Published in final edited form as:

Dev Comp Immunol. 2017 September ; 74: 110–124. doi:10.1016/j.dci.2017.04.009.

Hemolymph proteins of *Anopheles gambiae* larvae infected by *Escherichia coli*

Xuesong He^{a,*}, Xiaolong Cao^{a,b,*}, Yan He^a, Krishna Bhattacharj^{a,b}, Janet Rogers^b, Steve Hartson^b, and Haobo Jiang^{a,b,†}

^aDepartment of Entomology and Plant Pathology, Oklahoma State University, Stillwater, OK 74078, USA

^bDepartment of Biochemistry and Molecular Biology, Oklahoma State University, Stillwater, OK 74078, USA

Abstract

Anopheles gambiae is a major vector of human malaria and its immune system in part determines the fate of ingested parasites. Proteins, hemocytes and fat body in hemolymph are critical components of this system, mediating both humoral and cellular defenses. Here we assessed differences in the hemolymph proteomes of water- and *E. coli*-pricked mosquito larvae by a gel-LC-MS approach. Among the 1,756 proteins identified, 603 contained a signal peptide but accounted for two-third of the total protein amount on the quantitative basis. The sequence homology search indicated that 233 of the 1,756 may be related to defense. In general, we did not detect substantial differences between the control and induced plasma samples in terms of protein numbers or levels. Protein distributions in the gel slices suggested post-translational modifications (*e.g.* proteolysis) and formation of serpin-protease complexes and high M_r immune complexes. Based on the twenty-five most abundant proteins, we further suggest that major functions of the larval hemolymph are storage, transport, and immunity. In summary, this study provided first data on constitution, levels, and possible functions of hemolymph proteins in the mosquito larvae, reflecting complex changes occurring in the fight against *E. coli* infection.

Graphical Abstract

[†]Corresponding author: Dr. Haobo Jiang, Regents Professor, 127 Noble Research Center, Department of Entomology and Plant Pathology, Oklahoma State University, Stillwater, OK 74078, haobo.jiang@okstate.edu.

^{*}These authors contribute equally to this work.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Keywords

insect immunity; hemolymph proteins; LC-MS/MS; label-free quantification; serine protease; gel mobility shift

1. Introduction

The mosquito *Anopheles gambiae* is one of the major vector species transmitting deadly diseases, which impact millions of human lives each year (World Malaria Report 2015). Since its genome sequence became available (Holt et al. 2002), continuous efforts have been made to improve gene annotation, profile transcript levels, and elucidate protein functions, especially those related to immunity (Yassine and Osta, 2010; Clayton et al., 2014). The genetic makeup of the mosquito innate immune system is considered to be similar to those of other model insects such as *Drosophila melanogaster*. In these models, molecular patterns on the pathogen surface are recognized by pattern recognition receptors (PRRs) of the host to trigger serine protease (SP) cascades, melanization, antimicrobial protein (AMP) synthesis, and hemocyte responses. Families of defense proteins are implicated in parasite resistance, including thioester proteins (e.g. TEP1) (Blandin et al., 2004), Leu-rich repeat (LRR) proteins (e.g. Leu-rich immune molecule-1 (LRIM1), *A. plasmodium*-responsive LRR protein-1C (APL1C)) (Povelones et al., 2009), C-type lectins (CTLs) (e.g. CTL4, CTLMA2) (Osta et al., 2004), fibrinogen-related proteins (FREP1, FBN30) (Dong and Dimopoulos, 2009; Li et al., 2013), and clip-domain serine proteases (SPs) and serine protease homologs (SPHs) (e.g. CLIPs B3, B4, B8, B14, B15, B17, A2, A5, A7, A8 and SPCLIP1) (Volz et al., 2005 and 2006; Povelones et al., 2013; Barillas-Mury, 2007), serpins (e.g. serpin-2, 6) (Abraham et al., 2005; Michel et al., 2006; Suwanchaichinda and Kanost, 2009), and caspase-S2 (Ramphul et al., 2015). Four of these genes (APL1C, CLIPB15, serpin-2, caspase-S2) were also identified in an expression pattern analysis, along with TEP2, TEP4 and TEP15, LRR-7060, CTL2, CTLMA3, FBN9, peptidoglycan recognition protein (PGRP) S1, CLIPs B5, B7 and C2, SCRA-SPH3, Spätzle3 and 4, Toll6, 10 and 11, inhibitor of apoptosis (IAP) 5, caspase-S4, and adenosine deaminase genes (Li et al., 2013).

Previous proteomic studies of *A. gambiae* characterized peritrophic matrix, head, eggshell, saliva and salivary gland (Dinglasan et al. 2009; Lefevre et al., 2007; Amenya et al., 2010; Francischetti et al., 2002; Kalume et al., 2005), identifying peritrophins, digestive and metabolic enzymes, vitelline membrane and chorion proteins, oxidases, odorant binding proteins, and SG and D7 proteins. In the hemolymph of adult mosquitoes, Paskewitz and Shi (2005) described 280 spots on 2-D PAGE gels and identified 28 including those related to immunity (e.g. TEP15, CLIPB4, CLIPA6, serpin-2, serpin-15, phenoloxidase-6 (PO6), proPO2), lipid binding (e.g. apolipophorin III, MD2-like protein-3 (MDL3)), and iron metabolism (e.g. ferritin). Eight of the 28 proteins were differentially regulated upon wounding (e.g. glutathione S transferase-S1) or bacterial infection (PO6, chitinase-like BR-1 and BR-2). While these studies provided useful information about proteins in tissues or body fluids in the mosquito, the numbers of identified proteins were dwarfed by those of mRNAs detected in the corresponding tissue transcriptomes, due to limitations of the proteomic techniques previously used. To better understand the roles of hemolymph in physiological processes such as innate immunity, we chose the *E. coli* infection model which had been used in *A. gambiae* adults to show lethality at high dosage and induced production of defensin (Blandin et al., 2002; Coggins et al., 2012) and characterized *A. gambiae* larval hemolymph using a gel-LC-MS approach (Blagoev et al., 2004). We report the identification and quantification of plasma proteins, some of which showed challenge-associated changes in abundance or gel mobility, suggesting proteolytic processing and immune complex formation.

2. Materials and Methods

2.1. Mosquito rearing

A colony of *A. gambiae* G3 strain was obtained from Malaria Research and Reference Reagent Resource Center and maintained in an incubator at 27.5°C with 80% relative humidity in a 12 h light-dark cycle with gradual sunset and sunrise light transitions. As described previously (Benedict, 1997), newly hatched larvae (day 1–2) were fed in suspension of baker's yeast, the older larvae were fed a 1:2 (w/w) mixture of baker's yeast and ground fish food (Mike Reed Enterprises) in distilled water. Pupae were picked and placed side by side in cups with water prior to emergence. Newly emerged adults were maintained by 10% sucrose solution. To trigger embryonic development, adult females (days 6–10) were fed heparinized sheep blood (HemoStat Laboratories) using a membrane feeder (Hemotek). Laid eggs were collected on wet filter papers and transferred to distilled water for hatching.

2.2. Preparation of cell-free hemolymph from water- and bacteria-pricked larvae

Fourth instar larvae (20–25/group, 8 groups) were transferred to new cups with distilled water and placed on a filter paper to remove excess water. They were individually pricked in the thorax with a glass capillary tube dipped in distilled water (C for control) or a pellet of live *Escherichia coli* cells (I for induced). To prepare the bacterial pellet, a single fresh colony of *E. coli* BL21 was grown in 3.0 ml of LB medium at 37°C with shaking at 220 rpm until OD₆₀₀ reached 0.7–0.9. The cells were harvested by centrifugation at 4,500g for 20 min and resuspended in 1 ml of distilled water. This step was repeated twice to remove the

medium. After being transferred back to the same cup, the wounded larvae were provided with a small volume of the food suspension and incubated for 24 h. For hemolymph extraction, five insects from the same group (C or I) were blotted with a filter paper and placed together on a piece of paraffin film. Five microliters of a cocktail of protease inhibitors (cOmplete ULTRA, Roche Diagnostics) supplemented with 0.1% 1-phenyl-2-thiourea was added to the larvae. Each larva was torn slightly with forceps in the thorax and gently pressed with a pipette tip in the abdomen, so that released hemolymph instantly encountered the inhibitors. Pooled plasma (P) samples (*c.a.* 20 μ l per tube, 20–25 insects per pool) were centrifuged at 2000 $\times g$ for 5 min to remove tissues and cell debris. Protein concentrations of the samples and biological replicates (CP1–CP4, IP1–IP4 from different cohorts of the mosquito larvae) were determined by a modified Bradford assay using BSA as a standard.

2.3. SDS-PAGE separation of plasma proteins, in-gel trypsinolysis, and MS sample preparation

The eight protein samples (4 CPs and 4 IPs) were treated as described previously (He et al., 2016). Briefly, 40 μ g of each sample were separately on a 4–15% gradient SDS polyacrylamide gel (Bio-Rad), followed by staining with Coomassie blue. Subsequently, each of the eight lanes was divided into 12 gel slices based on the band patterns, generating a total of 96 gel samples. The gel pieces were reduced with tris(2-carboxyethyl)phosphine, alkylated with iodoacetamide, and digested with sequencing grade trypsin. Resulting trypsinolytic peptides were extracted from the gel pieces with 1% trifluoroacetic acid for LC-MS/MS analysis at Oklahoma State University Recombinant DNA/Protein Core Facility.

2.4. LC-MS/MS analysis

Each of the 40 samples from gel slices 1 through 5 (*i.e.* >80 kDa, Fig. 1A) was subjected to a single individual LC-MS/MS analysis on an LTQ-Orbitrap mass spectrometer (Thermo Fisher Scientific) as described previously (He et al., 2016), but using a 4-hr chromatography gradient of 2–40% ACN. The other 56 samples (gel slices 6 through 12, Fig. 1A) were each subjected to two LC-MS/MS analyses on an Orbitrap Fusion Tribrid mass spectrometer (Thermo Fisher Scientific), wherein peptides were separated (He et al., 2016) using a 100-min chromatography gradient of 2–40% ACN. The Fusion analyses employed a “Top Speed” data-dependent MS/MS strategy, wherein survey scans were performed in FT mode ($R=120,000$) at least every three seconds. In the interval between survey scans, up to 20 data-dependent MS/MS scans were conducted in the ion trap. The Fusion scan settings also included selection of precursor ions via the quadrupole mass filter, dynamic exclusion ($n=1$), and monoisotopic precursor selection.

2.5. Protein identification, quantification, and mobility examination

The spectra were searched against an *A. gambiae* protein database (PEST Peptides_AgamP4.2) downloaded on 9/1/2015 from VectorBase (<http://www.vectorbase.org/>). MaxQuant version 1.5.2.28 (Cox and Mann, 2008) was used to process raw data from the Orbitrap and Fusion spectrometers for database searching. In setting up the MaxQuant search, all of the MS files obtained for individual gel slices from a single lane were designated as “fractions” describing that lane’s sample. The parameters

used in the search were: no charge state deconvolution or deisotoping, trypsin digestion, maximum missed cleavage of 2, parental ion mass tolerance of 20 PPM, fragment ion mass tolerance of 0.50 Da, formylation or acetylation of protein N-termini, Met oxidation, cyclization of Gln to pyroGlu, iodoacetamide or acrylamide derivatization of Cys, no fixed modification, and all peptides (with or without modification) included in search. False discovery rates for proteins and peptides were set at 1%.

Proteins were quantified on the basis of their normalized label-free quantification (LFQ) intensities as a proxy for their relative abundances (Cox et al., 2014; Ahrné et al., 2013). The four samples from the challenged insects were designated as one protein group (*i.e.* “IP”), whereas the four samples from the control insects were designated as another (*i.e.* “CP”). The average LFQ intensity value for each protein’s expression in either group was calculated as the mean, and used to calculate the $\log_2(I/C)$ expression ratio. Individual protein LFQ intensities among the four biological replicates were also used to calculate the extent to which differences in protein abundance were significant, using Student’s t test and a significance threshold of 0.05.

