Bacterial infection, immune responses, and autophagy in *lutzomyia longipalpis* sand flies

by

Matthew C. Heerman

B.S., Kansas State University, 2007
B.S., Kansas State University, 2007
M.S., Kansas State University, 2012

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Entomology
College of Agriculture

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2016
Abstract

Microbial communities residing within the midgut of insect vectors play a critical role in the response to various zoonotic and human pathogens, and can directly alter the development and survival of the insects. Sand flies are the primary vector of Leishmania, the causative pathogen of leishmaniasis, a neglected tropical disease. Sand flies acquire many microbes from the soil where immature stages develop until emergence as adults. Gram-negative Pantoea agglomerans and gram-positive Bacillus subtilis are two bacteria commonly associated with sand fly populations. Here, I demonstrated that an EGFP- and a GFP-expressing version of these two bacteria localize to different compartments of the midgut; a phenomenon that is achieved, in part, to pH differences found across the length of the gut. Additionally, P. agglomerans is able to selectively induce midgut epithelial apoptosis while B. subtilis does not. This is accompanied by differential immune and homeostasis responses to both bacteria highlighted by immune pathway suppression via the Poor Immune Response upon Knock-in (Pirk) gene. These effects may actually be representative of a broader type of response to bacterial infection that might be present across several insect species. Finally, I demonstrated that during metamorphosis the sand fly relies, at least in part, upon the activation of multiple genes from the autophagy pathway to aid in generating adult tissues. More specifically, I demonstrate, using microscopy, the presence of ATG6 in the cytoplasm of developing midgut epithelial cells of the sand fly pupae.
Bacterial infection, immune responses, and autophagy in *lutzomyia longipalpis* sand flies

by

Matthew C. Heerman

B.S., Kansas State University, 2007
B.S., Kansas State University, 2007
M.S., Kansas State University, 2012

A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Entomology
College of Agriculture

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2016

Approved by:
Co-Major Professor
Dr. Marcelo Ramalho-Ortigao

Approved by:
Co-Major Professor
Dr. Kun Yan Zhu
Copyright

© Matthew C. Heerman 2016.
Abstract

Microbial communities residing within the midgut of insect vectors play a critical role in the response to various zoonotic and human pathogens, and can directly alter the development and survival of the insects. Sand flies are the primary vector of *Leishmania*, the causative pathogen of leishmaniasis, a neglected tropical disease. Sand flies acquire many microbes from the soil where immature stages develop until emergence as adults. Gram-negative *Pantoea agglomerans* and gram-positive *Bacillus subtilis* are two bacteria commonly associated with sand fly populations. Here, I demonstrated that an EGFP- and a GFP-expressing version of these two bacteria localize to different compartments of the midgut; a phenomenon that is achieved, in part, to pH differences found across the length of the gut. Additionally, *P. agglomerans* is able to selectively induce midgut epithelial apoptosis while *B. subtilis* does not. This is accompanied by differential immune and homeostasis responses to both bacteria highlighted by immune pathway suppression via the Poor Immune Response upon Knock-in (Pirk) gene. These effects may actually be representative of a broader type of response to bacterial infection that might be present across several insect species. Finally, I demonstrated that during metamorphosis the sand fly relies, at least in part, upon the activation of multiple genes from the autophagy pathway to aid in generating adult tissues. More specifically, I demonstrate, using microscopy, the presence of ATG6 in the cytoplasm of developing midgut epithelial cells of the sand fly pupae.
# Table of Contents

List of Figures ..................................................................................................................... ix
List of Tables ........................................................................................................................ xiii

Chapter 1 - Literature Review............................................................................................. 1
  1.1. Paratransgenesis ......................................................................................................... 1
  1.2. Commensal Dipteran Microbial Communities and Immunity .................................... 1
  1.3. Immunity and Autophagy .......................................................................................... 4
  1.4. Autophagy and development ....................................................................................... 5
  1.5. Research Goals and Objectives ................................................................................ 7
    1.5.1. Determination of distribution and persistence of paratransgenic bacterial candidates in the midgut of sand fly larvae ............................................................................. 7
    1.5.2. Determination of effects of bacterial infection on midgut epithelium cells and immune responses .................................................................................................................. 7
    1.5.3. Investigation of the role of autophagy during larval infection and development to adult stage ...................................................................................................................... 8

Chapter 2 - Bacterial Infection and Immune Responses in the Larval Midgut of *Lutzomyia longipalpis* Sand Flies ........................................................................................................... 9
Abstract ................................................................................................................................. 9

  2.1. Introduction ................................................................................................................ 9

  2.2. Materials and Methods ............................................................................................. 11
    2.2.1. Sand fly colony maintenance ................................................................................. 11
    2.2.2. Culturing and feeding of EGFP-expressing *B. subtilis* and GFP-expressing *P. agglomerans* .................................................................................................................. 11
    2.2.3. Immunocytochemistry .......................................................................................... 13
    2.2.4. Larval midgut pH ................................................................................................ 14
    2.2.5. Effect of pH on *in vitro* growth of EGFP-expressing *B subtilis* and GFP-expressing *P. agglomerans* .............................................................................................................. 14
    2.2.6. RNA extraction and reverse transcription .............................................................. 15
    2.2.7. Quantitative real time PCR ................................................................................ 15
2.3. Results

2.3.1. Bs and Pa localize to different regions of the sand fly larvae midgut ........................................ 16
2.3.2. Infection of the sand fly larval midgut by Pa or Bs leads to differential damage of the larval midgut epithelium ................................................................. 18
2.3.3. Effect of feeding agar on larval development ................................................................. 19
2.3.4. Bacterial feeding leads to differential expression profiles in sand fly larvae ........ 19

2.4. Discussion ................................................................................................................................. 20

Chapter 3 - The Role of Autophagy during Metamorphosis of the Sand Fly Lutzomyia longipalpis ................................................................. 41

Abstract ........................................................................................................................................ 41
3.1. Introduction ............................................................................................................................. 41

3.2. Material & Methods ............................................................................................................ 43
3.2.1. Sand fly colony maintenance ...................................................................................... 43
3.2.2. RNA extraction and reverse transcription ............................................................ 44
3.2.3. Real time quantitative PCR ....................................................................................... 44
3.2.4. Sequence alignments, phylogeny, and putative ecdysone response element prediction ................................................................................................................................. 45
3.2.5. Midgut immunocytochemistry & measurements .................................................... 45

3.3. Results .................................................................................................................................... 46
3.3.1. Change in expression profile of ATG1, ATG6, ATG8, and Vein during development ........................................................................................................................................... 46
3.3.2. Midgut specific autophagy during pupation ................................................................ 48
3.3.3. Evidence linking the role of ecdysone to the activation of autophagy .................. 49

3.4. Discussion ................................................................................................................................. 49

Chapter 4 - IMD Recognition, Toll Pathway, and the effect of Pirk Knockdown in Infected Lutzomyia longipalpis Larvae .................................................................................. 77

Abstract ........................................................................................................................................ 77
4.1. Introduction ............................................................................................................................. 78

4.2. Material & Methods ............................................................................................................ 79
4.2.1. Sand fly colony maintenance ...................................................................................... 79
4.2.2. Double stranded RNA synthesis and injection ........................................................ 79
List of Figures

Figure 2.1. Infection of sand fly larva midgut by *B. subtilis* and *P. agglomerans* .................. 31
Figure 2.2. pH of the sand fly larval gut. .................................................................................. 32
Figure 2.3. Effect of pH on in vitro growth of *B. subtilis* and *P. agglomerans* ..................... 33
Figure 2.4. Confocal images of *B. subtilis* and *P. agglomerans* infection of sand fly larvae midguts ......................................................................................................................... 34
Figure 2.5. mRNA expression profiles of 3rd instar *L. longipalpis* larvae post infection with *B. subtilis* and *P. agglomerans* .............................................................................................................. 35
Figure 2.6. Larval feeding of bacterial lawn ................................................................................. 36
Figure 2.7. Infection of sand fly larval midgut by *B. subtilis* or *P. agglomerans* .................... 37
Figure 2.8. Paraquat-induced apoptosis in sand fly larval midguts. ........................................... 38
Figure 2.9. Changes in sand fly larval gut length and width caused by diet ................................. 39
Figure 2.10. Localization of *B. subtilis* ..................................................................................... 40
Figure 2.11. Apical localization of *P. agglomerans* on midgut epithelia ................................. 40
Figure 3.1. mRNA expression profiles for feeding larvae (L3), non-feeding larvae (L4), early pupae (Pupa), and adult (Adult) sand flies .................................................................................................................. 57
Figure 3.2. Confocal images for ATG6 showing the anterior midgut of newly formed pupae... 58
Figure 3.3. Confocal images depicting the anterior sand fly midgut at different developmental stages stained with ATG6 antibodies ........................................................................................................ 59
Figure 3.4. Confocal images depicting the anterior sand fly midgut at larval, pupal, and adult developmental stages stained with ATG8 antibodies ................................................................. 60
Figure 3.5. Possible promoter sequence of region upstream of *ATG1* .................................. 61
Figure 3.6. The midgut morphological changes in the 3rd instar (L3), 4th instar (L4), pupa, and adult developmental stages .................................................................................................................. 62
Figure 3.7. Axial ratio for the midgut (L/W) for L3, L4, pupa, and adult sand flies..................... 63
Figure 3.8. Clustal Omega protein alignment of ATG1 sequences from *L. longipalpis*, *P. papatasi*, *D. melanogaster*, *A. aegypti*, *A. gambiae*, *C. quinquefasciatus*, and *G. morsitans*. ................................................................................................................................. 66
Figure 3.9. Neighbor-joining tree to infer the evolutionary relationship of ATG1 sequences using *T. castaneum* (TcATG1) as an outgroup ......................................................................................................... 67
Figure 3.10. Clustal Omega protein alignment of ATG6 sequences from L. longipalpis, P. papatasi, D. melanogaster, A. aegypti, A. gambiae, C. quinquefasciatus, and G. morsitans.

Figure 3.11. Neighbor-joining tree to infer the evolutionary relationship of ATG6 sequences using A. pism (ApATG6) as an outgroup.

Figure 3.12. Clustal Omega protein alignment of ATG8 sequences from L. longipalpis, P. papatasi, D. melanogaster, A. aegypti, A. gambiae, C. quinquefasciatus, and G. morsitans.

Figure 3.13. Neighbor-joining tree to infer the evolutionary relationship of ATG8 sequences.


Figure 3.15. Neighbor-joining tree to infer the evolutionary relationship of Vein sequences.

Figure 4.1. mRNA expression levels for L3 larvae fed normal diet (chow), control agar (Con), B. subtilis agar (Bs), and P. agglomerans agar (Pa).

Figure 4.2. mRNA expression levels for L3 larvae fed normal diet (chow), control agar (Con), B. subtilis agar (Bs), and P. agglomerans agar (Pa).

Figure 4.3. mRNA expression levels for L3 larvae that were uninjected, injected with dsGFP, or injected with dsPirk RNA.

Figure 4.4. mRNA expression levels for L3 larvae that were uninjected, injected with dsGFP, or injected with dsPirk RNA.

Figure 4.5. mRNA expression levels for L3 larvae that were uninjected, injected with dsGFP, or injected with dsPirk RNA.

Figure 4.6. mRNA expression levels for L3 larvae that were uninjected, injected with dsGFP, or injected with dsPirk RNA.

Figure 4.7. mRNA expression levels for L3 larvae that were uninjected, injected with dsGFP, or injected with dsPirk RNA.
Figure 4.8. mRNA expression levels for L3 larvae that were uninjected, injected with dsGFP, or injected with dsPirk RNA.

Figure 4.9. mRNA expression levels for L3 larvae that were uninjected, injected with dsGFP, or injected with dsPirk RNA.

Figure A.1. Putative CDS for mRNA encoding *L. longipalpis* Vein.

Figure A.2. Putative CDS for mRNA encoding *L. longipalpis* Pirk.

Figure A.3. Neighbor-joining tree to infer the evolutionary relationship of Pirk sequences.

Figure A.4. Clustal Omega protein alignment of Pirk sequences from *L. longipalpis*, *P. papatasi*, *D. melanogaster*, *D. buskii*, *A. albopictus*, *T. castaneum*, *S. oryzae*, and *L. migratoria*.

Figure A.5. Putative CDS for mRNA encoding *L. longipalpis* ATG6.

Figure A.6. Putative CDS for mRNA encoding *L. longipalpis* Attacin.

Figure A.7. Neighbor-joining tree to infer the evolutionary relationship of Attacin sequences.

Figure A.8. Clustal Omega protein alignment of Attacin sequences from *L. longipalpis*, *P. papatasi*, *D. melanogaster*, *A. aegyti*, and *T. castaneum*.

Figure A.9. Putative CDS for mRNA encoding *L. longipalpis* ATG1.

Figure A.10. Putative CDS for mRNA encoding *L. longipalpis* ATG8.

Figure A.11. Putative CDS for mRNA encoding *L. longipalpis* Duox.

Figure A.12. Neighbor-joining tree to infer the evolutionary relationship of Duox sequences.

Figure A.13. Clustal Omega protein alignment of Duox sequences from *L. longipalpis*, *P. papatasi*, *D. melanogaster*, *A. gambiae*, *A. aegyti*, and *C. quinquefasciatus*.

Figure A.14. Putative CDS for mRNA encoding *L. longipalpis* USP36.

Figure A.15. Neighbor-joining tree to infer the evolutionary relationship of USP36 sequences.

Figure A.16. Clustal Omega protein alignment of USP36 sequences from *L. longipalpis*, *P. papatasi*, *D. melanogaster*, *A. gambiae*, *A. aegyti*, and *C. quinquefasciatus*.

Figure A.17. Putative CDS for mRNA encoding *L. longipalpis* IMD.

Figure A.18. Neighbor-joining tree to infer the evolutionary relationship of IMD sequences.

