Protease Gene Duplication and Proteolytic Activity in *Drosophila* Female Reproductive Tracts

Erin S. Kelleher† and James E. Pennington‡*

*Department of Ecology and Evolutionary Biology, University of Arizona; †Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY; and ‡Department of Biochemistry and Molecular Biophysics, Center for Insect Science, University of Arizona

Secreted proteases play integral roles in sexual reproduction in a broad range of taxa. In the genetic model *Drosophila melanogaster*, these molecules are thought to process peptides and activate enzymes inside female reproductive tracts, mediating critical postmating responses. A recent study of female reproductive tract proteins in the cactophilic fruit fly *Drosophila arizonae*, identified pervasive, lineage-specific gene duplication amongst secreted proteases. Here, we compare the evolutionary dynamics, biochemical nature, and physiological significance of secreted female reproductive serine endoproteases between *D. arizonae* and its congener *D. melanogaster*. We show that *D. arizonae* lower female reproductive tract (LFRT) proteins are significantly enriched for recently duplicated secreted proteases, particularly serine endoproteases, relative to *D. melanogaster*. Isolated lumen from *D. arizonae* LFRTs, furthermore, exhibits significant trypsin-like and elastase-like serine endoprotease acitivity, whereas no such activity is seen in *D. melanogaster*. Finally, trypsin- and elastase-like activity in *D. arizonae* female reproductive tracts is negatively regulated by mating. We propose that the intense proteolytic environment of the *D. arizonae* female reproductive tract relates to the extraordinary reproductive physiology of this species and that ongoing gene duplication amongst these proteases is an evolutionary consequence of sexual conflict.

Introduction

In internally fertilizing organisms, sexual reproduction is mediated by an elaborate series of interactions between the male ejaculate and the female reproductive tract. This interface extends far beyond gamete fusion, playing essential roles in sperm fate (reviewed in Neubaum and Wolfner 1999) as well as female behavior and physiology (Reviewed in Wolfner 2007; Robertson 2007). Although reproductive tract interactions are fundamental to fertilization and organismal fitness, male ejaculates and female reproductive tracts are observed to evolve rapidly at both the morphological (Pitnick et al. 1999; Brennan et al. 2007; Marshall 2007) and biochemical levels (reviewed in Swanson and Vacquier 2002; Clark et al. 2006; Panhuis et al. 2006). This exceptional divergence often is hypothesized to be a consequence of a coevolutionary chase between males and females driven by sexual conflict or a difference in the reproductive interests of the two sexes (Parker 1979; Rice 1996; Gavrilets 2000).

The molecular underpinnings of ejaculate–female dynamics remain poorly understood; however, proteases have emerged as prominent reproductive players in both insects (Swanson et al. 2001, 2004; Braswell et al. 2006; Sirot et al. 2008) and mammals (reviewed in Dacheux et al. 2003). In *Drosophila melanogaster*, proteolysis is thought to modulate the female postmating response by processing or activating male-derived peptides and enzymes (Monsma et al. 1990; Park and Wolfner 1995; Peng et al. 2005; Ravi Ram et al. 2006; Pilpel et al. 2008). Population-genetic and divergence-based analyses, furthermore, reveal a high frequency of adaptive evolution amongst both male and female reproductive tract proteases and protease homologs, suggesting an exciting role for this class of enzymes in intersexual coevolution (Swanson et al. 2004; Panhuis and Swanson 2006; Haerty et al. 2007; Lawniczak and Begun 2007; Findlay et al. 2008; Prokupek et al. 2008; Wong et al. 2008).

A recent expressed sequence tag (EST) screen of the *Drosophila arizonae* lower female reproductive tract (LFRT: uterus, spermathecae, seminal receptacle, parovaria, common oviduct) identified five lineage-specific protease gene families in which two or more paralogs are expressed in the LFRT (Kelleher et al. 2007). Recurrent duplication of independent loci with similar biochemical functions, in conjunction with evidence of positive selection in three of these gene families, points to an adaptive expansion of proteolytic capacity in the *D. arizonae* lineage (Kelleher et al. 2007). It also may suggest intense sexual conflict as mathematical models have shown that rapid diversification is an important female "strategy" in sexually antagonistic coevolution (Gavrilets and Waxman 2002; Hayashi et al. 2007).

Enhanced proteolytic capacity in the LFRT may underlie two specialized physiological processes in *D. arizonae* females. First, *D. arizonae* incorporate significant quantities of male-derived protein into somatic tissues and oocytes (Markow and Ankney 1988; Pitnick et al. 1997). Proteases could play a critical role in this process by degrading sperm and/or seminal proteins into smaller peptides that are more easily absorbed. Second, *D. arizonae* females form an insemination reaction, an opaque white mass of unknown biochemical composition, after every copulation (Patterson 1946). Females must degrade this mass in order to oviposit or remate (Knowles and Markow 2001), a process which could involve proteolysis.

In this study, we compare the evolutionary history, biochemical nature, and physiological significance of secreted female reproductive serine endoproteases (SFRSEs) between *D. arizonae* and its congener *D. melanogaster*. *Drosophila melanogaster* exhibits neither ejaculate incorporation nor an insemination reaction (Markow and Ankney 1984, 1988; Pitnick et al. 1997), making it ideal for interspecific comparison with *D. arizonae*. First, we explicitly test the hypothesis that secreted proteases expressed in *D. arizonae* LFRTs have experienced a high frequency

Key words: gene duplication, protease, reproductive protein, female reproductive tract, sexual conflict.

E-mail: esk72@cornell.edu.

Mol. Biol. Evol. 26(9):2125–2134. 2009
doi:10.1093/molbev/msp121
Advance Access publication June 22, 2009

© The Author 2009. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution. All rights reserved.
For permissions, please e-mail: journals.permissions@oxfordjournals.org

of recent gene duplication when compared with *D. melanogaster*. We show that *D. arizonae* LFRTs are significantly enriched for recently duplicated secreted proteases, particularly serine endoproteases. Serine endoproteases comprise an enzymatic class that is particularly well studied in terms of catalytic function (Reviewed in Polgar 2005), key residues that determine substrate specificity (Perona and Craik 1995), and availability of synthetic substrates and inhibitors for biochemical assays. We therefore explore differences in serine endoprotease complement between *D. arizonae* and *D. melanogaster* LFRTs using both bioinformatic approaches and in vitro assays. *Drosophila arizonae* female reproductive tracts are shown to encode a greater number of enzymes in a broader range of specificities relative to *D. melanogaster* as well as enhanced proteolytic activity that is regulated by mating. We discuss our results in terms of differences in reproductive biology between *D. arizonae* and *D. melanogaster*.

