a,c). Overall gene flow between the populations will tend to be lower when j employs passive counter-adaptations than when i employs passive counter-adaptations, because between-population conflict-interactions will be decided in favor of j. Coupled with high divergence costs, these conflict situations may be the most likely candidates for 'engines of speciation.'

Discussion

Case studies

In this section I will discuss my predictions in relation to known examples of reproductive conflict. I hope to provide useful illustrations of how various reproductive conflicts may be interpreted using my conceptual framework, and also to explore possible generalizations regarding the nature of reproductive conflicts and the likelihood of certain evolutionary outcomes. Finally, I will suggest testable predictions that may be used to guide further research.

Intragenomic conflict

Intragenomic conflicts in particular have been linked to reproductive isolation because of their apparent potential to create genetic incompatibilities in hybrids. Many selfish genetic elements have been discovered as a direct result of observed hybrid disadvantage (Hurst and Schilthuizen, 1998). These deleterious effects include reduced fecundity (Beeman et al., 1992), skewed sex-ratios (Merçot et al., 1995), malformed gonads (Kidwell and Lisch, 1997), and sterility (Hurst and Pomiankowski, 1992). The manner of transmission varies greatly among intragenomic conflicts, in some cases being purely parasitic (i.e., the genome must reproduce for the selfish genetic element to reproduce, but not vice versa). In other cases, transmission includes an independent aspect – for example, the cytoplasmic bacterial symbiont *Wolbachia* can be horizontally as well as vertically transmitted between hosts (Werren et al., 1995). In such cases, because reproductive interdependence is weaker between the two parties, the link between the outcome of the conflict and gene flow between populations is also expected to be weaker.

Such variability makes it difficult to find a representative example of intragenomic conflict. However, there are several reasons why centromeric drive makes a useful example. Firstly, centromeric drive is an example of an intragenomic conflict in which the two parties are reproductively co-dependent, such that reproduction by one party entails reproduction by the other party.

Centromeric drive is a specialized case of meiotic drive in which centromeres are able to bias oogenesis in their favor (Henikoff and Malik, 2002). In doing so, they will also pull all other alleles on the same chromosome along with them, as well as any centromere-binding histones that happen to be bound to them.

Intragenomic conflicts are often described as an evolutionary conflict of interest between a selfish genetic element and the genome to which it belongs. In this case, we can see the conflict of interest as one between the centromeres and centromere-binding histones in which the evolutionary interests of the centromeric histones coincide with the interests of the genome. The centromeres (i) will increase their fitness by increasing the number of histones they bind with, thereby increasing the number of microtubules pulling them towards the developing oocyte. Centromeric histones (j), on the other hand, will maximize their fitness by balancing themselves more evenly between chromosomes (Malik and Bayes, 2006).

Another reason centromeric drive makes a useful example is because the phenotypic mechanisms thought to mediate the conflict are relatively well understood. Centromeric satellites are selected to expand, attracting a greater number of centromere-binding histones. Selection will then favor centromeric histones that decrease their binding specificity to restore parity among chromosomes and centromeres during meiosis (Malik and Henikoff, 2002). Decreasing DNA-binding specificity is an example of passive counter-adaptation by centromeric histones, making it relatively easy for histones to counteract the effect of expanding centromeric satellites. This low cost suggests that conflict between centromeric satellites and centromeric histones may be likely to lead to ongoing arms races between the two. In fact, there is widespread evidence that centromeric DNA along with centromeric histones have undergone rapid antagonistic coevolution in both plants and animals (Malik and Bayes, 2006). In contrast, yeast, which have no potential for biases during meiosis, have relatively simple centromeres (Malik and Henikoff, 2002). Despite evidence of past arms races, it appears that in most cases centromeric histones have won the conflict

(Malik and Bayes, 2006), perhaps due to expected asymmetries in the costs of decreasing binding-specificity versus centromeric expansion.

The expansion of centromeric satellites is an example of a quantitative counter-adaptation – the longer centromeric satellites are, the more centromeric histones they will attract, and the higher rate of transmission they will have in female meiosis. If, then, a population in which centromeric drive is absent is crossed with a population in which centromeres and centromeric histones have undergone antagonistic coevolution, should we expect the two to be reproductively isolated? When a centromeric histone (j) from the first population is paired with a centromere (i) from the second population, segregation during female meiosis will be strongly biased in favor of the driving centromere (Table 12, Case Study 1a). This will allow a centromere from the second population to quickly spread through the first population – at the same time, centromeric histones from the first population will do relatively poorly in the second popuation. However, unless pairings between centromeres and centromeric histones from different populations have particularly high fitness costs (k), I do not predict any strong barriers to gene flow (Table 12, Case Study 1b).

These predictions are supported by empirical evidence of centromeric drive in hybrids of *Mimulus nasutus* and *Mimulus guttatus* monkey flowers (Fishman and Willis, 2005). Because selfing reduces the transmission advantage of driving centromeres, we would expect outcrossing *M. guttatus* to be more likely to evolve centromeric drive than inbreeding *M. nasutus* (Malik, 2005; Hurst and Werren, 2001). When *M. nasutus*/*M. guttatus* F1 hybrids are crossed with one another, the genotypic ratio of F2 hybrids is strongly skewed towards an allele from the *M. guttatus* parental population. In backcrosses against either parental species, a 100% transmission bias occurred, but only when the F1 hybrid was the female parent. Fishman and Willis (2005) concluded that the transmission bias was due to female meiotic drive, and that the locus in question was either a centromere itself, or closely linked to a centromere. Despite the strength of centromeric drive in this system, there were not any observed fitness disadvantages to either male or female hybrids. Additionally, after multiple

generations of backcrossing hybrid descendents to the *M. nasutus* parental line, the *M. guttatus* driving allele and other closely linked alleles persisted in the population. Intragenomic conflict over centromeric drive therefore does not seem likely to lead to reproductive isolation between these two species of *Mimulus*.

Intergenomic, intraspecific conflict

Intrasexual conflicts stand out among other intraspecific reproductive conflicts in receiving a great deal of recent attention (Tregenza et al., 2006). Evidence that sexual conflict promotes speciation due to the divergence of reproductive characters has come from experimental, comparative, and theoretical avenues (Gage et al., 2004; Martin and Hosken, 2003; Arnqvist et al., 2000; Gavrilets 2000; Parker and Partridge, 1998). However, other sources suggest there is equal evidence that sexual conflict 1) has no effect on speciation rates, 2) slows speciation, or even 3) speeds extinction (Bacigalupe et al., 2007; Morrow et al., 2003; Gage et al., 2002; Parker and Partridge, 1998). I argue that which one of these outcomes ends up being the case will depend, to a large extent, upon the nature of the sexual counter-adaptations, as well as the conflict (Table 11).

In diving beetles (Dytiscidae), there is widespread conflict between the sexes over mating rate (Miller, 2003). Immediately after male beetles mount, females attempt to dislodge the would-be mates with erratic swimming behaviour (Bergsten and Miller, 2007). In order to maintain their grip on females, males have large, sucker-shaped setae that allow them to attach to the female's dorsal surface. In many species, females, as well as males have smooth elytra, an ideal surface for suction cup-like male setae to attach to. However, in several clades, females have evolved modified dorsal cuticles that make suction less efficient. These dorsal modifications are examples of quantitative counter-adaptations – female dorsal cuticles in these species include an amazing variety of ridged and stippled surfaces, suggesting that there are a variety of ways in which male suction may be disrupted (Bergsten and Miller, 2007; Hardling and Bergsten, 2006; Miller, 2003). Male setae show antagonistic coevolution with female dorsal

surfaces such that the number and positioning of the suction cups correspond to contours on the female's dorsal surface (Bergsten and Miller, 2007; Hardling and Bergsten, 2006). Male (i) and female (j) diving beetles in these species thus appear to be engaged in a qualitative arms race with one another (Miller, 2003).

Whereas modifications of male setae serve to increase specificity between males and females, modifications of female dorsal surfaces decrease specificity between the two sexes. Because males are specifically adapted to females from their own population, they will do relatively poorly with females from other populations. Consequently, females will do relatively well with males from populations other than their own. Although females themselves will have higher fitness in foreign populations, they will also tend to limit mating rates between the two populations. Along with the reduced fitness of male migrants, this reduction in gene flow may guide the two populations towards reproductive isolation (Table 12, Case study 3a). In the case of mating conflict between male and female diving beetles, there does not seem to be any intrinsic cost to divergence (factor k). However, wild populations have shown considerably asymmetries in arms level between recently diverged species (Bergsten and Miller, 2007). Therefore, if unknown divergence costs do exist (e.g., specialized male setae damage the dorsal cuticle of non-coevolved females, reducing female, and maybe male fitness), such fitness costs would make the evolution of reproductive isolation between populations more likely (Table 12, Case study 3b).

Intergenomic, interspecific conflict

When the reproduction of heterospecifics is closely linked, factors that influence gene flow in one species may also influence gene flow in the second. Although cases of such reproductive reciprocity are rare, when such systems do occur they are likely to involve both mutualistic and antagonistic components, making them useful opportunities for the study of reproductive conflicts. One well-known example of interspecific reproductive co-dependence is that of the obligate fig-fig wasp host-pollinator system. Fig plants are exclusively pollinated by female fig wasps attempting to oviposit in fig ovules (Kiester et al., 1984).

Each ovule is capable of producing either a single seed, or a single wasp, depending on whether or not it received pollen and/or an egg from its pollinator (Anstett et al., 1996). It is in the fig plant's best interest to balance male function (production of pollinators) with female function (production of eggs; Yu et al., 2004). Analogous to a selfish sex ratio distorter, it's in the fig wasp's best interest to bias reproduction in favor of male function.

It appears that monoecious fig plants are able to limit oviposition by making some flowers more costly to parasitize than others. Shorter-styled flowers are easier for female wasps to access than longer-styled flowers. Longer-styled flowers also tend to have larger stigmatic surfaces, allowing them to collect pollen more easily. Thus, the flowers that are the most likely to be pollinated are also the least likely to be parasitized (Jousselin et al., 2004). In some wasp species, females are prevented from ovipositing into longer-styled flowers by the length of their ovipositor. In other species, female wasps have sufficiently long ovipositors, but will parasitize any unoccupied shorter-styled flowers first, perhaps because a shorter handling time makes these flowers less costly. In these species, both long-styled and short-styled flowers are frequently parasitized, although long-styled flowers to a lesser extent (Yu et al., 2004). Considering that competition among foundresses for ovipositioning sites is typically high (Anstett et al., 1996), selection should favor females with longer ovipositors that are able to parasitize both longer-styled and shorter-styled flowers. In the case of the species with shorter ovipositors, it may be that an unknown evolutionary constraint makes the evolution of longer ovipositors too costly for female wasps (Yu et al., 2004).

Increases in style and ovipositor length represent quantitative counter-adaptations that should maintain their functional significance across populations. When costs prevent female wasps from evolving longer ovipositors, fig plants are able to successfully limit ovule parasitism to only the shorter-styled flowers (Population 1, Table 12, Case study 2). If, on the other hand, wasps are free to evolve longer ovipositors, fig plants will need to evolve even longer styles in order to keep ovule parasitism in check. However, the fact that some wasps are

able to evolve ovipositors long enough to parasitize even the longer-styled flowers suggests that there is an upper limit on style length. Fig wasps (i) and fig plants (j) thus seem to have reached a stalemate – fig plants cannot make styles long enough to prevent the parasitism of longer-styled flowers, but the costly handling time of longer styles prevents fig wasps from parasitizing all ovules. The result is that some, but not all of the longer-styled flowers will be parasitized each generation (Population 2, Table 12, Case study 2).

Because the arms race between fig plants and fig wasps is so close in population 2, both wasp and fig plant phenotypes from population 1 will do relatively poorly in the second population. Migrant wasps from population 1 will have shorter ovipositors, limiting parasitism to shorter-styled flowers, and making the migrants less reproductively competitive compared to residents. Fig plants in population 2 that originated from population 1 seeds will generally have shorter-styled flowers than resident fig plants, which will make these migrants more susceptible to seed parasitism. Despite the fact that female wasps from population 2 will do relatively well in population 1, I predict that gene flow between the two populations will tend to be constricted (Table 12, Case study 2). Part of the reason for this is that, although females with longer ovipositors will, on average, have a greater number of offspring than residents with shorter ovipositors, those additional offspring are expected to have relatively lower fitness – i.e., there is a divergence cost to interactions between the two populations. When, at maturity, female wasps emerge from their galls, those in the lower layers (i.e., in the longer-styled flowers) will be less likely to find mates before leaving the fig. Wasps developing in the lower layer may also be overcrowded during development relative to those in the upper layers, leading to a decrease in fitness (Anstett et al., 1996). It thus seems plausible that fig-fig wasp reproductive conflict will contribute to the reproductive isolation of both fig plant, and fig wasp populations. In particular, if fig wasps are selected to avoid fig plants from other populations due to the costs of reproductive conflict, pollen flow and therefore hybridization between fig plant populations will be unlikely. In support of this prediction, phylogenetic analyses indicate that fig wasp pollinators

have co-speciated with their host plants more often than completely parasitic fig wasp species (Weiblen and Bush, 2002).

Generalizations, and suggestions for future research

Hypotheses regarding the role of reproductive conflicts in speciation have centered around expectations of rapid divergence due to antagonistic coevolution, and the generation of hybrid incompatibilities as a by-product (Hayashi et al, 2007; Hurst and Schilthuizen, 1998; Kiester, 1984). Parker and Partridge (1998) argue that, in the case of reproductive conflicts between the sexes, the outcome of conflict-interactions will directly influence gene flow between populations. In this paper, I suggest that the predictive ability of either of these models, taken in isolation, is limited. It is reasonable to expect that, if there are high intrinsic fitness costs to hybridization, these costs will limit gene flow between populations in spite of other mechanisms influencing gene flow. However, antagonistic coevolution within populations will not necessarily lead to rapid divergence between populations. Furthermore, when it does, it will not necessarily be the sort of divergence to result in hybrid incompatibilities. Instead, the influence of divergence upon gene flow will be mediated by its influence on between-population interactions.

One example is the case of what I call quantitative armaments. When the efficacy of a counter-adaptation is correlated with its magnitude, I predict that directional selection will be more likely to cause parallel evolution in two sister populations. For example, in water striders, the speed with which females are able to dislodge males is positively correlated with the length of female abdominal spines (Arnqvist and Rowe, 1995). Supporting my prediction, several water strider species show similar elongation of abdominal spines in females (Andersen, 1993). Furthermore, phylogenetic-comparative analysis shows strong convergent coevolution between male and female body shape – which, in females, is partially determined by spine length (Arnqvist and Rowe, 2002b). Whether or not my predictions hold true over a wide range of taxa requires further research, both experimental and comparative. If it is the case that

quantitative counter-adaptations are more likely to follow similar evolutionary paths in allopatry, reproductive conflict may cause sister populations to be more divergent from their ancestral population than they are from one another. In these situations, I predict that reproductive conflicts will often promote gene flow more than they impede it. If there is any asymmetry between the two populations in arms level, individuals (or genetic elements) from the population with the higher arms level will be competitively superior to individuals from the population with the lower arms level. Additionally, individuals from the population with the lower arms level are more likely to be reproductively exploited by individuals from the population with the higher arms level, than vice versa. I therefore predict that gene flow will tend to be from the high arms level population to the low arms level population. Specifically, the alleles that confer the increased armament levels should be the ones to show the most introgression.

In contrast to quantitative armaments, I predict that qualitative armaments from different populations are more likely to evolve along orthogonal axes to one another. Thus, in cases of qualitative counter-adaptation, antagonistic coevolution should be more likely to result in true divergence between sister populations. However, the impact of that divergence on gene flow will depend upon the specificity of counter-adaptations within populations. Both parties may be counter-adapted to one another in specific ways, or one party may be more generally counter-adapted to the other party. In particular, if the reproductively *exploited* party is counter-adapted more generally, but the reproductively *exploitative* party is counter-adapted more specifically (Table 11, Scenario 7), the first party will tend to win between-population conflict interactions. Thus, when parties from different populations meet, reproductive exploitation, and therefore gene flow, will be less likely to occur.

Both quantitative and qualitative armaments are expected to be common in a variety of reproductive conflicts, across all levels of biological organization (Gage, 2004; Hosken and Stockley, 2004; Summers et al., 2003). If an armament shows little within-population variation, experimental and/or comparative studies may be necessary in order to determine whether or not the armament is

quantitative or qualitative. Similarly, determining whether a counter-adaptation is passive or active requires an understanding of the morphological, physiological, behavioural, or molecular mechanisms by which the two parties interact. Although I hesitate to make generalizations regarding the evolutionary outcome of intragenomic conflicts versus intergenomic conflicts, the following predictions can be made based upon the nature of the counter-adaptations in question:

1) When conflicts are mediated by qualitative signal-receptor interactions in which elaboration of the signal represents counter-adaptation by the reproductively exploitative party (i), and modification of the receptor represents counter-adaptation by the reproductively exploited party (j), reproductive isolation between sister populations will be more likely. As long as the receptor is not under strong natural selection in another context, the reproductively exploited party will be able to counter-adapt passively, decreasing gene flow between sister populations. Examples of reproductive conflicts in which this prediction would apply include: conflict between males and females over the rate at which sperm penetrate the egg coat (Swanson and Vacquier, 2002); conflict between queens and workers over worker reproduction in eusocial insects (Malka et al., 2007); and the exploitation of non-adaptive sensory biases in females by males (Hill, 1994).
2) When conflicts are mediated by quantitative offense-defense or toxin-antidote counter-adaptations, such that the greater the difference in armament levels between the two parties, the greater the mutual cost of their interaction, reproductive isolation between sister populations will be more likely. Such counter-adaptations will result in large costs of divergence (k) between sister populations with different average armament levels. Examples of reproductive conflicts in which this prediction would apply include: conflict between maternally- and paternally-imprinted genes over resource allocation to developing embryos (Haig, 2004); and conflict between male and female bed bugs over traumatic insemination (Morrow and Arnqvist, 2003).

The verbal model I have discussed is limited in that it makes few predictions regarding the relative importance of various barriers to gene flow. In some reproductive conflicts, the outcome of the contest interaction between the two parties will have a very definite influence of whether gene flow occurs between their respective populations – for example, in the case of conflicts between males and females over hybridization (Parker and Partridge, 1998). In many other cases, the divergence costs (k) of the interaction will be more important, leading to reproductive isolation through hybrid incompatiabilies. Nonetheless, by placing reproductive conflicts within a common framework in relation to gene flow, I hope to have drawn attention to the importance of understanding the nature of the mechanisms mediating the conflict in order to predict that conflict's evolutionary outcome.

Conclusion

Reproductive conflicts manifest themselves in a wide variety of contexts. Despite obvious differences, it is possible to draw useful parallels between conflicts operating at different levels of biological organization. In doing so, I suggest that the nature of the counter-adaptive mechanisms mediating a conflict may often give a better indication of how likely the conflict is to lead to speciation, than whether the conflict is intragenomic, intraspecific, or interspecific. Additionally, divergence of reproductive characters, on its own, does not imply reproductive divergence (i.e., reproductive isolation). Instead, such divergence may permit, or even, paradoxically, promote gene flow. Further exploration of these ideas will benefit from their development into a more formal model. In particular, taking account of the manner in which genetic transmission from one population to another actually occurs will likely have important consequences for predictions regarding gene flow, and reproductive isolation.

CHAPTER 4: CONCLUSION

This thesis is largely a synthesis of previous work regarding the causes and consequences of adaptive evolution in reproductive traits. In bringing this knowledge together, I hope to have narrowed existing gaps in the discipline, as well as to have highlighted areas that are priorities for future investigation. In both Chapter 2 and 3, I stress the importance of better characterizing the way in which reproductive traits function. In the majority of primate reproductive proteins, we have a very poor understanding of the relationship between changes at the molecular level and at the phenotypic level. It may thus be misleading to draw conclusions regarding the causes of rapid evolution in these proteins based solely on the presence or absence of correlations with d_N/d_S estimates. In many cases, establishing a stronger functional link between sequence variation and fitness variation will both complement d_N/d_S-based estimates of positive selection, and provide a more convincing argument for the role of selection in molecular evolution (Jensen-Seaman and Li, 2003; Podlaha and Zhang, 2003). In order to do so, however, the function of known reproductive proteins must be better characterized.

Developing a more mechanistic understanding is equally important in relating the evolution of populations to the formation of species. The manner in which two populations have diverged from one another will be a determining factor in predicting gene flow between the populations. I argue that, when that divergence is driven by conflict over reproduction, the nature of the counter-adaptations mediating the conflict will determine the nature of the resulting divergence. In order to predict whether or not conflict-driven speciation is likely to occur, therefore, it is necessary to understand the way in which counter-adaptations function. The functions of obvious morphological counter-adaptations are often easily observed (Morrow and Arnqvist, 2003; Arnqvist and Rowe,

2002b), or else are assumed (Anstett et al., 1996). Behavioural and physiological counter-adaptations, on the other hand, are usually less easily characterized. In intragenomic conflicts in particular, the mechanisms by which many selfish genetic elements manipulate their hosts are poorly understood. Better characterizing the functional significance of polymorphism in these systems will help to explain their role in the evolution of reproductive isolation (Hurst and Schilthuizen, 1998).

Are primate reproductive proteins likely to play a role in the evolution of reproductive isolation? The fact that both male and female sperm-egg interaction proteins appear to evolve more quickly in species with high sperm competition suggests that the evolution of these proteins is driven by intersexual conflict. Male-expressed sperm proteins will be selected to increase binding affinity with specific female-expressed sperm-receptors, in order to increase their chances of being the first to fertilize the egg. Female-expressed receptors will be selected to alter their sequence in order to decrease binding affinity with sperm, and prevent polyspermy (Swanson and Vacquier, 2002a). I argue that, because alterations in female-expressed sperm receptors represent passive counter-adaptations, the cost to females of maintaining the arms race will be relatively low, and ongoing antagonistic coevolution and divergence will be likely to result. Because males will have adapted to what are likely to be population-specific sperm-receptors, it will be difficult for males to successfully fertilize females that are not from their own population. It is thus likely that pre-zygotic reproductive isolation would occur as a result of conflict between males and females over fertilization. Sperm-egg interaction proteins therefore are particularly important, both as examples of adaptively evolving reproductive proteins, and as potential factors in the evolution of species.

FIGURES

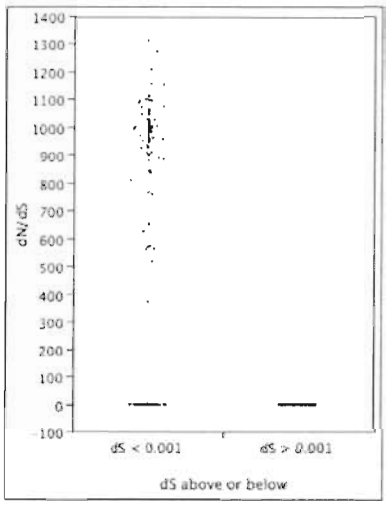
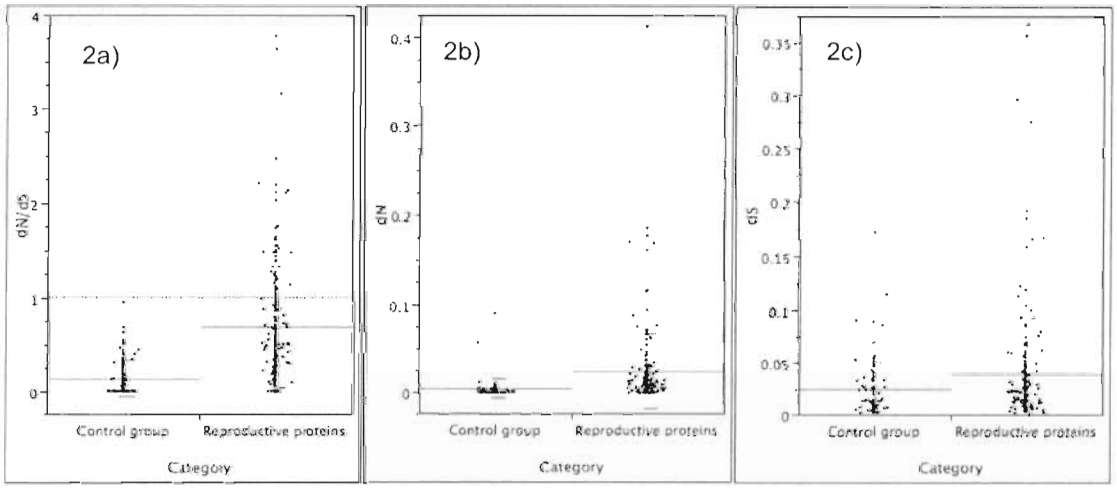


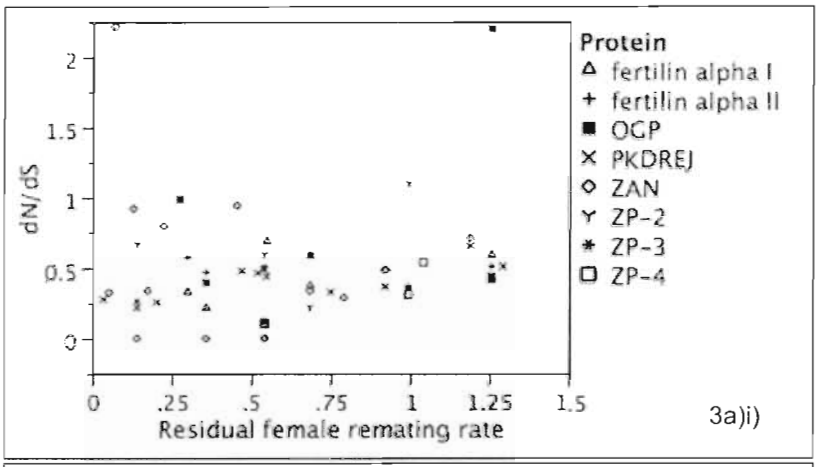
Figure 1. Results from a t-test comparing d_N/d_S estimates when $d_S < 0.001$ to those when $d_S > 0.001$.