Figure A.19. Clustal Omega protein alignment of IMD sequences from *L. longipalpis*, *P. papatasi*, *D. melanogaster*, *A. gambiae*, *A. aegyti*, *C. quinquefasciatus*, *T. Castaneum*, and *G. morsitans*.

Figure A.20. Putative CDS for mRNA encoding *L. longipalpis* Nos.
Figure A.21. Neighbor-joining tree to infer the evolutionary relationship of Nos sequences.... 142
Figure A.22. Clustal Omega protein alignment of Nos sequences from L. longipalpis, P. papatasi, D. melanogaster, A. gambiae, and A. aegyti.......................................................... 145
Figure A.23. Putative CDS for mRNA encoding L. longipalpis Domeless. ...................... 147
Figure A.24. Neighbor-joining tree to infer the evolutionary relationship of Domeless sequences. ........................................................................................................................................................................... 148
Figure A.25. Clustal Omega protein alignment of Domeless sequences from L. longipalpis, P. papatasi, D. melanogaster, A. gambiae, A. aegyti, and C. quinquefasciatus .......... 152
Figure A.26. Putative CDS for mRNA encoding L. longipalpis IMPer ............................ 154
Figure A.27. Neighbor-joining tree to infer the evolutionary relationship of IMPer sequences. 155
Figure A.28. Clustal Omega protein alignment of IMPer sequences from L. longipalpis, P. papatasi, D. melanogaster, A. gambiae, A. aegyti, and C. quinquefasciatus .................. 159
List of Tables

Table 2.1. *P. agglomerans* and *B. subtilis* infection rates................................................................. 27
Table 2.2. Persistence of bacterial infection between continuous and non-continuous feedings. 28
Table 2.3. Colony forming units (CFU) from continuous and non-continuous fed larvae........ 29
Table 2.4. List of primer sequence used in this study............................................................................. 30
Table 3.1. The accession numbers, identity, score, and expected value for each query protein (*D. melanogaster*) against each subject protein (*L. longipalpis*). ................................................................. 54
Table 3.2. The sequence of RT-qPCR primers used for each transcript measured in this study.. 55
Table 3.3. The species names, gene symbols, accession numbers, and databases for all sequences used for multiple sequence alignment and phylogenetic analysis in this study. *H. sapiens* ATG8 sequence is not derived from a database and may be procured from the company website via that catalog number................................................................. 56
Table 4.1. The accession numbers, identity, score, and expected value for each query protein (*D. melanogaster*) against each subject protein (*L. longipalpis*). ................................................................. 89
Table 4.2. The sequence of RT-qPCR primers used for each transcript measured in this study.. 90
Table 4.3. The species name, gene symbol, accession number, and database for all sequences used for multiple sequence alignment and phylogenetic analysis. ............................................. 91
Chapter 1 - Literature Review

1.1. Paratransgenesis

Current genetic control programs and techniques have been summarized concisely in a recent review [1]. This covers all topics ranging from transgenesis to paratransgenesis. Herein I choose to focus on strategies in place or in development directed towards the use of paratransgenesis as a means for vector and disease control as is currently reviewed [2].

Paratransgenesis is defined as using genetically modified symbiotic bacteria and/or fungi to express molecules that will abrogate the ability of the vector to transmit pathogens. This is a complicated process that involves selecting a microbe that can persist, or be driven into the vector, to diminish the vector competence of the insect without increasing its susceptibility to other pathogens. One example of Chryseobacterium meningosepticum showed that the bacteria were able to colonize and become the predominant gut microbe in Anopheles gambiae [3]. Furthermore Escherichia coli expressing anti Plasmodium antibodies were able to block Plasmodium berghei progression in Anopheles stephensi [4]. Wolbachia is a long standing popular intracellular candidate for vector control [5]. However, there is one reported case in Culex tarsalis where Wolbachia actually increased the susceptibility of the vector for West Nile virus [6]. Taking these things into account we wished to evaluate sand fly commensal bacteria for their ability to be putative paratransgenic platforms as more completely described below.

1.2. Commensal Dipteran Microbial Communities and Immunity

Bacterial symbionts significantly influence many aspects of the physiology of their host. In insects, both pathogenic and non-pathogenic bacteria have been shown to modulate immune response, homeostasis, development, and overall health of midgut physiology for both larval and
adult stages. Both Gram-positive (G+) and Gram-negative (G-) bacteria are commonly associated with the midgut tissue of Diptera, including several disease vectors. Many bacilli and enterobacter such as *Lactobacillus* and *Pantoea* have been identified from the midgut and other tissues of the fruit fly *Drosophila melanogaster* [7]. In field collected *Anopheles stephensi*, *Anopheles gambiae*, and *Aedes aegypti* from laboratory colonies, a number of G+ and G- bacteria were identified [8-11]. Among the G- bacteria, *Pantoea agglomerans* was also the most common genus identified from all cultivable bacteria in both male and female *Aedes albopictus* collected from two out of four sites in Madagascar [12]. In sand flies, several studies have focused on regional and potential species-specific variability in the microbial community of both *Phlebotomus* and *Lutzomyia* species. *P. agglomerans* and *Bacillus* spp. were commonly found in both natural and laboratory-reared sand fly populations. *Bacillus subtilis* and non-*Pantoea* members of the Enterobacter family were shown to be present in populations of *Phlebotomus papatasi* from, India, Turkey, Tunisia, and Egypt [13, 14]. *Bacillus* spp, *Serratia marcescens*, and *P. agglomerans* were identified in natural and laboratory populations of adult *Lutzomyia longipalpis* [15-18]. In laboratory colonies, Bacilli and Enterobacter were also identified in the larval stages of *L. longipalpis* [19, 20]. However, to date, most studies were focused on microorganisms present in the adult sand fly with little or no attention paid to physiological effects brought about by these microorganisms on developing stages. Here, we investigated the effects of EGFP-expressing *B. subtilis* (*Bs*) or GFP-expressing *P. agglomerans* (*Pa*) on the midgut innate immunity and epithelial homeostasis of 3rd instar larvae of *L. longipalpis*. With particular respect paid toward the ability of the immature vector to undergo metamorphosis and retain the bacteria tested.
The distribution of these bacteria within the midgut of sand fly larvae is likely driven in part by the gut pH gradient [21], and perhaps may display a cytotoxic effect on the midgut epithelium by infection with Pa or Bs like what is observed in Drosophila [22, 23]. These two phenomenon were investigated in these studies. Additionally, we were interested in the activation and regulation of the immune response in the sand fly. It has been previously shown that Caspar is involved in negatively regulating the immunodeficiency response in sand flies, and knocking it down via RNAi leads to lower parasite numbers in adult flies [18]. Here, we investigate a differential and suppressive response to infection with respect to G+ and G- bacteria, possibly influenced by the gene encoding another negative regulator of immunity, "poor immune response upon knock-in" or Pirk. Up-regulation of Pirk transcript level would theoretically lead to a significant depletion of the transcripts encoding the antimicrobial peptides.

There is no work conducted in sand flies to measure any phenomenon related to Toll pathway regulation or IMD recognition of foreign organisms and molecules. Additionally, the agar feeding regimens used to determine the effect of Pirk on effector molecule expression are quite different than the natural detritus material that sand flies are generally exposed to in the lab and nature. While there is data on the efficiency of RNAi mediated knockdown in adult sand flies [18, 24], no published data is currently aimed to elucidate the efficiency of RNAi mediated knockdown of sand fly transcripts in the immature larval stages. Here we investigate the negative regulator of Toll signaling, Cactus, as a marker for measuring Toll pathway activity. Although Cactus activity is generally described in hemocyte interactions with gram-positive bacteria and fungal organisms, or dorso-ventral development [25, 26], we wished to see if it may have so effect in the midgut against gram-positive B. subtilis. PGRP-LB is a well-studied recognition
molecule for the IMD pathway, known for its ability to protect *W. wigglesworthia* within the bacteriome of the tsetse fly midgut [27].

While a primary focus of these studies is to determine the viability of these bacteria as a paratransgenic platform, we also wanted to gain insights into how these microbes are related to intestinal immunity (explained above), and how they may engage in cross talk with other pathways of interest (explained below) that drive physiological processes and development.

1.3. Immunity and Autophagy

Growing numbers of studies are aimed at teasing out the complex relationships between immune and autophagic processes. The term ‘antimicrobial autophagy’ was recently used to describe this process and is reviewed thoroughly in Moy et al. 2013 [28], and is categorized as an innate immune response in *D. melanogaster*. It is interesting to see a process typically implicated in development and nutritional status play a role in immunity. Within *D. melanogaster* it is demonstrated to mediate a Toll and IMD independent autophagic response against the intracellular pathogenic bacteria *Listeria* [29]. Combating vesicular stomatitis virus in the fruit fly also occurs through a nutritional sensing phosphatidyl inosital kinase autophagic response (PI3K) [30]. Beyond the realm of insects lays the potential mechanisms and benefits that must be understood for higher animals, specifically mammals. The role of autophagy in immunity, aging, inflammation, and development is reviewed for higher animals in Mizushima et al. 2008 [31].

With respect to our research, we were interested in evaluating the potential of autophagy during immune responses to our putative paratransgenic commensal bacteria. Moreover, we wanted to probe at the putative linkage of autophagy to development of sand flies from immature stages to the adult life cycle (explained below).
1.4. Autophagy and development

Autophagy is a complex process that involves the processing and recycling of cellular macromolecules and larger organelles. This can be associated with a multitude of phenomenon including stresses like starvation, or intriguingly the decision of a cell to undergo programmed cell death or survival. Within higher mammals this process is intimately linked to numerous disease pathologies such as neurological degeneration and cancer. Interestingly, there is mounting evidence to suggest that autophagy is paramount in regulating the varied and finely orchestrated developmental events that occur in various insect life stages. A majority of the work involves the use of the model organism *Drosophila melanogaster*. Here the role of autophagy in recycling tissues and cellular reprogramming is apparent. During fruit fly embryogenesis autophagy plays a role in removing excess sera from developing insects [32]. Additionally, degradation of larval salivary glands during metamorphosis is required to produce the final adult tissues, and is autophagy dependent [33, 34]. Over-proliferation of the neuromuscular junction of larval stages of the fly is prevented via functioning autophagy [35]. Breakdown of fat body during development of not only fruit flies, but *Bombyx mori*, is also a crucial process that has been linked to autophagy [36, 37].

Larval midgut removal and the tissue remodeling and differentiation of the adult gut are a hallmark of metamorphosis. Autophagy plays a center role in nutrient recycling of midgut cells fated to undergo apoptosis and survival of gut cells that will ultimately comprise the adult tissue. A number of more recent studies implicate the function of autophagy in both pro-death and pro-survival in fruit flies and silk worms [38-41].

Within hematophagous insects there are few studies delineating the role of autophagy during development. Studies suggest that autophagy in the fat body of *Aedes aegypti* is important
in maintaining gonadotrophic cycles [42]. Additionally, the presence of electron dense autophagosomes has been detected in the developing midgut of the mosquito during metamorphosis [43].

Ecdysone is the major hormone employed by insects to progress through various developmental stages. Its complex set of interactions with ecdysone receptors and primary response genes is vital in finely tuning spatial and temporal developmental events. Briefly and specifically, E93 and Broad (Br-C) have been identified as 2 primary ecdysone response genes that signal the removal of the larval midgut of *D. melanogaster* before completion of the adult gut via autophagy [44-46]. Further, it is suggested that the action of these molecules may be conserved across holometabolous insects [47]. However, insights into the direct mode of action of ecdysone, primary response genes, and autophagy in affecting successful metamorphosis have only very recently begun to be understood. In *B. mori* it has been shown that autophagy-related 1 (ATG1) is a response gene to ecdysone. E93 is able to bind to ATG1 promoter region and induce its expression [48]. While these finding help in understanding how the larval gut may be removed during pupation, little is known about how the new adult gut is formed. It is suggested that Vein, an epidermal growth factor, may be involved in signaling intestinal stem cells of adult *D. melanogaster* to differentiate into healthy adult epithelial cells [23, 49, 50].

To our knowledge there is no study published on the role of autophagy and its possible response to ecdysone, or the place of epidermal growth factors during metamorphosis of the sand fly *L. longipalpis*. Here we investigate whether the key components of the autophagy pathway are conserved in sand flies and expressed during metamorphosis of the fly. We also take a preliminary glimpse into the midgut specificity for expression of an autophagy component in the cytoplasm of developing epithelial cells. Additionally, we investigate a possible candidate for the
generation of an adult midgut tissue in sand flies. Finally, tenable evidence supportive of a link between ATG1 and ecdysone is provided.

1.5. Research Goals and Objectives

1.5.1. Determination of distribution and persistence of paratransgenic bacterial candidates in the midgut of sand fly larvae

To evaluate the potential of *P. agglomerans* and *B. subtilis* as delivery systems for paratransgeneic microbes, we plan to examine GFP-labeled bacteria in the midgut of the sand fly larvae by using an agar based feeding system. Owing to their GFP or EGFP (respectively) signals, we expect to visualize the distribution of the bacteria within the alimentary canal using confocal microscopy. Additionally, we propose to determine midgut pH gradients, which may influence the localization of bacteria *in vivo* for the entire sand fly larval alimentary canal by using two different pH indicators. We plan to further confirm this via *in vitro* measurements of bacterial growth on selected pH media. The persistence of *P. agglomerans* and *B. subtilis* in the midgut will be determined by confocal microscopy visualization or by counting bacteria colonies on selected media. Our goal is to determine basic physiological traits exhibited within the gut of the fly larvae, and determine their ability to survive therein.

1.5.2. Determination of effects of bacterial infection on midgut epithelium cells and immune responses

To evaluate possible effects of *P. agglomerans* and *B. subtilis* on the midgut of sand fly larvae, we propose to employ qRT-PCR to determine the mRNA levels of immune effector and regulatory molecules within the gut. To visualize the direct effects of microbial presence on the
midgut epithelium we plan to use confocal microscopy to see cells specifically undergoing programmed cell death. We also propose to evaluate the efficiency of RNAi targeting candidate genes by injection in larvae, which have not been explored in sandy fly larvae. Our goal is to understand basic physiological effects of *P. agglomerans* and *B. subtilis* infection on developing larval gut tissues, and analyze the efficiency of RNAi as a common reverse genetic tool on larval stages.