To examine the electrophoretic mobility of individual proteins, samples were quantified as described above, but instead designating each gel slice as an individual sample when setting up the MaxQuant search (that is, individual gel slices were not designated as collective sample “fractions”). Changes in electrophoretic mobility were analyzed by comparing protein LFQ protein intensities among gel slices.

2.6. Protein nomenclature

Most of the identified protein sequences from the database already had names that suggest function. For uncharacterized or hypothetical proteins, their sequences were used as queries to search non-redundant protein sequences at NCBI by BLASTP (E-value $< 10^{-10}$). For sequences that still remained uncharacterized, BLAST2GO searches were performed using default settings. Based on the results, some were named after the domain(s) they contain and the rest referred by their VectorBase IDs (*i.e.* AGAP numbers). Signal peptides were predicted by SignalP 4.0 (Petersen et al., 2011) and Signal-3L (Shen and Chou, 2007).

2.7. Effects of sterile and septic wounding

To confirm *E. coli* infection, 4th instar larvae of *A. gambiae* were individually pricked in the thorax as described in Section 2.2 with two minor modifications. *E. coli* BL21 strain carrying a GFP expression plasmid was used to prepare bacterial pellet. The bacteria were washed in sterile phosphate-buffered saline (PBS, 138 mM NaCl, 2.7 mM KCl, 10 mM phosphate, pH 7.4). The control larvae were pricked with the same buffer. The live and dead larvae were counted at 24 h after treatment to calculate survival rates of the PBS_p and *E. coli*_p groups. To compare survival and AMP expression caused by pricking and injection, two more groups (PBS_i and *E. coli*_i) of larvae at the same stage were injected at 69 nl/larva with PBS or *E. coli* suspended in PBS (OD₆₀₀ = 1.0, *c.a.* 5.5×10^4 cells/insect), respectively. After scanning for *E. coli* GFP signals under a BX51 fluorescence microscope, 24 total RNA samples were prepared from live whole larvae (10 insects per sample, 2 PBS pricked and 2 *E. coli* pricked, 2 PBS injected, and 2 *E. coli* injected) at 2, 6 and 24 h using TRIZOL

Reagent (Life Technologies, Inc). Two total RNA samples were isolated from naïve larvae (10 insects per sample) and used as negative control. cDNA synthesis and quantitative real-time PCR were performed with three technical replicates for each of the 26 samples according to Yang et al (2016). Each cDNA sample was equivalent to starting with 250 ng total RNA. The primers were: j1723 (5'-AAGAAGCTGACTGGCCGTGA) and j1724 (5'-GTAGCTGCTGCAAACTTCGG) for rpS7; j1737 (5'-TCTGCTGGAACCATCATCGG) and j1738 (5'-ATCTCGTAAACTGCACCGCA) for gambicin-1; j1733 (5'-ATCTTTGTCGTGCTGGCAGT) and j1734 (5'-CTGCCTTGAACACTCGCTTG) for cecropin A. Relative mRNA levels were calculated as: $(1 + E_{rpS7})^{Ct, rpS3} / (1 + E_X)^{Ct, X}$ (Rieu and Powers, 2009), where $E_{rpS7} = 107.8\%$, $E_{gambicin-1} = 109.3\%$, and $E_{cecropin A} = 99.3\%$. The living and dead larvae pricked with PBS or *E. coli* were examined at 24 h by bright-field and fluorescence microscopy under the microscope with a DP-71 imaging system (Olympus).

3. Results and Discussion

3.1. Overview of the proteomics analysis

To identify and quantify proteins in *A. gambiae* larval plasma, we collected hemolymph samples from the larvae pricked with water or live *E. coli* and separated them by SDS-PAGE. There was no clear difference in band pattern between the CP and IP lanes (Fig. 1A). After gel cutting, in-gel trypsinolysis and LC-MS/MS analysis, we found the numbers of proteins identified in 12 groups of gel slices were comparable between CP and IP (Fig. 1B). These slices contained similar amounts of proteins (based on the staining pattern), ensuring that abundant proteins do not overwhelm signals of scarce ones and apparent M_r ranges of the identified proteins are available. To assess the samples' reproducibility, pairwise Pearson correlation coefficients of protein LFQ values were examined, yielding r^2 values of 0.897–0.978 (0.937 ± 0.029) among pairs of the CP samples and 0.959–0.991 (0.976 ± 0.011) among pairs of the IPs. Correlations between the CP and IP samples were 0.908–0.954 (0.941 ± 0.020), indicative of a minor global change in protein *amounts* after the immune challenge.

We identified a total of 1,756 proteins: 1688 from CP, 1695 from IP, and 1627 common to both conditions. Each protein had two or more matching MS/MS spectra (Table S1). Among them, 1,361 (78%) had descriptive names, 305 (17%) were named after their homologs found in the BLASTP search, 18 (1%) were designated with the domain(s) they contain, and the other 72 (4%) were represented by their VectorBase IDs. We manually divided the proteins into nine categories: immunity (233), metabolism (524), DNA/RNA synthesis (105), ion binding (147), cytoskeleton/motor (73), ATP/NAD(P)H binding (181), sensory and cuticle (52), ribosomal (78), and others (363) (Fig. 2A). Among the total, 603 (34.3%) were predicted to have a signal peptide (Table S1); the other 1,153 (65.7%) were predicted to be intracellular but detected in the plasma. This apparent contradiction was also observed in *Manduca sexta* plasma but the percentage (38.7%) was lower (He et al., 2016). We ascribed this anomaly to hemocyte rupture, tissue tearing, and incomplete removal of cellular debris. The contamination was substantial, as the LFQ sums of the extracellular proteins (34.3%) only accounted for 69.7% (CP) and 64.6% (IP) of the grand LFQ totals. An

alternative explanation for these hemolymph proteins lacking an apparent signal peptide is that some of them might be secreted into plasma via unconventional pathways (Nickel, 2010).

3.2. Abundant proteins and their functions in the larval hemolymph

We examined protein distributions based on their relative abundances (Fig. 3A). There were 61 proteins identified in CP but not IP (LFQ = 0) and 68 other were only found in IP. Most proteins (88% in both CP and IP) fell into the LFQ range of 10^6 to 10^9 . However, when total abundances of all proteins within the individual LFQ ranges were compared (Fig. 3B), we found that 23 CP and 21 IP proteins (LFQ $>10^{10}$) accounted for 62.2% and 54.3% of the LFQ grand totals, respectively. Similarly, 140 CP and 140 IP proteins in the LFQ range of 10^9 – 10^{10} represented 25.5% and 31.1% of the total protein amounts. Note that LFQ percentages presented here should only be treated as estimates and large deviations may exist, especially when individual proteins are compared.

We examined 25 most abundant proteins (Table 1). Five hexamerins accounted for 27.8% and 24.5% of the total protein amounts (Σ LFQ_{1-1,756}) in the CP and IP samples on average. These storage proteins support metamorphosis in the pupal stage. Lipophorins (9.6% CP; 8.3% IP) and vitellogenin (4.7% CP; 4.1% IP) transport lipids; larval serum protein-1 β , chemosensory protein and odorant binding protein-9 (5.9% CP and 4.7% IP) may also bind lipids. Ferritins (3.2% CP; 2.3% IP) store and transport iron (Larade and Storey, 2004; Ong et al., 2005). To our surprise, proPO2 and proPO3 (3.3% CP; 3.1%) are among the top 25 most abundant proteins, along with imaginal disc growth factor (a homolog of hemocyte aggregation inhibitor protein), TEPI5 and secreted gelsolin (a component of hemolymph clot) (3.3% CP; 2.8% IP). These defense proteins may participate in melanization, pathogen recognition, coagulation, and hemocyte aggregation (Pesch et al., 2016; Blandin et al., 2004; Levashina et al., 2001; Karlsson et al., 2004). We also correlated the intense bands (Fig. 1A) with the most abundant protein(s) in the corresponding gel slices. Slice 3 (230–250 kDa) contained mainly hexamerins (57% CP and 47% IP), and so did slice 6 (70–80 kDa, 53% CP and 34% IP). We understood the dissociation of hexamerins to 80 kDa monomers, but not why some of the 480 kDa complexes migrated to 240 kDa position. Functions of these covalently linked trimers are unclear. Apolipophorin III (74% CP and 67% IP) represented the intense band in slice 10 (20–22 kDa). In summary, identification of the abundant proteins suggests that the major functions of hemolymph are storage, transport, and immunity. This agrees with the findings on the adult hemolymph proteome (Paskewitz and Shi, 2005).

3.3. Hemolymph response to the bacterial treatment

Among the 1,756 proteins identified, 158 or 9% demonstrated significant ($p < 0.05$) changes in abundance between CP and IP. As a part of the response, 71 proteins were up-regulated more than 1.5 fold in insects infected with *E. coli* (Table 2). We classified them into 6 groups: immunity (4), cytoskeleton (3), DNA/RNA synthesis (13), metabolism (24), ATP/NAD(P)H binding (5), and others (22). In contrast, 30 proteins were down-regulated (IP/CP < 0.67) after the bacterial challenge (Table 3). They belong to three groups: immunity (9), metabolism (10), and others (11). Increases in the 67 immunity-unrelated proteins might reflect a higher demand for transcription, translation, and energy production after exposure

to the bacteria, decreases in the 21 immunity-unrelated ones suggest a down-regulation of other cellular processes.

Surprisingly, only four defense proteins (PGRP-LB, TEP2, SRPN10, E3 SUMO-protein ligase RanBP2) were up-regulated 1.5–2.2 fold (Table 2). Seven (Gram-negative bacteria-binding protein (GNBP), CLIPs B1, B8 & A7, SP219, GP71, MDL2) were down-regulated to 50–61% of the control levels ($p < 0.05$) (Table 3). Lysozyme-4 and proPO1–3 levels significantly reduced to 37, 34, 85, and 80% of the control, respectively. While the ratios were high for proPO2 and 3, the absolute amounts of reduction based on LFQs were 25- and 59-fold higher than that of proPO1 (Table S1). These observations in the larval hemolymph are drastically different from the transcriptome data of adults, which indicate that more immunity-related genes are up-regulated than down-regulated after an immune challenge (Aguilar et al., 2005). One possible explanation is that certain immune molecules (*e.g.* proPOs) have been heavily used during defense responses of the larvae but not fully replenished at 24 h after *E. coli* inoculation.

To further explain the disparity, we repeated the experiment using the same bacterial strain but carrying a GFP expression plasmid to test if the bacteria were still alive at 24 h after pricking, whether there was a difference in survival between the control and test groups, and what the dynamics of AMP gene expression was during the experiment. We found 80% of the PBS-pricked larvae survived whereas 63% of *E. coli*-pricked ones lived (Fig. 4A). This result suggests that *E. coli* pricking caused infection and infection led to more deaths than the buffer control did. In the other two groups, survival rates for PBS- and *E. coli*-injection were 83% and 66%, respectively, showing small variations between pricking and injection. The latter is a well-established infection model for adult mosquitos. Consistent with active infection, some of the bacteria were alive and mostly localized around the melanized wound site in the live larvae, as revealed by their movement during imaging (Fig. 4B). There was no fluorescent bacterium in the control larvae. In some dead larvae pricked with *E. coli*, GFP signals were distributed in various parts of the body, some near melanin tumors. The induced gambicin-1 expression peaked at 6 h in the *E. coli*-injected larvae and higher than those in the PBS-injected ones (Fig. 4C). The mRNA peak appeared at about 24 h in *E. coli*-pricked ones and, again, higher than those pricked with PBS. Similar expression patterns were observed for cecropin A. The gambicin and cecropin transcripts were hardly detected in the naïve larvae. Taken together, sterile and septic wounding both increased the AMP expression to various levels with different dynamics in larvae of this aquatic species. The survival of larvae was governed mainly by the presence/absence of *E. coli* in thousands and there was no clear difference between pricking and injection under the current conditions. The AMP transcript levels in *E. coli*-pricked larvae were only a few fold higher than PBS-pricked ones. In contrast to sterile wounding, the newly synthesized AMPs and other defense proteins (*e.g.* PPOs) were mostly consumed in the ongoing battle against *E. coli*, alive even at 24 h after inoculation. Based on the consistent results from these control experiments, we are confident that the proteomic data have faithfully reflected the complex changes in plasma proteins on the battleground. It would be interesting to compare hemolymph samples from naïve larvae and larvae pricked/injected with killed bacteria to confirm a consistent increase in mRNA and protein levels of defense molecules.