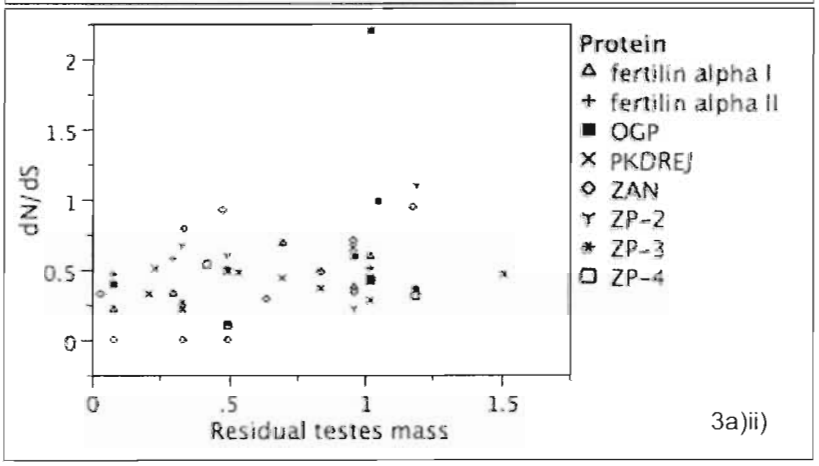
$d_S < 0.001$: n = 207, $d_S > 0.001$: n = 390; t = -11.12, df = 206; p < 0.0001.



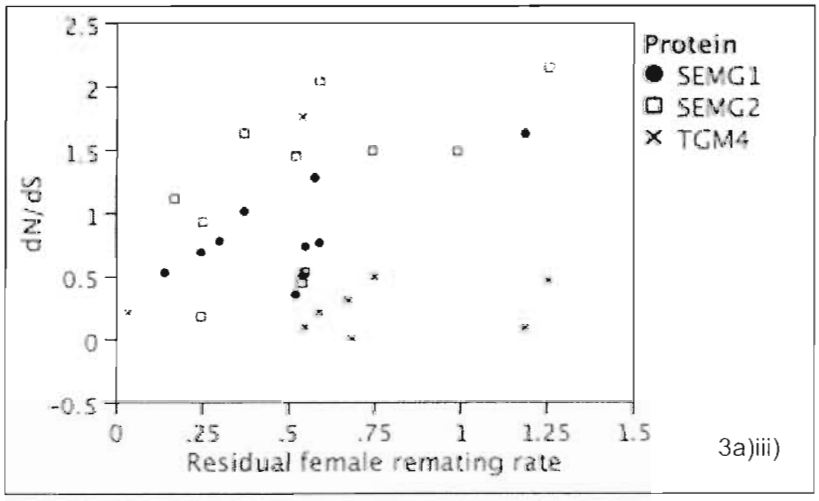
Figures 2a-c. Results of t-tests comparing terminal branch d_N/d_S, d_N and d_S estimates for control group and reproductive proteins.
Control group: n = 94, reproductive proteins: n = 188. Long lines represent means for each group (X); short lines mark one standard deviation (σ) from the mean in each direction.
a) t = 10.7, df = 243.36, p < 0.0001; control group X = 0.13, σ = 0.19; reproductive proteins X = 0.68, σ = 0.65. Dotted line indicates $d_N/d_S = 1$.
b) t = 6.05, df = 230.48, p < 0.0001; control group X = 0.003, σ = 0.011; reproductive proteins X = 0.023, σ = 0.043.
c) t = 2.98, df = 280, p = 0.0031; control group X = 0.024, σ = 0.027; reproductive proteins X = 0.039, σ = 0.054.

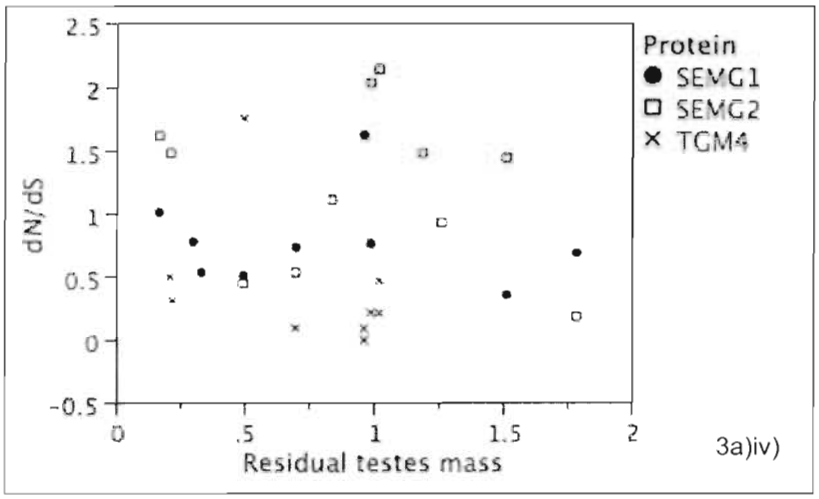
Figures 3a-d. Distribution of dN/dS estimates in relation to relative female remating rate, and relative testes mass.
Figures 3a,b: reproductive proteins; figures 3c,d: control group proteins.



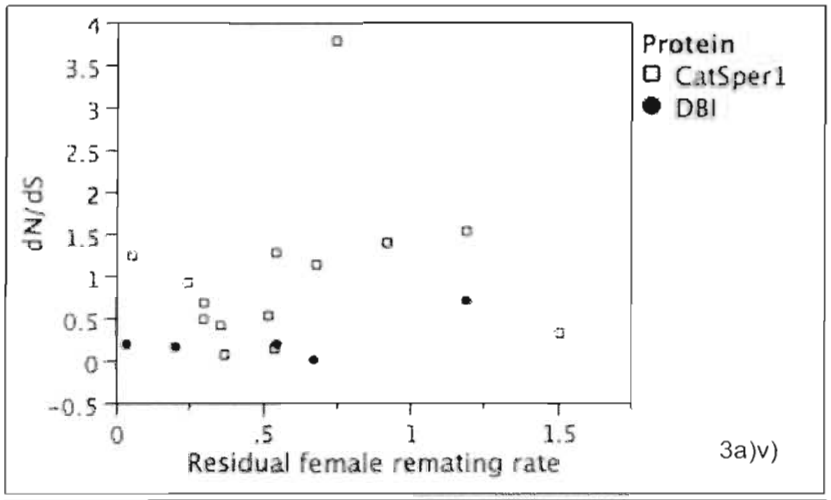


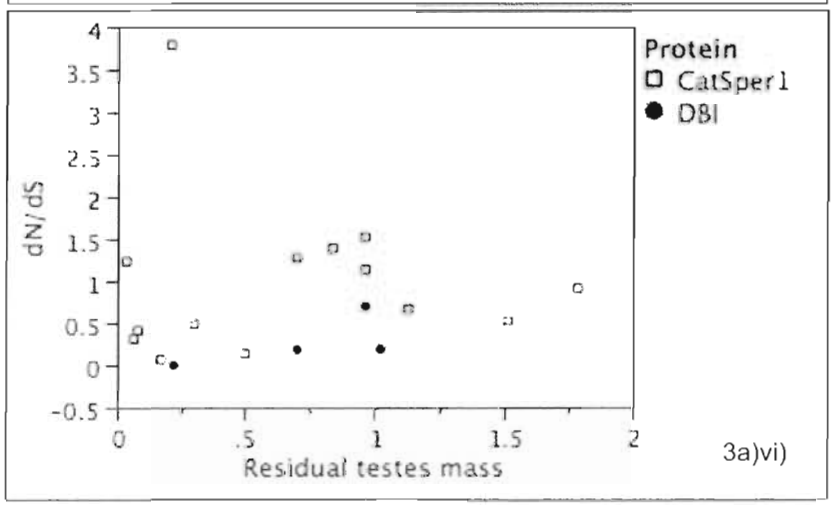
3a) i-ii) Proteins involved in sperm-egg interactions.



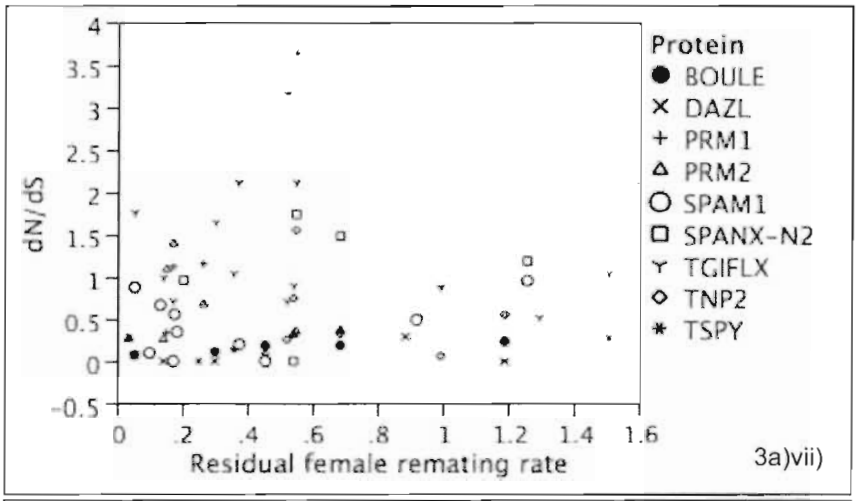


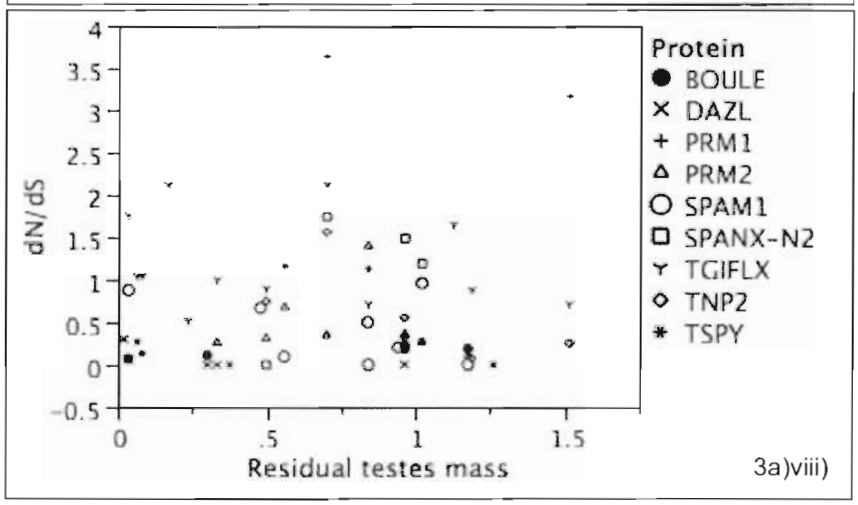
3a) iii-iv) Proteins involved in seminal coagulation.



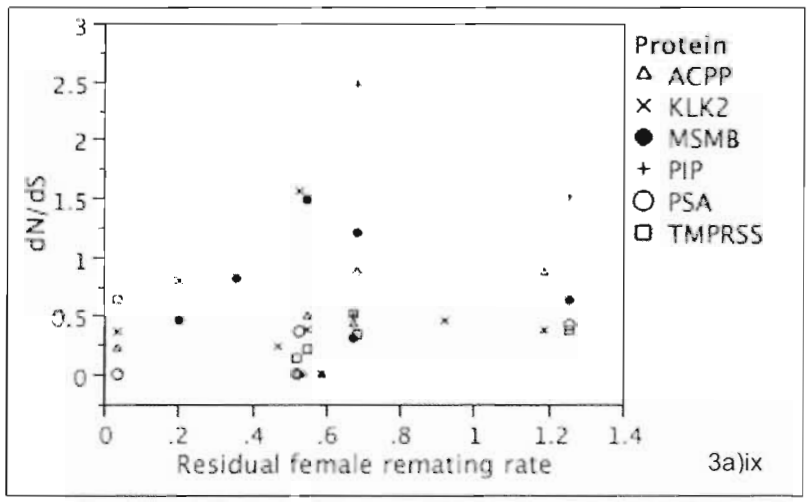


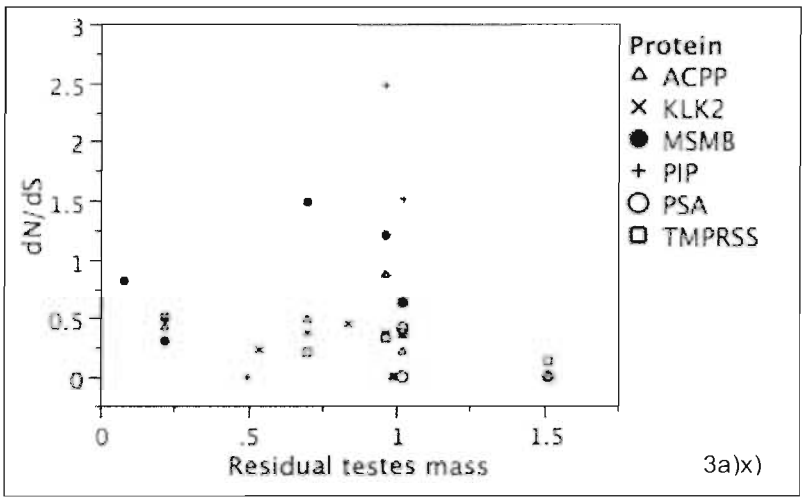
3a) v-vi) Proteins involved in sperm motility.



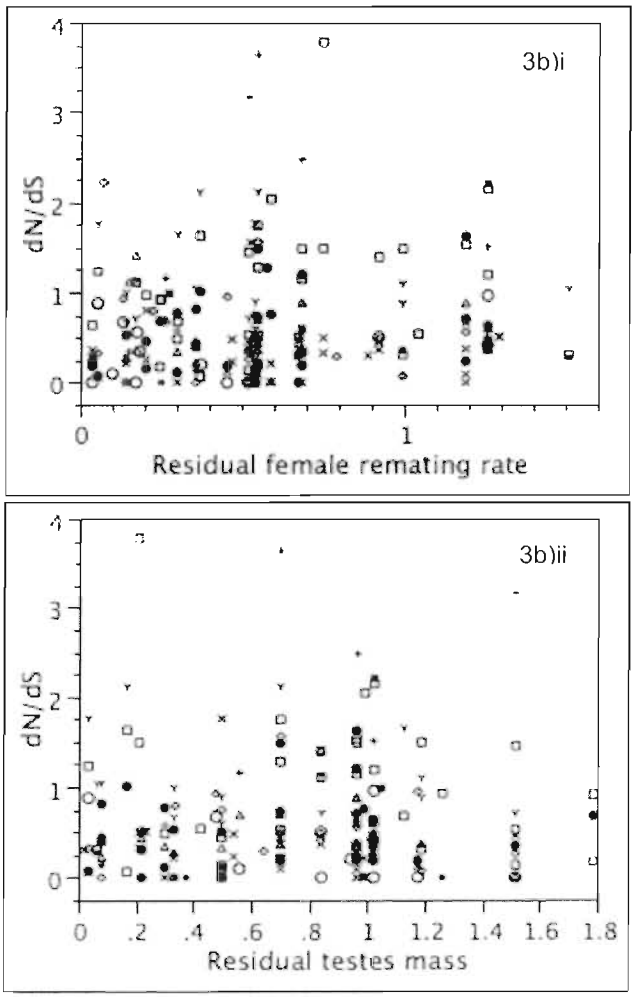


3a) vii-viii) Proteins involved in spermatogenesis.

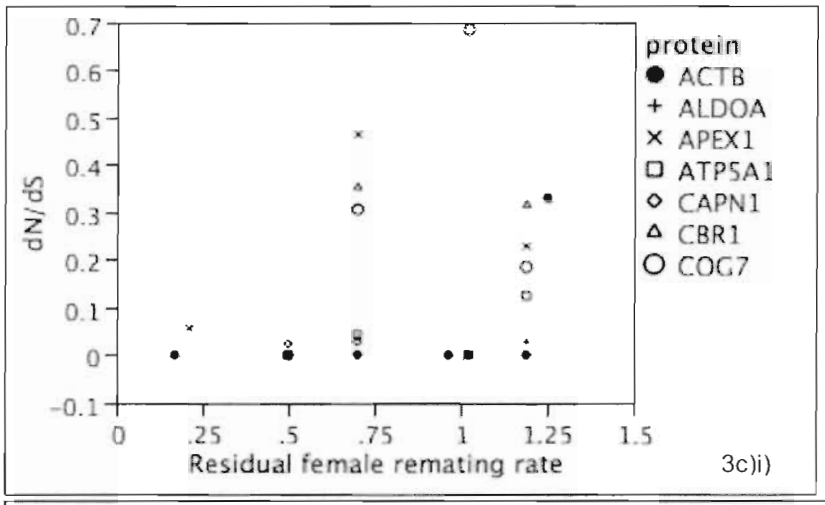




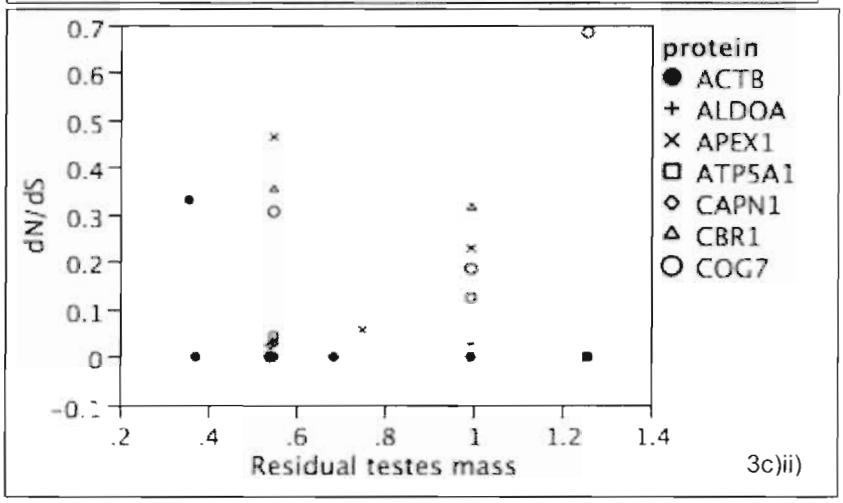
3a) ix-x) Proteins involved in the dissolution of seminal coagulum. TMPRSS is included in these figures as a seminal protein, although its function is poorly characterized.



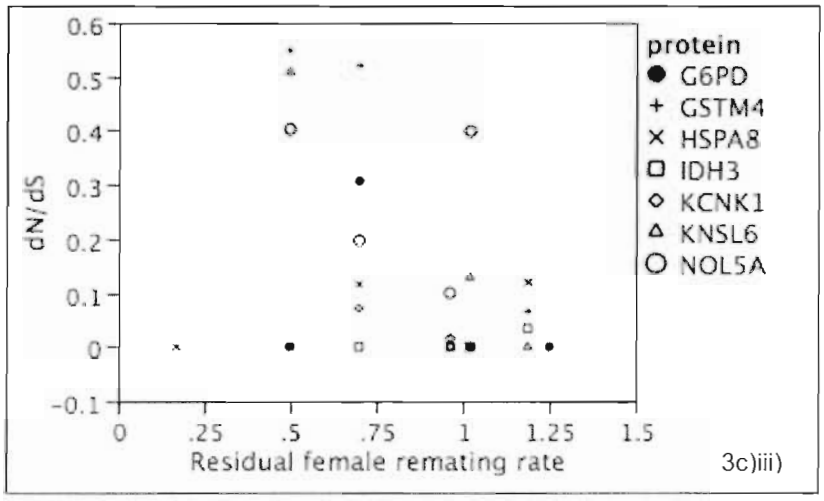
3b) i-ii) Pooled d_N/d_S estimates of all reproductive proteins; symbols are as above.



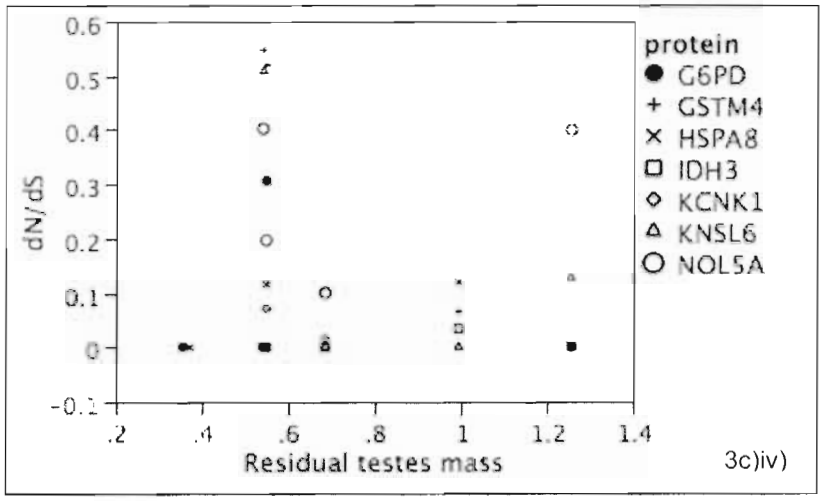
3c)i)

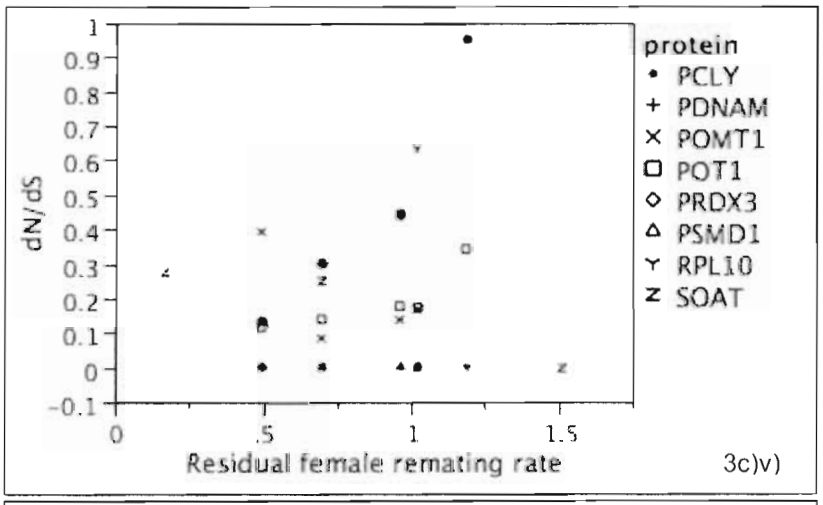


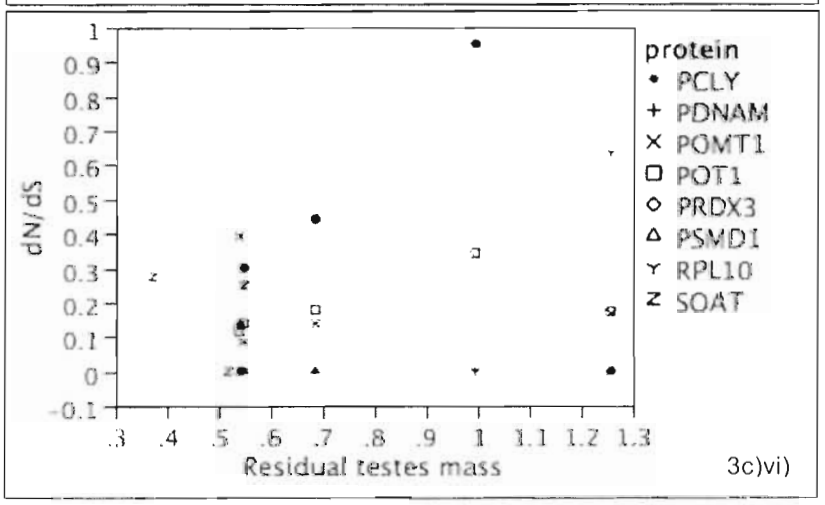
3c)ii)



