

RESEARCH ARTICLE

Open Access



# Combining RNA-seq and proteomic profiling to identify seminal fluid proteins in the migratory grasshopper *Melanoplus sanguinipes* (F)

Martha L. Bonilla<sup>1,2</sup>, Christopher Todd<sup>2</sup>, Martin Erlandson<sup>3</sup> and Jose Andres<sup>2\*</sup>

## Abstract

**Background:** Seminal fluid proteins control many aspects of fertilization and in turn, they play a key role in post-mating sexual selection and possibly reproductive isolation. Because effective proteome profiling relies on the availability of high-quality DNA reference databases, our knowledge of these proteins is still largely limited to model organisms with ample genetic resources. New advances in sequencing technology allow for the rapid characterization of transcriptomes at low cost. By combining high throughput RNA-seq and shotgun proteomic profiling, we have characterized the seminal fluid proteins secreted by the primary male accessory gland of the migratory grasshopper (*Melanoplus sanguinipes*), one of the main agricultural pests in central North America.

**Results:** Using RNA sequencing, we characterized the transcripts of ~8,100 genes expressed in the long hyaline tubules (LHT) of the accessory glands. Proteomic profiling identified 353 proteins expressed in the long hyaline tubules (LHT). Of special interest are seminal fluid proteins (SFPs), such as *EJAC-SP*, *ACE* and prostaglandin synthetases, which are known to regulate female oviposition in insects.

**Conclusions:** Our study provides new insights into the proteomic components of male ejaculate in Orthopterans, and highlights several important patterns. First, the presence of proteins that lack predicted classical secretory tags in accessory gland proteomes is common in male accessory glands. Second, the products of a few highly expressed genes dominate the accessory gland secretions. Third, accessory gland transcriptomes are enriched for novel transcripts. Fourth, there is conservation of SFPs' functional classes across distantly related taxonomic groups with very different life histories, mating systems and sperm transferring mechanisms. The identified SFPs may serve as targets of future efforts to develop species-specific genetic control strategies.

**Keywords:** Transcriptome, Proteomics, Next generation sequencing, Agricultural pest, Male reproductive genes, *Orthoptera*

## Background

Insect seminal fluid proteins (SFPs), produced by the male accessory gland (AG), not only contribute to spermatophore formation (the capsules containing the ejaculate) but they also influence individual fitness by modifying different aspects of the females' reproductive physiology and behavior [1–10]. For this reason, SFPs

have been the focus of an increasing number of evolutionary studies on sexual conflict, post-mating sexual selection, and speciation [4, 5, 11–18]. A less explored aspect of the relationship between SFPs and fitness is the potential for the development of new pest control strategies aimed at disrupting the reproductive cycle of insect vectors and pests [19]. In insects, many SFP-encoding genes are highly divergent and show signs of positive selection [20–23]. Therefore, these loci are potential targets for the future development of species-

\* Correspondence: jose.andres@usask.ca

<sup>2</sup>Department of Biology, University of Saskatchewan, 112 Science Pl., Saskatoon, SK S7N-5E2, Canada

Full list of author information is available at the end of the article



specific, gene silencing biopesticides based on specific genetic control strategies.

Increasingly sophisticated molecular tools and the ability to generate massive amounts of genomics and proteomics data makes it possible to identify the proteins that are transferred to females during copulation. During the last decade, we have witnessed a rapid increase in the number of studies dissecting the structure and function of SFPs in a broad range of insects [6, 9, 19, 24–44]. However, these studies have only characterized the protein composition of the seminal fluid in a handful of taxonomic groups and the SFPs of most of the disease vectors and agricultural pest insects have yet to be identified.

Orthopterans are a large order of insects with more than 20,000 species, including a number of economically significant pests [9]. Only five species of crickets, however, (4 *Grylloidea* and 1 *Tettigoniodea*) have been studied thus far [14, 16, 17, 25, 42, 45–47]. Swarming and migratory grasshoppers (Family *Acrididae*) have been long recognized as crop pests, since the origins of cultivation ~10,000 years ago. In western North America, grasshopper feeding results in an estimated annual loss of \$1.25 billion per year [48]. Here, as a first step to identify the repertoire of SFPs in economically relevant species of grasshoppers, we aim to characterize the protein secretome of the long hyaline tubules in the migratory grasshopper (*Melanoplus sanguinipes*), one of the main pest species feeding on cereal and crucifer crops in central North America [47].

*M. sanguinipes* males are promiscuous and mate frequently, transferring up to fourteen small tubular spermatophores with each mating. During copulation, male spermatophores penetrate a short distance into the female's spermathecal duct where the seminal fluid is discharged before they are withdrawn and become lodged between the female's genital valvulae [49]. Both spermatophore building proteins and SFPs are produced in the complex group of multi-paired male accessory glands. These include one pair of long hyaline tubules, four pairs of white glands, ten pairs of short hyaline tubules and one pair of seminal vesicles [50]. Each of these type of glands secretes different proteins at specific times as the male develops [51]. Of special relevance are the long hyaline tubules (LHT) that produce large amounts of a partially identified protein (Oviposition Stimulating Protein, OSP) that is transferred from males to females during copulation and stimulates oviposition [50–53]. OSP is the major component that it is discharged with the spermatophores. Other LHT-secreted proteins seem to contribute to the formation and correct uncoiling of spermatophores [54]. Using a combination of next-generation RNA sequencing and proteomic analysis, we have identified and characterized the SFPs expressed in

this gland. These include not only OSP but also novel transcripts and protein classes previously described in insects with very different mating and seminal fluid transferring systems.

## Methods

### Sampling of insects and tissue dissection

Migratory grasshopper males are sexually mature a week after their last moult [55]. All experiments were conducted using 10-day-old males from the nondiapause colony of *M. sanguinipes*, maintained at Agriculture and Agri-Food Canada's Saskatoon Research Centre (Saskatoon, SK). Rearing methods are those described in [56]. All male accessory glands were dissected in Ringer's buffer and the LHT was carefully isolated (Fig. 1).

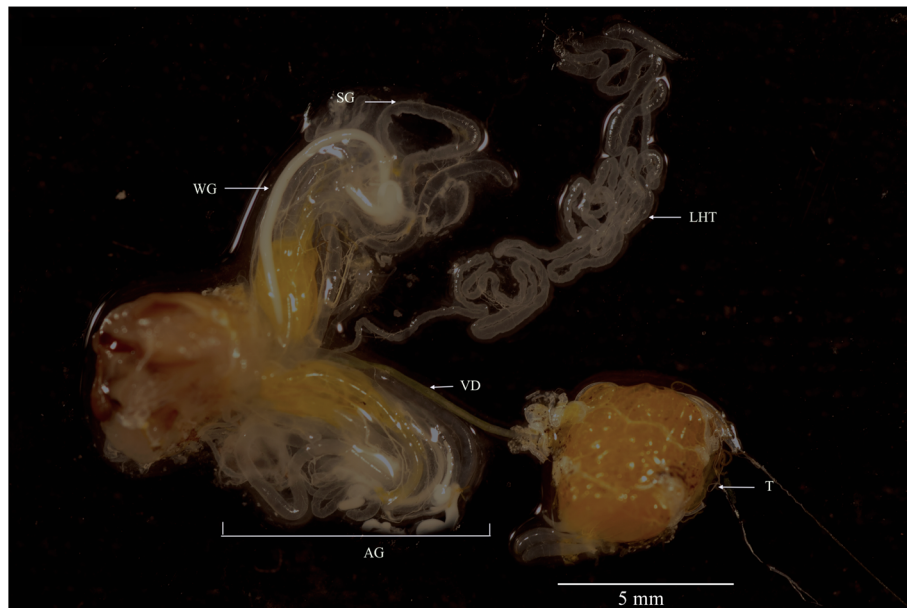
### cDNA library construction and assembly

Total RNA was extracted from the LHT of 9 mated and 9 virgins' adult males using TRIZOL (Invitrogen, Carlsbad, CA). Immediately after, the mRNA fraction was isolated using a Poly (A) Purist kit (Ambion), and checked its quality using an RNA 6000 Nano Kit and a Bioanalyzer 2100 (Agilent). Then the mRNA sample was fragmented and size selected (300 bp–1Kb) using the high temperature  $Zn^{2+}$  method and a cDNA Rapid Library Prep kit (Roche). Size fractionated samples were then sequenced at the National Research Council Plant Biotechnology Institute Canada (NRC-Saskatoon) using 454 GS-FLX platform and Titanium™ chemistry (Roche). Reads were *de novo* assembled using CLC Genomics Workbench software v 5.0 (similarity = 0.90, length fraction = 0.5, Mismatch cost = 2, indel cost = 3), into contigs with a minimum size cutoff of 200 bp.

### Transcriptome annotation and homology-based functional analyses

Assembled transcripts were submitted for homology, annotation searching and Gene Ontology (GO) categorization, using the online version of Blast2go program ([www.blast2go.com](http://www.blast2go.com)) [57]. Homology searches were done against non-redundant databases of the National Center for Biotechnology information (NCBI) using BLASTx algorithm. The annotations step was generated retrieving the keywords from the BLASTx descriptions. Gene Ontology categorizations of the functional annotations were carried out on the top BLASTx hits (1e-5 cutoff).

Search of putative open reading frames (ORFs) of at least 30 codons (90 bp) was done using the function Getorf from EMBOSS software [58]. The presence of conserved protein domains was tested using the online version of the program Pfam (<http://pfam.sanger.ac.uk/>) [59].