Materials and Methods

Identification of Annotated LFRT Proteins

Protein sequences from candidate LFRT proteins for *D. melanogaster* (150 annotated candidates, Swanson et al. 2004) and *Drosophila mojavensis* (234 annotated candidates, Kelleher et al. 2007) were obtained from flybase (http://www.flybase.org). It was necessary to use *D. mojavensis*, the closely related sister species of *D. arizonae* (most common recent ancestor = ~1.5 Ma, Matzkin 2004), for this analysis as no fully sequenced genome is available for *D. arizonae*. Swanson et al. (2004) and Kelleher et al. (2007) used almost identical experimental approaches for identifying candidate LFRT proteins, and therefore present comparable data sets between *D. arizonae* and *D. melanogaster*.

Identification of Annotated Serine Endoproteases

Drosophila melanogaster serine endoproteases and serine endoprotease homologs (147 proteases and 57 protease homologs, Ross et al. 2003) were obtained from flybase (http://www.flybase.org). Serine endoprotease homologs contain a recognizable protease domain, but substitutions have occurred in the amino acids forming the catalytic triad, likely rendering these proteases noncatalytic (Polgar 2005). We identified an additional two serine endoproteases, CG30025 and CG30031, as well as three serine endoprotease homologs, sphinx2, CG31780, and CG21827, based on close homology to at least one of the serine proteases or serine protease homologs described in Ross et al. (2003).

It was necessary to identify candidate serine endoproteases in the *D. mojavensis* genome de novo, using the same approach as Ross et al. (2003). Briefly, *Manduca sexta* PAP (Jiang et al. 1998) was used to query the GLEANR protein annotations of *D. mojavensis* (http://rana.lbl.gov/drosophila/) using PSI-Blast (*e* value = 1, Altschul et al. 1997). Every 20th sequence was retained for a second iteration of PSI-Blast. Conserved serine endoprotease domains were confirmed with hmmpfam (Eddy 1998). The complete list of 129 candidate *D. mojavensis* serine endoproteases and 38 *D. mojavensis* serine endoprotease homologs identified in this study is presented in supplementary table 1 (Supplementary Material online).

Identification of Recent Duplicates

To examine the frequency of recent duplicates among both candidate LFRT proteins and candidate serine endoproteases, additional paralogs were identified in the genomes of *D. mojavensis* and *D. melanogaster* using BlastP ($e = 0.001$, Altschul et al. 1990). For each protein and blast hit pair, coding sequences were aligned in ClustalW (Thompson et al. 1994), and % protein identity and corrected synonymous divergence (d_s) were calculated in PAML (Yang 1997). Recent duplicates were defined as proteins with greater than 50% identity, where $d_s < 0.5$ and are presented in supplementary table 2 (LFRT proteins) and supplementary table 3 (candidate proteases; Supplementary Materials online).

Functional Enrichment

Significantly overrepresented gene ontology terms (GO terms, Ashburner et al. 2000) in recently duplicated *D. arizonae*/*D. mojavensis* LFRT proteins were identified in Fatigo (Al-Shahrour et al. 2004, 2007). GO annotations for the *D. melanogaster* homolog of each LFRT protein was used as there is no existing GO annotation data set for *D. mojavensis*. Overrepresented GO terms were identified with Fisher's exact test after correcting for multiple measure based on the false discovery rate (Benjamini and Hochberg 1995).

SFRSE Annotation

We searched data sets from previous expression studies of *D. melanogaster* (Swanson et al. 2004; Mack et al. 2006; Lawniczak and Begun 2007) and *D. arizonae* (Kelleher et al. 2007) LFRTs to identify SFRSEs in both these species (table 2). Conservation of the catalytic triad, necessary for proteolytic function (Polgar 2005), was verified in *D. arizonae* ESTs where possible or in the ortholog of its sister species, *D. mojavensis* (http://rana.lbl.gov/drosophila/) when the relevant sequence was not present in the EST. Secondary domains in these proteases were identified previously (Kelleher et al. 2007), and CLIP domains were identified by eye as in Jiang and Kanost (2000). CLIP domains are cysteine-rich regions that are thought to play and important role in protein–protein interactions that regulate proteolytic cascades (Jiang and Kanost 2000). *Drosophila arizonae* female reproductive tract protease ESTs were translated and aligned to porcine elastase to identify key substrate specificity residues, as in Perona and Craik (1995). Catalytic function, secondary domains, and substrate specificity for *D. melanogaster* female reproductive tract proteases were adapted from Ross et al. (2003).

Stocks and Fly Husbandry

The *D. melanogaster* Oregon-R strain was obtained from T.A. Hartl at the University of Arizona and reared

on standard cornmeal media. The *D. arizonae* strain was collected in Tucson, AZ, in 12/2005 by E.S.K. and reared on opuntia banana media (http://stockcenter.arl.arizona.edu/).

Tissue Harvesting

For assays of proteolytic activity in *D. arizonae* and *D. melanogaster* LFRTs and *D. arizonae* male seminal vesicles and accessory glands (SVAGs), tissue was harvested from adults reared in population bottles in order to achieve the maximum diversity of mating states. LFRTs were removed from *D. melanogaster*, $\geq$1 day posteclosion, whereas LFRTs and SVAGs were removed from *D. arizonae* $\geq$9 days posteclosion to ensure reproductive maturity (reviewed in Markow 1996).

For comparisons of proteolytic activity between virgin and mated *D. arizonae* LFRTs, virgin males and females were isolated within 24 h of eclosion and aged separately for 9–12 days. For each cohort of females, 50% were mated at densities of approximately 10 females and 20 males per vial, whereas the remaining 50% were retained as virgins. After 2 h of unrestricted mating, the females were separated and their LFRTs removed within 2 h. We did not verify whether all females had mated; however, most dissected females exhibited an insemination reaction indicative of recent copulation (Patterson 1946). Virgin females were dissected concurrently to minimize differences between the two treatments.

For all assays, comparisons between treatments were made by standardizing to the total number of reproductive tracts dissected rather than the total extracted soluble protein. This approach was employed to minimize the effect of dilution of female proteases by male seminal proteins in mated females, which could lead to spurious differences in proteolytic activity.

All dissections were performed in 1× phosphate buffer solution on a glass slide. Tissue was harvested directly into trypsin assay buffer on ice (50 mM Tris, 10 mM $CaCl_2$, pH 7) and stored at −20 °C. Dissections were performed with care to prevent contamination from closely associated gut tissue (see supplementary fig. 1, Supplementary Material online).