3.4. Immunity-related proteins

We have identified a total of 233 putative defense proteins in the larval hemolymph (Table 4), 166 of which are probably secreted. The percentage (71%), which is much higher than 29% (or 437) of the 1,521 other proteins, features the role of humoral responses. Besides, some defense proteins (*e.g.* proPOs) lack a signal peptide and may come from ruptured cells or by other unconventional paths (Kanost and Gorman, 2008; Nickel et al., 2010). We divide them into eight subgroups: 53 PRRs, 8 TEPs, 9 AMPs, 59 SPs, 38 SPHs, 17 SP inhibitors, 7 proPOs, and 42 others (Fig. 2B).

The PRRs include 15 LRR proteins, 9 CTLs, 8 GNBPs, 4 PGRPs, and 13 FBNs among others. APL1C and LRIM1 are critical determinants of the mosquito antimalarial resistance (Povelones et al., 2009) and are more abundant than the other LRR proteins, except for LRR-7059 and LRR-7060. We detected low CTL4 and no CTLMA2 in the larval plasma. These two lectins positively impact the parasite development in adult females (Osta et al., 2004). CTL8, PGRP-S3, Nimrod, and GNBP-B1, most abundant in their respective families, may be important for antibacterial immunity in this stage. TEP1, TEP4 and TEP15 existed at higher levels than the other TEPs and, along with LRR-7060, may participate in antiparasitic responses (Blandin et al., 2004; Li et al., 2013). We did not detect any TEP in the proteomic analysis of *M. sexta* larval hemolymph (He et al., 2016). In contrast, the mosquito TEPs account for about 1% of total LFQs.

The detection of 59 SPs and 38 SPHs in these samples (Table 4) indicates a contamination of the plasma samples by gut contents. To address this problem, we examined their domain structures and expression profiles and considered 14 SPs and 17 SPHs as midgut proteins (data not shown). The GPs (*i.e.* gut SPs) likely digest dietary proteins but the GPHs' function is unclear. It is common that catalytically inactive SPHs are expressed in the digestive tract (Cao et al., 2015; Lin et al., 2015; Zhao et al., 2010).

Most of the other 45 SPs and 21 SPHs may participate in processes unrelated to digestion, as additional structural modules are present in these proteins for interacting with other molecules and their expression is not as synchronized as the digestive enzymes (data not shown). For instance, SP208, SP213/GRAAL, SP217, SP219, CLIPB47, and CLIPD6 contain multiple regulatory domains; 13 CLIPBs, 7 CLIPCs, 2 CLIPDs and 3 CLIPes have a clip domain and a protease domain; 14 SPHs contain a clip domain and a protease-like domain (Table 4). CLIPs B4, B8, B13, B9, B10, B3, B5, and C4 were more abundant than the other clip-domain SPs. We detected low levels of three multi-domain SPHs (SPH220, CLIPA2 and CLIPE12) and found ten of the fifteen clip-domain SPHs were more abundant, representing 0.02–0.50% of the LFQ totals. Identification and quantification of the nondigestive SPs and SPHs in the larval plasma provided useful clues for studying a putative SP-SPH network that coordinates some of the defense mechanisms (Barillas-Mury, 2007).

Biochemical elucidation of the mosquito SP-SPH system poses a fierce challenge, since less than 0.5 µl of hemolymph can be collected from one adult. Identification of the 66 nondigestive SPs and SPHs in larval plasma hints at a similarly complex system that regulates immune responses in adult hemolymph (Volz et al., 2005 and 2006; Povelones et al., 2013). Another problem is that fifteen of the identified 97 SPs and SPHs had minor to

major flaws in their sequences (data not shown). Nonetheless, the proteome analysis at least validate them as authentic gene products. More notably, this study greatly confines the system exploration from nearly 350 SP-related genes in the genome (Christophides et al., 2002). Data on their levels should allow us to further focus on the abundant ones with complex domain structures as candidates of the immune SP-SPH pathways.

We have identified 14 serpins (including two variants of SRPN4 and SRPN10). SRPN2, 3, 4, 7, 9, 10A, 11, 12 and 16 were relatively abundant (0.03–0.30%). Biological functions of several *A. gambiae* serpins have been characterized in the adult, which may regulate proPO activation (Danielli et al., 2005; Abraham et al., 2005; Michel et al., 2006; Suwanchaichinda and Kanost, 2009). Together, nondigestive SPs, SPHs and serpins are speculated to constitute an enzyme-cofactor-inhibitor system to coordinate humoral and cellular immunity, as demonstrated in other insects (Jiang et al., 2010; Park et al., 2010; Kanost and Jiang, 2015).

As putative substrates of clip-domain SPs, seven of the nine proPOs were detected at different levels in the larval plasma. ProPO2 and 3 were most abundant (1.2% and 2.0% of the total LFQs), followed by proPO6, 8, 4 (0.04–0.09%), and then proPO1 and 7 (0.01%). Based on their sheer amounts, we suggest that proPO2 and proPO3 are more important in antimicrobial responses of the larvae. Other immune effectors such as defensins, gambicin and lysozymes existed at 0.01%. The high basal level of transferrin (0.42%) had a 1.61-fold increase to 0.79% ($p = 0.055$) in plasma.

3.5. Gel distribution of the defense proteins

SDS-treated proteins should migrate in reducing gel to positions corresponding to their theoretical masses (M_r 's). However, post-translational modifications (e.g. glycosylation, proteolysis, covalent crosslinking) alters M_r 's and such changes are common during insect defense responses (He et al., 2016). Thus, to characterize the extent to which this might occur in the mosquito as a result of immune challenge, we examined distributions of the 233 defense proteins in the gel slices and identified fifteen showing major M_r decreases consistent with proteolytic processing (Table 5). While nature of these cleavages (e.g. sites and accountable enzymes) is unclear, these altered mobilities are notable because of their potential importance.

Due to the importance of CLIPs and their inhibitory regulation by serpins, we closely examined the gel slices containing them, and found eleven clip-domain SPs and eight serpins in gel slice 6 (70–80 kDa). Since these proteins are typically 45–55 kDa, the observed mobility suggests the formation of SDS-stable serpin-protease complexes (Table 6). Such acyl-enzyme complexes usually contain a 40–45 kDa serpin fragment and a 30 kDa SP catalytic domain, consistent with the previous observations (An et al., 2011; Tong et al., 2005). Interestingly, SRPN1, 2, 7, 8, 10B, CLIPs B1, B2, B4, B5, B8, B11, C2, C4, C9 and D6 were also detected at a higher M_r range (80–350 kDa) (Table 7). This result suggests the proteins are parts of larger covalent complexes, which may contain other defense proteins.

Our previous study of *M. sexta* larval hemolymph (He et al., 2016) revealed that, in general, abundant proteins are distributed among more gel slices than scarce ones do. As shown in

Fig. 5, the same correlation was observed in this project and applied in studying protein spreading in the gel slices. We found, like the CLIPs and serpins, 24 other defense proteins displayed substantial decreases in gel mobility (Table 7). They were GNBPA1, TEP6, FBN, 5 LRRs, Trynity, 6 SPs, 2 SPHs, and 7 proPOs. Since most of them were not abundant, their low-mobility species did not seem to be artifacts. We interpret these and some other slow migrating proteins (Table 6) as components of high M_r immune complexes. Covalent crosslinking by PO-generated compounds may have stabilized them under the reducing and denaturing conditions of SDS-PAGE; such formation of PO-generated macromolecular complexes has been previously observed in several studies (Yu et al., 2003; Zou and Jiang, 2005). Recently the role of POs in this process was clearly demonstrated (Clark and Strand, 2013). We found that 32 of the 66 proteins (10–80 kDa) identified in >80 kDa positions were immunity-related, at a ratio twice as high as 171 of the 702 total proteins identified in the *M. sexta* plasma proteome (He et al., 2016). Based on these, we are generally confident about the model of formation of high M_r complexes enriched with defense proteins in *A. gambiae*.

4. Conclusions

In this study, we have categorized 1,756 proteins in hemolymph of *A. gambiae* larvae, which represents a major increase in the hemolymph proteome coverage. In spite of the contaminating intracellular and gut enzymes, identification of *bona fide* plasma proteins provides an overview of the physiological functions of the mosquito larval hemolymph. Future studies of adult hemolymph samples should yield insights into composition changes in plasma proteins, which may directly interact with malaria parasites. There were no global differences between the control and induced samples and, instead of detecting increases in defense proteins, we observed a few more cases of relative decrease, possibly as a result of the imbalance between higher protein consumption and higher protein synthesis after bacterial infection than after sterile wounding. The active infection after pricking with *E. coli*, similar survival rates between pricking and injection, and transcription up-regulation of the AMP genes supported that the observed proteomic changes were not artifacts. The comparison of theoretical and observed M_r 's allowed us to detect possible posttranslational modifications of proteins (e.g. proteolysis and serpin-protease complex formation), occurring in the control and induced samples. Assembling and crosslinking of macromolecular complexes appears to be a common feature of innate immunity, now supported at the level of proteome in an important vector of human malaria.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by National Institutes of Health Grants AI112662 and GM58634 (to HJ). We would like to express our gratitude to Dr. Li Ma at National Institute for Microbial Forensics & Food and Agricultural Biosecurity, who kindly provided the GFP expressing plasmid. Mass spectrometry analyses were performed in the DNA/Protein Resource Facility at Oklahoma State University, wherein the OrbitrapXL instrument was supported by the NSF MRI and EPSCoR programs (award DBI/0722494). This article was approved for publication by the Director of the Oklahoma Agricultural Experiment Station and supported in part under project OKL02450.

Abbreviations

CP	control plasma or cell-free hemolymph from sterilely pricked larvae
IP	induced plasma from <i>E. coli</i> pricked larvae
AMP	antimicrobial protein
APL1C	<i>Anopheles plasmodium</i> -responsive LRR protein-1C
CTL	C-type lectin
FBN	fibrinogen
GFP	green fluorescent protein
GNBP	Gram-negative bacteria-binding protein
IAP	inhibitor of apoptosis
iBAQ	intensity-based absolute quantification
LDLp	low density lipophorin
LFQ	label-free quantification intensity
LRIM1	Leu-rich immune molecule-1
LRR	leucine-rich repeat
MDL	myeloid differentiation factor-2 (MD2) like protein
MS	mass spectrometry
PGRP	peptidoglycan recognition protein
PRR	pattern recognition receptor
PO and proPO	phenoloxidase and its precursor
SP and SPH	serine protease and its non-catalytic homolog
CLIP	clip-domain SP or SPH, SRPN, serpin
TEP	thioester protein or its homolog

References

- Abraham EG, Pinto SB, Ghosh A, Vanlandingham DL, Budd A, Higgs S, Kafatos FC, Jacobs-Lorena M, Michel K. An immune-responsive serpin, SRPN6, mediates mosquito defense against malaria parasites. *Proc Natl Acad Sci USA*. 2005; 102:16327–16332. [PubMed: 16260729]
- Aguilar R, Jedlicka AE, Mintz M, Mahairaki V, Scott AL, Dimopoulos G. Global gene expression analysis of *Anopheles gambiae* responses to microbial challenge. *Insect Biochem Mol Biol*. 2005; 35:709–719. [PubMed: 15894188]

