Insect Biochemistry and Molecular Biology 39 (2009) 366–371

Contents lists available at ScienceDirect

Insect Biochemistry and Molecular Biology

journal homepage: www.elsevier.com/locate/ibmb

Proteomic analysis of *Drosophila mojavensis* male accessory glands suggests novel classes of seminal fluid proteins

Erin S. Kelleher [a,b,*], Thomas D. Watts [a], Brooke A. LaFlamme [a,b], Paul A. Haynes [c,d], Therese A. Markow [a,e]

[a] Department of Ecology and Evolutionary Biology, University of Arizona, Tucson, AZ 85721, USA
[b] Department of Molecular Biology and Genetics, Cornell University, Ithaca, NY 14850, USA
[c] Department of Chemistry and Biomolecular Sciences, Macquarie University, North Ryde, NSW 2109, Australia
[d] Department of Biochemistry and Molecular Biophysics, University of Arizona, Tucson, AZ 85721, USA
[e] Section of Ecology, Behavior and Evolution, Division of Biological Sciences, University of California at San Diego, La Jolla, CA, USA

ARTICLE INFO

Article history:
Received 31 December 2008
Received in revised form
26 February 2009
Accepted 6 March 2009

Keywords:
Drosophila
Seminal protein
Reproductive protein
Ejaculate

ABSTRACT

Fruit-flies of the genus *Drosophila* are characterized by overwhelming variation in fertilization traits such as copulatory plug formation, sperm storage organ use, and nutritional ejaculatory donation. Despite extensive research on the genetic model *Drosophila melanogaster*, little is known about the molecular underpinnings of these interspecific differences. This study employs a proteomic approach to pin-point candidate seminal fluid proteins in *Drosophila mojavensis*, a cactophilic fruit-fly that exhibits divergent reproductive biology when compared to *D. melanogaster*. We identify several classes of candidate seminal fluid proteins not previously documented in the *D. melanogaster* male ejaculate, including metabolic enzymes, nutrient transport proteins, and clotting factors. Conversely, we also define 29 SFPs that are conserved despite >40 million years of *Drosophila* evolution. We discuss our results in terms of universal processes in insect reproduction, as well as the specialized reproductive biology of *D. mojavensis*.

1. Introduction

In internally fertilizing organisms, sexual reproduction is mediated by a complex series of biochemical interactions between the male ejaculate and the female reproductive tract. These interactions extend beyond sperm and egg, encompassing a broad range of physiological processes within mated females. Sperm must navigate through the female reproductive tract, remain viable in this environment, and ultimately fertilize female gametes (reviewed in Neubaum and Wolfner, 1999). Ejaculate–female interactions also are known to have important impacts on female behavior and physiology, such as upregulating immune response, stimulating ovulation, and preparing the uterus for embryonic implantation (Reviewed in Wolfner, 2007; Robertson, 2007).

In insects, the male ejaculate comprises a complex cocktail of inorganic solutes (Smedley and Eisner, 1995; Markow et al., 2001), lipids (Butterworth, 1969; Brieger and Butterworth, 1970), and seminal proteins (Swanson et al., 2001; Braswell et al., 2006; Sirot et al., 2008). The physiological and biochemical function of these molecules, particularly seminal fluid proteins (SFPs), has been most extensively studied in the genetic model *D. melanogaster*. In these animals, no fewer than 138 unique proteins in an array of biochemical classes are passed from males to females during copulation (Swanson et al., 2001; Mueller et al., 2005; Walker et al., 2006; Findlay et al., 2008). *D. melanogaster* SFPs play integral roles in the female post-mating response by modulating oogenesis, ovulation, immune response, sperm storage, female refractoriness, and feeding behavior (Reviewed in Wolfner, 2007).

Despite the essential nature of ejaculate–female interactions for fertilization and fitness, it frequently is observed that reproductive traits are amongst the most divergent between closely related organisms. In particular, the genus *Drosophila* exhibits overwhelming variation in sperm size, sperm number, sperm storage organ utilization, female refractoriness, female incorporation of ejaculate derived protein, and copulatory plug and insemination reaction formation (reviewed in Markow, 1996, 2002). These extensively documented biological differences, as well as the recent advent of genetic and genomic tools to 12 *Drosophila* species (Drosophila 12 Genomes Consortium, 2007; Zhang et al., 2007), makes *Drosophila* an exciting system in which to elucidate the molecular underpinnings of interspecific differences in reproductive tract interactions.

This study seeks to identify candidate seminal fluid proteins in *Drosophila mojavensis*, a cactophilic *Drosophila* that exhibits

* Corresponding author. Department of Molecular Biology and Genetics, 403 Biotech, Cornell University, Ithaca, NY 14850, USA. Tel.: +1 607 254 4569; fax: +1 607 255 6249.
E-mail address: esk72@cornell.edu (E.S. Kelleher).

