Temporally Variable Selection on Proteolysis-Related Reproductive Tract Proteins in *Drosophila*

Alex Wong,*†,[1] Michael C. Turchin,[1] Mariana F. Wolfner,[1] and Charles F. Aquadro[1]

[1]Department of Molecular Biology and Genetics, Cornell University

†Present address: Department of Biology, Carleton University, Ottawa, Ontario, Canada.

***Corresponding author:** E-mail: alex_wong@carleton.ca.

Associate editor: Willie Swanson

Abstract

In order to gain further insight into the processes underlying rapid reproductive protein evolution, we have conducted a population genetic survey of 44 reproductive tract–expressed proteases, protease inhibitors, and targets of proteolysis in *Drosophila melanogaster* and *Drosophila simulans*. Our findings suggest that positive selection on this group of genes is temporally heterogeneous, with different patterns of selection inferred using tests sensitive at different time scales. Such variation in the strength and targets of selection through time may be expected under models of sexual conflict and/or host–pathogen interaction. Moreover, available functional information concerning the genes that show evidence of selection suggests that both sexual selection and immune processes have been important in the evolutionary history of this group of molecules.

Key words: seminal proteins, sexual conflict, sperm competition, proteolysis, coevolution, immunity.

Introduction

Comparative and population genetic studies have documented positive selection on reproductive tract proteins in a broad range of organisms, including invertebrates, vertebrates, and plants (Clark et al. 2006; Panhuis et al. 2006). Several hypotheses have been put forward to explain the phenomenon of rapid reproductive tract protein evolution in different taxa. These hypotheses include sperm competition, sexual conflict, host–pathogen interactions, avoidance of selfing, or avoidance of interspecific fertilization (reviewed in Nasrallah 2002; Swanson and Vacquier 2002; Takebayashi et al. 2003; Clark et al. 2006; Panhuis et al. 2006; Turner and Hoekstra 2008). Determining the contributions of these different mechanisms toward rapid reproductive protein evolution will require a combination of functional, comparative, and population genetic approaches.

In this study, we investigated patterns of reproductive protein evolution in Drosophila, focusing particularly on proteolysis regulators and targets of proteolysis. Proteolysis regulators include proteases as well as proteins that modulate protease activity, for example, protease inhibitors (PIs) and protease homologs (proteins that resemble proteases in sequence and structure but are thought to have regulatory noncatalytic functions; Ross et al. 2003).

Proteolysis regulators and their targets are of interest for several reasons. First, proteolysis is thought to play an important role in regulating the activity of other reproductive proteins, such as the male accessory gland proteins of Drosophila (Acps; Peng et al. 2005; Ravi Ram et al. 2006), as well as a number of mammalian seminal proteins (Szecsi and Lilja 1993; Malm et al. 2000; de Lamirande 2007). For example, proteolytic cleavage of Drosophila sex peptide (SP), a well-studied Acp, releases its bioactive C-terminal peptide from sperm, allowing this peptide to reach its targets and mediate several postmating responses over the long term (Peng et al. 2005). In addition, several putative cleavage products of the egg-laying prohormone ovulin are capable of inducing ovulation, suggesting that proteolysis of ovulin releases active peptide hormones (Heifetz et al. 2005). At least one protease produced in the accessory gland is necessary for ovulin cleavage (Ravi Ram et al. 2006), and it is thought that female factors are also required (Park and Wolfner 1995).

A second reason for our focus on proteolysis regulators is that expression and proteomic screens have identified transcripts or peptides of many of these genes in both the male and the female reproductive tracts in Drosophila (Swanson et al. 2001, 2004; Lawniczak and Begun 2004; Mueller et al. 2004; Mack et al. 2006; Kelleher et al. 2007; Allen and Spradling 2008; Findlay et al. 2008, 2009; Kapelnikov et al. 2008; Prokupek et al. 2008; Almeida and Desalle 2009). The diversity of these molecules in reproductive tracts suggests that they play important functional roles. Moreover, such proteins represent a promising set of molecules for the study of male–female coevolution due to the presence of both male- and female-derived proteolysis regulators and targets in the female reproductive tract following mating.

Finally, evolutionary considerations point at proteolysis regulators and their targets as interesting objects of study. Several previous studies have documented positive selection and/or rapid rates of duplication for reproductive tract proteolysis regulators (Panhuis et al. 2003; Swanson et al. 2004; Kelleher et al. 2007; Lawniczak and Begun 2007; Almeida and Desalle 2008; Findlay et al. 2008; Wong, Turchin, et al. 2008; Findlay et al. 2009; Kelleher and Markow 2009; Kelleher and Pennington 2009; Kelleher et al. 2011), suggesting that they

may be involved in processes, such as sperm competition, sexual conflict, and host–pathogen interactions. Interestingly, although there is evidence for rapid evolution of individual proteolytic proteins, proteolysis regulators are components of seminal fluid in a broad range of taxa, including not only Drosophila but also other insects (Andrés et al. 2006; Braswell et al. 2006; Sirot et al. 2008) and mammals (Szecsi and Lilja 1993; Malm et al. 2000; Veveris-Lowe et al. 2007). This conservation of protein classes suggests an important reproductive function for proteolysis throughout animals.

Here, we investigate patterns of reproductive protein evolution using polymorphism data from 41 proteolysis regulators and three targets of proteolysis in *Drosophila melanogaster* and *Drosophila simulans*. We find evidence that selection on this group of genes is temporally heterogeneous, suggesting that the processes underlying selection are not constant through time. Such temporal variation can be generated in evolutionary arms races, such as are thought to occur in host–pathogen interactions and in sexual conflict. Interestingly, several genes subject to positive selection have documented or suspected roles in immunity or in sperm storage, further suggesting an important role for host–pathogen interactions and sperm competition in reproductive tract protein evolution.

Materials and Methods

Loci

We surveyed polymorphism at 3 loci encoding known targets of proteolysis, and 41 loci encoding proteolysis regulators—predicted proteases, PIs, or protease homologs. Protease homologs resemble proteases in primary sequence and tertiary structure but carry one or more catalytic site mutations such that they probably lack normal catalytic activity. Nonetheless, protease homologs have been reported to modulate protease activity either as agonists or as antagonists (e.g., Kwon et al. 2000; Lee et al. 2002; Asgari et al. 2003; Zhang et al. 2004; Gupta et al. 2005). Of the 44 loci that we studied, 17 (7 PIs and 10 protease/protease homologs) are known from an expressed-sequence tag screen by Swanson et al. (2004) to be expressed in the somatic portion of the female reproductive tract. Twenty-nine genes (9 PIs, 17 protease/protease homologs, and 3 targets) have strongly male accessory gland-biased expression and were initially identified as accessory gland specific by Swanson et al. (2001). Two genes (*CG10363* and *CG9456*) were identified in both male and female reproductive tract screens. It should be noted that the degree of tissue specificity differs substantially between the male and female samples: The male accessory gland genes were selected on the basis of strong expression bias in the accessory glands (Swanson et al. 2001), and microarray studies examining 14 adult tissues support their high specificity (Chintapalli et al. 2007; FlyAtlas.org). The female reproductive tract genes, by contrast, were selected for expression in the uterus, oviducts and/or sperm storage organs but not for strongly biased expression and thus have varying degrees of tissue specificity. As such, we calculated the tissue specificity measure τ (Yanai et al. 2005) for each gene examined in this study using publicly available microarray data for 14 adult tissues (FlyAtlas.org). τ ranges between 0 and 1, with higher values indicating a higher degree of tissue specificity.

