NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH Public Access
Author Manuscript
Dev Comp Immunol. Author manuscript; available in PMC 2013 April 01.

Published in final edited form as:
Dev Comp Immunol. 2013 April ; 39(4): 388–398. doi:10.1016/j.dci.2012.10.004.

A comprehensive analysis of the *Manduca sexta* immunotranscriptome

Ramesh T. Gunaratna and **Haobo Jiang**
Department of Entomology and Plant Pathology, Oklahoma State University, Stillwater, OK 74078, USA

Abstract

As a biochemical model, *Manduca sexta* has substantially contributed to our knowledge on insect innate immunity. The RNA-Seq approach was implemented in three studies to examine tissue immunotranscriptomes of this species. With the latest and largest focusing on highly regulated process- and tissue-specific genes, we further analyzed the same set of data using BLAST2GO to explore functional aspects of the larval fat body (F) and hemocyte (H) transcriptomes with (I) or without (C) immune challenge. Using immunity-related sequences from other insects, we found 383 homologous contigs and compared them with those discovered based on relative abundance changes. The major overlap of the two lists validated our previous research designed for gene discovery and transcript profiling in organisms lacking sequenced genomes. By concatenating the contigs, we established a repertoire of 232 immunity-related genes encoding proteins for pathogen recognition (16%), signal transduction (53%), microbe killing (13%), and others (18%). We examined their transcript levels along with attribute classifications and detected prominent differences in nine of the thirty level 2 gene ontology (GO) categories. The increase in extracellular proteins (155%) was consistent with the highly induced synthesis of defense molecules (*e.g.*, antimicrobial peptides) in fat body after the immune challenge. We identified most members of the putative Toll, IMD, MAPK-JNK-p38, and JAK-STAT pathways and small changes in their mRNA levels. Together, these findings set the stage for on-going analysis of the *M. sexta* immunogenome.

Keywords

insect immunity; hemolymph proteins; RNA-Seq; gene discovery; transcript profiling

1. Introduction

Insects possess a pristine form of the metazoan antimicrobial defense known as innate immunity (Hultmark, 1993), together with a facet of adaptive immunity via phagocyte-mediated immune memory (Pham et al., 2007). However, they lack the luxury of B and T cell-mediated adaptive immunity found in vertebrates (Agaisse, 2007). Insect immunity, comprising humoral and cellular responses, is rapid and effective in identifying and eliminating invading pathogens and parasites (Brey and Hultmark, 1998; Jiang et al., 2010; Lemaitre and Hoffmann, 2007). The general process of insect immunity, before deploying

Send correspondence to: Haobo Jiang, Department of Entomology and Plant Pathology, Oklahoma State University, Stillwater, OK 74078, Telephone: (405)-744-9400, Fax: (405)-744-6039, haobo.jiang@okstate.edu.

To get the CIFH contigs, readers can search the database “CIFH_contigs (March, 2010)” (http://darwin.biochem.okstate.edu/blast/blast.html) and retrieve sequences from a text file named “MsextaCIFHcontigs.fna” (http://entoplp.okstate.edu/profiles/jiang.htm). According to GenBank policy, contigs from more than one tissue or treatment are not allowed to be deposited there. Only CF, IF, CH, and IH contigs can be found in Transcriptome Shotgun Assembly Sequence Database at GenBank.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

killing mechanisms, consists of pathogen recognition via specific binding molecules (Kurata et al., 2006; Sansonetti, 2006; Yu et al., 2002), signal transduction and modulation via plasma serine proteinases and serine proteinase inhibitors (Gillespie et al., 1997; Kanost, 1999; Kanost et al., 2001; Marmaras and Lampropoulou, 2009), and receptor-mediated intracellular signaling via Toll (Valanne et al., 2011), IMD (Silverman and Maniatis, 2001), JNK (Ramet et al., 2002a), JAK-STAT (Baeg et al., 2005; Hou and Perrimon, 1997; Kisseleva et al., 2002), and MAPK-JNK-p38 (Han et al., 1998; Ragab et al., 2011) pathways. Signal transduction regulates both humoral and cellular immune responses. The former includes various antimicrobial peptides (AMPs) (Engstrom, 1999; Jiang, 2008), complement-like molecules (Aoun et al., 2011), and proteins involved in enzyme cascades that regulate melanin formation (Jang et al., 2008; Kanost and Gorman, 2008), which are synthesized and released into the plasma to entrap and kill invading pathogens or parasites (Gillespie et al., 1997; Hoffmann, 2003). In contrast, cellular immunity takes place in hemocytes and is comprised of phagocytosis, nodulation, and encapsulation (Fauvarque and Williams, 2011; Lavine and Strand, 2002; Strand, 2008; Zhuang et al., 2005).

Innate immunity plays a role in making insects the most diverse and abundant group of metazoans in the world (Chapman et al., 2006; Hultmark, 2003). This makes the immune system worth investigating in its own right. On the other hand, the common ancestry and similarities among insects and mammals make insects excellent model organisms (Hoffmann and Reichhart, 1997; Hultmark, 1993, 2003). These permit discovering evolutionary roots and features of animal immunity (Hoffmann et al., 1999; Khush and Lemaitre, 2000; Vilmos and Kurucz, 1998) and allow functional comparisons between diverse metazoan systems to identify shared and unique aspects of innate immunity (Khush and Lemaitre, 2000; Rolff and Reynolds, 2009; Wajant and Scheurich, 2004).

The advent of microarrays and next generation sequencing technologies coupled with bioinformatics tools has generated a large amount of immunotranscriptome data from insects with known genome sequence, such as *Drosophila* sp. (De Gregorio et al., 2001; Irving et al., 2001; Sackton et al., 2007), *Anopheles gambiae* (Christophides et al., 2002), *Apis mellifera* (Evans et al., 2006), *Aedes aegypti* (Waterhouse et al., 2007), *Tribolium castaneum* (Zou et al., 2007), *Bombyx mori* (Tanaka et al., 2008), and *Acyrthosiphon pisum* (Gerardo et al., 2010). Most of the immunotranscriptomic studies so far, for insects without sequenced genomes, lack quantitative levels of transcripts (Altincicek and Vilcinskas, 2007; Vogel et al., 2011; Zhang et al., 2010). As a member of economically important lepidopterans, *Manduca sexta* has been studied extensively in the field of insect physiology for decades (Jiang et al., 2010). Despite its prominent role, the *M. sexta* genome sequence is not yet published. Recently, transcriptomes of fat body, hemocytes, and midgut, in which many immunity-related genes are expressed, were determined using 454 pyrosequencing and Sanger sequencing technology (Pauchet et al., 2010; Zhang et al., 2011; Zou et al., 2008). The quantitative nature of the most recent study allowed us to analyze immune inducible and tissue specific gene expression. Although genome- and homology-independent discovery of new genes is possible, stringent thresholds set in the exploration hindered complete immunotranscriptomic analysis (Zhang et al., 2011). Therefore, the current work intended to extend the analysis by identifying most of the immunity-related genes in *M. sexta*, as a step towards the annotation of its immunogenome.

2. Methods and Materials

2.1. Construction, sequencing, and assembling of cDNA libraries

Insect rearing, bacterial injection, RNA isolation, cDNA synthesis, and library sequencing were described previously (Zhang et al., 2011). Briefly, fat body (F) and hemocytes (H) were prepared as controls (C) from sixty naïve larvae (5th instar, day 3) for total RNA

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

isolation and mRNA purification. Similarly, the same tissues were obtained from sixty induced (I) larvae (5th instar, day 3, injected with a mixture of bacteria 24 h before) for mRNA isolation and cDNA synthesis. After the CF, CH, IF, and IH cDNA libraries were separately run on a 454 GS-FLX pyrosequencer, reads were assembled to 19,020 CIFH contigs. For each contig, numbers of the CF, CH, IF, and IH reads incorporated were extracted from the Newbler Assembler output and tabulated using Microsoft Excel. As the tissues were pooled from sixty insects, read numbers are expected to faithfully represent the naïve and induced states of fat body and hemocytes.

2.2. Homologous sequence search, GO mapping, annotation and InterProScan search

The contigs were analyzed using the BLAST2GO software (Conesa et al., 2005; Gotz et al., 2008). In search for homologous sequences, the non-redundant protein database at NCBI was searched using BLASTX (Altschul et al., 1990) with a cutoff E-value of 10^{-15}. The BLAST hits were mapped to their corresponding GO annotations using the gene ontology database and several additional data files (Gotz et al., 2008). Subsequent annotation of contigs, to link information on cellular component (CC), molecular function (MF), and biological process (BP), was done by applying the annotation rule to all the GO terms. However, certain evidence code weights were changed from their default values to: EXP = IDA = IPI = 5, IMP = IGI = 4, and IEP = 3. Annotations were examined to remove broad or level 1 annotation. Additionally, the GO term known as auxin biosynthesis process was removed from the list of GO terms as the process does not exist in insects. Annex-based GO term augmentation was performed afterwards to, firstly, obtain extra annotations and, secondly, further validate annotations (Gotz et al., 2008; Myhre et al., 2006). Protein domain and signal peptide were predicted using InterProScan (Quevillon et al., 2005), which enabled further sequence annotation (Gotz et al., 2008). In order to obtain more refined annotations, level 1 annotation removal and Annex-based GO term augmentation were repeated.

2.3. Local BLASTX, domain search, and multiple sequence alignment

We downloaded immunity-related genes from *D. melanogaster* (462 genes from FlyBase using the keyword "immunity"), *B. mori* (205 genes from Tanaka et al. (2008)), and *A. mellifera* (184 genes from Evans et al. (2006)). Amino acid sequences of these genes were incorporated into a sequence database for local BLASTX analysis of the CIFH contigs. Domain prediction was performed in parallel search runs using batchwise domain search web utilities of web CD-search tool, Pfam and InterProScan. Sequence alignments and manual curation of the alignments were performed using MUSCLE (Edgar, 2004) implemented in MEGA 5 (Tamura et al., 2011).

2.4. Calculation of relative abundance of transcripts under immune challenge

Since each contig was assembled from reads in the four libraries, normalized read numbers (NRNs) were calculated as: actual reads number in library X × ($LNF_{CF} + LNF_{IF} + LNF_{CH} + LNF_{IH})/LNF_x$, where X is CF, IF, CH, or IH. Library normalization factors (LNFs) for CF (825), CH (3,980), IF (1,618), and IH (3,352) are the sums of read numbers for rpS2-rpS5, rpL4 and rpL8 in the corresponding libraries (Zhang et al., 2011). NRNs were then used to calculate relative abundance ($RA_{x/y} = NRN_x/NRN_y$). When a particular reads number was zero, an adjusted reads number ($ARN_{x/y}$ = actual read # in library X × LNFy/LNFx) was calculated instead. When multiple contigs encode a single gene, the particular contigs were concatenated and their NRNs summed in individual libraries for calculating RA or ARN values of the gene.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

2.5. Statistical analysis

Statistical differences in numbers of contigs or total normalized reads from a GO category were analyzed by Student's t-test. For instance, *p* values of immune inducibility (IC comparison) were derived from sums of number of contigs of IF and IH versus those of CF and CH. IC comparisons were also performed using sums of NRNs of IF and IH versus those of CF and CH. Similarly, *p* values of tissue specificity (FH comparison) were derived from sums of number of contigs and sums of NRNs of CF and IF versus those of CH and IH.