- Ahrné E, Molzahn L, Glatter T, Schmidt A. Critical assessment of proteome-wide label-free absolute abundance estimation strategies. *Proteomics*. 2013; 13:2567–2578. [PubMed: 23794183]
- Amenya DA, Chou W, Li J, Yan G, Gershon PD, James AA, Marinotti O. Proteomics reveals novel components of the *Anopheles gambiae* eggshell. *J Insect Physiol*. 2010; 56:1414–1419. [PubMed: 20433845]
- An C, Budd A, Kanost MR, Michel K. Characterization of a regulatory unit that controls melanization and affects longevity of mosquitoes. *Cell Mol Life Sci*. 2011; 68:1929–1939. [PubMed: 20953892]
- Asara JM, Christofk HR, Freemark LM, Cantley LC. A label-free quantification method by MS/MS TIC compared to SILAC and spectral counting in a proteomics screen. *Proteomics*. 2008; 8:994–999. [PubMed: 18324724]
- Barillas-Mury C. CLIP proteases and *Plasmodium* melanization in *Anopheles gambiae*. *Trends Parasitol*. 2007; 23:297–299. [PubMed: 17512801]
- Benedict, MQ. Care and maintenance of Anopheline mosquito colonies. In: Crampton, JM, Beard, CB., Louis, C., editors. *Molecular Biology of Insect Disease Vectors: A Methods Manual*. Chapman & Hall; New York: 1997. p. 3-12.
- Blandin S, Moita LF, Kocher T, Wilm M, Kafatos FC, Levashina EA. Reverse genetics in the mosquito *Anopheles gambiae*: targeted disruption of the defensin gene. *EMBO Rep*. 2002; 3:852–856. [PubMed: 12189180]
- Blandin S, Shiao SH, Moita LF, Janse CJ, Waters AP, Kafatos FC, Levashina EA. Complement-like protein TEPI is a determinant of vectorial capacity in the malaria vector *Anopheles gambiae*. *Cell*. 2004; 116:661–670. [PubMed: 15006349]
- Cao X, He Y, Hu Y, Zhang X, Wang Y, Zou Z, Chen Y, Blissard GW, Kanost MR, Jiang H. Sequence conservation, phylogenetic relationships, and expression profiles of nondigestive serine proteases and serine protease homologs in *Manduca sexta*. *Insect Biochem Mol Biol*. 2015; 62:51–63. [PubMed: 25530503]
- Christophides GK, Zdobnov E, Barillas-Mury C, Birney E, Blandin S, et al. Immunity-related genes and gene families in *Anopheles gambiae*. *Science*. 2002; 298:159–165. [PubMed: 12364793]
- Clark KD, Strand MR. Hemolymph melanization in the silkworm *Bombyx mori* involves formation of a high molecular mass complex that metabolizes tyrosine. *J Biol Chem*. 2013; 288:14476–14487. [PubMed: 23553628]
- Clayton AM, Dong Y, Dimopoulos G. The *Anopheles* innate immune system in the defense against malaria infection. *J Innate Immun*. 2014; 6:169–181. [PubMed: 23988482]
- Coggins SA, Estévez-Lao TY, Hillyer JF. Increased survivorship following bacterial infection by the mosquito *Aedes aegypti* as compared to *Anopheles gambiae* correlates with increased transcriptional induction of antimicrobial peptides. *Dev Comp Immunol*. 2012; 37:390–401. [PubMed: 22326457]
- Cox J, Hein MY, Lubner CA, Paron I, Nagaraj N, Mann M. Accurate proteome-wide label-free quantification by delayed normalization and maximal peptide ratio extraction, termed MaxLFQ. *Mol Cell Proteomics*. 2014; 13:2513–2526. [PubMed: 24942700]
- Cox J, Mann M. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. *Nat Biotechnol*. 2008; 26:1367–1372. [PubMed: 19029910]
- Danielli A, Barillas-Mury C, Kumar S, Kafatos FC, Loukeris TG. Overexpression and altered nucleocytoplasmic distribution of Anopheles ovalbumin-like SRPN10 serpins in *Plasmodium*-infected midgut cells. *Cell Microbiol*. 2005; 7:181–190. [PubMed: 15659062]
- Dinglasan RR, Devenport M, Florens L, Johnson JR, McHugh CA, Donnelly-Doman M, Carucci DJ, Yates JR 3rd, Jacobs-Lorena M. The *Anopheles gambiae* adult midgut peritrophic matrix proteome. *Insect Biochem Mol Biol*. 2009; 39:125–134. [PubMed: 19038338]
- Dong Y, Dimopoulos G. *Anopheles* fibrinogen-related proteins provide expanded pattern recognition capacity against bacteria and malaria parasites. *J Biol Chem*. 2009; 284:9835–9844. [PubMed: 19193639]
- Francischetti IM, Valenzuela JG, Pham VM, Garfield MK, Ribeiro JM. Toward a catalog for the transcripts and proteins (sialome) from the salivary gland of the malaria vector *Anopheles gambiae*. *J Exp Biol*. 2002; 205:2429–2451. [PubMed: 12124367]

- He Y, Cao X, Zhang S, Rogers J, Hartson S, Jiang H. Changes in the plasma proteome of *Manduca sexta* larvae in relation to the transcriptome variations after an immune challenge: evidence for high molecular weight immune complex formation. *Mol Cell Proteomics*. 2016; 15:1176–1187. [PubMed: 26811355]
- Holt RA, Subramanian GM, Halpern A, Sutton GG, Charlab R, et al. The genome sequence of the malaria mosquito *Anopheles gambiae*. *Science*. 2002; 298:129–149. [PubMed: 12364791]
- Jiang H, Vilcinskis A, Kanost MR. Immunity in lepidopteran insects. *Adv Exp Med Biol*. 2010; 708:181–204. [PubMed: 21528699]
- Kalume DE, Okulate M, Zhong J, Reddy R, Suresh S, Deshpande N, Kumar N, Pandey A. A proteomic analysis of salivary glands of female *Anopheles gambiae* mosquito. *Proteomics*. 2005; 5:3765–3777. [PubMed: 16127729]
- Kanost, MR., Gorman, MJ. Phenoloxidases in insect immunity. In: Beckage, NE., editor. *Insect Immunology*. Elsevier; San Diego: 2008. p. 69–96.
- Kanost MR, Jiang H. Clip-domain serine proteases as immune factors in insect hemolymph. *Curr Opin Insect Sci*. 2015; 11:47–55. [PubMed: 26688791]
- Karlsson C, Korayem AM, Scherfer C, Loseva O, Dushay MS, Theopold U. Proteomic analysis of the *Drosophila* larval hemolymph clot. *J Biol Chem*. 2004; 279:52033–52041. [PubMed: 15466469]
- Larade K, Storey KB. Accumulation and translation of ferritin heavy chain transcripts following anoxia exposure in a marine invertebrate. *J Exp Biol*. 2004; 207:1353–1360. [PubMed: 15010486]
- Lefevre T, Thomas F, Schwartz A, Levashina E, Blandin S, Brizard JP, Le Bourlignu L, Demetree E, Renaud F, Biron DG. Malaria *Plasmodium* agent induces alteration in the head proteome of their *Anopheles* mosquito host. *Proteomics*. 2007; 7:1908–1915. [PubMed: 17464940]
- Levashina EA, Moita LF, Blandin S, Vriend G, Lagueux M, Kafatos FC. Conserved role of a complement-like protein in phagocytosis revealed by dsRNA knockout in cultured cells of the mosquito, *Anopheles gambiae*. *Cell*. 2001; 104:709–718. [PubMed: 11257225]
- Li J, Wang X, Zhang G, Githure JJ, Yan G, James AA. Genome-block expression-assisted association studies discover malaria resistance genes in *Anopheles gambiae*. *Proc Natl Acad Sci USA*. 2013; 110:20675–20680. [PubMed: 24297936]
- Lin H, Xia X, Yu L, Vasseur L, Gurr GM, Yao F, Yang G, You M. Genome-wide identification and expression profiling of serine proteases and homologs in the diamondback moth, *Plutella xylostella* (L.). *BMC Genomics*. 2015; 16:1054. [PubMed: 26653876]
- Michel K, Suwanchaichinda C, Morlais I, Lambrechts L, Cohuet A, Awono-Ambene PH, Simard F, Fontenille D, Kanost MR, Kafatos FC. Increased melanizing activity in *Anopheles gambiae* does not affect development of *Plasmodium falciparum*. *Proc Natl Acad Sci USA*. 2006; 103:6858–6863.
- Nickel W. Pathways of unconventional protein secretion. *Curr Opin Biotechnol*. 2010; 21:621–626. [PubMed: 20637599]
- Ong DS, Wang L, Zhu Y, Ho B, Ding JL. The response of ferritin to LPS and acute phase of *Pseudomonas* infection. *J Endotoxin Res*. 2005; 11:267–280. [PubMed: 16262999]
- Osta MA, Christophides GK, Kafatos FC. Effects of mosquito genes on *Plasmodium* development. *Science*. 2004; 303:2030–2032. [PubMed: 15044804]
- Park JW, Kim CH, Rui J, Park KH, Ryu KH, Chai JH, Hwang HO, Kurokawa K, Ha NC, Söderhäll I, Söderhäll K, Lee BL. Beetle immunity. *Adv Exp Med Biol*. 2010; 708:163–180. [PubMed: 21528698]
- Paskewitz SM, Shi L. The hemolymph proteome of *Anopheles gambiae*. *Insect Biochem Mol Biol*. 2005; 35:815–824. [PubMed: 15944078]
- Pesch YY, Riedel D, Patil KR, Loch G, Behr M. Chitinases and imaginal disc growth factors organize the extracellular matrix formation at barrier tissues in insects. *Sci Rep*. 2016; 6:18340. [PubMed: 26838602]
- Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods*. 2011; 8:785–786. [PubMed: 21959131]
- Povelones M, Bhagavatula L, Yassine H, Tan LA, Upton LM, Osta MA, Christophides GK. The CLIP-domain serine protease homolog SPCLIP1 regulates complement recruitment to microbial surfaces

- in the malaria mosquito *Anopheles gambiae*. PLoS Pathog. 2013; 9:e1003623. [PubMed: 24039584]
- Povelones M, Waterhouse RM, Kafatos FC, Christophides GK. Leucine-rich repeat protein complex activates mosquito complement in defense against *Plasmodium* parasites. Science. 2009; 324:258–261. [PubMed: 19264986]
- Ramphul UN, Garver LS, Molina-Cruz A, Canepa GE, Barillas-Mury C. *Plasmodium falciparum* evades mosquito immunity by disrupting JNK-mediated apoptosis of invaded midgut cells. Proc Natl Acad Sci USA. 2015; 112:1273–1280. [PubMed: 25552553]
- Rieu I, Powers SJ. Real-time quantitative RT-PCR: design, calculations, and statistics. Plant Cell. 2009; 21:1031–1033. [PubMed: 19395682]
- Shen HB, Chou KC. Signal-3L: A 3-layer approach for predicting signal peptides. Biochem Biophys Res Commun. 2007; 363:297–303. [PubMed: 17880924]
- Suwanchaichinda C, Kanost MR. The serpin gene family in *Anopheles gambiae*. Gene. 2009; 442:47–54. [PubMed: 19394412]
- Volz J, Müller HM, Zdanowicz A, Kafatos FC, Osta MA. A genetic module regulates the melanization response of *Anopheles* to *Plasmodium*. Cell Microbiol. 2006; 8:1392–1405. [PubMed: 16922859]
- Volz J, Osta MA, Kafatos FC, Müller HM. The roles of two clip domain serine proteases in innate immune responses of the malaria vector *Anopheles gambiae*. J Biol Chem. 2005; 280:40161–40168. [PubMed: 16188883]
- WHO. World Malaria Report 2015. 2015.
- Yang F, Wang Y, He Y, Jiang H. In search of a function of *Manduca sexta* hemolymph protease-1 in the innate immune system. Insect Biochem Mol Biol. 2016; 76:1–10. [PubMed: 27343384]
- Yassine H, Osta MA. *Anopheles gambiae* innate immunity. Cell Microbiol. 2010; 12:1–9. [PubMed: 19804484]
- Yu XQ, Jiang H, Wang Y, Kanost MR. Nonproteolytic serine proteinase homologs are involved in prophenoloxidase activation in the tobacco hornworm, *Manduca sexta*. Insect Biochem Mol Biol. 2003; 33:197–208. [PubMed: 12535678]
- Zhao P, Wang GH, Dong ZM, Duan J, Xu PZ, Cheng TC, Xiang ZH, Xia QY. Genome-wide identification and expression analysis of serine proteases and homologs in the silkworm *Bombyx mori*. BMC Genomics. 2010; 11:405. [PubMed: 20576138]
- Zou Z, Jiang H. *Manduca sexta* serpin-6 regulates immune serine proteinases PAP-3 and HP8: cDNA cloning, protein expression, inhibition kinetics, and function elucidation. J Biol Chem. 2005; 280:14341–14348. [PubMed: 15691825]

Identify 1,756 proteins in the larval hemolymph of *Anopheles gambiae*
Indicate storage, transport, and immunity as major functions of the hemolymph
Detect 233 proteins for immune recognition, signaling, regulation, and execution
Suggest immune complex formation and other dynamic changes during infection

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Fig. 1. Schematic overview of the sample treatment and data analysis

(A) Hemolymph from the larvae challenged with *E. coli* and water was separately pooled (20–25 insects per group) to generate four induced (IP) and four control (CP) plasma samples (see Materials and Methods). The eight plasma samples were separated by SDS-PAGE on a 4–15% gradient gel that was later cut into 96 pieces. After in-gel trypsinolysis, peptides were separately extracted and analyzed by LC-MS/MS. Protein identification and quantification were performed as described in Materials and Methods. Sizes and positions of the M_r markers (lane M) are labeled on the left, while gel slice numbers (#'s) and corresponding M_r ranges of the gel slices are indicated on the right. (B) Distributions of the number of proteins identified in each gel slice. Black = CP; pink = IP.