Previous studies have identified six Acps that undergo proteolysis following transfer to the female: SP, ovulin, the sperm storage protein Acp36DE, the protease CG11864, the protease homolog CG9997, and the PI CG9334. Here, we use the term "targets" to refer only to the first three of the six known targets of proteolysis, with the latter three considered under proteases or PIs, respectively. Table 1 lists all 44 loci, with predicted molecular functions, known biological roles, tissue specificity, sample sizes, and the sex in which reproductive tract expression was initially identified.

Drosophila Strains and DNA Sequencing

For polymorphism analyses in *D. melanogaster*, we used chromosome extraction lines for the X, second, and third chromosomes, isolated from isofemale lines derived from an Ugandan population (population samples are described in Pool and Aquadro 2006. *Drosophila simulans* sequences were collected from isofemale lines derived from a Madagascar population. Populations were chosen to reflect ancestral variation in *D. melanogaster* (Uganda: Pool and Aquadro 2006 and *D. simulans* (Madagascar: e.g., Kopp et al. 2006) in order to minimize the confounding effects of population bottlenecks associated with recent colonization events (e.g., Haddrill et al. 2005). For heterozygous sites in sequences from the *D. simulans* isofemale lines, one of the two bases was randomly selected. Phasing of multiple heterozygous bases in a single gene was not required since no gene harbored more than one heterozygous base.

DNA was extracted using the Puregene DNA purification kit (Gentra Systems, Minneapolis, MN). Loci were amplified by polymerase chain reaction (PCR), and PCR products were sequenced using BigDye chemistry (Applied Biosystems, Foster City, CA) on an ABI 3730 automated sequencer at the Cornell University Life Sciences Core Laboratories Center. PCR and sequencing primer sequences are given in the supplementary data, Supplementary Material online. Sequence alignments were performed using the ClustalW algorithm as implemented in CodonCode Aligner (CodonCode Corp., Dedham, MA).

Molecular Population Genetics

Summary statistics (π, θ) were calculated using the Analysis software package, which is based on the libsequence C++ libraries (Thornton 2003). For inferences of selection at loci encoding putative proteolysis regulators and targets of proteolysis, we used two classes of method: those that infer selection on a recent timescale (~0.1 Ne—Przeworski 2003) from the site–frequency spectrum (SFS) and those that infer historical selection using both polymorphism and divergence data. For inference of recent selection, we used Tajima's *D* (Tajima 1989) and Fay and Wu's *H* (Fay and Wu 2000), both of which test predictions concerning specific subsets of the SFS. We also used the clsw

Table 1. Genes Surveyed in this Study.

Gene	Type	Function/Effects	Max. Tissue	τ	Sex	Sample Size *D. melanogaster*	Sample Size *D. simulans*
CG8982 (Acp26Aa, ovulin)	Target	Ovulation	Male acc. gland	1	M	13	14
CG7157 (Acp36DE)	Target	Sperm storage	Male acc. gland	1	M	11	10
CG17673 (Acp70A, sexpeptide)	Target	Remating, eggproduction and laying, feeding	Male acc. gland	1	M	14	12
CG1262 (Acp62F)	PI	Sperm competition, toxic	Male acc. gland	1	M	16	13
CG1342	PI	Unknown	Male acc. gland	NA	M	14	11
CG8137	PI	Toxic	Male acc. gland	1	M	11	14
CG10956	PI	Unknown	Male acc. gland	1	M	14	8
CG31902	PI	Unknown	Male acc. gland	1	M	12	11
CG32203	PI	Unknown	Male acc. gland	1	M	14	13
CG33121	PI	Unknown	Male acc. gland	1	M	14	14
CG9997	Prot. hom.	Remating, eggproduction and laying, sperm release	Male acc. gland	1	M	13	12
CG11864	Prot.	Cleavage of ovulin	Male acc. gland	1	M	12	15
CG6168	Prot.	Immunity	Male acc. gland	1	M	18	12
CG32382	Prot. hom.	Immunity	Male acc. gland	1	M	14	11
CG32383	Prot. hom.	Immunity	Male acc. gland	1	M	14	8
CG1895	Prot.	Unknown	Male acc. gland	1	M	15	14
CG6069	Prot. hom.	Unknown	Male acc. gland	1	M	20	12
CG9806	Prot.	Unknown	Male acc. gland	0.99	M	14	11
CG10586	Prot.	Unknown	Male acc. gland	1	M	14	15
CG10587	Prot.	Unknown	Male acc. gland	1	M	16	13
CG11037	Prot.	Unknown	Male acc. gland	1	M	16	15
CG11664	Prot. hom.	Unknown	Male acc. gland	1	M	14	14
C13518	Prot.	Unknown	Male acc. gland	NA	M	14	13
CG17242	Prot.	Unknown	Male acc. gland	1	M	13	12
CG18557	Prot.	Unknown	Male acc. gland	0.85	M	14	13
CG4847	Prot.	Unknown	Male acc. gland	0.96	M	10	9
CG32833	Prot.	Unknown	Male acc. gland	1	M	14	12
CG9456	PI	Unknown	Male acc. gland	0.96	F + M	14	14
CG10363 (TepIV)	PI	Immunity	Hindgut	0.77	F + M	13	15
CG1857 (necrotic)	PI	Immunity	Fat body	0.8	F	12	12
CG11331	PI	Immunity	Crop	0.86	F	15	14
CG1865	PI	Unknown	Spermathecae	0.7	F	12	13
CG3604	PI	Unknown	Hindgut	0.99	F	16	15
CG18525	PI	Unknown	Spermathecae	0.8	F	15	14
CG3066	Prot.	Immunity	Crop	0.68	F	14	12
CG3074	Prot.	Eggshell matrix	Crop	0.8	F	14	12
CG3097	Prot.	Unknown	Crop	0.91	F	15	13
CG9849	Prot.	Unknown	Salivary gland	0.5	F	13	12
CG9897	Prot. hom.	Unknown	Spermathecae	1	F	7	13
CG13318	Prot. hom.	Unknown	Spermathecae	0.95	F	15	12
CG14642	Prot.	Unknown	Spermathecae	0.86	F	11	12
CG18125	Prot.	Unknown	Spermathecae	1	F	12	12
CG31199	Prot.	Unknown	Eye	0.98	F	13	12
CG31681	Prot.	Unknown	Spermathecae	1	F	14	14

NOTE. —Targets are proteins known to undergo proteolysis following mating. PI: Predicted protease inhibitors. Prot.: Predicted catalytic proteases. Prot. hom.: Predicted protease homologs. "Max. tissue" indicates tissue of highest expression, and τ is the tissue specificity measure of Yanai et al. (2005), with $\tau = 1$ indicating absolute specificity and $\tau = 0$ indicating equal expression in all tissues. "NA" in the τ column indicates that τ cannot be calculated because of low expression levels. "Sex" indicates whether a gene was included because of its identification in screens of the male (M) or female (F) reproductive tracts (Swanson et al. 2001; Swanson et al. 2004).

method of Kim and Stephan (2002), which uses several features of the SFS to increase power and reduce the chance of false positives. We also applied to our data a multilocus version of the Hudson–Kreitman–Aguadé (HKA) test (Hudson et al. 1987; Wright and Charlesworth 2004), which can detect reductions in sequence variability following selective sweeps. As presumably neutral reference loci, we used four noncoding intergenic regions sequenced in the Uganda population of *D. melanogaster* by Pool and Aquadro (2006) (see also Wong, Turchin, et al. 2008) or five noncoding regions sequenced in the Madagascar population of *D. simulans* by Nolte and Schlötterer (2008).