Colorimetric Assays of Proteolytic Activity in *D. arizonae* and *D. melanogaster* Female Reproductive Tissues

Chromogenic *p*-nitroanilide substrate for trypsin, Bz-DL-Arg-*p*NA · HCl (DL-BApNA, Sigma), was prepared as a 100-mM stock solution in dimethyl sulfoxide. Colorimetric *p*-nitroanilide substrate for elastase, Boc-Ala-Ala-Pro-Ala-*p*NA (BAAPApNA, Calbiochem), was prepared as a 2-mM stock solution in trypsin assay buffer. Diisoflourophosphate (DFP, Calbiochem) serine protease inhibitor was prepared as a 1-M stock solution in isopropyl alcohol. 4-(2-Aminoethyl) benzenesulfonyl fluoride hydrochloride (AEBSF, Sigma-Aldrich) serine protease inhibitor was prepared as a 1-M stock solution in deionized water.

For both species, nine replicates of 100 individually dissected LFRTs were centrifuged at 1000 × g for 3 min, to release only the soluble fraction. The supernatant of all nine replicates was pooled and then split into nine replicate aliquots. These aliquots formed three technical replicates of three treatments: 1) chromogenic substrate at final concentration 3.3 mM (trypsin) or 1 mM (elastase); 2) 60-s preincubation with AEBSF at final concentration 6.66 mM, followed by addition of the chromogenic substrate at final concentration 3.3 mM (trypsin) or 1 mM (elastase); and 3) 60-s preincubation with DFP at final concentration 6.66 mM, followed by addition of the chromogenic substrate at final concentration 3.3 mM (trypsin) or 1 mM (elastase).

Trypsin assays were allowed to incubate for 20 min at room temperature, whereas elastase assays were allowed to incubate for 10 min at room temperature. For all experiments, activity was measured as an increase in absorbance at 405 nm, as detected by a Cary 50 Bio UV spectrophotometer (Varian, Palo Alto, CA), compared with a standard control of 3.3 mM trypsin substrate or 1 mM elastase substrate in assay buffer.

Colorimetric Assays of Proteolytic Activity in *D. arizonae* Male Reproductive Tissues

Reagents, protein isolation, and reaction conditions were as in assays of LFRTs (above). Supernatant from 10 replicates of 100 individually dissected SVAGs was pooled and split into 10 replicate aliquots. These 10 aliquots formed three technical replicates of three different treatments (as above) plus a control containing only reproductive tract protein in assay buffer. This control was necessary, as *D. arizonae* testes are pigmented. Activity of all nine assays was measured as an increase in absorbance at 405 nm above this control.

Colorimetric Assays of Proteolytic Activity Virgin versus Mated *D. arizonae* LFRTs

Stock solutions, reaction conditions, and activity measurements were as in other assays (above); however, both the DL-BApNA (ICN Biomedicals) and the BAAPApNA (Bachem) were ordered from a different supplier. Supernatant from four biological replicates of 100 virgin LFRTs and 100 mated LFRTs were compared for trypsin- and elastase-like activity.

Evolutionary Analyses

Maximum likelihood estimates of pairwise d_N/d_S between *D. melanogaster* and *Drosophila simulans* coding sequences and between *D. arizonae* ESTs and *D. mojavensis* coding sequences were generated in PAML (Yang 1997). Although the divergence times between *D. melanogaster* and *D. simulans* (~3 Ma, Hey and Kliman 1993) and *D. arizonae* and *D. mojavensis* (~1.5 Ma, Matzkin 2004) are slightly different, this should not affect our estimate of d_N/d_S as the difference in divergence time will effect both site classes equally.

Table 1
Recent Duplicates in *Drosophila melanogaster* and *Drosophila mojavensis* LFRT Proteins

Candidate Female Reproductive Tract Protein	Functional Class
D. melanogaster	
IM10-PA	Defense response
CG30035-PB	Carbohydrate transport
scpr-C-PA	CRISP
D. mojavensis	
Dmoj\GLEANR_12010	Serine endoprotease
Dmoj\GLEANR_12324	Serine protease
Dmoj\GLEANR_12325	Serine protease
Dmoj\GLEANR_1234	Protease inhibitor
Dmoj\GLEANR_12931	Metalloprotease
Dmoj\GLEANR_13880	Sulfate transport
Dmoj\GLEANR_2575	Serine endoprotease
Dmoj\GLEANR_2703	Metalloprotease
Dmoj\GLEANR_3081	Unknown function
Dmoj\GLEANR_4546	Glycosyl hydrolase
Dmoj\GLEANR_5037	Unknown function
Dmoj\GLEANR_6725	Unknown function
Dmoj\GLEANR_6984	Serine endoprotease
Dmoj\GLEANR_7051	Lipase
Dmoj\GLEANR_778	Metalloprotease
Dmoj\GLEANR_896	Serine endoprotease
Dmoj\GLEANR_897	Serine endoprotease
Dmoj\GLEANR_898	Serine endoprotease
Dmoj\GLEANR_9617	Serine protease

NOTE.—Annotated candidate LFRT proteins from *D. melanogaster* (Swanson et al. 2004) and *Drosophila arizonae* (Kelleher et al. 2007) with recent duplicates in the *D. melanogaster* and *D. mojavensis* genomes are identified. Functional class is based on GO terms from flybase (http://flybase.org/) and conserved domains.

Results

Drosophila arizonae SFRSEs Are Enriched for Recently Duplicated Serine Endoproteases

To explicitly test the hypothesis that the *D. arizonae**D. mojavensis* lineage has experienced exceptional duplication of SFRSEs, we first compared the frequency of recent duplicates between *D. arizonae*/*D. mojavensis* and *D. melanogaster* LFRT proteins. Whereas only three (of 150, Swanson et al. 2004) *Drosophila melanogaster* LFRT proteins have a highly similar paralog ($d_s < 0.5$) in the *D. melanogaster* genome, a total of 19 *D. arizonae*/*D. mojavensis* LFRT proteins (of 234, Kelleher et al. 2007) have a highly similar paralog in the *D. mojavensis* genome (table 1, supplementary table 2, Supplementary Material online). *Drosophila arizonae*/*D. mojavensis* LFRT proteins as a whole, therefore, are considerably enriched for recent duplicates relative to *D. melanogaster* (two-tailed Fisher's exact test, $P = 0.01$). We note this is likely a conservative estimate as six recent duplicates identified in Kelleher et al. (2007) remain unannotated and thus were excluded from the comparison. There is no evidence that *D. mojavensis* experiences elevated turnover in gene families with respect to other *Drosophila* species, including *D. melanogaster* (Hahn et al. 2007). It is unlikely, therefore, that the increased frequency of recent duplicates is a genome-wide phenomenon in *D. mojavensis* or *D. arizonae*.