Fig. 2. Categorization of the total (1,756) and putative defense (233) proteins

(A) The 1,756 proteins classified into nine arbitrary groups with their names and number of proteins in each group indicated. (B) The 233 defense proteins were subdivided into seven groups: pattern recognition receptor (PRR), serine protease (SP), serine protease homolog (SPH), serpin, antimicrobial protein (AMP), prophenoloxidase (proPO), thioester protein (TEP), and others, with their counts marked.

Fig. 3. Distributions of protein numbers (A) and LFQ percentages (B) in various LFQ ranges
 Within each range, sums of LFQs of the proteins identified in CP (black) and IP (pink) are used to calculate their percentages in the grand totals ($\Sigma\text{LFQ}_{\text{CP}}$ and $\Sigma\text{LFQ}_{\text{IP}}$), respectively.

Fig. 4. Effects of sterile and septic wounding and injection on *A. gambiae* larvae
(**A**) Survival statistics; (**B**) Localization of GFP-labeled bacteria in the living and dead larvae 24 h after pricking with PBS or *E. coli*. Blurring of the signals was caused by their movements during imaging. *upper*: bright field; *lower*: fluorescence. Due to movement of *E. coli* cells in the culture (*right*), their positions in the bright field no longer corresponded to those in the fluorescence microscope image obtained a few minutes later. (**C**) Expression profiles of gambicin-1 (*upper*) and cecropin A (*lower*) genes after PBS and *E. coli* pricking and injection. Two pools of naïve larvae (10 insect per group) were as used as negative controls. The AMP mRNA levels relative to *A. gambiae* rpS7 are shown for each sample as mean ± standard deviation (n = 3). Two biological replicates were analyzed for all these samples.

Fig. 5. Relationships of protein abundances and number of the gel slices they were identified Relative abundances, log₂LFQ, for proteins identified in CP (open circles, panel **A**) or IP (pink dots, panel **B**) and numbers of gel slices they were detected were plotted. The number of proteins within each group is marked above the data series.

Table 1

A list of 25 most abundant proteins

ID	Name	M _r (kDa)	CP%*	IP%*	IP/CP
AGAP001826-PA	Apolipoprotein-I&II	371	2.50	1.10	0.38
AGAP013365-PA	Apolipoprotein-III	22	7.07	7.16	0.86
AGAP008369-PA	Vitellogenin	170	4.71	4.06	0.74
AGAP008054-PD	Chemosensory protein	15	1.68	1.23	0.63
AGAP000278-PA	Odorant binding protein-9	16	2.98	1.83	0.52
AGAP002464-PA	Ferritin G subunit	26	2.05	1.45	0.6
AGAP002465-PA	Ferritin heavy chain	25	1.13	0.8	0.61
AGAP001659-PA	Hexamerin	84	7.99	6.41	0.69
AGAP001657-PA	Hexamerin	84	9.10	6.36	0.6
AGAP005768-PA	Hexamerin	82	1.15	2.75	2.04
AGAP005766-PA	Hexamerin A	83	2.71	3.70	1.17
AGAP001345-PA	Hexamerin A	83	6.8	5.28	0.66
AGAP010657-PA	Larval serum protein Iβ	24	1.26	1.68	1.14
AGAP008060-PA	Imaginal disc growth factor	48	0.88	0.94	0.91
AGAP006258-PA	proPO2	78	1.19	1.18	0.85
AGAP004975-PA	proPO3	79	2.10	1.96	0.8
AGAP008364-PA	TEP15	164	1.66	1.41	0.73
AGAP011369-PA	Gelsolin	43	0.78	0.49	0.54
AGAP007059-PA	LRR-7059	124	0.72	0.67	0.79
AGAP000651-PC	Actin	42	0.64	0.84	1.12
AGAP005627-PD	Creatine kinase	40	1.09	1.37	1.07
AGAP002564-PE	Fructose-1,6BP aldolase, class I	39	0.97	1.43	1.26
AGAP009623-PA	GAP dehydrogenase	36	0.77	0.72	0.8
AGAP013400-PA	Fatty acid-binding protein	15	0.67	1.34	1.72
AGAP012057-PA	RNA polymerase-associated RTF1	88	0.93	0.49	0.45
Σ			63.53	56.65	

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

* : average percentage of each protein's LFQs in the total LFQ protein intensity for the four CP or IP samples. Proteins shaded light blue are intracellular proteins. The IP/CP ratios shaded green are statistically significant ($p < 0.05$).

Table 2

A list of 71 up-regulated proteins ($I/C > 1.5, p < 0.05$)

Group	ID	Name	RA _C [*]	RA _I [*]	IP/CP
Immunity	AGAP002982-PA	E3 SUMO-protein ligase RanBP2	14	32	2.00
	AGAP001212-PB	PGRP-LB	27	72	2.24
	AGAP005246-PD	SRPN10	435	983	1.93
	AGAP008366-PA	TEP2	4	6	1.51
Cytoskeleton/motor	AGAP012185-PA	Erythrocyte membrane protein band 4.1	52	109	1.80
	AGAP001315-PE	Microtubule-associated protein 7 family	11	26	2.10
	AGAP001799-PA	Tropomyosin 1	79	141	1.52
DNA/RNA synthesis	AGAP002945-PA	Bifunctional glutamyl/prolyl-tRNA synthetase	14	44	2.61
	AGAP006125-PA	Density-regulated protein	6	13	1.78
	AGAP001883-PA	ELAV-like 1	4	10	2.30
	AGAP004725-PA	Eukaryotic translation initiation factor 3C	11	27	2.16
	AGAP003486-PA	General transcriptional corepressor trfa	13	28	1.9
	AGAP005015-PA	Heterogeneous nuclear ribonucleoprotein K	5	11	1.75
	AGAP007299-PA	Importin-7	23	43	1.64
	AGAP012013-PA	Nuclear factor of activated T-cells 5	6	13	1.69
	AGAP002351-PA	Nuclear pore complex protein Nup98-Nup96	2	10	4.67
	AGAP002654-PB	Poly(A)-binding protein 1	9	20	1.87
Metabolism	AGAP010553-PA	Poly(U)-binding-splicing factor PUF60	2	4	1.63
	AGAP002655-PA	RNA binding protein	2	7	2.88
	AGAP010640-PA	Translation initiation factor	186	367	1.68
	AGAP006227-PA	Alpha esterase	0	7	∞
	AGAP004236-PA	Beta-lactamase-like protein 2 homolog	0	2	∞
	AGAP009405-PA	CPAP3-E	8	14	1.59
	AGAP005627-PE	Creatine kinase	1	6	10.27
	AGAP003124-PA	Dihydropyrimidinase	2	6	2.47
	AGAP000513-PB	Dipeptidase E	47	99	1.81
	AGAP013400-PA	Fatty acid-binding protein	6664	13411	1.72

Group	ID	Name	RA _C [*]	RA _I [*]	IP/CP
ATP/NAD(P)H binding	AGAP004071-PB	Fimbrin	59	113	1.63
	AGAP006670-PA	Gamma-glutamyl hydrolase	109	244	1.91
	AGAP003077-PB	Glutamyl aminopeptidase	80	161	1.72
	AGAP004383-PA	GSTD10	9	32	3.07
	AGAP003257-PA	GSTU2	20	46	2.03
	AGAP006353-PA	Histidine triad nucleotide binding protein 1	156	289	1.58
	AGAP004747-PA	Ion binding and proteolysis	60	111	1.58
	AGAP012008-PA	Na ⁺ /H ⁺ exchange regulatory cofactor NHE-RF1	20	48	2.05
	AGAP000500-PD	NADPH-ferrihemoprotein reductase	0	1	∞
	AGAP009172-PA	Prolyl oligopeptidase	101	185	1.56
	AGAP004758-PB	Proteasomal ubiquitin receptor adrm1 homolog	65	136	1.80
	AGAP006171-PA	Protein phosphatase	0	6	∞
	AGAP004093-PA	Sterol carrier protein-2	910	2399	2.25
	AGAP003052-PA	Tetrapeptide repeat-containing protein α	116	219	1.62
	AGAP011872-PA	Ubiquitin-activating enzyme E1	256	481	1.61
	AGAP009841-PA	UBX domain-containing protein 1	15	30	1.67
Other	AGAP009648-PA	Ureidoimidazole decarboxylase	219	408	1.59
	AGAP005981-PA	DnaJ homolog subfamily A	10	23	1.87
	AGAP001690-PA	Regulating synaptic exocytosis protein 2	0	4	∞
	AGAP003153-PD	V-type proton ATPase catalytic subunit A	591	1315	1.90
	AGAP009486-PA	V-type proton transporting ATPase 54 kDa	58	135	1.97
	AGAP002884-PA	V-type proton transporting ATPase subunit B	348	718	1.76
	AGAP001467-PA	AGAP001467-PA	9	20	1.87
	AGAP013060-PA	AGAP013060-PA	1380	4381	2.71
	AGAP011762-PA	BAG domain-containing protein Samui	10	19	1.59
	AGAP010557-PA	B-cell receptor-associated protein 31	13	26	1.66
	AGAP005316-PA	Charged multivesicular body protein 4	36	71	1.71
	AGAP010251-PA	Coatomer protein complex alpha subunit	4	10	2.37
	AGAP010900-PA	Cuticular protein 1 from fifty-one aa family	43	87	1.72
	AGAP006103-PA	Farnesic acid o-methyl transferase-like	81	173	1.81

Group	ID	Name	RA _C [*]	RA _i [*]	IP/CP
	AGAP009738-PA	Glutaredoxin	93	181	1.66
	AGAP000941-PB	Heat shock protein beta-1 isoform x2	182	376	1.77
	AGAP000941-PA	Heat shock protein beta-1 isoform x2	15	31	1.76
	AGAP007310-PA	Klaroid	5	10	1.59
	AGAP005291-PA	Lupus la ribonucleoprotein	44	84	1.63
	AGAP003238-PC	N-myc downstream regulated protein	49	92	1.62
	AGAP005369-PA	NOLC1-like isoform x2	4	14	3.27
	AGAP008747-PA	Nsp1p	5	12	2.02
	AGAP008046-PA	PACSLIN2	1	5	3.91
	AGAP004310-PA	Perq amino acid-rich protein 2	8	14	1.59
	AGAP012746-PA	Phyhd1 protein	17	32	1.58
	AGAP004520-PA	Ran-binding protein 3	8	17	1.83
	AGAP004273-PB	Synapse-associated protein	1	4	2.96
	AGAP00626-PA	Vesicle-associated membrane protein B	10	19	1.67

* : Relative abundance (RA_C or RA_i) is defined as each protein's LFQ × 1,000,000 ÷ total LFQs for the CP or IP samples.