0965-1748/$ – see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ibmb.2009.03.003

divergent reproductive biology when compared to *D. melanogaster* (Markow, 1996, 2002). *D. mojavensis* ejaculates are known to contain a nutritional donation that is incorporated through the mated female reproductive tract, while no such incorporation is seen in *D. melanogaster* (Markow and Ankney, 1984). *D. mojavensis* males also induce an insemination reaction in mated females, an opaque, clot-like mass that fills the uterus until degraded (Patterson, 1946). *D. mojavensis* sperm are stored only in the female seminal receptacle, while *D. melanogaster* females use both the seminal receptacle and the paired spermathecae for sperm storage (Pitnick et al., 1999). Finally, *D. mojavensis* females copulate daily, on average, while *D. melanogaster* females intersperse copulations by several days (Markow, 1996).

We employ a proteomic approach to isolate 786 unique proteins from the *D. mojavensis* male accessory gland, the primary source of *Drosophila* SFPs (Reviewed in Wolfner, 2002). 240 of these molecules are identified as candidate seminal fluid proteins based on expression data, bioinformatics analyses, and peptide abundance. Conserved SFPs between *D. melanogaster* and *D. mojavensis*, as well as novel components of the *D. mojavensis* male ejaculate are furthermore identified. We discuss our results in terms of universal processes in insect reproduction, as well as the specialized reproductive biology of *D. mojavensis*.

2. Materials and methods

2.1. Drosophila cultures

D. mojavensis used in this study were collected as a multi-female line by T.A.M. in the area of San Carlos, Sonora, Mexico in November of 2001. They were cultured in the lab under low-density conditions in opuntia/banana medium (http://flyfood.arl.arizona.edu/opuntia.php3) in half pint milk bottles.

2.2. Tissue collection

Adults were collected from half pint milk bottles on the day of eclosion and kept in yeasted vials of opuntia/banana food separated by sex until sexually mature (9–12 days). Males were then anesthetized under CO_2 and four replicates of ~50 paired accessory glands were excised in cold de-ionized water, and pooled. Two replicates were snap frozen in liquid N_2 and stored at −80 °C, while the other two replicates were spun at 6000 × g for 5 min and supernatants were removed and assayed as soluble protein fractions.

2.3. Protein fractionation and digestion

100 μg of protein was fractionated on a 10% linear SDS-PAGE minigel (Bio-Rad) according to established methods. Proteins present in the gel lane were visualized with silver staining, and the gel lane was then cut into 32 equal-sized pieces of approximately 2 mm each with a scalpel and transferred to a 96-well plate (ABgene) for further processing (Breci et al., 2005).

Silver stained gel bands were destained (Gharahdaghi et al., 1999) and digested (Wilm et al., 1996) using a Multiprobe-II liquid handling system (Perkin Elmer, Shelton, CT). Following digestion, tryptic peptides were extracted from the gel pieces with 5% formic acid/50% CH_3CN. The combined extracted peptides were concentrated to 10 μl using a Speedvac vacuum centrifuge (Savant, Farmingdale, NY).

2.4. Nanoflow High performance liquid chromatography – tandem mass spectrometry

A microbore HPLC system (Surveyor, Thermo, San Jose, CA) was modified to operate at capillary flow rates using a simple T-piece flow-splitter. Columns (8 cm × 100 μm I.D.) were prepared by packing 100 Å, 5 μm Zorbax C18 resin at 500 psi pressure into columns with integrated electrospray tips made from fused silica, pulled to a 5-μm tip using a laser puller (Sutter Instrument Co., Novato, CA). Electrospray voltage of 1.8 kV was applied using a gold electrode via a liquid junction upstream of the column. Samples were introduced onto the analytical column using a Surveyor autosampler (Surveyor, Thermo, San Jose, CA). The HPLC column eluent was eluted directly into the electrospray ionization source of a Thermo LCQ Deca XP Plus ion trap mass spectrometer (Thermo, San Jose, CA).

Peptides were eluted in a gradient using buffer A (0.1% formic acid) and buffer B (acetonitrile, 0.1% formic acid), at a flow rate of 500 nL/min. Following an initial wash with buffer A for 10 min, peptides were eluted with a linear gradient from 0 to 50% buffer B over a 60-min interval, followed by 50–98%B over 5 min and a 5-min wash at 98%B. Automated peak recognition, dynamic exclusion, and daughter ion scanning of the top three most intense ions were performed using the Xcalibur software as previously described: i) full mass survey scan 400–1500 amu, ii) MS/MS of most abundant ion from survey scan, iii) MS/MS of 2nd most abundant ion from survey scan, iv) MS/MS of 3rd most abundant ion from survey scan. Other instrument parameters included: collision energy 39%, activation Q 0.25, activation time 30 ms, isolation width 2.0 amu, dynamic exclusion enabled with repeat count 2, duration 0.5 min, exclusion duration 5 min, exclusion mass width low 1.5 amu, high 1.5 amu (Hattrup et al., 2007).