**Fig. 1** Male reproductive system of *M. sanguinipes*. LHT, Long hyaline tubules; T, testes; VD, vas deferens; WG, white glands SW, Short hyaline glands, AG, accessory gland complex

#### Identification of SFPs through mass-spectrometry

In insect species in which males transfer a single external spermatophore, SFPs can be identified using mass-spectrometry by comparing the spectra of peptides directly isolated from the spermatophore to that of a species-specific protein database (e.g. [40]). However this is not always possible. For example, in species such as the migratory grasshopper (in which males sequentially transfer up to fourteen spermatophores during a single copulation) it is impossible to collect all spermatophores without disrupting the mating pairs [49, 60]. In such cases one can possibly use isotopic labeling of female proteins to identify male transferred SFPs [19, 61–63]. However, this is not always possible. An alternative strategy in such cases is to identify the peptides secreted by the male accessory glands of interest. Although this approach has limitations and cannot distinguish between spermatophore-building proteins and those transferred to the females, it has been successfully applied to identify SFPs in other insect species [34, 40, 42, 64]. Thus, a single pool of ten LHT excised from individual males was prepared as described above. These glands were transferred to ice-cold insect Ringer buffer (153 mM NaCl, 2.68 mM KCl, 1.36 mM CaCl<sub>2</sub>), and gently centrifuged at 1,000 g for 1 minute to separate the secretion of the gland contents from the tissue. The resulting supernatant was stored at -80 °C. We considered these supernatant samples as “SFP-enriched” as they contained not only secreted proteins, but also likely LHT tissue. In-gel digestion, tryptic peptide extractions and Nano liquid chromatography with tandem mass spectrometry

(LC-MS/MS) were conducted by Genome BC Proteomics Centre at the University of Victoria, Victoria, British Columbia, Canada. Proteins from the viscous secretions were separated by electrophoresis on a one-dimensional 5–15 % gradient SDS polyacrylamide gel (Additional file 1: Figure S1) and visualized using Simply-Blue SafeStain (Invitrogen, Carlsbad, CA). The entire running lane of the gel was then sliced into 16 different equal size fractions. Each fraction was then transferred to a Genomics Solutions Progest perforated digestion tray, de-stained (50/45/5 (v/v) methanol/water/acetic acid) prior to reduction with 10 mM dithiothreitol (Sigma) and alkylation with 100 mM iodoacetamide (Sigma), and digested for 5 h at 37 °C using sequencing grade porcine trypsin solution (20 ng/μL, Promega, Madison, WI) at an enzyme: protein ratio of 1:50. Each sample was then lyophilized following acid extraction (50/40/10 acetonitrile/water/formic acid) and stored at -80 °C prior to mass spectrometry analysis.

The peptide mixtures were separated by on-line reversed phase chromatography using a Thermo Scientific EASY-nLC II system with a reversed-phase pre-column Magic C-18AQ (100 μm I.D., 3 cm length, 5 μm, 100 Å, Michrom BioResources Inc, Auburn, CA) and a reversed phase nano-analytical column Magic C-18AQ (75 μm I.D., 15 cm length, 5 μm, 100 Å, Michrom BioResources Inc, Auburn, CA). The chromatography system was coupled to an LTQ Orbitrap Velos mass spectrometer equipped with a Nanospray Flex source (Thermo Fisher Scientific). MS/MS spectra were analyzed with Proteome Discoverer 1.4.0.228 software suite (Thermo Scientific).

Peak lists generated of the Collision Induced Dissociation (CID) spectra were submitted to an in-house Mascot 2.4 server to identify proteins by searching against two protein databases (Uniprot tremble and Swissprot) and a six-reading frame translation of the LHT-cDNA library previously generated, with 260,330 potential open reading frames (ORFs) to confirm correct existing transcripts by peptide spectrum matches. The default search settings used for protein identification were: MS/MS accuracies were set to <math>< 0.6\text{ Da}</math>, and two missed cleavages for full trypsin with fixed modifications Carbamidomethyl (C); variable modifications: deamidation (N, Q); oxidation (M) and propionamide (C). A protein was positively identified if the ion score value of at least two different peptides exceeded the significance threshold ( $p < 0.05$ ). Proteins matching only one peptide ( $p < 0.05$ ) were only positively identified if the ion score value of the matching peptide was at least double the significance threshold [65]. The relative quantitation of identified proteins was then estimated using the exponentially modified protein abundance index (emPAI, Additional file 2: Table S1) [66]. For each identified protein a single emPAI value was obtained by adding the data from all gel slices.

To identify which proteins are part of the LHT-secretome first the program SignalP 4.0 (<http://www.cbs.dtu.dk/services/SignalP>) [67] was used to predict secreted proteins based on the presence and location of a signal peptide. Then, a neuronal network analysis (SecretomeP 2.0 (<http://www.cbs.dtu.dk/services/SecretomeP/>)) was implemented to further identify proteins putatively secreted by non-classical secretory pathways (*i.e.* Scoring function  $> 0.5$ ) [68]. All putatively secreted proteins were used as queries in local BLASTP searches against a combined database including annotated SFPs from *Drosophila melanogaster* [29], *Aedes aegypti* [65], *Heliconius* butterflies [32], *Tribolium castaneum* [24] and *Allonemobius* and *Gryllus* crickets [14, 40, 46, 69]. Following [35], Pairs of sequences that had reciprocal best Blast hits (RBBHs) with *e*-values  $< 1 \times 10^{-3}$ , identities  $\geq 30\%$  and bit score  $\geq 100$  were considered putative orthologous.

### Phylogenetic analyses

The secreted peptide known as OSP, is the most abundant protein secreted by the LHT and the only oviposition factor identified in *M. sanguinipes*. To gain further insight on its function, we carried out a multiple alignment of the published amino acid sequences of insect's takeout/juvenile hormone-binding proteins (JHBPs) including several SFPs that have been tentatively included in this superfamily [24, 48, 54, 67, 70–74]. Phylogenetic analyses were conducted using Neighbor-joining methods and the Jones–Taylor–Thornton (JTT [75])

substitution matrix, as implemented in PAUP\* b1.0 [76]. The confidence of the tree topology was assessed by a bootstrap ( $n = 1,000$  replicates).

### Expression patterns of putative SFPs

For the subset of putative *M. sanguinipes* SFPs' that have a canonical signal peptide and/or showed significant orthology with SFPs described in other insects we investigated the expression patterns analysis using a qualitative reverse transcription polymerase chain reaction (RT-PCR) [31, 32, 40]. Total RNA was extracted from adult virgin individuals (7–10 days old;  $n = 5$  individuals/sex) and five different tissues: male accessory gland (AG), testes (T), male head and leg (MHL) female spermatheca and oviduct (RTF), and female head and leg (FHL). Abdominal and thorax tissues were not included in this analysis to avoid potential contamination with reproductive tissues. For each tissue, cDNA was synthesized from 1  $\mu\text{g}$  of total RNA of the pooled individual samples using a Quantitect Reverse Transcription kit (Qiagen). Consequently, it was not possible to capture individual variation in the expression levels, and further studies are needed to address this point. All RT-PCRs were performed using the touchdown PCR protocol (see Additional file 3: Table S2 for PCR and primers details).

## Results

### cDNA library annotation and characterization

Sequencing using 454 GS FLX titanium technology generated 259,565 high quality reads (92,250,778 bases) that assembled into 82 singletons and 8,056 contigs. Read length of transcripts ranged from 62 to 6,208 bp with average contig length of 649 bp. We did not find a significant correlation between transcripts length and number of reads ( $r = 0.08$ ,  $p > 0.05$ ;  $n = 8,138$ ).

Analysis of the expressed sequence tags (ESTs) frequency spectrum revealed that most of the transcripts ( $\sim 66\%$ ) occurred as either singletons or contigs that included only a small number of reads ( $n \leq 5$ ). Only a small group ( $n = 36$ ) was represented by a high ( $n > 500$ ) number of reads. Nucleotide-based Blast analyses (BLASTx) revealed that  $\sim 60\%$  of the contigs show significant similarities with either annotated gene products and/or known protein domains (*E*-value  $\leq 10^{-5}$ ). Only a small fraction (4.4 %) showed significant homology to the same annotated transcript.

Gene Ontology (GO) assignments were used to classify the functions of the predicted genes based on contigs with significant BLASTx (*E*-value  $\leq 10^{-5}$ ). Contigs were assigned to 23 biological processes, 9 cell components and 14 molecular functions based on GO level II (Additional file 4: Figure S2A–C). Some contigs were associated with multiple GO annotations because a single