We used the McDonald–Kreitman (MK) test (McDonald and Kreitman 1991) to make inferences about historical selection using a combination of polymorphism and divergence data. We also used our MK data to estimate the rate of adaptive amino acid substitution using the method of Bierne and Eyre-Walker (2004).

Interlocus linkage disequilibrium (LD) parameters were estimated using a custom Perl script (available upon request). For every pair of loci, the correlation coefficient r^2 was calculated for each pair of polymorphic amino acid sites (each with the minor allele represented at least twice in the sample). For each pair of loci, ZnS, the average of all

pairwise values of r^2, was used as a summary measure of interlocus LD (Kelly 1997). We did not analyze pairs of genes sequenced in fewer than five strains. The significance of each interlocus ZnS value was assessed via permutation tests, whereby 10,000 permutations were generated by swapping labels (strain names) on haplotypes. In this way, intralocus haplotype structure was maintained for the permutations, but interlocus LD was randomized.

Results and Discussion

A variety of hypotheses have been proposed to explain the observation that an unusual proportion of reproductive tract proteins in Drosophila and other organisms evolve rapidly and adaptively. Male–female coevolution, sperm competition, and host–pathogen interactions are among the leading proposals (Civetta 2003; Clark et al. 2006; Panhuis et al. 2006; Chapman 2008; Turner and Hoekstra 2008), but it has proven difficult to distinguish between these potential mechanisms. Here, we present results from a molecular population genetic survey of reproductive tract proteolysis regulators and targets of proteolysis in Drosophila. Our findings have implications for the broad understanding of the molecular evolution of reproductive tract proteins.

Patterns of Diversity

We sequenced 44 loci encoding known targets of proteolysis and putative proteolysis regulators in population samples of *D. melanogaster* and *D. simulans*. An average of 13.7 and 12.5 alleles were sequenced for each locus in the *D. melanogaster* and *D. simulans* samples, respectively. Over all loci, average diversity in *D. melanogaster* (π) was 0.007 (standard deviation [SD] = 0.005). Diversity was substantially higher in *D. simulans*, where mean $\pi = 0.015$ (SD = 0.006). Both estimates are similar to those previously documented in the literature (e.g., Andolfatto 2005; Begun et al. 2007). The difference in diversity between the two species was highly significant (paired *t*-test $P = 3.3 \times 10^{-8}$), consistent with a larger effective population size in *D. simulans*. Systematic differences were also observed in the SFS between the two species, with a significantly lower Tajima's *D* in *D. simulans* indicating a relative excess of rare alleles in that species (mean $D = -0.36$ in *D. melanogaster*, -0.87 in *D. simulans*; paired *t*-test $P = 0.00013$).

Inferences of Recent Selection

We used several methods to infer the action of recent directional selection in *D. melanogaster* and *D. simulans*. Tajima's *D* and Fay and Wu's *H* detect departures from the neutral equilibrium model using specific portions of the SFS (an excess of rare alleles and high frequency–derived alleles, respectively), and the more recent compositive likelihood method of Kim and Stephan (2002) (clsw) uses a spatially explicit model of selection to test several features of the standard hitchhiking model. Additionally, the HKA test detects local reductions in polymorphism; in the multilocus version implemented here (Wright and Charlesworth 2004), variation at a gene of interest is compared with variation at multiple presumably neutral noncoding loci.

Tests of the SFS find virtually no evidence of recent selection on any of the 44 loci tested in either *D. melanogaster* or *D. simulans* (fig. 1). At a false discovery rate (FDR; Storey and Tibshirani 2003) of 0.1, no locus in either species shows significant deviations from neutrality using Tajima's *D* test, Fay and Wu's *H* test, or Kim and Stephan's clsw test. These data suggest that selection has not acted on the sampled loci on a time scale that leaves a signature in the SFS.

In *D. simulans*, HKA tests similarly suggest a paucity of recent selection on the sampled genes with no locus rejecting neutrality at an FDR of 0.1. By contrast, we find that 7 loci of 44 show evidence for a local reduction in variation in *D. melanogaster* using the HKA test, consistent with the action of recent selection (fig. 1, table 2). Our inference of selection using the HKA test but not with methods that detect skews in the SFS toward rare alleles (D, clsw) is consistent with very recent selection such that there are insufficient rare variants present in the sequenced region to have power to detect selection.

Inferences of Ancestral Selection

We used MK tests to infer historical selection on protein sequences in each species individually. This test is capable of detecting an excess of amino acid substitutions between two species, suggesting a history of repeated positive selection on protein sequence. Using this approach, we find evidence for selection at multiple loci in each species: At an FDR of 0.1, four loci reject neutrality in *D. melanogaster* (table 2) and nine loci reject neutrality in *D. simulans* (table 3). Thus, although patterns of very recent selection appear to differ for the two species, reproductive tract proteolysis regulators and targets have been subject to selection on a deeper time scale in both species. Notably, the sets of loci showing evidence for selection using the MK test on the one hand and the HKA test on the other are mutually exclusive—that is, no locus rejects neutrality using both tests (table 2).

In order to make a quantitative comparison of the rate of adaptive substitution between species and between sexes, we used the method of Bierne and Eyre-Walker (2004) (fig. 2). Combining data from multiple genes, this method estimates α, the proportion of amino acid substitutions that have been adaptive, under the assumption that synonymous polymorphisms and substitutions are neutral. In our data set, estimates of α show differences both between species and between sexes. Averaging over all loci (i.e., for genes expressed in either sex), α is significantly higher in *D. simulans* than in *D. melanogaster* (*simulans*: $\alpha = 0.52$, 95% confidence interval [CI]: 0.44, 0.59; *melanogaster*: $\alpha = 0.27$, 95% CI: 0.12, 0.40), reflecting the finding that more than twice as many single loci reject neutrality in single gene MK tests. Within *D. melanogaster*, α is higher for male-specific genes than for female reproductive tract genes, although not significantly so (males: $\alpha = 0.34$, 95% CI: 0.18, 0.47; females: $\alpha = -0.04$, 95% CI: -0.47, 0.27). No difference is apparent in *D. simulans* (males: $\alpha = 0.52$,