2.5. Genome sequences and protein annotations

The release 5 complete genomic sequence from *D. melanogaster* was downloaded from the Berkley *Drosophila* Genome Project (http://www.fruitfly.org/sequence/release5genomic.shtml), and the annotated proteins from release 5 were downloaded from flybase (ftp://ftp.flybase.net/releases/FB2008_05/dmel_r5.8/fasta). The CAF1 assemblies of the *D. mojavensis*, *Drosophila grimshawi*, and *Drosophila virilis*, genomes, as well as the GLEANR predicted proteins for these genomes were obtained from the Assembly, Annotation and Alignment of 12 *Drosophila* species (http://rana.lbl.gov/~venky/AAA/freeze_20061030/).

2.6. Database searching and result filtering and validation

MS/MS data were analyzed using SEQUEST run under Bioworks 3.1 (Thermo, San Jose, CA). Dynamic peptide modification by oxidation of methionine and static modification by carbamidomethylation of cysteine were considered. The peptide and fragment mass tolerance were 2 Da and 0.2 Da, respectively. Database search results were filtered and organized using DTAselect and Contrast (Tabb et al., 2002).

Experimental mass spectra were searched against a database of theoretical mass spectra generated from all annotated proteins in the *D. mojavensis* genome. To identify additional proteins not present in the gene predictions, we used nr6frame (D. states, unpublished program) to generate a complete 6-frame translation genome sequence. Filtering these sequences for short ORFs (<11 residues), ORFs that were exclusively monoresidue repeats, and ORFs that could not produce a tryptic peptide, resulted in a second database of ~9.1 million ORFS. Proteins identified in the sample that matched the 6-frame translation database were compared to all predicted proteins using blastP (*e*-value = 10^{-10}, Altschul, 1990), and queries with a significant hit were removed. The remaining unannotated proteins, were concatenated with the identified GLEANR proteins, to make a consensus set of 786 unique proteins.

In this work, the initial criteria for a preliminary positive peptide identification for a doubly-charged peptide were a correlation factor (Xcorr) greater than 2.5, a delta cross-correlation factor (dCn) greater than 0.1, a minimum of one tryptic peptide terminus, and a high preliminary scoring (Andon et al., 2003). For triply-charged peptides the correlation factor threshold was set at 3.5, and for singly-charged peptides the threshold was set at 1.8. At least two independent tryptic fragments were required for positive peptide identification.

For all annotated proteins, we calculated and ranked the mean normalized spectral abundance factor (NSAF) across all experiments in which the protein was found (Florens et al., 2006; Paoletti et al., 2006). Briefly, the NSAF is the total number of peptides of a given protein identified in an experimental sample, normalized by the total length of the protein, normalized by the sum of this value for all proteins in the sample. Ranked mean NSAF therefore estimates the relative abundance of a given protein across all experiments.

2.7. Bioinformatic analyses

All 786 proteins were annotated for seven criteria: 1) predicted signal sequence from SignalP (Bendtsen et al., 2004), 2) predicted signal anchor from SignalP, 3) predicted transmembrane domain from tmhmm (Sonnhammer et al., 1998) 4) male-biased gene expression (Zhang et al., 2007) 5) presence in both soluble fractions (see tissue collection) 6) *D. melanogaster*, *D. grimshawi*, and *D. virilis* annotated protein homologs (blastP $e = 0.001$, Altschul, 1990), and 7) *D. melanogaster*, *D. grimshawi*, and *D. virilis* genomics homology (tblastn $e = 0.001$, Alschul, 1995). These annotations, as well as the mean NSAF for each protein, can be found in Supplementary Table 1.

240 candidate SFPs were additionally functionally annotated based on data from three sources 1) gene ontology (Ashburner et al., 2000) terms for the *D. melanogaster* homolog obtained from flybase (http://flybase.org/static_pages/downloads/ID.html), 2) conserved pfam domains, predicted by hmmpfam (Eddy, 1998) 3) annotated sperm proteins for *D. melanogaster* (Dorus et al., 2006). Functional annotations of candidate SFPs can be found in Supplementary Table 2.

3. Results

3.1. Protein identification and abundance

In four experiments, a total of 786 unique proteins were identified from harvested *D. mojavensis* male accessory glands (Supplementary Table 1). 766 of these proteins corresponded to GLEANR gene predictions for the *D. mojavensis* genome, while an additional twenty proteins were identified by searching against a 6-frame translation of the complete genome sequence. Although gene structure of proteins identified by 6-frame translation requires experimental validation, this approach greatly enriched the efficacy of our shot-gun proteomic approach for accurate peptide identification.