To identify classes of proteins that are prevalent among recent duplicates, we tested for overrepresentation of molecular function GO terms (Ashburner et al. 2000) relative to our complete list of annotated and unannotated *D. arizonae*/*D. mojavensis* LFRT proteins (241 total genes, Kelleher et al. 2007). Five interrelated terms were significantly overrepresented in recent duplicates after correction for multiple testing: hydrolase activity, peptidase activity, serine-type peptidase activity, endopeptidase activity, and serine-type endopeptidase activity. Recently duplicated *D. arizonae* LFRT proteins, therefore, are significantly enriched for secreted proteases, particularly serine endoproteases. *Drosophila arizonae* LFRT proteins as a whole, moreover, are not enriched in recent duplicates relative to *D. melanogaster* when all proteases are excluded from the data (two-tailed Fisher's exact test, $P = 0.75$). Thus, the high frequency of recent duplicates observed in *D. arizonae* LFRT protein largely is due to preferential duplication of secreted proteases in this lineage.

The observed preferential duplication could be exclusive to those serine endoproteases that are expressed in LFRTs or could be general to all serine endoproteases in the *D. mojavensis* genome. We therefore examined whether there was a higher frequency of recent duplicates ($d_s < 0.5$) among *D. mojavensis* serine endoproteases (129 total, supplementary table 1, Supplementary Material online) relative to *D. melanogaster* (149 total, Ross et al. 2003). *Drosophila mojavensis* serine endoproteases are significantly enriched for recent duplicates (29 of 129) relative to *D. melanogaster* (10 of 149, two-tailed Fisher's exact test, $P = 2 \times 10^{-4}$). This enrichment is considerably less significant; however, when LFRT proteins and their close paralogs are excluded from the data set (two-tailed Fisher's exact test, $P = 0.018$), suggesting that the enrichment of recent duplicates largely is driven by the preferential duplication of LFRT proteins. Indeed, recently duplicated *D. mojavensis* serine endoproteases are significantly enriched for LFRT proteins and their close paralogs (two-tailed Fisher's exact test, $P = 1.8 \times 10^{-4}$).

An elevated frequency of recent duplicates among serine endoproteases points to an adaptive expansion of proteolytic capacity in *D. arizonae* LFRTs. As an enzymatic class, serine endoproteases are exceedingly well described in terms of defining how key amino acid residues affect catalytic function (reviewed in Polgar 2005) and substrate specificity (Perona and Craik 1995). Synthetic substrates and inhibitors for these proteases, furthermore, are readily available. The remainder of this study, therefore, focuses on a comparison of the SFRSE complement between *D. arizonae* and *D. melanogaster*.

Drosophila arizonae LFRTs are Enriched for Digestive Serine Endoproteases

Comparisons of the nature, number, and specificity of SFRSEs suggest dramatic enhancement of *D. arizonae* proteolytic capacity relative to *D. melanogaster* (table 2). Almost twice as many SFRSEs are found in *D. arizonae* LFRTs (15) as in *D. melanogaster* LFRTs (8), despite multiple examinations of female reproductive tract proteins in the latter species including two high-throughput transcriptional studies (Swanson et al. 2004; Mack et al. 2006; Panhuis and Swanson 2006; Lawniczak and Begun 2007; Prokupek et al. 2008). All but two of these

Table 2
SFRSEs in *Drosophila melanogaster* and *Drosophila arizonae*

coding sequence (CDS)	189	216	226	Predicted Specificity	Secondary Domain
D. arizonae					
Dari/anon-EST:Kelleher5	Lys	Lys	Thr	Elastase?	
Dari/anon-EST:Kelleher6	Thr	Gly	Ala	Chymotrypsin	
Dari/anon-EST:Kelleher7	Ser	Gly	Arg	Unknown	
Dari/anon-EST:Kelleher8	Ser	Val	Asn	Elastase	
Dari/anon-EST:Kelleher10	Thr	Gly	Ala	Chymotrypsin	
Dari/anon-EST:Kelleher82	Thr	?	?	Unknown	
Dari/anon-EST:Kelleher267	?	?	?	Unknown	2 CLIP
Dari/anon-EST:Kelleher318	Asp	Gly	Thr	Unknown	
Dari/anon-EST:Kelleher361	Asp	?	?	Unknown	
Dari/anon-EST:Kelleher472	Gly	Gly	Gly	Unknown	CUB
Dari/anon-EST:Kelleher506	Met	Gly	Asp	Elastase?	
Dari/anon-EST:Kelleher580	Lys	?	?	Unknown	
Dari/anon-EST:Kelleher594	Asp	Gly	Gly	Trypsin	
Dari/anon-EST:Kelleher595	Asp	Gly	Gly	Trypsin	
Dari/anon-EST:Kelleher596	Gly	Ala	Ala	Unknown	
D. melanogaster					
Dmel/CG3066	Asp	Gly	Gly	Trypsin	CLIP
Dmel/Tequila	Asp	Gly	Gly	Trypsin	CBM_14\SCSR\Ldl_recept_a
Dmel/CG16705	Asp	Gly	Gly	Trypsin	CLIP
Dmel/CG17012	Gly	Thr	Thr	Unknown	
Dmel/CG17240	Asp	Gly	Gly	Trypsin	
Dmel/CG17239	Asp	Gly	Gly	Trypsin	
Dmel/CG17234	Ser	Val	Arg	Unknown	
Dmel/CG14642	Ser	Gly	Ser	Trypsin	

NOTE.—For each protease, key residues for substrate specificity 189, 216, and 226 as well as predicted specificity as in Perona and Craik (1995). Secondary protein–protein interaction domains were identified by eye (CLIP domains) or from previous reports (Ross et al. 2003; Kelleher et al. 2007). More details on protein domains can be found at (http://pfam.sanger.ac.uk/). The symbol "?" indicates that the relevant site was not included in the EST sequence.

D. arizonae SFRSEs, furthermore, lack secondary protein–protein interaction domains (table 2). The presence of such domains is important as they are common to insect serine endoproteases involved in physiological responses and developmental cascades and generally are absent in proteases whose primary function is nutritional digestion (Ross et al. 2003).

Serine endoproteases make effective digestive enzymes because they exhibit no absolute specificity in terms of recognizing the three-dimensional structure of their substrate. Rather, these enzymes show preferences for cleaving the scissile bond of a specific amino acid or set of amino acids, as determined by three key residues in the substrate-binding pocket (Perona and Craik 1995). Examination of these residues in *D. arizonae* SFRSEs suggests a broad range of specificities including all three major classes of digestive enzymes, trypsin, chymotrypsin, and elastase as well as several proteases with unpredictable specificity. *Drosophila melanogaster* SFRSEs, by comparison, present no evidence for chymotrypsin- or elastase-like activity, suggesting a narrower range of putative substrates.