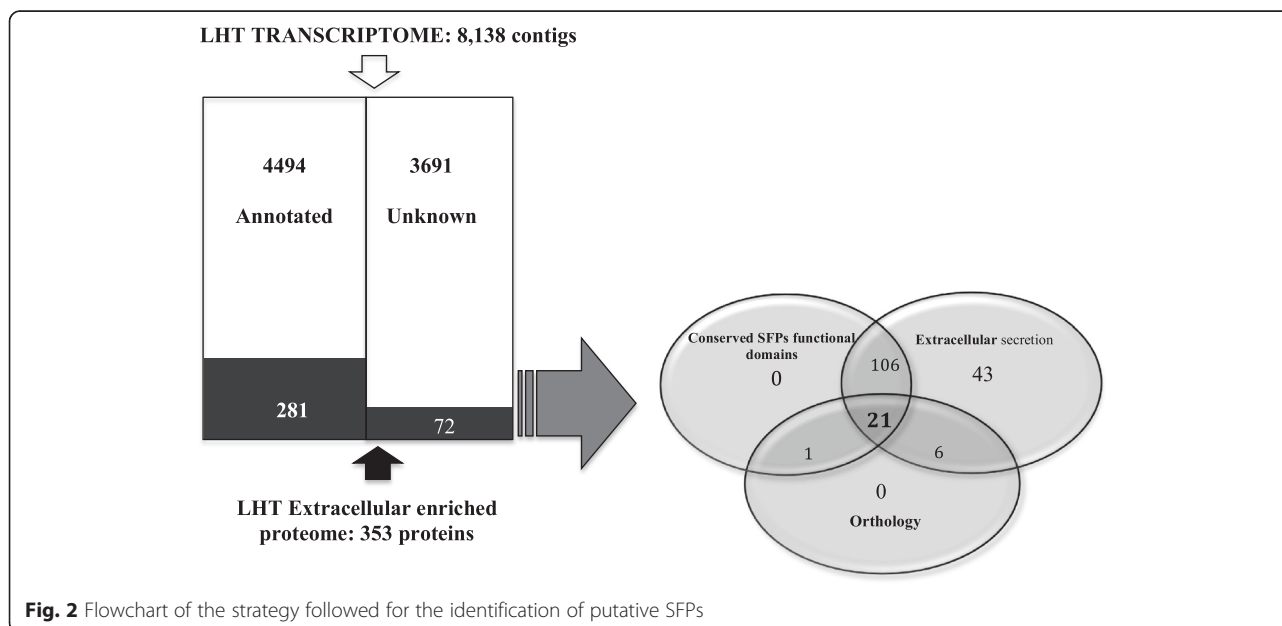


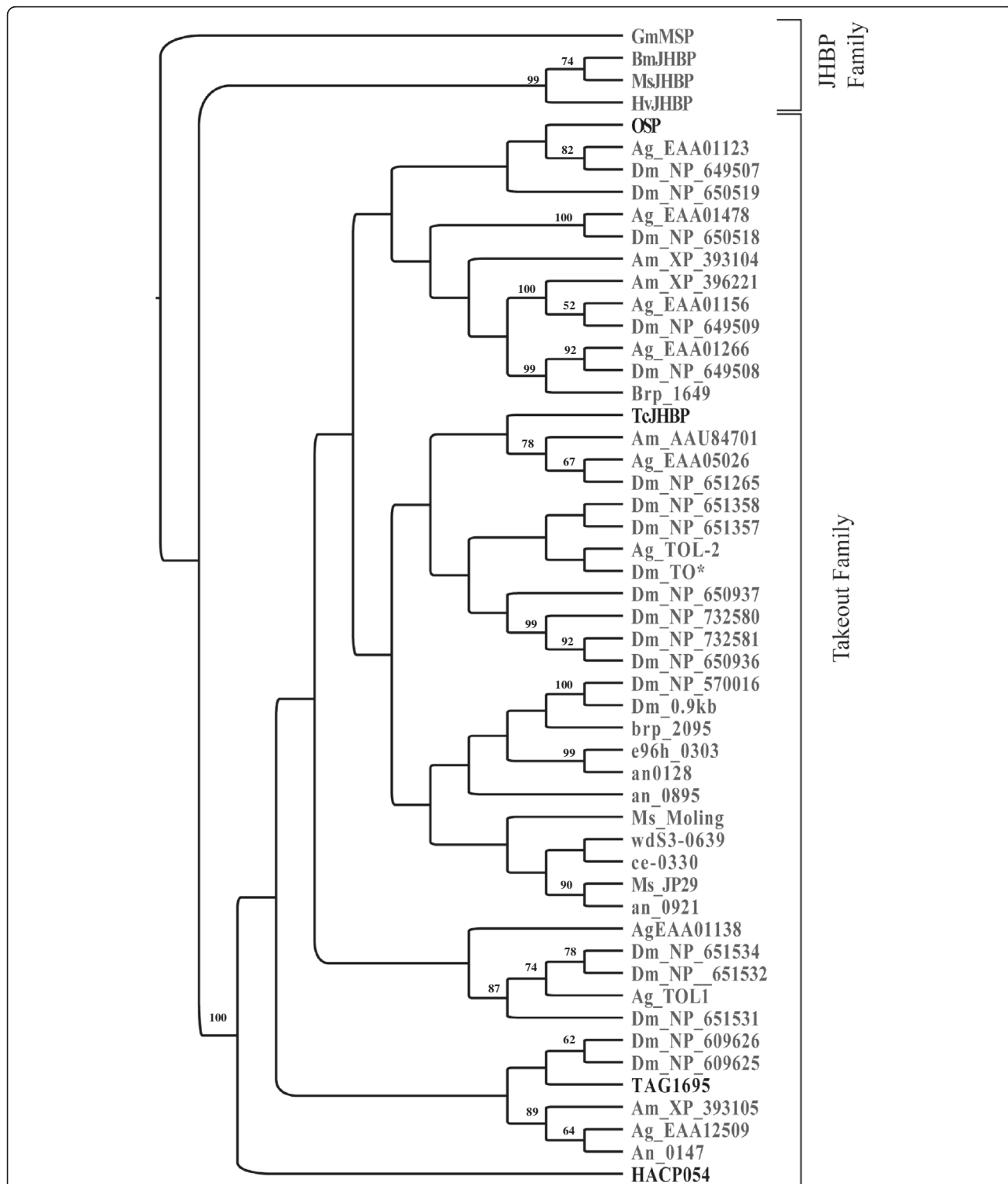
sequence may be annotated in any or all categories within a single category, giving more “GO” annotations than sequences annotated [33]. Within the biological Processes, 42 % of the contigs were assigned to metabolic and basic cellular process. Remaining contigs were involved in a broad range of biological process such as: biological regulation (10 %), developmental process (7 %), signal transduction (5 %), localization (7 %), reproductive process (2 %), cellular adhesion (1 %), response to stress (7 %) and immune response (1 %).

**Identification of putative SFPs**

Identification of candidate SFPs was based on combination of both, transcriptome and proteomic analyses (Fig. 2). Using annotation analyses we initially identified 4,497 LHT-expressed genes (Additional file 5: Table S3). We further identified LHT seminal proteins using a combination of mass spectrometry and bioinformatics as follows: First, we compared the peptide sequences from tryptic digests of our long hyaline tubule SFP-enriched fractions with peptide sequences generated *in silico* from a translation of the LHT transcriptome. This search resulted in the identification of 353 gene products (Additional file 2: Table S1), of which 28 % ( $n = 99$ ) were only matched by a single-peptide. Most of these products correspond to a broad range of diverse protein families, including metabolic and structural proteins, as expected if the protein sample contained traces of LHT tissue. However, as expected if most of these proteins indeed represent SFP both the average number of reads and exponentially modified protein abundance values (emPAI) were higher in this group than in the rest of the transcriptome ( $N_{PUTATIVE\ SFPs} = 744.9 \pm 429.3$ ,  $N_{TRANSCRIPTOME} = 22.2 \pm 5.44$ , permutation test  $p < 0.001$ ;  $emPAI_{PUTATIVE\ SFPs} = 2207.7 \pm 2192.8$ ,  $emPAI_{TRANSCRIPTOME} = 3.6 \pm 0.65$ , permutation test  $p < 0.001$ ). Also, for this group of proteins we found a positive correlation between gene and protein expression levels (Spearman-Rho = 0.29,  $p = 0.01$ ). In an attempt to identify *bona fide* SFPs we first selected those transcripts potentially encoding extracellular, secreted, proteins (*i.e.* contigs with a predicted signal peptide or secreted via the non-classical pathway). As expected, such analyses revealed a significant number of gene products (176/353) that are potentially transferred to the female during mating. Homology-based functional analyses revealed that the most abundant LHT-protein (EmPAI value 151323.39) is approximately 460x more abundant than the second-most abundant protein. This transcript contains a single putative juvenile hormone (JH) binding protein domain and the N-terminal of this product corresponds (100 % similarity) with the Oviposition Stimulating Protein (OSP) [53] (Additional file 6: Figure S3). BLASTp alignments showed significant homology between OSP and described proteins of the take out/Juvenile hormone binding proteins (TO/JHBP). Phylogenetic analysis showed OSP clustering with a large group of highly diversified takeout (TO) proteins including other known insect SFPs containing TO/JHBP domains. However, these SFPs do not form a monophyletic group (Fig. 3). Approximately 26 % ( $n = 46/176$ ) of the secreted transcripts did not show any significant similarities with annotated domains, suggesting that the LHT secretome is enriched in novel SFPs.

Because previous studies in spermatophore producing insects have shown that there is some functional





**Fig. 3** Phylogenetic tree based on published amino acid sequences of takeout/JHPB family proteins. Phylogenetic analyses were conducted using Neighbor-joining methods and the Jones–Taylor–Thornton substitution matrix (JTT, Jones et al. 1992). The numbers on each branch represent bootstrap values ( $n = 1,000$ ). Known SFPs with predicted JHPB binding domain are shown in bold. Species codes and number of individual sequences included in the analysis *Drosophila melanogaster*, Dm [112, 113]; *Anopheles gambiae*, Ag [72]; *Manduca sexta*, Ms [114, 115]; *Apis mellifera*, Am [80]; *Bombyx mori*, BmJHBP, ce- 0303, brp\_2095, e96h\_0303, an\_128, an\_0895, wdS3-0639, an\_0921, an\_0147 [75]; *Heliothis virescens*, Hv [116]; *Galleria mellonella*, Gm [117]; *Tribolium castanenum*, Tc [24]; *Ceratitis capitata*, TAG1695 [73]; *Heliconius melpomene*, HAC054 [32]

conservation of SFPs across distantly related taxa [14, 32, 40, 42], we cross-referenced the list of 176 predicted secreted proteins (Additional file 7: Table S4) with a database containing putative SFPs previously identified in other insect species (see Methods). Using reciprocal BLAST we were able to find putative orthologous for (36) proteins described in this study, but only 28 of them had a bit score  $>100$ . As expected, most orthologous ( $\sim 43\%$ ,  $n = 12$ ) were found in other orthopterans and only a few were found in more distantly related taxa ( $n_{Heliconious} = 5$ ,  $n_{Tribolium} = 4$ ,  $n_{Aedes} = 7$ ; Table 1).

Although previous studies suggest that not all SFP encoding genes are exclusively expressed in male reproductive tissues [28, 29] male-biased expression is likely to be expected. Thus, we looked at the expression patterns in the group of putative SFPs that had a canonical signal peptide and/or showed significant orthology with previously reported reproductive proteins (see Fig. 4, Table 1). Our results showed that approximately 30% (19/64) showed male expression bias (i.e. expressed in males but not in females), of which 12 corresponded to proteolysis regulators, two are lipases, three have unknown functions, and one encodes OSP.

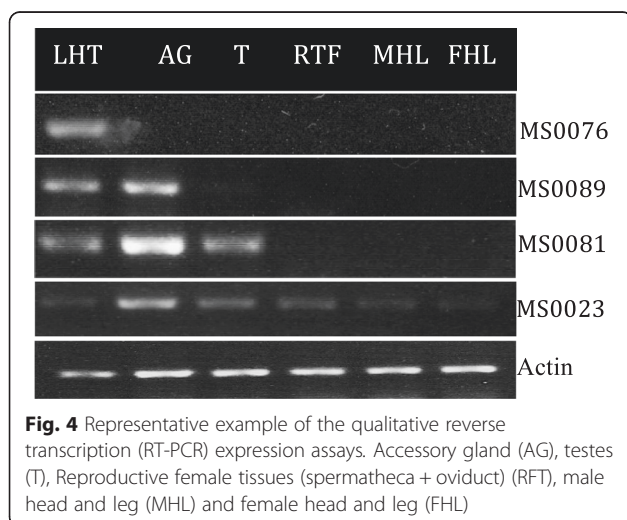
## Discussion

Male's seminal fluid contains peptides that modify almost all aspects of female reproductive physiology and behavior. Here, we have capitalized on genomic and proteomic techniques to characterize the SFPs secreted by the LHT gland of *M. sanguinipes* [14, 19, 24, 28, 31, 35, 65]. Our analyses revealed that the secretome of the LHT is complex.