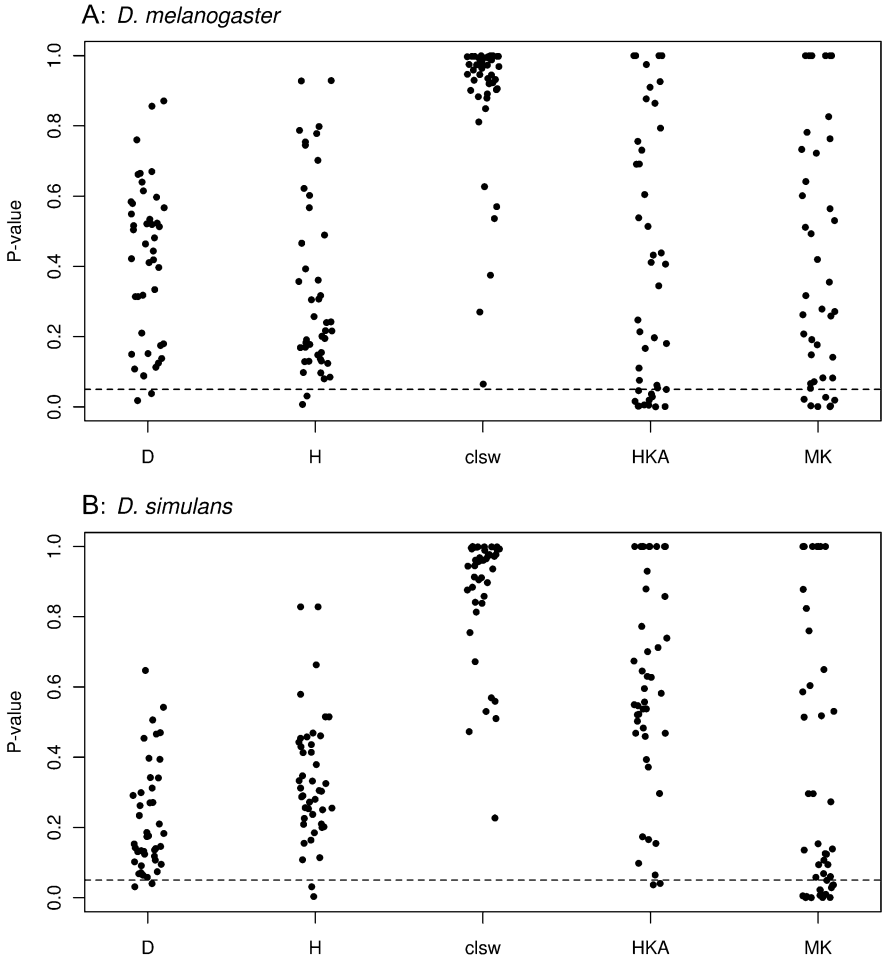

FIG. 1. Temporally variable selection on reproductive tract proteolysis regulators and targets. *P* values are given for all 44 genes surveyed in each species for a variety of neutrality tests (Tajima's *D*, Fay and Wu's *H*, Kim and Stephan's clsw, the HKA test, and the McDonald–Kreitman test). The dotted line represents $P = 0.05$.

95% CI: 0.43, 0.60; females: $\alpha = 0.54$, 95% CI: 0.44, 0.59). Thus, these estimates also suggest variation in the rate of adaptive substitution, particularly for genes expressed in the female reproductive tract.

Temporally Variable Selection

By using statistical tests that are sensitive to selection at different timescales, we can begin to make inferences concerning the consistency—or lack thereof—of selection over time (e.g., Jensen and Bachtrog 2010). Our population genetic data suggest that selection on reproductive tract proteolysis regulators and targets is temporally heterogeneous. On a recent timescale, we find evidence for recent selection on three accessory gland proteins, three female reproductive tract proteins, and one protein present in the reproductive tracts of both sexes in *D. melanogaster* as indicated by a reduction in polymorphism in HKA tests (table 2). By contrast, our *D. simulans* population sample shows no

Table 2. *P* Values for Statistical Tests of Selective Neutrality on Genes Evidence for Positive Selection Along the *D. melanogaster* Lineage.

Gene	Sex	Ontology	Tajima's *D*	Fay and Wu's *H*	clsw	HKA	MK
CG10587	M	Prot.	0.44	0.93	0.98	**0.0049**	0.56
CG11037	M	Prot.	0.11	0.62	0.94	**0.0023**	0.51
CG14642	M	Prot.	0.09	0.2	0.57	**0.00026**	1
CG17242	M	Prot.	0.58	0.13	0.95	0.93	**0.00087**
CG32382	M	Prot. hom.	0.51	0.097	0.85	0.97	**0.0033**
CG8137	M	PI	0.21	0.18	0.97	0.88	**0.00065**
CG8982 (ovulin)	M	Target	0.14	0.031	0.63	0.43	**0.0026**
CG9456	F + M	PI	0.41	0.75	0.89	**0.00093**	0.083
CG18125	F	Prot.	0.15	0.12	0.91	**0.0056**	0.42
CG1865	F	PI	0.11	0.098	0.38	**0.016**	0.067
CG3066	F	Prot.	0.52	0.16	0.96	**0.0022**	1

NOTE. —*P* values in bold are significant at a FDR of 0.1.

Table 3. *P* Values for Statistical Tests of Selective Neutrality on Genes Evidence for Positive Selection Along the *D. simulans* Lineage.

Gene	Sex	Ontology	Tajima's D	Fay and Wu's H	clsw	HKA	MK
CG10363	M/F	PI	0.14	0.47		1	**0.00034**
CG17242	M	Prot.	0.47	0.38	0.99	0.56	**0.022**
CG32203	M	PI	0.039	0.93	1	0.39	**3.1×10^{6}**
CG32833	M	Prot.	0.10	0.27	0.97	0.70	**0.00031**
CG4847	M	Prot.	0.21	0.21	0.81	0.47	**0.00015**
CG6069	M	Prot. hom.	0.18	0.26	0.91	1	**0.022**
CG8137	M	PI	0.40	0.33	0.98	1	**0.0051**
CG8982	M	Target	0.18	0.20	0.88	0.47	**0.0065**
CG9997	M	Prot. hom.	0.26	0.31	0.91	0.67	**0.010**
CG3066	M	Prot.	0.30	0.25	0.67	0.74	**0.0033**

NOTE. —*P* values in bold are significant at a FDR of 0.1.

evidence for recent selection at any locus by either the HKA test or the neutrality tests based on the SFS. On a deeper timescale, estimates of the rate of adaptation suggest similar patterns of adaptive evolution for male and female reproductive tract genes in *D. simulans* but a slower rate for female reproductive tract genes in *D. melanogaster*.

Models of sexual selection and sexual conflict generate a variety of predictions with respect to the frequency, direction, and extent of trait evolution (e.g., Iwasa and Pomiankowski 1995; Gavrilets 2000; Gavrilets and Hayashi 2006). Predictions of a long-term coevolutionary chase are frequently derived, and indeed, such an evolutionary regime could generate some of the patterns described in this study (e.g., consistently strong selection on male reproductive tract proteins). Gavrilets and Hayashi (2006) emphasized that initial values of key parameters, such as genetic variance in male and female traits and the strength of selection on males and females, can have important consequences for the outcome of sexual conflict. For example, differences in these parameters can determine whether a population undergoes a long-term arms race or if it ultimately evolves toward the male optimum. We therefore suggest that differences in such parameters may underlie our inference of temporally variable selection on reproductive tract proteins.

Functional Characteristics of Positively Selected Loci

The functional characteristics of positively selected genes suggest roles for immunity and sexual selection/sexual conflict in driving reproductive tract protein evolution. At least three genes showing evidence for positive selection in this study have documented or suspected roles in immunity (CG32382—Kambris et al. 2006; CG3066—Castillejo-López and Häcker 2005; CG10363—De Gregorio et al. 2001). As such, host–pathogen interactions may underlie their rapid evolution. Naturally occurring sexually transmitted diseases have not to our knowledge been documented in Drosophila, but the risk of pathogen introduction during mating has been demonstrated (Miest and Bloch-Qazi 2008). Several Acps appear to have antibacterial activity (Lung et al. 2001; Mueller et al. 2007), and genes with known roles in immunity are expressed in the reproductive tracts of both males and females (table 1). Mating alters the expression levels of several antimicrobial peptides in females (Lawniczak and Begun 2004; McGraw et al. 2004; Peng, Zipperlen, et al. 2005; Mack et al. 2006; Domanitskaya et al. 2007; Kapelnikov et al. 2008; Winterhalter and Fedorka 2009), although the physiological consequences of these gene expression changes are not clear (Fedorka et al. 2007; Wigby et al. 2008;). Together, these observations raise the possibility that host–pathogen interactions in the female reproductive tract could also contribute to rapid Acp evolution (see also Lawniczak et al. 2007).