Drosophila arizonae LFRTs Exhibit Significant Trypsin- and Elastase-Like Serine Endoprotease Activity

Our evolutionary and bioinformatic analyses suggest that recent gene duplication has enriched *D. arizonae* LFRTs for digestive serine endoproteases with a broad range of specificities including trypsin, chymotrypsin, and elastase (table 1). To test this hypothesis directly, we used chromogenic *p*-Nitroanilide substrates to detect proteolytic activity in LFRT lumens isolated from females in a mixture of mating states. Although chymotrypsin activity was not detected in *D. arizonae* LFRTs (data not shown), significant levels of trypsin- and elastase-like activity were exhibited by lumen isolated from these tissues (fig. 1). This activity decreased when isolated lumen was preincubated with the serine endoprotease inhibitors AEBSF (trypsin: $F_{1,6} = 102.57$, $P = 5.29 \times 10^{-5}$; elastase: $F_{1,6} = 41.04$, $P = 6.82 \times 10^{-4}$) and DFP (trypsin: $F_{1,6} = 184.64$, $P = 9.86 \times 10^{-6}$; elastase: $F_{1,6} = 4140.83$, $P = 9.47 \times 10^{-10}$), as expected if trypsin- and elastase-like activities are due to serine endoproteases (fig. 1).

To determine if trypsin- and elastase-like serine endoproteases could be derived from males during mating, we assayed *D. arizonae* SVAGs for serine endoprotease activity. Although the spectrophotometer detects absorbance at 405 nm, this value was not significantly different in assays preincubated with serine endoprotease inhibitors. Because these assays were not controlled for the inherent yellow pigment of *p*-Nitroanilide stock solution (see Materials and Methods), we conclude that this represents background absorbance from the chromogenic substrate rather than enzyme activity. These absorbance values, furthermore, are similar to values seen in blank solution containing only assay buffer and chromogenic substrate (data not shown). Although male-derived proteases could become activated only inside females (Ravi Ram et al. 2006), our data provide no evidence that trypsin- or elastase-like activity in *D. arizonae* female reproductive tracts originates in the male ejaculate.

Drosophila melanogaster LFRTs exhibit fewer serine endoproteases than *D. arizonae* and no predicted elastase-like serine endoproteases (table 1). Consistent with this

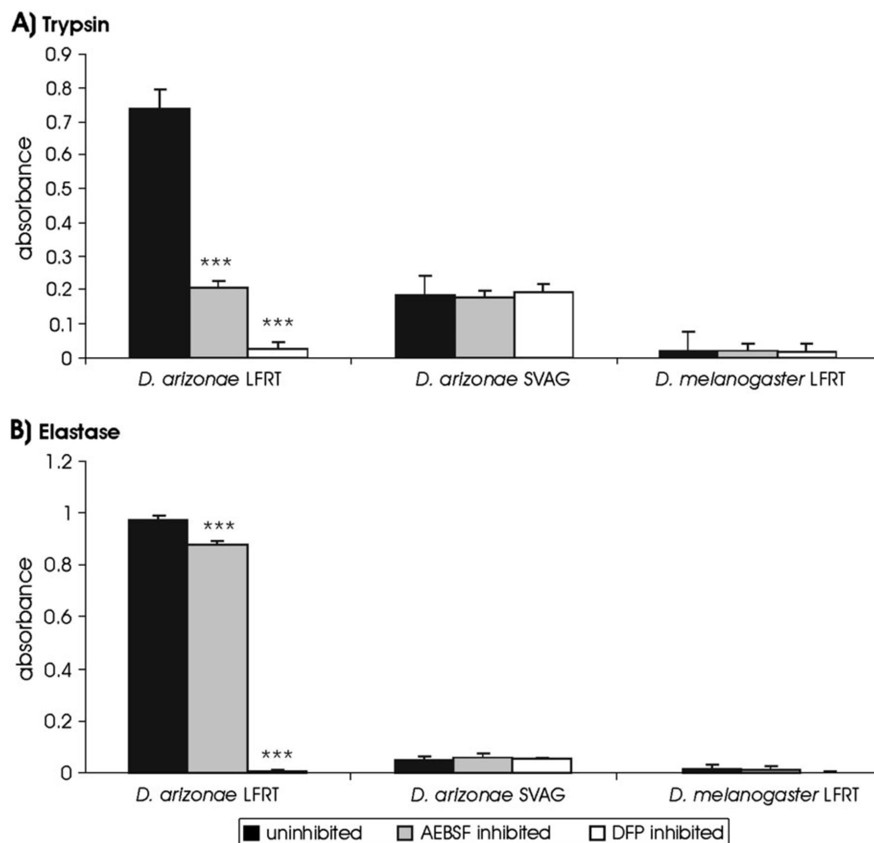

FIG. 1.—Serine endoprotease activity in the reproductive tissues of *Drosophila arizonae* females and males and *Drosophila melanogaster* females. Activity is measured as absorbance of the chromogenic (A) trypsin and (B) elastase substrate at 405 nm. Enzyme activity is decreased by preincubation with serine endoprotease inhibitors indicating the active proteases utilize serine in their active sites. * $P > 0.05$; **$P > 0.01$; ***$P > 0.001$.

observation, our enzyme assays detect minimal trypsin- or elastase-like activity in isolated LFRT lumen (fig. 1). Enzyme activity, furthermore, was not significantly reduced upon preincubation with serine endoprotease inhibitors (fig. 1), providing no evidence for serine endoprotease activity. Although it remains possible that the relative magnitude of detected activity would differ under other assay conditions, these data suggest that proteolytic capacity may present a significant physiological difference between *D. arizonae* and *D. melanogaster*.

Reduced proteolytic activity in mated females when compared with virgins suggests that SFRSEs are negatively regulated by the male ejaculate. Although it is possible that reduced activity could reflect competition between male-derived substrates and synthetic substrates for access to proteases, the magnitude of the observed decrease, particularly for trypsin-like enzymes, makes this explanation unlikely. Synthetic substrates are expected to be in considerable molar excess to proteases and endogenous substrates, minimizing the effect of dilution by endogenous molecules.