Percentage increases in numbers of contigs from a GO category were calculated as: I/C = [(sum of contig numbers of IF and IH − sum of contig numbers of CF and CH)/lower of the sums of contig numbers of the two groups] × 100 and F/H = [(sum of contig numbers of CF and IF − sum of contig numbers of CH and IH)/lower of the sums of contig numbers of the two groups] × 100. Similar calculations were performed using sums of NRNs from CF, IF, CH, and IH. Percentage increases in numbers of reads from a specific tissue (*e.g.*, fat body) were calculated as I/C_F = [(sum of IF NRNs − sum of CF NRNs)/lower sum of the two NRNs] × 100.

Data generated in above steps were merged, and mining of specific data were performed using SQL scripts in Microsoft Access, and default functions in Microsoft Excel.

3. Results

3.1. Distribution of M. sexta immunity-related genes

Here we report the repertoire of and changes in transcripts involved in multiple facets of innate immunity in *M. sexta*, such as pathogen recognition, signal transduction/modulation, and hemocyte adhesion. We identified 129 additional immunity-related genes (*i.e.* 204 contigs) in this study, apart from 103 highly regulated genes (*i.e.* 179 contigs) found in the previous study (Zhang et al., 2011). Taken together, genes for intracellular signal transduction account for 31% of the entire set; extracellular signaling molecules and their modulators make up 22% (Fig. 1). Gene products for pathogen recognition constitute 16%, whereas highly induced AMPs represent 13% of the total.

3.2. Global changes in level 2 GO categories

At GO level 2 (Gotz et al., 2008), expression of immunity-related genes is variable in fat body and hemocytes from naïve (C) and injected (I) larvae in terms of cellular component (CC), molecular function (MF) and biological process (BP). Since total numbers of the identified genes in each category do not significantly change between control and induced fat body (F) or hemocytes (H) (data not shown), we took advantage of the known read numbers for each contig in our datasets (Zhang et al., 2011), calculated summation of normalized read numbers (NRNs) for each gene (some concatenated from two or more contigs), and compared the sums of CF, IF, CH, and IH NRNs in each of level 2 GO categories (Fig. 2). In twelve of the thirty categories, their totals of all NRN sums were lower than 10% of the single highest NRN sum in the respective CC (51,074), MF (67,960), or BP (65,296) group and, therefore, omitted for statistical analysis. Five of the remaining eighteen had significant differences ($t < 0.05$): enzyme regulator activity (MF, $p = 0.001$, IC), molecular transducer activity (MF, $p = 0.008$, FH), cellular component organization (BP, $p = 0.001$, FH), developmental process (BP, $p = 0.024$, FH), and signaling (BP, $p = 0.015$, FH). Differences in the following five groups are less pronounced but worth mentioning, since level 2 GO terms are so general that a higher *p* value (*e.g.*, 0.05~0.20) may still reflect important changes: cell (CC, $p = 0.107$, FH), extracellular region (CC, $p =$

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

0.143, IC), macromolecular complex (CC, $p = 0.060$, FH), immune system (BP, $p = 0.145$, IC), and response to stimulus (BP, $p = 0.161$, IC).

We further inspected percentage changes of NRNs in the eighteen level 2 GO categories. When IC and FH comparisons were performed, we observed >50% changes in the following fifteen categories: extracellular region (CC, I > C: 106%), macromolecular complex (CC, H > F: 160%), catalytic activity (MF, I > C: 93%), enzyme regulator activity (MF, I > C: 74%), molecular transducer activity (MF, H > F: 381%), biological adhesion (BP, C > I: 72%, H > F: 89%), biological regulation (BP, H > F: 65%), cell wall organization or biogenesis (BP, I > C, 300%; F > H, 767%), cellular component organization (BP, F > H, 403%), developmental process (BP, F > H: 83%), immune system (BP, I > C: 300%), localization (BP, I > C: 95%), metabolism (BP, I > C: 118%, F > H: 90%), response to stimulus (BP, I > C: 392%, F > H: 59%), and signaling (BP, H > F: 276%).

While differences were observed in more categories between fat body and hemocytes, it is perhaps more interesting from the perspective of immunity to document major increases in total NRNs in either tissue before and after the immune challenge. Therefore, we studied the dataset and detected over 50% changes in extracellular region (CC, F: 155%, H: 59%), binding (MF, F: 95%), catalytic activity (MF, F: 243%), enzyme regulator activity (MF, F: 77%; H: 71%), biological regulation (BP, F: 111%), cell wall organization or biogenesis (BP, F: 297%; H: 329%), cellular process (BP, F: 69%), immune system (BP, F: 756%, H: 99%), localization (BP, F: 265%), metabolism (BP, F: 246%), and response to stimulus (BP, F: 1015%, H: 126%). The most dramatic increases in NRN occurred in the categories of extracellular region (CC, F: 31,038, 155%) and catalytic activity (MF, F: 48,149, 243%). The increase in extracellular protein transcripts was consistent with the highly induced synthesis of defense molecules (*e.g.*, AMPs) in fat body after the immune challenge.

3.3. Pathogen recognition

Pathogen detection is essential in subsequent measures taken to counteract the invasion. In insects, recognition proteins sense the pathogen presence by binding to their surface components known as pathogen-associated molecular patterns. We previously reported highly regulated β-1,3-glucan recognition proteins (βGRPs) and peptidoglycan recognition proteins (PGRPs) among others (Zhang et al., 2011). Here we report sixteen new genes coding for putative pattern recognition proteins: leureptin-2, Dscam, thioester-containing protein (TEP)-1 and -2, galectin-2 and -4, nimrod A, Draper, PGRP-L2, -L5, -LC, -S2, βGRP-3 and -4, immulectin (IML)-3a and - 3b (Table S1).

Apart from a role in pathogen recognition, these proteins are involved in phagocytosis as well as activation of signaling cascades. Leureptin (Zhu et al., 2010), Dscam (Watson et al., 2005), TEPs (Blandin and Levashina, 2004), galectins (Pace and Baum, 2004), nimrod A, Draper (Fauvarque and Williams, 2011) and IML-3 promote phagocytosis whereas PGRPs, βGRPs and IML-3 activate signaling cascades (Jiang et al., 2010). Among phagocytosis promoters, Dscam, nimrod A, and Draper showed mRNA level increases of >1.5-fold while leureptin and IML-3b transcript levels were up-regulated more than two-fold. In contrast, galectin-4 showed more than two-fold down-regulation in hemocytes. Transcript levels of galectin-2, TEP-1, and TEP-2 were low and their changes small. Draper and nimrod A showed preferred expression in hemocytes, whereas leureptins were highly expressed in fat body. Among activators of signaling cascades, PGRP-S5 levels were highly up-regulated in fat body while PGRP-L2 was down-regulated by two-fold. PGRP-L5 and βGRP-3 mRNA levels did not change much after the immune challenge. Unlike βGRPs involved in immunity, βGRP-4 contains a signal peptide and a GH16 domain with the catalytic residues (Glu, Asp, and Glu) but no RGD motif.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

3.4. Extracellular enzymes and their regulation

Many members of the serine proteinase family have been cloned and characterized from *M. sexta*. Some comprise an extracellular enzyme system that leads pathogen recognition to killing mechanisms. These proteinases are sequentially activated and later down-regulated by inhibitors in the plasma (Jiang et al., 2005). We have identified contigs encoding putative signal mediators and modulators (Table S2), including nineteen hemolymph proteinases (HPs), prophenoloxidase-activating proteinases (PAPs), scolexin, serine proteinase homologs (SPHs), Zn proteinase, twelve serpins and two other proteinase inhibitors, as well as enzymes involved in melanization (*e.g.* punch, Phe hydroxylase, Tyr hydroxylase, dopa decarboxylase, phenoloxidase) (Jiang et al., 2010; Krishnakumar et al., 2000; Zhang et al., 2011).

In *M. sexta*, a serine proteinase cascade produces active phenoloxidase (PO) that catalyzes the formation of quinones and melanin (Kanost et al., 2004). HP14, an initiator of the pathway, was up-regulated along with the next cascade component, HP21 (Wang and Jiang, 2007). Both genes were predominantly expressed in fat body. Up-regulation of HP14 transcription occurred in fat body and hemocytes and this was true for HP21 only in hemocytes. Another branch of the proPO activation system stems from HP6. HP6 mRNA was slightly up-regulated in both tissues whereas proHP8 (an HP6 substrate) was mainly synthesized in fat body. Transcripts of proPAP1 (another HP6 substrate) showed a 3.3-fold up-regulation in fat body. Five *M. sexta* serpins are known to regulate the proteinase system at multiple steps (Jiang et al., 2010). Serpin-1 transcript level is high (NRN >7000) in fat body and does not change after immune challenge. Serpin-3, -4, and -6 mRNA levels experienced more than two-fold increases. Serpin-3 and -4 transcripts increased only in fat body and serpin-6 mRNA was up-regulated in hemocytes. PO is one member of the enzyme system for melanization. Other members include up-regulated Phe hydroxylase and slightly up-regulated GTP cyclohydrolase I (*i.e.* punch). In addition, we have identified HP2, HP3 (IH/CH: 0.5), HP4 (IH/CH: 0.8), HP5 (IH/CH: 5.8), HP12 (IF/CF: 1.3), HP13 (IH/CH: 1.0), HP15 (IH/CF: 6.5), HP20 (IF/CF: 4.4), SPH2, serpin-7 (IF/CF 2.0), and serpin-11.

3.5. Signal transduction via major signaling pathways

The Toll, IMD, JAK-STAT, JNK, and p38 signal transduction pathways (Table 1) govern the production of effector molecules to eliminate pathogens (Fig. 3) and, hence, have been in the limelight of insect innate immunity research (Boutros et al., 2002; Dostert et al., 2005; Han et al., 1998; Kallio et al., 2005; Kim and Kim, 2005). We have identified orthologs of the pathway members and assume their functions and modes of action are conserved among insects.

3.5.1. Toll pathway—Components of Gram-positive bacteria and fungi activate the *Drosophila* Toll pathway through a cytokine, Spätzle (Lemaitre et al., 1996; LeMosy et al., 1999) (Fig. 3A). *M. sexta* Spätzle-1B (An et al., 2010) had a four-fold increase in mRNA levels in both tissues. The Toll receptor showed 2.2- and 1.5-fold up-regulation in fat body and hemocytes, respectively. All members of a complex formed after the Toll activation (MyD88, Tube, and Pelle) (Weber et al., 2003) were up-regulated in both tissues. The increase in *Ms*MyD88 mRNA level was small (1.5-fold) while *Ms*Pelle transcripts elevated 5.1-fold in fat body. *Ms*Cactus was highly up-regulated in fat body and slightly in hemocytes after the injection but *Ms*Dorsal transcript abundance did not change significantly.

Sumoylation plays a regulatory role in immunity by covalent modification of proteins in the NF-κB signaling pathways (Mabb and Miyamoto, 2007). We have identified one ubiquitin-conjugating (Ubc) protein similar to *Drosophila* Lesswright or Ubc9, a conjugating enzyme

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

that stabilizes Cactus (Table 1) (Abraham, 2007; Huang et al., 2005). *Ms*Lesswright mRNA showed 4.3-fold up-regulation in fat body whereas Smt3, a possible activator of Dorsal (Bhaskar et al., 2002; Xu et al., 2010), had a 1.6-fold increase in both tissues. Uba2 and Aos1, whose transcript levels showed slight increase of 1.5-fold in hemocytes, may activate Smt3 (Bhaskar et al., 2000).