Table 3

A list of 30 down-regulated proteins ($I/C < 0.67$, $p < 0.05$)

Group	IDs	Names	RA _C [*]	RA _I [*]	IP/CP
Immunity	AGAP009110-PA	GNBP	302	185	0.52
	AGAP005663-PA	SP-5663	394	280	0.61
	AGAP011792-PA	CLIPA7-like	1548	944	0.52
	AGAP003251-PA	CLIPB1	210	123	0.50
	AGAP003057-PA	CLIPB8	1012	620	0.52
	AGAP005072-PA	SP-IgG-2LamD	35	22	0.53
	AGAP007385-PA	Lysozyme 4 (c-type)	9	4	0.37
	AGAP002857-PB	MDL2	58	34	0.50
	AGAP002825-PA	proPO1	110	43	0.34
Metabolism	AGAP003490-PA	Alanine-glyoxylate aminotransferase	15	11	0.64
	AGAP008783-PA	Arginase	12	8	0.57
	AGAP002465-PA	Ferritin heavy chain	11332	8040	0.61
	AGAP008798-PA	Guanine nucleotide exchange factor MSS4	5	1	0.16
	AGAP007237-PA	Heme peroxidase	11	0	0.00
	AGAP009033-PA	Heme peroxidase	272	178	0.56
	AGAP001826-PA	Apolipoprotein-I&II	24991	11018	0.38
	AGAP004654-PA	Phosphoadenylate 3'-nucleotidase	9	4	0.39
	AGAP000439-PA	Tetrahydrobiopterin dehydratase	187	139	0.63
	AGAP008064-PA	Uroporphyrinogen-III synthase	6	5	0.64
Other	AGAP010846-PA	AGAP010846	18	6	0.29
	AGAP028095-PC	AGAP028095	42	23	0.47
	AGAP004108-PB	Amalgam	311	213	0.59
	AGAP008052-PA	Chemosensory protein	1072	760	0.61
	AGAP002822-PA	Condensin-2 complex subunit H2	12	2	0.16
	AGAP008013-PA	Flaggrin-2 isoform x1	5122	3798	0.63
	AGAP001768-PB	Gamma-interferon-inducible protein IP-30	37	21	0.49
	AGAP001657-PA	Hexamerin	90951	63631	0.60

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Group	IDs	Names	RA _C [*]	RA _I [*]	IP/CP
	AGAP001127-PA	P37NB protein	16	8	0.42
	AGAP012057-PA	RNA polymerase-associated protein RTF1	9333	4944	0.45
	AGAP001989-PA	Secreted salivary gland protein	1525	737	0.41

* : Relative abundance (RA_C or RA_I) is defined as each protein' s LFQs × 1000,000 ÷ total LFQs for the CP or IP samples.

Table 4

A list of 233 defense proteins

Group	ID	Name	M_r (kDa)	RA_C^*	RA_I^*	IP/CP	p-value
PRR	AGAP007036-PA	APL1A	49.4	15	0	0	0.356
	AGAP007035-PA	APL1B	63.9	26	22	0.74	0.342
	AGAP007033-PA	APL1C	82.4	1577	1018	0.55	0.137
	AGAP004811-PA	CTL1	21.8	31	29	0.79	0.402
	AGAP004810-PA	CTL3	20.8	213	179	0.72	0.276
	AGAP005335-PA	CTL4	19.8	13	4	0.28	0.103
	AGAP003625-PA	CTL8	21.5	1160	1345	0.99	0.981
	AGAP006430-PB	CTL-GA2	24.9	13	21	1.43	0.396
	AGAP010193-PA	CTL-GA3	27	142	170	1.02	0.929
	AGAP007412-PA	CTL-MA1	20.1	21	14	0.57	0.097
	AGAP007411-PA	CTL-MA3	19.4	119	94	0.68	0.075
	AGAP002911-PA	CTL-MA9	17.5	2	2	0.89	0.939
	AGAP010021-PA	Dumpy	172.4	13	71	4.71	0.414
	AGAP010024-PA	Dumpy	345.1	26	115	3.75	0.411
	AGAP003027-PA	Dumpy-like	43.6	11	22	1.63	0.31
	AGAP009106-PA	GNBP	32.1	249	155	0.53	0.064
	AGAP009110-PA	GNBP	42	302	185	0.52	0.004
	AGAP009146-PA	GNBP	33.5	122	47	0.33	0.084
	AGAP006761-PA	GNBP-A1	55.7	9	7	0.67	0.176
	AGAP004455-PA	GNBP-B1	44.1	5003	4409	0.75	0.338
	AGAP002798-PA	GNBP-B2	43.7	211	265	1.07	0.829
	AGAP002799-PA	GNBP-B3	43.1	0	5	146.34	0.356
	AGAP002796-PA	GNBP-B4	46.7	21	41	1.63	0.232
	AGAP006327-PA	LRIM (short)	39.6	89	74	0.71	0.095
	AGAP006348-PA	LRIM1	57.3	2032	1324	0.56	0.138
	AGAP007039-PA	LRIM4	59.9	21	12	0.46	0.172
	AGAP005693-PA	LRIM17	48.8	1784	1307	0.63	0.081
	AGAP006644-PA	LRRP	77	4	2	0.5	0.359

Group	ID	Name	M_r (kDa)	RA_C^*	RA_I^*	IP/CP	p-value
	AGAP011503-PA	LRRP	32	29	27	0.79	0.392
	AGAP005962-PA	LRRP shoc-2	90.7	49	46	0.79	0.242
	AGAP004832-PA	LRRP-1	117.8	875	719	0.7	0.03
	AGAP003878-PA	LRRP-15	63.2	59	93	1.34	0.013
	AGAP007030-PA	LRRP-7030	115.4	1	0	0	0.356
	AGAP007059-PA	LRRP-7059	124	7232	6670	0.79	0.106
	AGAP007060-PA	LRRP-7060	132.9	4086	3653	0.76	0.251
	AGAP009762-PA	Nimrod	141.6	614	545	0.76	0.497
	AGAP001212-PB	PGRP-LB	23.3	27	72	2.24	0.04
	AGAP000536-PA	PGRP-S1	22.4	51	34	0.58	0.078
	AGAP006343-PA	PGRP-S2	20	6	0	0	0.356
	AGAP006342-PA	PGRP-S3	20	290	334	0.98	0.942
	AGAP011239-PA	FBN7	30.4	80	76	0.81	0.468
	AGAP009184-PA	FBN8	35.9	635	361	0.49	0.146
	AGAP010775-PA	FBN8	23.3	0	0	0	0.356
	AGAP011223-PA	FBN8	24.8	33	21	0.53	0.067
	AGAP011225-PA	FBN8	34.5	60	37	0.52	0.084
	AGAP009556-PA	FBN8	22.4	270	242	0.76	0.161
	AGAP004918-PA	Fibrinogen	35	32	39	1.03	0.835
	AGAP004996-PA	Fibrinogen	46.8	44	36	0.7	0.609
TEP	AGAP006743-PA	Fibrinogen	37.4	34	31	0.79	0.042
	AGAP006790-PA	Fibrinogen	30.8	4	1	0.29	0.165
	AGAP011197-PA	Fibrinogen	32.3	97	111	0.98	0.941
	AGAP004917-PA	Fibrinogen-related pr. 1	34.2	71	69	0.83	0.251
	AGAP006914-PA	Fibrinogen-related pr. 1	31.3	77	76	0.84	0.343
	AGAP010815-PA	TEP1	152.1	908	428	0.4	0.071
	AGAP008366-PA	TEP2	154.6	4	6	1.51	0.021
	AGAP010812-PA	TEP4	149.4	237	278	1	0.993
	AGAP010814-PA	TEP6	151.3	30	4	0.12	0.108
	AGAP010830-PA	TEP9	151.5	5	2	0.31	0.537
	AGAP008654-PA	TEP12	96.3	47	17	0.32	0.127

Group	ID	Name	M _r (kDa)	RA _C *	RA _I *	IP/CP	p-value
AMP	AGAP008368-PA	TEP14	139.4	5	1	0.17	0.091
	AGAP008364-PA	TEP15	163.6	16555	14077	0.73	0.045
	AGAP004632-PA	Defensin	10	0	1	27.53	0.356
	AGAP007199-PA	Defensin	7	4	9	2.08	0.341
	AGAP008645-PA	Gambicin	8.8	9	29	2.67	0.089
	AGAP007347-PA	Lysozyme 1 (c-type)	15.3	14	13	0.76	0.64
	AGAP007345-PA	Lysozyme 3 (c-type)	16.6	103	134	1.12	0.817
	AGAP007385-PA	Lysozyme 4 (c-type)	17.4	9	4	0.37	0.046
	AGAP007344-PA	Lysozyme 8 (c-type)	16.5	0	0	13.93	0.356
	AGAP011119-PA	Lysozyme 3	18	93	106	0.98	0.916
proPO	AGAP000376-PA	Transferrin precursor	69.2	4208	7912	1.61	0.055
	AGAP002825-PA	proPO1	79.3	110	43	0.34	0.031
	AGAP006258-PA	proPO2	78.1	11866	11762	0.85	0.045
	AGAP004975-PA	proPO3	78.6	21025	19615	0.8	0.024
	AGAP004981-PA	proPO4	78.5	448	442	0.84	0.357
	AGAP004977-PA	proPO6	79	710	880	1.06	0.827
	AGAP004980-PA	proPO7	79.6	4	19	3.66	0.414
	AGAP004976-PA	proPO8	79.3	521	629	1.03	0.914
SP	AGAP001798-PA	SP217	72.6	2	4	1.31	0.464
	AGAP005625-PA	SP213	146.8	128	154	1.03	0.903
	AGAP001365-PA	SP208	68.6	19	15	0.68	0.149
	AGAP005072-PA	SP219	96.4	35	22	0.53	0.002
	AGAP003686-PA	CLIPB47	39.8	165	160	0.83	0.315
	AGAP003251-PA	CLIPB1	40.9	210	123	0.5	0.004
	AGAP003246-PA	CLIPB2	38.4	21	9	0.35	0.078
	AGAP013487-PA	CLIPB3a	34.2	101	170	1.44	0.331
	AGAP003249-PA	CLIPB3b	40.1	218	173	0.68	0.087
	AGAP003250-PA	CLIPB4	39.4	4906	4007	0.7	0.076
	AGAP004148-PA	CLIPB5	41.3	147	175	1.02	0.891
	AGAP003057-PA	CLIPB8	44.8	1012	620	0.52	0.048

Group	ID	Name	M_r (kDa)	RA_C^*	RA_I^*	IP/CP	p-value
	AGAP013442-PB	CLIPB9,10	42.0	351	276	0.67	0.054
	AGAP009214-PA	CLIPB11	39.8	6	7	0.95	0.824
	AGAP004855-PA	CLIPB13	44.8	303	263	0.74	0.036
	AGAP009844-PA	CLIPB15	40.6	54	51	0.82	0.182
	AGAP011325-PA	CLIPB45	34.2	46	41	0.76	0.131
	AGAP008835-PA	CLIPC1	42.6	83	80	0.83	0.182
	AGAP004317-PA	CLIPC2	41.5	49	46	0.8	0.346
	AGAP004318-PA	CLIPC3	43.1	97	99	0.87	0.296
	AGAP000573-PB	CLIPC4	39.4	180	166	0.79	0.086
	AGAP000315-PA	CLIPC6	39.6	106	100	0.81	0.316
	AGAP004719-PA	CLIPC9	40.6	76	77	0.87	0.354
	AGAP000572-PA	CLIPC10	40.9	33	37	0.96	0.821
	AGAP002422-PA	CLIPD1	48.5	48	53	0.94	0.581
	AGAP002813-PA	CLIPD6	52.8	30	53	1.51	0.222
	AGAP012022-PA	CLIPE24	97	8	13	1.41	0.146
	AGAP008403-PA	CLIPE15	99.3	97	99	0.87	0.593
	AGAP012614-PA	CLIPB3b-like	43.4	103	186	1.55	0.261
	AGAP001198-PA	GP13	29.4	3	0	0	0.078
	AGAP005663-PA	GP71	33.8	394	280	0.61	0.026
	AGAP005670-PA	GP49	32.2	866	746	0.74	0.277
	AGAP005671-PA	GP65	32.2	2575	2765	0.92	0.709
	AGAP005686-PA	SP31	31.9	95	109	0.98	0.95
	AGAP007252-PA	GP51	32.9	43	37	0.74	0.266
	AGAP009121-PA	GP92	27.7	48	52	0.93	0.785
	AGAP005687-PA	SP33	32.1	14	16	1.01	0.968
	AGAP006674-PA	GP66	32.4	790	793	0.86	0.444
	AGAP006675-PA	GP60	32.1	9	4	0.41	0.214
	AGAP001246-PA	GP72	30.3	33	45	1.14	0.792
	AGAP001248-PA	GP70	28.9	12	16	1.12	0.814
	AGAP001249-PA	GP52	27.1	1283	1331	0.89	0.666
	AGAP006539-PA	SP63	28.8	37	67	1.54	0.369