300 proteins were identified in multiple experiments, while the remaining 486 were present in only a single sample. To estimate the abundance of each protein in *D. mojavensis* male accessory glands, we calculated the normalized spectral abundance factor (NSAF) for each protein (Supplementary Table 1). The total number of experiments in which a protein was sampled was highly correlated with the mean NSAF across all experiments in which that protein was found ($r^2 = 0.56$, $F_{1,764} = 984.88$, $p < 0.0001$).

3.2. Conserved Drosophila SFPs

To pin-point conserved components of the *Drosophila* male ejaculate, we identified *D. mojavensis* accessory gland proteins (ACPs) that exhibited the highest blast identity to a known *D. melanogaster* SFP (Findlay et al., 2008). Conservation of these SFPs across > 40 million years of *Drosophila* evolution implies an integral role in male reproduction, providing excellent candidates for future genetic studies. We discovered 31 *D. mojavensis* ACPs that corresponded to 29 *D. melanogaster* SFPs (two SFPs are duplicated in *D. mojavensis*), in an array of functional classes including proteolysis, lipid metabolism, carbohydrate interactions, and defense/immunity (Table 1). To determine if specific functional classes were more or less likely to be conserved, we performed 2 × 2 contingency tests for the representation of a given class in conserved versus unconserved SFPs. No functional classes were significantly over or under-represented, implying similar rates of SFP turnover.

Table 1
Conserved SFPs between *D. melanogaster* and *D. mojavensis*. *D. melanogaster* SFPs – total number of *D. melanogaster* SFPs in each functional class from Findlay et al. (2008). Conserved – number of *D. mojavensis* SFPs with homologs identified in *D. mojavensis* male accessory glands.

	D. melanogaster SFPs	Conserved
unknown	47	9
protease	15	6
protease inhibitor	14	1
defense immunity	12	1
lipid metabolism	9	2
carbohydrate interaction	7	2
odorant binding	7	0
other functions	7	3
post-mating behavior	5	1
sperm protein	5	2
chitin binding	4	0
DNA interactions	3	1
protein modification	3	1
Total	138	29

3.3. Identification of Candidate D. mojavensis SFPs

To gain a more comprehensive view of the *D. mojavensis* male ejaculate, we sought to identify additional candidate SFPs that are not shared with *D. melanogaster*. Each of the 786 proteins were annotated for five possible indicators of inclusion in the male ejaculate 1) male-biased expression 2) secreted signal sequence 3) signal anchor 4) transmembrane domain, and 5) presence in both soluble fractions (Supplementary Table 2). Contingency tests were then employed to indicate which criteria enriched for presence of the 31 conserved seminal proteins (Table 2). Proteins with secreted signal-sequences, male-biased expression, and representation in both soluble fractions were significantly enriched for conserved seminal proteins, while proteins with signal anchors and transmembrane domains were not (Table 2). Candidate SFPs, therefore, are categorized as all proteins that exhibit a signal sequence, male-biased expression, or are found in both soluble fractions. Obvious

Table 2
Enrichment for conserved SFPs based on five criteria. Total # of proteins = number of candidate genes that exhibit the given criteria. # of conserved SFPs = number of candidate genes that exhibit the given criteria and are amongst the 31 conserved SFPs between in *D. mojavensis*. *p*-value = significance of the Fisher's exact test that the proteins exhibiting a given criteria are enriched for conserved SFPs relative to those who do not.

criteria	Total # of proteins	# of conserved SFPs	*p*-value
male-biased expression	104	17	9.73×10^{-9}
signal sequence	186	18	2.81×10^{-5}
signal anchor	27	2	0.31
transmembrane domain	114	2	0.29
both soluble fractions	40	10	2.64×10^{-6}

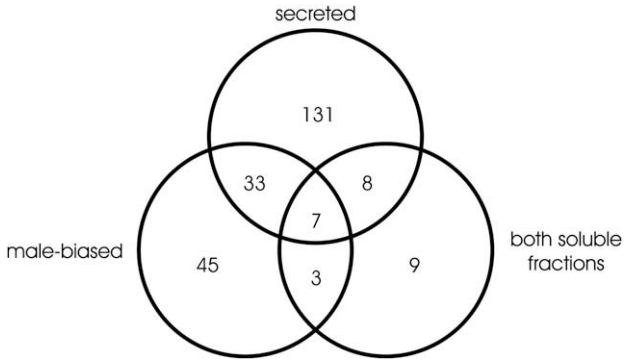

Fig. 1. Venn Diagram of criteria for *D. mojavensis* SFP candidates. Male-biased expression was determined by Zhang et al. (2007). Secreted signal-sequences were identified with SignalP (Bendtsen et al., 2004).

ribosomal protein (7) and sperm protein (25) contaminants were excluded from the candidate genes, yielding a total of 240 proteins (Fig. 1).