A potential approach to identify *bona fide* SFPs is to consider only those proteins that either showed orthology with SFPs described in other insects and/or that have canonical signal peptide. Using exclusively these

overstrict criteria, we have been able to identify as many as 64 different putative SFPs (Table 1). However, this number does not include rapidly evolving SFPs and genes with unknown functions with non-canonical secretion signals, which may represent a significant fraction of the ejaculate [33, 40, 77]. Assuming that our non-canonical secretion analyses generated only a few ( $\sim 5\%$ ) false positives, and that all secreted proteins are part of the seminal fluid, the number of identified putative SFPs raises to 176 (Fig. 2). This latter number is likely to be a better estimate of total number of SFPs secreted by the LHT, which seems to be higher than that described in other insects using similar methods (mean  $N_{\text{Secreted-SFPs}}$ : 38; range: 13-138 (reviewed in [7, 24, 77])). Although this difference might just reflect the variation in experimental design (or in the interpretation of the results) among published studies it is also possible that the LHT secretome is particularly heterogeneous. Interestingly, a large number of putative SFPs have also been described in *Gryllus* and *Allonemobius* crickets, suggesting that Orthopterans ejaculates may be especially complex [14, 40, 42, 46, 77].

The characterization of the LHT secretome revealed several important patterns. First, the presence of proteins that lack canonical signal peptides in accessory gland proteomes is indeed common [19, 35]. In the LHT of *M. sanguinipes* only  $\sim 28\%$  (49/176) of the proteins predicted to be secreted contain a signal peptide. This result highlights the importance of non-standard routes in the secretion of SFPs. Second, the products of a few highly expressed genes dominate the secretion of accessory glands. Previous studies in crickets have suggested that the highly expressed genes of the accessory glands encode structural (i.e. spermatophore building) proteins [40, 42]. However, this is not necessarily the case in the LHT of the migratory grasshopper. The most abundant product of the LHT was a protein identified as OSP, known to stimulate oviposition in migratory grasshopper females [53]. Third, accessory gland transcriptomes are enriched in novel transcripts. 26% (46/176) predicted secreted SFPs identified in the proteomics experiment lack annotated domains, suggesting that this gland is enriched with highly abundant novel seminal fluid proteins coding genes whose function in reproductive behavior has yet to be evaluated. Many of the other male-expressed and LHT secreted proteins described in this study belong to the same functional classes as previously described SFPs including proteins involved in processing and degradation of proteins [26, 27, 36, 78], odorant/hormone binding-like proteins [31, 60, 79], immune and stress responses [19, 34, 35], and metabolic pathways and/or structural proteins with known functions related to reproduction. In what



**Table 1** Putative SFPs identified in the LHT of *Melanoplus sanguinipes*

Contig	Expression patterns	Functional summary	Reads	EmPAI	Genome matched (Reference)	BLASTP description	Orthology			Signal secretion
							E-value	Hit Rank	Bit score	
<i>Proteolysis regulators</i>										
MS0199	a	Dipeptidase	58	6.07	<i>A. fasciatus</i> (2)	EG018903.1 Af_ARG_10E12_sp6 Accessory gland	6.00E-63	2	208	SP
MS1698	a	Serine protease inhibitor	60	1.06	<i>H. melpomne</i> (2)	HM023856.2  SFP HACPO59	4.00E-45	1	139	SP
MS3398	a	Serpin b6	21	2.15	<i>H. melpomene</i> (2)	HM023856.2 SFP HACPO59	9.00E-23	4	134	SP
MS1783	a	Angiotensin-converting enzyme	93	0.96	<i>L. migratoria</i> (3)	AY487174.1  Angiotensin converting enzyme	0	1	803	SP
MS0381	a	Metallo-endopeptidase (nepriyisin)	36	0.98	<i>G. firmus</i> (1)	EB409838.1 Gf_AcGl1_07A05_SP6 AcGl1	2.00E-34	1	127	SP
MS0081	a	Serine protease	481	6.21	<i>A. fasciatus</i> (1)	GQ911573.1 ejaculate serine protease (ejac-sp)	5.00E-80	1	168	NC
MS0089	a	Serine protease	481	28.8	<i>H. melpomene</i> (2)	HM023850.1 SFP HACPO49	7.00E-61	1	119	NC
MS3857	a	Serine carboxypeptidase	6	1.5	<i>Ae. Aegypti</i> (6)	XM_001660689.1 Zinc carboxypeptidase	6.00E-48	1	116	NC
MS2174	b	Serine protease	29	2.17	<i>H. melpomene</i> (2)	HM023834.1 SFP HACPO27	3.00E-36	1	148	NC
MS1693	a	Serine protease	175	0.29	<i>H. melpomene</i> (2)	HM023850.1 SFP	2.00E-21	1	40.5	SP
MS3819	b	Transmembrane protease serine	6	0.6	<i>G. firmus</i> (1)	DQ520137.1 Accessory gland protein (AG-0159F)	1.00E-10	4	46.4	SP
MS2043	a	Serine protease	8	0.75	<i>G. firmus</i> (1)	EB409712.1Gf_AcGl1_08D06_SP6	46.0E-10	4	45.5	SP
MS5974	b	Serine protease inhibitor	4	0.2	-	-	-	-	-	SP
MS0294	a	Serine carboxypeptidase	146	9.86	-	-	-	-	-	SP
MS6504	a	Serine carboxypeptidase	8	1.26	-	-	-	-	-	SP
<i>Environmental stress proteins</i>										
MS0126	c	Heat shock protein	84	0.79	<i>T. casteanum</i> (5)	XM_965476.2 70kD protein cognate (LOC659147)	1.00E-155	1	538	SP
MS0131	b	Heat shock protein	17	1.59	<i>T. casteanum</i> (5)	XM_961518.2. 70kD protein cognate 70, LOC657158	0	1	1216	-
<i>Iron storage proteins</i>										
MS1628	b	Transferrin	33	4.26	<i>G. firmus</i> (1)	EB409584.1 Gf_AcGl1_03G10_SP6 AcGl1	2.00E-69	1	256	SP
MS0380	b	Ferritin subunit	51	0.14	-	-	-	-	-	SP
<i>Oxidative stress proteins</i>										
MS1782	b	Catalase	4	1	<i>Ae. Aegypti</i> (6)	XM_001663550.1 Aedes aegypti catalase partial	4.00E-115	1	290	SP
MS0023	b	Aldo-keto reductase	323	2.56	<i>H. erato</i> (2)		2.00E-57	1	102	NC



**Table 1** Putative SFPs identified in the LHT of *Melanoplus sanguinipes* (Continued)

						HM023814.1 Seminal fluid protein HACP058				
MS2301	c	Phenoloxidase	26	1.69	<i>T. casteanum</i> (5)	NM_001039433.1  Tribolium castaneum pro-phenol oxidase subun...	0	1	286	NC
MS0678	c	Glutathione s-transferase	12	2.88	<i>Ae. Aegypti</i> (6)	XM_001661819.1 glutathione s-transferase	4.00E-55	1	203	-
MS0031	b	Endoplasmic	142	0.17	-	-	-	-	-	SP
MS4425	b	Aldehyde oxidase	7	0.3	-	-	-	-	-	SP
<i>Lipases</i>										
MS5352	c	Lipase	10	4.85	<i>Ae. Aegypti</i> (6)	XM_001651399.1  vitellogenin partial	1.00E-24	1	109	SP
MS1414	a	Lipase	19	1.53	-	-	-	-	-	SP
MS0358	a	Lipase	119	21.37	-	-	-	-	-	SP
<i>Cell adhesion protein, membrane-bound and structural proteins</i>										
MS2129	b	Moesin ezrin radixin	42	3.91	<i>Ae. Aegypti</i> (6)	XM_001652925. Moesin/ezrin/radixin	0	1	903	
MS0111	b	Actin	446	2.56	<i>G. firmus</i> (1)	EB409566 Gf_AcG1_03F03_SP6 5	1.00E-92	1	333	NC
MS0014	b	Annexin isoform a	85	0.11	<i>T. casteanum</i> (5)	XM_962838.2 PREDICTED: similar to Annexin IX CG5730-PC	9.00E-147	1	515	-
MS0238	b	Elongation factor 1-alpha	238	1.67	<i>A. fasciatus</i> (2)	EG018577.1 Af_ARG_09B03_sp6 Accessory gland	3.00E-127	1	449	-
MS0234	b	Elongation factor 2	864	0.46	<i>G. firmus</i> (1)	EB409600.1 Gf_AcG1_01A02_SP6 AcG11	2.00E-128	1	453	-
<i>Others</i>										
MS0219	b	Calreticulin	343	0.37	-	-	-	-	-	SP
MS0105	b	Midline fasciclin	287	4.08	<i>A. fasciatus</i> (2)	EG018790.1 Af_ARG_02F01_sp6 Accessory gland	2.00 E-75	1	169	SP
MS0076	a	Hypothetical EAI_15996	6173	73.01	<i>G. firmus</i> (1)	EB409920.1 Gf_AcG1_06A06_SP6	2.00E-47	1	184	SP
MS0157	b	Nucleoside diphosphate kinase	154	1.43	<i>Ae. Aegypti</i> (6)	XM_001662462.1 Aedes aegypti nucleoside-diphosphate kinase	3.00E-87	1	310	SP
MS1368	b	Protein 5nuc-like	99	0.66	<i>G. firmus</i> (1)	EB409815.1 Gf_AcG1_09F10_SP6	3.00E-30	1	110	SP
MS0047	b	Peptidyl-prolyl cis-trans isomerase 5	106	4.97	<i>G. firmus</i> (1)	EU669817.1 Peptidyl-prolyl isomerase-1	3.00E-45	2	174	SP
MS1669	c	Aminopeptidase -like	12	0.18	<i>Ae. Aegypti</i> (6)	XM_001656177.1 Aedes aegypti leucyl amino peptidase partial	2.00E-79	1	142	-
MS1026	b	Peptidoglycan recognition protein	23	2.22	-	-	-	-	-	SP
MS2092	b	Hypoxia up-regulated protein	12	0.73	-	-	-	-	-	SP
MS7155	a	Juvenile hormone-inducible	7	0.12	-	-	-	-	-	SP
MS0013	a	Oviposition Stimulating Protein (OSP)	28685	151323.3	-	-	-	-	-	SP
MS0682	b	Proactivator polypeptide	209	0.08	-	-	-	-	-	SP
MS0237	c	Carbonic anhydrase 2-like	18	1.56	-	-	-	-	-	SP
MS0221	c	Neutral alpha-glucosidase ab-like	24	0.45	-	-	-	-	-	SP