Furthermore, three additional positively selected genes may have roles in sperm storage or sperm competition: The Acp PI CG8137 localizes to the sperm storage organs (SSO) following mating (Ravi Ram et al. 2005), and the predicted PI CG1865 and the predicted protease CG18125 have biased expression in the female sperm storage organs (table 1; FlyAtlas.org). Previous studies have found roles in sperm

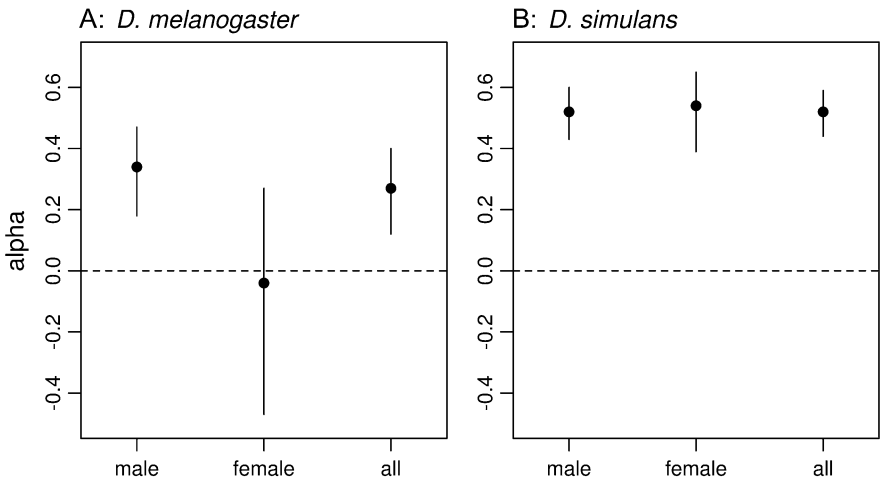

FIG. 2. Estimates of the rate of adaptive amino acid substitution (α) for 44 reproductive tract genes in *D. melanogaster* and *D. simulans* using the method of Bierne and Eyre-Walker (2004). Genes are separated according to their expression in the male or female reproductive tracts, with "all" representing all genes regardless of site of expression.

Table 4. Pairs of Loci with High Levels of LD.

Locus 1	Locus 2	Sex—Locus 1	Sex—Locus 2	Ontology—Locus 1	Ontology—Locus 2	Species	ZnS	P[a]
CG32203	CG13318	M	F	PI	Prot. hom.	*D. melanogaster*	0.47	0.0033
CG9806	CG7157 (Acp36DE)	M	M	Prot.	Target	*D. simulans*	0.17	0.0016
CG8982 (ovulin)	CG33121	M	M	Target	PI	*D. simulans*	0.18	0.0055
CG8982	CG18525	M	F	Target	PI	*D. simulans*	0.21	0.0088
CG1342	CG1857	M	F	PI	PI	*D. simulans*	0.27	0.0091

[a] *P* values were calculated from 10,000 permutations.

storage for several Acps that localize to the SSO (Bertram et al. 1996; Ravi Ram and Wolfner 2007; Wong, Albright, et al. 2008). Thus, effects of these genes on sperm competition and/or sperm preference may underlie selection.

Finally, we also found evidence for positive selection on two genes with known effects on female egg laying, CG8982 (which encodes the ovulation hormone Ovulin) and the predicted protease homolog CG9997 (table 3). CG9997 is necessary for the maintenance of SP (Ravi Ram and Wolfner 2009), a small peptide hormone that induces egg production and egg laying (Chen et al. 1988; Aigaki et al. 1991; Chapman et al. 2003; Liu and Kubli 2003), increases female feeding postmating (Carvalho et al. 2006) and decreases female sleep and lifespan (Wigby and Chapman 2005; Isaac et al. 2010). Given its effects on the mated female, SP is an excellent candidate as a molecular agent of sexual conflict. Nonetheless, SP shows little evidence for positive selection; it may be that interactions with multiple receptors impose substantial constraints on its evolution (Yapici et al. 2008; Ja et al. 2009). We suggest that selection may instead act on molecules that modulate SP activity, such as CG9997. In this regard, it will be interesting to investigate the molecular evolution of other proteins that interact with SP and CG9997 (Ravi Ram and Wolfner 2009).

Linkage Disequilibrium

Models of mate choice predict LD between trait and preference loci (e.g., Kirkpatrick 1982). Importantly, this LD arises solely as a consequence of biased mating such that physical linkage is not a prerequisite. If females bearing a preference allele P mate preferentially with males bearing trait allele T, then we should expect to see an excess of offspring carrying both the P and the T alleles. By analogy, we predicted that loci involved in postcopulatory sexual selection might show elevated LD as has been shown to be the case in abalone (Clark et al. 2009). Thus, in an attempt to identify coevolving genes in our sample, we calculated the LD summary statistic ZnS (Kelly 1997) for pairs of genes. We considered only LD between amino acid polymorphisms since these are most likely to reflect protein coevolution. Furthermore, we limited our analysis to pairs of genes sequenced in five or more strains. Significance of individual ZnS values was assessed using a permutation test (10,000 permutations).

No pair of loci showed significant LD at an FDR of 0.1. We note, however, that our data set is not ideal for this analysis because of our relatively low sample size (between 5 and 20 chromosomes sampled per locus) and a large number of tests. Several pairs of loci do show elevated ZnS when using a less stringent cut-off of $P < 0.01$ (table 4) and may represent promising candidates for future biochemical, genetic, and population genetic studies. Of particular interest are three gene pairs with high levels of LD between targets of proteolysis and proteolysis regulators: The proteolysis target Acp36DE (CG7157) shows relatively high LD with the predicted protease CG9806 in *D. simulans*, and ovulin (Acp26Aa/CG8982) is in high LD with the predicted PIs CG33121 and CG18525 (also in *D. simulans*).

Segregating and Fixed Putative Loss-of-function Alleles

Previous studies have suggested that Acps tend to turn over rapidly between species. Orthologs to many *D. melanogaster* Acps are not detected in distantly related species (Mueller et al. 2005; Wagstaff and Begun 2005a; Haerty et al. 2007) (although in some cases, high levels of sequence divergence may preclude detection using reciprocal blast), and many Acps from other species of Drosophila are similarly lineage specific (Holloway and Begun 2004; Wagstaff and Begun 2005b; Begun et al. 2006; Kelleher et al. 2007; Findlay et al. 2008, 2009). The population samples that we sequenced in this study harbored a number of putative loss-of-function alleles at multiple loci (table 5) that may represent loci becoming pseudogenized. In two cases (CG31681 and CG32383), a single allele was sequenced with a premature stop codon. The low frequencies of these alleles may be consistent with mutation–selection balance. However, in the case of CG14642, a female-expressed protease, 6 of 16 *D. melanogaster* alleles carried single base pair frameshifts due to at least three independent mutational

Table 5. Putative LOF Aalleles Observed in *D. melanogaster* at 3 of the 41 Proteolysis Regulators and Three Targets of Proteolysis.

Gene	# Uunique LOF Alleles	Type	Frequency
CG31681	1	Premature stop	1/14
CG32383	1	Premature stop	1/14
CG14642	3	Frameshift, premature stop	4/16, 1/16,1/16

NOTE. —Frequency indicates the number of LOF alleles observed at each gene relative to the number of chromosomes sampled from the population. LOF = loss of function.