Besides the aforementioned components, Pellino, Tollip-d, Tollip-v, TRAF, atypical PKC (aPKC), Ref(2)P, and ECSIT are associated with the Toll pathway as well (Valanne et al., 2011). We identified all their transcripts but TRAF. *Ms*Pellino mRNA level increased 2.5-fold in fat body after the challenge, other five members had <1.5-fold changes in mRNA levels. *Ms*Tollip-d was preferentially expressed in fat body whereas *Ms*Tollip-v mRNA was abundant in hemocytes.

3.5.2. IMD pathway—Diaminopimelic acid-PG of Gram-negative bacteria activates the IMD pathway via PGRP-LC, -LE, and IMD in *Drosophila* (Choe et al., 2002; Gottar et al., 2002; Kaneko et al., 2004 and 2006; Ramet et al., 2002b) (Fig. 3B). Members of a putative IMD-FADD-Dredd complex in *M. sexta* showed a two-fold up-regulation in fat body except for FADD (IF/CH: 0.6) (Table 1). *Ms*IAP2 mRNA decreased more than three-fold in fat body. Functionally critical *Ms*Uev1 and *Ms*Ubc13 showed a constitutive level of expression. Despite its key role in the IMD pathway, *Ms*TAK1 mRNA level was low. The TAK1 partner Tab2, however, had a 2.6-fold increase in transcript level in fat body. The *Drosophila* IKK complex may cleave Relish's ankyrin (ANK) repeats and cause it to translocate into the nucleus (Stoven et al., 2003). Transcript levels of *Ms*IKKβ were low but *Ms*IKKγ's increased more than two-fold in fat body. *Ms*Relish expression was also highly up-regulated in fat body.

A homolog of *Dm*Serpent, which activates AMP gene expression by binding to cognate *cis* regulatory elements (Petersen et al., 1999), was down-regulated more than three-fold in fat body. In contrast, *Dm*Ntf-2 prevents Dorsal, Dif, or Relish from nuclear translocation when mutated (Bhattacharya and Steward, 2002), and its homolog showed two-fold up-regulation in hemocytes (Table 1). Among regulators of the IMD pathway, Sickie, Caspar, and POSH homologs are found: *Dm*Sickie induces the Dredd-mediated Relish cleavage (Foley and O'Farrell, 2004) and *Ms*Sickie transcript level is down-regulated in fat body but highly up-regulated in hemocytes; *Dm*Caspar inhibits IMD signaling (Lee and Ferrandon, 2011) and *Ms*Caspar mRNA level is up-regulated in both tissues; *Dm*POSH governs the IMD pathway activation and termination as well as JNK pathway activation via regulation of TAK1 degradation (Lee and Ferrandon, 2011) and *Ms*POSH did not show any marked change in transcript levels.

3.5.3. MAPK-JNK-p38 pathway—In *Drosophila*, components of the Ras/MAPK pathway (Fig. 3D) activate JNK and p38, down-regulate the IMD pathway, and induce lamellocyte formation and hemocyte proliferation (Dong et al., 2002; Lee and Ferrandon, 2011; Ragab et al., 2011). *M. sexta* PDGF/VEGF receptor (PVR) and Alk receptor, which may trigger the MAPK pathway and Rac1 activation (Zettervall et al., 2004), did not show much change in transcript levels. The Alk mRNA level was particularly low and limited to naïve hemocytes (Table S6). Among members of the MAPK pathway were Cdc42, Dsor1, Rac1, and Ras85D. Except for a two-fold up-regulation of *Ms*Rac1 in fat body and slight increase of *Ms*Ras85D in hemocytes, changes in mRNA levels were small. Among identified members that may trigger p38 branch of the MAPK pathway were homologs of *Dm*Licrone/MKK3 and *Dm*MEKK1 (Han et al., 1998; Inoue et al., 2001). Both *Ms*MKK3 and *Ms*MEKK1 showed two-fold down-regulation in hemocytes despite slight up-regulation of p38 mRNA levels in both tissues.

JNK, a branch of the MAP kinase pathway, mediates stress-related responses and controls AMP gene expression (Ragab et al., 2011). In *Drosophila*, TAK1, Rac1, mixed-lineage kinases (MLKs), or MKK4 initiates the JNK pathway (Gallo and Johnson, 2002; Park et al., 2004; Silverman et al., 2003; Williams et al., 2006). *Ms*MLK1 showed a slight increase in hemocytes, while *Ms*MKK4 mRNA level decreased two-fold (Table 1). We identified all the members of the JNK branch. *Ms*JNK mRNA was highly up-regulated in hemocytes, FOS and Jra had a less than two-fold up-regulation in fat body, and FOS mRNA level increased 1.9-fold in hemocytes. Contig 3082 encodes a part of *Ms*JNK that has a protein kinase domain. JNK activates the transcription factor Aop (Anterior open) that mediates lamellocyte formation in *Drosophila* (Zettervall et al., 2004) and *Ms*Aop exhibited an up-regulation in fat body. Multiple ankyrin repeats single KH domain (MASK) is involved in signal relaying during the above processes (Bokoch, 2005; Hall, 1998; Kleino et al., 2005), and *Ms*MASK is slightly up-regulated in fat body and hemocytes.

3.5.4. JAK-STAT pathway—The *Drosophila* JAK-STAT pathway is involved in antiviral immune responses (Dostert et al., 2005) (Fig. 3D). We identified all members of this pathway, except for the cytokine Unpaired (upd). All of them contain corresponding functional domains and are highly similar in structure to their homologs in *B. mori*. Most of them had lower mRNA levels in fat body than hemocytes. *Ms*JAK (Hopscotch) showed a two-fold down-regulation in hemocytes.

3.6. Hemocyte adhesion

During an infection, usually non-adherent hemocytes tend to aggregate to trigger cellular immune responses against invading pathogens (Lavine and Strand, 2002). We identified three hemocyte-specific integrin α subunits, two β subunits, and one integrin-linked protein kinase (Table S3). Integrin α subunit mRNAs showed minimal changes except for α1, which was down-regulated in fat body. Integrin β subunits and the kinase were slightly up-regulated in hemocytes. We identified two other cell surface molecules, neuroglian and tetraspanin, that contribute to the integrin-mediated hemocyte aggregation (Nardi et al., 2006; Zhuang et al., 2007a; Zhuang et al., 2007b). Both proteins were highly expressed in hemocytes. Neuroglian had a more than 1.5-fold up-regulation in hemocytes while tetraspanin is slightly up-regulated in fat body. Among three paralytic peptide binding proteins (PPBP-1, -2, and -3), PPBP-1 had a slight increase in fat body while PPBP-3 expression was up-regulated more than two-fold in hemocytes. PPBP-3 mRNA level was higher than those of PPBP-1 and -2.

3.7. Autophagy

Autophagy governs the lysosome-dependent turnover of proteins or organelles and plays key roles in other cellular processes as well as human diseases (Shintani and Klionsky, 2004). Among nine different autophagy-related (Atg) molecules found are two ubiquitin-like proteins (Atg8 and Atg12), E1-like Atg5, Cys proteinase Atg4, and Atg4-like proteins (Table S4). These proteins are implicated in the process of macroautophagy (Geng and Klionsky, 2008). Except for up-regulated Atg8, all the other members are either expressed at low levels or down-regulated in fat body.

3.8. AMPs

We previously found 25 unique AMPs encoded by 61 highly up-regulated contigs (Zhang et al., 2011). Despite near complete coverage in that study, we identified other antimicrobial molecules, namely *M. sexta* lysozyme-like protein 1 (LLP1), four WAP-domain proteins, and attacin-3 through -6 (Table S5). LLP1 is slightly up-regulated in fat body. Two of the WAP-containing proteins showed increased expression in fat body, despite these genes are

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

poorly expressed. Attacins are highly expressed in fat body except for attacin-6, which is highly expressed only in hemocytes. All other attacins are also highly up-regulated in hemocytes with the exception of attacin-5, which is expressed only in the infected fat body. A closer look at the multiple sequence alignment of attacin-coding contigs revealed that the attacin family of AMPs comprise six members as opposed to two reported previously in *M. sexta* (Table S5, Fig. 4). There is a cluster of two attacin genes in the *B. mori* genome, closely similar to *Ms*Attacin-2. Similar gene duplications gave rise to 2–3 attacin genes in other lepidopteran insects. In *M. sexta*, a different gene expansion yielded five other genes (*Ms*Attacin-1, -3, -4, -5, and -6) in a lineage-specific way. A monophylatic group of four *D. melanogaster* attacin genes as well as *T. castaneum* attacin-1, -2 and -3 was probably generated in a similar way.

3.9. Others

This category comprises other genes involved in signaling, hemocyte proliferation and development, reactive molecular species synthesis and regulation, and gene silencing (Table S6). Focal adhesion kinase mediates signals from integrin receptors to MAPK pathway and, hence, plays a central role in regulating cellular immunity (Sieg et al., 1999). It has characteristic functional domains, such as FERM-M, PTK, and Focal-AT. The ligand Serrate and its receptor Notch mediate signal transduction to control hematopoiesis (Williams, 2007). *Ms*Serrate and *Ms*Notch are among the most highly expressed after AMPs. *Ms*Serrate mRNA level reduced in fat body while *Ms*Notch transcripts in hemocytes remained unchanged after the challenge. A transcription factor known as Brahma in *D. melanogaster* may also control hematopoiesis (Remillieux-Leschelle et al., 2002), and *Ms*Brahma shows no change in expression.

Reactive oxygen species (ROS) play a part in cytotoxic defense against microbes via activating AMPs or enhancing melanogenesis (Lavine and Strand, 2002). Nitric oxide synthase (NOS) generates nitric oxide (Nappi et al., 2000). However, the expression of *Ms*NOS is low in both the tissues even after the induction. Thioredoxin peroxidases and peroxiredoxins regulate amount of ROS, especially after an oxidative burst in the case of an infection, to maintain cellular homeostasis (Christensen et al., 2005; Nappi and Christensen, 2005). We found three thioredoxin peroxidases and one peroxiredoxin. Thioredoxin POD1 is up-regulated in hemocytes while thioredoxin POD3 and peroxiredoxin in fat body. However, thioredoxin POD2 is slightly down-regulated in fat body.

Homology-based gene silencing is involved in the *Drosophila* antiviral response (Wang et al., 2006). We identified homologs of Argonaute-1 and Dicer-2 that compose a part of the RNA interference silencing complex (RISC) (Ding et al., 2004).

4. Discussion

This extended study of the quantitative transcriptome data unveiled 95 new immunity-related genes in *M. sexta*. Along with the 137 reported previously (Jiang et al., 2010; Zhang et al., 2011), the number of such genes summed up to 232. In comparison to the 205, 462 and 184 genes retrieved from *B. mori*, *D. melanogaster* and *A. mellifera*, our studies, not based on annotated genome, revealed a similar number of genes. The deep analysis of fat body and hemocyte transcriptomes did uncover a large portion of the complete set of immunity-related genes that would come from a genome analysis. This is a valuable piece of information for researchers doing similar transcriptome studies in organisms that lack sequenced genomes. The extensiveness and depth of our transcriptome data are further supported by the discovery and analysis of six attacin genes in *M. sexta* (Fig. 4). Another important aspect of our transcriptomic data is that the inducibility of certain genes reported conformed well to known expression data on each of those genes (Zhang et al., 2011).

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Along the same line, we found significant similarities between inducibility of genes involved in the *Drosophila* Toll pathway (De Gregorio et al., 2001), further corroborating the data generated by Zhang *et al.* (2011) and utilized in this study.