1.5.3. Investigation of the role of autophagy during larval infection and development to adult stage

To explore the role of autophagy during bacterial infection or metamorphosis of the sand fly larvae we propose to measure relative transcript levels of selected markers of autophagy signaling. We will dissect gut tissues and specifically visualize autophagic expression within the epithelium of the insect undergoing metamorphosis by using confocal microscopy. The goal of these studies is to better understand the role of bacterial infection with respect to immune response and autophagy during the larval developmental stages. Moreover, we expect to gain insights as to how autophagy may play crucial roles in the metamorphosis of the sand flies as a disease vector.
Chapter 2 - Bacterial Infection and Immune Responses in the Larval Midgut of *Lutzomyia longipalpis* Sand Flies

Abstract

The midgut microbial community in insect vectors of disease is crucial for an effective immune response against infection with various human and animal pathogens. Depending on the aspects of their development, insects can acquire microbes present in soil, water, and plants. Sand flies are major vectors of leishmaniasis, and shown to harbor a wide variety of Gram-negative and Gram-positive bacteria. Sand fly larval stages acquire microorganisms from the soil, and the abundance and distribution of these microorganisms may vary depending on the sand fly species or the breeding site. Here, we assess the distribution of two bacteria commonly found within the gut of sand flies, *Pantoea agglomerans* and *Bacillus subtilis*. We demonstrate that these bacteria are able to differentially infect the larval digestive tract, and regulate the immune response in sand fly larvae. Moreover, colonization of the gut is driven, at least in part, a gradient of pH present in the gut.

2.1. Introduction

Bacterial symbionts significantly influence many aspects of the physiology of their host. In insects, both pathogenic and non-pathogenic bacteria have been shown to modulate immune response, homeostasis, development, and overall health of midgut physiology for both larval and adult stages. Both Gram-positive (G+) and Gram-negative (G-) bacteria are commonly associated with the midgut tissue of Diptera, including several disease vectors. Many bacilli and enterobacter such as *Lactobacillus* and *Pantoea* have been identified from the midgut and other tissues of the fruit fly *Drosophila melanogaster* [7]. In field collected *Anopheles stephensi,*
Anopheles gambiae, and Aedes aegypti from laboratory colonies, a number of G+ and G-
bacteria were identified [8-11]. Among G-, Pantoea agglomerans was also the most common
genus identified from all cultivable bacteria in both male and female Aedes albopictus collected
from two out of four sites in Madagascar [12]. In sand flies, several studies have focused on
regional and potential species-specific variability in the microbial community of both
Phlebotomus and Lutzomyia species. P. agglomerans and Bacillus spp. were commonly found in
both natural and laboratory-reared sand fly populations. Bacillus subtilis and non-Pantoea
members of the Enterobacter family were shown to be present in populations of Phlebotomus
papatasi from, India, Turkey, Tunisia, and Egypt [13, 14]. Bacillus spp, Serratia marcescens,
and P. agglomerans were identified in natural and laboratory populations of adult Lutzomyia
longipalpis [15-18]. In laboratory colonies, Bacilli and Enterobacter were also identified in the
larval stages of L. longipalpis [19, 20]. However, to date, most studies were focused on
microorganisms present in the adult sand fly with little or no attention paid to physiological
effects brought about by these microorganisms on developing stages. Here, we investigated the
effects of EGFP-expressing B. subtilis (Bs) or GFP-expressing P. agglomerans (Pa) on the
midgut innate immunity and epithelial homeostasis of 3rd instar larvae of L. longipalpis.
Additionally, we determined that the distribution of these bacteria within the midgut of sand fly
larvae is in part driven by the gut pH, and demonstrate a cytotoxic effect on the midgut
epithelium by infection with Pa. Strikingly, we show evidence suggesting a differential and
suppressive response to infection with respect to G+ and G- bacteria, likely influenced by the
gene encoding the negative regulator of immunity, "poor immune response upon knock-in" or
Pirk. Up-regulation of Pirk transcript level leads to a significant depletion of the transcripts
encoding the antimicrobial peptide Attacin and the immunomodulatory peroxidase IMPer. With
respect to the organization of these microbes across the length and distal spacing of the midgut, G+ Bs were distributed throughout the entire alimentary canal in larvae, whereas G- Pa were found primarily in the posterior midgut. Our results strongly suggest that the range of pHs present within the sand fly larval gut likely is a driving force defining the ability of certain bacteria such as P. agglomerans to infect areas of the gut. The results presented here may have implications beyond the sand fly system and may explain how the distribution (and possibly colonization) of bacteria and other microbes may occur within the guts of insects.

2.2. Materials and Methods

2.2.1. Sand fly colony maintenance

L. longipalpis (Jacobina strain – LLJB) were the colony maintained in the Department of Entomology, Kansas State University. Larvae were maintained in 250 or 500 ml plastic jars (Nalgene) with an approximately 2 cm-thick bed made of dental plaster (Schein), and fed on either larval chow (a mixture of 50 % rabbit droppings and 50 % rabbit food) or on 1.5 % agar in LB, with or without the (E)GFP-expressing bacteria (see below).

2.2.2. Culturing and feeding of EGFP-expressing B. subtilis and GFP-expressing P. agglomerans

EGFP-expressing B. subtilis (strain 1012 transformed with plasmid pAD43-25, obtained from the Bacillus Genetic Stock Center) and GFP-expressing P. agglomerans (strain EPA-E325 transformed with plasmid pT-3078-5, a gift from Dr. David Lampe) were cultured at 30 °C overnight in LB medium supplemented with 5 µg/ml chloramphenicol (Alfa Aesar, A Jonson Mattey Co.) or with 50 µg/ml Carbenicillin (Teknova), respectively. Bacterial cultures were
centrifuged at 2500 rpm for 20 minutes at room temperature and the pellets were washed twice with 1X PBS. Bacteria were then suspended in PBS for a final concentration of $10^9$ bacteria in 50-to-80 µl that was spread on a plate containing a thin layer (2-3 mm thick) of LB-agar (no antibiotics) and grown overnight. The following day, fluorescence of the bacterial lawn was confirmed and the LB-agar was cut into 3-to-5 mm$^2$ pieces to be fed to early L3 *L. longipalpis* larvae (depicted in Supporting Figure 2.6.), and lawns were replaced every 48 h. Prior to feeding on the LB-agar with *Bs* or *Pa*, all larvae were starved for 6 to 10 hours to allow for excretion of midgut contents, and rinsed in sterile water. As controls, L3 larvae were fed on LB-agar plus 5 mM paraquat, an herbicide that strongly induces apoptosis [51], and on LB-agar plus kanamycin (50 µg/ml) (we were unable to use plain LB-agar due to contamination). Larvae fed *ad libitum* for up to 48 h at 26 °C and 80 % humidity, with a 12:12 h light-dark cycle. Groups of (n=20) larvae were collected at 12, 24, 36, and 48 h post feeding with three biological replicates. Food intake was determined by examining each larva under a dissecting microscope (10X). All larval feedings were done according to feeding groups using 500 ml Nalgene pots with a 2-3 cm layer of dental cement.

Alternatively, larvae were fed for 12 h on bacteria-containing LB-agar and transferred to pots with plain LB-agar (no bacteria and no antibiotics). Larvae in groups of 3 to 5 were collected every 3 h and assessed for GFP signal using a Zeiss confocal LSM microscope. CFUs were also measured for larvae collected at 12 h and 24 h, by surface sterilizing each larva, dissecting and grinding each whole gut using a hand-held homogenizer in 60 µl 1X PBS, and plating the homogenate on selective media (5 µg/ml chloramphenicol for *B. subtilis* or 50 µg/ml carbenicillin for *P. agglomerans*) and incubating at 28 °C to 30 °C.
To assess for any effects of diet on midgut development, the length and width of the larval midgut were measured using the LSM 510 using the software ZEISS LSM Image Browser (Zeiss International). Midgut length was determined by measuring from the beginning of the anterior midgut to the posterior region of the midgut. Width measurements were obtained from three regions of each midgut: the anterior (ant), the middle (mid), and the posterior (pos) regions.

### 2.2.3. Immunocytochemistry

Whole guts from *L. longipalpis* L3 larvae were dissected (n=3) from three separate treatments of L3 into PBS and fixed for 20 minutes at room temperature with 4 % paraformaldehyde in PBS. Tissues were washed 4 times for 30 minutes with PBS containing 0.3 % Triton X-100 (PBST), then blocked with PBS containing 1 % bovine serum albumin for 30 minutes at room temperature. Tissues were then incubated overnight at 4 °C with primary antibodies for rabbit anti cleaved caspase3 (Cell Signaling Technology) diluted 1:500 in PBST. Tissues were washed 3 times for 30 minutes with PBST, and incubated overnight at 4 °C with Alexa Fluor® 594 goat anti-rabbit (Invitrogen) diluted 1:1000 in PBST. Tissues were washed 3 times for 30 minutes with PBST, and nuclei were stained for 5 minutes with 10 µg/mL of DAPI (Invitrogen). Samples were mounted in Vectashield® (Vector Laboratories) anti-photo bleaching reagent, and images were obtained with a LSM 510 confocal microscope using the software ZEISS LSM Image Browser. In addition, measurements of larval guts length and width were obtained for each feeding treatment using the LSM510 confocal microscope.
2.2.4. Larval midgut pH

We assessed the pH within the midgut by feeding *L. longipalpis* L3 larvae with LB-agar containing 0.4 % of the pH indicators Bromothymol blue and Phenol red. LB-agar medium adjusted to pH 7 added prior to sterilization. Each indicator agar medium was fed to a group to 50 L3 larvae following six hours of food deprivation. Feeding of the larvae was performed by placing the larvae and fragments of approximately 3-5 mm$^2$ of the dye-containing agar inside a 500 ml Nalgene pot with a 1 cm layer of plaster, maintained at room temperature and with a relative humidity of 80-85 %. Larvae were allowed to feed *ad libitum* for 20 hours. The pH indicator dyes were visualized through the translucent cuticle of the larvae. The pH inside the larval gut was determined by comparing the color and intensity shown within the gut with those from 0.4 % solutions of both dyes made in 8 ml LB medium (with one added drop of chloroform to prevent bacterial growth), and with pH ranging from pH 4 to 10 in 0.5 increments.

2.2.5. Effect of pH on *in vitro* growth of EGFP-expressing *B subtilis* and GFP-expressing *P. agglomerans*

Overnight cultures of fluorescent *Bs* and *Pa* were diluted to OD$_{600}$ = 0.1 and further diluted 1:10$^4$ prior to plating onto LB-agar plates supplemented with either 50 µg/ml of carbenicillin (CAR) or 5 µg/ml of chloramphenicol (CAM) for selection of *Pa* or *Bs*, respectively. The pH of plates ranged from pH 6 to 9.5 in 0.5 increments. Plates were incubated overnight at 37 °C for *Bs*, and at room temperature for *Pa*, and each experiment was performed in triplicate, and repeated twice. The following day bacterial colonies were counted, and colonies growing at each pH were observed under fluorescent microscope. One-way ANOVA with a Tukey test was performed to determine differences between pH.
2.2.6. RNA extraction and reverse transcription

Midguts from *L. longipalpis* L3 larvae were dissected under a stereoscope microscope in Hyclone® (Thermo Scientific) phosphate buffered saline (PBS) at 12, 24, and 36 h post feeding in either the *Bs* or *Pa*. Sterile, 1.5 % agar in LB was used as feeding control. Total RNA was isolated from pools of 20 midguts using TRIzol® (Invitrogen). For each group of larval midguts, RNA isolation was done in triplicate. RNA quality was assessed by electrophoresis on 1 % agarose-5 % formaldehyde in 1x MOPS, and stored at -80 °C. First strand cDNA synthesis was conducted using Superscript™ III reverse transcription kit (Invitrogen) as described [24].

2.2.7. Quantitative real time PCR

mRNA levels were quantified with iQ™ SYBER® Green Supermix (Bio-Rad) using 95 °C melting, 57 °C annealing, and 72 °C extension temperature for 40 cycles using a Realplex Master cycler (Eppendorf). Relative fold changes were assessed using the ∆∆Ct method [24, 52], and calibrated against the expression observed for same stage larvae fed on the plain LB-agar control. Sequences for *Attacin* (*Att*), *IMPer*, *Vein*, *Domeless*, *IMD*, *Pirk*, *USP36*, and *Duox* were obtained using the tBLASTN algorithm from corresponding annotated sequences found in *D. melanogaster* blasted against *L. longipalpis* contigs. Predicted full-length transcripts were made using GENSCAN (http://genes.mit.edu/GENSCAN.html), and primers sequences were generated using Primer3 (http://biotools.umassmed.edu/bioapps/primer3_www.cgi). Primers for RPS6 were previously described in [16]. Primers for *Def1* were based on the sequences described in an earlier study [17]. All other primers used in this study were designed from gene sequences in the NCBI database and their accession numbers are as follows: IMPer, AJWK01035414.1; DUOX,
AJWK01035414.1; IMD, AJWK01008032.1; Domeless, AJWK01008028.1; Attacin, AJWK01017071.1; Pirk, AJWK01015539.1; USP36, AJWK01027563.1; and Vein, AJWK01005322.1. All primer sequences used in this study are summarized in Table 2.4.