Group	ID	Name	M_r (kDa)	RA_C^*	RA_I^*	IP/CP	p-value
	AGAP011920-PA	GP59	26.3	522	323	0.53	0.11
	AGAP012946-PA	SP53	35.5	285	271	0.81	0.359
	AGAP013221-PA	SP7	35	101	74	0.63	0.085
	AGAP004566-PA	SP25	35.7	11	13	0.95	0.837
	AGAP006673-PA	GP81	33.1	21	21	0.88	0.535
	AGAP002543-PA	SP71	29.8	1	0	0	0.142
	AGAP012328-PA	SP45	36.5	7	7	0.77	0.636
	AGAP008808-PA	SP130	67.5	67	79	1.02	0.784
	AGAP010240-PA	GP97	28.2	161	110	0.59	0.063
	AGAP003960-PA	SP122	64.6	391	340	0.74	0.171
	AGAP013252-PA	SP134	66.6	111	121	0.94	0.398
	AGAP011427-PA	SP111	96.2	60	28	0.39	0.061
	AGAP012504-PA	CLIP33	93.9	53	80	1.28	0.283
	AGAP027981-PA	CLIP33-like	98	19	41	1.91	0.09
	AGAP012269-PA	SP127	72.3	2	4	1.69	0.139
	AGAP007043-PA	SP112	59.9	186	148	0.68	0.166
SPH	AGAP001979-PA	SPH220	226	4	10	2.21	0.422
	AGAP012505-PA	CLIP23-like	31.5	0	1	24.8	0.356
	AGAP003691-PA	CLIP12	94.4	533	335	0.54	0.088
	AGAP000290-PA	CLIP27	54	1	2	1.07	0.938
	AGAP009216-PA	CLIPB43	33.9	0	0	0	0.356
	AGAP010730-PA	CLIPA28	28.2	865	701	0.69	0.072
	AGAP011791-PA	CLIPA1	48.4	582	392	0.58	0.112
	AGAP011790-PB	CLIPA2	55.9	701	669	0.82	0.508
	AGAP011780-PA	CLIPA4	45.9	1140	1127	0.84	0.473
	AGAP011789-PA	CLIPA6	45.7	4565	4961	0.93	0.561
	AGAP011792-PA	CLIPA7	80.9	1548	944	0.52	0.011
	AGAP010731-PA	CLIPA8	40.9	143	117	0.7	0.248
	AGAP011781-PA	CLIPA12	40.9	157	163	0.88	0.458
	AGAP011788-PA	CLIPA14	30.4	2296	2343	0.87	0.34
	AGAP002270-PA	CLIPB7	43.6	51	68	1.12	0.741

Group	ID	Name	M_r (kDa)	RA_C^*	RA_I^*	IP/CP	p-value
	AGAP013184-PA	CLIPB36	42.5	20	21	0.9	0.573
	AGAP003689-PA	CLIPC7	67	282	311	0.94	0.695
	AGAP005642-PA	GPH46	33	30	34	0.97	0.916
	AGAP011608-PA	GPH48	36.4	6	9	1.38	0.417
	AGAP011919-PA	GPH76	28.1	244	156	0.55	0.401
	AGAP011917-PA	GPH64	26.2	10	10	0.89	0.812
	AGAP005707-PA	GPH77	32.3	44	28	0.54	0.201
	AGAP005709-PA	GPH90	28.7	2	2	0.88	0.864
	AGAP006676-PA	GPH50	28.6	452	578	1.09	0.719
	AGAP004740-PA	GPH47	27.8	28	23	0.71	0.489
	AGAP005703-PA	GPH61	31.4	1	0	0	0.356
	AGAP005708-PA	GPH69	29.6	3	0	0	0.165
	AGAP003248-PA	SPH6	33.2	1	1	1.02	0.986
	AGAP004638-PA	SPH47	37.3	281	260	0.79	0.41
	AGAP006486-PA	GPH53	30.8	43	45	0.9	0.817
	AGAP006485-PA	GPH55	30.5	584	621	0.91	0.858
	AGAP013117-PA	SPH48	33.5	48	36	0.64	0.105
	AGAP001245-PA	GPH57	28.7	283	244	0.74	0.326
	AGAP006677-PA	GPH67	29.6	29	35	1.01	0.969
	AGAP009122-PA	GPH93	29.2	34	64	1.62	0.084
	AGAP006487-PA	GPH54	30.4	429	486	0.97	0.945
	AGAP013164-PA	SPH34	28.2	373	516	1.18	0.659
	AGAP003626-PA	GPH36	34.5	2511	2094	0.71	0.215
Serp	AGAP006909-PA	SRPN1	47.7	40	72	1.56	0.234
	AGAP006911-PA	SRPN2	46.5	2162	1987	0.79	0.199
	AGAP006910-PA	SRPN3	47.1	520	720	1.18	0.373
	AGAP009670-PA	SRPN4	68.9	338	340	0.86	0.596
	AGAP009670-PB	SRPN4	61.8	1341	1253	0.8	0.314
	AGAP007693-PA	SRPN7	44.3	368	408	0.95	0.792
	AGAP003194-PA	SRPN8	48.8	95	98	0.88	0.383
	AGAP003139-PA	SRPN9	50.4	2009	2956	1.26	0.365

Group	ID	Name	M _r (kDa)	RA _C *	RA _I *	IP/CP	p-value
	AGAP005246-PD	SRPN10	42.6	435	983	1.93	0.044
	AGAP005246-PE	SRPN10	42.2	0	4	124.84	0.356
	AGAP001377-PA	SRPN11	57.1	2342	2844	1.04	0.905
	AGAP001375-PA	SRPN12	64.8	891	1445	1.39	0.467
	AGAP009213-PA	SRPN16	61.1	1471	1349	0.78	0.314
	AGAP001376-PA	SRPN17	53.7	60	71	1.01	0.939
	AGAP008968-PA	Kazal domain protein	6.5	337	346	0.88	0.757
	AGAP011482-PA	Kazal domain protein	8.5	23	34	1.25	0.514
	AGAP006813-PA	SP inhibitor	13.4	2	4	1.5	0.377
	AGAP002585-PA	Cys-rich protein	175.6	6	15	2.35	0.34
	AGAP004631-PA	Clotting factor deficiency 2	26.1	1	2	1.4	0.553
	AGAP003987-PA	Clq binding protein	29.6	38	22	0.48	0.209
	AGAP002878-PA	Cystatin-like protein	11	66	83	1.08	0.49
	AGAP011460-PA	Cys-rich protein (salivary)	11.2	258	119	0.39	0.11
	AGAP006253-PA	Cys-rich venom protein	9.5	323	431	1.14	0.761
	AGAP012970-PA	Cys-rich venom protein	8.8	5	3	0.64	0.45
	AGAP011832-PA	Death-associated protein 1	10.3	97	68	0.6	0.395
	AGAP008878-PA	Defense protein	17.7	484	529	0.93	0.799
	AGAP010884-PA	DsCAM A	214.7	74	41	0.47	0.645
Others	AGAP000025-PA	E3 SUMO-pr. ligase 2	150.4	0	0	0	0.356
	AGAP002982-PA	E3 SUMO-pr. Lig. RanBP2	308.1	14	32	2	0.021
	AGAP010822-PA	Fasciclin	26.3	294	319	0.93	0.749
	AGAP010823-PA	Fasciclin isoform c	52.4	1	0	0	0.356
	AGAP001708-PA	Gd-domain protein	30.9	68	65	0.82	0.44
	AGAP008797-PA	Ig (CD79A) ₁ binding pr.	42.2	0	0	14.42	0.356
	AGAP000032-PA	Integrin alpha-ps2 x1	166.7	7	10	1.36	0.279
	AGAP007629-PB	Laminin gamma 1	179.6	10	11	1.01	0.965
	AGAP004993-PA	Laminin subunit alpha	412.1	13	12	0.76	0.207
	AGAP002857-PB	MDL2	18.1	58	34	0.5	0.02

Group	ID	Name	M_r (kDa)	RA _C *	RA _I *	IP/CP	p-value
	AGAP011319-PA	Pacifastin-related peptide	25.3	1342	1150	0.73	0.105
	AGAP008804-PB	Peroxin-19	33.2	3	1	0.36	0.25
	AGAP001325-PA	Peroxioredoxin 5, atypical	20.6	429	428	0.85	0.438
	AGAP004674-PA	Phenoloxidase inhibitor	36.3	1199	1294	0.92	0.727
	AGAP010477-PB	Phosducin-like 3	26.3	6	8	1.25	0.179
	AGAP005531-PA	PCD6-interacting protein	94.1	25	36	1.26	0.091
	AGAP000378-PA	PCD protein 4	47.4	23	31	1.15	0.857
	AGAP005432-PA	PCD protein 5	14.8	9	12	1.23	0.642
	AGAP003476-PA	Protein BCP1	33.6	26	29	0.93	0.772
	AGAP004333-PA	Multidomain TMP	173.7	45	14	0.27	0.125
	AGAP003012-PA	3PAN-ZP-TM	78.6	9	15	1.42	0.61
	AGAP000305-PA	SPARC	22.2	18	24	1.16	0.579
	AGAP011765-PA	Spondin-1	87	153	145	0.81	0.403
	AGAP003338-PA	Thioredoxin	15.5	3	1	0.32	0.384
	AGAP007201-PA	Thioredoxin	15.6	24	35	1.27	0.468
	AGAP009584-PA	Thioredoxin	12.1	361	270	0.64	0.119
	AGAP000396-PA	Thioredoxin peroxidase	26	25	24	0.82	0.433
	AGAP011054-PA	Thioredoxin peroxidase	22	747	949	1.09	0.209
	AGAP011824-PA	Thioredoxin peroxidase	25	777	865	0.95	0.703
	AGAP005462-PA	Thioredoxin-like pr. 1	31.6	18	19	0.89	0.528
	AGAP001613-PA	Thioredoxin-like TMP1	38.9	7	11	1.38	0.455
	AGAP003615-PA	Toll-interacting protein	30.4	3	5	1.42	0.516

* : Relative abundance (RA_C or RA_I) is defined as each protein's LFQs $\times$ 1000,000 $\div$ total LFQs for the CP or IP samples.