A major goal of this research was not limited to discover sequences similar to the queries; rather it was to identify genes most likely related to immunity. For instance, BLAST search using aPKC as a query revealed 34 contigs at a commonly used cutoff E-value of 10^{-5}, but there is only one ortholog (contigs 5708 and 74333) in *M. sexta*. Twenty-eight of the hits were identified because they encode a kinase domain commonly found in genes, which may not be related to immunity. As such, many studies may have yielded inflated lists of homologous genes with limited value in orthology-based function predictions. Contrary to that practice, we took measures to reduce false positives, such as adopting a stringent threshold in the initial BLASTX analysis, searching for domain structures, and placing more weight on experimentally proven ontology in the GO annotation. Our initial BLASTX search (E-value $<10^{-15}$) against NCBI NR database resulted in 411 hits, dominated by lectins (80), proPO subunits (32), attacins (25), serine proteinases (30), and serine proteinase inhibitors (37). The parallel, local BLASTX analysis using known immunity-related genes from *B. mori, A. mellifera* and *D. melanogaster*, along with domain searches and multiple sequence alignments, yielded 379 highly scrutinized contigs. Although the number difference was only 18, the second list overlapped with the first only in 197 cases. Over 50% or 214 of the positives in the first list were incorrect: the use of a stringent threshold did not greatly reduce false positives; it, instead, yielded a lower number of valuable hits. Therefore, we adopted the 2nd list and improved it by merging 379 contigs into 232 groups, each of which represents one or more contigs putatively encoded by a single gene (Tables 1, S1–S6).

Based on the categorization of immune functions, we found genes for signal transduction and modulation account for 54% or 179 of the 232 genes whose products form pathways which crosstalk in multiple steps (Fig. 3). Genes for pathogen recognition and execution account for 16% and 10% of the gene set and, unlike signaling proteins, their products exert similar functions by extensively complementing each other to cope with a broad spectrum of infectious agents. The remaining 20% are involved in other processes, such as cell adhesion and autophagy. While this functional classification provided a good overview of the immune system, general GO analysis at level 2 did not yield clear differences in gene counts in the I-C and F-H comparisons (data not shown). Only after we took mRNA levels into consideration could significant differences be observed in certain categories of CC, MF, and BP at GO level 2 (Fig. 2). Six of the thirty groups are significantly different ($p < 0.20$) between fat body and hemocytes, whereas four categories are in the I-C comparison. Considering the high level of generalization in GO terms at level 2, we believe $p < 0.20$ is remarkable, especially when a large percentage of increase or decrease (>50%) is observed. The most dramatic changes occur in the categories of extracellular region (31,038, CC), catalytic activity (48,149, MF), and metabolism (46,433, BP) in fat body after the immune challenge. Highly induced expression of AMPs and other plasma defense proteins is partly responsible for the increase in total mRNA levels of extracellular molecules.

Phagocytosis, nodulation, and encapsulation as a result of pathogen recognition and cell adhesion comprise insect cellular immunity. Except for slightly up-regulated nimrod A, Dscam and Draper, up-regulated leureptin and IML-3, and down-regulated galectin-4, other genes showed no major change at the transcription level (Table S1), suggesting a complex regulation of phagocytosis. A majority of the genes involved in autophagy showed low transcript levels (Table S4). Atg-8, which plays a critical role in autophagy, had a 3.1-fold mRNA level increase in fat body, whereas Atg-3, -4, and -5 transcripts reduced to 1/3 in the same tissue after the immune challenge. Total NRNs of the Atg genes were 4.4-fold higher

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

in fat body than hemocytes. Since active engulfment of microbes occurs in the latter, the regulation of autophagy seems complicated, like phagocytosis, nodulation, and encapsulation. Hemolectin had a 2.5-fold down-regulation of mRNA level in hemocytes. In contrast, the increases in Reeler (IF^{ARN} 97.9; IH/CH: 3.0) and proPSP (IF/IH: 2.6) transcripts may enhance nodulation.

Melanogenesis plays a key role in immunity by participating in killing of entrapped microbes and wound healing. In *M. sexta*, an extracellular network of serine proteinases generates active PO, PSP, and Spätzle by proteolytic processing (Jiang et al., 2010). Many HP-related proteins were up-regulated after the immune challenges, including HP14, HP6, PAPs, and SPH1 (Table S2). Serpin-3, 4, 5, and 6, whose mRNAs became more abundant, negatively regulate some of these HPs that activate proPO. PO catalyzes the key steps for quinone and melanin formation, whereas other enzymes (*e.g.* Phe and Tyr hydroxylases, Punch, and dopa decarboxylase) also contribute to melanization. Substantial increases in their transcript levels (Table S2) further indicate the enzyme system for melanogenesis is highly coordinated and regulated at that level. In addition to proPAP1 activation, *M. sexta* HP6 generates active HP8 that processes Spätzle precursor (IF/CF 3.6, IH/CH 4.0) to initiate the Toll pathway (An et al., 2009 and 2010) for AMP induction. The increases in HP6 (IF/CF 1.6) and HP8 (IF/CF 1.9) mRNA levels indicate that reverse transcription-PCR is less quantitative than deep sequencing in detecting less than twofold induction.

Massively parallel pyrosequencing of transcripts from larval fat body and hemocytes allowed us to identify most components of the intracellular signaling pathways and quantify changes in their transcript levels after the immune challenges (Fig. 3). Although Dif, TRAF, PGRP-LE, and Wengen are missing in our contigs, evidence for the existence of Toll, IMD, MAPK-JNK-p38, and JAK-STAT pathways is compelling. We plan to search the genome sequence for these genes and profile their expression in these two tissues in the future. The current data did not show dramatic mRNA level changes, consistent with the fact that many members of the intracellular signal transduction pathways are activated via posttranslational modifications (*i.e.* phosphorylation and dephosphorylation). However, this does not exclude a possible regulation at the transcription level. In fact, major increases in mRNA levels were observed for Tube (IF/CF: 8.4), Cactus (IF/CF: 10.0), and Relish (IF/CF: 7.0), and the induced production of Cactus and Relish are probably related to their cleavage during immune signaling and the need to be replenished for a secondary response. After excluding these three genes, we calculated the averages and ranges of induction for Toll, IMD, MAPK-JNK-p38 pathways as 1.8 (1.3–2.2), 1.2 (0.3–2.0), and 1.2 (0.8–1.5) in fat body, as well as 1.4 (0.8–2.2), 1.1 (0.2–2.1), and 1.1 (0–1.9) in hemocytes, respectively. These increases, although small, may substantially contribute to the induction of AMP synthesis (Table S5). Interestingly, those values for the antiviral JAK-STAT pathway were 0.6 (0.2–0.9) in fat body and 0.8 (0.5–1.1) in hemocytes, consistent with the fact that we did not use any elicitor to mimic viral infection. It would be interesting to compare effects of virulent and incompatible viruses on transcription of the genes in the antiviral signaling pathway.

Concluding remarks

The next-generation sequencing approach we adopted has yielded a set of 19,020 contigs and corresponding read numbers from control and induced larval fat body and hemocytes of *M. sexta.* The long average size (923 bp), known immunity-related genes of other insects, and extensive sequence comparisons have facilitated the identification of 232 genes (or 383 contigs), assignment of GO terms and immune processes, and examination of transcriptional regulation of the entire system. The results validated our previous study, uncovered genes (*e.g.*, components of signaling pathways), demonstrated the practicality of genome-independent expression profiling of a complex process, and paved the way for annotation of the immunogenome.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Drs. Ulrich Melcher and Jack Dill with for their critical comments on the manuscript. This work was supported by National Institutes of Health Grants GM58634 (to H. Jiang). This article was approved for publication by the Director of the Oklahoma Agricultural Experiment Station and supported in part under project OKL02450.

Abbreviations

CF and CH	control (C) fat body (F) and hemocytes (H) from naïve larvae
IF and IH	induced (I) fat body and hemocytes from larvae injected with bacteria
Alk	anaplastic lymphoma kinase
AMP	antimicrobial peptide
ANK	ankyrin
aPKC	atypical protein kinase C
ARN	adjusted read number
AtgX	autophagy-related protein X
BP	biological process
CC	cellular component
CRD	carbohydrate recognition domain
CTL	C-type lectin
Dscam	Down syndrome cell adhesion molecule
ECSIT	evolutionarily conserved intermediate in Toll pathway
EGF	NIM and EMI, epidermal growth factor, nimrod and emilin
EST	expressed sequence tag
FN	fibronectin
GO	gene ontology
HAIP	hemocyte aggregation inhibitor protein
Hem	hemipterous
HP	hemolymph proteinase
IAP	inhibitor of apoptosis
Ig	immunoglobulin
IKK	IκB kinase
IMD	immune deficiency
IML	immulectin
JAK-STAT	Janus kinase-signal transducer and activator of transcription
JNK	Jun N-terminal kinase

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Jra	Jun-related antigen
LNF	library normalization factor
LPS	lipopolysaccharide
LRR	leucine-rich repeat
MAPK	mitogen-activated protein kinase
MASK	multiple ankyrin repeats single KH domain
MEKK	MAP kinase kinase kinase
MF	molecular function
MLK	mixed-linage kinase
NFκB and IκB	nuclear factor-κB and its inhibitor
NO and NOS	nitric oxide and its synthase
NRN	normalized read number
NTF	nuclear translocation
ORF	open reading frame
PAP	proPO-activating proteinase
PDGF and VEGF	platelet-derived and vascular endothelial growth factors
PG and PGRP	peptidoglycan and its recognition protein
PIAS	protein inhibitor of activated STAT
PO and proPO	phenoloxidase and its precursor
POSH	plenty of SH3 domains
PPBP	paralytic peptide-binding protein
PSP	plasmatocyte spreading peptide
PVR	PDGF/VEGF receptor
RA	relative abundance
RISC	RNA interference silencing complex
ROS	reactive oxygen species
SH2/3	src homology 2/3 domain
SOCS	suppressor of cytokine signaling
SP and SPH	serine proteinase and its homolog
SUMO	small ubiquitin-like modifier
TAK	transforming growth factor β (TGFβ) activated kinase
TEP	thioester-containing protein
TIR	Toll/interleukin-1 receptor
TRAF	tumor necrosis factor (TNF) receptor-associated factor
UBC	ubiquitin-conjugating domain
VWD	Von Willebrand disease factor

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

WAP	whey acidic protein
ZnF	zinc finger
α2M	α2-macrogobulins
βGRP	β-1, 3-glucanas related protein

References

Abraham, J. Ohio University; Ohio: 2007. The role of *Drosophila* SUMO conjugating enzyme Lesswright in larval hematopoiesis: Effects on Cactus, Dorsal and Dorsal-related immunity factor (Dif). LINK. http://rave.ohiolink.edu/etdc/view?acc_num=ohiou1187275172

Agaisse H. An adaptive immune response in *Drosophila*? Cell Host Microbe. 2007; 1:91–93. [PubMed: 18005687]

Altincicek B, Vilcinskas A. Identification of immune-related genes from an apterygote insect, the firebrat *Thermobia domestica*. Insect Biochem Mol Biol. 2007; 37:726–731. [PubMed: 17550828]

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990; 215:403–410. [PubMed: 2231712]

An C, Jiang H, Kanost MR. Proteolytic activation and function of the cytokine Spatzle in the innate immune response of a lepidopteran insect, *Manduca sexta*. FEBS J. 2010; 277:148–162. [PubMed: 19968713]

An C, Ishibashi J, Ragan EJ, Jiang H, Kanost MR. Functions of Manduca sexta hemolymph proteinases HP6 and HP8 in two innate immune pathways. J Biol Chem. 2009; 284:19716–19726. [PubMed: 19487692]