2.3. Results

2.3.1. Bs and Pa localize to different regions of the sand fly larvae midgut

A scheme representing the anatomy of *L. longipalpis* sand fly 3rd instar (L3) gut is depicted in Figure 1.1.A. Following continuous feeding of LB-agar containing EGFP-expressing *Bs* to larvae, a pervasive signal was found across the entire length of the midgut for infected insects as depicted by the fluorescent signal of full length images of the gut (Figure 2.1.B, 2.7.A, and 2.10.). However, when the GFP-expressing *Pa* was fed to larvae in a similar manner the bacteria were mostly localized to a narrow area of the posterior portion of the midgut (Figure 2.1.C and 2.7.B), and were only found on the apical surface of the midgut lumen (Figure 2.11.). Also observed were areas of high intensity GFP signal in *Pa*-infected guts, suggesting the presence of biofilm (Figure 2.1.C). A less pronounced GFP signal in *Pa* fed was also observed in the proventriculus of the gut (Figure 2.1.C). Infection rates as determined by a qualitative assessment of the GFP signal within the midgut of the larvae following continuous feeding are described in Table 2.1.

To further investigate what could be driving such a distinct microbial distribution, we assessed the gut pH range within larvae *in vivo*, and compared it to pH growth assays for the two GFP-expressing bacteria *in vitro*. LB-agar supplemented with pH indicators Bromothymol blue or Phenol red were fed to larvae, and the color gradient generated was visible through the translucent cuticle of larvae using light microscopy. Intensity of colors varied between larvae due
to the initial ingestion time, load size, and bolus movement across the gut. Fifteen larvae of each treatment were compared to the pH references. The bromothymol blue dye has a range of pH from 6 to 7.6, and phenol red has a range of 6.8 to 8.4. Previous results on *L. longipalpis* larval gut pH indicated a basic pH >9 in the anterior portion and an more acidic pH >6.5 in the posterior portion of the midgut [21]. The results shown in Figure 2.2 confirm such a pH gradient in the *L. longipalpis* larvae, clearly pointing to a basic pH for the anterior part of the midgut, including the proventriculus (PV), and an acidic pH in the posterior part of the midgut.

Both *Bs* and *Pa* bacteria were grown on antibiotic supplemented LB-agar plates with pH ranging from 6 to 9.5. CFU counts were obtained in triplicate to assess the viability of the two strains at different pH (Figure 2.3.A). Colonies of EGFP-expressing *Bs* did not show a significant difference in CFU counts at any given pH. Additionally, the colony size for *Bs* fed was smaller at low (6-6.5), and high (9.5) pH. However, *Pa* showed a significant difference in CFU counts at pH ranging from 6-7 with respect to pH 9.5 (Figure 2.3.B). Also, at pH 8-9 colony size and fluorescent intensity began to decrease, and by pH 9.5 there was no visible growth.

Following the continuous feeding experiment described above, we tested whether similar results could be obtained by feeding larvae once with either *Bs* or *Pa*. Larvae were fed for 12 h on LB-agar with the respective bacteria and transferred to pots with fresh LB-agar (no antibiotics). *Pa* bacteria were cleared by 21 h after infection, whereas *Bs* were cleared by 24 h (Table 2.2.). CFU counts (Table 2.3.) were generally in agreement with the results observed for the GFP signals assessed.
2.3.2. Infection of the sand fly larval midgut by *Pa* or *Bs* leads to differential damage of the larval midgut epithelium

*Bs* and *Pa* were assessed for their ability to infect the sand fly L3 larvae midgut following feeding, and their effect on induction of apoptosis. Monoclonal antibodies targeting the cleaved caspase3 were used as an immunocytochemical marker to identify epithelial cells undergoing caspase-dependent programmed cell death, and to assess the integrity of the midgut. In order to test if this was a viable approach, we fed larvae with LB-agar supplemented with the apoptosis inducer paraquat, and compared its effects to larvae fed on LB-agar alone. The LB-agar fed larvae displayed a well-defined midgut epithelium with little background staining for active caspase3 (Figure 2.8.A). In contrast, larvae fed on LB-agar supplemented with paraquat showed midgut epithelia with significant loss of integrity, that were also severely flattened after mounting on the slide with reduced luminal space detectable by looking at nuclei alone (Figure 2.8.B). The apoptotic effect of the paraquat was further confirmed by the presence of a large population of cells showing heavy cytoplasmic specific staining for caspase3 (Figure 2.8.C and D). After 12h of infection with *Bs* an extensive amount of luminal nucleic material is observed using DAPI staining (Figure 2.4.A), and a massive infection can be seen in Figure 2.4.B. Very little background caspase staining is observed in *Bs*-fed compared to the paraquat treated controls (Figures 2.4.C, 2.4.D, and 2.8.). In contrast, when *Pa* was used for infection, the Gram-negative bacteria induced staining comparable to that of the paraquat control (Figures 2.4.G, H, and 2.8.). However, we did not observe a similar breakdown in midgut superstructure.
2.3.3. Effect of feeding agar on larval development

Effects of feeding agar to the developing sand fly L3 larvae were assessed by comparing the length of the whole midgut and the width of three areas within the midgut to those obtained from larvae raised on regular larvae food (50 % rabbit feces + 50 % rabbit food). The same parameters were also measured from guts of *Pa* and *Bs* infected larvae. A significant difference was found for midgut length when comparing regular sand fly larval food and the agar fed, except at 48 h. For the three measurements of gut width (anterior, middle and posterior), significant differences were only observed between regular food and agar (Supporting Figure 2.9).

2.3.4. Bacterial feeding leads to differential expression profiles in sand fly larvae

We compared the mRNA expression profiles of nine genes likely involved in various physiological processes ranging from innate immunity, homeostasis, and epithelial regeneration in midguts of *L. longipalpis* L3 larvae fed on *Bs* and *Pa* bacteria. Among the transcripts assessed were *Att*, *Def1*, *Duox*, *IMPer*, *Vein*, *Domeless*, *IMD*, *Pirk*, and *USP36*. Results from qRT-PCR indicate that, compared to control fed, larvae fed on agar containing either *Bs* or *Pa* showed significant difference in expression for a number of genes analyzed.

At 12 h post infection, transcript levels for *Att* were downregulated by nearly 75 % for both *Bs* and *Pa* infected larvae, and *IMPer* was downregulated by 25 % in *Pa* fed (Figure 2.5.A). *Pirk* showed a 2.5-fold change (~150 % increase) in expression following infection with *Pa* (Figure 2.5.B).

At 24 h post infection, *Domeless* and *IMD* followed a similar profile and were upregulated in both infections, albeit *Domeless* was not significantly different between the
control and Bs fed. Pirk was also upregulated at in both bacterial infections, and with a profile that also was similar to Domeless and IMD (Figure 2.5.E). USP36 was downregulated only in Bs infection, and Def1 was upregulated only in Pa (Figure 2.5.F).

At 36 h post infection, Domeless and IMD again displayed similar profiles, however both were downregulated in comparison to control (Figure 2.5.H). For Pirk, whereas expression if Bs fed returned to control levels, in Pa infection it remained significantly higher (nearly 2-fold) (Figure 2.5.H). Finally, USP36 expression was reduced by roughly 10 % in both infections, but for Bs the significance was 0.078 (Figure 2.5.I).

2.4. Discussion

In insects, gut bacteria have been shown to significantly contribute to nutrition, modulation of the immune response, and protection from parasites and other pathogens. The insect gut varies greatly in terms of morphology and physicochemical properties, and these may influence the distribution and structure of the microbial community in the gut. During development, sand fly larvae are exposed to a wide variety of soil bacteria and other microorganisms that are able to colonize the insect gut [13-20, 53]). However, as we are aware, no studies have focused on the mechanisms by which bacteria are able to develop within the sand fly gut, or the types of specific responses induced by the colonization. Here, we assessed the ability of two bacteria previously identified from the guts of insects, including sand flies, to infect the gut of sand fly larvae, and investigated the specific responses (innate immunity, epithelia regeneration, homeostasis) induced by these bacteria.

When fed to L. longipalpis larvae, EGFP-expressing Bs bacteria were distributed throughout the entirety of the alimentary canal, residing within the peritrophic matrix and along
the lumen. In contrast, GFP-expressing Pa bacteria were mostly localized to the posterior midgut, and only at the apical surface (although we did observe GFP signal for Pa at the proventriculus of the gut, it is possible that these bacteria had not yet been killed by the alkaline conditions). We speculated that this phenomenon must be driven by pH and/or by specific cell types that line the midgut lumen. A pH driven effect on the colonization of the gut was suggested by the use of pH indicators. It has been previously determined in sand fly larvae that the pH of the midgut is highly alkaline at the anterior portion and decreases towards the posterior region [21]. Our approach allowed us to confirm the location within the larval gut wherein the pH ranges between 6 and 7, which also coincides with the location where Gram-negative Pa preferentially infects the larval gut. These results were indirectly confirmed by in vitro growth assays obtained for Pa in which these bacteria clearly favor a pH in the range of 6-7. The ability of Bs to sporulate under unfavorable growth conditions may have further contributed to its survival along the gut of the larvae.

In addition, we observed a marked difference between the rates and the persistence of infections of Pa and Bs in sand fly larvae. Following continuous feeding on bacterial lawn, most of the Pa bacteria are cleared within 24 h whereas infection with Bs remained for up to 48 h. However, if the bacterial lawn is replaced at 48 h during continuous feeding, larvae do re-infect. In contrast, with the non-continuous feeding on the bacterial lawn led to clearing of Pa by 21 h and Bs by 24 h. Hence, the data indicate that the sand fly larvae are able to clear bacterial infection if exposure is not maintained. Another possibility if that the loss of Pa and Bs during non-continuous feeding may also be caused by competition with other microorganisms present in the larval gut. And in spite of differences known to exist in the half-lives of GFP and EGFP
proteins [54, 55], the CFU counts reported support the clearing of Pa from the midgut during loss of GFP signal.

Using a caspase3 antibody [22, 23] to detect apoptotic activity in L. longipalpis, we also were able to clearly identify differences between Bs and Pa infection of the sand fly larval gut. The microscopy data strongly suggest that only Pa induces caspase activity within the midgut of the sand fly, while Bs causes little to no staining.

Quantitative RT-PCR analyses were used to assess changes in the expression profiles of nine selected genes chosen based on their roles in insect midgut immunity and homeostasis. Related to midgut immunity, selected genes included those coding for effector molecules such as the antimicrobial peptides (AMPs) attacin (Att) and defensin (Def1), as well as Duox and IMPer. Also included in this category was the immunodeficiency regulatory gene encoding “poor immune response upon knock-in” or Pirk.

Attacin has long been implicated in bacteria killing from a number of studies pertaining to its role in innate immunity [56]. Def1 was shown to be upregulated in adult L. longipalpis after bacterial challenge [17]. It has been shown that Defensin A, acting in concert with Cecropin A, blocks Plasmodium transmission in A. aegypti [57]. The effector molecules Duox and IMPer have been demonstrated to have effects on the midgut peritrophic matrix structure and parasite killing in A. gambiae mosquitoes [58-60]. For Def1, Att, DUOX, and IMPer, there are multiple studies suggesting that these effector molecules are regulated by the immunodeficiency pathway [59, 61, 62]. Pirk has been previously shown to be a negative regulator of IMD activity [63, 64]. While Pirk acts to suppress IMD at the level of signal transduction, Caspar negatively regulates IMD at the level of transcription. Studies in A. gambiae implicate the knockdown of IMD in
increased infectivity of mosquitoes [65, 66]. In sand flies, Caspar knockdown led to a decrease in Leishmania mexicana load in L. longipalpis [18].

*Domeless, Vein,* and *USP36* were selected based on their roles in pathways related to innate immunity to midgut regeneration. When the innate immune response is activated in the midgut, there are associated energy costs and damage to healthy epithelial cells that can negatively affect the insect. Artificially activating ROS production in *A. stephensi* led to reduction in infective lifespan, and deleterious effects associated with mitochondria [67]. Domeless is a receptor in the JAK/STAT pathway that is crucial for recognizing damage to healthy epithelial cells. JAK/STAT signaling reaches intestinal stem cells (ISCs) and enteroblasts (EBs) leading to the secretion of an epidermal growth factor (Vein) ending in regeneration of midgut epithelia via proliferation and differentiation of ISCs and EBs [23, 68-70]. Additionally, the deubiquitinating enzyme USP36 is a negative regulator of IMD and provides a route of cross-talk between IMD and JAK/STAT pathways [71, 72]. USP36 is also involved in controlling selective autophagy [72].

Our results suggest that *Pirk* may be acting to suppress *Att* and *IMPer* activity at 12 h post infection for *Pa*-infected insects, however, another still unidentified mechanism likely is involved in the reduction of *Att* levels for *Bs*-infected. Although *Pirk* was significantly upregulated during *Bs* and *Pa* infection at 24 h, there was also an increase for the immune transcription factor *IMD*. Such *IMD* increase can be linked to an associated upregulation of *Def1* in *Pa* infected, but no significant difference was found in *Def1* for *Bs* infected. The upregulation of *IMD* may be explained by the concomitant down regulation of *USP36* in the *Bs* infected larvae, but not *Pa* infected. Additionally, the nearly two-fold increase of *Domeless* in *Pa* fed
larvae suggests the possibility of homeostatic response to damage by the larvae immune response that occurs within the first 12 h of infections as our data have demonstrated.

By 36 h, upregulation of Pirk continued in Pa infected individuals, but no effect was observed for the expression of the effector molecules. With the downregulation of USP36, we expected an upregulation of IMD. However, the opposite was detected: IMD was downregulated. Interestingly, Domeless was also downregulated at 36 h, possibly due to lack (or clearing) of bacteria in the gut as indicated by the non-continuous feeding experiments. It remains to be investigated whether downregulation of IMD, when USP36 levels were also lowered, is associated with increased autophagy during bacterial clearing. The expression analyses data corroborates what was observed with regards to the progress of infection in Bs and Pa. Of significance, our results indicate that sand fly larvae are able to differentially regulate (or suppress, as the case here) their immune response according to the bacterial challenge they are exposed to.