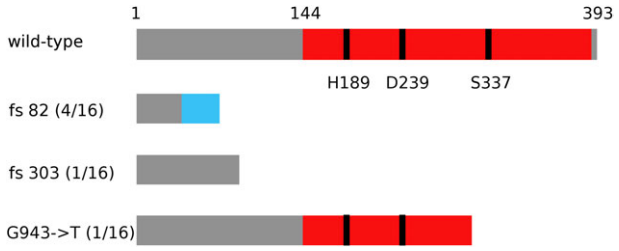

FIG. 3. Putative loss-of-function alleles of *D. melanogaster* CG14642. In this schematic of the protein, red represents the predicted proteolytic domain of this protein, with the catalytic residues H189, D239, and S337 indicated as black bars. Two frame shift mutations (fs 82 and fs 303) as well as an allele with a premature stop codon (G943→T) were detected in our population sample of 16, with frequencies indicated in parentheses.

events (fig. 3). It is tempting to posit that CG14642 is in transit toward becoming a pseudogene.

Summary

Our population genetic survey of 44 reproductive tract proteolysis regulators and targets of proteolysis revealed evidence that positive selection on these genes is not only frequent but is also variable through time and between species. For this group of genes, very recent selection appears to be more common in *D. melanogaster* than in *D. simulans* (tables 2 and 3), whereas on a deeper time scale, the rate of adaptive substitution is higher in *D. simulans* (fig. 2). These findings suggest that the strength and targets of selection change over time, consistent with an ongoing arms race between sexes and/or between host and pathogen. Moreover, available functional data on genes subject to positive selection suggest roles for both sexual selection and immunity in driving their evolution.

Supplementary Material

Supplementary data are available at *Molecular Biology* and *Evolution* online (www.mbe.oxfordjournals.org/).

Acknowledgments

We thank Andrew Clark, Rick Harrison, and an anonymous reviewer for helpful comments on this manuscript. We acknowledge funding support from National Institute of Health grants to M.F.W. (R01-HD038921) and CFA (R01-GM036431) and from an National Science Foundation Dissertation Improvement Grant to A.W. (DEB 0508152). A.W. was a Howard Hughes Medical Institute predoctoral fellow during this work. M.C.T. received support from the Cornell-HHMI Scholar Summer Fellowship Program.

References

Aigaki T, Fleischmann I, Chen PS, Kubli E. 1991. Ectopic expression of sex peptide alters reproductive behavior of female *D. melanogaster*. *Neuron* 7:557–563.

Allen AK, Spradling AC. 2008. The Sf1-related nuclear hormone receptor Hr39 regulates *Drosophila* female reproductive tract development and function. *Development* 135:311–321.

Almeida FC, Desalle R. 2008. Evidence of adaptive evolution of accessory gland proteins in closely related species of the *Drosophila repleta* group. *Mol Biol Evol* 25:2043–2053.

Almeida FC, Desalle R. 2009. Orthology, function and evolution of accessory gland proteins in the *Drosophila repleta* group. *Genetics* 181:235–245.

Andolfatto P. 2005. Adaptive evolution of non-coding DNA in *Drosophila*. *Nature* 437:1149–1152.

Andrés JA, Maroja LS, Bogdanowicz SM, Swanson WJ, Harrison RG. 2006. Molecular evolution of seminal proteins in field crickets. *Mol Biol Evol* 23:1574–1584.

Asgari S, Zhang G, Zareie R, Schmidt O. 2003. A serine proteinase homolog venom protein from an endoparasitoid wasp inhibits melanization of the host hemolymph. *Insect Biochem Mol Biol* 33:1017–1024.

Begun DJ, Holloway AK, Stevens K, et al. (13 co-authors) 2007. Population genomics: whole-genome analysis of polymorphism and divergence in *Drosophila simulans*. *PLoS Biol* 5:e310.

Begun DJ, Lindfors HA, Thompson ME, Holloway AK. 2006. Recently evolved genes identified from *Drosophila yakuba* and *D. erecta* accessory gland expressed sequence tags. *Genetics* 172:1675–1681.

Bertram MJ, Neubaum DM, Wolfner MF. 1996. Localization of the Drosophila male accessory gland protein Acp36DE in the mated female suggests a role in sperm storage. *Insect Biochem Mol Biol* 26:971–980.

Bierne N, Eyre-Walker A. 2004. The genomic rate of adaptive amino acid substitution in Drosophila. *Mol Biol Evol* 21:1350–1360.

Braswell WE, Andrés JA, Maroja LS, Harrison RG, Howard DJ, Swanson WJ. 2006. Identification and comparative analysis of accessory gland proteins in Orthoptera. *Genome* 49:1069–1080.

Carvalho GB, Kapahi P, Anderson DJ, Benzer S. 2006. Allocrine modulation of feeding behavior by the sex peptide of Drosophila. *Curr Biol* 16:692–696.

Castillejo-López C, Häcker U. 2005. The serine protease Sp7 is expressed in blood cells and regulates the melanization reaction in *Drosophila*. *Biochem Biophys Res Commun* 338:1075–1082.

Chapman T. 2008. The soup in my fly: evolution, form and function of seminal fluid proteins. *PLoS Biol* 6:e179.

Chapman T, Bangham J, Vinti G, Seifried B, Lung O, Wolfner MF, Smith HK, Partridge L. 2003. The sex peptide of *Drosophila melanogaster*: female post-mating responses analyzed by using RNA interference. *Proc Nat Acad Sci U S A* 100:9923–9928.

Chen PS, Stumm-Zollinger E, Aigaki T, Balmer J, Bienz M, Böhlen P. 1988. A male accessory gland peptide that regulates reproductive behavior of female *D. melanogaster*. *Cell* 54:291–298.

Chintapalli VR, Wang J, Dow JA. 2007. Using FlyAtlas to identify better *Drosophila melanogaster* models of human disease. *Nat Genet* 39:715–720.

Civetta A. 2003. Shall we dance or shall we fight? Using DNA sequence data to untangle controversies surrounding sexual selection. *Genome* 929:925–929.

Clark NL, Aagaard JE, Swanson WJ. 2006. Evolution of reproductive proteins from animals and plants. *Reproduction* 131:11–22.

Clark NL, Gasper J, Sekino M, Springer SA, Aquadro CF, Swanson WJ. 2009. Coevolution of interacting fertilization proteins. *PLoS Genet* 5:e1000570.

De Gregorio E, Spellman PT, Rubin GM, Lemaitre B. 2001. Genome-wide analysis of the *Drosophila* immune response by using oligonucleotide microarrays. *Proc Nat Acad Sci U S A* 98:12590–12595.

de Lamirande E. 2007. Semenogelin, the main protein of the human semen coagulum, regulates sperm function. *Semin Thromb Hemost* 33:60–68.

Domanitskaya EV, Liu H, Chen S, Kubli E. 2007. The hydroxyproline motif of male sex peptide elicits the innate immune response in *Drosophila* females. *FEBS J* 274:5659–5668.