Using flybase annotations and conserved protein domains, our 240 candidate genes were further categorized by functional class (Supplementary Table 2, Fig. 2). Functional classes such as proteases, CRISPs, carbohydrate interaction are found not only in *D. melanogaster* SFPs, they also are common to mammalian ejaculates (Mueller et al., 2004). Proteins involved in nutrient transport and metabolism, however, have not previously been identified as SFPs in *D. melanogaster* (Swanson et al., 2001; Findlay et al., 2008). These molecules therefore point to a potential difference in the biological function of the *D. mojavensis* and *D. melanogaster* male ejaculates.

3.4. Novel SFPs in D. mojavensis

Rapid gain and loss of individual SFPs is a widespread phenomenon of ejaculate evolution in *Drosophila* (Begun and Lindfors, 2005; Mueller et al., 2005; Begun et al., 2006; Findlay et al., 2008). Although some lineage-specific SFPs may simply have non-reproductive functions in outgroup species, novel SFPs also are known to evolve both through duplication and divergence (Wagstaff and Begun, 2007; Findlay et al., 2008; Almeida and Desalle, 2008a,b), and from unidentified and possibly non-coding sequence (Begun et al., 2006; Findlay et al., 2008). Because novel SFPs are more likely involved in reproductive processes that are unique to *D. mojavensis*, we sought to identify *D. mojavensis* accessory gland proteins with no homology to any other *Drosophila* proteins, as well as lineage-specific duplicates.

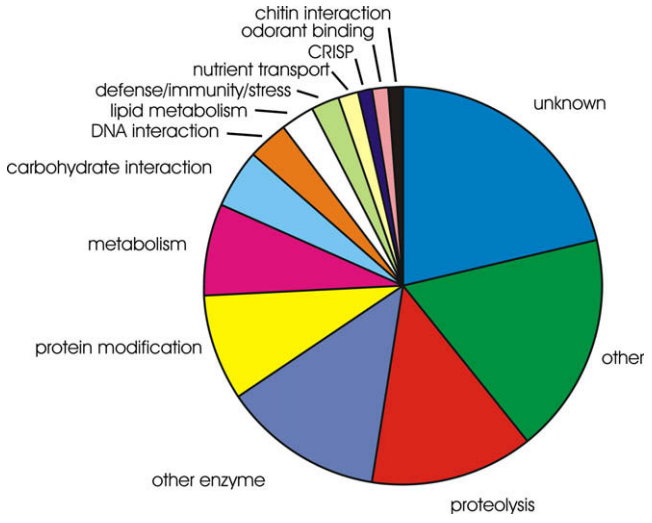

Fig. 2. Functional classes in 240 candidate D. mojavensis SFPs. Functional class was assigned based on conserved proteins domains, and gene ontology terms for the *D. melanogaster* homolog (Supplementary Table 2).

Table 3
Lineage-specific gene families in *D. mojavensis* candidate SFPs. Function – annotated function based on conserved proteins domains, and gene ontology terms for the *D. melanogaster* homolog (Supplementary Table 2). *D. mojavensis* candidate paralogs – paralogs that are candidate *D. mojavensis* SFPs identified in this study. Additional paralogs – paralogs.

annotated function	*D. mojavensis* candidate paralogs	*D. mojavensis* non-candidate paralogs	*D. virilis* ortholog
cysteine protease inhibitor	GLEANR_10569, GLEANR_10570	GLEANR_10572, GLEANR_10571	GLEANR_4152
acetyltransferase	GLEANR_1147, GLEANR_2327		GLEANR_3032
protein metabolism	GLEANR_13509, GLEANR_13510		GLEANR_13785
serine endoprotease	GLEANR_3287	GLEANR_3286, GLEANR_3285	GLEANR_6167
unknown	GLEANR_4380	GLEANR_4381	GLEANR_6576
metabolism	GLEANR_830, GLEANR_831		GLEANR_899
unknown	GLEANR_8821	GLEANR_9082	GLEANR_9298

We identified 15 (of 786) *D. mojavensis* ACPs with no detectable blast identity to either the GLEANR annotated proteins or complete coding sequences for the *D. melanogaster*, *D. grimshawi*, or *D. virilis* genomes (http://rana.lbl.gov/~venky/AAA/freeze_20061030/). These proteins are excellent candidates for novel genes of unknown origin; however, it remains possible that they have evolved so rapidly that all detectable homology has been obscured. Interestingly, only one of these proteins is present in the GLEANR gene predictions for *D. mojavensis*, while the remaining 14 proteins were identified from our 6-frame translation of the *D. mojavensis* genome. Although, the authenticity of these proteins will need to be verified experimentally, our data and others (Findlay et al., 2008) suggest that 6-frame translation is an important tool for identifying novel coding sequences.