We have previously shown that feeding different bacteria to L. longipalpis larvae affects survival and development [20]. In the current study, we demonstrate a selective localization of bacteria in the larvae driven by gut pH and downstream effects on the larval gut. This provides a link between the type of bacteria colonizing the gut, physicochemical aspects of the gut, and overall insect health. With regards to mechanisms driving the localization of bacteria, it is also likely that different cell types lining the gut epithelia are involved. In support of this hypothesis, concentrated pockets of Pa binding to the posterior end of the larval gut were observed, indicating the presence of a biofilm. However, the presence of a preferred cell type or membrane receptor involved in binding of bacteria cannot be discarded. Gram-negative bacteria are known to form biofilm within the gut of vectors [73]. P. agglomerans form intestinal biofilms in the
Mediterranean fruit fly *Ceratitis capitata* [74] that resemble what we observed in sand fly larvae. Taken together, these data suggest a pH-dependent localization or growth of bacteria within the insect midgut previously reported to be a random event [74].

With regards to cell type, Fernandes et al [43] reported the presence of different cell types in *A. aegypti* during development and metamorphosis, and a precedent for favored microbial binding was previously demonstrated for *Leishmania major* binding to the midgut epithelial cell lining of the sand fly *P. papatasi* [75]. Thus, it is conceivable that at least one of these events may also be involved in dictating the success of bacterial colonization within the sand fly gut. Nevertheless, mechanisms such as autophagy may also play a role in bacterial removal (reviewed in Huang et al. [76]). Further, both *Pa* and *Bs* do not survive metamorphosis.

It is important to note that the agar based feeding system used in our experiments does not replicate the natural conditions faced by sand fly larvae and agar does not provide the necessary nutrition for normal larvae development. As shown by our analyses, the midgut length and width differed significantly between larvae fed on agar versus those fed on regular sands fly larvae chow. However, no differences in such parameters of the midgut morphology were observed between agar fed larvae and the agar plus bacteria fed larvae. Interestingly, differences observed for the midgut parameters tested only lasted until either the agar or the bacterial lawn were replaced. Additionally, larvae were not able to sustain a GFP-positive signal when fed on LB-agar plus bacteria for 12 h and then transferred to plain LB-agar. These data were also supported by CFU counts. Notwithstanding, this method was proven useful for specifically delivering the EGFP- or GFP-expressing bacteria. Similar approaches may be used to deliver selected microbes to sand fly larvae in paratransgenic applications to control sand fly populations [77, 78].
In conclusion, this study demonstrates that bacteria selectively infect the sand fly larval midgut, (possibly) leading to epithelial damage. In addition, the data also point to a modulation of the innate immune response likely controlled by expression of *Pirk*. We also show for the first time that the insect midgut pH is a factor driving microbial organization of the gut. Our results contribute towards understanding of midgut responses to infections and provide new insights for development of vector control approaches using paratransgenesis.
Table 2.1. *P. agglomerans* and *B. subtilis* infection rates

<table>
<thead>
<tr>
<th>Time post infection</th>
<th><em>P. agglomerans</em></th>
<th><em>B. subtilis</em></th>
</tr>
</thead>
<tbody>
<tr>
<td>12 h</td>
<td>9 of 14 (64%)</td>
<td>14 of 15 (93%)</td>
</tr>
<tr>
<td>24 h</td>
<td>0 of 9</td>
<td>15 of 15 (100%)</td>
</tr>
<tr>
<td>36 h</td>
<td>1 of 9 (11%)</td>
<td>7 of 7 (100%)</td>
</tr>
<tr>
<td>48 h</td>
<td>0 of 9</td>
<td>3 of 8 (37.5%)</td>
</tr>
</tbody>
</table>

*P. agglomerans* and *B. subtilis* were continuously fed to sand fly larvae. At time points indicated (left column), the infection rates in sand fly larvae were determined by a qualitative assessment (presence or absence) of GFP signal using a Zeiss 510 confocal microscope.

doi:10.1371/journal.pntd.0003923.t001
Table 2.2. Persistence of bacterial infection between continuous and non-continuous feedings

<table>
<thead>
<tr>
<th>Time</th>
<th>Continuous feeding</th>
<th>Non-continuous feeding</th>
<th>GFP positive signal</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>P. agglomerans</td>
<td>B. subtilis</td>
<td>P. agglomerans</td>
</tr>
<tr>
<td>12 h</td>
<td>8 of 10⁺</td>
<td>10 of 10⁺</td>
<td>-⁺</td>
</tr>
<tr>
<td>15 h</td>
<td>-</td>
<td>-</td>
<td>1 of 5</td>
</tr>
<tr>
<td>18 h</td>
<td>-</td>
<td>-</td>
<td>1 of 5</td>
</tr>
<tr>
<td>21 h</td>
<td>-</td>
<td>-</td>
<td>0 of 5</td>
</tr>
<tr>
<td>24 h</td>
<td>1 of 5</td>
<td>5 of 5</td>
<td>0 of 4</td>
</tr>
<tr>
<td>27 h</td>
<td>-</td>
<td>-</td>
<td>0 of 5</td>
</tr>
<tr>
<td>30 h</td>
<td>-</td>
<td>-</td>
<td>0 of 5</td>
</tr>
<tr>
<td>36 h</td>
<td>1 of 5</td>
<td>4 of 4</td>
<td>0 of 3</td>
</tr>
<tr>
<td>42 h</td>
<td>-</td>
<td>-</td>
<td>0 of 5</td>
</tr>
<tr>
<td>48 h</td>
<td>0 of 5</td>
<td>3 of 5</td>
<td>0 of 5</td>
</tr>
</tbody>
</table>

Persistence of bacterial infection in sand fly larvae was determined by a qualitative assessment (presence or absence) of GFP signal using a Zeiss 510 confocal microscope. GFP positive signal observed for larvae fed for up to 48 h on bacterial lawns (continuous feeding), of for larvae fed 12 h on bacterial lawn of each bacteria, and transferred to plain LB-agar (non-continuous feeding). Larvae were collected at different time points and the midguts were dissected and prepared for confocal analyses as indicated. Results observed for the continuous and non-continuous feeding are shown as number of larvae displaying a GFP signal per total larval guts.

*As larvae in both groups fed for 12h, there is no difference between continuous and non-continuous for that time point; all larvae were grouped into the continuous feeding group.

doi:10.1371/journal.pntd.0003923.k002
<table>
<thead>
<tr>
<th>Time</th>
<th>Continuous feeding</th>
<th>Non-continuous feeding</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>P. agglomerans</td>
<td>B. subtilis</td>
</tr>
<tr>
<td>12 h</td>
<td>0–48 (n = 5) (10.8 ± 20.9)</td>
<td>30–2600 (n = 5) (647.6 ± 1095.6)</td>
</tr>
<tr>
<td>24 h</td>
<td>4–6 (n = 3) (6.0 ± 2.0)</td>
<td>0–480 (n = 3) (330.2 ± 286.2)</td>
</tr>
</tbody>
</table>

Range (AVG±Stdev) distribution of CFUs present in gut larva at 12 h (n = 5) and 24 h (n = 3) after continuous or non-continuous feeding. In continuous feeding, larvae were fed on LB-agar plus bacteria for 24 h; for non-continuous feeding, larvae were fed on LB-agar plus bacteria for 12 h and transferred to LB-agar. Guts were dissected from surface sterilized larvae and homogenized in 60 μl of 1X PBS followed by plating on selective LB media. After overnight incubation at 37°C plates were scored for the presence of colonies assessed according to morphology and GFP signal. Results shown are representative of three experiments.

†CFUs present in Pa-fed larva at 24 h were likely due to cannibalism observed in this group.

doi:10.1371/journal.pntd.0003823.k003
Table 2.4. List of primer sequence used in this study

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Symbol</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dual Oxidase F</td>
<td>DUOX</td>
<td>5'-GGCAAGACGGAAGACAAG-3'</td>
</tr>
<tr>
<td>Dual Oxidase R</td>
<td>DUOX</td>
<td>5'-TCAACAGGAGCAGCATC-3'</td>
</tr>
<tr>
<td>Immuno-Modulatory Peroxidase F</td>
<td>IMPer</td>
<td>5'-CTGTGTGCAGTAAATGTC-3'</td>
</tr>
<tr>
<td>Immuno-Modulatory Peroxidase R</td>
<td>IMPer</td>
<td>5'-GTGGTAGGTGTGCTGGTAAG-3'</td>
</tr>
<tr>
<td>Immunodeficiency F</td>
<td>IMD</td>
<td>5'-GGTGAAACAACACTCAAGCT-3'</td>
</tr>
<tr>
<td>Immunodeficiency R</td>
<td>IMD</td>
<td>5'-GTTACTCTGGTGCTGGGGA-3'</td>
</tr>
<tr>
<td>Domeless F</td>
<td>Dome</td>
<td>5'-TCAACACACCCCCAAATAC-3'</td>
</tr>
<tr>
<td>Domeless R</td>
<td>Dome</td>
<td>5'-ACGCCTCTCAATCACGATA-3'</td>
</tr>
<tr>
<td>Attacin A F</td>
<td>AttA</td>
<td>5'-AGGATGAGAGAGGAAGACAG-3'</td>
</tr>
<tr>
<td>Pirk F</td>
<td>Pirk</td>
<td>5'-AAAGATGAGGGGTAGGGAAG-3'</td>
</tr>
<tr>
<td>Pirk R</td>
<td>Pirk</td>
<td>5'-CCAACATACGCAAATCAG-3'</td>
</tr>
<tr>
<td>USP36 F</td>
<td>USP36</td>
<td>5'-CTACGAAACTGGAAGAGTTG-3'</td>
</tr>
<tr>
<td>USP36 R</td>
<td>USP36</td>
<td>5'-GATTTTGCTCTGCTGGATG-3'</td>
</tr>
<tr>
<td>Vein F</td>
<td>Vein</td>
<td>5'-CGCAATGGATGAGAAC-3'</td>
</tr>
<tr>
<td>Vein R</td>
<td>Vein</td>
<td>5'-TGAGCAATACCTACGCTGAC-3'</td>
</tr>
<tr>
<td>Defensin 1 F</td>
<td>Def1</td>
<td>5'-GCCTGTGTTTGTAGTTCT-3'</td>
</tr>
<tr>
<td>Defensin 1 R</td>
<td>Def1</td>
<td>5'-GCTTCCCTCCATCGCTTT-3'</td>
</tr>
<tr>
<td>Ribosomal Protein s6</td>
<td>LIS6</td>
<td>5'-TCCCTGGGTGATGAGTTG-3'</td>
</tr>
<tr>
<td>Ribosomal Protein s6</td>
<td>LIS6</td>
<td>5'-CCCTTGTGCGTGGTCTTC-3'</td>
</tr>
</tbody>
</table>

Supporting Table S1 – Primer list

Primers were based on sequence information obtained from various databases. F and R in each primer indicate the forward and reverse primers used, respectively.
Figure 2.1. Infection of sand fly larva midgut by *B. subtilis* and *P. agglomerans*.

EGFP- or GFP-expressing Bs and Pa were grown on LB-agar plates with selective media and fed to 3rd instar sand fly larvae. Larvae guts were dissected and assessed for the distribution of each bacterium. In A, a schematic representation of the sand fly larval gut. Ingested food is moved from right (proventriculus – pv) to left, towards to posterior midgut and hindgut. Confocal images (1024x1024 per tile pixel resolution) of the distribution of EGFP-expressing Bs-infected and GFP-expressing Pa-infected midguts are shown in panels B and C. Posterior (pos) and anterior (ant) portions of midguts are indicated. SG, salivary glands. Inset in 1C: blow up of area of gut delineated by a rectangle (asterisk) showing biofilm formed in Pa infection. Bars = 100 µm.
Third instar sand fly larvae were fed with pH indicators bromothymol blue (A) and phenol red (B). Shown is the distribution of each indicator within live sand fly larval guts, with the predicted pH for each area of the gut indicated. Live 3rd instar sand fly larvae are shown from left (posterior or caudal setae) to right (anterior or head).
Figure 2.3. Effect of pH on in vitro growth of *B. subtilis* and *P. agglomerans*.

Cultured bacteria were grown on LB-agar plates of pH varying from 6-9.5. The colony forming units are measured for Bs (A) and Pa (B). One way ANOVA, with a post hoc Tukey test, was performed to assess significance. pHs 6, 6.5, and 7 were statistically different than pH 9.5 (P<0.01 for pH 6 and 6.6; and P<0.05 for pH 7).
Figure 2.4. Confocal images of *B. subtilis* and *P. agglomerans* infection of sand fly larvae midguts.

Anterior midgut image 12h post feeding depicting differential distribution of bacteria and apoptotic responses. A) DAPI staining; B) Shows Bs distributed throughout the anterior larval gut; C) Immuno-staining for cleaved caspase3 along the lumen of the midgut. D) Merge. 12h post infection with GFP-expressing Pa localized to the posterior region of the midgut epithelium and induces apoptotic activity. E) DAPI staining depicting the midgut epithelium. F) Shows Pa localized on the apical portion of the lumen of the midgut. G) Immuno-staining for cleaved caspase3 along the lumen of the midgut depicting high levels of caspase3 activity. H) Merge. Bars = 50 µm.
Figure 2.5. mRNA expression profiles of 3rd instar *L. longipalpis* larvae post infection with *B. subtilis* and *P. agglomerans*.

*B. subtilis* (Bs) or *P. agglomerans* (Pa) bacteria were fed to larvae and the expression of selected transcripts relative to agar fed control larvae was assessed. Expression profiles were obtained for the effector molecules Att and IMPer, in addition to the epithelial growth factor Vein are shown in A, D, and G. Expression profiles for the JAK/STAT receptor Domeless, the transcription factor for immunodeficiency IMD, and the negative regulator of IMD pathway Pirk are shown in B, G, and H. Expression profiles for the E3 Ubiquitin ligase associated with IMD, and the effector molecules dual oxidase DUOX and Def1 are shown in C, F, and I. Following bacterial feeding, total RNA was obtained at 12 h (A, B, and C), at 24 h (D, E, and F), and at 36 h (G, H, and I) post infection (PI). All Ct values were normalized to the ribosomal protein S6 (RPS6). Error bars are represented as the standard deviation determined from three biological replicates each with n=20 midguts. Significance was determined using student t-tests. * denotes *P*<0.05 and ** denotes *P*<0.01. Y-axis, fold change.
Figure 2.6. Larval feeding of bacterial lawn.