Fay JC, Wu CI. 2000. Hitchhiking under positive Darwinian selection. *Genetics* 155:1405–1413.
Fedorka KM, Linder JE, Winterhalter W, Promislow D. 2007. Post-mating disparity between potential and realized immune response in *Drosophila melanogaster*. *Proc Biol Soc* 274:1211–1217.
Findlay GD, MacCoss MJ, Swanson WJ. 2009. Proteomic discovery of previously unannotated, rapidly evolving seminal fluid genes in *Drosophila*. *Genome Res* 19:886–896.
Findlay GD, Yi X, Maccoss MJ, Swanson WJ. 2008. Proteomics reveals novel Drosophila seminal fluid proteins transferred at mating. *PLoS Biol* 6:e178.
Gavrilets S. 2000. Rapid evolution of reproductive barriers driven by sexual conflict. *Nature* 403:886–889.
Gavrilets S, Hayashi TI. 2006. The dynamics of two- and three-way sexual conflicts over mating. *Philos Trans R Soc Lond B Biol Sci* 361:345–354.
Gupta S, Wang Y, Jiang H. 2005. Manduca sexta prophenoloxidase (proPO) activation requires proPO-activating proteinase (PAP) and serine proteinase homologs (SPHs) simultaneously. *Insect Biochem Mol Biol* 35:241–248.
Haddrill P, Thornton KR, Charlesworth B, Andolfatto P. 2005. Multilocus patterns of nucleotide variability and the demographic and selection history of Drosophila melanogaster populations. *Genome Res* 15:790–799.
Haerty W, Jagadeeshan S, Kulathinal RJ, et al. (11 co-authors) 2007. Evolution in the fast lane: rapidly evolving sex-related genes in *Drosophila*. *Genetics* 177:1321–1335.
Heifetz Y, Vandenberg LN, Cohn HI, Wolfner MF. 2005. Two cleavage products of the *Drosophila* accessory gland protein ovulin can independently induce ovulation. *Proc Nat Acad Sci U S A* 102:743–748.
Holloway AK, Begun DJ. 2004. Molecular evolution and population genetics of duplicated accessory gland protein genes in *Drosophila*. *Mol Biol Evol* 21:1625–1628.
Hudson RR, Kreitman M, Aguadé M. 1987. A test of neutral molecular evolution based on nucleotide data. *Genetics* 116:153–159.
Isaac RE, Li C, Leedale AE, Shirras AD. 2010. *Drosophila* male sex peptide inhibits siesta sleep and promotes locomotor activity in the post-mated female. *Proc Biol Soc* 277:65–70.
Iwasa Y, Pomiankowski A. 1995. Continual change in mate preferences. *Nature* 377:420–422.
Ja WW, Carvalho GB, Madrigal M, Roberts RW, Benzer S. 2009. The *Drosophila* G protein-coupled receptor, Methuselah, exhibits a promiscuous response to peptides. *Protein Sci* 18:2203–2208.
Jensen JD, Bachtrog D. 2010. Characterizing recurrent positive selection at fast-evolving genes in *Drosophila miranda* and *Drosophila pseudoobscura*. *Genome Biol Evol* 2:371–378.
Kambris Z, Brun S, Jang I-H, Nam H-J, Romeo Y, Takahashi K, Lee W-J, Ueda R, Lemaitre B. 2006. *Drosophila* Immunity: a large-scale in vivo RNAi screen identifies five serine proteases required for Toll activation. *Curr Biol* 16:808–813.
Kapelnikov A, Zelinger E, Gottlieb Y, Rhrissorrakrai K, Gunsalus KC, Heifetz Y. 2008. Mating induces an immune response and developmental switch in the *Drosophila* oviduct. *Proc Nat Acad Sci U S A* 105:13912–13917.
Kelleher ES, Clark NL, Markow TA. 2011. Diversity enhancing selection acts on a female reproductive protease family in four sub-species of *Drosophila* mojavensis. *Genetics* 187:865–876.
Kelleher ES, Markow TA. 2009. Duplication, selection and gene conversion in a *Drosophila mojavensis* female reproductive protein family. *Genetics* 181:1451–1465.
Kelleher ES, Pennington JE. 2009. Protease gene duplication and proteolytic activity in *Drosophila* female reproductive tracts. *Mol Biol Evol* 26:2125–2134.
Kelleher ES, Swanson WJ, Markow TA. 2007. Gene duplication and adaptive evolution of digestive proteases in *Drosophila arizonae* female reproductive tracts. *PLoS Genet* 3:e148.
Kelly JK. 1997. A test of neutrality based on interlocus associations. *Genetics* 146:1197–1206.
Kim Y, Stephan W. 2002. Detecting a local signature of genetic hitchhiking along a recombining chromosome. *Genetics* 160:765–777.
Kirkpatrick M. 1982. Sexual selection and the evolution of female choice. *Evolution* 36:1–12.
Kopp A, Frank A, Fu J. 2006. Historical biogeography of *Drosophila simulans* based on Y-chromosomal sequences. *Mol Phy Evol* 38:355–362.
Kwon TH, Kim MS, Choi HW, Joo CH, Cho MY, Lee BL. 2000. A masquerade-like serine proteinase homologue is necessary for phenoloxidase activity in the coleopteran insect, *Holotrichia diomphalia* larvae. *Eur J Biochem* 267:6188–6196.
Lawniczak MKN, Barnes AI, Linklater JR, Boone JM, Wigby S, Chapman T. 2007. Mating and immunity in invertebrates. *Trends Ecol Evol* 22:48–55.
Lawniczak MKN, Begun DJ. 2004. A genome-wide analysis of courting and mating responses in *Drosophila melanogaster* females. *Genome* 47:900–910.
Lawniczak MKN, Begun DJ. 2007. Molecular population genetics of female-expressed mating-induced serine proteases in *Drosophila melanogaster*. *Mol Biol Evol* 24:1944–1951.
Lee KY, Zhang R, Kim MS, Park JW, Park HY, Kawabata S, Lee BL. 2002. A zymogen form of masquerade-like serine proteinase homologue is cleaved during pro-phenoloxidase activation by Ca2+ in coleopteran and *Tenebrio molitor* larvae. *Eur J Biochem* 269:4375–4383.
Liu H, Kubli E. 2003. Sex-peptide is the molecular basis of the sperm effect in *Drosophila melanogaster*. *Proc Nat Acad Sci U S A* 100:9929–9933.
Lung O, Kuo L, Wolfner MF. 2001. *Drosophila* males transfer antibacterial proteins from their accessory gland and ejaculatory duct to their mates. *J Insect Physiol* 47:617–622.
Mack PD, Kapelnikov A, Heifetz Y, Bender M. 2006. Mating-responsive genes in reproductive tissues of female *Drosophila melanogaster*. *Proc Nat Acad Sci U S A* 103:10358–10363.
Malm J, Hellman J, Hogg P, Lilja H. 2000. Enzymatic action of prostate-specific antigen (PSA or hK3): substrate specificity and regulation by Zn(2+), a tight-binding inhibitor. *Prostate* 45:132–139.
McDonald JH, Kreitman M. 1991. Adaptive protein evolution at the Adh locus in *Drosophila*. *Nature* 351:652–654.
McGraw LA, Gibson G, Clark AG, Wolfner MF. 2004. Genes regulated by mating, sperm, or seminal proteins in mated female *Drosophila melanogaster*. *Curr Biol* 14:1509–1514.
Miest TS, Bloch-Qazi MC. 2008. Sick of mating: sexual transmission of a pathogenic bacterium in *Drosophila melanogaster*. *Fly (Austin)* 2:215–219.
Mueller JL, Page JL, Wolfner MF. 2007. An ectopic expression screen reveals the protective and toxic effects of Drosophila seminal fluid proteins. *Genetics* 175:777–783.
Mueller JL, Ravi Ram K, McGraw LA, Bloch-Qazi MC, Siggia ED, Clark AG, Aquadro CF, Wolfner MF. 2005. Cross-species comparison of *Drosophila* male accessory gland protein genes. *Genetics* 171:131–143.
Mueller JL, Ripoll DR, Aquadro CF, Wolfner MF. 2004. Comparative structural modeling and inference of conserved protein classes in *Drosophila* seminal fluid. *Proc Nat Acad Sci U S A* 101:13542–13547.
Nasrallah JB. 2002. Recognition and rejection of self in plant reproduction. *Science* 296:305–308.