Serine Endoprotease Activity in *D. arizonae* Female Reproductive Tracts Is Negatively Regulated by Mating

To further elucidate the interaction between female proteases and the male ejaculate, we measured differences in trypsin- and elastase-like activity in matched cohorts of virgin and recently mated (<4 h postcopulation) *D. arizonae* females. Virgin females exhibit significant trypsin- and elastase-like acitivity, suggesting that the proteolytic activity detected here does not primarily originate in the male ejaculate. Both trypsin- and elastase-like activity, furthermore, were significantly reduced in mated female LFRT lumens when compared with virgins (trypsin: $F_{1,6} = 100.18$, $P = 5.76 \times 10^{-5}$; elastase: $F_{1,6} = 8.44$, $P = 0.027$; fig. 2), the opposite relationship of what would be expected if proteolytic acitivity was derived from males.

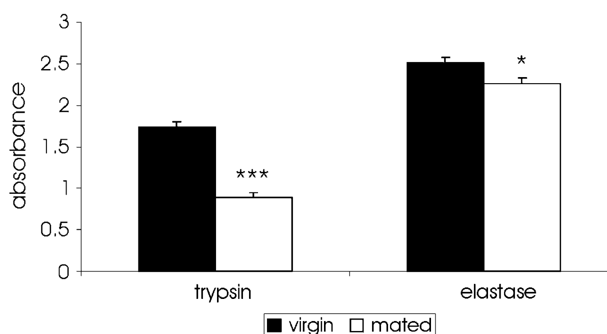

FIG. 2.—Serine endoprotease activity in *Drosophila arizonae* lower reproductive tracts is dependent on female mating status. Activity is absorbance of the chromogenic substrate at 405 nm. *$P > 0.05$; **$P > 0.01$; ***$P > 0.001$.

Table 3
Protein Evolution of SFRSEs

Drosophila arizonae EST	*Drosophila mojavensis* CDS	d_N	d_S	d_N/d_S
Dari\anon-EST:Kelleher5	Dmoj\anon-EST:Kelleher5	0.05	0.04	1.20
Dari\anon-EST:Kelleher5	Dmoj\anon-EST:Kelleher6	0.08	0.17	0.44
Dari\anon-EST:Kelleher8	Dmoj\anon-EST:Kelleher8	0.14	0.31	0.47
Dari\anon-EST:Kelleher7	Dmoj\anon-EST:Kelleher7	0.03	0.07	0.36
Dari\anon-EST:Kelleher10	No ortholog			
Dari\anon-EST:Kelleher82	Dmoj\GLEANR_12010	0.00	0.02	0.13
Dari\anon-EST:Kelleher267	Dmoj\GLEANR_17341	0.01	0.03	0.24
Dari\anon-EST:Kelleher318	Dmoj\GLEANR_2575	0.07	0.14	0.48
Dari\anon-EST:Kelleher361	Dmoj\GLEANR_3606	0.01	0.04	0.32
Dari\anon-EST:Kelleher472	Dmoj\GLEANR_5738	0.01	0.06	0.12
Dari\anon-EST:Kelleher506	Dmoj\GLEANR_6984	0.01	0.03	0.46
Dari\anon-EST:Kelleher580	DmojGLEANR_8733	0.03	0.07	0.39
Dari\anon-EST:Kelleher594	Dmoj\GLEANR_896	0.11	0.12	0.89
Dari\anon-EST:Kelleher596	Dmoj\GLEANR_898	0.05	0.12	0.44
Dari\anon-EST:Kelleher595	Dmoj\GLEANR_897	0.10	0.13	0.83
			Mean d_N/d_S = 0.48 ± 0.075	
Drosophila melanogaster CDS	*Drosophila simulans* CDS	d_N	d_S	d_N/d_S
Dmel/CG3066	Dsim/GLEANR_3734	0.02	0.12	0.16
Dmel/Tequila	Dsim/GLEANR_14168,14169	0.02	0.14	0.12
Dmel/CG16705	Dsim/GLEANR_4787	0.02	0.18	0.09
Dmel/CG17012	Dsim/GLEANR_6593	0.13	0.13	0.92
Dmel/CG17240	Dsim/GLEANR_6596	0.07	0.12	0.60
Dmel/CG17239	Dsim/GLEANR_6595	0.08	0.11	0.69
Dmel/CG17234	Dsim/GLEANR_6882	0.07	0.10	0.73
Dmel/CG14642	Dsim/GLEANR_3486	0.03	0.14	0.18
			Mean d_N/d_S = 0.44 ± 0.10	

NOTE.—Evolutionary rates were calculated between *D. melanogaster* and *D. arizonae* and their orthologs in the *D. simulans* and *D. mojavensis* genomes in PAML (Yang 1997). d_N, nonsynonymous substitutions per nonsynonymous site; d_S, synonymous substitutions per nonsynonymous site; and d_N/d_S, ratio nonsynonymous substitutions per nonsynonymous site to synonymous substitutions per nonsynonymous site.

Some *D. melanogaster* and *D. arizonae* SFRSEs Evolve Rapidly

Evolutionary rates of SFRSEs could serve as a metric to detect important differences in SFRSE dynamics between *D. arizonae* and *D. melanogaster*. We therefore estimated the ratio of replacement to silent substitutions (d_N/d_S) in both *D. arizonae* and *D. melanogaster* SFRSEs by comparing to their ortholog in the *D. simulans* and *D. mojavensis* genomes, respectively (table 3). Modest discrepancies between our results and previously reported values (Swanson et al. 2004) likely arise from the use of a *D. simulans* EST rather than the full length coding sequence in the previous study. We find no evidence for a difference in d_N/d_S between *D. melanogaster* and *D. arizonae* SFRSEs ($F_{1,22}$ = 0.13, P = 0.72), suggesting similar selective regimes in both lineages. We furthermore note that both data sets exhibit a high average d_N/d_S (*D. melanogaster* = 0.43, *D. arizonae* = 0.48) and several proteases with $d_N/d_S > 0.5$, suggestive of adaptive evolution (Swanson et al. 2004). Indeed, several of these proteins have been shown to experience positive selection in previous studies (Panhuis and Swanson 2006; Kelleher et al. 2007; Lawniczak and Begun 2007; Kelleher and Markow 2009).

Discussion

Our previous observation of lineage-specific gene families of secreted proteases in *D. arizonae* LFRT proteins suggested a recent, adaptive expansion of female reproductive proteolytic capacity (Kelleher et al. 2007; Kelleher and Markow 2009). The data presented here indicate that *D. arizonae* LFRT proteins are enriched for recent duplicates relative to its congener *D. melanogaster* and that this enrichment reflects preferential duplication of secreted proteases, particularly serine endoproteases. We furthermore show that *D. arizonae* female reproductive tracts exhibit a larger more diverse complement of serine endoproteases in their LFRTs as well as considerable trypsin- and elastase-like serine endoprotease activity that is regulated by mating. Collectively, our data suggest that SFRSEs exhibit divergent evolutionary dynamics and physiological functions between these two lineages.