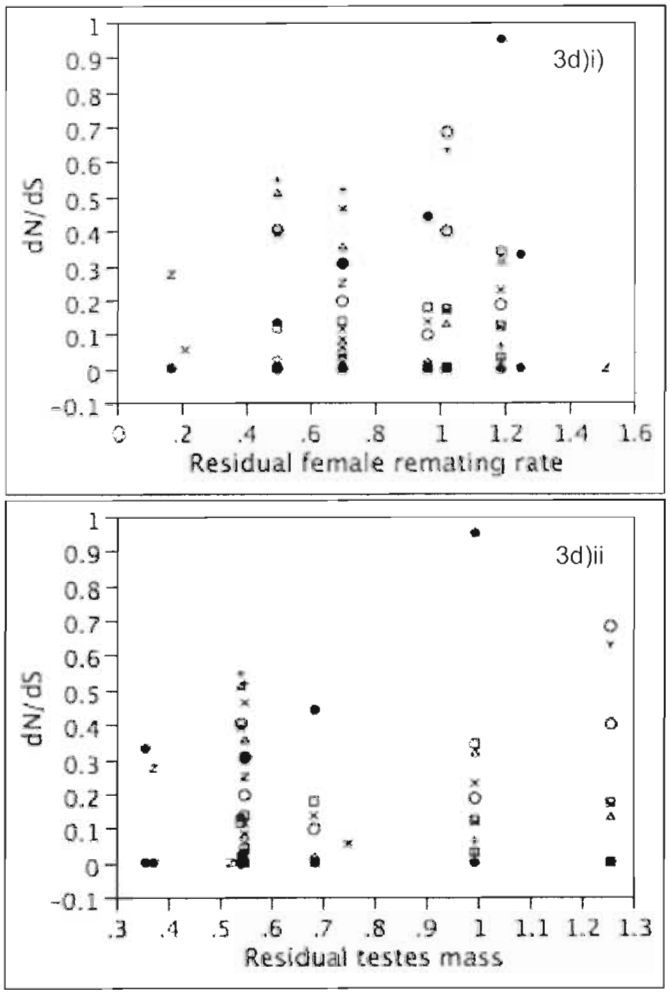
3c)iii)





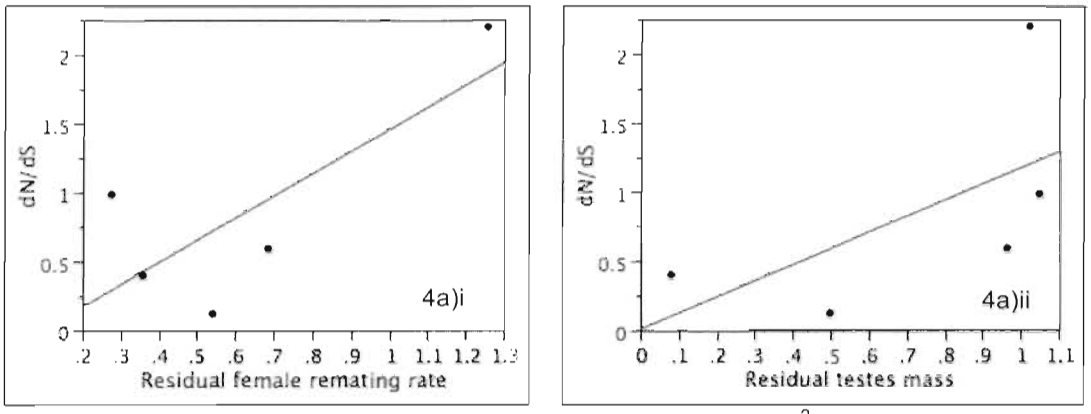


3c) i-vi) Control group proteins, presented in groups of 7-8 proteins. Groupings are unrelated to protein function.

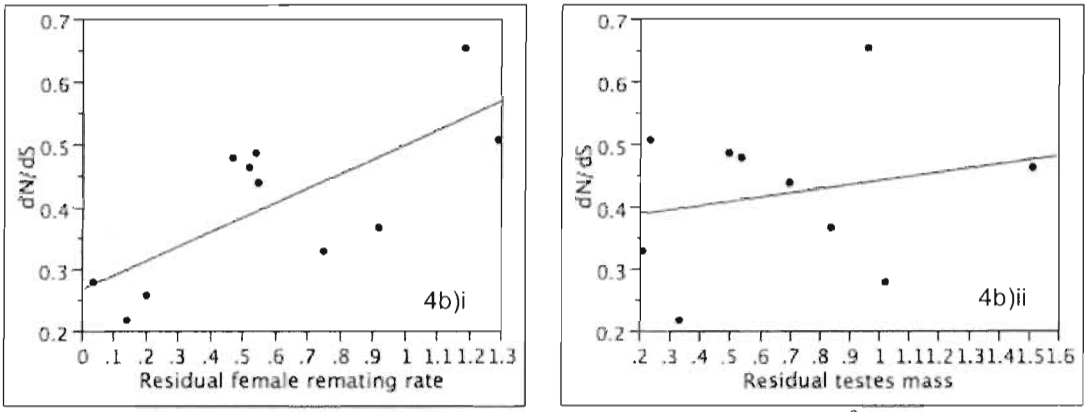


3d) i-ii) Pooled d_N/d_S estimates of all control group proteins; symbols are as above.

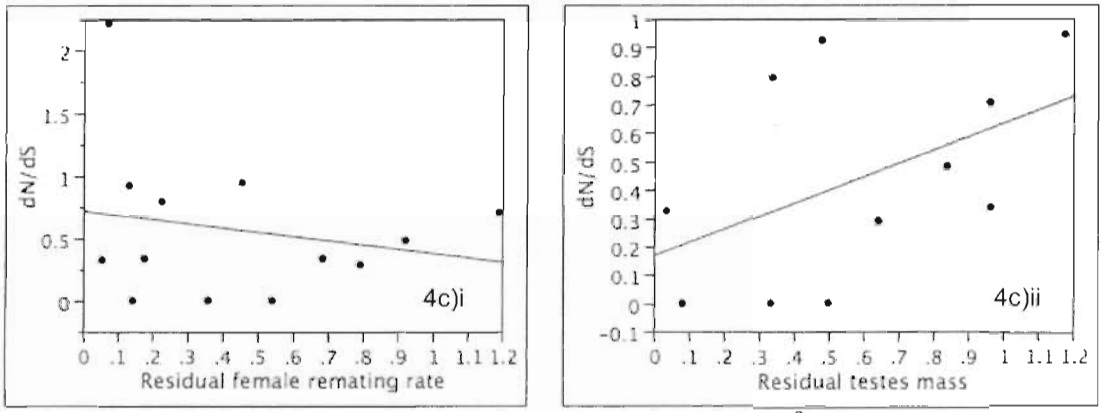
Figures 4a-i. Examples of results from species-level regressions.



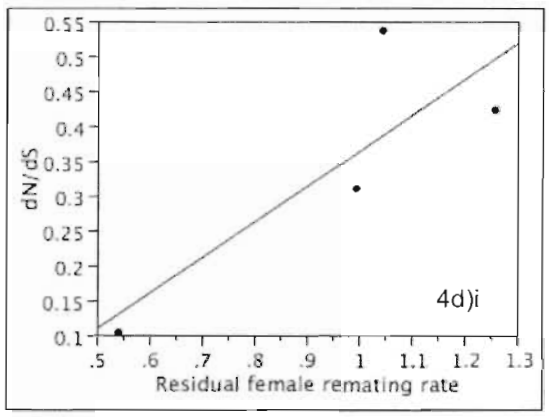
4a) OGP d_N/d_S estimates compared to i) female remating rate: $r^2 = 0.59$, $p = 0.13$; ii) testes mass: $r^2 = 0.37$, $p = 0.27$.

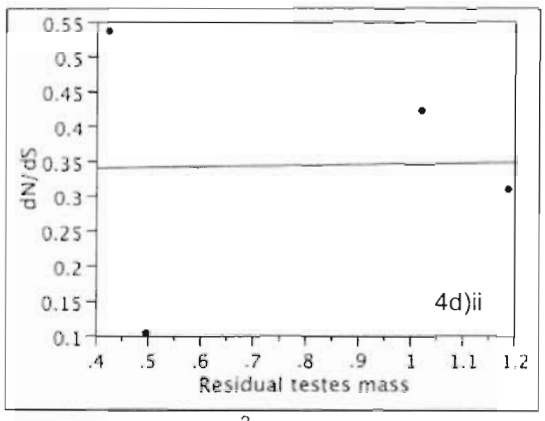


4b) PKDREJ d_N/d_S estimates compared to i) female remating rate: $r^2 = 0.54$, $p = 0.010$; ii) testes mass: $r^2 = 0.045$, $p = 0.55$.

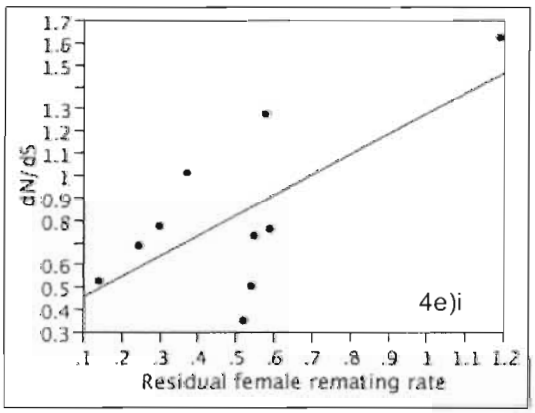


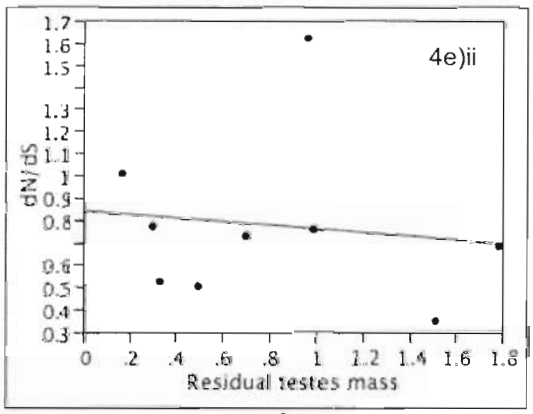
4c) ZAN d_N/d_S estimates compared to i) female remating rate: $r^2 = 0.043$, $p = 0.50$; ii) testes mass: $r^2 = 0.23$, $p = 0.13$.



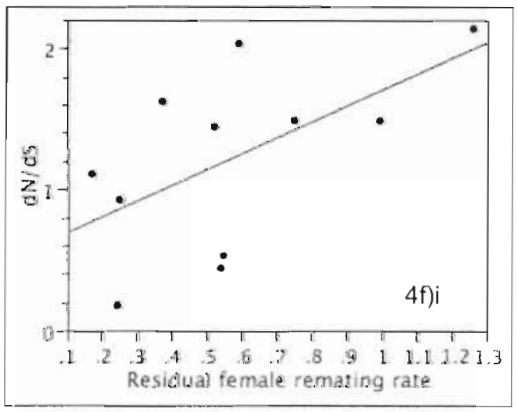


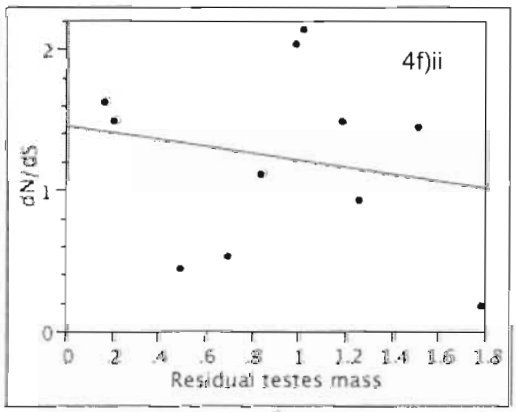
4d) ZP-4 d_N/d_S estimates compared to i) female remating rate: $r^2 = 0.70$, $p = 0.17$; ii) testes mass: $r^2 = 5.23 \times 10^{-4}$, $p = 0.98$.



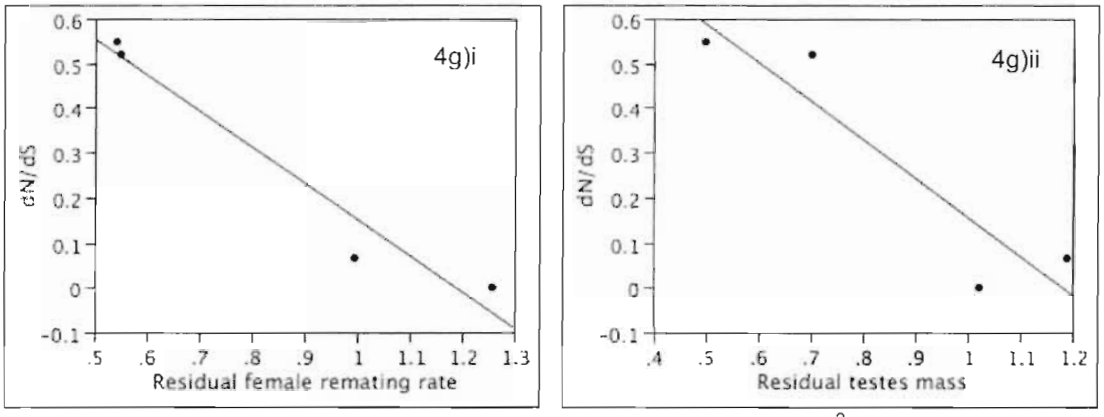


4e) SEMG1 d_N/d_S estimates compared to i) female remating rate: $r^2 = 0.47$, $p = 0.029$; ii) testes mass: $r^2 = 0.014$, $p = 0.76$.



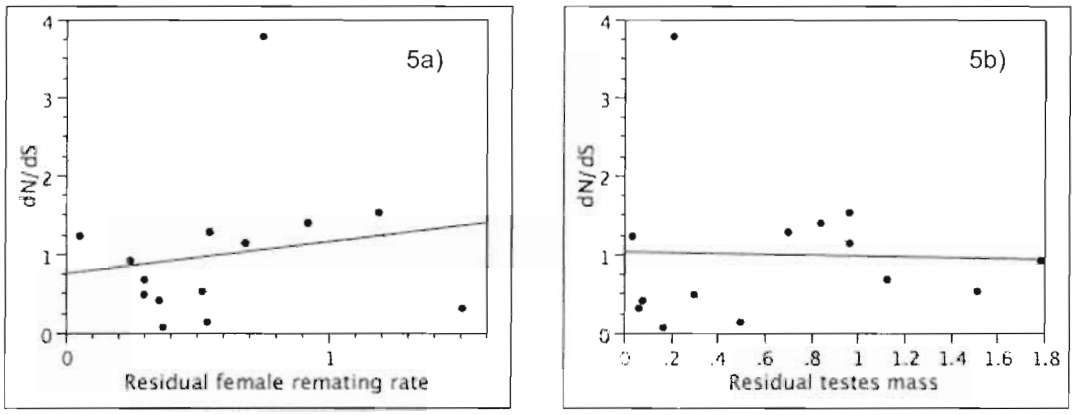


4f) SEMG2 d_N/d_S estimates compared to i) female remating rate: $r^2 = 0.34$, $p = 0.062$; ii) testes mass: $r^2 = 0.039$, $p = 0.56$.

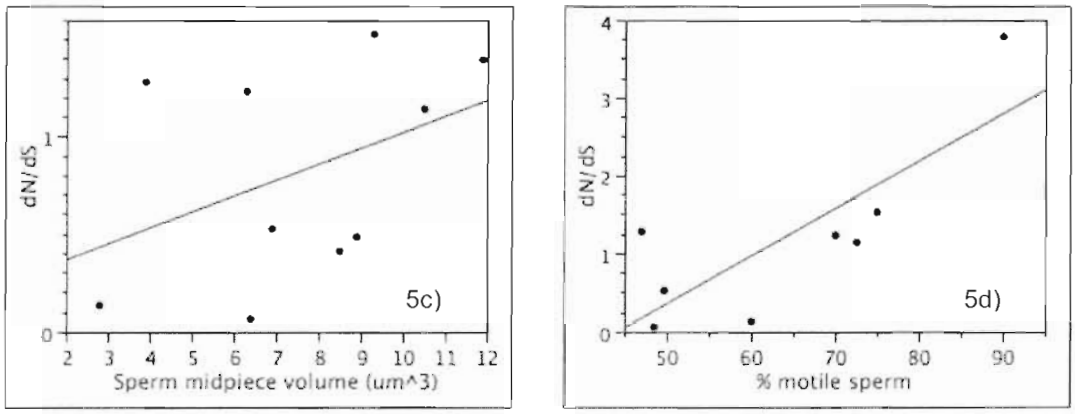


4g) GSTM4 d_N/d_S estimates compared to i) female remating rate: $r^2 = 0.95$, $p = 0.024$; ii) testes mass: $r^2 = 0.85$, $p = 0.075$.

Figures 5a-d. Results of species-level regressions for CatSper1.

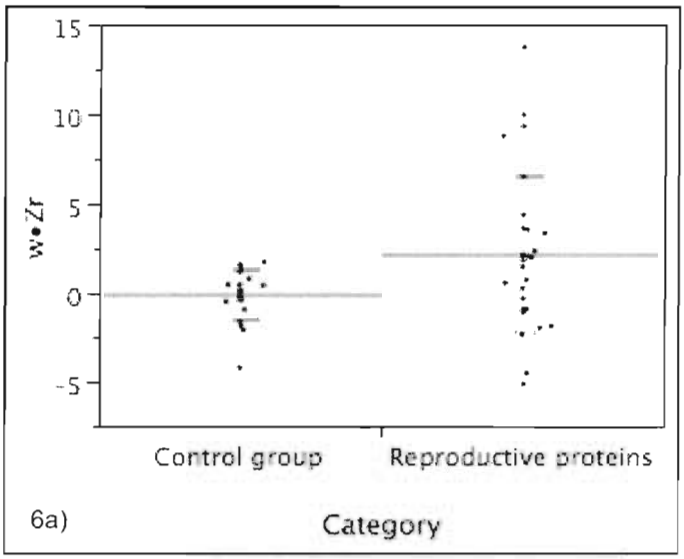


5 a,b. Species-level, linear regressions comparing CatSper1 d_N/d_S estimates to a) female mating rate: $r^2 = 0.029$, $p = 0.56$; b) testes mass: $r^2 = 0.0011$, $p = 0.91$.

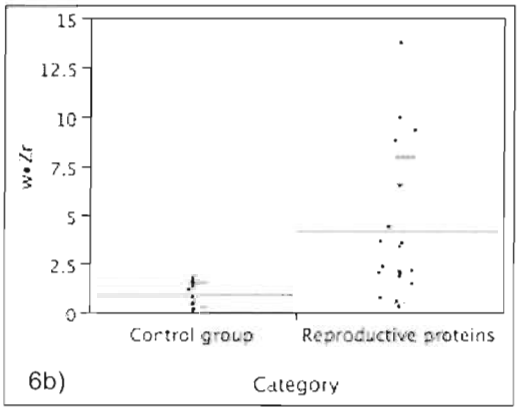


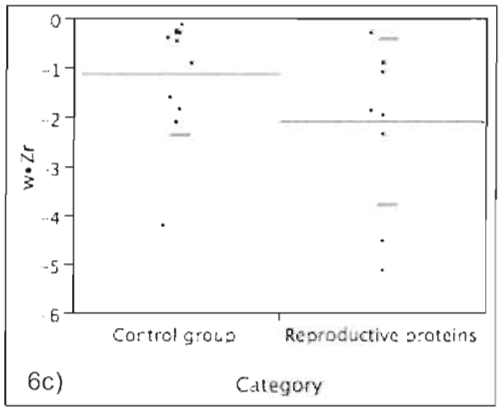
5 c,d. Species-level, linear regressions comparing CatSper1 d_N/d_S estimates to c) sperm midpiece volume (μm) : $r^2 = 0.18$, $p = 0.22$; d) percent motile sperm (%) : $r^2 = 0.63$, $p = 0.018$.

Figures 6a-c. Results of t-tests comparing weighted, standardized correlation coefficients (wZ_r) from control group and reproductive protein phylogenetic-comparative analyses.
Control group: n = 22, reproductive proteins: n = 28. Long lines represent means for each group; short lines mark one standard deviation from the mean in each direction. wZ_r were summarized by using the comparison with the lowest p-value from each protein.



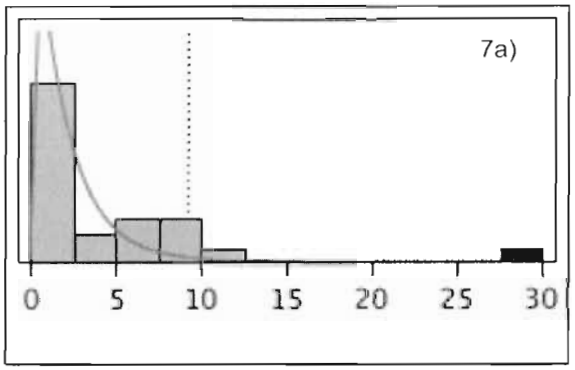
6a) t = 2.58, df = 33.86, two-tailed p = 0.014; control group X = -0.13, σ = 1.41; reproductive proteins X = 2.14, σ = 4.37.



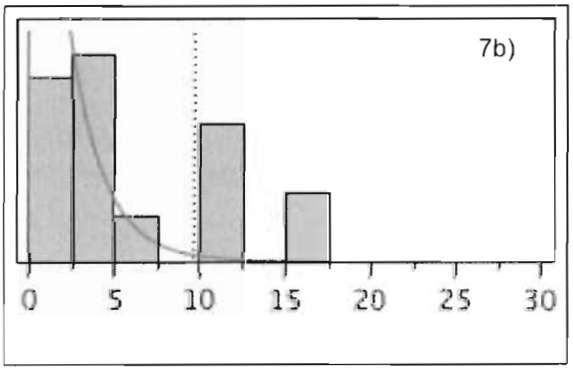


6b, c) Comparison of b) positive wZ_r and c) negative wZ_r between control group and reproductive proteins.
b) t = 3.71, df = 19.62, one-tailed p = 0.0007; control group X = 0.89, n = 11, σ = 0.61; reproductive proteins X = 4.15, n = 19, σ = 3.75.
c) t = -1.46, df = 14.47, one-tailed p = 0.92; control group X = -1.14, n = 11, σ = 1.24; reproductive proteins X = -2.12, n = 9, σ = 1.67.

Figures 7a,b. Frequency distributions of Likelihood Ratio Test (LRT) statistics measuring the likelihood of a correlation between species-specific d_N/d_S estimates and female promiscuity.
Red lines indicate the expected distribution of the test statistic. Dotted lines indicate the threshold beyond which likelihoods are significant at the Bonferroni-corrected α-level, according to the number of comparisons in each group: a) Control group proteins, n = 22; the black bar represents potential outlier, GSTM4. b) Reproductive proteins, n = 28. LRT statistics were summarized by using the statistic with the lowest p-value (i.e., the statistic with the greatest magnitude) from each protein.

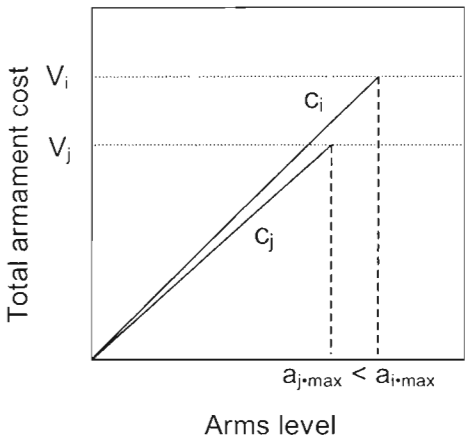


7a) Control group proteins, distribution of LRT statistics (Table 4b).
The distribution is significantly different from the expected null distribution ($W^2 = 0.53$, $p = 0.036$). However, if GSTM4 is excluded as an outlier, the distribution ceases to be significantly different from the expected null distribution ($W^2 = 0.42$, $p = 0.066$).