Aoun RB, Hetru C, Troxler L, Doucet D, Ferrandon D, Matt N. Analysis of thioester-containing proteins during the innate immune response of *Drosophila melanogaster*. J Innate Immunity. 2011; 3:52–64. [PubMed: 21063077]

Baeg GH, Zhou R, Perrimon N. Genome-wide RNAi analysis of JAK/STAT signaling components in *Drosophila*. Genes Dev. 2005; 19:1861–1870. [PubMed: 16055650]

Bhaskar V, Smith M, Courey AJ. Conjugation of Smt3 to dorsal may potentiate the *Drosophila* immune response. Mol Cell Biol. 2002; 22:492. [PubMed: 11756545]

Bhaskar V, Valentine SA, Courey AJ. A functional interaction between dorsal and components of the Smt3 conjugation machinery. J Biol Chem. 2000; 275:4033–4040. [PubMed: 10660560]

Bhattacharya A, Steward R. The *Drosophila* homolog of NTF-2, the nuclear transport factor-2, is essential for immune response. EMBO Rep. 2002; 3:378–383. [PubMed: 11943764]

Blandin S, Levashina EA. Thioester-containing proteins and insect immunity. Mol Immunol. 2004; 40:903–908. [PubMed: 14698229]

Bokoch GM. Regulation of innate immunity by Rho GTPases. Trends Cell Biol. 2005; 15:163–171. [PubMed: 15752980]

Boutros M, Agaisse H, Perrimon N. Sequential activation of signaling pathways during innate immune responses in *Drosophila*. Dev Cell. 2002; 3:711–722. [PubMed: 12431377]

Brey, PT.; Hultmark, D. Molecular mechanisms of immune responses in insects. Chapman & Hall; London; New York: 1998.

Chapman, AD. Australian Biodiversity Information, S., Australia. Numbers of living species in Australia and the world. Australian Govt., Dept. of Environ. Heritage; Canberra, ACT: 2006.

Choe KM, Werner T, Stoven S, Hultmark D, Anderson KV. Requirement for a peptidoglycan recognition protein (PGRP) in Relish activation and antibacterial immune responses in *Drosophila*. Science. 2002; 296:359–362. [PubMed: 11872802]

Christensen BM, Li J, Chen CC, Nappi AJ. Melanization immune responses in mosquito vectors. Trends Parasitol. 2005; 21:192–199. [PubMed: 15780842]

Christophides GK, Zdobnov E, Barillas-Mury C, Birney E, Blandin S, Blass C, Brey PT, Collins FH, Danielli A, Dimopoulos G, Hetru C, Hoa NT, Hoffmann JA, Kanzok SM, Letunic I, Levashina EA, Loukeris TG, Lycett G, Meister S, Michel K, Moita LF, Muller HM, Osta MA, Paskewitz

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

SM, Reichhart JM, Rzhetsky A, Troxler L, Vernick KD, Vlachou D, Volz J, von Mering C, Xu JN, Zheng LB, Bork P, Kafatos FC. Immunity-related genes and gene families in *Anopheles gambiae*. Science. 2002; 298:159–165. [PubMed: 12364793]

Conesa A, Gotz S, Garcia-Gomez JM, Terol J, Talon M, Robles M. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. Bioinformatics. 2005; 21:3674–3676. [PubMed: 16081474]

De Gregorio E, Spellman PT, Rubin GM, Lemaitre B. Genome-wide analysis of the *Drosophila* immune response by using oligonucleotide microarrays. Proc Natl Acad Sci USA. 2001; 98:12590–12595. [PubMed: 11606746]

Ding SW, Li H, Lu R, Li F, Li WX. RNA silencing: a conserved antiviral immunity of plants and animals. Virus Res. 2004; 102:109–115. [PubMed: 15068886]

Dong C, Davis RJ, Flavell RA. MAP kinases in the immune response. Annu Rev Immunol. 2002; 20:55–72. [PubMed: 11861597]

Dostert C, Jouanguy E, Irving P, Troxler L, Galiana-Arnoux D, Hetru C, Hoffmann JA, Imler JL. The Jak-STAT signaling pathway is required but not sufficient for the antiviral response of *Drosophila*. Nat Immunol. 2005; 6:946–953. [PubMed: 16086017]

Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 2004; 32:1792–1797. [PubMed: 15034147]

Engstrom Y. Induction and regulation of antimicrobial peptides in *Drosophila*. Dev Comp Immunol. 1999; 23:345–358. [PubMed: 10426427]

Evans JD, Aronstein K, Chen YP, Hetru C, Imler JL, Jiang H, Kanost M, Thompson GJ, Zou Z, Hultmark D. Immune pathways and defence mechanisms in honey bees *Apis mellifera*. Insect Mol Biol. 2006; 15:645–656. [PubMed: 17069638]

Fauvarque MO, Williams MJ. *Drosophila* cellular immunity: a story of migration and adhesion. J Cell Sci. 2011; 124:1373–1382. [PubMed: 21502134]

Foley E, O'Farrell PH. Functional dissection of an innate immune response by a genome-wide RNAi screen. PLoS Biol. 2004; 2:E203. [PubMed: 15221030]

Gallo KA, Johnson GL. Mixed-lineage kinase control of JNK and p38 MAPK pathways. Nat Rev Mol Cell Biol. 2002; 3:663–672. [PubMed: 12209126]

Geng JF, Klionsky DJ. The Atg8 and Atg12 ubiquitin-like conjugation systems in macroautophagy. EMBO Rep. 2008; 9:859–864. [PubMed: 18704115]

Gerardo NM, Altincicek B, Anselme C, Atamian H, Barribeau SM, De Vos M, Duncan EJ, Evans JD, Gabaldon T, Ghanim M, Heddi A, Kaloshian I, Latorre A, Moya A, Nakabachi A, Parker BJ, Perez-Brocal V, Pignatelli M, Rahbe Y, Ramsey JS, Spragg CJ, Tamames J, Tamarit D, Tamborindeguy C, Vincent-Monegat C, Vilcinskas A. Immunity and other defenses in pea aphids, *Acyrthosiphon pisum*. Genome biology. 2010; 11:R21. [PubMed: 20178569]

Gillespie JP, Kanost MR, Trenczek T. Biological mediators of insect immunity. Annu Rev Entomol. 1997; 42:611–643. [PubMed: 9017902]

Gottar M, Gobert V, Michel T, Belvin M, Duyk G, Hoffmann JA, Ferrandon D, Royet J. The *Drosophila* immune response against Gram-negative bacteria is mediated by a peptidoglycan recognition protein. Nature. 2002; 416:640–644. [PubMed: 11912488]

Gotz S, Garcia-Gomez JM, Terol J, Williams TD, Nagaraj SH, Nueda MJ, Robles M, Talon M, Dopazo J, Conesa A. High-throughput functional annotation and data mining with the Blast2GO suite. Nucleic Acids Res. 2008; 36:3420–3435. [PubMed: 18445632]

Hall A. Rho GTPases and the actin cytoskeleton. Science. 1998; 279:509. [PubMed: 9438836]

Han ZQS, Enslen H, Hu XD, Meng XJ, Wu IH, Barrett T, Davis RJ, Ip YT. A conserved p38 mitogen-activated protein kinase pathway regulates *Drosophila* immunity gene expression. Mol Cell Biol. 1998; 18:3527–3539. [PubMed: 9584193]

Hoffmann JA. The immune response of *Drosophila*. Nature. 2003; 426:33–38. [PubMed: 14603309]

Hoffmann JA, Kafatos FC, Janeway CA, Ezekowitz RAB. Phylogenetic perspectives in innate immunity. Science. 1999; 284:1313–1318. [PubMed: 10334979]

Hoffmann JA, Reichhart JM. *Drosophila* immunity. Trends Cell Biol. 1997; 7:309–316. [PubMed: 17708965]

Hou XS, Perrimon N. The JAK-STAT pathway in *Drosophila*. Trends Genet. 1997; 13:105–110. [PubMed: 9066269]

Huang L, Ohsako S, Tanda S. The lesswright mutation activates Rel-related proteins, leading to overproduction of larval hemocytes in *Drosophila melanogaster*. Dev Biol. 2005; 280:407–420. [PubMed: 15882582]

Hultmark D. Immune reactions in *Drosophila* and other insects: a model for innate immunity. Trends Genet. 1993; 9:178–183. [PubMed: 8337755]

Hultmark D. Drosophila immunity: paths and patterns. Curr Opin Immunol. 2003; 15:12–19. [PubMed: 12495727]

Inoue H, Tateno M, Fujimura-Kamada K, Takaesu G, Adachi-Yamada T, Ninomiya-Tsuji J, Irie K, Nishida Y, Matsumoto K. A *Drosophila* MAPKKK, D-MEKK1, mediates stress responses through activation of p38 MAPK. EMBO J. 2001; 20:5421–5430. [PubMed: 11574474]

Irving P, Troxler L, Heuer TS, Belvin M, Kopczynski C, Reichhart JM, Hoffmann JA, Hetru C. A genome-wide analysis of immune responses in *Drosophila*. Proc Natl Acad Sci USA. 2001; 98:15119–15124. [PubMed: 11742098]

Jang IH, Nam HJ, Lee WJ. CLIP-domain serine proteases in *Drosophila* innate immunity. BMB Rep. 2008; 41:102–107. [PubMed: 18315944]

Jiang H. The biochemical basis of antimicrobial responses in. Insect Sci. 2008; 15:53–66.

Jiang H, Vilcinskas A, Kanost MR. Immunity in lepidopteran insects. Adv Exp Med Biol. 2010; 708:181–204. [PubMed: 21528699]

Jiang H, Wang Y, Gu Y, Guo X, Zou Z, Scholz F, Trenczek TE, Kanost MR. Molecular identification of a bevy of serine proteinases in *Manduca sexta* hemolymph. Insect Biochem Mol Biol. 2005; 35:931–943. [PubMed: 15944088]

Kallio J, Leinonen A, Ulvila J, Valanne S, Ezekowitz RA, Ramet M. Functional analysis of immune response genes in *Drosophila* identifies JNK pathway as a regulator of antimicrobial peptide gene expression in S2 cells. Microbes Infect. 2005; 7:811–819. [PubMed: 15890554]

Kaneko T, Goldman WE, Mellroth P, Steiner H, Fukase K, Kusumoto S, Harley W, Fox A, Golenbock D, Silverman N. Monomeric and polymeric Gram-negative peptidoglycan but not purified LPS stimulate the *Drosophila* IMD pathway. Immunity. 2004; 20:637–649. [PubMed: 15142531]

Kaneko T, Yano T, Aggarwal K, Lim JH, Ueda K, Oshima Y, Peach C, Erturk-Hasdemir D, Goldman WE, Oh BH. PGRP-LC and PGRP-LE have essential yet distinct functions in the *Drosophila* immune response to monomeric DAP-type peptidoglycan. Nat Immunol. 2006; 7:715–723. [PubMed: 16767093]

Kanost MR. Serine proteinase inhibitors in arthropod immunity. Dev Comp Immunol. 1999; 23:291–301. [PubMed: 10426423]

Kanost, MR.; Gorman, MG. Phenoloxidases in insect immunity. In: Beckage, N., editor. Insect Immunology. Academic Press/Elsevier; San Diego: 2008. p. 69-96.

Kanost MR, Jiang HB, Wang Y, Yu XQ, Ma CC, Zhu YF. Hemolymph proteinases in immune responses of *Manduca sexta*. Phylogenetic Perspectives on the Vertebrate Immune System. 2001; 484:319–328.