All instances in which multiple proteins in the *D. mojavensis* genome, including at least one of our 786 identified proteins, exhibited higher blast identity to each other than to their homolog in *D. virilis*, were categorized as lineage-specific duplicates (Supplementary Table 3). 16 *D. mojavensis* ACPs are associated with lineage-specific gene duplications, of which 11 are candidate SFPs (Table 2). Candidate SFPs, therefore, are greatly enriched in lineage-specific duplicates when compared to non-candidates in our data set (Fisher's Exact Test $p = 0.002$). SFP gene families are involved in proteolysis and metabolism, as well as unknown and likely novel functions (Table 3).

4. Discussion

Our study presents the first comprehensive analysis of *D. mojavensis* male accessory glands, the primary site of SFP production in *Drosophila* (Reviewed in Wolfner, 2002). The divergence between *D. mojavensis* and its congener, the genetic model *D. melanogaster* empowered us to identify integral components of the *Drosphila* male ejaculate; 29 SFPs conserved despite >40 million years of *Drosophila* evolution. These proteins make exciting candidates for future genetic studies, and will likely yield important insights into the biochemical underpinnings of insect reproduction. Conversely, the 209 additional candidate SFPs identified here are of

interest in light of the comparative reproductive biology of the genus *Drosophila*, as well as the specialized functions of the *D. mojavensis* ejaculate.

Metabolic enzymes and related proteins were prominent in *D. mojavensis* male accessory glands, comprising almost 8% of candidate SFPs (Fig. 2). Intriguingly, two of six identified lineage-specific candidate SFP families also encode metabolic enzymes. These classes of enzymes were not prominent amongst *D. melanogaster* SFPs (Findlay et al., 2008), nor amongst secreted proteins expressed in *D. melanogaster* male accessory glands (Swanson et al., 2001), pointing to an important biological difference between these two species. Whether metabolic processes are specific to the male accessory gland, occur in female reproductive tracts, or both, remains unknown.

Similar to metabolism, four candidate *D. mojavensis* SFPs have annotated functions in nutrient transport, while no such proteins have been implicated in the *D. melanogaster* male ejaculate (Swanson et al., 2001; Findlay et al., 2008). The candidate genes include 1 phosphate transporter, 1 lipid transporter, and 2 transporters with nutrient reservoir activity. Intriguingly, both phosphorus (Markow et al., 2001), and lipids (Butterworth, 1969; Brieger and Butterworth, 1970) have been described as components of the *D. melanogaster* male ejaculates, although little is known about the role of these biomolecules in ejaculates of other *Drosophila* species. The presence of secreted transport proteins for phosphorus and lipids may represent a novel molecular mechanism by which *D. mojavensis* males provision females. Similarly, nutrient reservoirs may serve as carrier molecules for the nutritional ejaculatory donation observed in *D. mojavensis* (Markow and Ankney, 1984, 1988).

D. mojavensis males are known to induce a clot-like insemination reaction in the female uterus after mating, the biochemical nature of which remains unknown (Patterson, 1946). Intriguingly, three of our candidate seminal fluid proteins showed blast homology to three clotting factors from the *D. melanogaster* larval hemolymph: CG1106, CG11064, and CG1469 (Karlsson et al., 2004). Similarly, two candidates exhibit conserved fibrinogen domains, a common clotting factor in mammalian blood (reviewed in Giangrande, 2003). All five clotting factors were present in two or more experiments, and four of five were amongst the top 200 (of 786) most abundant proteins (Supplementary Table 1), suggesting they do not represent contaminants from the hemolymph. Their potential role in the insemination reaction makes these proteins exciting candidates for functional and evolutionary studies.

Despite striking functional differences between *D. melanogaster* and *D. mojavensis* male ejaculates, the evolutionary dynamics of SFPs appear similar between these two species. Rapid evolution of SFPs, sometimes by positive selection, is well documented in both species (Begun et al., 2000; Swanson et al., 2001; Wagstaff and Begun, 2005; Almeida and Desalle, 2008a). Additionally, both ejaculates exhibit lineage-specific changes in SFP content by gene duplication (Wagstaff and Begun, 2007; Findlay et al., 2008; Almeida and Desalle, 2008a,b; Table 2), and the acquisition of novel genes with unknown evolutionary origins (Begun et al., 2006; Findlay et al., 2008; Supplementary Table 1). These data are consistent with the hypothesis that rapid evolution of male ejaculates is a universal byproduct of sexual reproduction.