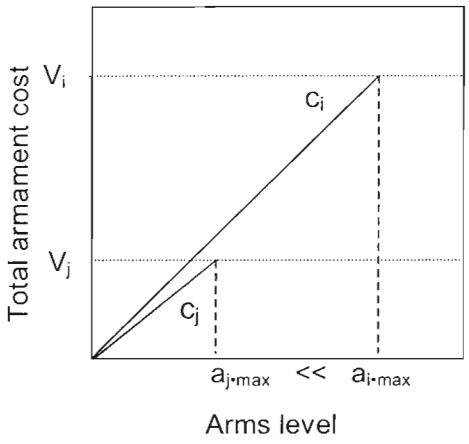


7b) Reproductive group proteins, distribution of LRT statistics (Table 4a).
The distribution is significantly different from the expected null distribution ($W^2 = 2.13$, $p = 0.0010$).

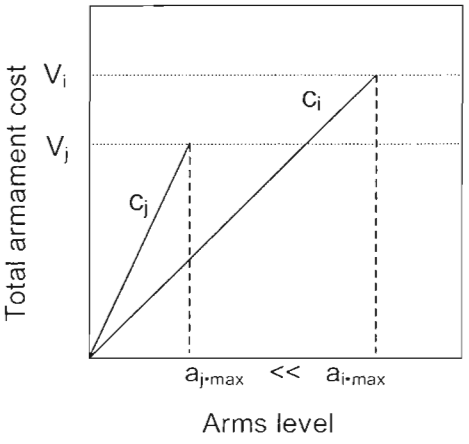
Figures 81.-8.5. Predicted outcomes of antagonistic coevolution between two parties. V_i and V_j are the values of winning a conflict-interaction for parties i and j respectively. These values are in the same currency as total armament cost – when the value of winning exceeds the total cost of maintaining the necessary arms level (represented by the horizontal dotted lines), winning will no longer be profitable. As the value of winning decreases, or as the costs of upgrading to a higher arms level (slope $c_{i,j}$) increase, the maximum possible arms level (a_{max}) will be reached more quickly. The party with the highest arms level ($a_{i,j}$) at any point in time will win the conflict-interaction. Based on Parker's arm race model (2006).



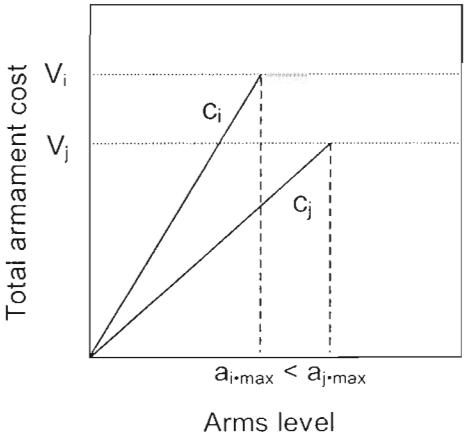
8.1a.
Antagonistic resolution in favor of i.



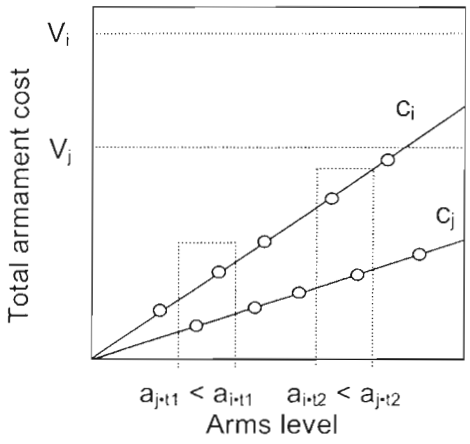
8.1b.
Antagonistic resolution in favor of i; conflict is quickly resolved.



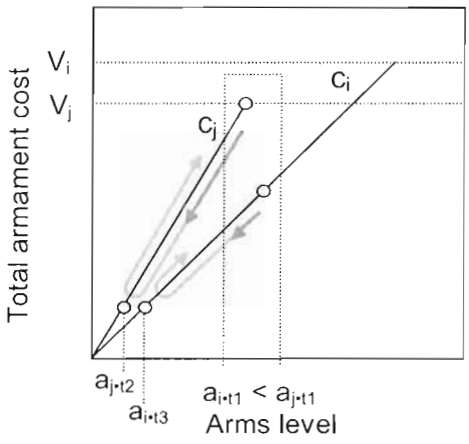
8.1c.
Antagonistic resolution in favor of i; conflict is quickly resolved.



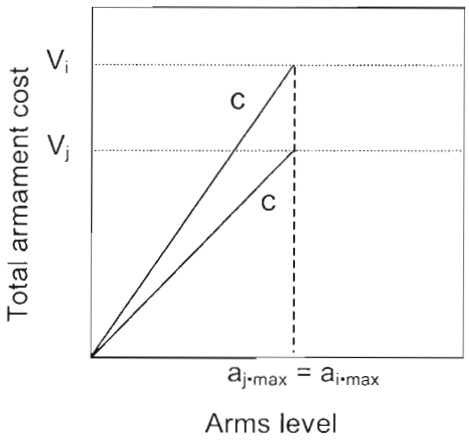
8.1d.
Antagonistic resolution in favor of j.



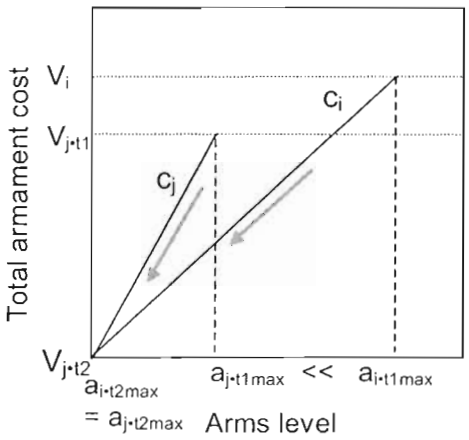
8.2.
Ongoing arms race; winning party alternates through time between i and j.



8.3.
Cycling arms race; i wins conflict (t1), relaxing selection on j to maintain a high arms level; j decreases arms level (t2), decreasing selection on i to maintain a high arms level; i decreases arms level (t3), restoring selection on j to increase arms level.



8.4.
The parties reach a stalemate – sometimes i will win, sometimes j will win.



8.5.
Mutualistic resolution; j is able to decrease the costs of losing completely ($V_j = 0$), which selects for a decrease in the arms level of j, which in turn selects for a decrease in the arms level of i.

TABLES

Table 1. Previous studies linking the adaptive evolution of primate reproductive proteins to mating system.

Function Reproductive protein	Taxa	Method	Evidence of an association between d_N/d_S and mating system	References
Seminal coagulation proteins				
SEMG1, semenogelin I	Hominoids	Within-species sequence polymorphism	YES - Chimpanzees (high sperm competition) show low polymorphism, suggestive of a selective sweep; Gorillas (low sperm competition) show high polymorphism, loss of function	Kingan et al., 2003
	Hominoids, Old World monkeys, New World monkeys, Strepsirrhines	Terminal branch-specific d_N/d_S; Species-level, linear regression	NO - No significant correlation between d_N/d_S and mating system; Increasing trend between d_N/d_S and female remating rate in hominoids: $r^2 = 0.01$, $p > 0.05$	Hurle et al., 2007
SEMG2, semenogelin II	Hominoids, Gibbons, Old World monkeys	Terminal branch-specific d_N/d_S; Closely-related taxa combined to avoid unreliable d_N/d_S estimates caused by short branches; Species-level, linear regression	YES - d_N/d_S are positively correlated with both relative testes mass and female remating rate: $r^2 = 0.52$, $p = 0.035$; $r^2 = 0.98$, $p < 0.0001$; Increasing trend between dN/dS and degree of seminal coagulation	Dorus et al., 2004a
	Hominoids, Gibbons, Old World monkeys	Published d_N/d_S from Dorus et al., (2004); High male:female body mass dimorphism used to indicate low sperm competition; Species-level, linear regression	NO - No significant correlation between d_N/d_S and sexual dimorphism: $r^2 = 0.350$, $p = 0.123$	Herlyn and Zischler, 2007
	Hominoids, Old World monkeys,	Terminal branch-specific d_N/d_S; Species-level, linear regression	NO - No significant correlation between d_N/d_S and mating system	Hurle et al., 2007

	New World monkeys, Strepsirrhines			
Sperm-egg receptor candidates				
Zonadhesin, ZAN	Hominoids, gibbons, Old World monkeys, New World monkeys, Strepsirrhines	Terminal branch-specific d_N/d_S; High male:female body mass dimorphism used to indicate low sperm competition; Species-level, linear regression; Repeated analysis excluding Old World monkeys due to suspected phylogentic correlation	YES - d_N/d_S are negatively correlated with sexual dimorphism: $r^2 = 0.436$, $p = 0.005$; Excluding Old World monkeys: $r^2 = 0.471$, $p = 0.028$	Herlyn and Zischler, 2007
PKDREJ	Lemur, hominoids, Old World monkeys, New World monkeys	Maximum Likelihood-based: Lineage-specific d_N/d_S vs. d_N/d_S of lineages grouped according to expected sperm competition	NO - Model that estimates d_N/d_S according to sperm competition class *not* significantly more likely than model that estimates an individual d_N/d_S for each branch	Hamm et al., 2007

Table 2. Descriptions of primate proteins included in this study.

Protein symbol	Protein name [synonyms]	Expression and localization	Function	References
Reproductive proteins				
Sperm-egg interactions				
OGP	Oviductal Glycoprotein [MUC9]	Secreted by oviduct epithelial cells; associates with oocytes, and developing embryos	Involved in fertilization and protection of oocytes and early embryos; binds to head and midpiece of capacitated sperm	Kadam et al. (2006)
PKDREJ	Polycystic Kidney Disease (polycystin) and sperm Receptor for Egg Jelly homolog, sea urchin (REJ)-like	Testes-specific; localized to the exterior acrosome in spermatozoa	Possibly binds to Zona Pellucida glycoproteins (ZPs), inducing the acrosome reaction	Hamm et al. (2007)
ZAN	Zonadhesin (Zen)	Testes, primarily in spermatocytes; localized to mature spermatozoa head	Binds to Zona Pellucida glycoproteins (ZPs); possibly involved in inducing the acrosome reaction	Lea et al. (2001)
ZP-4	Zona-Pellucida glycoprotein 4 [ZPB]	Oocyte; localized in egg coat	Binds to the acrosomal cap of capacitated spermatozoa; cross-links Zona Pellucida glycoproteins 2 and 3 (ZP2 and ZP3)	Gahlay et al. (2002)
fertilin alpha I	ADAM1 - isoform I [A Disintegrin and a Metalloprotease domain alpha - isoform I]	Testes, during spermatogenesis; localized to spermatozoa head	Possibly involved in sperm-egg interactions; required for fertilization	Evans (2002)
ZP-2	Zona-Pellucida glycoprotein 2 [ZPA]	Oocyte; localized in egg coat	Binds first to the principal segment of capacitated spermatozoa, then to the inner acrosomal membrane and midpiece of acrosome-reacted spermatozoa	Gahlay et al. (2002)
ZP-3	Zona-Pellucida glycoprotein 3 [ZPC]	Oocyte; localized in egg coat	Binds to the head of capacitated (but not acrosome-reacted) spermatozoa, acting as receptor for induction of the acrosome reaction	Gahlay et al. (2002)
fertilin alpha II	ADAM1 - isoform II [A Disintegrin and a Metalloprotease domain alpha - isoform II]	Testes, during spermatogenesis; localized to spermatozoa head	Possibly involved in sperm-egg interactions; required for fertilization	Evans (2002)

Sperm motility				
DBI	Diazepam Binding Inhibitor [Acyl-CoA-binding protein (ACBP); Endozepine]	Prostate (found in seminal fluid); late spermatogenesis sperm, localizes to mitochondrial rich midpiece; also many nonreproductive tissues	Possibly involved in sperm motility; involved in lipid metabolism, steroidogenesis, insulin secretion, receptor modulation	Clark and Swanson (2005); Kolmer et al. (1997)
CatSper1	Cation Channel 1, Sperm [Ca2+ Channel 1, Sperm]	Spermatid membrane; localized in principal piece of spermatozoa tail	Directly influences individual sperm progressive motility; required for male fertility	Li et al. (2006)
Seminal coagulation				
SEMG1	Semenogelin I	Prostate, found in seminal fluid; seminal vesicles; vas deferens; epididymis; trachea	Binds with Semenogelin II (SEMG2) to form seminal coagulum	Jensen-Seaman and Li (2003)
SEMG2	Semenogelin II	Prostate, found in seminal fluid; seminal vesicles; vas deferens; epididymis; trachea	Binds with Semenogelin I (SEMG1) to form seminal coagulum	Jensen-Seaman and Li (2003)
TGM4	Prostate-specific Transglutaminase 4	Prostate-specific	Cross-links Semenogelin I and II (SEMG1 and 2) during seminal coagulation; inhibits Prostate-Specific Antigen (PSA)-mediated dissolution of seminal coagulum	Peter et al. (1998)
Spermatogenesis				
Boule	Boule Protein [BOL-Like; BOLL]	Spermatocyte- and spermatid-specific	Controls the translation of Meiotic Cell Division Cycle 25 (Cdc25) Phosphatase; required for sperm production	Reynolds and Cooke (2005)
SPANX-N2	Sperm Protein Associated with the Nucleus on the X-chromosome	Post-meiotic sperm; localized to nuclear envelope of spermatids; localizes to sperm head and midpiece in mature spermatozoa	Involved in spermiogenesis	Westbrook et al. (2006); Salemi et al. (2004)
DAZL	Deleted in Azoospermia-Like Protein [DAZ-Like]	Germ cell- (mostly spermatid-) specific	RNA-binding; required for sperm production	Reynolds and Cooke (2005)
PRM2	Protamine 2 (P2)	Post-meiotic sperm	Condenses and protects DNA during late spermiogenesis and in mature spermatozoa, replacing Transition Proteins (TNPs)	Cho et al. (2003)
SPAM1	Sperm Adhesion Molecule 1 (PH20)	Spermatid- and epididymis-specific; localizes to sperm surface	Involved in sperm maturation and storage; increases the ability of sperm to penetrate the oocyte's cumulus	Martin-DeLeon (2006)
TGIFLX	TG-interacting Factor	Male germ cell-specific	Regulates transcription during spermatogenesis	Wang and Zhang (2004)

	(TGIF)-Like X-chromosome Protein			
TNP2	Transition Protein 2 (TNP-2)	Post-meiotic sperm	Condenses and protects DNA during early spermiogenesis, replacing histones; necessary for proper processing of Protamines and for proper spermiogenesis	Tseden et al. (2007)
PRM1	Protamine 1 (P1)	Post-meiotic sperm	Condenses and protects DNA during late spermiogenesis and in mature spermatozoa, replacing Transition Proteins (TNPs); necessary for proper processing of Protamine 2	Cho et al. (2003)
TSPY	Testes-specific Protein Y-encoded	Testes-specific; primarily expressed in spermatogonia, and to a lesser extent in spermatids	Involved in spermatogenesis	Kido and Lau (2006)
Dissolution of seminal coagulum/Host defense				
PIP	Prolactin-Induced Protein	High-prostate, found in seminal fluid; also other exocrine glands	Inhibits T-cell apoptosis	Gaubin et al. (1999)
ACPP	Prostatic Acid Phosphatase [PAP; ACP3; ACP-3; PSAP]	Prostate-specific, large amounts are found in seminal fluid	Cleaves Semenogelin I (SEMG1) - cleaved semenogelins protect spermatozoa against microbial attack; cleaves PSA substrates	Brillard-Bourdet et al. (2002)
KLK2	Kallikrein 2	Essentially prostate-specific, found in seminal fluid	Activates Prostate-Specific Antigen (PSA)	Olsson et al. (2005); Lilja (2003)
MSMB	Beta-Microseminoprotein [MSP; Prostate Secretory Protein of 94 Amino Acids (PSP94; PRPS)]	High-prostate (found in seminal fluid); also other mucous tissues	Possible sperm-motility inhibitor; immunoglobulin-binding	Chao et al. (1996); Reeves et al. (2005)
PSA	Prostate-specific Antigen [Kallikrein 3]	Essentially prostate-specific	Dissolves seminal coagulum by cleaving SEMG1 and 2; cleaved semenogelins protect spermatozoa against microbial attack	Olsson et al. (2005); Brillard-Bourdet et al. (2002)
Unclassifiable				
TMPRSS	Transmembrane Serine Protease 2	High-prostate, localized in prostate epithelium	Possibly involved in cell-cell interactions; protease domain is secreted	Vaarala et al. (2001)

Protein symbol	Protein name [synonyms]	Expression and localization	Function	References
Control group proteins				Online Mendelian Inheritance in Man - NCBI (www.ncbi.nlm.nih.gov/)
ACTB	Actin, Beta	Ubiquitous: component of cytoplasmic actin	Regulates cell division, vesicle trafficking and secretion; modulates cell migration during embryogenesis and differentiation	
ALDOA	Aldolase A, fructose-bisphosphate [ALDA; Fructose 1,6-Bisphophate Aldolase A]	Regulated during development; highly -expressed in adult muscle tissue; repressed expression in liver, kidney and intestine	Involved in carbohydrate metabolism	
APEX	Apex Nuclease [Apurinic Endonuclease (APE); Human Apurinic Endonuclease 1 (HAP1); Apurinic/ Apyrimidinic Endonuclease Redox Factor 1 (REF1)]	Ubiquitous	DNA repair enzyme	
ATP5A1	ATP Synthase H+ transporting Mitochondrial F1 Complex Alpha Subunit, isoform 1 [Mitochondrial ATP Synthetase (ATPM); Mitochondrial ATP Synthetase Oligomycin-resistant (OMR); ATP5A]	Ubiquitous: component of mitochondrial ATP Synthase, specifically the catalytic core	Catalyzes ATP synthesis	
CAPN1	Calpain 1 (Calpain large polypeptide L1; Calcium-dependent Protease; Calcium-activated neutral proteinase 1; CANP 1; Capa1)	Ubiquitous	Intracellular protease; mediates cell apoptosis particularly in muscle tissue	
CBR1	Carbonyl Reductase 1 [CBR]	Wide	Involved the metabolism of many carbonyl compounds	
COG7	Component of Oligomeric	Ubiquitous	Critical for the structure and function of the Golgi apparatus; influences	

	Gogli Complex 7		intracellular membrane trafficking
G6PD	Glucose-6-Phosphate Dehydrogenase	Ubiquitous	Critical for the regeneration of NADPH, which protects against and repairs oxidative damage
GSTM4	Glutathione S-Transferase, MU-4 [GSTM4-4; GST-Mu2; MGC131945; MGC9247]	Wide	Involved in cellular detoxification
HSPA8	Heat-Shock 70-KD Protein 8 [Heat-Shock Cognate Protein, 71-KD; HSC71; HSP73; HSC70; formerly HSPA10]	Induced by heat-shock	Chaperone protein
IDH3A	Isocitrate Dehydrogenase 3, Alpha Subunit [Isocitrate Dehydrogenase, NAD(+)-specific, Mitochondrial, Alpha subunit]	Ubiquitous	Key enzyme of the citric acid (Krebs) cycle
KCNK1	Potassium Channel, subfamily K, member 1 [Potassium Channel, Weakly Inward-rectifying with Twin P domains, 1 (TWIK1)]	High-nervous tissue; also peripheral tissues, mostly stomach and small intestine tissue	Involved in muscle contraction, maintenance of action potential, hormone secretion, osmotic regultion, and ion flow
KNSL6	Kinesin-Like 6 [Mitotic Centromere-Associated Kinesin (MCAK); Kinesin Family Member 2C (KIF2C)]	High-thymus, -testis; low-small intestine, -colon, -placenta; very low-spleen and -ovary - found in tissues containing dividing cells	Regulates spindle assembly and chromosome segregation during mitosis and meiosis
NOL5A	Nucleolar Protein 5A [NOP56P; NOP56]	Component of the ribonucleoprotein methylation complex	Critical for nucleotide methylation
PCLY	Prenylcysteine Lyase [PCL1; Prenylcysteine Oxidase 1 (PCYOX1)]	Ubiquitous	Involved in prenylcysteine metabolism

PCNA	Proliferating Cell Nuclear Antigen [DNA Polymerase Delta Auxiliary Protein; Cyclin]	Proliferating cells; highest during the S-Phase of the cell cycle	Involved in DNA replication, particularly repair synthesis
POMT1	Protein O-Mannosyl Transferase 1 [Rotated Abdomen, Drosophila, Homolog of (RT)]	Ubiquitous: maximum-testis, high fetal brain and pituitary; localized to the endoplasmic reticulum	Involved in O-mannosylation protein modification; required for cell wall rigidity
POT1	Protection of Telomeres 1	Ubiquitous	Directly protects chromosome ends
PRDX3	Peroxiredoxin 3 [Antioxidant Protein 1 (AOP1); Mer5]	Ubiquitous: localizes within the mitochondria	Antioxidant protein; required to maintain normal mitochondrial function
PSMD1	Proteasome 26S subunit, Non-ATPase, 1	Ubiquitous: component of proteasome 26S	Involved in the degradation of un-needed proteins
RPL10	Ribosomal Protein L10 [QM gene]	Ubiquitous: high-liver, -spleen, -testis, and -adrenal gland	Involved in transcription regulation
SOAT	Sterol O-Acly Transferase [SOAT1; Sterol Acyltransferase; Acyl-CoA: Cholesterol Acyl Transferase (ACACT; ACAT1)]	Wide: includes macrophages, adrenal glands, and liver	Converts cholesterol to cholesterol esters for storage

Tables 3a,b. Results from species-level regressions, uncorrected for phylogenetic dependency.

p-values marked * are significant at the 0.05 α-level. Negative correlation coefficients are shaded gray.