Drosophila arizonae LFRT proteins are enriched for recently duplicated serine endoproteases when compared with those of *D. melanogaster*. This pattern reflects preferential duplication of serine endoproteases expressed in the LFRT rather than an elevated duplication rate in this enzymatic class as a whole. Intriguingly, male seminal proteins in the *repleta* species group also exhibit a high frequency of recent duplicates, although these paralogs are not clearly biased toward a particular functional class (Wagstaff and Begun 2007; Almeida and DeSalle 2008a; 2008b). Accelerated gene duplication rates, therefore, may be an important aspect of reproductive protein evolution within the *repleta* species group.

Although the selective force that underlies the exceptional frequency of gene duplications among *repleta*

species group reproductive proteins remains unclear, it is interesting to speculate that this pattern may arise from sexual conflict. Mathematical models of sexually antagonistic coevolution between interacting male and female molecules have predicted that it is adaptive for females to diversify in the face of pursuit by a male locus, and that male proteins may in turn diversify in response to females (Gavrilets and Waxman 2002; Hayashi et al. 2007). Although these models predict the rise of two divergent alleles at a single locus (Gavrilets and Waxman 2002; Hayashi et al. 2007), duplication and diversification of such loci would produce the same ultimate result. Intriguingly, *D. arizonae* females are three to five times more promiscuous than *D. melanogaster* (Reviewed Markow 1996), indicating that this lineage will experience comparatively more intense sexual conflict (Parker 1979).

The adaptive significance of preferential duplication of SFRSEs in the *D. arizonae*/*D. mojavensis* lineage is yet unclear. The bioinformatics analysis presented in this study, however, indicates that *D. arizonae* presents a larger number of SFRSEs in a broader range of predicted specificities than *D. melanogaster*. The majority of these proteins lack secondary protein–protein interaction domains, furthermore, suggesting their primary function is digestive (Ross et al. 2003). Consistent with this hypothesis, isolated lumen from *D. arizonae* LFRTs exhibits considerable trypsin- and elastase-like serine endoprotease activity reminiscent of gastrointestinal tracts (Billingsley and Hecker 1991; Oppert et al. 2002; Zhu et al. 2003), whereas no such activity is detected in *D. melanogaster*. The intense proteolytic environment presented by the *D. arizonae* female reproductive tract, therefore, may represent an important physiological difference from *D. melanogaster*.

Mated *D. arizonae* LFRTs exhibited significantly lower enzyme activity than virgin LFRTs, particularly for trypsin-like enzymes. This result appears counterintuitive; if female proteases cleave or degrade substrates in the male ejaculate, mating is predicted to be a positive regulator of proteolytic activity. If it is adaptive for males to avoid degradation of ejaculatory components due to sexual conflict, however, they may seek to negatively regulate female proteases. Mechanistically, this could be accomplished at either the transcriptional level or through protease inhibitors in the male ejaculate (Wagstaff and Begun 2005; Kelleher et al. 2009). Intriguingly, two protease inhibitors in the *D. mojavensis* ejaculate have experienced recent, lineage-specific gene duplication (Kelleher et al. 2009).

We previously have hypothesized that duplicated digestive proteases in *D. arizonae* LFRTs may be required to facilitate incorporation of ejaculate-derived protein, degradation of the insemination reaction, or both, in mated *D. arizonae* females (Kelleher et al. 2007). Adaptive male avoidance of female proteases is easy to envision in the context of this specialized reproductive physiology. If females are digesting important seminal proteins or sperm for their own nutritional purposes, this could be extremely costly to males. Alternatively, it may be adaptive for males to encumber female degradation of the ejaculate-induced insemination reaction. Indeed, the reaction mass is thought to be a male strategy to delay female remating and ensure paternity (Markow and Ankney 1984, 1988; Pitnick et al. 1997), and male–female conflict over the size and duration of the insemination reaction previously has been proposed (Knowles and Markow 2001).

Supplementary Material

Supplementary tables 1-3 and figure 1 are available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

The authors would like to acknowledge Roger Miesfeld for generous use of equipment and reagents and Therese Markow, Willie Swanson, Jeremy Bono, Vanessa Corby-Harris, and three anonymous reviewers for helpful comments that significantly improved the manuscript. This research was funded by a Doctoral Dissertation Improvement Grant to E.S.K. and National Institutes of Health grant AI31951 to R.L.M. E.S.K. was supported by an National Science Foundation–interdisciplinary graduate education and research trainning research traineeship in Evolutionary, Functional and Computational Genomics at the University of Arizona and a Dissertation Fellowship from the American Association of University Women.

Literature Cited

Al-Shahrour F, Díaz-Uriarte R, Dopazo J. 2004. FatiGO: a web tool for finding significant associations of gene ontology terms with groups of genes. Bioinformatics. 20:578–580.

Al-Shahrour F, Minguez P, Tárraga J, Medina I, Alloza E, Montaner D, Dopazo J. 2007. FatiGO +: a functional profiling tool for genomic data. Integration of functional annotation, regulatory motifs and interaction data with microarray experiments. Nucleic Acids Res. 35:W91–W96.

Almeida FC, DeSalle R. 2008a. 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. 2008b. Orthology, function, and evolution of accessory gland proteins in the *Drosophila repleta* group. Genetics. 181:235–245.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J Mol Biol. 215:403–410.

Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI BLAST: a new generation of protein database search programs. Nucleic Acids Res. 25:3389–3402.

Ashburner M, Ball CA, Blake JA, et al. (20 co-authors). 2000. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet. 25:25–29.

Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J Roy Stat Soc B. 57:289–300.

Billingsley PF, Hecker H. 1991. Blood digestion in the mosquito, *Anopheles stephensi* Liston (Diptera: Culicidae): activity and distribution of trypsin, aminopeptidase, and alpha-glucosidase in the midgut. J Med Entomol. 28:865–871.

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.

Brennan PL, Prum RO, McCracken KG, Sorenson MD, Wilson RE, Birkhead TR. 2007. Coevolution of male and female genital morphology in waterfowl. PLoS ONE. 2:e418.

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

Dacheux JL, Gatti JL, Dacheux F. 2003. Contribution of epididymal secretory proteins for spermatozoa maturation. Microsc Res Tech. 61:7–17.

Eddy SR. 1998. Profile hidden Markov models. Bioinformatics. 14:755–763.