Acknowledgements

The authors would like to thank Geoff Findlay and Willie Swanson for helpful discussion and advice on analysis on identification of unannotated peptides, two anonymous reviewers for comments on the manuscript, and Linda Breci for SEQUEST database searching. This research was funded by the state of Arizona. B.A.L was supported by the Undergraduate Biology Research Program at the University of Arizona. E.S.K. was supported by an NSF-IGERT fellowship in Functional, Evolutionary, and Computational Genomics at the University of Arizona, and a dissertation fellowship from the American Association of University Women.

Appendix. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.ibmb.2009.03.003.

References

Almeida, F.C., 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, F.C., Desalle, R., 2008b. Orthology, function, and evolution of accessory gland proteins in the *Drosophila repleta* species group. Genetics (online early).
Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410.
Andon, N.L., Eckert, D., Yates III, J.R., Haynes, P.A., 2003. High-throughput functional affinity purification of mannose binding proteins from *Oryza sativa*. Proteomics 3, 1270–1278.
Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., Harris, M.A., Hill, D.P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J.C., Richardson, J.E., Ringwald, M., Rubin, G.M., Sherlock, G., 2000. Gene ontology: tool for the unification of biology. The gene ontology Consortium. Nat. Genet. 25, 25–29.
Begun, D.J., Lindfors, H.A., 2005. Rapid evolution of genomic Acp complement in the melanogaster subgroup of *Drosophila*. Mol. Biol. Evol. 22, 2010–2021.
Begun, D.J., Whitley, P., Todd, B.L., Waldrip-Dail, H.M., Clark, A.G., 2000. Molecular population genetics of male accessory gland proteins in *Drosophila*. Genetics 156, 1879–1888.
Begun, D.J., Lindfors, H.A., Thompson, M.E., Holloway, A.K., 2006. Recently evolved genes identified from *Drosophila yakuba* and *D. erecta* accessory gland expressed sequence tags. Genetics 172, 1675–1681.
Bendtsen, J.D., Nielsen, H., von Heijne, G., Brunak, S., 2004. Improved prediction of signal peptides: SignalP 3.0. J. Mol. Biol. 340, 783–795. 57.
Braswell, W.E., Andrés, J.A., Maroja, L.S., Harrison, R.G., Howard, D.J., Swanson, W.J., 2006. Identification and comparative analysis of accessory gland proteins in 5. Genome 49, 1069–1080.
Breci, L., Hattrup, E., Keeler, M., Johnson, R., Letarte, J., Haynes, P.A., 2005. Comprehensive proteomics in yeast using chromatographic fractionation, gas phase fractionation, protein gel electrophoresis, and isoelectric focusing. Proteomics 8, 2018–2028.
Brieger, G., Butterworth, F.M., 1970. *Drosophila melanogaster*: identity of male lipid in reproductive system. Science 167, 1262.
Butterworth, F.M., 1969. Lipids of *Drosophila*: a newly detected lipid in the male. Science 163, 1356–1357.
Dorus, S., Busby, S.A., Gerike, U., Shabanowitz, J., Hunt, D.F., Karr, T.L., 2006. Genomic and functional evolution of the *Drosophila melanogaster* sperm proteome. Nat. Genet. 38, 1440–1445.
Drosophila 12 Genomes Consortium, 2007. Evolution of genes and genomes on the *Drosophila* phylogeny Drosophila 12 Genomes Consortium Drosophila 12 Genomes Consortiumy. Nature 450, 203–218.
Eddy, S.R., 1998. Profile hidden Markov models. Bioinformatics 14, 755–763.
Findlay, G.D., Yi, X., Maccoss, M.J., Swanson, W.J., 2008. Proteomics reveals novel *Drosophila* seminal fluid proteins transferred at mating. PLoS Biol. 6, e178.
Florens, L., Carozza, M.J., Swanson, S.K., Fournier, M., Coleman, M.K., Workmak, J.L., Washburn, M.P., 2006. Analyzing chromatin remodeling complexes using shotgun proteomics and normalized spectral abundance factors. Methods 40, 303–311.
Gharahdaghi, F., Weinberg, C.R., Meagher, D.A., Imai, B.S., Mische, S.M., 1999. Mass spectrometric identification of proteins from silver-stained polyacrylamide gel: a method for the removal of silver ions to enhance sensitivity. Electrophoresis 20, 601–605.
Giangrande, P.L., 2003. Six characters in search of an author: the history of the nomenclature of coagulation factors. Br. J. Haematol. 121, 703–712.
Hattrup, E., Neilson, K.A., Breci, L., Haynes, P.A., 2007. Proteomic analysis of shade-avoidance response in tomato leave. J. Agric. Food Chem. 55, 8310–8318.
Karlsson, C., Korayem, A.M., Scherfer, C., Loseva, O., Dushay, M.S., Theopold, U., 2004. Proteomic analysis of the *Drosophila* larval hemolymph clot. J. Biol. Chem. 279, 52033–52041.
Markow, T.A., Ankney, P.F., 1984. *Drosophila* Males Contribute to Oogenesis in a Multiple Mating Species. Science 224, 302–303.
Markow, T.A., Ankney, P.F., 1988. Insemination Reaction in *Drosophila* found in species whose males contribute material to oocytes before fertilization. Evolution 42, 1097–1101.
Markow, T.A., 1996. Evolution of *Drosophila* mating systems. Evol. Biol. 29, 73–106.
Markow, T.A., Coppola, A., Watts, T.D., 2001. How *Drosophila* males make eggs: it is elemental. Proc. Biol. Sci. 268, 1527–1532.
Markow, T.A., 2002. Perspective: female remating, operational sex ratio, and the arena of sexual selection in *Drosophila* species. Evolution 56, 1725–1734.