Table 3a. Reproductive proteins

	Comparison with:							
	Number of mates/periovulatory period				Testes mass			
Protein	Two-tailed	r	r^2	n	Two-tailed	r	r^2	n
Functional category	*p*-value				*p*-value			
Sperm-egg interactions								
OGP	0.13	0.77	0.59	5	0.27	0.61	0.37	5
ln(OGP)	*0.36*	0.53	*0.28*	-	*0.24*	0.64	*0.41*	-
PKDREJ	0.010 *	0.73	0.54	11	*0.55*	0.21	0.045	10
ln(PKDREJ)	*0.011* *	0.73	*0.54*	-	*0.54*	0.22	*0.048*	-
ZAN	0.50	-0.21	0.043	13	0.13	0.48	0.23	11
ln(ZAN)	*0.74*	0.10	*0.011*	-	*0.13*	0.49	*0.24*	-
ZP-4	0.17	0.84	0.70	4	0.98	0.023	0.00052	4
ln(ZP-4)	*0.091*	0.91	*0.83*	-	*0.77*	0.23	*0.053*	-
fertilin alpha I	0.32	0.57	0.32	5	0.22	0.67	0.45	5
ln(fertilin alpha I)	*0.28*	0.60	*0.36*	-	*0.15*	0.75	*0.56*	-
ZP-2	0.93	0.052	0.0027	5	0.82	0.14	0.019	5
ln(ZP-2)	*0.98*	-0.016	*0.00025*	-	*0.94*	-0.051	*0.0026*	-
ZP-3	0.58	0.42	0.18	4	0.84	0.16	0.025	4
ln(ZP-3)	*0.50*	0.50	*0.25*	-	*0.75*	0.25	*0.065*	-
fertilin alpha II	0.58	-0.41	0.17	4	0.45	-0.55	0.30	4
ln(fertilin alpha II)	*0.59*	-0.41	*0.17*	-	*0.44*	-0.57	*0.32*	-
Sperm motility								
DBI	0.21	0.68	0.46	5	0.35	0.66	0.43	4
ln(DBI)	*0.98*	-0.018	*0.00032*	-	*0.067*	0.93	*0.87*	-
CatSper1	0.56	0.17	0.029	14	0.91	-0.033	0.0011	14
ln(CatSper1)	*0.66*	0.13	*0.017*	-	*0.41*	0.24	*0.056*	-
Seminal coagulation								
SEMG1	0.029 *	0.69	0.47	10	0.76	-0.12	0.014	9
ln(SEMG1)	*0.097*	0.56	*0.31*	-	*0.58*	-0.21	*0.045*	-
SEMG2	0.062	0.58	0.34	11	0.56	-0.20	0.039	11
ln(SEMG2)	*0.12*	0.49	*0.24*	-	*0.34*	-0.32	*0.10*	-
TGM4	0.84	-0.079	0.0063	9	0.31	-0.39	0.15	9
ln(TGM4)	*0.90*	-0.055	*0.0030*	-	*0.30*	-0.41	*0.17*	-
Spermatogenesis								
BOULE	0.021 *	0.93	0.87	5	0.034 *	0.91	0.82	5
ln(BOULE)	*0.043* *	0.89	*0.79*	-	*0.019* *	0.94	*0.88*	-
SPANX-N2	0.78	0.18	0.031	5	0.37	0.62	0.39	4
ln(SPANX-N2)	*0.79*	0.16	*0.026*	-	*0.20*	0.80	*0.64*	-
DAZL	0.55	0.31	0.10	6	0.61	0.26	0.070	6

ln(DAZL)	*0.99*		0.0091	*0.000083*	-	*0.79*	-0.14	*0.020*	-
PRM2	0.39		-0.36	0.13	8	0.69	0.19	0.035	7
ln(PRM2)	*0.49*		-0.28	*0.081*	-	*0.72*	0.17	*0.029*	-
SPAM1	0.32		0.35	0.12	10	0.27	-0.44	0.19	8
ln(SPAM1)	*0.67*		0.15	*0.024*	-	*0.23*	-0.48	*0.23*	-
TGIFLX	0.24		-0.36	0.13	12	0.52	-0.20	0.042	12
ln(TGIFLX)	*0.20*		-0.40	*0.16*	-	*0.51*	-0.21	*0.044*	-
TNP2	0.47		-0.37	0.14	6	0.15	-0.66	0.44	6
ln(TNP2)	*0.45*		-0.39	*0.15*	-	*0.14*	-0.68	*0.46*	-
PRM1	0.0032	*	0.98	0.96	5	0.57	0.44	0.19	5
ln(PRM1)	*0.032*	*	0.91	*0.83*	-	*0.52*	0.48	*0.23*	-
TSPY	0.59		0.41	0.17	4	0.68	-0.25	0.063	5
ln(TSPY)	*0.68*		0.10	*0.010*	-	*0.50*	-0.40	*0.16*	-
Dissolution of seminal coagulum/Host defense									
PIP	0.32		0.37	0.14	4	0.20	0.79	0.63	4
ln(PIP)	*0.46*		0.54	*0.29*	-	*0.57*	0.44	*0.19*	-
ACPP	0.15		0.61	0.37	7	0.48	-0.33	0.11	7
ln(ACPP)	*0.67*		0.20	*0.039*	-	*0.18*	-0.57	*0.32*	-
KLK2	0.76		-0.13	0.016	8	0.86	-0.10	0.0092	6
ln(KLK2)	*0.90*		-0.055	*0.0030*	-	*0.51*	-0.33	*0.11*	-
MSMB	0.96		-0.025	0.00060	6	0.50	0.41	0.17	5
ln(MSMB)	*0.99*		0.0045	*0.000020*	-	*0.46*	0.44	*0.19*	-
PSA	0.24		0.76	0.58	4	0.67	0.50	0.25	3
ln(PSA)	*0.29*		0.71	*0.51*	-	*0.67*	0.50	*0.25*	-
Unclassifiable									
TMPRSS	0.54		-0.32	0.10	6	0.38	-0.45	0.20	6
ln(TMPRSS)	*0.78*		-0.14	*0.021*	-	*0.27*	-0.54	*0.29*	-

Table 3b. Control group proteins

	Comparison with:							
	Number of mates/periovulatory period				Relative testes mass			
Protein	Two-tailed	r	r2	n	Two-tailed	r	r2	n
Functional category	p-value				p-value			
ACTB	0.35	-0.47	0.22	6	0.28	0.32	0.10	6
ln(ACTB)	*0.39*	-0.44	*0.19*	-	*0.25*	0.56	*0.31*	-
ALDOA	0.69	-0.31	0.094	4	0.79	0.21	0.044	4
ln(ALDOA)	*0.73*	-0.27	*0.072*	-	*0.72*	0.28	*0.081*	-
APEX1	0.27	-0.73	0.54	4	0.91	0.09	0.008	4
ln(APEX)	*0.18*	-0.82	*0.67*	-	*0.77*	-0.23	*0.052*	-
ATP5A1	0.88	0.12	0.015	4	0.35	0.36	0.13	4
ln(ATP5A1)	*0.88*	-0.12	*0.014*	-	*0.56*	0.44	*0.19*	-
CAPN1	0.20	-0.80	0.64	4	0.11	-0.89	0.79	4
ln(CAPN1)	*0.041* *	-1.00	*1.00*	*3*	*0.30*	-0.89	*0.80*	*3*
CBR1	0.83	-0.13	0.018	5	0.76	0.19	0.035	5
ln(CBR1)	*0.78*	-0.22	*0.050*	*4*	*0.67*	0.33	*0.11*	*4*
COG7	0.24	0.75	0.57	4	0.51	0.49	0.24	4
ln(COG7)	*0.39*	0.61	*0.37*	-	*0.27*	0.73	*0.54*	-

G6PD	0.79		-0.21	0.045	4	0.67		-0.33	0.11	4
ln(G6PD)	*0.79*		-0.21	*0.045*	-	*0.65*		-0.35	*0.12*	-
GSTM4	0.024	*	-0.97	0.95	4	0.075		-0.92	0.85	4
ln(GSTM4)	*0.083*		-0.92	*0.84*	-	*0.44*		-0.57	*0.32*	-
HSPA8	0.85		0.10	0.010	6	0.44		0.39	0.15	6
ln(HSPA8)	*0.90*		0.075	*0.0056*	*5*	*0.37*		0.52	*0.27*	*5*
IDH3	0.74		0.26	0.067	4	0.28		0.72	0.52	4
ln(IDH3)	*0.93*		0.11	*0.013*	*3*	*0.47*		0.74	*0.55*	*3*
KCNK1	0.53		-0.47	0.22	4	0.83		-0.17	0.03	4
ln(KCNK1)	*0.48*		-0.52	*0.27*	-	*0.93*		0.07	*0.00*	-
KNSL6	0.49		-0.51	0.26	4	0.062		-0.94	0.88	4
ln(KNSL6)	-		*NA*[1]	-	*2*	-		*NA*[1]	-	*2*
NOL5A	0.59		0.41	0.17	4	0.71		-0.29	0.085	4
ln(NOL5A)	*0.64*		0.36	*0.13*	-	*0.65*		-0.35	*0.12*	-
PCLY	0.016		0.98	0.97	4	0.052		0.95	0.90	4
ln(PCLY)	*0.10*		0.90	*0.81*	-	*0.015*	*	0.98	*0.97*	-
PDNAM	1.00		0.0049	0.000024	4	0.30		0.70	0.49	4
ln(PDNAM)	*0.96*		0.045	*0.0020*	-	*0.27*		0.73	*0.53*	-
POMT1	0.78		-0.21	0.046	4	0.35		-0.65	0.42	4
ln(POMT1)	*0.96*		-0.045	*0.0020*	-	*0.52*		-0.48	*0.23*	-
POT1	0.40		0.49	0.24	5	0.085		0.82	0.68	5
ln(POT1)	*0.33*		0.56	*0.31*	-	*0.039*	*	0.90	*0.81*	-
PRDX3	0.58		-0.42	0.18	4	0.72		0.28	0.078	4
ln(PRDX3)	*0.58*		-0.42	*0.18*	-	*0.76*		0.24	*0.059*	-
PSMD1	0.16		0.84	0.70	4	0.23		0.77	0.59	4
ln(PSMD1)	*0.20*		0.81	*0.65*	-	*0.22*		0.78	*0.61*	-
RPL10	0.20		0.80	0.64	4	0.64		0.36	0.13	4
ln(RPL10)	*0.20*		0.80	*0.64*	-	*0.62*		0.39	*0.15*	-
SOAT	0.36		-0.63	0.40	4	0.12		-0.88	0.77	4
ln(SOAT)	*0.41*		-0.59	*0.35*	-	*0.13*		-0.87	*0.76*	-

NA[1]: too few datapoints

Tables 4a,b. Results from phylogenetic-comparative analyses.

p-values marked * are significant and the 0.05 α-level ; p-values marked ** are significant at the Bonferroni-corrected α-level (reproductive proteins: α' = 0.0018; control group proteins: α' = 0.0023). Negative correlation coefficients are shaded gray. Control group proteins GSTM4 is marked as a potential outlier.

Table 4a. Reproductive proteins

	Comparison with:											
	Number of mates/periovulatory period						Testes mass					
Protein	LRT	Two-tailed		r	r^2	n	LRT	Two-tailed		r	r^2	n
Functional category	statistic	p-value					statistic	p-value				
Sperm-egg interactions												
OGP	7.92	0.0049	*	0.89	0.79	5	8.82	0.003	*	0.91	0.83	5
ln(OGP)	*8.37*	*0.0038*	*	*0.90*	*0.81*	-	*10.35*	*0.0013*	**	*0.93*	*0.87*	-
PKDREJ	12.43	0.00042	**	0.82	0.68	11	8.70	0.0032	*	0.76	0.58	10
ln(PKDREJ)	*8.046*	*0.0046*	*	*0.72*	*0.52*	-	*5.99*	*0.014*	*	*0.67*	*0.45*	-
ZAN	2.27	0.13		0.40	0.16	13	7.98	0.0047	*	0.72	0.52	11
ln(ZAN)	*11.24*	*0.00080*	**	*0.76*	*0.58*	-	*16.34*	*$5.3x10^{-5}$*	**	*0.88*	*0.77*	-
ZP-4	6.54	0.011	*	0.90	0.81	4	2.44	0.12		0.68	0.46	4
ln(ZP-4)	*10.89*	*0.00097*	**	*0.97*	*0.93*	-	*4.23*	*0.040*	*	*0.81*	*0.65*	-
fertilin alpha I	0.17	0.68		0.08	0.01	5	0.030	0.86		0.08	0.01	5
ln(fertilin alpha I)	*0.70*	*0.40*		*0.36*	*0.13*	-	*0.39*	*0.53*		*0.27*	*0.08*	-
ZP-2	0.10	0.75		0.14	0.02	5	0.057	0.81		0.11	0.01	5
ln(ZP-2)	*0.082*	*0.77*		*0.13*	*0.02*	-	*0.038*	*0.85*		*0.086*	*0.01*	-
ZP-3	2.97	0.085		-0.72	0.52	4	2.72	0.099		-0.70	0.49	4
ln(ZP-3)	*2.00*	*0.16*		*-0.63*	*0.39*	-	*1.79*	*0.18*		*-0.60*	*0.36*	-
fertilin alpha II	2.35	0.13		-0.67	0.44	4	3.66	0.056		-0.77	0.60	4
ln(fertilin alpha II)	*2.62*	*0.11*		*-0.69*	*0.48*	-	*4.060*	*0.044*	*	*-0.80*	*0.64*	-
Sperm motility												
DBI	15.16	0.00010	**	0.98	0.95	5	4.74	0.029	*	0.83	0.69	4
ln(DBI)	*2.36*	*0.12*		*0.61*	*0.38*	-	*0.22*	*0.64*		*0.23*	*0.05*	-
CatSper1	2.50	0.11		0.40	0.16	14	2.39	0.12		0.40	0.16	14
ln(CatSper1)	*4.62*	*0.032*	*	*0.53*	*0.28*	-	*1.75*	*0.19*		*0.34*	*0.12*	-
Seminal coagulation												
SEMG1	15.67	$7.56x10^{-5}$	**	0.89	0.79	10	9.47	0.0021	*	0.81	0.65	9
ln(SEMG1)	*13.30*	*0.00027*	**	*0.86*	*0.74*	-	*10.40*	*0.0013*	**	*0.83*	*0.68*	-
SEMG2	11.48	0.00070	**	0.80	0.64	11	4.00	0.046	*	0.55	0.30	11
ln(SEMG2)	*7.82*	*0.0052*	*	*0.71*	*0.51*	-	*3.77*	*0.052*		*0.54*	*0.29*	-
TGM4	5.91	0.015	*	-0.69	0.48	9	3.30	0.069		-0.55	0.31	9
ln(TGM4)	*0.24*	*0.63*		*-0.16*	*0.03*	-	*0.90*	*0.34*		*-0.31*	*0.10*	-
Spermatogenesis												
BOULE	11.58	0.00067	**	0.95	0.90	5	7.46	0.0063	*	0.88	0.78	5
ln(BOULE)	*8.91*	*0.0028*	*	*0.91*	*0.83*	-	*10.026*	*0.0015*	**	*0.93*	*0.87*	-
SPANX-N2	3.34	0.068		0.70	0.49	5	2.13	0.14		0.64	0.41	4
ln(SPANX-N2)	*5.63*	*0.018*	*	*0.83*	*0.68*	-	*3.76*	*0.053*		*0.78*	*0.61*	-

DAZL	2.51	0.11		0.58	0.34	6	0.0061	0.94		-0.03	0.00	6
ln(DAZL)	*0.36*	*0.55*		*0.24*	*0.06*	-	*0.0070*	*0.93*		*0.034*	*0.00*	-
PRM2	0.40	0.53		0.22	0.05	8	0.08	0.78		-0.11	0.01	7
ln(PRM2)	*0.68*	*0.41*		*0.29*	*0.08*	-	*0.0025*	*0.96*		*-0.019*	*0.00*	-
SPAM1	0.064	0.80		0.39	0.15	10	0.064	0.8		0.52	0.27	8
ln(SPAM1)	*0.92*	*0.34*		*0.30*	*0.09*	-	*0.064*	*0.8*		*0.48*	*0.23*	-
TGIFLX	2.62	0.11		-0.44	0.20	12	$5.8x10^{-5}$	0.99		-0.0022	0.00	12
ln(TGIFLX)	*2.91*	*0.088*		*-0.46*	*0.22*	-	*0.0041*	*0.95*		*0.018*	*0.00*	-
TNP2	0.53	0.47		-0.29	0.08	6	0.0060	0.94		-0.032	0.00	6
ln(TNP2)	*0.40*	*0.53*		*-0.25*	*0.06*	-	*0.010*	*0.92*		*-0.041*	*0.00*	-
PRM1	3.86	0.049	*	-0.73	0.54	5	0.38	0.54		-0.27	0.07	5
ln(PRM1)	*3.69*	*0.055*		*-0.72*	*0.52*	-	*0.58*	*0.45*		*-0.33*	*0.11*	-
TSPY	0.22	0.64		0.23	0.05	4	3.71	0.54		-0.72	0.52	5
ln(TSPY)	*$2.6x10^{-5}$*	*1*		*-0.0025*	*0.00*	-	*4.20*	*0.041*	*	*-0.75*	*0.57*	-
Dissolution of seminal coagulum/Host defense												
PIP	8.43	0.0037	*	0.94	0.88	4	11.47	0.00071	**	0.97	0.94	4
ln(PIP)	*6.10*	*0.014*	*	*0.88*	*0.78*	-	*8.15*	*0.0043*	*	*0.93*	*0.87*	-
ACPP	4.97	0.026	*	0.71	0.51	7	4.58	0.032	*	0.69	0.48	7
ln(ACPP)	*1.79*	*0.18*	*	*0.48*	*0.23*	-	*1.37*	*0.24*		*0.42*	*0.18*	-
KLK2	3.35	0.067		-0.58	0.34	8	3.35	0.067		-0.65	0.43	6
ln(KLK2)	*1.44*	*0.23*		*-0.41*	*0.16*	-	*2.10*	*0.15*		*-0.54*	*0.30*	-
MSMB	0.00	1.00		0.00	0.00	6	0.11	0.74		-0.15	0.02	5
ln(MSMB)	*0.075*	*0.78*		*0.11*	*0.01*	-	*0.018*	*0.89*		*-0.060*	*0.00*	-
PSA	1.23	0.27		0.51	0.26	4	NA[1]	-		-	-	3
ln(PSA)	*0.93*	*0.34*		*0.46*	*0.21*	-	*NA[1]*	-		-	-	-
Unclassifiable												
TMPRSS	0.76	0.34		0.34	0.12	6	0.47	0.49		0.27	0.07	6
ln(TMPRSS)	*2.15*	*0.14*		*0.55*	*0.30*	-	*1.95*	*0.16*		*0.53*	*0.28*	-

NA[1]: < 4 species; phylogenetic analysis not possible

Table 4b. Control group proteins

	Comparison with:										
	Number of mates/periovulatory period						Relative testes mass				
Protein Functional category	LRT statistic	Two-tailed p-value		r	r^2	n	LRT statistic	Two-tailed p-value		r	r^2	n
ACTB	5.51	0.019	*	-0.78	0.60	6	2.55	0.11		0.59	0.35	6
ln(ACTB)	*5.21*	*0.022*	*	*-0.76*	*0.58*	-	*2.69*	*0.10*		*0.60*	*0.36*	-
ALDOA	0.21	0.65		-0.22	0.05	4	0.27	0.61		-0.25	0.06	4
ln(ALDOA)	*0.34*	*0.56*		*-0.29*	*0.08*	-	*0.36*	*0.55*		*-0.29*	*0.09*	-
APEX1	9.42	0.0022	*	-0.95	0.90	4	4.17	0.041	*	-0.80	0.65	4
ln(APEX)	*4.98*	*0.026*	*	*-0.84*	*0.71*	-	*3.13*	*0.077*		*-0.74*	*0.54*	-
ATP5A1	0.018	0.89		0.066	0.00	4	0.071	0.79		0.13	0.02	4
ln(ATP5A1)	*0.019*	*0.89*		*-0.069*	*0.00*	-	*0.020*	*0.89*		*-0.070*	*0.00*	-
CAPN1	3.78	0.052		-0.79	0.63	4	7.03	0.008	*	-0.92	0.85	4
ln(CAPN1)	*2.99*	*0.084*		*-0.73*	*0.53*	-	*1.42*	*0.23*		*-0.55*	*0.30*	-
CBR1	0.059	0.81		0.11	0.01	5	0.011	0.92		0.047	0.00	5
ln(CBR1)	*0.22*	*0.64*		*0.21*	*0.04*	-	*0.11*	*0.74*		*0.15*	*0.02*	-

COG7	7.37	0.0066	*	0.92	0.84	4	5.37	0.020	*	0.86	0.74	4
ln(COG7)	*5.28*	*0.022*	*	*0.86*	*0.73*	-	*4.12*	*0.042*	*	*0.80*	*0.64*	-
G6PD	0.37	0.54		-0.30	0.09	4	0.78	0.38		-0.42	0.18	4
ln(G6PD)	*0.42*	*0.52*		*-0.32*	*0.10*	-	*0.85*	*0.36*		*-0.44*	*0.19*	-
GSTM4	28.12	$1.1x10^{-7}$	**	-1.00	1.00	4	15.56	$7.98x10^{-5}$	**	-0.99	0.98	4
ln(GSTM4)	*11.21*	*0.00081*	**	*-0.97*	*0.94*	-	*8.22*	*0.0041*	*	*-0.93*	*0.87*	-
HSPA8	0.13	0.72		0.15	0.02	6	0.093	0.76		0.12	0.02	6
ln(HSPA8)	*0.078*	*0.78*		*-0.11*	*0.01*	-	*0.0022*	*0.96*		*0.019*	*0.00*	-
IDH3	0.82	0.37		0.43	0.18	4	0.53	0.47		0.35	0.12	4
ln(IDH3)	*0.63*	*0.43*		*0.38*	*0.14*	-	*0.27*	*0.60*		*0.26*	*0.07*	-
KCNK1	0.39	0.53		-0.31	0.09	4	0.62	0.43		-0.38	0.14	4
ln(KCNK1)	*0.31*	*0.57*		*-0.28*	*0.08*	-	*0.38*	*0.54*		*-0.30*	*0.09*	-
KNSL6	8.92	0.0028	*	-0.94	0.89	4	11.41	0.00073	**	-0.97	0.94	4
ln(KNSL6)	*0.051*	*0.82*		*-0.11*	*0.01*	-	*0.19*	*0.66*		*-0.22*	*0.05*	-
NOL5A	0.04	0.83		0.10	0.01	4	0.02	0.88		0.078	0.01	4
ln(NOL5A)	*0.012*	*0.91*		*0.055*	*0.00*	-	$6.8x10^{-5}$	*0.99*		*0.0041*	*0.00*	-
PCLY	4.01	0.045	*	0.80	0.63	4	1.93	0.16		0.62	0.38	4
ln(PCLY)	*4.59*	*0.032*	*	*0.83*	*0.68*	-	*2.21*	*0.14*		*0.65*	*0.42*	-
PDNAM	1.02	0.31		NA^2	-	-	1.54	0.21		NA^2	-	-
ln(PDNAM)	*1.60*	*0.21*		*0.57*	*0.33*	-	*2.27*	*0.13*		*0.66*	*0.43*	-
POMT1	2.99	0.084		-0.73	0.53	4	2.38	0.12		-0.67	0.45	4
ln(POMT1)	*1.69*	*0.19*		*-0.59*	*0.34*	-	*1.27*	*0.26*		*-0.52*	*0.27*	-
POT1	1.20	0.27		0.46	0.21	5	1.05	0.30		0.44	0.19	5
ln(POT1)	*2.10*	*0.15*		*0.59*	*0.34*	-	*1.86*	*0.17*		*0.56*	*0.31*	-
PRDX3	0.07	0.79		-0.13	0.02	4	0.043	0.84		-0.10	0.01	4
ln(PRDX3)	*0.014*	*0.91*		*-0.059*	*0.00*	-	*0.018*	*0.89*		*-0.068*	*0.00*	-
PSMD1	7.58	0.0059	*	0.91	0.84	4	5.49	0.019	*	NA^2	-	-
ln(PSMD1)	*7.22*	*0.0072*	*	*0.91*	*0.84*	-	*5.28*	*0.022*	*	*0.86*	*0.73*	-
RPL10	8.70	0.0032	*	0.94	0.89	4	6.51	0.011	*	0.90	0.80	4
ln(RPL10)	*8.74*	*0.0031*	*	*0.94*	*0.89*	-	*6.48*	*0.011*	*	*0.90*	*0.80*	-
SOAT	0.26	0.61		-0.25	0.06	4	0.014	0.91		-0.059	0.00	4
ln(SOAT)	*0.12*	*0.72*		*-0.17*	*0.03*	-	*0.0024*	*0.96*		*0.024*	*0.00*	-

NA^2: Variance of untransformed $d_N/d_S = 0$

Table 5. Summary of mean correlation coefficients (r_z) from species-level regressions.

'd_N/d_S' versus '$\ln(d_N/d_S)$' indicates whether or not the natural logarithms of the d_N/d_S estimates were used in the comparisons. 'M' versus 'T' indicates whether or not the d_N/d_S estimates were compared to residual(female remating rate), or residual(testes mass). 'lowest *p*-value' summarizes results across these four possible comparisons by using the comparison with the lowest *p*-value from each protein. Rows in bold script indicate that the average r_z is significant.

Category (no. proteins)	mean r_z	upper 95% CI	lower 95% CI	r_z^2	two-tailed *p*-value
Reproductive proteins (28)					
$d_N d_S$•M (28)	**0.28**	**0.44**	**0.11**	**0.08**	**0.0011**
Sperm-egg interactions	**0.36**	**0.63**	**-0.0049**	**0.13**	**0.027**
Seminal coagulation	**0.47**	**0.73**	**0.083**	**0.22**	**0.0095**
Sperm motility	0.27	0.67	-0.26	0.071	0.17
Spermatogenesis	0.17	0.47	-0.17	0.028	0.16
Dissolution of seminal coagulum/Host defense	0.25	0.65	-0.27	0.060	0.18
$\ln(d_N d_S)$•M (28)	**0.22**	**0.39**	**0.038**	**0.048**	**0.0090**
Sperm-egg interactions	**0.44**	**0.69**	**0.094**	**0.19**	**0.0070**
Seminal coagulation	**0.38**	**0.68**	**-0.030**	**0.14**	**0.034**
Sperm motility	0.11	0.57	-0.41	0.012	0.35
Spermatogenesis	0.028	0.35	-0.30	0.00077	0.44
Dissolution of seminal coagulum/Host defense	0.14	0.58	-0.36	0.021	0.30
$d_N d_S$•T (27)	0.021	0.21	-0.17	0.00042	0.42
Sperm-egg interactions	**0.34**	**0.64**	**-0.043**	**0.12**	**0.040**
Seminal coagulation	-0.23	0.19	-0.58	0.052	0.86
Sperm motility	0.035	0.54	-0.49	0.0012	0.45
Spermatogenesis	-0.028	0.31	-0.36	0.00080	0.56
Dissolution of seminal coagulum/Host defense (4)	0.029	0.57	-0.53	0.00085	0.46
$\ln(d_N d_S)$•T (27)	0.0060	0.20	-0.19	0.000036	0.48
Sperm-egg interactions	**0.36**	**0.65**	**-0.023**	**0.13**	**0.032**
Seminal coagulation	-0.31	0.10	-0.64	0.10	0.93
Sperm motility	0.35	0.73	-0.20	0.12	0.11
Spermatogenesis	-0.065	0.28	-0.39	0.0042	0.64
Dissolution of seminal coagulum/Host defense (4)	-0.22	0.38	-0.69	0.047	0.76
lowest *p*-value (28)	**0.31**	**0.46**	**0.13**	**0.093**	**0.00048**
Sperm-egg interactions	**0.60**	**0.79**	**0.30**	**0.36**	**0.00018**
Seminal coagulation	**0.039**	**0.68**	**-0.019**	**0.15**	**0.0310**
Sperm motility	**0.43**	**0.76**	**-0.081**	**0.19**	**0.048**

Spermatogenesis	-0.043	0.28	-0.36	0.0018	0.60
Dissolution of seminal coagulum/Host defense	0.41	0.77	-0.15	0.17	0.074
Control group proteins (22)					
$d_N d_S$•M (22)					
with outlier	-0.039	0.32	-0.39	0.0015	0.42
without outlier	0.040	0.39	-0.33	0.0016	0.58
ln($d_N d_S$)•M (19)					
with outlier	-0.022	0.38	-0.41	0.00049	0.46
without outlier	0.051	0.45	-0.36	0.0026	0.59
$d_N d_S$•T (22)					
with outlier	0.13	0.46	-0.23	0.017	0.76
without outlier	0.19	0.52	-0.18	0.038	0.85
ln($d_N d_S$)•T (19)					
with outlier	0.40	0.68	0.013	0.16	0.98
without outlier	0.44	0.71	0.052	0.19	0.99
lowest *p*-value (21)					
with outlier	0.23	0.55	-0.16	0.052	0.88
without outlier	0.32	0.62	-0.067	0.10	0.95

Table 6. Summary of mean correlation coefficients (r_z) from phylo-comparative analyses.