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, Waxman D. 2002. Sympatric speciation by sexual conflict. Proc Natl Acad Sci USA. 99:10533–10538.

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.

Hahn MW, Han MV, Han SG. 2007. Gene family evolution across 12 *Drosophila* genomes. PLoS Genet. 3:e197.

Hayashi TI, Vose M, Gavrilets S. 2007. Genetic differentiation by sexual conflict. Evolution. 61:516–529.

Hey J, Kliman RM. 1993. Population genetics and phylogenetics of DNA sequence variation at multiple loci within the *Drosophila melanogaster* species complex. Mol Biol Evol. 10:804–822.

Jiang H, Kanost MR. 2000. The clip-domain family of serine proteinases in arthropods. Insect Biochem Mol Biol. 30: 95–105.

Jiang H, Wang Y, Kanost MR. 1998. Pro-phenol oxidase activating proteinase from an insect, Manduca sexta: a bacteria-inducible protein similar to Drosophila easter. Proc Natl Acad Sci USA. 95:12220–12225.

Kelleher ES, Markow TA. 2009. Duplication, selection and gene conversion in a *Drosophila mojavensis* female reproductive protein family. Genetics. 181:1451–1465.

Kelleher ES, Swanson WJ, Markow TA. 2007. Gene duplication and adaptive evolution of digestive proteases in *Drosophila arizonae* female reproductive tracts. PLoS Genet. 3:1541–1549.

Kelleher ES, Watts TD, LaFlamme BA, Haynes PD, Markow TA. 2009. Proteomic analysis of *Drosophila mojavensis* male accessory glands suggests novel classes of seminal fluid proteins. Insect Biochem Mol Biol. 39: 366–371.

Knowles LL, Markow TA. 2001. Sexually antagonistic coevolution of a postmating prezygotic reproductive character in desert Drosophila. Proc Nat Acad Sci USA. 98:8692–8696.

Lawniczak MK, Begun DJ. 2007. Molecular population genetics of female-expressed mating-induced serine proteases in *Drosophila melanogaster*. Mol Biol Evol. 24:1944–1951.

Mack PD, Kapelnikov A, Heifetz Y, Bender M. 2006. Mating-responsive genes in reproductive tissues of female Drosophila melanogaster. Proc Natl Acad Sci USA. 103:10358–10363.

Markow TA. 1996. Evolution of *Drosophila* mating systems. Evol Biol. 29:73–106.

Markow TA, Ankney PF. 1984. *Drosophila* males contribute to oogenesis in a multiple mating species. Science. 224: 302–303.

Markow TA, Ankney PF. 1988. Insemination reaction in *Drosophila* found in species whose males contribute material to oocytes before fertilization. Evolution. 42:1097–1101.

Marshall JL. 2007. Rapid evolution of spermathecal duct length in the *Allonemobius socius* complex of crickets: species, population and *Wolbachia*effects. PLoS ONE. 2:e720.

Matzkin LM. 2004. Population genetics and geographic variation of alcohol dehydrogenase (Adh) paralogs and glucose-6-phosphate dehydrogenase (G6pd) in *Drosophila mojavensis*. Mol Biol Evol. 21:276–285.

Monsma SA, Harada HA, Wolfner MF. 1990. Synthesis of two *Drosophila* male accessory gland proteins and their fate after transfer to the female during mating. Dev Biol. 142: 465–475.

Neubaum DM, Wolfner MF. 1999. Wise, winsome, or weird? Mechanisms of sperm storage in female animals. Curr Top Dev Biol. 41:67–97.

Oppert B, Hartzer K, Zuercher M. 2002. Digestive proteinases in *Lasioderma serricorne* (Coleoptera: Anobiidae). Bull Entomol Res. 92:331–336.

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. 2006. Molecular evolution and population genetic analysis of candidate female reproductive genes in *Drosophila*. Genetics. 173:2039–2047.

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.

Parker GA. 1979. Sexual selection and sexual conflict. In: Blum MS, Blum NA, editors. Sexual selection and reproductive competition in insects. London: Academic Press. p. 123–166.

Patterson JT. 1946. A new type of isolating mechanism in Drosophila. Proc Nat Acad Sci USA. 32:202–208.

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

Perona JJ, Craik CS. 1995. Structural basis of substrate specificity in the serine proteases. Protein Sci. 4:337–360.

Pilpel N, Nezer I, Applebaum SW, Heifetz Y. 2008. Mating-increases trypsin in female *Drosophila* hemolymph. Insect Biochem Mol Biol. 38:320–330.

Pitnick S, Markow TA, Spicer GS. 1999. Evolution of multiple kinds of female sperm-storage organs in *Drosophila*. Evolution. 53:1804–1822.

Pitnick S, Spicer GS, Markow TA. 1997. Phylogenetic examination of female incorporation of ejaculate in *Drosophila*. Evolution. 51:833–845.

Polgar L. 2005. The catalytic triad of serine peptidases. Cell Mol Life Sci. 62:2161–2172.

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

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

Rice WR. 1996. Sexually antagonistic male adaptation triggered by experimental arrest of female evolution. Nature. 381:232–234.

Robertson SA. 2007. Seminal fluid signaling in the female reproductive tract: lessons from rodents and pigs. J Anim Sci. 85:E36–E44.

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.

Swanson WJ, Clark AG, Waldrip-Dail HM, Wolfner MF, Aquadro CF. 2001. Evolutionary EST analysis identifies

rapidly evolving male reproductive proteins in *Drosophila*. Proc Natl Acad Sci USA. 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.
Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22:4673–4680.
Wagstaff BJ, Begun DJ. 2005. Molecular population genetics of accessory gland protein genes and testis-expressed genes in *Drosophila mojavensis* and *D. arizonae*. Genetics. 171: 1083–1101.
Wagstaff BJ, Begun DJ. 2007. Adaptive evolution of recently duplicated accessory gland protein genes in desert *Drosophila*. Genetics. 177:1023–1030.
Wolfner MF. 2007. "S.P.E.R.M." (seminal proteins (are) essential reproductive modulators): the view from Drosophila. Soc Reprod Fertil. (Suppl. 65):183–199.
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.
Yang Z. 1997. PAML: a program package for phylogenetic analysis by maximum likelihood. Comput Appl Biosci. 13: 555–556.
Zhu YC, Zeng F, Oppert B. 2003. Molecular cloning of trypsin-like cDNAs and comparison of proteinase activities in the salivary glands and gut of the tarnished plant bug *Lygus lineolaris* (Heteroptera: Miridae). Insect Biochem Mol Biol. 33:889–899.

Jody Hey, Associate Editor

Accepted June 12, 2009