Mueller, J.L., Ripoll, D.R., Acquadro, C.F., Wolfner, M.F., 2004. Comparative structural modeling and inference of conserved protein classes in *Drosophila* seminal fluid. Proc. Nat. Acad. Sci. U.S.A 101, 13542–13547.

Mueller, J.L., Ravi Ram, K., McGraw, L.A., Bloch-Qazi, M.C., Siggia, E.D., Clark, A.G., Aquadro, C.F., Wolfner, M.F., 2005. Cross-species comparison of *Drosophila* male accessory gland protein genes. Genetics 171, 131–143.

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

Paoletti, A.C., Parmely, T.J., Tomomori-Sato, C., Sato, S., Zhu, D., Conaway, R.C., Conaway, J.W., Florens, L., Washburn, M.P., 2006. Quantitative proteomic analysis of distinct mammalian Mediator complexes using normalized spectral abundance factors. Proc. Natl. Acad. Sci. U.S.A 103, 18928–18933.

Patterson, J.T., 1946. A new type of isolating mechanism in Drosophila. Proc. Nat. Acad. Sci. U.S.A 32, 202–208.

Pitnick, S., Markow, T.A., Spicer, G., 1999. Evolution of multiple kinds of female sperm storage organs in *Drosophila*. Evolution 53, 1804–1822.

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

Sirot, L.K., Poulson, R.L., McKenna, M.C., Girnary, H., Wolfner, M.F., Harrington, L.C., 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.

Smedley, S.R., Eisner, T., 1995. Sodium uptake by puddling in a moth. Science 270, 1816–1818.

Sonnhammer, E.L., von Heijne, G., Krogh, A., 1998. A hidden Markov model for predicting transmembrane helices in protein sequences. Proc. Int. Conf. Intell. Syst. Mol. Biol. 6, 175–182.

Swanson, W.J., Clark, A.G., Waldrip-Dail, H.M., Wolfner, M.F., Aquadro, C.F., 2001. Evolutionary EST analysis identifies rapidly evolving male reproductive proteins in *Drosophila*. Proc. Natl. Acad. Sci. U. S. A. 98, 7375–7379.

Tabb, D.L., McDonald, W.H., Yates III, J.R., 2002. DTASelect and Contrast: tools for assembling and comparing protein identifications from shotgun proteomics. J. Proteome. Res. 1, 21–26.

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

Wagstaff, B.J., Begun, D.J., 2007. Adaptive evolution of recently duplicated accessory gland protein genes in desert *Drosophila*. Genetics 177, 1023–1030.

Walker, M.J., Rylett, C.M., Keen, J.N., Audsley, N., Sajid, M., Shirras, A.D., Isaac, R.E., 2006. Proteomic identification of *Drosophila melanogaster* male accessory gland proteins, including a pro-cathepsin and a soluble γ-glutamyl transpeptidase. Proteome Sci. 4 (9).

Wolfner, M.F., 2002. The gifts that keep on giving: physiological functions and evolutionary dynamics of male seminal proteins in *Drosophila*. Heredity 88, 85–93.

Wolfner, M.F., 2007. "S.P.E.R.M." (seminal proteins (are) essential reproductive modulators): the view from *Drosophila*. Soc. Reprod. Fertil. Suppl., 183–199.

Wilm, M., Shevchenko, A., Houthaeve, T., Breit, S., Schweigerer, L., Fotsis, T., Mann, M., 1996. Femtomole sequencing of proteins from polyacrylamide gels by nano-electrospray mass spectrometry. Nature 379, 466–469.

Zhang, Y., Sturgill, D., Parisi, M., Kumar, S., Oliver, B., 2007. Constraint and turnover in sex-biased gene expression in the genus *Drosophila*. Nature 450, 233–237.