'd_N/d_S' versus '$\ln(d_N/d_S)$' indicates whether or not the natural logarithms of the d_N/d_S estimates were used in the comparisons. 'M' versus 'T' indicates whether or not the d_N/d_S estimates were compared to residua(female remating rate), or residual(testes mass). 'lowest *p*-value' summarizes results across these four possible comparisons by using the comparison with the lowest *p*-value from each protein. Rows in bold script indicate that the average r_z is significant.

Category (no. proteins)	mean r_z	upper 95% CI	lower 95% CI	r_z^2	two-tailed p-value
Reproductive proteins (28)					
$d_N d_S$•M (28)	**0.40**	**0.54**	**0.23**	**0.16**	**3.6×10^{-6}**
Sperm-egg interactions	**0.55**	**0.76**	**0.23**	**0.30**	**0.0007**
Seminal coagulation	**0.57**	**0.79**	**0.22**	**0.33**	**0.0015**
Sperm motility	**0.61**	**0.85**	**0.16**	**0.37**	**0.006**
Spermatogenesis	0.13	0.44	-0.20	0.02	0.22
Dissolution of seminal coagulum/Host defense	0.18	0.61	-0.33	0.03	0.25
$\ln(d_N d_S)$•M (28)	**0.43**	**0.57**	**0.27**	**0.18**	**6.0×10^{-7}**
Sperm-egg interactions	**0.67**	**0.83**	**0.40**	**0.44**	**1.5×10^{-5}**
Seminal coagulation	**0.62**	**0.82**	**0.29**	**0.38**	**0.00047**
Sperm motility	**0.54**	**0.82**	**0.07**	**0.30**	**0.014**
Spermatogenesis	0.078	0.39	-0.25	0.0061	0.32
Dissolution of seminal coagulum/Host defense	0.15	0.59	-0.35	0.023	0.28
$d_N d_S$•T (27)	**0.36**	**0.52**	**0.18**	**0.13**	**6.5×10^{-5}**
Sperm-egg interactions	**0.60**	**0.80**	**0.28**	**0.36**	**0.00036**
Seminal coagulation	**0.41**	**0.70**	**0.0030**	**0.16**	**0.024**
Sperm motility	**0.45**	**0.78**	**-0.081**	**0.20**	**0.047**
Spermatogenesis	0.11	0.43	-0.24	0.012	0.27
Dissolution of seminal coagulum/Host defense (4)	0.28	0.72	-0.32	0.078	0.18
$\ln(d_N d_S)$•T (27)	**0.40**	**0.55**	**0.23**	**0.16**	**8.5×10^{-6}**
Sperm-egg interactions	**0.69**	**0.85**	**0.41**	**0.47**	**2.0×10^{-5}**
Seminal coagulation	**0.49**	**0.74**	**0.10**	**0.24**	**0.007**
Sperm motility	0.33	0.72	-0.22	0.11	0.11
Spermatogenesis	0.14	0.46	-0.21	0.02	0.21
Dissolution of seminal coagulum/Host defense (4)	0.15	0.65	-0.44	0.023	0.31
lowest *p*-value (28)	**0.50**	**0.62**	**0.34**	**0.25**	**5.8×10^{-9}**
Sperm-egg interactions	**0.77**	**0.88**	**0.57**	**0.59**	**6.0×10^{-8}**
Seminal coagulation	**0.57**	**0.79**	**0.22**	**0.33**	**0.0015**
Sperm motility	**0.68**	**0.88**	**0.29**	**0.47**	**0.0013**
Spermatogenesis	0.068	0.38	-0.26	0.0046	0.34
Dissolution of seminal	0.32	0.73	-0.26	0.10	0.14

coagulum/Host defense					
Control group proteins (22)					
$d_N d_S$•M (21)					
with outlier	-0.22	0.16	-0.53	0.047	0.13
without outlier	-0.065	0.31	-0.42	0.0043	0.37
ln($d_N d_S$)•M (22)					
with outlier	-0.034	0.32	-0.38	0.0011	0.43
without outlier	0.042	0.40	-0.32	0.0018	0.59
$d_N d_S$•T (20)					
with outlier	-0.079	0.30	-0.43	0.0062	0.34
without outlier	0.023	0.39	-0.35	0.00053	0.55
ln($d_N d_S$)•T (22)					
with outlier	0.15	0.48	-0.22	0.021	0.78
without outlier	0.21	0.53	-0.16	0.045	0.87
lowest *p*-value (22)					
with outlier	-0.20	0.17	-0.52	0.04	0.15
without outlier	-0.051	0.31	-0.40	0.0026	0.40

Table 7. Results of independent *t*-tests comparing wZ_r of control group versus reproductive proteins.

7 a. Phylo-comparative analysis

Comparison	Control group mean wZ_r (n)	Reproductive proteins mean (n)	*t* ratio	df	*p*-value
dNdS•M	**-0.28 (21)**	**1.70 (28)**	**2.58**	**37.92**	**0.014**
ln(dNdS)•M	**-0.043 (22)**	**1.84 (28)**	**2.74**	**34.48**	**0.0098**
dNdS•T	**-0.10 (20)**	**1.43 (27)**	**2.46**	**35.97**	**0.019**
ln(dNdS)•T	**0.19 (22)**	**1.60 (27)**	**2.27**	**31.16**	**0.030**
lowest *p*-value	**-0.13 (22)**	**2.14 (28)**	**2.58**	**33.86**	**0.014**

7 b. Species-level, linear regression analysis

Comparison	Control group mean wZ_r (n)	Reproductive proteins mean (n)	*t* ratio	df	*p*-value
dNdS•M	**0.050 (22)**	**1.16 (28)**	**2.35**	**37.57**	**0.024**
ln(dNdS)•M	**-0.026 (19)**	**0.89 (28)**	**2.060**	**38.70**	**0.046**
dNdS•T	0.17 (22)	0.074 (28)	-0.25	47.16	0.80
ln(dNdS)•T	0.51 (19)	0.022 (28)	-1.12	43.85	0.27
lowest *p*-value	0.20 (22)	1.24 (28)	1.66	39.84	0.11

d_N/d_S' versus 'ln(d_N/d_S)' indicates whether or not the natural logarithms of the d_N/d_S estimates were used in the comparisons. 'M' versus 'T' indicates whether or not the d_N/d_S estimates were compared to residual(female remating rate), or residual(testes mass). 'lowest *p*-value' summarizes results across these four possible comparisons by using the comparison with the lowest *p*-value from each protein. Rows in bold script indicate that the average r_z is significant.

Table 8. Evidence of positive selection for reproductive proteins included in this study.

Proteins are grouped according to known functions. Results from this study indicating a correlation between female promiscuity and protein divergence are summarized in the right-hand column: * $p < 0.05$; ** $p < 0.0015$; ( ) – negative correlation.

Functional category	Reproductive protein	Evidence of positive selection	Taxa	References	Significance of correlation
Sperm-egg interactions					
	OGP	Yes - site-specific	Mammals (including Old World monkeys, human)	Swanson et al. (2001)	**
	PKDREJ	Yes - correlated with functional domains	Lemur, New World monkeys, Old World monkeys, hominids	Hamm et al. (2007)	**
	ZAN	Yes - site-specific	Mammals (including human)	Swanson et al. (2003)	**
		Yes - post-translational modifications	Strepsirrhine, New World monkeys, Old World monkeys, human, mammal-outgroup	Herlyn and Zischler (2005)	
		Yes	New World monkeys, Old World monkeys, hominids	Gasper and Swanson (2006)	
		Yes - site-specific	Strepsirrhines, New World monkeys, Old World monkeys, human, mammal-outgroup	Herlyn and Zischler (2007)	
	ZP-4	No data	-	-	**
	fertilin alpha I	Yes - site-specific, correlated with functional domains	Mammals (including Old World monkey)	Swanson et al. (2003)	
		Suggestive correlation with functional domains	Mammals (including New World monkey, Old World monkeys, hominid)	Civetta (2003)	
	ZP-2	Yes - site-specific	Mammals (including New World monkey, Old World monkey, human)	Swanson et al. (2001)	
	ZP-3	Yes - site-specific, correlated with functional domains	Mammals (including New World monkey, Old World monkey, human)	Swanson et al. (2001)	
		No clear evidence	Fish, birds, mammals (including New World monkey, Old World monkey, human)	Berlin and Smith (2005)	
	fertilin alpha II	Suggestive correlation with functional domains	Mammals (including New World monkey, Old world monkeys)	Civetta (2003)	(*)

Sperm motility					
	DBI	Yes - site specific	New World monkeys, Old World monkeys, gibbons, hominids	Clark and Swanson (2005)	**
	CatSper1	Yes - also positive selection for indels	Lemur, New World monkeys, Old World monkeys, hominids	Podlaha and Zhang (2003)	*
Seminal coagulation					
	SEMG1	Inter- and intra-specific variation in sequence length; dN/dS not significant	Gibbon, hominids	Jensen-Seaman and Li (2003)	**
		Human-chimpanzee: elevated divergence, reduced polymorphism	Hominids	Kingan et al. (2003)	
		Yes	Strepsirrhines, New World monkeys, Old World monkeys, hominids	Hurle et al. (2007)	
	SEMG2	Inter- and intra-specific variation in sequence length; dN/dS not significant	Gibbon, hominids	Jensen-Seaman and Li (2003)	**
		Yes - correlated with functional domains	New World monkey, Old World monkeys, gibbons, hominids	Dorus et al. (2004a)	
		Yes	Strepsirrhines, New World monkeys, Old World monkeys, hominids	Hurle et al. (2007)	
	TGM4	Yes - nonsignificant if human and chimp excluded, also branch-specific	New World monkeys, Old World monkeys, gibbons, hominids	Clark and Swanson (2005)	(*)
Spermatogenesis					
	BOULE	No	Strepsirrhine, New World monkeys, Old World monkey, hominids	Tung et al. (2006)	**
	SPANX-N2	Suggestive increase in both dN and dS compared to non-coding regions	New World monkey, Old World monkey, hominids	Kouprina et al. (2004)	*
	DAZL	No	Strepsirrhine, New World monkeys, Old World monkey, hominids	Tung et al. (2006)	
	PRM2	No clear evidence	Primates (Old World monkeys, hominids), rodents, pecoran ruminants (deer , bovids)	Rooney and Zhang (1999)	
		Yes	Human, chimpanzee	Wyckoff et al. (2000)	
		Yes - site-specific	Mammals (including New World monkeys, Old World monkey, gibbon, hominids)	Swanson et al. (2001)	
	SPAM1	Yes - site-specific	Mammals (including Old World monkey,	Swanson et al. (2003)	

			human)		
	TGIFLX	Yes	Primates	Wang and Zhang (2004)	
	TNP2	Yes	Human, chimpanzee	Wyckoff et al. (2000)	
		Yes - site-specific	Mammals (including Old World monkey, hominids)	Swanson et al. (2001)	
	PRM1	Yes	Primates (Old World monkeys, hominids), rodents, pecoran ruminants (deer and bovids)	Rooney and Zhang (1999)	(*)
		Yes	Human, chimpanzee	Wyckoff et al. (2000)	
		Yes - site-specific	Mammals (including New World monkey, gibbon, hominids)	Swanson et al. (2001)	
	TSPY	Yes	Human, chimpanzee	Wyckoff et al. (2000)	(*)
Dissolution of seminal coagulum/Host defense					
	PIP	Yes	Old World monkey, gibbons, hominids	Clark and Swanson (2005)	**
		Yes	Mammals (including Old World monkeys, hominids)	Kitano et al. (2006)	
	ACPP	Yes - site-specific, suggestive correlation with functional domains, also branch-specific	Old World monkeys, gibbon, hominids	Clark and Swanson (2005)	*
	KLK2	Yes - correlation with functional domains	Lemur, New World monkey, Old World monkeys, hominids	Clark and Swanson (2005)	
	MSMB	Yes - correlation with functional domains	New World monkeys, Old World monkeys, gibbons, hominids	Clark and Swanson (2005)	
	PSA	Yes - branch-specific only	Old World monkeys, gibbon, hominids	Clark and Swanson (2005)	
Unclassifiable					
	TMPRSS	Yes - also suggestive correlation with functional domains	Old World monkeys, gibbons, hominids	Clark and Swanson (2005)	

Table 9. Summary of theoretical work regarding the outcomes of evolutionary conflicts.

Outcome Alternative terms	Type of conflict	Predictions	Assumptions	Reference
Antagonistic resolution				
'One sex wins, other loses'	Intersexual	• sex with the greater arms level wins; if sexes are closely matched outcome is less certain	• arms level will be determined by relative value of winning and costs of manipulation	Parker, 2006
Stable equilibrium	Intersexual	• equilibrium will be near female optimum	• natural selection on female stronger than sexual selection on male	Gavrilets et al., 2001
Population extinction	Intersexual	• male wins, increasing extinction risk	• male-imposed costs associated with sexual conflict will lower overall population fitness	Parker & Partridge, 1998
Joint annihilation	Intragenomic	• driver wins, and population eventually goes extinct	• driver is either sex-linked (eliminating one sex from the population), or has other strongly deleterious effects at fixation	Hurst et al., 1996
	Interlocus (general)	• individual for whom manipulation is more costly will eventually lose	• manipulation is costly	Higashi and Yamamura, 1994
Compromised (i.e., forced) altruism	Parent-offspring	• ESS will be closer to the optimum of the party that is able to manipulate the other at a lower cost, but will always be intermediate	• each party pays both direct and indirect costs because of relatedness	Yamamura and Higashi, 1992
Mutualistic resolution				
	Intersexual	• females evolve tolerance towards male-imposed costs	• natural selection on female sensory system is weak enough that it can be modified	Rowe et al., 2005
	Intragenomic	• asymmetries in genetic transmission	• it is possible to completely removed	Hurst et al., 1996

		that originally created the conflict are removed	asymmetries in genetic transmission	
Extinction of conflict	Intragenomic	• suppressor succeeds in preventing drive, selfish driver is lost; alternatively, driver wins and eventually becomes fixed	• no other selfish genetic driver is able to take advantage of asymmetries in genetic transmission; there are no fitness costs to the genome once the driver is fixed	Hurst et al., 1996
Compromise	Intersexual, parent-offspring	• the optima of both parties will adjust in response to each other's manipulation until the two converge on a common optimum	• the greater the difference in optima between the two parties, the greater the cost of manipulation	Hardling et al., 2001
	Parent-offspring	• equilibrium will be intermediate between parent and offspring optima	• the closer parent and offspring come to one another's optima, the weaker the selction for manipulation	Parker and Macnair, 1979
Stalemate				
	Intersexual	• neither sex is able manipulate each other any further	• natural selection prevents further coevolution	Gavrilets et al., 2001
Stalemate	Intragenomic	• polymorphism at both driver and suppressor loci (there is neither fixation nor loss of the driver)	• action of the driver is not associated with any great fitness costs on the part of the genome, or complete suppression is not possible due to the nature of the driver	Hurst et al., 1996
Ongoing arms race				
Sexual arms race	Intersexual	• arms levels will escalate until the conflict becomes too costly for one of the two sexes	• arms levels depend upon the rate at which the cost of the conflict increases with increases in arms level	Parker, 2006
Coevolutionary arms race	Intersexual	• exaggerated counter-adaptations	• selection on female sensory system is strong, preventing the evolution of female tolerance	Rowe et al., 2005
	Intersexual	• rapid sexually antagonistic	• natural selection is negligible (cost of	Gavrilets et al., 2001

		coevolution	increasing armaments is relatively low)	
Escalation	Intragenomic	• repeated genetic conflicts	• the upper limit of the fitness cost on the population is relatively low	Hurst et al., 1996
Extreme elaboration of display traits	Intersexual	• males decrease cost of indicator traits (making them less honest); females modify preference in a way that increases the cost of indicator traits (making them more honest)	• once honest indicator has been fixed, males are able to adopt strategies that reduce the cost of character expression	Hill, 1994
Evolutionary escalation of manipulation	Parent-offspring	• each party will increase their defense or resistance as long as the benefit of winning is greater than the cost of counter-adaptation	• each party can slightly increase defense or resistance at a low cost	Yamamura and Higashi, 1992
Cycling arms race				
	Intersexual	• periods of escalation followed by periods of de-escalations, which are in turn followed by periods of re-escalation	• once the conflict has been won, there is no longer selection for the loser to maintain arms levels	Parker, 2006
Stable limit cycle	Intersexual	• multiple equilibria pull the conflict in different directions; the outcome may depend upon the starting point	• natural selection limits the range of the conflict	Gavrilets et al., 2001
	Intersexual, parent-offspring	• cyclically fluctuating arms levels	• costs associated with armaments increase slowly	Hardling, 1999
Cyclic evolution	Interlocus (general)	• there is no ESS	• after the conflict has been won, the winner forgets' and loses the armament, beginning the resolution process anew	Higashi and Yamamura, 1994

Table 10. Influence of conflict parameters on the outcome of conflict interactions.

Vi and Vj indicate the value of winning the conflict for parties i and j, respectively; ci and cj indicate the rate at which costs to each party accrue with increasing arms level. The arms level of a counter-adaptation indicates its contest potential; the party with the higher arms level will win the conflict. The relative value of winning versus the costs of armament for each party will determine the maximum arms level each party is capable of - which will in turn determine the eventual winner of the conflict. If counter-adaptations by j increase specificity between i and j, those counter-adaptations are said to be active; if counter-adaptations by j act to decrease the specificity between i and j, those counter-adaptations are said to be passive.

Scenario	**Value of winning conflict** V_i	V_j	**Cost of achieving arms level necessary to win conflict** c_i	c_j	**Value vs. cost of winning** $V_i/c_i : V_j/c_j$	**Maximum arms level** $a_{i\text{-}max} : a_{j\text{-}max}$	**Counter-adaptations by j** Active vs. passive	**Outcome**
1) Antagonistic resolution								
a)	high	high	low	low	>	>	both	i eventually wins
b)	high	low	low	low	>	>	"	i wins; conflict is resolved quickly
c)	high	high	low	high	>	>	"	" "
d)	high	high	high	low	<	<	"	j eventually wins
2) Ongoing arms race (Red Queen process)								
	very high	high	low	very low	?	?	passive	Conflict is not immediately resolved; over time outcome will alternate between i winning and j winning
3) Cycling arms race								
	high	high	low	high	>	>	active	j loses the conflict, relaxing selection for costly counter-adpatations; once j decreases arms level, selection on the arms level of i will decrease, de-escalating the conflict
4) Stalemate								
	high	high	high	high	=	=	both	There is no clear winner: sometimes i will win, sometimes j will win

5) Mutualistic resolution							
high	0 (previously high)	low	high	>	>	passive	j is able to decrease the costs of losing to the point that the two parties are no longer in conflict; may be more likely when it is not possible to de-escalate the conflict

Table 11. Predictions regarding the effect of various conflict parameters on gene flow between populations.

Relative arms levels (a_i and a_j) determine the outcome of a conflict between individuals i and j, such that the party that wins the conflict is said to have the higher arms level. Quantitative counter-adaptations are those for which increases in the magnitude of the trait translate to increases in arms level; if there is no correlation between magnitude and arms level, the counter-adaptation is considered to be qualitative. Because qualitative counter-adaptations are more likely to be population specific, low (L) or high (H) arms levels are given for each population, rather than a comparison of arms levels between populations. If counter-adaptations by j increase specificity between i and j, those counter-adaptations are said to be active; if counter-adaptations by j act to decrease the specificity between i and j, those counter-adaptations are said to be passive. Divergence cost (k) measures the fitness cost of interactions between parties from different populations due to disparity in arms levels - k_i measures the cost when i has the higher arms level; k_j measures the cost when j has the higher arms level. Gene flow indicates the extent and direction of expected gene flow between the two populations, assuming the equal importance of all determining factors (data shown in the middle column - interactions marked by '>' will tend to promote gene flow, interactions marked by '<' will tend to discourage gene flow, and those marked by '=' will neither promote nor discourage gene flow). The predictions in this table assume that i and j are reproductively co-dependent on one another; if only i is reproductively dependent on j, but not vice versa, the effect of 'Contest outcomes' on gene flow will be slightly reduced.

Scenario		**Relative arms levels - outcomes**			**Contest outcomes**		**Migrant vs. resident Contest success**				**Migrant vs. resident Hybridization success**				**Divergence cost**		**Gene flow?**	**Reproductive isolation?**
	Nature of counter-adaptations	a_{i1}:a_{j1}	a_{i2}:a_{j2}	a_1:a_2			Pop' 1		Pop' 2		Pop' 1		Pop' 2					
		i_1 x j_1	i_2 x j_2	1 x 2	i_2 x j_1	i_1 x j_2	i_2:i_1	j_2:j_1	i_1:i_2	j_1:j_2	i_2:i_1	j_2:j_1	i_1:i_2	j_1:j_2	k_i	k_j		Speciation?
	Quantitative/ similar across populations																	
1)	a)	>	>	=	>	>	=	=	=	=	=	=	=	=	<	<	2 <••••••> 1	No
	b)	<	<	=	<	<	=	=	=	=	=	=	=	=	<	<	2 <••> 1	No
2)	a)	>	>	<	>	<	=	>	<	=	=	=	=	=	0	0	2 •••> 1	No
	b)	<	<	<	>	<	>	=	=	<	=	=	=	=	0	0	2 •••> 1	No
	c)	<	>	<	>	<	>	=	<	=	=	=	=	=	0	0	2 •••> 1	No

	d)	>	<	<	>	<	=	>	=	<	=	=	=	=	0	0	2 •••> 1	No
3)	a)	>	>	<	>	<	=	>	<	=	=	=	=	=	0	0	2 •••> 1	No
	b)	?	?	<	>	<	≥	>	<	≤	=	=	=	=	0	0	2 •••> 1	No
	c)	=	=	<	>	<	>	>	<	<	=	=	=	=	0	0	2 ••> 1	No
4)	a)	>	>	<	>	<	=	>	<	=	=	=	=	=	0	0	2 •••> 1	No
	b)	>	>	<	>	<	=	>	<	=	<	=	=	<	<	0	2 •> 1	Maybe
	c)	>	>	<	>	<	=	>	<	=	<	<	<	<	<	<	No	Yes
Scenario	**Nature of counter-adaptations**	**Relative arms levels - outcomes** $a_{i1}:a_{j1}$ $i_1 \times j_1$	$a_{i2}:a_{j2}$ $i_2 \times j_2$	a_1, a_2	**Contest outcomes** $i_2 \times j_1$	$i_1 \times j_2$	**Migrant vs. resident Contest success** Pop' 1 $i_2:i_1$	$j_2:j_1$	Pop' 2 $i_1:i_2$	$j_1:j_2$	**Migrant vs. resident Hybridization success** Pop' 1 $i_2:i_1$	$j_2:j_1$	Pop' 2 $i_1:i_2$	$j_1:j_2$	**Divergence cost** k_i	k_j	**Gene flow?**	**Reproductive isolation?**
	Qualitative/ population-specific																	
	j active, i active																	
5)	a)	>	>	L, L	=	=	<	=	<	=	=	=	=	=	<	<	2 <••> 1	No
	b)	<	<	L, L	=	=	=	<	=	<	=	=	=	=	<	<	2 <••> 1	No
	c)	>	>	L, H	=	=	<	=	<	=	<	<	<	<	<	<	No	Yes
	d)	<	<	L, H	=	=	=	<	=	<	<	<	<	<	<	<	No	Yes
	j active (i passive)																	
6)	a)	>	>	L, L	>	>	=	=	=	=	=	=	=	=	<	<	2 <••••••> 1	No
	b)	<	<	L, L	>	>	>	<	>	<	=	=	=	=	<	<	2 <••••> 1	No
	c)	>	>	L, H	>	>	=	=	=	=	<	<	<	<	<	<	2 <•> 1	Maybe
	d)	<	<	L, H	>	>	>	<	>	<	<	<	<	<	<	<	No	Yes
	j passive (i active)																	
7)	a)	>	>	L, L	<	<	<	>	<	>	=	=	=	=	>	>	2 <•> 1	Maybe
	b)	<	<	L, L	<	<	=	=	=	=	=	=	=	=	>	>	2 <••> 1	No
	c)	>	>	L, H	<	<	<	>	<	>	<	<	<	<	>	>	No	Yes

d)	<	<	L, H	<	<	=	=	=	=	<	<	<	<	>	>	No	Yes

Table 12. Application of predictions to specific case studies.