**Table 1** Putative SFPs identified in the LHT of *Melanoplus sanguinipes* (Continued)

MS0644	b	Profilin	149	0.54	-	-	-	-	-	SP
MS2078	c	Plasma alpha-l-fucosidase	21	1.3	-	-	-	-	-	SP
MS0168	c	Esterase fe4-like	32	1.81	-	-	-	-	-	SP
MS0673		AaeL_AAEL001498	24	0.55						SP
MS3488	b	Hemomucin	10	1.07	-	-	-	-	-	SP
MS0097	b	Pdgf- and vegf-related factor 1-like precursor	6	0.26	-	-	-	-	-	SP
MS1006	b	Imaginal disc growth factor 4	75	4.26	-	-	-	-	-	SP
MS1117	c	Calumenin	19	0.23	-	-	-	-	-	SP
<i>Unknown</i>										
MS0327	c	Unknown	164	1.54	-	-	-	-	-	SP
MS1333	a	Unknown	55	3.39	-	-	-	-	-	SP
MS0039	a	Unknown	681	1.43	-	-	-	-	-	SP
MS2957	b	Unknown	13	0.18	-	-	-	-	-	SP
MS0784	b	Unknown	7	3.09	-	-	-	-	-	SP
MS0278	c	Unknown	375	0.51	-	-	-	-	-	SP
MS0281	c	Unknown	167	0.44	-	-	-	-	-	SP
MS5071	b	Unknown	2	1.92	-	-	-	-	-	SP
MS1809	c	Unknown	33	1.71	-	-	-	-	-	SP

References ((1) Andres et al. [14, 40]; (2) Walters and Harrison. [32] (3) Macours et al. [107]; (4) Findlay et al. [29]; (5) Xu et al. [24]; (6) Sirot et al. [58]). Expression patterns using qualitative RT-PCR ((a) Male expression exclusively; (b) Expression in both sexes and (c) Not amplified. Signal secretion [Signal peptide (SP), Non-canonical secretion (NP)

follows we discuss the potential roles of the LHT-SFPs that were identified using both the proteomics and genomics experiments.

#### Take-out (TO) and Juvenile hormone binding proteins (JHBPs)

*TO/JHBPs* are small proteins (~240 amino acids) found exclusively in insects, and are involved in the transportation of hydrophobic ligands [80]. Putative SFPs containing TO/JBHP domains have been previously described in the male accessory glands of *Heliconious* butterflies, the Mediterranean fruit fly (*Ceratitis capitata*) and the flour beetle, *Tribolium castaneum* [19, 24, 32], and may be involved in the transfer of small hydrophobic molecules during copulation.

Homology based analyses revealed that OSP [53], the most abundant protein in the LHT, belongs to this group. Juvenile hormone (JH) is known to stimulate oogenesis [81]. Thus, there is the possibility that OSP acts as a carrier for JH. However, our distance-based phylogenetic tree clustered OSP with the highly diversified family of *TO* proteins instead of with known JH-binding proteins (Fig. 3). This result suggests that OSP is not likely to be involved in JH transportation. In fact, extensive biochemical and molecular functional

characterization of this protein have revealed that it has little affinity for JH [82].

#### SFPs involved in the synthesis of prostaglandins

Prostaglandins (PGs), prostaglandin-precursors and prostaglandin-synthesizing enzymes have been found in the testes, and the accessory glands of different insect's groups [83–87]. Our bioinformatics analyses revealed the presence of prostaglandin F synthase ( $\text{PGF}_{2\alpha}$  MS4100) and prostaglandin E synthase 2 ( $\text{PGE}_2$ , MS5577) in the LHT. In at least two crickets species (*Acheta domesticus* and *Teleogryllus commodus*) short-term oviposition is stimulated by the post-copulatory synthesis of prostaglandins in the female's reproductive tract [84, 85, 88]. Thus, LHT-PG synthetases may play a similar role in *M. sanguinipes*. However, this function is not necessarily conserved across orthopterans and prostaglandins do not appear to affect egg-laying in the migratory locust, *Locusta migratoria* [86].

The synthesis of prostaglandins also involves the oxidation of lipids and the release of reactive oxygen species (ROS), which might cause DNA damage, membrane degradation and premature activation of the sperm. Therefore, antioxidant SFPs may be particularly important in those species in which prostaglandins are stored and/o synthesized in the sperm storage organs of the

females [89–91]. Accordingly, in our study we have identified 16 putative SFPs with antioxidant properties, 5 of which are orthologous of SFPs found in other insect species [Glutathione S-transferase (GST) contigs (MS4391, MS0677 and MS678), catalase (CAT) contig (MS1782), phenoloxidase, contig (MS2301), aldo keto reductase contig (MS0023)] [19, 92].

Three different transcripts (MS358, MS1414 and MS5352) that have predicted signal peptides and which expression is restricted to the LHT showed significant homology with different lipases found in the seminal fluid of other insects [41, 66, 84, 85]. Although an obvious function of these enzymes is to provide energy to sperm by the hydrolysis of triglycerides [34], lipases are also known regulators of the complexes that catalyze the conversion of arachidonic acid to prostaglandins inside the female's reproductive tract [93]. Thus, it is possible that these LHT-SFPs may help to regulate the synthesis of prostaglandins in the spermatheca of the mated females.

#### Proteolysis regulators

Proteolysis regulators are common key modulators of insects' reproductive physiology [19, 24, 27, 35, 38, 46]. Our results showed that this functional class represents a significant fraction of the LHT secretome. In total, 23 putative SFPs correspond to proteases or their inhibitors. These include 12 serine proteases, 5 serine protease inhibitors (SERPINs), 4 carboxypeptidases, 2 neprilysins metalloendopeptidases, and 1 dipeptidase. Proteolysis regulators secreted by the LHT are highly conserved, and for all but one we have found orthologous proteins in distantly related species of insects. Among them, we have identified a gene (contig MS0081) that encodes a highly expressed trypsin-like serine protease, which is the putative orthologous protein of the ejaculate serine protease *ejac-sp* gene. Along with *OSP*, *EJAC-SP* is one of the most abundant proteins in the seminal fluid of the ground cricket *Allonemobious socius*, and is one of the few peptides known to mediate oviposition in orthopterans [46].

Zinc-dependent metallopeptidases, including neprilysins and angiotensin-like converting enzymes, are important proteolysis regulators. The predicted MS381 product shows functional homology with different neprilysins, a family of proteins known to play important roles in both spermatogenesis and fertilization [94]. In insects, different isoforms of a highly conserved angiotensin-converting enzyme (ACE) are expressed in male accessory glands [24, 38, 95, 96]. The contig MS1783 shows significant orthology with *D. melanogaster* ANCE and *T. castaneum* Lom-ACE. While ANCE is involved in spermatid differentiation [95], the knockdown of *Lom-ACE* results in significant reduction in egg production by mated females and

production of abnormal sperm [24]. It is possible that MS1783 plays similar functional roles in *M. sanguinipes*.

#### Iron storage proteins

Two contigs (MS2388 and MS1628) correspond to a single predicted transferrin. Transferrins are iron binding proteins that have been reported in the male reproductive tract of blood-feeding insects [65, 80], but also in the accessory glands of the field cricket *T. oceanicus* [25]. Insect transferrins are multifunctional proteins [97, 98]. While in blood-feeding insects seminal transferrins may be related to blood utilization, in other insects groups such as orthopterans they may contribute to vitellogenesis, to immune functions, and /or to prevent oxidative stress [97]. In vertebrates, quantitative variation in transferrin in the seminal plasma correlates with sperm numbers, sperm motility, and male fertility [99–101], suggesting that transferrins may also play a role in sperm capacitation.

Secreted ferritins are known to be present in the hemolymph, the gut, and the ovaries of insects [98, 102]. These proteins have antioxidant activity [103] and are involved in innate immune responses [104, 105] and iron homeostasis [106], including iron store in eggs [102]. However, to the best of our knowledge they have not yet been reported in the seminal fluid of insects. Our proteomics analyses showed presence of a secreted ferritin homolog (contig MS0381) in the LHT. This result strongly suggests that ferritins are present in the seminal fluid of the migratory grasshopper. Though the function of ferritin-like proteins in the seminal fluid is not known, knockdown experiments have shown that the expression of these proteins affects oviposition and egg hatching rates in at least one species of tick [107].

#### Metabolism-related proteins

In mammals, there is ample evidence that different (extra) cellular substrates and metabolic pathways are required to support the energetic requirements of sperm activation and fertilization. For example, while the acrosome reaction requires lactate or pyruvate for ATP production by oxidative phosphorylation, gamete fusion requires glucose to produce NADPH by the pentose phosphate pathway (reviewed in [108, 109]). In the LHT we have identified several secreted glycolytic enzymes, 5 of which [dehydrogenase, isocitrate dehydrogenase,  $\alpha$ -enolase, glyceraldehyde-3-phosphate dehydrogenase, and nucleoside diphosphate kinase] have been found in the seminal fluid of the field cricket *Teleogryllus oceanicus* [77]. The roles and mode of action of most of these enzymatic SFPs in insects are not yet known. However, glucose dehydrogenases are required for sperm storage and utilization in *Drosophila* [110], and Neutral- $\alpha$ -glucosidase (contig MS0221) is required during sperm maturation in humans [111].

## Conclusions

In summary, this report is the first attempt at the identification of SFPs in the migratory grasshopper, *M. sanguinipes*. Using a combination of transcriptome and proteomic analyses we were able to identify 64 putative SFPs. Of special interest are relatively conserved genes, such as *EJAC-SP*, *ACE* and prostaglandin synthetases, products that are known to regulate female oviposition rate. Gene silencing has considerable promise for developing novel pest control techniques. However, functional characterization experiments in acridid grasshoppers and locusts are needed to assess if these SFP loci are useful targets for the implementation of this type of strategy.