Nolte V, Schlötterer C. 2008. African *Drosophila melanogaster* and *D. simulans* populations have similar levels of sequence variability, suggesting comparable effective population sizes. *Genetics* 178:405–412.

Panhuis TM, Clark NL, Swanson WJ. 2006. Rapid evolution of reproductive proteins in abalone and *Drosophila*. *Philos Trans R Soc Lond B Biol Sci* 361:261–268.

Panhuis TM, Swanson WJ, Nunney L. 2003. Population genetics of accessory gland proteins and sexual behavior in *Drosophila melanogaster* populations from evolution canyon. *Evolution* 57:2785–2791.

Park M, Wolfner MF. 1995. Male and female cooperate in the prohormone-like processing of a *Drosophila melanogaster* seminal fluid protein. *Dev Biol* 171:694–702.

Peng J, Chen S, Büsser S, Liu H, Honegger T. 2005. Gradual release of sperm bound sex-peptide controls female postmating behavior in *Drosophila*. *Curr Biol* 15:207–213.

Peng J, Zipperlen P, Kubli E. 2005. *Drosophila* sex-peptide stimulates female innate immune system after mating via the Toll and Imd pathways. *Curr Biol* 15:1690–1694.

Pool JE, Aquadro CF. 2006. History and structure of sub-Saharan populations of *Drosophila melanogaster*. *Genetics* 174:915–929.

Prokupek A, Hoffmann F, Eyun S, Moriyama E, Zhou M, Harshman L. 2008. An evolutionary expressed sequence tag analysis of *Drosophila* spermatheca genes. *Evolution* 62:2936–2947.

Przeworski M. 2003. Estimating the time since the fixation of a beneficial allele. *Genetics* 164:1667–1676.

Ravi Ram K, Ji S, Wolfner MF. 2005. Fates and targets of male accessory gland proteins in mated female *Drosophila melanogaster*. *Insect Biochem Mol Biol* 35:1059–1071.

Ravi Ram K, Sirot LK, Wolfner MF. 2006. Predicted seminal astacin-like protease is required for processing of reproductive proteins in *Drosophila melanogaster*. *Proc Nat Acad Sci U S A* 103:18674–18679.

Ravi Ram K, Wolfner MF. 2007. Sustained post-mating response in *Drosophila melanogaster* requires multiple seminal fluid proteins. *PLoS Genet* 3:e238.

Ravi Ram K, Wolfner MF. 2009. A network of interactions among seminal proteins underlies the long-term postmating response in Drosophila. *Proc Nat Acad Sci U S A* 106:15384–15389.

Ross J, Jiang H, Kanost MR, Wang Y. 2003. Serine proteases and their homologs in the *Drosophila melanogaster* genome: an initial analysis of sequence conservation and phylogenetic relationships. *Gene* 304:117–131.

Sirot LK, Poulson RL, McKenna MC, Girnary H, Wolfner MF, Harrington LC. 2008. Identity and transfer of male reproductive gland proteins of the dengue vector mosquito, *Aedes aegypti*: potential tools for control of female feeding and reproduction. *Insect Biochem Mol Biol* 38:176–189.

Storey JD, Tibshirani R. 2003. Statistical significance for genomewide studies. *Proc Nat Acad Sci U S A* 100:9440–9445.

Swanson WJ, Clark AG, Waldrip-Dail HM, Wolfner MF, Aquadro CF. 2001. Evolutionary EST analysis identifies rapidly evolving male reproductive proteins in *Drosophila*. *Proc Nat Acad Sci U S A* 98:7375–7379.

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

Swanson WJ, Wong A, Wolfner MF, Aquadro CF. 2004. Evolutionary expressed sequence tag analysis of *Drosophila* female reproductive tracts identifies genes subjected to positive selection. *Genetics* 168:1457–1465.

Szecsi PB, Lilja H. 1993. Gastricsin-mediated proteolytic degradation of human seminal fluid proteins at pH levels found in the human vagina. *J Androl* 14:351–358.

Tajima F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* 123:585–595.

Takebayashi N, Brewer PB, Newbigin E, Uyenoyama MK. 2003. Patterns of variation within self-incompatibility loci. *Mol Biol Evol* 20:1778–1794.

Thornton K. 2003. Libsequence: a C++ class library for evolutionary genetic analysis. *Bioinformatics* 19:2325–2327.

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

Veveris-Lowe TL, Kruger SJ, Walsh T, Gardiner RA, Clements JA. 2007. Seminal fluid characterization for male fertility and prostate cancer: kallikrein-related serine proteases and whole proteome approaches. *Semin Thromb Hemost* 33:87–99.

Wagstaff BJ, Begun DJ. 2005a. Comparative genomics of accessory gland protein genes in *Drosophila melanogaster* and *D. pseudoobscura*. *Mol Biol Evol* 22:818–832.

Wagstaff BJ, Begun DJ. 2005b. Molecular population genetics of accessory gland protein genes and testis-expressed genes in *Drosophila mojavensis* and *D. arizonae*. *Genetics* 171:1083–1101.

Wigby S, Chapman T. 2005. Sex peptide causes mating costs in female *Drosophila melanogaster*. *Curr Biol* 15:316–321.

Wigby S, Domanitskaya EV, Choffat Y, Kubli E, Chapman T. 2008. The effect of mating on immunity can be masked by experimental piercing in female *Drosophila melanogaster*. *J Insect Physiol* 54:414–420.

Winterhalter WE, Fedorka KM. 2009. Sex-specific variation in the emphasis, inducibility and timing of the post-mating immune response in *Drosophila melanogaster*. *Proc Biol Soc* 276:1109–1117.

Wong A, Albright SN, Giebel JD, Ravi Ram K, Ji S, Fiumera AC, Wolfner MF. 2008. A role for Acp29AB, a predicted seminal fluid lectin, in female sperm storage in *Drosophila melanogaster*. *Genetics* 180:921–931.

Wong A, Turchin MC, Wolfner MF, Aquadro CF. 2008. Evidence for positive selection on *Drosophila melanogaster* seminal fluid protease homologs. *Mol Biol Evol* 25:497–506.

Wright SI, Charlesworth B. 2004. The HKA test revisited: a maximum-likelihood-ratio test of the standard neutral model. *Genetics* 168:1071–1076.

Yanai I, Benjamin H, Shmoish M, et al. (12 co-authors) 2005. Genome-wide midrange transcription profiles reveal expression level relationships in human tissue specification. *Bioinformatics* 21:650–659.

Yapici N, Kim YJ, Ribeiro C, Dickson BJ. 2008. A receptor that mediates the post-mating switch in *Drosophila* reproductive behaviour. *Nature* 451:33–37.

Zhang G, Lu Z-Q, Jiang H, Asgari S. 2004. Negative regulation of prophenoloxidase (proPO) activation by a clip-domain serine proteinase homolog (SPH) from endoparasitoid venom. *Insect Biochem Mol Biol* 34:477–483.