1a. Centromeric histones 'win' in both populations. Centromeric satellites are relatively explanded in Pop' 2 compared to Pop' 1; there are no divergence costs.
1b. Same as above, with divergence costs when centromeric drive is very strong.
2. Fig plants 'win' in Pop' 1; in Pop' 2, fig plants and fig wasps reach a stalemate. Female wasp ovipositors and fig plant styles are relatively longer in Pop' 2. Only fig wasps experience the divergence cost of having longer ovipositors than what their hosts are coevolved to – the cost to fig plants is compensated by an abundance of high fitness pollinators produced by neighbouring fig plants.
3a. Males and females are engaged in an arms race – at any one point, either sex may be in the winning role. The complexity of male setae and female dorsal surfaces is relatively greater in Pop' 2; there are no divergence costs.
3b. Same as above, with divergence costs (k) when males have extremely derived setae.

Case study	**Relative arms levels - outcomes**			**Contest outcomes**		**Migrant vs. resident Contest success**				**Migrant vs. resident Hybridization success**				**Divergence cost**		**Gene flow?**	**Reproductive isolation?**
						Pop' 1		Pop' 2		Pop' 1		Pop' 2					
Nature of counter-adaptations	a_{i1}:a_{j1}	a_{i2}:a_{j2}	a_1:a_2														
	i_1 x j_1	i_2 x j_2	1 x 2	i_2 x j_1	i_1 x j_2	i_2:i_1	j_2:j_1	i_1:i_2	j_1:j_2	i_2:i_1	j_2:j_1	i_1:i_2	j_1:j_2	k_i	k_j		Speciation?
Quantitative/ similar across populations																	
1) Centromeric drive Centromere: i; centromeric histone: j																	
a)	<	<	<	>	<	>	=	=	<	=	=	=	=	0	0	2 •••> 1	No
b)	<	<	<	>	<	>	=	=	<	<	=	=	<	<	0	2 •> 1	Maybe
2) Fig-fig wasp system Fig wasp: i; fig: j	<	=	<	>	<	>	=	<	<	<	=	=	=	<	0	2 •> 1	Maybe

Qualitative/ population-specific																		
3) Diving beetles Male: i; female: j																		
a)	?	?	L, H	<	<	<	>	<	>	=	=	=	=	0	0	2 <•> 1	Maybe
b)	?	?	L, H	<	<	<	>	<	>	<	=	=	<	>	0	No	Yes

APPENDIX

Appendix 1. Genbank accession numbers for sequences used in this study.

Category	Protein	Species	Accession #
Control group proteins			
	ACTB	*Cercopithecus aethiops*	AB004047
		Homo sapiens	NM_001101
		Macaca fascicularis	AB170391
		Macaca fuscata	AF209434
		Macaca mulatta	NM_001033084
		Pongo pygmaeus	CR860530
	ALDOA	*Homo sapiens*	NM_184041
		Macaca fascicularis	AB066558
		Pan troglodytes	XR_023822
		Pongo pygmaeus	CR925940
	APEX1	*Ateles geoffroyi*	DQ976593
		Macaca fascicularis	AB171333
		Pan troglodytes	NM_001081485
		Pongo pygmaeus	DQ977483
	ATP5A1	*Homo sapiens*	AK092735
		Macaca fascicularis	AB170693
		Pan troglodytes	XR_023218
		Pongo pygmaeus	CR861028
	CAPN1	*Homo sapiens*	NM_005186
		Macaca fascicularis	AF284440
		Macaca mulatta	XM_001114172
		Pongo pygmaeus	CR925924
	CBR1	*Homo sapiens*	NM_001757
		Macaca fascicularis	AB059654
		Macaca mulatta	XM_001088120
		Pan troglodytes	XM_531449
		Pongo pygmaeus	CR858173
	COG7	*Homo sapiens*	NM_153603
		Macaca fascicularis	AB070114
		Pan troglodytes	XM_001161673

	Pongo pygmaeus	CR857446
G6PD	*Homo sapiens*	NM_000402
	Macaca fuscata	AF208984
	Pan troglodytes	XM_001146640
	Pongo pygmaeus	DQ173570
GSTM4	*Homo sapiens*	BC108729
	Macaca fascicularis	AF200709
	Pan troglodytes	XM_513625
	Pongo pygmaeus	CR859804
HSPA8	*Cercopithecus aethiops*	X73685
	Homo sapiens	NM_006597
	Macaca fascicularis	AB072749
	Macaca mulatta	XM_001108049
	Pan troglodytes	XM_508830
	Pongo pygmaeus	CR861166
IDH3	*Macaca fascicularis*	X87172
	Macaca mulatta	XM_001106839
	Pan troglodytes	XM_001149155
	Pongo pygmaeus	CR860617
KCNK1	*Homo sapiens*	NM_002245
	Macaca mulatta	XM_001112053
	Pan troglodytes	XM_525096
	Pongo pygmaeus	CR858111
KNSL6	*Homo sapiens*	AY026505
	Macaca fascicularis	AB072747
	Macaca mulatta	XM_001093746
	Pan troglodytes	XM_001151208
NOL5A	*Homo sapiens*	NM_006392
	Macaca mulatta	XM_001110561
	Pan troglodytes	XM_514472
	Pongo pygmaeus	CR859194
PCLY	*Homo sapiens*	BC051891
	Macaca fascicularis	AB062961
	Macaca mulatta	XM_001098650
	Pongo pygmaeus	CR860270
PCNA	*Homo sapiens*	NM_182649
	Macaca fascicularis	AF347680
	Macaca mulatta	XM_001115756
	Pan troglodytes	XM_514499
POMT1	*Homo sapiens*	NM_007171
	Macaca mulatta	XM_001118542
	Pan troglodytes	XM_528446
	Pongo pygmaeus	CR857448

POT1	*Homo sapiens*	NM_015450
	Macaca fascicularis	AB066545
	Macaca mulatta	XM_001087702
	Pan troglodytes	XM_519345
	Pongo pygmaeus	CR860078
PRDX3	*Homo sapiens*	BC111397
	Macaca mulatta	BQ807861
	Pan troglodytes	XM_001154135
	Pongo pygmaeus	CR857380
PSMD1	*Homo sapiens*	BC005036
	Macaca mulatta	BQ807960
	Pan troglodytes	XM_526057
	Pongo pygmaeus	CR860782
RPL10	*Homo sapiens*	NM_006013
	Macaca mulatta	XM_001089131
	Pan troglodytes	XM_001158531
	Pongo pygmaeus	CR859565
SOAT	*Cercopithecus aethiops*	AF053336
	Gorilla gorilla	AF354622
	Pan troglodytes	XM_514030
	Pongo pygmaeus	AF354623
Reproductive proteins		
ACPP	*Erythrocebus patas*	DQ150476
	Gorilla gorilla	DQ150471
	Hylobates syndactylus	DQ150473
	Macaca mulatta	DQ150475
	Pan paniscus	DQ150470
	Papio anubis	DQ150474
	Pongo pygmaeus	DQ150472
BOULE	*Macaca mulatta*	XM_001086915
	Microcebus murinus	AJ746579
	Pan paniscus	AJ717405
	Saguinus oedipus	AJ717406
	Saimiri sciureus	AJ717408
CatSper1	*Aotus trivirgatus*	AAQ95776
	Ateles geoffroyif	AAQ95774
	Cercopithecus aethiops	AAQ95780
	Colobus guereza	AAQ95782
	Gorilla gorilla	AAQ95786
	Homo sapiens	AAH32950
	Lemur catta	AAQ95788
	Macaca mulatta	AAQ95779

	Miopithecus talapoin	AAQ95778
	Pan paniscus	AAQ95784
	Papio hamadryas	AAQ95781
	Pongo pygmaeus	AAQ95787
	Saguinus oedipus	AAQ95775
	Saimiri sciureus	AAQ95777
DAZL	*Callithrix jacchus*	AF144131
	Cebus apella	AF053609
	Microcebus murinus	AJ746580
	Pan paniscus	AJ717409
	Saguinus oedipus	AJ717410
	Saimiri sciureus	AJ717411
DBI	*Erythrocebus patas*	DQ150448
	Hylobates syndactylus	DQ150442
	Pan paniscus	DQ150439
	Pongo pygmaeus	DQ150441
	Saguinus labiatus	DQ150450
fertilin alpha I	*Macaca mulatta*	XM_001109414
	Pan troglodytes	XM_509380
	Papio hamadryas	Y15519
	Pongo pygmaeus	Y15491
	Saguinus oedipus	Y15511
fertilin alpha II	*Macaca fascicularis*	X79809
	Pan troglodytes	XR_023088
	Papio hamadryas	Y15520
	Saguinus oedipus	Y15512
KLK2	*Cercopithecus cephus*	DQ150459
	Erythrocebus patas	DQ150458
	Lemur catta	N. Clark, personal communication
	Macaca nigra	DQ150456
	Pan paniscus	DQ150453
	Papio anubis	DQ150457
	Pongo pygmaeus	DQ150454
	Saguinus labiatus	DQ150460
MSMB	*Hylobates syndactylus*	DQ150466
	Macaca mulatta	DQ150467
	Pan troglodytes	DQ150461
	Papio hamadryas	U49786
	Pongo pygmaeus	DQ150464
	Saguinus oedipus	AJ010154, AJ010158, AJ010158
OGP	*Homo sapiens*	NM_002557

	Macaca mulatta	AAB70664
	Macaca radiata	AAQ17078
	Pan troglodytes	XM_513641
	Papio hamadryas	M59903
PIP	*Homo sapiens*	NM_002652
	Hylobates syndactylus	DQ150514
	Macaca mulatta	DQ150516
	Pan troglodytes	DQ150509
PKDREJ	*Ateles geoffroyif*	EF517287
	Callithrix jacchus	EF517290
	Erythrocebus patas	EF517286
	Gorilla gorilla	EF517281
	Homo sapiens	EF517278
	Lagothrix lagothricha	EF517288
	Lemur catta	EF517291
	Macaca nigra	EF517285
	Pan paniscus	EF517280
	Pongo pygmaeus	EF517282
	Saguinus labiatus	EF517289
PRM1	*Ateles seniculus*	L14592
	Gorilla gorilla	AF215709
	Hylobates lar	L14588
	Pongo pygmaeus	AF215710
	Semnopithecus entellus	AF294851
PRM2	*Ateles seniculus*	X71335
	Callithrix jacchus	X85371
	Erythrocebus patas	AF195644
	Homo sapiens	AF215713
	Hylobates lar	X71339
	Macaca mulatta	X71338
	Pongo pygmaeus	X71337
	Semnopithecus entellus	AH010090
PSA	*Cercopithecus cephus*	DQ150484
	Erythrocebus patas	DQ150483
	Gorilla gorilla	DQ150479
	Hylobates gabriellae	DQ150481
	Pan troglodytes	DQ150477
SEMG1	*Callithrix jacchus*	AJ005842
	Cercopithecus aethiops	DP000048
	Colobus guereza	DP000038
	Gorilla gorilla	AY256472
	Homo sapiens	AY256465
	Hylobates klossi	AY256474

	Pan paniscus	AY256471
	Papio anubis	DP000036
	Pongo pygmaeus	AY256473
	Saguinus oedipus	AJ002153
SEMG2	*Ateles geoffroyif*	AY781393
	Cercopithecus aethiops	DP000048
	Colobus guereza	DP000038
	Gorilla gorilla	DP000041
	Homo sapiens	NM_003008
	Hylobates lar	AY781389
	Macaca fascicularis	AY781390
	Macaca nemestrina	AY781391
	Pan troglodytes	DP000037
	Papio anubis	DP000036
	Pongo pygmaeus	DP000045
SPAM1	*Cercopithecus mitis*	DQ437094
	Hylobates lar	DQ437098
	Lemur catta	DQ437084
	Microcebus murinus	DQ437087
	Otolemur crassicaudatus	DQ437082
	Pan troglodytes	XM_527873
	Perodicticus potto	DQ437083
	Propithecus verreauxi	DQ437086
	Pygathrix nemaeus	DQ437091
	Saimiri sciureus	DQ437088
	Varecia variegata	DQ437085
SPANX-N2	*Homo sapiens*	DQ336115
	Macaca mulatta	XM_001086432
	Pan troglodytes	NM_001042629
	Pongo pygmaeus	AY457942
	Saguinus labiatus	AY457945
TGIFLX	*Aotus trivirgatus*	AY449639
	Callithrix jacchus	AY449637
	Cercopithecus aethiops	AY449635
	Gorilla gorilla	AJ345074
	Homo sapiens	NM_139214
	Hylobates lar	AJ345076
	Lagothrix lagothricha	AY449641
	Macaca fascicularis	AJ345079
	Miopithecus talapoin	AJ345077
	Papio hamadryas	AJ345080
	Pongo pygmaeus	AJ345075
	Saimiri sciureus	AY449640

TGM4	*Erythrocebus patas*	DQ150495
	Homo sapiens	NM_003241
	Hylobates syndactylus	DQ150490
	Macaca mulatta	DQ150493
	Pan paniscus	DQ150487
	Pan troglodytes	DQ150486
	Pongo pygmaeus	DQ150489
TMPRSS	*Erythrocebus patas*	DQ150508
	Gorilla gorilla	DQ150501
	Hylobates syndactylus	DQ150504
	Macaca mulatta	DQ150506
	Pan troglodytes	DQ150499
	Pongo pygmaeus	DQ150502
TNP2	*Gorilla gorilla*	AF215718
	Homo sapiens	NM_005425
	Macaca fascicularis	AB169384
	Macaca mulatta	AF215720
	Pan paniscus	AF215717
	Pongo pygmaeus	AF215719
TSPY	*Allenopipthecus nigroviridis*	AY048053
	Miopithecus talapoin	AY048065
	Pan troglodytes	AY958082
	Theropithecus gelada	AF284278
ZAN	*Alouatta belzebul*	DQ910892
	Aotus azarae	DQ910894
	Callicebus cupreus	DQ910893
	Callithrix jacchus	AY428846
	Cercopithecus mitis	DQ910896
	Eulemur fulvus	DQ910886
	Homo sapiens	AF332975
	Lemur catta	DQ910887
	Macaca mulatta	DQ910898
	Microcebus murinus	DQ910889
	Pan paniscus	AY739235
	Papio hamadryas	DQ910899
	Pygathrix nemaeus	DQ910895
	Saguinus fuscicollis	DQ910891
	Saimiri sciureus	DQ910890
	Varecia variegata	DQ910888
ZP-2	*Callithrix jacchus*	Y10767
	Homo sapiens	M90366
	Macaca fascicularis	AY222645
	Macaca mulatta	XM_001091029

	Pan troglodytes	XM_510869
ZP-3	*Callithrix jacchus*	S71825
	Homo sapiens	M60504
	Macaca fascicularis	AY222644
	Pan troglodytes	XM_528035
ZP-4	*Homo sapiens*	NM_021186
	Macaca fascicularis	AY222647
	Pan troglodytes	XM_525105
	Papio cynocephalus	AY222646

REFERENCE LIST

Aguade, M., Miyashita, N. & Langley, C. (1992) Polymorphism and divergence in the *Mst 355* male accessory gland gene region. *Genetics* **132**, 755-770.

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. & Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research* **25**, 3389-3402.

Andersen, N. M. (1993) Classification, phylogeny, and zoogeography of the pond skater genus *Gerris* Fabricius (Hemiptera: Gerridae). *Canadian Journal of Zoology* **71**, 2473-2508.

Anderson, M. J., Nyholt, J. & Dixson, A. F. (2005) Sperm competition and the evolution of sperm midpiece volume in mammals. *Journal of the Zoological Society of London* **267**, 135-142.

Anstett, M.-C., Bronstein, J. L. & Hossaert-McKey, M. (1996) Resource allocation: a conflict in the fig/fig wasp mutualism. *Journal of Evolutionary Biology* **9**, 417-428.

Arnqvist, G. (2004) Sexual conflict and sexual selection: lost in the chase. *Evolution* **58**, 1383-1388.

Arnqvist, G., Edvardsson, M., Friberg, U. & Nilsson, T. (2000) Sexual conflict promotes speciation in insects. *Proceedings of the National Academy of Sciences of the United States of America* **97**, 10460-10464.

Arnqvist, G. & Rowe, L. (2002) Antagonistic coevolution between the sexes in a group of insects. *Nature* **415**, 787-789.

Arnqvist, G. & Rowe, L. (1995) Sexual conflict and arms races between the sexes: a morphological adaptation for control of mating in a female insect. *Proceedings: Biological Sciences* **261**, 123-127.

Beeman, R. W., Friesen, K. S. & Denell, R. E. (1992) Maternal-effect selfish genes in flour beetles. *Science* **256**, 89-92.

Arnqvist, G., Rowe, L. & Krebs, J. (2005) Sexual conflict. Princeton University Press

Arnqvist, G. & Rowe, L. (2002b) Correlated evolution of male and female morphologies in water striders. *Evolution* **56**, 936-947.

Bacigalupe, L. D., Crudgington, H. S., Hunter, F., Moore, A. J. & Snook, R. R. (2007) Sexual conflict does not drive reproductive isolation in experimental populations of *Drosophila pseudoobscura*. *Journal of Evolutionary Biology* **20**, 1763-1771.

Benjamini, Y. & Hochberg, Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B (Methodological)* **57**, 289-300.

Bergsten, J. & Miller, K. B. (2007) Phylogeny of diving beetles reveals a coevolutionary arms race between the sexes. *Public Library of Science ONE* **6**, 1-6.

Berlin, S. & Smith, N. G. C. (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. *BioMed Central Evolutionary Biology* **5**

Boinski, S. (1987) Mating patterns in squirrel monkeys (*Saimiri oerstedi*). *Behavioral Ecology and Sociobiology* **21**, 13-21.

Bonferroni, C. E. (1936) Teoria statistica delle classi e calcolo delle probabilita. *Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commerciali di Firenze* **8**, 3-62.

Brillard-Bourdet, M., Rehault, S., Juliano, L., Ferrer, M., Moreau, T. & Gauthier, F. (2002) Amidolytic activity of prostatic acid phosphatase on human semenogelins and semenogelin-derived synthetic substrates. *European Journal of Biochemistry* **269**, 390-395.

Bustamante, C. D., Fledel-Alon, A., Williamson, S., Nielsen, R., Todd Hubisz, M., Glanowski, S., Tanenbaum, D. M., White, T. J., Sninsky, J. J., Hernandez, R. D., Civello, D., Adams, M. D., Cargill, M. & Clark, A. G. (2005) Natural selection on protein-coding genes in the human genome. *Nature* **437**, 1153-1157.

Campbell, C. J. (2006) Copulation in free-ranging black-handed spider monkeys (*Ateles geoffroyi*). *American Journal of Primatology* **68**, 507-511.

Carlson, A. E., Westenbrock, R. E., Quill, T., Ren, D., Clapham, D. E., Hille, B., Garbers, D. L. & Babcock, D. F. (2003) CatSper1 required for evoked Ca2+ entry and control of flagellar function in sperm. *Proceedings of the*

National Academy of Sciences of the United States of America **100**, 14864-14868.

Chao, C.-F., Chiou, S.-T., Jeng, H. & Chang, W.-C. (1996) The porcine sperm motilty inhibitor is identical to beta-microseminoprotein and is a competitive inhibitor of Na+, K+-ATPase. *Biochemical and Biophysical Research Communications* **218**, 623-628.

Chapman, T., Arnqvist, G., Bangham, J. & Rowe, L. (2003) Sexual Conflict. *Trends in Ecology & Evolution* **1**, 41-47.

Cho, C., Jung-Ha, H., Willis, W. D., Goulding, E. H., Stein, P., Xu, Z., Schultz, R. M., Hecht, N. B. & Eddy, E. M. (2003) Protamine 2 deficiency leads to sperm DNA damage and embryo death in mice. *Biology of Reproduction* **69**, 211-217.

Civetta, A. (2003) Positive selection within sperm-egg adhesion domains of fertilin: an ADAM gene with a potential role in fertilization. *Molecular Biology and Evolution* **20**, 21-29.

Clark, N. L. & Swanson, W. J. (2005) Pervasive adaptive evolution in primate seminal proteins. *Public Library of Science Genetics* **1**, 335-342.

Clark, N. L., Aagaard, J. E. & Swanson, W. J. (2006) Evolution of reproductive proteins from animals and plants. *Reproduction* **131**, 11-22.

Cohen, J. (1992) A power primer. *Quantitative methods in psychology* **112**, 155-159.

Corey, D. M., Dunlap, W. P. & Burke, M. J. (1998) Averaging correlations: expected values and bias in combined Pearson *r*s and Fisher's *z* transformations. *The Journal of General Psychology* **125**, 245-261.

Dawkins, R. (1976) The Selfish Gene (Oxford University Press, New York).

de Ruiter, J. R. (2004) Genetic markers in primate studies: elucidating behavior and its evolution. *International Journal of Primatoloty* **25**, 1173-1189.

Dixson, A. F. (1998) *Primate sexuality: comparative studies of the prosimians, monkeys, apes, and human beings* (Oxford University Press, Oxford, New York).

Dixson, A. F. & Anderson, M. J. (2002) Sexual selection, seminal coagulation and copulatory plug formation in primates. *Folia Primatologica* **73**, 63-69.

Dixson, A. F. & Anderson, M. J. (2004) Sexual behavior, reproductive physiology and sperm competition in male mammals. *Physiology and Behavior* **83**, 361-371.

Dobzhansky, T. (1940) Speciation as a stage in evolutionary divergence. *The American Naturalist* **74**, 312-321.

Dorus, S., Evans, P. D., Wyckoff, G. J., Choi, S. S. & Lahn, B. T. (2004a) Rate of molecular evolution of the seminal protein gene *SEMG2* correlates with levels of female promiscuity. *Nature Genetics* **36**, 1326-1329.

Dorus, S., Vallender, E. J., Evans, P. D., Anderson, J. R., Gilbert, S. L., Mahowald, M., Wyckoff, G. J., Malcom, C. M. & Lahn, B. T. (2004b) Accelerated evolution of nervous system genes in the origin of *Homo sapiens*. *Cell* **119**, 1027-1040.

Evans, J. P. (2002) The molecular basis of sperm-oocyte membrane interactions during mammalian fertilization. *Human Reproduction Update* **8**, 297-311.

Faul, F., Erdfelder, E., Lang, A.-G. & Buchner, A. (2007) G*Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavioral Research Methods* **39**, 175-191.

Felsenstein, J. (1985) Phylogenies and the comparative method. *The American Naturalist* **125**, 1-15.

Fisher, R. A. (1930) *The Genetical Theory of Natural Selection.* Oxford University Press, Oxford

Fishman, L. & Willis, J. H. (2005) A novel meiotic drive locus almost completely distorts segregation in Mimulus (monkeyflower) hybrids. *Genetics* **169**, 347-353.

Frigge, M., Hoaglin, D. C. & Iglewicz, B. (1989) Some implementations of the boxplot. *The American Statistician* **43**, 50-54.

Fuentes, A. (2000) Hylobatid communities: changing views on pair bonding and social organization in hominoids. *Yearbook of Physical Anthropology* **43**, 33-60.

Gage, M. J. G. (2004) Evolution: sexual arms races. *Current Biology* **14**, R378-R380.

Gage, M. J. G., Parker, G. A., Nylin, S. & Wiklund, C. (2002) Sexual selection and speciation in mammals, butterflies and spiders. *Proceedings of the Royal Society of London Series B-Biological Sicences* **269**, 2309-2316.

Gahlay, G. K., Srivastava, N., Govind, C. K. & Gupta, S. K. (2002) Primate recombinant zona pellucida proteins expressed in *Escherichia coli* bind to spermatozoa. *Journal of Reproductive Immunology* **53**, 67-77.

Gasper, J. & Swanson, W. J. (2006) Molecular population genetics of the gene encoding the human fertilization protein zonadhesin reveals rapid adaptive evolution. *The American Journal of Human Genetics* **79**, 820-830.