## Availability of supporting data

DNA sequences of the seminal fluid proteins have been deposited in GenBank. (Accession numbers: KU218647-KU218708). Phylogenetic trees, and all other data are available through Dryad Digital Repository doi 10.5061/dryad.t80d3 (<http://dx.doi.org/10.5061/dryad.t80d3>).

## Additional files

**Additional file 1: Figure S1.** Protein analysis of Long Hyaline Tubule (LHT) luminal secretions isolated from *M. sanguinipes* males on day 10 post-eclosion on 12 % SDS-PAGE. The entire gel lane was cut in 16 equal slices of approximately the same size (see Methods). (PDF 196 kb)

**Additional file 2: Table S1.** *Melanoplus sanguinipes* Seminal fluid proteins identified by mass-spectrometry. Predicted function based on the annotation and BLAST2GO. (XLSX 84 kb)

**Additional file 3: Table S2.** Primers used in RT-PCR and PCR conditions. PCR was performed with the following thermal conditions 60 s at 95 °C, 50 s at 94 °C, 30 s at 62 °C, 3 cycles /°C declining over a range at 1 °C intervals of 30 s, 72 s at 60 °C, 50 s at 95 °C, 30 s at 60 °C, 6 cycles /°C declining over a range at 0.5 °C intervals of 30 s, 50 s at 95 °C, 30 s at 56 °C, 30 cycles of 30 s and 60 s at 72 °C. Relative expression was not quantified. (XLSX 49 kb)

**Additional file 4: Figure S2.** Analyses of the LHT transcriptome from *M. sanguinipes* based on GO level II. A) Biological process. B) Component cellular. C) Molecular function. (PDF 452 kb)

**Additional file 5: Table S3.** Transcriptome Annotation of the closest BLASTx hits (1e-5 cutoff). (XLSX 676 kb)

**Additional file 6: Figure S3.** Protein sequence alignment of contig (MS0013) and N-terminal sequence protein of OSP obtained from Yi and Gillott. [53]. Signal peptide sequence is showed within the square. (PDF 73 kb)

**Additional file 7: Table S4.** Predict secreted proteins based on the presence of a signal peptide (SP) or by non-classical secretory pathways (NC) (Scoring function > 0.5). Expression patterns using qualitative RT-PCR to proteins with SP [(a) Male expression exclusively; (b) Expression in both sexes and (c) Not amplified. (XLSX 47 kb)

## Abbreviations

LHT: Long hyaline tubules; EJAC-SP: Ejaculate serine protease; ACE: Angiotensin-converting enzyme; SFPs: Seminal fluid proteins; AG: Accessory gland; OSP: Oviposition Stimulating Protein; NRC-Saskatoon: National Research Council Plant Biotechnology Institute Canada; GO: Gene Ontology; NCBI: Non-redundant databases of the National Center for Biotechnology information; ORFs: Open reading frames; LC-MS/MS: Liquid chromatography with tandem mass spectrometry; RBBHs: Reciprocal best Blast hits; JHBP: Juvenile hormone binding proteins; T: Testes; MHL: Male head and leg; RTF: Female spermatheca and oviduct; FHL: Female head and leg; RT-PCR: Reverse

transcription polymerase chain reaction; EST: Expressed Sequence Tags; EmPAI: Exponentially modified protein abundance values; JH: Juvenile hormone; TO: Take-out; PGF<sub>2</sub>: Prostaglandin F synthase; PGE<sub>2</sub>: Prostaglandin E synthase 2; PG: Prostaglandins; GST: Glutathione S-transferase Glutathione S-transferase; CAT: Catalase; ROS: Reactive oxygen species; SERPINS: Serine protease inhibitors; HSPs: Heat-shock proteins.

## Competing interests

The authors declare that they have no competing interests.

## Authors' contributions

JA, CT and ME design the experiments; MLB and JA analyzed the data. JA and MLB wrote the paper. All authors have read and approve this manuscript.

## Authors' information

JA is an evolutionary biologist interested in the genetic architecture of adaptative traits and reproductive isolation. ME is an insect molecular physiologist interested in pest management. MLB is an agronomist and CT is a plant molecular physiologist.

## Acknowledgements

We thank two anonymous reviewers for their thoughtful comments on an early version of this manuscript. This study was funded by a NSERC Discovery grant to JA and an ADF grant from the Saskatchewan Ministry of Agriculture to CT, JA and ME.

## Author details

<sup>1</sup>Facultad de Ciencias Agropecuarias, Universidad Nacional de Colombia, A.237. Palmira, Valle del Cauca, Colombia. <sup>2</sup>Department of Biology, University of Saskatchewan, 112 Science Pl., Saskatoon, SK S7N-5E2, Canada. <sup>3</sup>Saskatoon Research Centre, Agriculture and Agri-Food Canada, 107 Science Pl., Saskatoon, SK S7N-0X2, Canada.

Received: 15 May 2015 Accepted: 15 December 2015

Published online: 22 December 2015

## References

1. Prout T, Clark AG. Seminal fluid causes temporarily reduced egg hatch in previously mated females. *Proc R Soc Lond Ser B Biol Sci.* 2000;267(1439):201–3.
2. Harshman LG, Prout T. Sperm displacement without sperm transfer in *Drosophila melanogaster*. *Evolution.* 1994;758–766.
3. Wigby S, Sirot LK, Linklater JR, Buehner N, Calboli FC, Bretman A, et al. Seminal fluid protein allocation and male reproductive success. *Curr Biol.* 2009;19(9):751–7.
4. Simmons LW. Sperm competition and its evolutionary consequences in the insect. Princeton: Princeton University Press; 2001.
5. Wong A. The molecular evolution of animal reproductive tract proteins: what have we learned from mating-system comparisons? *Int J Evol Biol.* 2011;2011.
6. Wong A, Rundle H. Selection on the *Drosophila* seminal fluid protein Acp62F. *Ecol Evol.* 2013.
7. Avila FW, Sirot LK, LaFlamme BA, Rubinstein CD, Wolfner MF. Insect Seminal Fluid Proteins: Identification and Function. *Annu Rev Entomol.* 2011;56(1):21–40.
8. Chen PS. The Functional Morphology and Biochemistry of Insect Male Accessory Glands and their Secretions. *Annu Rev Entomol.* 1984;29(1):233–55.
9. Gillott C. MALE ACCESSORY GLAND SECRETIONS: Modulators of Female Reproductive Physiology and Behavior. *Annu Rev Entomol.* 2003;48(1):163–84.
10. Leopold RA. The Role of Male Accessory Glands in Insect Reproduction. *Annu Rev Entomol.* 1976;21(1):199–221.
11. Singh RS, Kulathinal RJ. Sex gene pool evolution and speciation: a new paradigm. *Gene Genet Syst.* 2000;75(3):119–30.
12. Wolfner M. Tokens of love: functions and regulation of *Drosophila* male accessory gland products. *Insect Biochem Mol Biol.* 1997;27:179–92.
13. Chapman T, Arnqvist G, Bangham J, Rowe L. Sexual conflict. *Trends Ecol Evol.* 2003;18(1):41–7.
14. Andrés JA, Maroja LS, Harrison RG. Searching for candidate speciation genes using a proteomic approach: seminal proteins in field crickets. *Proc R Soc B Biol Sci.* 2008;275(1646):1975–83.
15. Fiumera AC, Dumont BL, Clark AG. Sperm Competitive Ability in *Drosophila melanogaster* Associated With Variation in Male Reproductive Proteins. *Genetics.* 2005;169(1):243–57.