Kanost MR, Jiang HB, Yu XQ. Innate immune responses of a lepidopteran insect, *Manduca sexta*. Immunol Rev. 2004; 198:97–105. [PubMed: 15199957]

Khush RS, Lemaitre B. Genes that fight infection - what the *Drosophila* genome says about animal immunity. Trends Genet. 2000; 16:442–449. [PubMed: 11050330]

Kim T, Kim YJ. Overview of innate immunity in *Drosophila*. J Biochem Mol Biol. 2005; 38:121–127. [PubMed: 15826489]

Kisseleva T, Bhattacharya S, Braunstein J, Schindler CW. Signaling through the JAK/STAT pathway, recent advances and future challenges. Gene. 2002; 285:1–24. [PubMed: 12039028]

Kleino A, Valanne S, Ulvila J, Kallio J, Myllymaki H, Enwald H, Stoven S, Poidevin M, Ueda R, Hultmark D, Lemaitre B, Ramet M. Inhibitor of apoptosis 2 and TAK1-binding protein are components of the *Drosophila* Imd pathway. EMBO J. 2005; 24:3423–3434. [PubMed: 16163390]

Krishnakumar S, Burton D, Rasco J, Chen XY, O'Donnell J. Functional interactions between GTP cyclohydrolase I and tyrosine hydroxylase in *Drosophila*. J Neurogenet. 2000; 14:1–23. [PubMed: 10938545]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Kurata S, Ariki S, Kawabata S. Recognition of pathogens and activation of immune responses in *Drosophila* and horseshoe crab innate immunity. Immunobiol. 2006; 211:237–249.

Lavine MD, Strand MR. Insect hemocytes and their role in immunity. Insect Biochem Mol Biol. 2002; 32:1295–1309. [PubMed: 12225920]

Lee KZ, Ferrandon D. Negative regulation of immune responses on the fly. EMBO J. 2011; 30:988–990. [PubMed: 21407251]

Lemaitre B, Hoffmann J. The host defense of *Drosophila melanogaster*. Annu Rev Immunol. 2007; 25:697–743. [PubMed: 17201680]

Lemaitre B, Nicolas E, Michaut L, Reichhart JM, Hoffmann JA. The dorsoventral regulatory gene cassette spatzle/Toll/cactus controls the potent antifungal response in *Drosophila* adults. Cell. 1996; 86:973–983. [PubMed: 8808632]

LeMosy EK, Hong CC, Hashimoto C. Signal transduction by a protease cascade. Trends Cell Biol. 1999; 9:102–107. [PubMed: 10201075]

Mabb AM, Miyamoto S. SUMO and NF-κB ties. Cell Mol Life Sci. 2007; 64:1979–1996. [PubMed: 17530464]

Marmaras VJ, Lampropoulou M. Regulators and signalling in insect haemocyte immunity. Cell Signal. 2009; 21:186–195. [PubMed: 18790716]

Myhre S, Tveit H, Mollestad T, Laegreid A. Additional gene ontology structure for improved biological reasoning. Bioinformatics. 2006; 22:2020–2027. [PubMed: 16787968]

Nappi A, Christensen B. Melanogenesis and associated cytotoxic reactions: applications to insect innate immunity. Insect Biochem Mol Biol. 2005; 35:443–459. [PubMed: 15804578]

Nappi AJ, Vass E, Frey F, Carton Y. Nitric oxide involvement in *Drosophila* immunity. Nitric Oxide Biol Chem. 2000; 4:423–430.

Nardi JB, Pilas B, Bee CM, Zhuang S, Garsha K, Kanost MR. Neuroglian-positive plasmatocytes of *Manduca sexta* and the initiation of hemocyte attachment to foreign surfaces. Dev Comp Immunol. 2006; 30:447–462. [PubMed: 16171863]

Pace KE, Baum LG. Insect galectins: roles in immunity and development. Glycoconj J. 2004; 19:607–614. [PubMed: 14758086]

Park JM, Brady H, Ruocco MG, Sun H, Williams D, Lee SJ, Kato T Jr, Richards N, Chan K, Mercurio F, Karin M, Wasserman SA. Targeting of TAK1 by the NF-κB protein Relish regulates the JNK-mediated immune response in *Drosophila*. Genes Dev. 2004; 18:584–594. [PubMed: 15037551]

Pauchet Y, Wilkinson P, Vogel H, Nelson DR, Reynolds SE, Heckel DG, ffrench-Constant RH. Pyrosequencing the *Manduca sexta* larval midgut transcriptome: messages for digestion, detoxification and defence. Insect Mol Biol. 2010; 19:61–75. [PubMed: 19909380]

Petersen UM, Kadalayil L, Rehorn KP, Hoshizaki DK, Reuter R, Engstrom Y. Serpent regulates *Drosophila* immunity genes in the larval fat body through an essential GATA motif. EMBO J. 1999; 18:4013–4022. [PubMed: 10406806]

Pham LN, Dionne MS, Shirasu-Hiza M, Schneider DS. A specific primed immune response in *Drosophila* is dependent on phagocytes. Plos Pathog. 2007; 3:e26. [PubMed: 17352533]

Quevillon E, Silventoinen V, Pillai S, Harte N, Mulder N, Apweiler R, Lopez R. InterProScan: protein domains identifier. Nucleic Acids Res. 2005; 33:W116–120. [PubMed: 15980438]

Ragab A, Buechling T, Gesellchen V, Spirohn K, Boettcher AL, Boutros M. *Drosophila* Ras/MAPK signalling regulates innate immune responses in immune and intestinal stem cells. EMBO J. 2011; 30:1123–1136. [PubMed: 21297578]

Ramet M, Manfruelli P, Pearson A, Mathey-Prevot B, Ezekowitz RA. Functional genomic analysis of phagocytosis and identification of a *Drosophila* receptor for *E. coli*. Nature. 2002a; 416:644–648. [PubMed: 11912489]

Ramet M, Lanot R, Zachary D, Manfruelli P. JNK signaling pathway is required for efficient wound healing in *Drosophila*. Dev Biol. 2002b; 241:145–156. [PubMed: 11784101]

Remillieux-Leschelle N, Santamaria P, Randsholt NB. Regulation of larval hematopoiesis in *Drosophila melanogaster*: a role for the multi sex combs gene. Genetics. 2002; 162:1259–1274. [PubMed: 12454071]

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

Rolff, J.; Reynolds, SE. Insect infection and immunity: Evolution, ecology, and mechanisms. Oxford University Press; Oxford: 2009.

Sackton TB, Lazzaro BP, Schlenke TA, Evans JD, Hultmark D, Clark AG. Dynamic evolution of the innate immune system in *Drosophila*. Nat Genet. 2007; 39:1461–1468. [PubMed: 17987029]

Sansonetti PJ. The innate signaling of dangers and the dangers of innate signaling. Nat Immunol. 2006; 7:1237–1242. [PubMed: 17110939]

Shintani T, Klionsky DJ. Autophagy in health and disease: A double-edged sword. Science. 2004; 306:990–995. [PubMed: 15528435]

Sieg DJ, Hauck CR, Schlaepfer DD. Required role of focal adhesion kinase (FAK) for integrin-stimulated cell migration. J Cell Sci. 1999; 112:2677. [PubMed: 10413676]

Silverman N, Maniatis T. NF-κB signaling pathways in mammalian and insect innate immunity. Genes Dev. 2001; 15:2321–2342. [PubMed: 11562344]

Silverman N, Zhou R, Erlich RL, Hunter M, Bernstein E, Schneider D, Maniatis T. Immune activation of NF-κB and JNK requires *Drosophila* TAK1. J Biol Chem. 2003; 278:48928–48934. [PubMed: 14519762]

Stoven S, Silverman N, Junell A, Hedengren-Olcott M, Erturk D, Engstrom Y, Maniatis T, Hultmark D. Caspase-mediated processing of the Drosophila NF-κB factor Relish. Proc Natl Acad Sci USA. 2003; 100:5991–5996. [PubMed: 12732719]

Strand MR. The insect cellular immune response. Insect Sci. 2008; 15:1–14.

Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol. 2011; 28:2731–2739. [PubMed: 21546353]

Tanaka H, Ishibashi J, Fujita K, Nakajima Y, Sagisaka A, Tomimoto K, Suzuki N, Yoshiyama M, Kaneko Y, Iwasaki T, Sunagawa T, Yamaji K, Asaoka A, Mita K, Yamakawa M. A genome-wide analysis of genes and gene families involved in innate immunity of *Bombyx mori*. Insect Biochem Mol Biol. 2008; 38:1087–1110. [PubMed: 18835443]

Valanne S, Wang JH, Ramet M. The *Drosophila* Toll signaling pathway. J Immunol. 2011; 186:649–656. [PubMed: 21209287]

Vilmos P, Kurucz E. Insect immunity: evolutionary roots of the mammalian innate immune system. Immunol Lett. 1998; 62:59–66. [PubMed: 9698099]

Vogel H, Altincicek B, Glockner G, Vilcinskas A. A comprehensive transcriptome and immune-gene repertoire of the lepidopteran model host *Galleria mellonella*. BMC Genomics. 2011; 12:308. [PubMed: 21663692]

Wajant H, Scheurich P. Analogies between *Drosophila* and mammalian TRAF pathways. Prog Mol Subcell Biol. 2004; 34:47–72. [PubMed: 14979664]

Wang XH, Aliyari R, Li WX, Li HW, Kim K, Carthew R, Atkinson P, Ding SW. RNA interference directs innate immunity against viruses in adult *Drosophila*. Science. 2006; 312:452–454. [PubMed: 16556799]

Wang Y, Jiang H. Reconstitution of a branch of the *Manduca sexta* prophenoloxidase activation cascade *in vitro*: snake-like hemolymph proteinase 21 (HP21) cleaved by HP14 activates prophenoloxidase-activating proteinase-2 precursor. Insect Biochem Mol Biol. 2007; 37:1015–1025. [PubMed: 17785189]

Waterhouse RM, Kriventseva EV, Meister S, Xi ZY, Alvarez KS, Bartholomay LC, Barillas-Mury C, Bian GW, Blandin S, Christensen BM, Dong YM, Jiang HB, Kanost MR, Koutsos AC, Levashina EA, Li JY, Ligoxygakis P, MacCallum RM, Mayhew GF, Mendes A, Michel K, Osta MA, Paskewitz S, Shin SW, Vlachou D, Wang LH, Wei WQ, Zheng LB, Zou Z, Severson DW, Raikhel AS, Kafatos FC, Dimopoulos G, Zdobnov EM, Christophides GK. Evolutionary dynamics of immune-related genes and pathways in disease-vector mosquitoes. Science. 2007; 316:1738–1743. [PubMed: 17588928]

Watson FL, Puttmann-Holgado R, Thomas F, Lamar DL, Hughes M, Kondo M, Rebel VI, Schmucker D. Extensive diversity of Ig-superfamily proteins in the immune system of insects. Science. 2005; 309:1874–1878. [PubMed: 16109846]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

Weber AN, Tauszig-Delamasure S, Hoffmann JA, Lelievre E, Gascan H, Ray KP, Morse MA, Imler JL, Gay NJ. Binding of the *Drosophila* cytokine Spatzle to Toll is direct and establishes signaling. Nat Immunol. 2003; 4:794–800. [PubMed: 12872120]

Williams MJ. *Drosophila* hemopoiesis and cellular immunity. J Immunol. 2007; 178:4711–4716. [PubMed: 17404248]