After overnight incubation on LB-agar plus antibiotics, each bacterial lawn was cut and fed to 3rd instar larvae. On the left (A) the larvae are feeding on EGFP-expressing Bs, while on the right (B) the larvae are feeding on GFP-expressing Pa.
Figure 2.7. Infection of sand fly larval midgut by *B. subtilis* or *P. agglomerans*.

Larval guts were imaged using a resolution of 512 x 512 (number of pixels per tile). Ingested food is moved from right (proventriculus – pv) to left, towards to posterior midgut and hindgut. EGFP-expressing *Bs*-infected (A) and GFP-expressing *Pa*-infected (B) midguts are shown. Posterior (pos) and anterior (ant) midgut are marked. Arrowheads indicate the separation between midgut and hindgut. Bars = 100 μm.
Figure 2.8. Paraquat-induced apoptosis in sand fly larval midguts.

Ingestion of paraquat induces a detectable and systemic apoptotic response in the cytoplasm of midgut epithelial cells 12h post feeding. A) Merge of caspase3 and DAPI stained nuclei for larvae fed only LB-agar medium in the anterior midgut versus B) larvae with LB-agar supplemented with paraquat visualized with DAPI 12h post infection (anterior midgut). C) Immuno-staining for cleaved caspase3 in paraquat fed larvae. D) Merge of B and C. Bars = 50 µm.
Figure 2.9. Changes in sand fly larval gut length and width caused by diet.

*L. longipalpis* 3rd instar larvae were fed on either regular sand fly larval food (50% rabbit feces + 50% rabbit food), or on LB-agar with or without bacteria, for up to 72 hours. Larval midguts were dissected and measured for length, and width in the anterior, middle, and posterior regions of the gut.
Figure 2.10. Localization of B. subtilis.
Refer to Supporting Information Figure S1 Video found at http://journals.plos.org/plosntds/article?id=10.1371%2Fjournal.pntd.0003923

Figure 2.11. Apical localization of P. agglomerans on midgut epithelia.
Refer to Supporting Information Figure S2 Video found at http://journals.plos.org/plosntds/article?id=10.1371%2Fjournal.pntd.0003923
Chapter 3 - The Role of Autophagy during Metamorphosis of the Sand Fly Lutzomyia longipalpis

Abstract

The development of holometabolous insects is a complex process, especially with respect to the stages where the organism does not acquire nutrients from the environment. Undergoing the metamorphosis from larvae to adult requires the insect to manage and recycle its own available resources to develop fully. Autophagy is crucial for non-feeding larvae to undergo successful pupation and metamorphosis into the adult form. Here, we were able to measure 3 genes related to the pre-autophagosome, nucleation, and sequestration involved in successful autophagy. We found that during pupation and metamorphosis that these transcripts were significantly upregulated. Additionally, we provide evidence for one candidate epidermal growth factor Vein for production of adult tissues within the pupae of sand flies. We were also able to visualize autophagosomes residing within the cytoplasmic compartment of pupal stages using confocal microscopy. Finally, we present preliminary evidence from the sand fly genome that ecdysone secretion during metamorphosis may be directly linked to the activation of autophagy. Understanding the underlying mechanisms of this process will provide invaluable data in the future management of this neglected tropical vector.

3.1. Introduction

Autophagy is a complex process that involves the processing and recycling of cellular macromolecules and larger organelles. This can be associated with a multitude of phenomenon including stresses like starvation, or intriguingly the decision of a cell to undergo programmed cell death or survival. Within higher mammals this process is intimately linked to numerous disease pathologies such as neurological degeneration and cancer. Interestingly, there is
mounting evidence to suggest that autophagy is paramount in regulating the varied and finely orchestrated developmental events that occur in various insect life stages. A majority of the work involves the use of the model organism *Drosophila melanogaster*. Here the role of autophagy in recycling tissues and cellular reprogramming is apparent. During fruit fly embryogenesis autophagy plays a role in removing excess sera from developing insects [32]. Additionally, degradation of larval salivary glands during metamorphosis is required to produce the final adult tissues, and is autophagy dependent [33, 34]. Over-proliferation of the neuromuscular junction of larval stages of the fly is prevented via functioning autophagy [35]. Breakdown of fat body during development of not only fruit flies, but *Bombyx mori*, is also a crucial process that has been linked to autophagy [36, 37].

Larval midgut removal and the tissue remodeling and differentiation of the adult gut are a hallmark of metamorphosis. Autophagy plays a center role in nutrient recycling of midgut cells fated to undergo apoptosis and survival of gut cells that will ultimately comprise the adult tissue. A number of more recent studies implicate the function of autophagy in both pro-death and pro-survival in fruit flies and silk worms [38-41].

Within hematophagous insects there are few studies delineating the role of autophagy during development. Studies suggest that autophagy in the fat body of *Aedes aegypti* is important in maintaining gonadotrophic cycles [42]. Additionally, the presence of electron dense autophagosomes has been detected in the developing midgut of the mosquito during metamorphosis [43].

Ecdysone is the major hormone employed by insects to progress through various developmental stages. Its complex set of interactions with ecdysone receptors and primary response genes is vital in finely tuning spatial and temporal developmental events. Briefly and
specifically, E93 and Broad (Br-C) have been identified as 2 primary ecdysone response genes that signal the removal of the larval midgut of *D. melanogaster* before completion of the adult gut via autophagy [44-46]. Further, it is suggested that the action of these molecules may be conserved across holometabolous insects [47]. However, insights into the direct mode of action of ecdysone, primary response genes, and autophagy in affecting successful metamorphosis have only very recently begun to be understood. In *B. mori* it has been shown that autophagy-related 1 (ATG1) is a response gene to ecdysone. E93 is able to bind to ATG1 promoter region and induce its expression [48]. While these finding help in understanding how the larval gut may be removed during pupation, little is known about how the new adult gut is formed. It is suggested that Vein, an epidermal growth factor, may be involved in signaling intestinal stem cells of adult *D. melanogaster* to differentiate into healthy adult epithelial cells [23, 49, 50].

To our knowledge there is no study published on the role of autophagy and its possible response to ecdysone, or the place of epidermal growth factors during metamorphosis of the sand fly *L. longipalpis*. Here we present evidence that key components of the autophagy pathway are conserved in sand flies and expressed during metamorphosis of the fly. We also demonstrate the midgut specificity for expression of an autophagy component in the cytoplasm of developing epithelial cells. Additionally, we provide a possible candidate for the generation of an adult midgut tissue in sand flies. Finally, we provide a preliminary link that suggests that the promoter region of ATG1 may contain the necessary elements for interaction with primary response genes associated with ecdysone.

### 3.2. Material & Methods

#### 3.2.1. Sand fly colony maintenance
*L. longipalpis* (Jacobina strain – LLJB) colonies were reared in the Department of Entomology, Kansas State University. Larvae were maintained in 250 or 500 ml plastic jars (Nalgene) with an approximately 2 cm-thick bed made of dental plaster (Schein), and fed on larval chow (a mixture of 50% rabbit droppings and 50% rabbit food).

**3.2.2. RNA extraction and reverse transcription**

Whole insects from *L. longipalpis* L3, wandering L4, pupae, or adult insects were dissected under a stereoscope microscope in Hyclone (Thermo Scientific) phosphate buffered saline (PBS). Total RNA was isolated from pools of 5 insects using TRIzol (Invitrogen). For each group of midguts, RNA was isolated in triplicate. RNA quality was assessed by electrophoresis on 1% agarose-5% formaldehyde in 1x MOPS gel, and stored at -80 °C. First strand cDNA synthesis was conducted using Superscript™ III reverse transcription kit (Invitrogen) as described [24].

**3.2.3. Real time quantitative PCR**

mRNA levels for different developmental stages were quantified with iQ SYBER Green Supermix (Bio-Rad) using 95° C melting, 57° C annealing, and 72° C extension temperature for 40 cycles using a Realplex Master cycler (Eppendorf). Relative fold changes were assessed using the ∆∆Ct method [24, 52, 79]. Sequences for *ATG1, ATG6*, and *ATG8* were obtained by taking sequences from *D. melanogaster* and using tBlastn and/or Blastp against *L. longipalpis* sequences found at VectorBase (https://www.vectorbase.org/blast). Primer sequences were generated using Primer3 software (http://biotools.umassmed.edu/bioapps/primer3_www.cgi). Primers for *Vein* were based on the sequences used in a previous study [79]. PCR amplicon
sequence was verified using direct sequencing at Kansas State University. All Blastp data and primer sequences used in this study are summarized in Table 3.1. & 3.2.

3.2.4. Sequence alignments, phylogeny, and putative ecdysone response element prediction

Sequences for ATG1, ATG6, ATG8, and Vein were obtained for various dipterans and humans using Blastp at NCBI or VectorBase (summary of organisms, accession numbers, and databases are provided in supplementary table 3.3.). In the case of Homo sapiens GABARALP1 (ATG8) the sequence was provided by the manufacturer (Abcam). Multiple sequence alignments at the protein level were conducted using Clustal Omega software (http://www.ebi.ac.uk/Tools/msa/clustalo/). Alignments were then visualized using BoxShade (http://www.ch.embnet.org/software/BOX_form.html). To determine phylogeny, neighbor joining trees were generated with 1000 bootstrap replicates using Mega7 (www.megasoftware.net). To determine putative ecdysone response element regions upstream of ATG1, genomic sequence (AJWK01026160.1) was obtained from the European Nucleotide Archive (http://www.ebi.ac.uk/ENA/). This sequence was searched for the canonical and non-canonical ecdysone responsive sequence elements described previously [80-82].

3.2.5. Midgut immunocytochemistry & measurements

Whole guts from *L. longipalpis* larvae (n=5) were dissected from three separate treatments of L3, wandering L4, pupae, or adult into PBS and fixed for 20 minutes at room temperature with 4% paraformaldehyde in PBS. Tissues were washed 4 times for 30 minutes with PBS containing 0.3% Triton X-100 (PBST), then blocked with PBS containing 1% bovine
serum albumin for 30 minutes at room temperature. Tissues were then incubated overnight at 4 °C with primary antibodies for rabbit anti- *Drosophila melanogaster* Beclin1 (ATG6) rabbit anti-Human GABARAPL (ATG8) from Abcam diluted 1:500 in PBST. Tissues were washed 3 times for 30 minutes with PBST, and incubated overnight at 4 °C with Alexa Fluor® 594 goat anti-rabbit (Invitrogen) diluted 1:1000 in PBST. Tissues were washed 3 times for 30 minutes with PBST, and nuclei were stained for 5 minutes with 10 µg/mL DAPI (Invitrogen). Samples were mounted in FluorSave (CalBiochem) anti-photo bleaching reagent, and images were obtained with a LSM 700 confocal microscope (Zeiss). In addition, images were manipulated and midgut length, width, and axial ratios [83] were measured using the Zen confocal microscope software (Zeiss).

### 3.3. Results

#### 3.3.1. Change in expression profile of ATG1, ATG6, ATG8, and Vein during development

There are multiple stages associated with autophagy including the early autophagosome kinase activity, membrane trafficking, and elongation. We sought to quantify elements of these stages by measuring transcript markers for each. Additionally, there is little known about how the formation of the midgut is signaled, so we measured a putative epidermal growth factor (EGF). We began by selecting feeding 3rd instar (L3) larvae, non-feeding 4th instar (L4) larvae, newly formed pupae, and adult sand flies for performing qRT-PCR to measuring fold changes for the transcripts *ATG1* (serine/threonine kinase), *ATG6* (vesicle trafficking), *ATG8* (elongation), and *Vein* (putative EGF) calibrated against the L3 stage. We found 1 to 1 orthology for all of these transcripts against other dipterans and designed primers accordingly (summarized
in Figures 3.8.-3.16. and Tables 3.1.-3.3.). For ATG1 we observed a greater than 10 fold increase in transcript abundance for pupae compared to all other developmental stages (Figure 3.1.A). We saw a modest but non-significant increase for ATG6 at L4, and a significant 200 plus fold change for the transcript at the pupal stage versus all other stages (Figure 3.1.A). The significant peak stage for ATG8 expression occurred during L4 versus L3 and adult. Pupa showed a mean upregulation of expression but not significantly different using one way ANOVA with Tukey multiple comparisons to any other stage due to the high degree of variability in L4 (Figure 3.1.B). If the same statistical method is repeated sans L4 data, pupae is indeed significantly (P<0.05) different from L3 and adult samples. In the case of Vein an increase of ~30 fold is observed for both L4 and pupa versus L3 and adult (Figure 3.1.B. However, this change is again insignificant (P<0.06) using one way ANOVA with Tukey multiple comparisons due to, again, the large degree of variability in the L4 samples obtained. When conducting this same statistical test using only L3, Pupa, and adult we observe a significant (P<0.001) difference with respect to pupa versus L3 and adult. For both ATG8 and Vein it is of importance to note that there was a high degree of variability in biological replicates obtained from L4 larvae.

We also measured via confocal microscopy the differences in midgut dimensions and axial ratios (L/W) at all four stages as a means to identify developmental timing. There is no significant difference in length between L3, pupa, and adult sand fly guts. However, the L4 gut is significantly ~200 microns longer than pupa and adult on average (Figure 3.6.A). In addition to length, the average width of the L4 midgut is ~100 microns wider than any other stage observed (Figure 3.6.B). This correlates to a significantly lower value for the axial ratio observed in L4 larvae in comparison to pupae (Figure 3.7.). These observations allow us to see specific midgut morphologies as the insect approaches, proceeds through, and exits metamorphosis.
3.3.2. Midgut specific autophagy during pupation

Since our differences in RT-qPCR were generated from the entire insect undergoing development, we wished to further probe the midgut in order to visualize using immunocytochemistry specifically what was occurring in that tissue. To do so we applied antibodies targeting c-terminal end of D. melanogaster ATG6 and n-terminal end of Homo sapiens ATG8. The n-terminal antigen sequence used to generate antibodies against HsATG8 is similar to that of L. longipalpis as shown by multiple sequence alignment (Figure 3.11.), and we expected cross-reaction between species.