16. Andrés JA, Larson EL, Bogdanowicz SM, Harrison RG. Patterns of Transcriptome Divergence in the Male Accessory Gland of Two Closely Related Species of Field Crickets. *Genetics*. 2013;193(2):501–13.
17. Larson EL, Andrés JA, Harrison RG. Influence of the Male Ejaculate on Post-Mating Prezygotic Barriers in Field Crickets. *PLoS ONE*. 2012;7(10):e46202.
18. Civetta A, Singh RS. Sex-related genes, directional sexual selection, and speciation. *Mol Biol Evol*. 1998;15(7):901–9.
19. Sirot LK, Hardstone MC, Helinski MEH, Ribeiro JMC, Kimura M, Deewatthanawong P, et al. Towards a Semen Proteome of the Dengue Vector Mosquito: Protein Identification and Potential Functions. *PLoS Negl Trop Dis*. 2011;5(3):e989.
20. Gavrilits S. Rapid evolution of reproductive barriers driven by sexual conflict. *Nature*. 2000;403(6772):886–9.
21. Swanson W, Clark A, Waldrip-Dail H, Wolfner M, Aquadro C. Evolutionary EST analysis identifies rapidly evolving male reproductive proteins in *Drosophila*. *Proc Natl Acad Sci U S A*. 2001;98:7375–9.
22. Swanson WJ, Vacquier VD. The rapid evolution of reproductive proteins. *Nat Rev Genet*. 2002;3(2):137–44.
23. Civetta A, Singh RS. High divergence of reproductive tract proteins and their association with postzygotic reproductive isolation in *Drosophila melanogaster* and *Drosophila virilis* group species. *J Mol Evol*. 1995;41(6):1085–95.
24. Xu J, Baulding J, Palli SR. Proteomics of *Tribolium castaneum* seminal fluid proteins: Identification of an angiotensin-converting enzyme as a key player in regulation of reproduction. *J Proteome*. 2013;78:83–93.
25. Baldini F, Gabrieli P, Rogers DW, Catteruccia F. Function and composition of male accessory gland secretions in *Anopheles gambiae*: a comparison with other insect vectors of infectious diseases. *Pathog Glob Health* 2012;106(2):82–93.
26. LaFlamme BA, Wolfner MF. Identification and function of proteolysis regulators in seminal fluid. *Mol Reprod Dev*. 2013;80(2):80–101.
27. LaFlamme BA, Ravi Ram K, Wolfner MF. The *Drosophila melanogaster* Seminal Fluid Protease “Seminase” Regulates Proteolytic and Post-Mating Reproductive Processes. *PLoS Genet*. 2012;8(1):e1002435.
28. Findlay GD, Swanson WJ. Proteomics enhances evolutionary and functional analysis of reproductive proteins. *Bioessays*. 2010;32(1):26–36.
29. Findlay GD, Yi X, MacCoss MJ, Swanson WJ. Proteomics Reveals Novel *Drosophila* Seminal Fluid Proteins Transferred at Mating. *PLoS Biol*. 2008;6(7):e178.
30. Sonenshine DE, Bissinger BW, Egekwu N, Donohue KV, Khalil SM, Roe RM. First Transcriptome of the Testis-Vas Deferens-Male Accessory Gland and Proteome of the Spermatophore from *Dermacentor variabilis* (Acari: Ixodidae). *PLoS ONE*. 2011;6(9):e24711.
31. South A, Sirot LK, Lewis SM. Identification of predicted seminal fluid proteins in *Tribolium castaneum*. *Insect Mol Biol*. 2011;20(4):447–56.
32. Walters JR, Harrison RG. Combined EST and proteomic analysis identifies rapidly evolving seminal fluid proteins in *Heliconius* butterflies. *Mol Biol Evol*. 2010;27(9):2000–13.
33. Walters J, Harrison R. EST analysis of male accessory glands from *Heliconius* butterflies with divergent mating systems. *BMC Genomics*. 2008;9(1):592.
34. Walker M, Rylett C, Keen J, Audsley N, Sajid M, Shirras A, et al. Proteomic identification of *Drosophila melanogaster* male accessory gland proteins, including a pro-cathepsin and a soluble gamma-glutamyl transpeptidase. *Proteome Sci*. 2006;4(1):9.
35. Baer B, Heazlewood JL, Taylor NL, Eubel H, Millar AH. The seminal fluid proteome of the honeybee *Apis mellifera*. *Proteomics*. 2009;9(8):2085–97.
36. Ram KR, Wolfner MF. Sustained Post-Mating Response in *Drosophila melanogaster* Requires Multiple Seminal Fluid Proteins. *PLoS Genet*. 2007;3(12):e238.
37. Ravi Ram K, Ji S, Wolfner M. Fates and targets of male accessory gland proteins in mated female *Drosophila melanogaster*. *Insect Biochem Mol Biol*. 2005;35:1059–71.
38. Rylett CM, Walker MJ, Howell GJ, Shirras AD, Isaac RE. Male accessory glands of *Drosophila melanogaster* make a secreted angiotensin I-converting enzyme (ANCE), suggesting a role for the peptide-processing enzyme in seminal fluid. *J Experimental Biol*. 2007;210(20):3601–6.
39. Reinhardt K, Wong CH, Georgiou AS. Detection of seminal fluid proteins in the bed bug, *Cimex lectularius*, using two-dimensional gel electrophoresis and mass spectrometry. *Parasitology*. 2009;136(03):283–92.
40. Andrés JA, Maroja LS, Bogdanowicz SM, Swanson WJ, Harrison RG. Molecular evolution of seminal proteins in field crickets. *Mol Biol Evol*. 2006;23(8):1574–84.
41. Davies SJ, Chapman T. Identification of genes expressed in the accessory glands of male Mediterranean Fruit Flies (*Ceratitis capitata*). *Insect Biochem Mol Biol*. 2006;36(11):846–56.
42. Braswell WE, Andrés JA, Maroja LS, Harrison RG, Howard DJ, Swanson WJ. Identification and comparative analysis of accessory gland proteins in Orthoptera. *Genome*. 2006;49(9):1069–80.
43. Chapman T, Bangham J, Vinti G, Seifried B, Lung O, Wolfner M, et al. The sex peptide of *Drosophila melanogaster*: female post-mating responses analyzed by using RNA interference. *Proc Natl Acad Sci U S A*. 2003;100:9923–8.
44. Chapman T, Davies S. Functions and analysis of the seminal fluid proteins of male *Drosophila melanogaster* fruit flies. *Peptides*. 2004;25:1477–90.
45. Lange AB, Loughton BG. An oviposition-stimulating factor in the male accessory reproductive gland of the locust, *Locusta migratoria*. *Gen Comp Endocrinol*. 1985;57(2):208–15.
46. Marshall JL, Huestis DL, Hiromasa Y, Wheeler S, Oppert C, Marshall SA, et al. Identification, RNAi Knockdown, and Functional Analysis of an Ejaculate Protein that Mediates a Postmating, Prezygotic Phenotype in a Cricket. *PLoS ONE*. 2009;4(10), e7537.
47. Olfert O, Weiss R. Impact of grasshopper feeding on selected cultivars of cruciferous oilseed crops. *J Orthoptera Res*. 2002;11(1):83–6.
48. Branson DH, Joern A, Sword GA. Sustainable management of insect herbivores in grassland ecosystems: new perspectives in grasshopper control. *Bioscience*. 2006;56(9):743–55.
49. Pickford R, Gillott C. Insemination in the migratory grasshopper, *Melanoplus sanguinipes* (Fabr.). *Can J Zool*. 1971;49(12):1583–8.
50. Pickford R, Ewen AB, Gillott C. Male accessory gland substance: an egg-laying stimulant in *Melanoplus sanguinipes* (F.) (Orthoptera:Acrididae). *Can J Zool*. 1969;47(6):1199–203.
51. Cheeseman MT, Gillott C. Identification and partial characterization of the major secretory protein of the long hyaline gland in the male grasshopper, *Melanoplus sanguinipes*. *Insect Biochem*. 1988;18(2):135–44.
52. Gillott C. Male insect accessory glands: Functions and control of secretory activity. *Invertebr Reprod Dev*. 1996;30(1-3):199–205.
53. Yi S-X, Gillott C. Purification and characterization of an oviposition-stimulating protein of the long hyaline tubules in the male migratory grasshopper, *Melanoplus sanguinipes*. *J Insect Physiol*. 1999;45(2):143–50.
54. Cheeseman M, Gillott C. Long hyaline gland discharge and multiple spermatophore formation by the male grasshopper, *Melanoplus sanguinipes*. *Physiol Entomol*. 1989;14(3):257–64.
55. Couche G, Gillott C, Tobe S, Feyerisen R. Juvenile hormone biosynthesis during sexual maturation and after mating in the adult male migratory grasshopper, *Melanoplus sanguinipes*. *Can J Zool*. 1985;63(12):2789–92.
56. Ewen AB, Hinks CF. Rearing a non-diapause strain of the migratory grasshopper, *Melanoplus sanguinipes* (F.) (Orthoptera: Acrididae) in the laboratory. In: Proceedings of the 4th Triennial Meeting Pan American Acridological Society: 1986. 1986. p. 169–73.
57. Götz S, García-Gómez JM, Terol J, Williams TD, Nagaraj SH, Nueda MJ, et al. High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Res*. 2008;36(10):3420–35.
58. Rice P, Longden I, Bleasby A. EMBOSS: The European Molecular Biology Open Software Suite. *Trends Genet*. 2000;16(6):276–7.
59. Bateman A, Coin L, Durbin R, Finn RD, Hollich V, Griffiths-Jones S, et al. The Pfam protein families database. *Nucleic Acids Res*. 2004;32 suppl 1:D138–41.
60. Friedel T, Gillott C. Male accessory gland substance of *Melanoplus sanguinipes*: An oviposition stimulant under the control of the corpus allatum. *J Insect Physiol*. 1976;22(3):489–95.
61. Dean MD, Findlay GD, Hoopmann MR, Wu CC, MacCoss MJ, Swanson WJ, et al. Identification of ejaculated proteins in the house mouse (*Mus domesticus*) via isotopic labeling. *BMC Genomics*. 2011;12(1):306.
62. Boes KE, Ribeiro JM, Wong A, Harrington LC, Wolfner MF, Sirot LK. Identification and characterization of seminal fluid proteins in the Asian tiger mosquito, *Aedes albopictus*. 2014.
63. Findlay GD, MacCoss MJ, Swanson WJ. Proteomic discovery of previously unannotated, rapidly evolving seminal fluid genes in *Drosophila*. *Genome Res*. 2009;19(5):886–96.
64. Wolfner MF, Harada HA, Bertram MJ, Stelick TJ, Kraus KW, Kalb JM, et al. New genes for male accessory gland proteins in *Drosophila melanogaster*. *Insect Biochem Mol Biol*. 1997;27(10):825–34.
65. Sirot LK, Poulson RL, Caitlin McKenna M, Girnary H, Wolfner MF, Harrington LC. 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.* 2008;38(2):176–89.
66. Ishihama Y, Oda Y, Tabata T, Sato T, Nagasu T, Rappsilber J, et al. Exponentially Modified Protein Abundance Index (emPAI) for Estimation of Absolute Protein Amount in Proteomics by the Number of Sequenced Peptides per Protein. *Mol Cell Proteomics.* 2005;4(9):1265–72.
  67. Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods.* 2011;8(10):785–6.
  68. Bendtsen JD, Jensen LJ, Blom N, Von Heijne G, Brunak S. Feature-based prediction of non-classical and leaderless protein secretion. *Protein Eng Des Sel.* 2004;17(4):349–56.
  69. Marshall JL, Huestis DL, Garcia C, Hiromasa Y, Wheeler S, Noh S, et al. Comparative Proteomics Uncovers the Signature of Natural Selection Acting on the Ejaculate Proteomes of Two Cricket Species Isolated by Postmating, Prezygotic Phenotypes. *Mol Biol Evol.* 2011;28(1):423–35.
  70. Noriega F, Ribeiro J, Koener J, Valenzuela J, Hernandez-Martinez S, Pham V, et al. Comparative genomics of insect juvenile hormone biosynthesis. *Insect Biochem Mol Biol.* 2006;36(4):366–74.
  71. Charles J-P, Wojtasek H, Lentz AJ, Thomas BA, Bonning BC, Palli SR, et al. Purification and reassessment of ligand binding by the recombinant, putative juvenile hormone receptor of the tobacco hornworm, *Manduca sexta*. *Arch Insect Biochem Physiol.* 1996;31(4):371–93.
  72. Justice R, Dimitratos S, Walter M, Woods D, Biessmann H. Sexual dimorphic expression of putative antennal carrier protein genes in the malaria vector *Anopheles gambiae*. *Insect Mol Biol.* 2003;12(6):581–94.
  73. Scolari F, Gomulski LM, Ribeiro JM, Siciliano P, Meraldi A, Falchetto M, et al. Transcriptional profiles of mating-responsive genes from testes and male accessory glands of the Mediterranean fruit fly, *Ceratitis capitata*. *PLoS ONE.* 2012;7(10), e46812.
  74. Wybrandt GB, Andersen SO. Purification and sequence determination of a yellow protein from sexually mature males of the desert locust, *Schistocerca gregaria*. *Insect Biochem Mol Biol.* 2001;31(12):1183–9.
  75. Saito K, Su ZH, Emi A, Mita K, Takeda M, Fujiwara Y. Cloning and expression analysis of takeout/JHBP family genes of silkworm, *Bombyx mori*. *Insect Mol Biol.* 2006;15(3):245–51.
  76. Olfert O, Weiss R, Kriticos D. Application of general circulation models to assess the potential impact of climate change on potential distribution and relative abundance of *Melanoplus sanguinipes* (Fabricius)(Orthoptera: Acrididae) in North America. *Psyche.* 2011;2010.
  77. Simmons L, Tan YF, Millar A. Sperm and seminal fluid proteomes of the field cricket *Teleogryllus oceanicus*: identification of novel proteins transferred to females at mating. *Insect Mol Biol.* 2013;22(1):115–30.
  78. Gubb D, Sanz-Parra A, Barcena L, Troxler L, Fullaondo A. Protease inhibitors and proteolytic signalling cascades in insects. *Biochimie.* 2010;92(12):1749–59.
  79. Nielsen H, Brunak S, von Heijne G. Machine learning approaches for the prediction of signal peptides and other protein sorting signals. *Protein Eng.* 1999;12(1):3–9.
  80. Hagai T, Cohen M, Bloch G. Genes encoding putative Takeout/Juvenile hormone binding proteins in the honeybee (*Apis mellifera*) and modulation by age and juvenile hormone of the takeout-like gene *GB19811*. *Insect Biochem Mol Biol.* 2007;37(7):689–701.
  81. Bajgar A, Jindra M, Dolezel D. Autonomous regulation of the insect gut by circadian genes acting downstream of juvenile hormone signaling. *Proc Natl Acad Sci.* 2013.
  82. Sieminska E, Andres J, Todd C, Erlandson M. Characterization of *Melanoplus sanguinipes* oviposition stimulating protein expression and re-examination of its potential role in stimulating oviposition. *J Insect Physiol.* 2015.
  83. Stanley-Samuels DW, Loher W. Arachidonic and other long-chain polyunsaturated fatty acids in spermatophores and spermathecae of *Teleogryllus commodus*: significance in prostaglandin-mediated reproductive behaviour. *J Insect Physiol.* 1983;29(1):41–5.
  84. Destephano DB, Brady UE. Prostaglandin and prostaglandin synthetase in the cricket, *Acheta domestica*. *J Insect Physiol.* 1977;23(7):905–11.
  85. Loher W, Ganjian I, Kubo I, Stanley-Samuels D, Tobe SS. Prostaglandins: their role in egg-laying of the cricket *Teleogryllus commodus*. *Proc Natl Acad Sci.* 1981;78(12):7835–8.
  86. Lange AB. The transfer of prostaglandin-synthesizing activity during mating in *Locusta migratoria*. *Insect Biochem.* 1984;14(5):551–6.
  87. Setty BY, Ramaiah T. Isolation and identification of prostaglandins from the reproductive organs of male silkworm, *Bombyx mori* L. *Insect Biochem.* 1979;9(6):613–7.
  88. Stanley-Samuels DW, Peloquin JJ, Loher W. Egg-laying in response to prostaglandin injections in the Australian field cricket, *Teleogryllus commodus*. *Physiol Entomol.* 1986;11(2):213–9.
  89. Collins A, Williams V, Evans J. Sperm storage and antioxidative enzyme expression in the honey bee, *Apis mellifera*. *Insect Mol Biol.* 2004;13(2):141–6.
  90. Shaw WR, Teodori E, Mitchell SN, Baldini F, Gabrieli P, Rogers DW, et al. Mating activates the heme peroxidase HPX15 in the sperm storage organ to ensure fertility in *Anopheles gambiae*. *Proc Natl Acad Sci.* 2014;111(16):5854–9.
  91. Prokupek AM, Kachman SD, Ladunga I, Harshman LG. Transcriptional profiling of the sperm storage organs of *Drosophila melanogaster*. *Insect Mol Biol.* 2009;18(4):465–75.
  92. Pascini TV, Ramalho-Ortigão M, Martins GF. Morphological and morphometrical assessment of spermathecae of *Aedes aegypti* females. *Mem Inst Oswaldo Cruz.* 2012;107(6):705–12.
  93. Stanley-Samuels DW. The biological significance of prostaglandins and related eicosanoids in insects. In: *Insect lipids: Chemistry, biochemistry and biology.* 1993. p. 45–97.
  94. Thomas J, RYLETT C, Carhan A, BLAND N, Bingham R, SHIRRAS A, et al. *Drosophila melanogaster* NEP2 is a new soluble member of the neprilysin family of endopeptidases with implications for reproduction and renal function. *Biochem J.* 2005;386:357–66.
  95. Hurst D, Rylett CM, Isaac RE, Shirras AD. The drosophila angiotensin-converting enzyme homologue Ance is required for spermiogenesis. *Dev Biol.* 2003;254(2):238–47.
  96. Ekbote U, Looker M, Isaac RE. ACE inhibitors reduce fecundity in the mosquito, *Anopheles stephensi*. *Comp Biochem Physiol B Biochem Mol Biol.* 2003;134(4):593–8.
  97. Chen S-S, Chang LS, Chen H-W, Wei Y-H. Polymorphisms of glutathione S-transferase M1 and male infertility in Taiwanese patients with varicocele. *Hum Reprod.* 2002;17(3):718–25.
  98. Dunkov B, Georgieva T. Insect iron binding proteins: Insights from the genomes. *Insect Biochem Mol Biol.* 2006;36(4):300–9.
  99. Ber A, Vardinon N, Yogev L, Yavetz H, Homonnai Z, Yust I, et al. Transferrin in seminal plasma and in serum of men: its correlation with sperm quality and hormonal status. *Hum Reprod.* 1990;5(3):294–7.
  100. Ashrafzadeh A, Karsani SA, Nathan S. Mammalian sperm fertility related proteins. *Int J Med Sci.* 2013;10(12):1649.
  101. Frenette G, Lessard C, Sullivan R. Polyol pathway along the bovine epididymis. *Mol Reprod Dev.* 2004;69(4):448–56.
  102. Georgieva T, Dunkov B, Dimov S, Ralchev K, Law J. *Drosophila melanogaster* ferritin: cDNA encoding a light chain homologue, temporal and tissue specific expression of both subunit types. *Insect Biochem Mol Biol.* 2002;32(3):295–302.
  103. Dunkov BC, Georgieva T, Yoshiga T, Hall M, Law JH. *Aedes aegypti* ferritin heavy chain homologue: feeding of iron or blood influences message levels, lengths and subunit abundance. *J Insect Sci.* 2002;2(1):7.
  104. Bartholomay LC, Cho W-L, Rocheleau TA, Boyle JP, Beck ET, Fuchs JF, et al. Description of the transcriptomes of immune response-activated hemocytes from the mosquito vectors *Aedes aegypti* and *Armigeres subalbatus*. *Infect Immun.* 2004;72(7):4114–26.
  105. Levy F, Bulet P, Ehret-Sabatier L. Proteomic analysis of the systemic immune response of *Drosophila*. *Mol Cell Proteomics.* 2004;3(2):156–66.
  106. Nichol H, Law JH, Winzerling JJ. Iron metabolism in insects. *Annu Rev Entomol.* 2002;47(1):535–59.
  107. Hajdusek O, Sojka D, Kopacek P, Buresova V, Franta Z, Sauman I, et al. Knockdown of proteins involved in iron metabolism limits tick reproduction and development. *Proc Natl Acad Sci.* 2009;106(4):1033–8.
  108. Miki K. Energy metabolism and sperm function. *Soc Reprod Fertil Suppl.* 2006;65:309–25.
  109. Smith G, Rothwell K, Wood S, Yeaman S, Bownes M. Specificity and localization of lipolytic activity in adult *Drosophila melanogaster*. *Biochem J.* 1994;304:775–9.
  110. Iida K, Cavener DR. Glucose dehydrogenase is required for normal sperm storage and utilization in female *Drosophila melanogaster*. *J Exp Biol.* 2004;207(4):675–81.
  111. Xu H-R, Lu J-C, Chen F, Huang Y-F, Yao B, Lu N-Q. The effect of chymotrypsin on the determination of total alpha-glucosidase activity in seminal plasma and the correlation between alpha-glucosidase level and semen parameters. *Syst Biol Reprod Med.* 2006;52(6):441–6.
  112. So WW, Sarov-Blat L, Kotarski CK, McDonald MJ, Allada R, Rosbash M. Takeout, a novel *Drosophilagene* under circadian clock transcriptional regulation. *Mol Cell Biol.* 2000;20(18):6935–44.