Table 5

A list of fifteen proteins likely cleaved by proteases

ID and name	M_r (kDa)	RA*	slice: M_r range	1 >350	2 250-350	3 230-250	4 140-230	5 80-140	6 70-80	7 45-70	8 30-45	9 22-30	10 20-22	11 15-20	12 <15
AGAP004993-PA	412.1	0.9	CP	0	0	0	0	0	44	24	16	3	0	14	0
Laminin subunit alpha		0.7	IP	0	0	0	0	0	30	27	23	13	0	7	0
AGAP010024-PA	345.1	1.2	CP	0	4	0	0	28	14	12	16	4	5	5	13
Dumpy		10.8	IP	0	0	0	0	9	7	5	47	17	5	7	3
AGAP002982-PA	308.1	0.5	CP	0	0	0	0	0	0	0	60	40	0	0	0
ligase RanBP2		1.5	IP	0	0	0	0	0	0	8	54	21	10	7	0
AGAP001979-PA	226	0.1	CP	0	0	0	0	0	0	0	32	23	46	0	0
SCRASPH3		0.5	IP	0	0	0	0	0	0	0	35	6	46	13	0
AGAP010021-PA	172.4	0.8	CP	0	0	0	0	0	14	12	31	20	10	12	0
Dumpy		6.8	IP	0	0	0	0	1	3	6	41	26	15	8	0
AGAP000032-PA	166.7	0.8	CP	0	0	0	0	0	0	0	3	31	18	48	0
Integrin α -ps2 x1		1.5	IP	0	0	0	0	0	0	0	0	27	32	41	0
AGAP009762-PA	141.6	40.2	CP	0	0	0	0	0	0	11	47	16	15	8	3
Nimrod		66.2	IP	0	0	0	0	0	0	14	25	23	24	10	4
AGAP005072-PA	96.4	2.3	CP	0	0	0	13	11	17	15	40	2	3	0	0
SP-IgG-2LamD		1.1	IP	0	0	0	27	0	12	5	53	3	0	0	0
AGAP011427-PA	96.2	1.2	CP	0	0	0	0	0	0	0	39	19	0	42	0
3SP		0.7	IP	0	0	0	0	0	27	0	32	41	0	0	0
AGAP004975-PA	78.6	0.5	CP	0	0	0	0	0	5	10	70	15	0	0	0
proPO3		1.2	IP	8	9	0	7	0	1	8	46	8	5	3	5
AGAP006644-PA	77	0.2	CP	0	0	0	0	0	0	0	71	29	0	0	0
LRR		0.2	IP	0	0	0	0	0	0	0	40	60	0	0	0
AGAP004674-PA	36.3	176.4	CP	0	0	0	0	1	0	0	2	5	7	49	34
PO inhibitor		180.1	IP	0	0	0	0	0	0	0	4	5	7	54	28
AGAP005707-PA	32.3	5.5	CP	3	2	0	3	4	0	1	3	8	57	17	3
SPH		3.2	IP	7	4	1	4	0	0	0	5	13	59	7	0
AGAP011319-PA	25.3	207.9	CP	0	0	0	0	0	0	0	1	4	19	63	14
Pacifastin-related		166.5	IP	0	0	0	0	0	0	0	1	4	15	58	21
AGAP004810-PA	20.8	23	CP	0	0	0	0	4	0	1	3	6	5	9	71
CTL3		28.7	IP	0	0	0	0	9	1	2	4	6	4	16	59

*: Relative abundance (RA) is defined as each protein's LFQs $\times$ 1000,000 $\div$ total LFQs of all proteins for the CP or IP samples. Value for a protein in a gel slice is the average percentage of that protein's total LFQ in the four CP or IP samples. These values are also shown in the gradient heat map from white (0) to red (100). The red boxes indicate the positions of the proteins based on their calculated M_r 's.

Table 6

Components of hypothesized high M_r serpin-protease complexes

ID and name	M_r (kDa)	RA*	slice: M_r range:	1 >350	2 250-350	3 230-250	4 140-230	5 80-140	6 70-80	7 45-70	8 30-45	9 22-30	10 20-22	11 15-20	12 <15
AGAP001376-PA SRPN17	53.7	3.6 5.1	CP IP	0 0	3 4	0 0	8 9	5 0	13 13	24 25	33 25	9 11	0 3	2 0	3 10
AGAP002813-PA CLIPD6	52.8	2.2 3.2	CP IP	0 3	4 7	7 13	8 8	12 13	18 7	14 15	16 16	4 3	0 4	5 2	12 9
AGAP003139-PA SRPN9	50.4	157.2 166.4	CP IP	3 2	3 2	4 3	5 4	11 13	14 17	28 44	29 11	2 2	0 1	1 1	1 1
AGAP003194-PA SRPN8	48.8	106.5 79.4	CP IP	7 9	10 13	6 8	18 22	13 12	10 8	28 20	3 3	1 1	0 2	1 1	1 1
AGAP006909-PA SRPN1	47.7	3.2 4	CP IP	0 4	0 8	0 0	14 11	10 5	20 15	28 40	29 17	0 0	0 0	0 0	0 0
AGAP006910-PA SRPN3	47.1	40.8 47.2	CP IP	5 5	4 4	4 4	7 8	7 6	10 11	18 26	20 11	19 18	2 2	2 3	3 3
AGAP006911-PA SRPN2	46.5	142 128.1	CP IP	3 3	3 4	4 4	8 10	7 4	19 22	25 34	24 11	2 3	0 2	1 2	3 2
AGAP003057-PA CLIPB8	44.8	88 72.3	CP IP	5 5	5 6	4 4	9 10	6 6	9 12	12 14	29 17	8 8	4 7	5 6	7 6
AGAP007693-PA SRPN7	44.3	23.2 26	CP IP	5 7	5 9	4 6	11 15	12 9	10 7	13 31	10 11	3 2	2 1	4 1	1 0
AGAP005246-PD SRPN10B	42.6	5 6.2	CP IP	8 7	7 6	6 5	10 9	7 8	10 10	15 8	8 6	5 6	0 3	19 27	4 3
AGAP008835-PA CLIPC1	42.6	29 68.5	CP IP	2 2	3 3	3 3	6 5	8 6	15 20	23 30	33 27	2 1	2 1	2 1	1 1
AGAP004317-PA CLIPC2	41.5	4.2 3.3	CP IP	4 4	3 2	1 0	7 13	9 11	9 14	11 0	19 16	7 8	7 5	5 5	19 22
AGAP004148-PA CLIPB5	41.3	9.3 11.2	CP IP	6 8	8 12	2 10	10 13	7 7	11 6	6 9	15 9	4 4	1 0	4 4	25 21
AGAP003251-PA CLIPB1	40.9	8.6 7.2	CP IP	6 10	5 9	6 8	9 15	14 12	13 9	21 19	25 17	0 0	0 2	0 0	0 0
AGAP004719-PA CLIPC9	40.6	1.8 2.1	CP IP	0 5	4 10	17 21	16 15	36 27	23 16	4 7	0 0	0 0	0 0	0 0	0 0
AGAP009214-PA CLIPB11	39.8	11.2 10.5	CP IP	7 9	7 8	9 10	6 8	12 9	15 17	16 17	6 7	9 9	6 7	4 0	4 1
AGAP003250-PA CLIPB4	39.4	12.1 11.8	CP IP	4 6	5 8	2 4	6 10	10 11	12 11	8 7	29 27	10 6	2 2	3 2	9 5
AGAP000573-PB CLIPC4	39.4	396.3 330.7	CP IP	7 7	6 7	7 7	10 11	9 8	15 13	10 11	13 10	4 4	7 8	5 5	8 10
AGAP003246-PA CLIPB2	38.4	1.5 0.4	CP IP	0 0	0 0	0 0	7 16	9 0	4 35	8 5	55 45	4 0	0 0	1 0	11 0

* : Relative abundance (RA) is defined as each protein's LFQs $\times 1000,000 \div$ total LFQs of all proteins for the CP or IP samples. Value for a protein in a gel slice is the average percentage of that protein's total LFQ in the four CP or IP samples. These values are also shown in the gradient heat map from white (0) to red (100). The red boxes indicate the positions of the proteins based on their calculated M_r 's.

Table 7

Components of hypothesized high M_r immune complexes

ID and name	M_r (kDa)	RA*	slice: M_r range:	1 >350	2 250-350	3 230-250	4 140-230	5 80-140	6 70-80	7 45-70	8 30-45	9 22-30	10 20-22	11 15-20	12 <15
AGAP010814-PA TEP6	151.3	1.3 0	CP IP	30 0	0 0	45 0	0 0	0 0	0 0	11 0	9 0	5 0	0 0	0 0	0 0
AGAP007060-PA LRR-7060	132.9	241.6 199.2	CP IP	17 20	30 36	22 23	22 16	6 4	2 1	1 1	0 0	0 0	0 0	0 0	0 0
AGAP007059-PA LRR-7059	124	387.9 383.5	CP IP	13 14	27 34	18 20	31 25	4 3	2 1	2 2	1 1	1 1	0 0	0 0	0 0
AGAP004832-PA LRR-1	117.8	49.8 38.8	CP IP	17 21	28 31	21 23	28 21	4 2	1 0	1 0	0 0	0 0	0 0	0 0	0 0
AGAP008403-PA SP-SPH-clip-SPH	99.3	6 4.8	CP IP	18 22	28 31	18 24	12 10	17 9	1 1	4 3	2 0	0 0	0 0	0 0	0 0
AGAP027981-PA SP-2SPH	98	1.1 2	CP IP	14 19	25 24	12 15	14 15	20 6	0 0	0 0	8 14	0 3	0 0	7 4	0 0
AGAP012022-PA SP-SPH-clip-SPH	97	0.3 0.4	CP IP	0 16	31 53	0 0	32 0	37 32	0 0	0 0	0 0	0 0	0 0	0 0	0 0
AGAP003691-PA 2SPH-clip-SPH	94.4	35.4 16.1	CP IP	13 18	24 25	16 19	13 12	21 14	6 7	3 2	2 1	1 1	0 1	0 0	2 0
AGAP012504-PA SP-2SPH	93.9	2.9 4.3	CP IP	10 15	19 25	9 15	13 12	23 14	2 2	8 1	9 14	5 2	0 0	0 0	0 0
AGAP005962-PA LRR shoc-2	90.7	2.6 2.1	CP IP	11 19	23 31	22 24	22 21	21 6	0 0	0 0	0 0	0 0	0 0	0 0	0 0
AGAP004980-PA proPO7	79.6	0.1 1.2	CP IP	0 15	0 15	61 60	39 11	0 0	0 0	0 0	0 0	0 0	0 0	0 0	0 0
AGAP002825-PA proPO1	79.3	7 2	CP IP	6 0	12 28	15 23	16 13	8 0	20 24	10 10	5 2	4 0	4 0	0 0	0 0
AGAP004976-PA proPO8	79.3	28.8 35.3	CP IP	5 8	14 16	19 34	22 26	14 5	13 3	4 1	3 1	4 5	0 1	2 0	1 0
AGAP004977-PA proPO6	79	45.9 57.6	CP IP	11 12	11 13	25 31	12 11	10 9	16 9	5 4	3 3	2 2	0 2	1 1	4 5
AGAP003012-PA SP71 isoform A	78.6	1531.9 1367.7	CP IP	7 10	9 11	21 23	10 11	8 6	24 16	6 5	4 4	3 3	3 5	3 3	2 2
AGAP004981-PA proPO4	78.5	33.4 26.9	CP IP	7 11	9 14	19 28	10 15	9 7	19 9	8 5	5 3	4 2	4 4	5 2	3 1
AGAP006258-PA proPO2	78.1	907.4 826.7	CP IP	11 12	10 12	17 18	9 11	8 7	21 14	6 5	4 4	3 3	3 4	5 6	3 4
AGAP008808-PA 3SPH	67.5	0.3 1	CP IP	0 72	0 0	0 28	100 0	0 0	0 0	0 0	0 0	0 0	0 0	0 0	0 0
AGAP013252-PA 2SP	66.6	6.8 8	CP IP	9 9	7 10	18 27	13 12	13 11	17 8	9 6	5 4	4 3	0 3	4 6	2 3
AGAP003960-PA 2SP	64.6	21.8 21	CP IP	17 15	16 17	14 27	18 18	13 9	4 3	11 5	3 2	3 2	0 2	0 0	1 0
AGAP007035-PA APL1B	63.9	0.8 0.5	CP IP	0 0	0 0	0 0	46 85	14 0	9 0	30 15	0 0	0 0	0 0	0 0	0 0
AGAP006761-PA GNBPA1	55.7	0.2 0.1	CP IP	0 0	12 0	0 25	64 75	24 0	0 0	0 0	0 0	0 0	0 0	0 0	0 0
AGAP006743-PA Fibrinogen	37.4	0.9 0.6	CP IP	0 0	25 0	0 0	29 52	45 48	0 0	0 0	0 0	0 0	0 0	0 0	0 0

* : Relative abundance (RA) is defined as each protein's LFQs $\times 1000,000 \div$ total LFQs of all proteins for the CP or IP samples. Value for a protein in a gel slice is the average percentage of that protein's total LFQ in the four CP or IP samples. These values are also shown in the gradient heat map from white (0) to red (100). The red boxes indicate the positions of the proteins based on their calculated M_r 's.