We processed tissues and looked for fluorescent signal using confocal microscopy. We observed that the anterior midgut region of pupae showed staining within only the epithelial cells (Ec) for ATG6 while no staining was observed in the midgut lumen (Lu) (Figure 3.2.A). More specifically the staining was limited to the cytoplasmic compartment of positively stained cells where one would expect autophagosomes to be present (Figure 3.2.B and C). This indicates that the midgut epithelial cells of pupae forming adult midgut tissues are expressing ATG6. The other stages, L3, L4, and adult, probed for ATG6 staining showed little to no red signal for the presence of antibodies and shown again by anterior region (Figure 3.3.A-C). We did see a non-significant upregulation for the ATG6 transcript in L4 larvae, but we were unable to see staining in the midgut for this protein.

According to qRT-PCR, we expected to see staining for ATG8 in the midgut epithelial cells of both L4 and pupae. However, in the case of this particular antibody we were unable to see any signal. Not only did L4 and pupae lack staining, but also L3 and adults (Figure 3.4.A-D). It is possible that the antibodies derived from the human sequence simply did not cross react with
sand fly tissues, or tissues other than the midgut at these developmental stages were expressing the ATG8 protein.

3.3.3. Evidence linking the role of ecdysone to the activation of autophagy

We previously mentioned that there is growing evidence to support the idea that ecdysone hormone action may be directly linked to autophagy through the promoter of response genes, specifically the ATG1. So we procured the promoter region for L. longipalpis ATG1 from VectorBase and searched the sequence upstream of the start codon for possible candidate sequences. These sequence elements are formed from imperfect palindromic sequences comprising two half sites separated by n nucleotides (-/n/+). One of the original and canonical negative (-) ecdysone responsive element sequence half sites identified as 5’-GTTTCA-3’[80] was found twice upstream of the start codon (Figure 3.4.). Another possible positive (+) half site, 5’-TGACAT-3’, is seen downstream of the negative sites (Figure 3.4.). Additionally, the GAGA motif previously described [82] as binding to primary response gene E93 is conserved in L. longipalpis (Figure 3.4.). There is a non-canonical difference in these results, as these half-sites contain a large number of nucleotides (>5) between them. While containing the correct sequence, this fact, makes them poor candidates for a canonical primary ecdysone response. These data will require rigorous scientific confirmation of these sequences role in activating ATG1 expression in response to ecdysone. However, there is observable preliminary evidence to suggest that ATG1 in L. longipalpis may be directly linked to ecdysone at the level of promoter activation and expression.

3.4. Discussion

Growing evidence in multiple orders of insects suggests an intimate link between developmental transitions and autophagy. This finely tuned interplay between death of the larvae
stages cells and tissues and survival of progenitor cells destined to become adult tissues during metamorphosis is only partially understood. On one hand autophagy has been roundly demonstrated to remove immature tissues. However, there is suggested a likely role of autophagy in pro survival for certain subsets of cells. Additionally, there is little published data demonstrating what types of signals are produced to guide these progenitor cells into differentiating into adult tissues. Moreover, traditional theory dictates that autophagy is a downstream event of target of rapamycin (TOR) regulation [33, 36, 37] in the presence of ecdysone. However, recent work in B. mori demonstrates a direct link between the molting hormone and the onset of autophagy that may skip the need for TOR signaling [48].

To date there are few studies within hematophagous insects dedicated to understanding the potential roles of autophagy during metamorphosis, and how the adult midgut may be signaled to form. Here we show preliminary evidence showing a conserved autophagy primary network genes present in the sand fly L. longipalpis and present a possible candidate for midgut formation. Additionally, we are the first within blood feeding insects to provide a putative direct genetic link between the ecdysone hormone and the expression of autophagy molecules.

When we measured the transcripts ATG1, ATG6, ATG8, and Vein during metamorphosis we saw that they were all up-regulated during the pupal stage. These three target transcripts have been demonstrated from yeast to mammalian systems to be markers for the various stages of autophagy. ATG1 is a crucial gene in establishing successful autophagy. The kinase activity of ATG1 allows for the phosphorylation of downstream targets leading to the formation of autophagosomes (reviewed in [84]). The n-terminal kinase domain in sand flies bears striking resemblance to that of other dipteras (Figure 3.8. and 3.9.), and likely serves a conserved function with respect to autophagy. Ubiquitin like protein ATG8 is extremely conserved from
insects to higher organisms (Figures 3.12. & 3.13. as its n-terminal region in *H. sapiens* almost exactly matches that of *L. longipalpis* (Figure 3.16.). It is anchored to the membrane of the autophagosome via a phosphatidylethanolamine addition and serves to extend the size of the isolation vesicle [85, 86]. It has also been noted in mammals for its ability to remove insoluble ubiquitinated proteins leading to survival [87]. The role of *ATG6* in autophagosome formation is quite interesting as it interacts with a number of different proteins, which can lead to programmed cell death or cell survival depending on its binding partners. In insects its function is best understood in *D. melanogaster*. It has a role in vesicle trafficking and leads to the degradation of cellular components and apoptosis when paired with vacuolar protein sorter (Vps34) [88]. However, it is also required for hematopoiesis and cellular differentiation in the fruit fly [89]. This pro-survival mode of action for *ATG6* is likely due to its association with the anti-apoptotic protein Bcl-2 [90]. It would be interesting to further probe what protein complex *ATG6* is involved in during metamorphosis within the sand fly to determine if it is playing a pro-survival or pro-apoptotic role or both and in what tissues, or specifically cells, is this role being carried out. *Vein* was our putative epidermal growth factor that may be signaling adult tissues to develop within the puparium. It contains conserved epidermal growth factor and immunoglobulin domains for EGFR signaling in vertebrates [91]. There are studies that pinpoint *Vein*’s action in the development of the adult fruit fly midgut [49, 50]. Additionally, it plays a role in the repair of adult midguts after infection [23]. We have previously describe *L. longipalpis Vein* in response to infection from *Bacillus subtilis* and *Pantoea agglomerans* in the sand fly larvae [79]. It’s presence in both L4 larvae and pupae suggests the beginning of adult tissue formation in sand fly puparium. We have also seen an up-regulation for the putative *Vein* in the Dengue fever mosquito *Aedes aegypti* (data not shown here). It would be interesting to see
if *Vein* up-regulation was being localized to intestinal stem cells lining the midgut, however, we did not have suitable antibodies to probe at this possibility.

While we saw an increase in autophagy and EGF expression in the whole insect during metamorphosis, we wanted to visualize their localization spatially in dissected midguts of sand flies. We used antibodies against *ATG6* and *ATG8* to determine localization within the guts. It was extremely interesting to note that we saw *ATG6* staining within the epithelial cells of developing adults within the puparium. To date, most studies implicate its role in the removal of the larval gut, or a pro-apoptotic response [38-41]. However, our results may be indicating a pro-survival and differentiation of newly formed epithelial cells. As mentioned before it would be interesting to see which proteins are interacting with *ATG6* in this particular situation to gain further insights into the mode of action of this protein within this context. It was unfortunate to not see similar staining for the *ATG8* antibodies in the pupal midgut. It is possible that *H. sapiens* *ATG8* antigen used to create these antibodies is different enough from that of *L. longipalpis* to prevent cross-reaction, however this is unlikely (Figure 3.16.). It is more likely that the concentration of antibodies (1:500) used to probe the gut was not sufficient to generate a signal. A previous paper in *A. aegypti* used *ATG8* primary antibodies at a 10-fold higher concentration and were able to very clearly observe autophagosomes with anchored *ATG8* molecules [42]. Another possibility remains that the increased *ATG8* expression seen in RT-qPCR was occurring in other tissue types within the puparium, and not within the midgut.

It is exciting to provide preliminary evidence to further the idea that ecdysone may be able to directly activate *ATG1*’s expression leading to its kinase activity and autophagosome assembly through the promoter directly. There are many studies that demonstrate *ATG1* response to ecdysone, but the mechanism is only now beginning to be unraveled. A very recent paper
delves into the biochemistry of ecdysone response genes and the activation of the ATG1 promoter for expression [48]. While sharing the canonical sequences demonstrated in multiple systems for decades, they do not share same distance between half-sites or putative secondary structure elements. It is likely that they are not the same imperfect palindromes sequences describe in previous work. However, this does not disqualify them as having structural proclivities as other elements previously described above. Here we are able to show that the primary response ecdysone responsive element and GAGA sequences are conserved within the promoter of L. longipalpis. Of course further experimentation would be required to determine if this region of the ATG1 gene is indeed responsive to ecdysone treatment in vivo. However, these preliminary results show promise that autophagy may be hardwired into the process of metamorphosis across different insect orders.
Table 3.1. The accession numbers, identity, score, and expected value for each query protein (*D. melanogaster*) against each subject protein (*L. longipalpis*).

<table>
<thead>
<tr>
<th>Gene Symbol</th>
<th>Accession # <em>D. melanogaster</em></th>
<th>Accession # <em>L. longipalpis</em></th>
<th>Identity</th>
<th>Score</th>
<th>E-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATG1</td>
<td>NP_648601.1</td>
<td>LLOJ007855</td>
<td>79.70%</td>
<td>1431</td>
<td>4.00E-157</td>
</tr>
<tr>
<td>ATG6</td>
<td>NP_651209.1</td>
<td>LLOJ007047</td>
<td>59.30%</td>
<td>1907</td>
<td>0</td>
</tr>
<tr>
<td>ATG8</td>
<td>NP_727447.1</td>
<td>LLOJ007649</td>
<td>96.60%</td>
<td>814</td>
<td>3.00E-80</td>
</tr>
<tr>
<td>Vein</td>
<td>NP_523942.2</td>
<td>LLOJ001534</td>
<td>61%</td>
<td>933</td>
<td>1.00E-84</td>
</tr>
</tbody>
</table>
Table 3.2. The sequence of RT-qPCR primers used for each transcript measured in this study.

<table>
<thead>
<tr>
<th>Gene Symbol</th>
<th>Primer 5’ - 3’ Forward</th>
<th>Primer 5’-3’ Reverse</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATG1</td>
<td>GGTGATTTGGGGGACTAC</td>
<td>ATTCTGGCTTGAGGTC</td>
</tr>
<tr>
<td>ATG6</td>
<td>GCTCCTGTTGATGGTCA</td>
<td>CGTAAGGAGGGAGTTTGATG</td>
</tr>
<tr>
<td>ATG8</td>
<td>GAAGTACCTTGGCCGCTCT</td>
<td>CAGAGATCCATTGATGC</td>
</tr>
<tr>
<td>Vein</td>
<td>CGCAATGGATGAGAACAC</td>
<td>TGAGCAATACCTACGCTGAC</td>
</tr>
</tbody>
</table>
Table 3.3. The species names, gene symbols, accession numbers, and databases for all sequences used for multiple sequence alignment and phylogenetic analysis in this study. *H. sapiens* ATG8 sequence is not derived from a database and may be procured from the company website via that catalog number.

<table>
<thead>
<tr>
<th>Organism</th>
<th>Gene Symbol</th>
<th>Accession Number</th>
<th>Database</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>L. longipalpis</em></td>
<td>ATG1</td>
<td>LLOJ007855</td>
<td>VectorBase</td>
</tr>
<tr>
<td></td>
<td>ATG6</td>
<td>LLOJ007047</td>
<td>VectorBase</td>
</tr>
<tr>
<td></td>
<td>ATG8</td>
<td>LLOJ007649</td>
<td>VectorBase</td>
</tr>
<tr>
<td></td>
<td>Vein</td>
<td>LLOJ001534</td>
<td>VectorBase</td>
</tr>
<tr>
<td><em>P. papatasi</em></td>
<td>ATG1</td>
<td>PPAI003929</td>
<td>VectorBase</td>
</tr>
<tr>
<td></td>
<td>ATG6</td>
<td>FG117924.1</td>
<td>VectorBase</td>
</tr>
<tr>
<td></td>
<td>ATG8</td>
<td>PPAI007833</td>
<td>VectorBase</td>
</tr>
<tr>
<td></td>
<td>Vein</td>
<td>PPAI005726</td>
<td>VectorBase</td>
</tr>
<tr>
<td><em>D. melanogaster</em></td>
<td>ATG1</td>
<td>NP_648601.1</td>
<td>NCBI</td>
</tr>
<tr>
<td></td>
<td>ATG6</td>
<td>NP_651209.1</td>
<td>NCBI</td>
</tr>
<tr>
<td></td>
<td>ATG8</td>
<td>NP_727447.1</td>
<td>NCBI</td>
</tr>
<tr>
<td></td>
<td>Vein</td>
<td>NP_523942.2</td>
<td>NCBI</td>
</tr>
<tr>
<td><em>A. aegypti</em></td>
<td>ATG1</td>
<td>XP_001657286.1</td>
<td>NCBI</td>
</tr>
<tr>
<td></td>
<td>ATG6</td>
<td>XP_001654535.1</td>
<td>NCBI</td>
</tr>
<tr>
<td></td>
<td>ATG8</td>
<td>XP_001652571.1</td>
<td>NCBI</td>
</tr>
<tr>
<td></td>
<td>Vein</td>
<td>XP_001662637.1</td>
<td>NCBI</td>
</tr>
<tr>
<td><em>A. gambiae</em></td>
<td>ATG1</td>
<td>XP_309350.4</td>
<td>NCBI</td>
</tr>
<tr>
<td></td>
<td>ATG6</td>
<td>XP_310418.5</td>
<td>NCBI</td>
</tr>
<tr>
<td></td>
<td>ATG8</td>
<td>XP_312238.3</td>
<td>NCBI</td>
</tr>
<tr>
<td></td>
<td>Vein</td>
<td>XP_001237118.2</td>
<td>NCBI</td>
</tr>
<tr>
<td><em>C. quinquefasciatus</em></td>
<td>ATG1</td>
<td>XP_001842942.1</td>
<td>NCBI</td>
</tr>
<tr>
<td></td>
<td>ATG6</td>
<td>XP_001861442.1</td>
<td>NCBI</td>
</tr>
<tr>
<td></td>
<td>ATG8</td>
<td>XP_001844428.1</td>
<td>NCBI</td>
</tr>
<tr>
<td></td>
<td>Vein</td>
<td>XP_001846605.1</td>
<td>NCBI</td>
</tr>
<tr>
<td><em>G. morsitans</em></td>
<td>ATG1</td>
<td>GMOY006001</td>
<td>VectorBase</td>
</tr>
<tr>
<td></td>
<td>ATG6</td>
<td>GMOY009498</td>
<td>VectorBase</td>
</tr>
<tr>
<td></td>
<td>ATG8</td>
<td>GMOY001732</td>
<td>VectorBase</td>
</tr>
<tr>
<td></td>
<td>Vein</td>
<td>GMOY004977</td>
<td>VectorBase</td>
</tr>
<tr>
<td><em>H. Sapiens</em></td>
<td>ATG8</td>
<td>See Abcam #86497</td>
<td>NA</td>
</tr>
</tbody>
</table>
Figure 3.1. mRNA expression profiles for feeding larvae (L3), non-feeding larvae (L4), early pupae (Pupa), and adult (Adult) sand flies.