Gaubin, M., Autiero, M., Basmaciogullari, S., Metivier, D., Misehal, Z., Culerrier, R., Oudin, A., Guardiola, J. & Piatier-Tonneau, D. (1999) Potent inhibition of CD4/TCR-mediated T cell apoptosis by a CD4-binding glycoprotein secreted from breast tumor and seminal vesicle cells. *The Journal of Immunology* **162**, 2631-2638.

Gavrilets, S. (2000) Rapid evolution of reproductive barriers driven by sexual conflict. *Nature* **403**, 886-889.

Gavrilets, S. (2003) Perspective: Models of speciation: What have we learned in 40 years. *Evolution* **57**, 2197-2215.

Gavrilets, S. & Waxman, D. (2002) Sympatric speciation by sexual conflict. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 10533-10538.

Gavrilets, S., Arnqvist, G. & Friberg, U. (2001) The evolution of female mate choice by sexual conflict. *Proceedings of the Royal Society of London Series B - Biological Sciences* **268**, 531-539.

Godfray, H. C. J. (1995) Evolutionary theory of parent-offspring conflict. *Nature* **376**, 133-138.

Gomulkiewicz, R., Nuismer, S. L. & Thompson, J. N. (2003) Coevolution in variable mutualisms. *American Naturalist* **162**, S80-S93.

Goossens, B., Setchell, M., James, S. S., Funk, S. M., Chikhi, L., Abulani, A., Ancrenaz, M., Lackman-Ancrenaz, I. & Bruford, M. W. (2006) Philopatry and reproductive success in Bornean orang-utans (*Pongo pygmaeus*). *Molecular Biology* **15**, 2577-2588.

Haig, D. (1993) Genetic conflicts in human pregnancy. *Quarterly Review of Biology* **68**, 495-532.

Haig, D. (2004) Genomic imprinting and kinship: How good is the evidence. *Annual Review of Genetics* **38**, 553-585.

Haldane, J. B. S. (1992) Disease and evolution (Reprinted from La Ricerca Scientifica Supplemento, Vol 19, pg 1-11, 1949). *Current Science* **63**, 599-604.

Hamm, D., Mautz, B. S., Wolfner, M. F., Aquadro, C. F. & Swanson, W. J. (2007) Evidence of amino acid diversity-enhancing selection within humans and among primates at the candidate sperm-receptor gene *PKDREJ*. *The American Journal of Human Genetics* **81**, 44-52.

Harcourt, A. H. (1991) Sperm competition and the evolution of non-fertilizing sperm in mammals. *Evolution* **45**, 314-328.

Harcourt, A. H., Harvey, P. H., Larson, S. G. & Short, R. V. (1981) Testis weight, body weight and breeding system in primates. *Nature* **293**, 55-57.

Harcourt, A. H., Purvis, A. & Liles, L. (1995) Sperm competition: mating system, not breeding season, affects testes size of primates. *Functional Ecology* **9**, 468-476.

Hardling, R. (1999) Arms races, conflict costs and evolutionary dynamics. *Journal of Theoretical Biology* **196**, 163-167.

Hardling, R. & Bergsten, J. (2006) Nonrandom mating preserves intrasexual polymorphism and stops population differentiation in sexual conflict. *The American Naturalist* **167**, 401-409.

Hardling, R. & Smith, H. G. (2005) Antagonistic coevolution under sexual conflict. *Evolutionary Ecology* **19**, 137-150.

Hardling, R., Smith, H. G., Jormalainen, V. & Tuomi, J. (2001) Resolution of evolutionary conflicts: costly behaviours enforce the evolution of cost-free competition. *Evolutionary Ecology Research* **3**, 829-844.

Harmon, L. J. & Losos, J. B. (2005) The effect of intraspecific sample size on Type I and Type II error rates in comparative studies. *Evolution* **59**, 2705-2710.

Harvey, P. H. & Rambaut, A. Comparative analyses for adaptive radiations. *Philosophical Transactions: Biological Sciences* **355**, 1599-1605.

Hayashi, T. I., Vose, M. & Gavrilets, S. (2007) Genetic differentiation by sexual conflict. *Evolution* 516-529.

Hayes, J. D. & Strange, R. C. (2000) Glutathione S-Transferase polymosphisms and their biological consequences. *Pharmacology* **61**, 154-166.

Henikoff, S. & Malik, H. S. (2002) Selfish drivers. *Nature* **417**, 227-228.

Herlyn, H. & Zischler, H. (2007) Sequence evolution of the sperm ligand zonadhesin correlates negatively with body weight dimorphism in primates. *Evolution* **61**, 289-298.

Higashi, M. & Yamamura, N. (1994) Resolution of evolutionary conflict: a general theory and its applications. *Researches on Population Ecology* **36**, 15-22.

Hill, G. E. (1994) Trait elaboration via adaptive mate choice: sexual conflict in the evolution of signals of male quality. *Ethology Ecology & Evolution* **6**, 351-370.

Hosken, D. J. & Stockley, P. (2004) Sexual selection and genital evolution. *Trends in Ecology and Evolution* **19**

Huang, H., Feng, J., Famulski, J., Rattner, J. B., Liu, S. T., Koo, G. D., Muschel, R., Chan, G. K. T. & Yen, T. J. (2007) Tripin/hSgo2 recruits MCAK to the inner centromere to correct defective kinetochore attachments. *The Journal of Cell Biology* **177**, 413-424.

Hurle, B., Swanson, W. J., Program, N. I. S. C. C. S. & Green, E. D. (2007) Comparative sequence analyses reveal rapid and divergent evolutionary changes of the WFDC locus in the primate lineage. *Genome Research* **17**, 276-286.

Hurst, L. D. & Pomiankowski, A. (1992) Speciation events. *Nature* **359**, 781.

Hurst, G. D. D. & Schilthuizen, M. (1998) Selfish genetic elements and speciation. *Heredity* **80**, 2-8.

Hurst, L. D., Atlan, A. & Bengtsson, B. O. (1996) Genetic conflicts. *Quarterly Review of BiologyY* **71**, 317-364.

Hurst, G. D. D. & Werren, J. H. (2001) The role of selfish genetic elements in eukaryotic evolution. *Nature Reviews Genetics* **2**, 597-606.

Jensen-Seaman, M. I. & Li, W.-H. (2003) Evolution of the hominoid semenogelin genes, the major proteins of ejaculated semen. *Journal of Molecular Evolution* **57**, 261-270.

Johnstone, R. A. & Grafen, A. (1993) Dishonesty and the handicap principle. *Animal Behaviour* **46**, 759-764.

Jousselin, E., Kjellberg, F. & Herre, E. A. (2004) Flower specialization in a passively pollinated monoecious fig: A question of style and stigma.

International Journal of Plant Sciences **165**, 587-593.

Kadam, K. M., Souza, S. J. D., Bandivdekar, A. H. & Natraj, U. (2006) Identification and characterization of oviductal glycoprotein-binding protein partner on gametes: epitopic similarity to non-muscle myosin IIA, MYH 9. *Molecular Human Reproduction* **12**, 275-282.

Kappeler, P. M. (1997) Intrasexual selection and testis size in strepsirhine primates. *Behavioral Ecology* **8**, 10-19.

Kappeler, P. M. (1991) Patterns of sexual dimorphism in body weight among prosimian primates. *Folia Primatologia* **57**, 132-146.

Kido, T. & Lau, Y.-F. C. (2006) The rat TSPY is preferentially expressed in elongated spermatids and interacts with the core histones. *Biochemical and Biophysical Research Communications* **350**, 56-67.

Kidwell, M. G. & Lisch, D. (1997) Transposable elements as sources of variation in animals and plants. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 7704-7711.

Kiester, A. R., Lande, R. & Schemske, D. W. (1984) Models of coevolution and speciation in plants and their pollinators. *The American Naturalist* **124**, 220-243.

Kim, I.-G., Jun, D. Y., Sohn, U. & Kim, Y. H. (1997) Cloning and expression of human mitotic centromere-associated kinesin gene. *Biochimica et Biophysica Acta* **1359**, 181-186.

Kingan, S. B., Tatar, M. & Rand, D. M. (2003) Reduced polymorphism in the chimpanzee semen coagulating protein, semenogelin I. *Journal of Molecular Evolution* **57**, 159-169.

Kitano, T., Tian, W., Umetsu, K., Yuasa, I., Yamazaki, K., Saitou, N. & Osawa, M. (2006) Origin and evolution of gene for prolactin-induced protein. *Gene* **383**, 64-70.

Kolliker, M., Brodie, E. D. I. I. I. & Moore, A. J. (2005) The coadaptation of parental supply and offspring demand. *The American Naturalist* **166**, 506-516.

Kolmer, M., Pelto Huikko, M., Parvinen, M., Hoog, C. & Alho, H. (1997) The transcriptional and translational control of diazepam binding inhibitor expression in rat male germ-line cells. *DNA and Cell Biology* **16**, 59-72.

Kondoh, M. (2001) Co-evolution of nuptial gift and female multiple mating

resulting in diverse breeding systems. *Evolutionary Ecology Research* **3**, 75-89.

Kondoh, M. & Higashi, M. (2000) Reproductive isolation mechanism resulting from resolution of intragenomic conflict. *American Naturalist* **156**, 511-518.

Kouprina, N., Mullokandov, M., Rogozin, I. B., Collins, N. K., Solom, G., Otstot, J., Risinger, J. I., Koonin, E. V., Barrett, J. C., Larionov, V. & Pastan, I. (2004) The SPANX gene family of cancer/testis-specific antigens: rapid evolution and amplification in African great apes and hominids. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 3077-3082.

Lea, I. A., Sivashanmugam, P. & O'Rand, M. G. (2001) Zonadhesin: characterization, localization, and zona pellucida binding. *Biology of Reproduction* **65**, 1691-1700.

Lee, Y.-H. & Vacquier, V. D. (1992) The divergence of species-specific abalone sperm lysin is promoted by positive Darwinian selection. *Biological Bulletin* **182**, 97-104.

Lee, Y. H., Ota, T. & Vacquier, V. D. (1995) Positive selection is a general phenomenon in the evolution of abalone sperm lysin. *Molecular Biology and Evolution* **12**, 231-238.

Lessells, C. M. (2006) The evolutionary outcome of sexual conflict. *Philosophical Transactions of the Royal Society B* **361**, 301-317.

Li, H.-G., Liao, A.-H., Ding, X.-F., Zhou, H. & Xiong, C.-L. (2006) The expression and significance of CATSPER1 in human testis and ejaculated spermatozoa. *Asian Journal of Andrology* **8**, 301-306.

Lilja, H. (2003) Biology of prostate-specific antigen. *Urology* **62**, 27-33.

MacCallum, C. & Hill, E. (2006) Being positive about selection. *Public Library of Science Biology* **4**, 293-295.

Malik, H. S. (2005) *Mimulus* finds centromeres in the driver's seat. *Trends in Ecology and Evolution* **20**

Malik, H. S. & Bayes, J. J. (2006) Genetic conflicts during meiosis and the evolutionary origins of centromere complexity. *Biochemical Society Transactions* **34**, 569-573.

Malik, H. S. & Henikoff, S. (2002) Conflict begets complexity: the evolution of centromeres. *Current Opinion in Genetics & Development* **12**, 711-718.

Malka, O., Shnieor, S., Hefetz, A. & Katzav-Gozansky, T. (2007) Reversible royalty in worker honeybees (*Apis mellifera*) under the queen influence. *Behavioral Ecology and Sociobiology* **61**, 465-473.

Martin, O. Y. & Hosken, D. J. (2003) The evolution of reproductive isolation through sexual conflict. *Nature* **423**, 979-982.

Martin-DeLeon, P. A. (2006) Epididymal SPAM1 and its impact on sperm function. *Molecular and Cellular Endocrinology* **250**, 114-121.

McDonald, J. H. & Kreitman, M. (1991) Adaptive protein evolution at the *Adh* locus in *Drosophila*. *Nature* **351**, 652-654.

Mercot, H., Atlan, A., Jacques, M. & Montchamp-Moreau, C. (1995) Sex-ratio distortion in *Drosophila simulans*: co-occurence of a meiotic driver and a supressor fo drive. *Journal of Evolutionary Biology* **8**, 283-300.

Metz, E. C. & Palumbi, S. R. (1996) Positive selection and sequence rearrangements generate extensive polymorphism in the gamete recognition protein bindin. *Molecular Biology and Evolution* **13**, 397-406.

Miller, K. B. (2003) The phylogeny of diving beetles (Coleoptera: Dytiscidae) and the evolution of sexual conflict. *Biological Journal of the Linnean Society* **79**, 359-388.

Moller, A. P. (1988) Ejaculate quality, testes size and sperm competition in primates. *Journal of Human Evolution* **17**, 479-488.

Morrow, E. H. & Arnqvist, G. (2003) Costly traumatic insemination and a female counter-adaptation in bed bugs. *Proceedings of the Royal Society of London Series B* **270**, 2377-2381.

Morrow, E. H., Pitcher, T. E. & Arnqvist, G. (2003) No evidence that sexual selection is an 'engine of speciation' in birds. *Ecology Letters* **6**, 228-234.

Nakagawa, S. (2004) A farewell to Bonferroni: the problems of low statistical power and publication bias. *Behavioral Ecology* **15**, 1044-1045.

Nielsen, R. & Yang, Z. (1998) Likelihood models for detecting positively selected amino acid sites and aplications to the HIV-1 envelope gene. *Genetics* **148**, 929-936.

Nunn, C. L. (2000) Promiscuity and the primate immune system. *Science* **290**, 1168-1170.

Nunn, C. L. (1999) in *Primate Males*, ed. Kappeler, P. M. (Cambridge University Press, Cambridge), pp. 192-204.

Olsson, A. Y., Bjartell, a., Lilja, H. & Lundwall, A. (2005) Expression of prostate-specific antigen (PSA) and human glandular kallikrein 2 (hK2) in ileum and other extraprostatic tissues. *International Journal of Cancer* **113**, 290-297.

Orr, H. A. & Presgraves, D. C. (2000) Speciation by postzygotic isolation: forces, genes and molecules. *Bioessays* **22**, 1085-1094.

Pagel, M. (1999) Inferring the historical patterns of biological evolution. *Nature* **401**, 877-884.

Pagel, M. (1997) Inferring evolutionary processes from phylogenies. *Zoologica Scripta* **26**, 331-348.

Parker, G. A. (1979) in *Sexual Selection and Reproductive Competition in Insects*, eds. Blum, M. S. & Blum, N. A. Academic Press London

Parker, G. A. (2006) Sexual conflict over mating and fertilization: an overview. *Philosophical Transactions of the Royal Society B* **361**, 235-259.

Parker, G. A. & Partridge, L. (1998) Sexual conflict and speciation. *Philosophical Transactions of the Royal Society of London Series B - Biological Sciences* **353**, 261-274.

Peter, A., Lilja, H., Lundwall, A. & Malm, J. (1998) Semenogelin I and semenogelin II, the major gel-forming proteins in human semen, are substrates for transglutaminase. *European Journal of Biochemistry* **252**, 216-221.

Plavcan, J. M. (2004) in *Sexual selection in primates*, eds. Kappeler, P. M. & van Schaik, C. P. pp. 230-252.

Plavcan, J. M. & van Schaik, C. P. (1997) Intasexual competition and body weight dimorphism in anthropoid primates. *American Journal of Physical Anthropology* **103**, 37-68.

Podlaha, O. & Zhang, J. (2003) Positive selection on protein-length in the evolution of a primate sperm ion channel. *Proceedings of the National Academy of Sciences of the United States of America* **100**, 12241-12246.

Rambaut, A. (2007) Se-Al sequence alignment editor v2.0a11 Carbon.

Reeves, J. R., Xuan, J. W., Arfanis, K., Morin, C., Garde, S. V., Ruiz, M. T., Wisniewski, J., Panchal, C. & Tanner, J. E. (2005) Identification,

purification and characterization of a novel human blood protein with binding affinity fro prostate secretory protein of 94 amino acids. *Biochemical Journal* **385**, 105-114.

Ren, D., Navarro, B., Perez, G., Jackson, A. C., Hsu, s., Shi, Q., Tilly, J. L. & Clapham, D. E. (2001) A sperm ion channel required for sperm motility and male fertility. *Nature* **413**, 603-609.

Reynolds, N. & Cooke, H. J. (2005) Role of the *DAZ* genes in male fertility. *Reproductive BioMedicine Online* **10**, 72-80.

Rice, W. R. & Holland, B. (1997) The enemies within: intergenomic conflict, interlocus contest evolution (ICE), and the intraspecific Red Queen. *Behavioral Ecology and Sociobiology* **41**, 1-10.

Ricklefs, R. E. & Starck, J. M. Applications of phylogenetically independent contrasts: a mixed progress report. *Oikos* **77**, 167-172.

Rohlf, F. J. (2006) A comment on phylogenetic correction. *Evolution* **60**, 1509-1515.

Rooney, A. P. & Zhang, J. (1999) Rapid evolution of a primate sperm protein: relaxation of functional constraint or positive Darwinian selection. *Molecular Biology and Evolution* **16**, 706-710.

Rowe, L. & Day, T. (2006) Detecting sexual conflict and sexually antagonistic coevolution. *Philosophical Transactions of the Royal Society B* **361**, 277-285.

Rowe, L., Cameron, E. & Day, T. (2005) Escalation, retreat, and female indifference as alternative outcomes of sexually antagonistic coevolution. *The American Naturalist* **165**, S5-S18.

Rowe, J. D., Nieves, E. & Listowsky, I. (1997) Subunit diversity and tissue distribution of human glutathione S-transferases: interpretations based on electrospray ionization-MS and peptide sequence-specific antisera. *Biochemical Journal* **325**, 481-486.

Salemi, M., Calogero, A. E., Benedetoo, D. D., Cosentino, A., Barone, N., Rappazo, G. & Vicari, E. (2004) Expression of SPANX proteins in human-ejaculated spermatozoa and sperm precursors. *International Journal of Andrology* **27**, 134-139.

Sawyer, S. A. & Hartl, D. L. (1992) Population genetics of polymorphism and divergence. *Genetics* **132**, 1161-1176.

Schulke, O., Kappeler, P. M. & Zischler, H. (2004) Small testes size despite high extra-pair paternity in the pair-living nocturnal primate *Phaner furcifer*. *Behavioral Ecology and Sociobiology* **55**, 293-301.

Shackleford, T. K. & Goetz, A. T. (2006) Comparative evolutionary psychology of sperm competition. *Journal of Comparative Psychology* **120**, 139-146.

Singh, M., Kumara, H. N., Kumar, M. A., Singh, M. & Cooper, M. (2006) Male influx, infanticide, and female transfer in *Macaca radiata radiata*. *International Journal of Primatology* **27**, 515-528.

Slatkin, M. & Maynard Smith, J. (1979) Models of coevolution. *Quarterly Review of Biology* **54**, 233-263.

Summers, K., McKeon, S., Sellars, J., Keusenkothen, M., Morris, J., Gloeckner, D., Pressley, C., Price, B. & Snow, H. (2003) Parasitic exploitation as an engine of diversity. *Biological Reviews* **78**, 639-675.

Swanson, W. J., Nielsen, R. & Yang, Q. (2003) Pervasive adaptive evolution in mammalian fertilization proteins. *Molecular Biology and Evolution* **20**, 18-20.

Swanson, W. J. & Vacquier, V. D. (2002a) Reproductive protein evolution. *Annual Review of Ecology and Systematics* **33**, 161-179.

Swanson, W. J. & Vacquier, V. D. (2002b) The rapid evolution of reproductive proteins. *Nature Reviews Genetics* **3**, 137-144.

Swanson, W. J., Yang, Z., Wolfner, M. F. & Aquadro, C. F. (2001) Positive Darwinian selection drives the evolution of several female reproductive proteins in mammals. *Proceedings of the National Academy of Sciences of the United States of America* **98**, 2509-2514.

Thoren, S., Lindenfors, P. & Kappeler, P. M. (2006) Phylogenetic analyses of dimorphism in primates: evidence for stronger selection on canine size than on body size. *American Journal of Physical Anthropology* **130**, 50-59.

Tregenza, T., Wedell, N. & Chapman, T. (2006) Introduction. Sexual conflict: a new paradigm. *Philosophical Transactions of the Royal Society B* **361**, 229-234.

Trivers, R. L. (1972) in *Sexual Selection and the Descent of Man, 1871-1971*, ed. Campbell, B. (Chicago, IL. Aldine Publishing Company, pp. 136-179.

Tsaur, S. C. & Wu, C. I. (1997) Positive selection and the molecular evolution of a gene of male reproduction, Acp26Aa of Drosophila. *Molecular Biology*

and Evolution **14**, 544-549.

Tseden, K., Topaloglu, O., Meinardt, A., Dev, A., Adham, I., Muller, C., Wolf, S., Bohm, D., Schluter, G., Engel, W. & Nayernia, K. (2007) Premature translation of transition protein 2 mRNA causes sperm abnormalities and male infertility. *Molecular Reproduction and Development* **74**, 273-279.

Tung, J. Y., Luetjens, C. M., Wistuba, J., Xu, E. Y., Pera, R. A. R. & Gromoll, J. (2006) Evolutionary comparison of the reproductive genes, *DAZL* and *BOULE*, in primates with and without *DAZ*. *Development Genes and Evolution* **216**, 158-168.

Vaarala, M. H., Porvari, K., Kyllonen, A., Lukkarinen, O. & Vihko, P. (2001) The *TMPRSS2* gene encoding transmembrane serine protease is overexpressed in a majority of prostate cancer patients: detection of mutated *TMPRSS2* form in a case of aggressive disease. *International Journal of Cancer* **94**, 705-710.

van Doorn, G. S. & Weissing, F. J. (2006) Sexual conflict and the evolution of female preferences for indicators of male quality. *The American Naturalist* **168**, 742-757.

van Noordwijk, M. A. & van Schaik, C. P. (2004) in *Sexual selection in primates*, eds. Kappeler, P. M. & van Schaik, C. P. (Cambridge University Press, Cambridge) pp. 208-229.

van Schaik, C. P., van Noordwijk, M. A. & Nunn, C. L. (1999) in *Comparative Primate Socioecology*, ed. Lee, P. C. (Cambridge University Press, Cambridge), pp. 204-240.

Wang, X. & Zhang, J. (2004) Rapid evolution of mammalian x-linked testis-expressed homeobox genes. *Genetics* **167**, 879-888.

Weiblen, G. D. & Bush, G. L. (2002) Speciation in fig pollinators and parasites. *Molecular Ecology* **11**, 1573-1578.

Werren, J. H., Windsor, D. & Guo, L. (1995) Distribution of *Wolbachia* among neotropical arthropods. *Proceedings of the Royal Society of London Series B* **262**, 197-204.

Westbrook, V. A., Schoppee, P. D., Vanage, G. R., Klotz, K. L., Diekman, A. B., Flickinger, C. J., Coppola, M. A. & Herr, J. C. (2006) Hominoid-specific *SPANXA/D* genes demonstrate differential expression in individuals and protein localization to a distinct nuclear envelope domain during spermatid morphogenesis. *Molecular Human Reproduction* **12**, 703-716.

West-Eberhard, M. J. (1983) Sexual selection, social competition, and speciation. *Quarterly Review of Biology* **58**, 155-183.

Westneat, D. F. & Sargent, R. C. (1996) Sex and parenting: the effects of sexual conflict and parentage on parental strategies. *Trends in Ecology & Evolution* **11**, 87-91.

Whitlock, M. C. (2005) Combining probability from independent tests: the weighted Z-method is superior to Fisher's approach. *Journal of Evolutionary Biology* **18**, 1368-1373.

Wyckoff, G. J., Wang, W. & Wu, C.-I. (2000) Rapid evolution of male reproductive genes in the descent of man. *Nature* **403**, 304-309.

Yamamura, N. & Higashi, M. (1992) An evolutionary theory of conflict resolution between relatives: altruism, manipulation, compromise. *Evolution* **46**, 1236-1239.

Yang, Z. (1998) Likelihood ratio tests for detecting positive selection and application to primate lysozyme evolution. *Molecular Biology and Evolution* **15**, 568-573.

Yu, D. W., Ridley, J., Jousselin, E., Herre, E. A., Compton, S. G., Cook, J. M., Moore, J. C. & Weiblen, G. D. (2004) Oviposition strategies, host coercion and the stable exploitation of figs by wasps. *Proceedings of the Royal Society of London Series B-Biological Sciences* **271**, 1185-1195.

Zinner, D., Krebs, E., Schrod, A. & Kaumanns, W. (2006) Early sexual maturity in male hamadryas baboons (*Papio hamadryas hamadryas*) and its reproductive implications. *American Journal of Physical Anthropology* **129**, 584-590.