Williams MJ, Wiklund ML, Wikman S, Hultmark D. Rac1 signalling in the *Drosophila* larval cellular immune response. J Cell Sci. 2006; 119:2015–2024. [PubMed: 16621891]

Xu H, Hao W, He D, Xu Y. Smt3 is required for the immune response of silkworm, *Bombyx mori*. Biochimie. 2010; 92:1306–1314. [PubMed: 20561559]

Yu XQ, Zhu YF, Ma C, Fabrick JA, Kanost MR. Pattern recognition proteins in *Manduca sexta* plasma. Insect Biochem Mol Biol. 2002; 32:1287–1293. [PubMed: 12225919]

Zettervall CJ, Anderl I, Williams MJ, Palmer R, Kurucz E, Ando I, Hultmark D. A directed screen for genes involved in *Drosophila* blood cell activation. Proc Natl Acad Sci USA. 2004; 101:14192–14197. [PubMed: 15381778]

Zhang FJ, Guo HY, Zheng HJ, Zhou T, Zhou YJ, Wang SY, Fang RX, Qian W, Chen XY. Massively parallel pyrosequencing-based transcriptome analyses of small brown planthopper (*Laodelphax striatellus*), a vector insect transmitting rice stripe virus (RSV). BMC Genomics. 2010; 11:303. [PubMed: 20462456]

Zhang S, Gunaratna RT, Zhang X, Najar F, Wang Y, Roe B, Jiang H. Pyrosequencing-based expression profiling and identification of differentially regulated genes from *Manduca sexta*, a lepidopteran model insect. Insect Biochem Mol Biol. 2011; 41:733–746. [PubMed: 21641996]

Zhu YF, Ragan EJ, Kanost MR. Leureptin: a soluble, extracellular leucine-rich repeat protein from *Manduca sexta* that binds lipopolysaccharide. Insect Biochem Mol Biol. 2010; 40:713–722. [PubMed: 20688162]

Zhuang SF, Kelo L, Nardi JB, Kanost MR. Neuroglian on hemocyte surfaces is involved in homophilic and heterophilic interactions of the innate immune system of *Manduca sexta*. Dev Comp Immunol. 2007a; 31:1159–1167. [PubMed: 17448535]

Zhuang SF, Kelo LS, Nardi JB, Kanost MR. An integrin-tetraspanin interaction required for cellular innate immune responses of an insect, *Manduca sexta*. J Biol Chem. 2007b; 282:22563–22572. [PubMed: 17553801]

Zhuang SF, Nardi JB, Kanost MR. An integrin, neuroglian, and a tetraspanin modulate adhesion in hemocytes of an insect, *Manduca sexta*. FASEB J. 2005; 19:A1061–A1062.

Zou Z, Evans JD, Lu ZQ, Zhao PC, Williams M, Sumathipala N, Hetru C, Hultmark D, Jiang HB. Comparative genomic analysis of the *Tribolium* immune system. Genome Biol. 2007; 8(8):R177. [PubMed: 17727709]

Zou Z, Najar F, Wang Y, Roe B, Jiang H. Pyrosequence analysis of expressed sequence tags for *Manduca sexta* hemolymph proteins involved in immune responses. Insect Biochem Mol Biol. 2008; 38:677–682. [PubMed: 18510979]

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

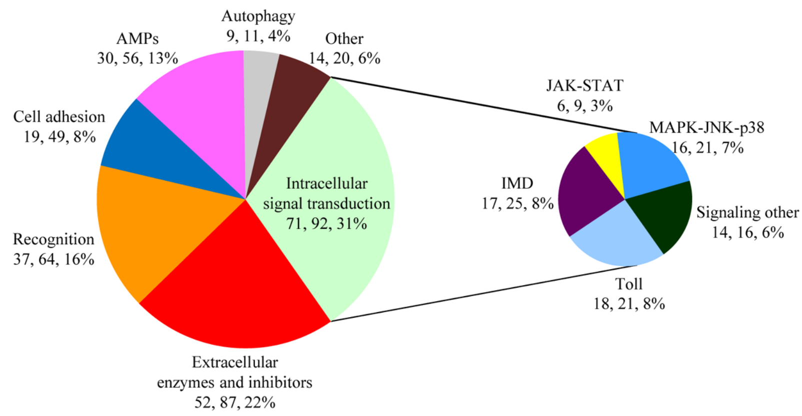

Fig. 1. Distribution of 232 *M. sexta* immunity-related genes

The pie chart shows gene number, contig number, and percentage of genes in each functional category relative to the entire set. The category of "intracellular signal transduction" is further divided into pathways (*right*).

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

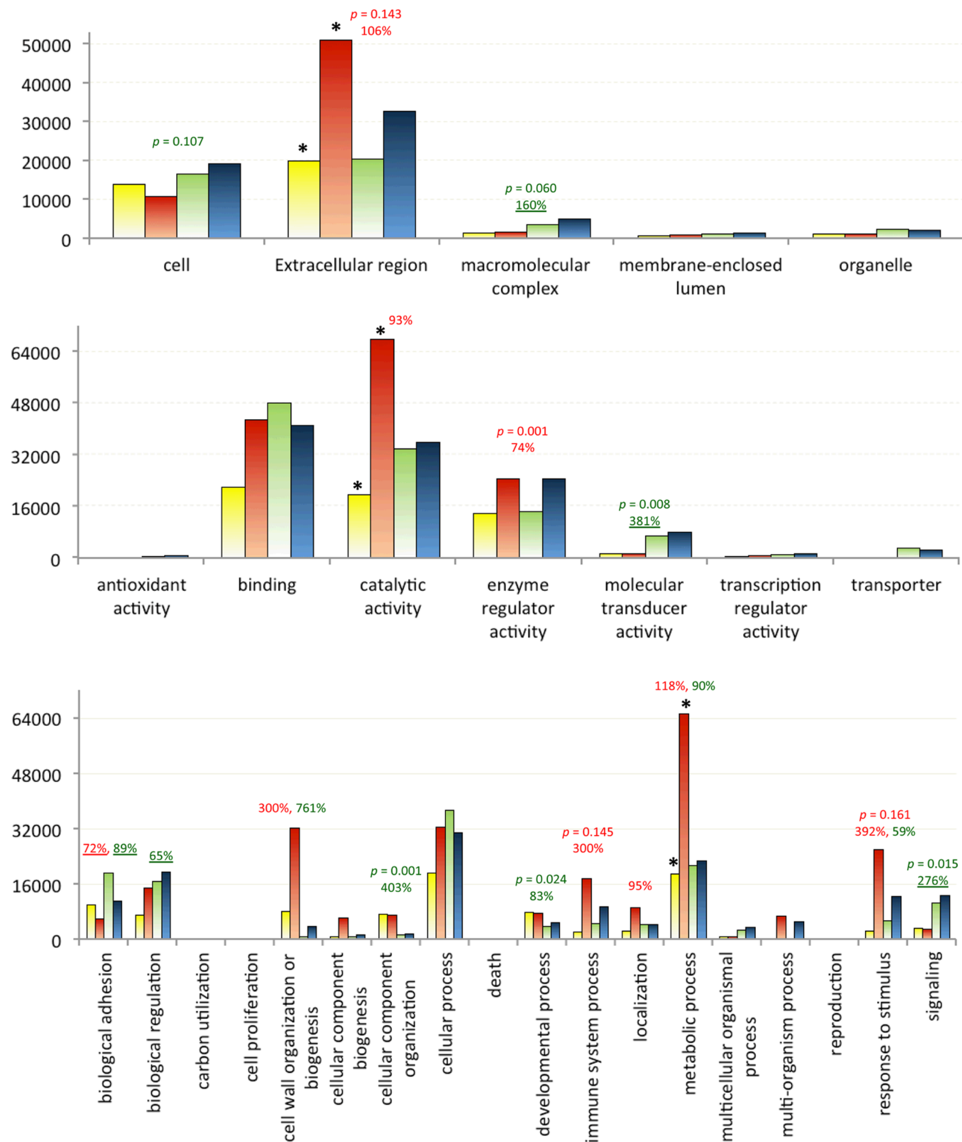

Fig. 2. Expression analysis of cellular components (CC, *top*), molecular functions (MF, *middle*), and biological processes (BP, *bottom*) at GO level 2

The bar graph is generated using data from the sum of NRNs for each annotated gene. Each GO term is comprised of four values, each from a particular library (CF: *yellow*, IF, *orange*, CH: *green*, and IH: *blue*). Bar height represents sum of sums of NRNs in a library within a specific GO group. *p* value (< 0.20) and percentage (>50%) increase or decrease (*underlined*) of immune inducibility (*red*, IF-IH vs. CF-CH) and tissue specificity (*green*, CF-IF vs. CH-IH) are indicated on the top. Of the eleven GO categories in which NRN sums have >50% differences in either or both tissues, three increase most dramatically and are marked with “*”.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

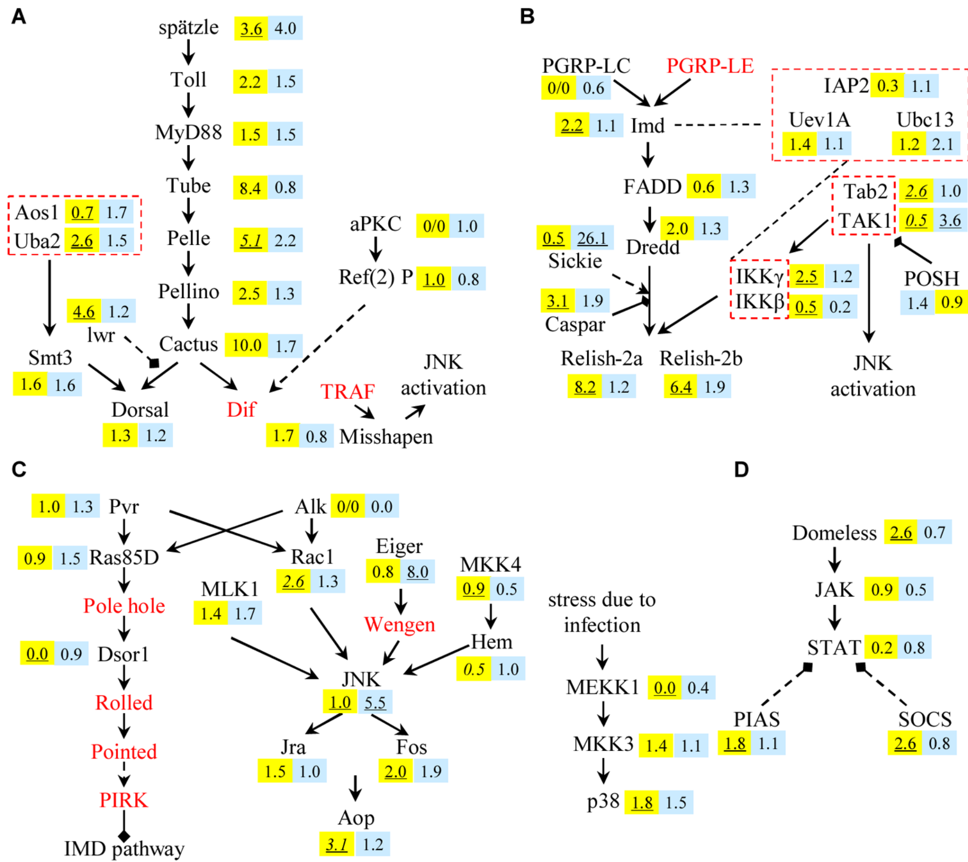

Fig. 3. Identification and profiling of transcripts involved in the Toll (A), IMD (B), JAK-STAT (C) and MAPK-JNK-p38 (D) signal transduction pathways

The intracellular signaling processes, based mostly on *Drosophila* research, are described in the text assuming the pathways are conserved among insects. Genes that are not found in our dataset are shown in *red*. Immune inducibilities (*i.e.*, NRN ratios or *ARN*s) in fat body (*yellow*) and hemocytes (*blue*) are indicated near the corresponding genes. *Underlined* number or 0/0 denotes low RN_{CF} or RN_{CH} (0 ~ 4) and, hence, less reliable NRN ratio.