113. Lorenz L, Hall JC, Rosbash M. Expression of a *Drosophila* mRNA is under circadian clock control during pupation. *Development*. 1989;107(4):869–80.
114. Palli SR, Touhara K, Charles J-P, Bonning BC, Atkinson JK, Trowell SC, et al. A nuclear juvenile hormone-binding protein from larvae of *Manduca sexta*: a putative receptor for the metamorphic action of juvenile hormone. *Proc Natl Acad Sci*. 1994;91(13):6191–5.
115. Du J, Hiruma K, Riddiford LM. A novel gene in the *takeout* gene family is regulated by hormones and nutrients in *Manduca* larval epidermis. *Insect Biochem Mol Biol*. 2003;33(8):803–14.
116. Wojtasek H, Prestwich GD. Key disulfide bonds in an insect hormone binding protein: cDNA cloning of a juvenile hormone binding protein of *Heliothis virescens* and ligand binding by native and mutant forms. *Biochemistry*. 1995;34(15):5234–41.
117. Han J, Lee CS, Yun CY, Lee BH, Ko YG, Kang CS, et al. Cloning and expression of male-specific protein (MSP) from the hemolymph of greater wax moth, *Galleria mellonella* L. *Arch Insect Biochem Physiol*. 2003;54(3): 110–20.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