NIH-PA Author Manuscript

NIH-PA Author Manuscript

NIH-PA Author Manuscript

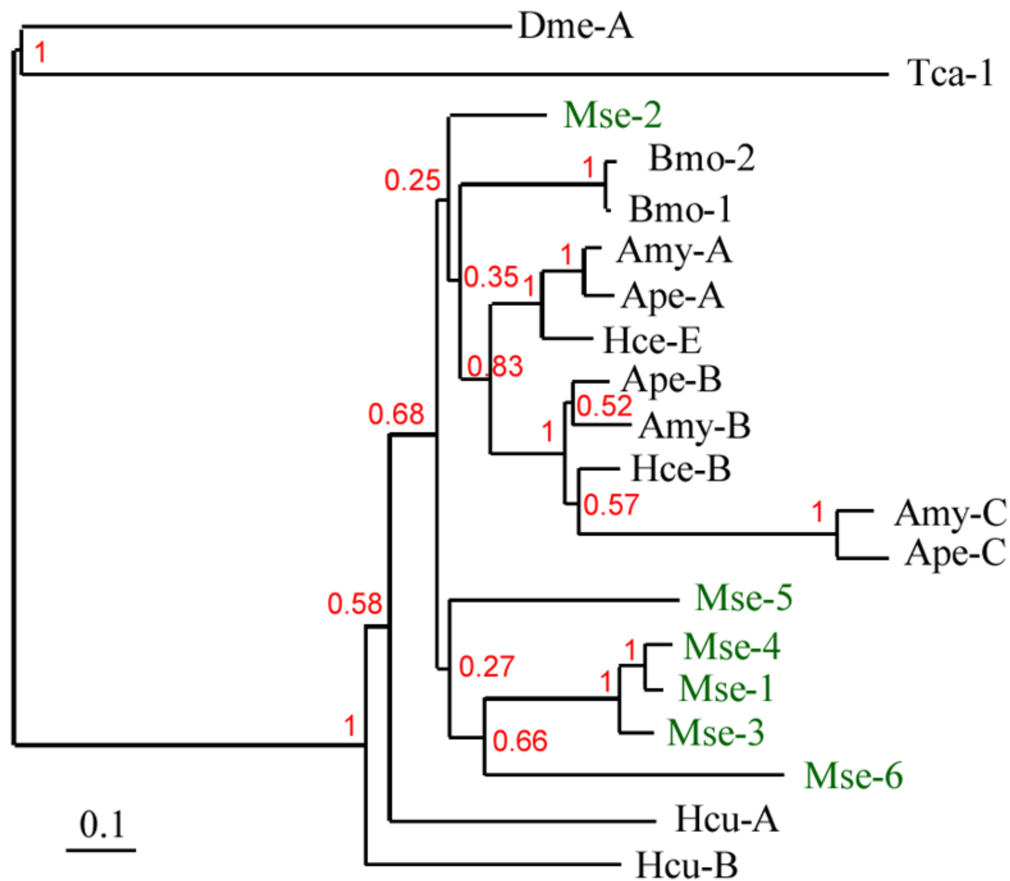

Fig. 4. Phylogenetic relationships among insect attacins

Amino acid sequences of *M. sexta* attacin-1 through -6 (Mse-1, 2, 3, 4, 5, 6, green); *Antheraea mylitta* attacin-A, B, C (Amy-A, B, C), *Antheraea pernyi* attacin-A, B, C (Ape-A, B, C), *B. mori* attacin-1 and 2 (Bmo-1, 2), *D. melanogaster* attacin-A (Dme-A), *Hyalophora cecropia* attacin-B and E (Hce-B, E), *Hyphantria cunea* attacin-A and B (Hcu-A, B), *T. castaneum* attacin-1 (Tca-1) are aligned using MUSCLE 3.7 at www.phylogeny.fr. The phylogenetic tree was constructed using JTT substitution matrix in ProtDist/FastDist+ neighbor with a bootstrap replicates of 1000. The bootstrap values (%) are indicated at nodes. Correct assembling of Mse-5 (17350 + 13563) and -6 (15744 + 16576 + 15159) from multiple contigs is confirmed by the draft genome sequence (data not shown).

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript

Table 1

Members of the intracellular signaling pathways

Gene name	Domain(s)	Contig(s)	nCF	nCH	nIF	nIH	nIF:nCF	nIH:nCH
		Toll pathway						
Spätzle		2287	11.8	68.8	42.3	274.1	3.6	4.0
Toll		5599, 6893, 14282	106.64	164.6	235.6	247.9	2.2	1.5
MyD88	Death, TIR	864	118.5	137.5	181.2	201.2	1.5	1.5
Tube	Death, PK	1313	23.7	127.7	199.4	102.1	8.4	0.8
Pelle	Death, PK	2038	0.0	17.2	60.4	37.9	*5.1*	2.2
Pellino		292	71.1	147.4	175.2	189.6	2.5	1.3
Cactus	ANK-2	1044, 3381, 15574	118.5	260.3	1190.2	452.0	10.0	1.7
Lesswright	UBC	4591	11.8	167.0	54.4	198.3	4.6	1.2
aPKC	PK, PK C-domain	5708, 7433	0.0	39.3	0.0	37.9	0/0	1.0
Ref(2)P	PBI, ZZ Zinc finger	5971	23.7	78.6	24.2	64.2	1.0	0.8
Rel/Dorsal	Rel domain	2384	82.9	198.9	108.7	242.0	1.3	1.2
ECSIT	ECSIT	4177	0.0	14.7	6.0	20.4	*0.5*	1.4
Tollip-d	CUE	4949	71.1	0.0	72.5	0.0	1.0	0/0
Tollip-v	C2, CUE	731	59.2	311.9	48.3	344.1	0.8	1.1
Smt3	Sumo	7946	59.2	167.0	96.7	262.5	1.6	1.6
Aos1	Aos1-SUMO	5438	35.5	56.5	24.2	93.3	0.7	1.7
Uba2	Uba2-SUMO	890	11.8	140.0	30.2	210.0	2.6	1.5
Misshapen	PK, CNH	289	106.6	311.9	181.2	259.5	1.7	0.8
		IMD pathway						
IMD	Death	2368	35.5	61.4	78.5	67.1	2.2	1.1
FADD	Death, DID	342	876.8	122.8	531.6	157.5	0.6	1.3
Dredd	Caspase (Peptidase-C14)	1615, 14535, 15028	71.1	149.8	139.0	201.2	2.0	1.3
IAP2	BIR/Inhibitor of apoptosis	1174, 7327, 8290, 9234	106.6	98.2	30.2	110.8	0.3	1.1
Ubc13/ben	UBC	2901	82.9	117.9	96.7	247.9	1.2	2.1
Uev1A	UBC	3326	154.0	368.4	217.5	411.2	1.4	1.1
TAK1	PTK	8422	0.0	4.9	6.0	17.5	*0.5*	3.6
Tab2	CUE	1637	0.0	44.2	30.2	43.7	*2.6*	1.0

Gene name	Domain(s)	Contig(s)	nCF	nCH	nIF	nIH	nIF:nCF	nIH:nCH
IKKβ	PK	5609	11.8	14.7	6.0	2.9	0.5	0.2
IKKγ		1049	11.8	93.3	30.2	107.9	2.6	1.2
Relish-2A	PEST	15531, 15532	11.8	61.4	96.7	72.9	8.2	1.2
Relish-2B	Rel homology domain	4802	23.7	103.2	151.0	192.5	6.4	1.9
Ntf2	NTF2	6033, 8947	118.5	169.5	157.1	332.4	1.3	2.0
Serpent	ZnF-GATA	4249, 17496	485.8	582.1	163.1	501.6	0.3	0.9
Sickie		5128, 7157	35.5	2.5	18.1	64.2	0.5	26.1
Caspar	UBX	2428	11.8	27.0	36.2	52.5	3.1	1.9
POSH	SH3	1777, 5429	47.4	338.9	66.5	309.1	1.4	0.9
		MAPK pathway with JNK and p38 branches						
Eiger	TNF	1020	497.6	9.8	380.6	78.7	0.8	8.0
Cdc42	Cdc42	647	130.3	579.6	114.8	621.1	0.9	1.1
Dsor1	PK	6185	11.8	34.4	0.0	29.2	0.0	0.9
Rac1	Rac1	3605	0.0	27.0	30.2	35.0	*2.6*	1.3
Ras85D	Ras	73, 132, 205, 1185	651.7	1350.8	555.8	1991.7	0.9	1.5
MLK1	PK, PTK, SH3	1825, 1841	59.2	117.9	84.6	198.3	1.4	1.7
MEKK1	PK	1947	11.8	49.1	0.0	20.4	0.0	0.4
Licrone/MKK3	PK-MKK3-6	2351	82.9	108.1	114.8	122.5	1.4	1.1
p38	PK	7214	23.7	63.9	42.3	93.3	1.8	1.5
MKK4	PK-MKK4	3655	35.5	63.9	30.2	29.2	0.9	0.5
Hem	PK-MKK7	4608	0.0	24.6	6.0	23.3	*0.5*	1.0
JNK	PK	3082	23.7	7.4	24.2	40.8	1.0	5.5
FOS	bZIP	4904	23.7	95.8	48.3	177.9	2.0	1.9
Jra	bZIP	13290, 13291	71.1	186.7	108.7	186.6	1.5	1.0
Aop	SAM-PNT, ETS	1136	0.0	125.3	36.2	154.6	*3.1*	1.2
MASK	ANK, KH-I	225, 4036	106.6	206.3	157.1	265.4	1.5	1.3
		JAK-STAT pathway						
Domeless	SH2	7557, 9588	11.8	27.0	30.2	17.5	2.6	0.7
JAK/Hopscotch	PTK, SH2, B41	20	272.5	498.6	241.7	250.8	0.9	0.5
STAT	SH2, Protein	2221, 14109	106.6	135.1	24.2	102.1	0.2	0.8
PIAS	Interacting, MIZ/SP RING	602	23.7	113.0	42.3	119.6	1.8	1.1

NIH-PA Author Manuscript NIH-PA Author Manuscript NIH-PA Author Manuscript

Gene name	Domain(s)	Contig(s)	nCF	nCH	nIF	nIH	nIF:nCF	nIH:nCH
SOCS	SH2, SOCS box	1187, 6886	11.8	159.6	30.2	125.4	2.6	0.8
Stam	VHS	13543	0.0	19.6	6.0	20.4	*0.5*	1.0

*Genes reported by Zhang et al. (2011) are *underlined*. nCF, nCH, nIF and nIH are normalized read numbers and, for genes with two or more contigs, they represent the total values. When nCF = 0, adjusted NRN for nIF (*italics*) is calculated as nIF × 825/9775; when nCH = 0, adjusted NRN for nIH (*italics*) is calculated as nIH × 3980/9775.

NIH-PA Author Manuscript
NIH-PA Author Manuscript
NIH-PA Author Manuscript