(A) The fold change and standard errors for the transcripts encoding ATG1 and ATG6 while. (B) The fold measurements and errors for ATG8 and Vein transcripts. One way ANOVA was performed with a Tukey posttest for multiple comparisons to measure statistical significance (P<0.05, a’ = ns P<0.06) [a’ & b’, represent ANOVA with Tukey test data taken without the biological replicate explained in the results section].
Figure 3.2. Confocal images for ATG6 showing the anterior midgut of newly formed pupae.

The empty lumen (Lu) space of the midgut, DAPI stained nuclei of epithelial cells (Ec), and red staining for ATG6 are shown in (A). The insets show red staining for ATG6 in the cytoplasmic compartments of epithelial cells (B), while a merge of the nuclei and cytoplasm can also be seen (C). Bar indicates 100 microns.
Figure 3.3. Confocal images depicting the anterior sand fly midgut at different developmental stages stained with ATG6 antibodies.

The nuclei of cells are visualized via DAPI staining while ATG6 is in red. The anterior portion of 3rd instar (L3) larvae (A), 4th instar (L4) larvae (B), and adult (C) midguts are shown. Bar indicates 100 microns.
Figure 3.4. Confocal images depicting the anterior sand fly midgut at larval, pupal, and adult developmental stages stained with ATG8 antibodies. The nuclei of cells are visualized using DAPI while ATG8 is in red. L3 (A), L4 (B), Pupae (C), and adult (D) are shown. Bar indicates 100 microns.
Figure 3.5. Possible promoter sequence of region upstream of ATG1

A 1000 nucleotide sequence before the start codon (green) contains a 3 putative ecdysone response elements (red) and a putative GAGA binding sequence (blue). GGTTCA and TGAACCT are canonical EcRE sequences reported in earlier work.
Figure 3.6. The midgut morphological changes in the 3rd instar (L3), 4th instar (L4), pupa, and adult developmental stages.

(A) The average length and standard deviation of insects in each stage. (B) The change in average width across stages. One way ANOVA was performed with a Tukey posttest for multiple comparisons to measure statistical significance (P<0.05).
Figure 3.7. Axial ratio for the midgut (L/W) for L3, L4, pupa, and adult sand flies.

This calculation is based on the length of the midgut divided by the average width of the midgut across its length. One way ANOVA was performed with a Tukey posttest for multiple comparisons for significance (P<0.05).
Figure 3.8. Clustal Omega protein alignment of ATG1 sequences from *L. longipalpis*, *P. papatasi*, *D. melanogaster*, *A. aegypti*, *A. gambiae*, *C. quinquefasciatus*, and *G. morsitans*. 
Figure 3.9. Neighbor-joining tree to infer the evolutionary relationship of ATG1 sequences using *T. castaneum* (TcATG1) as an outgroup.
DmATG6 335 FWDTKDAAMVAFDCLQEQFKEVERKHMPLLPPYMEKGK1LPSGTGNSYS1KIQFNSE
GmATG6 331 FWDTKDAAMVAFDCLQEQFKEVERKHMPLLPPYMEKGK1LPSGTGNSYS1KIQFNSE
LlATG6 359 FWDTKDAAMVAFDCLQQFKEVEGERGSGFCLPYKMK1GK1EDSATGNSYS1KIQFNSE
PpATG6 90 FWDTKDAAMVAFDCLQEQFKEVEGERGSGFCLPYKMK1GK1EDSATGNSYS1KIQFNSE
AgATG6 334 LWDIEAAMVAFDCLQEQFKEVRRDPFLCPYKMK1EDSATGSS1KIQFNSE
CqATG6 335 FWDTKDAAMVAFDCLQEQFKEVVKCFPFLCPYKMK1EDSATGNSYS1KIQFNSE
AaATG6 335 FWDTKDAAMVAFDCLQEQFKEVVKCFPFLCPYKMK1EDSATGNSYS1KIQFNSE

DmATG6 395 EQWTKALKFLLTNLKWGLNWVSQFSP----
GmATG6 391 EQWTKALKFLLTNLKWGLNWVSQFAWQ----
LlATG6 419 EQWTKALKFLLTNLKWGLNWVSQFAPDRLDD
PpATG6 150 EQWTKALKFLLTNLKWGLNWVSQFAPDRLDD
AgATG6 394 EQWTKALKFLLTNLKWGLNWVSQFAPDPDQR--
CqATG6 395 EQWTKALKFLLTNLKWGLNWVSQFAPDPDKH--
AaATG6 395 EQWTKALKFLLTNLKWGLNWVSQFAPDPDPSA

Figure 3.10. Clustal Omega protein alignment of ATG6 sequences from L. longipalpis, P. papatasi, D. melanogaster, A. aegypti, A. gambiae, C. quinquefasciatus, and G. morsitans.
Figure 3.11. Neighbor-joining tree to infer the evolutionary relationship of ATG6 sequences using *A. pisum* (ApATG6) as an outgroup.
Figure 3.12. Clustal Omega protein alignment of ATG8 sequences from *L. longipalpis*, *P. papatasi*, *D. melanogaster*, *A. aegypti*, *A. gambiae*, *C. quinquefasciatus*, and *G. morsitans*.
Figure 3.13. Neighbor-joining tree to infer the evolutionary relationship of ATG8 sequences.
Figure 3.15. Neighbor-joining tree to infer the evolutionary relationship of Vein sequences.
Figure 3.16. Clustal Omega protein alignment of ATG8 sequences from *L. longipalpis* and *H. sapiens* depicting N-terminal region from human antigen used to make antibodies (Abcam #86479) showing similarity to full length ATG8 from sand fly.
Chapter 4 - IMD Recognition, Toll Pathway, and the Effect of Pirk Knockdown in Infected *Lutzomyia longipalpis* Larvae

Abstract

Our previous efforts have elucidated a number of physiological mechanisms occurring within the sand fly with respect to immunity, homeostasis, and autophagy. We have sought to build on these observations in regards to the areas of IMD recognition of foreign particles and possible effects of the Toll pathway during infection with *P. agglomerans* and *B. subtilis*. We also sought to provide additional observations pertaining to the agar based feeding system compared to the normal laboratory chow that insect typically receive with respect to the expression of autophagic and immune molecules. Importantly, our lab was the first to demonstrate that RNAi can be applied to adult sand flies, and subsequently other studies have confirmed this phenomenon. Here I present evidence to suggest that RNAi mediated knockdown of ‘poor immune response upon knockin’ is achievable in the immature larval stage. We show that there is no change in Cactus levels suggesting that Toll pathway may not be playing any significant role in the midgut response to *P. agglomerans* or *B. subtilis*. Also, autophagy transcript levels do not significantly change, further encouraging the idea that agar based feeding methods may be comparable to normal dietary chow on the basis of nutritional value. Most importantly we show for the first time that it is possible to achieve almost 50% RNAi efficiency in the L3 larval stage of the sand fly. This has implications for controlling this vector of blood borne pathogens at the immature stages.
4.1. Introduction

IMD and Toll pathway are both crucial in the ability of insects to fight off microbial insults. Previous results in sand flies have suggested that the immature larval stages may suppress their IMD response via Pirk expression to allow *B. subtilis* and *P. agglomerans* to reside within the midgut of sand flies [79]. However, there is no work conducted in sand flies to measure any phenomenon related to Toll pathway or IMD recognition of foreign organisms and molecules. Additionally, the agar feeding regimens used to determine the effect of Pirk on effector molecule expression are quite different than the natural detritus material that sand flies are generally exposed to in the lab and nature. While there are data on the efficiency of RNAi mediated knockdown in adult sand flies [18, 24], no published data are currently aimed to elucidate the efficiency of RNAi mediated knockdown of sand fly transcripts in the immature larval stages. Here we investigate the negative regulator of Toll signaling, Cactus, as a marker for measuring Toll pathway activity. Although Cactus activity is generally described in hemocyte interactions with gram-positive bacterial and fungal organisms or dorso-ventral development [25, 26], we wished to see if it may have so effect in the midgut against gram-positive *B. subtilis*. PGRP-LB is a well-studied recognition molecule for the IMD pathway, known for its ability to protect *W. wigglesworthia* within the bacteriome of the tsetse fly midgut [27]. Here we measure the sand fly ortholog after infection to see any changes in expression level. With respect to chow versus agar fed insects we assess changes in the autophagic markers ATG1, 6, & 8; in addition to the immune regulators USP36 and IMD. Using Pirk as a target for injection mediated dsRNA delivery we measure its efficiency and effect on other immune transcripts of interest.
4.2. Material & Methods

4.2.1. Sand fly colony maintenance

*L. longipalpis* (Jacobina strain – LLJB) colonies were reared in the Department of Entomology, Kansas State University. Larvae were maintained in 250 or 500 ml plastic jars (Nalgene) with an approximately 2 cm-thick bed made of dental plaster (Schein), and fed on larval chow (a mixture of 50% rabbit droppings and 50% rabbit food).

4.2.2. Double stranded RNA synthesis and injection

Double stranded RNA targeting *Pirk* or *GFP* transcripts were synthesized using the primer sequences described to generate a 269 bp fragment for dsPirk and 300 bp fragment for dsGFP (Table 4.2., Figure A.2.). Templates for dsRNA synthesis were generated via GoTaq Green mastermix (Promega) PCR using 95º C for 2 min followed by 30 cycles of: 95º for 30s melting, 58º C annealing, and 72º C extension, with a final 10 min extension time at 72º C. One ug of template DNA was used to synthesize dsRNA overnight at 37º C. dsRNA molecules were column purified per the manufacturers instruction and eluted in 2 rounds of Hyclone (Thermo Scientific) water and concentrated to 3 µg/µL using the YM-100 filter columns (Millipore).

Third instar larvae were place in a cold dish, then injected between the head and 1st thoracic segment with 23 nL of dsRNA using a Nanojet II microinjector (Drummond). Larvae were then transferred to a humid Nalgene plaster container for approximately 12h to recover before being exposed to their respective diet treatments. Feeding post injection was taken hours after recovery from injection and being placed on diet.
4.2.3. RNA extraction and reverse transcription

Whole insects from *L. longipalpis* feeding on different treatments at 12, 24, and 36h post injection were selected under a stereoscope microscope in Hyclone (Thermo Scientific) phosphate buffered saline (PBS). Total RNA was isolated from pools of 5 midguts in the case of expression profile analysis, and 5 whole insects in the case of dsRNA injections using TRIzol (Invitrogen). For each group of midguts or insects, RNA isolation was conducted in triplicate. RNA quality was assessed by electrophoresis on 1% agarose-5% formaldehyde in 1x MOPS gel, and stored at -80 °C. First strand cDNA synthesis was conducted using Superscript™ III reverse transcription kit (Invitrogen) as described [24].

4.2.4. Real time quantitative PCR

mRNA levels for different developmental stages were quantified with iQ SYBER Green Supermix (Bio-Rad) using 95º C melting, 57º C annealing, and 72º C extension temperature for 40 cycles using a Realplex Master cycler (Eppendorf). Relative fold changes were assessed using the ∆∆Ct method [24, 52, 79], and calibrated against the expression of each transcript for the control fed and/or uninjected L3 larvae [52]. Sequences for obtained by taking sequences from *D. melanogaster* and using tBlastn and/or Blastp against *L. longipalpis* sequences found at VectorBase (https://www.vectorbase.org/blast). Primer sequences were generated using Primer3 software (http://biotools.umassmed.edu/bioapps/primer3_www.cgi). PCR amplicon sequence was verified using direct sequencing at Kansas State University. All Blastp data and primer sequences used in this study are summarized in Tables 3.1., 3.2., 4.1., and 4.2.
4.2.5. Predicted amplicon sequences, alignments, and phylogeny

Sequences were obtained for various orders of insects using Blastp at NCBI or VectorBase (summary of organisms, accession numbers, and databases are provided in supplementary Tables 3.3. and 4.3.). Multiple sequence alignments at the protein level were conducted using Clustal Omega software (http://www.ebi.ac.uk/Tools/msa/clustalo/). Alignments were then visualized using BoxShade (http://www.ch.embnet.org/software/BOX_form.html). To determine phylogeny, neighbor joining trees were generated with 1000 bootstrap replicates using Mega7 (www.megasoftware.net).

4.3. Results

4.3.1. Expression of Toll pathways, IMD recognition and signal transduction molecules after infection

To build on our previous studies concerning the effect of Bs and Pa infection on midgut immune molecule expression we assessed the transcript levels of a negative regulator of Toll signaling (Cactus), a negative regulator of IMD pathway (Caudal), and a putative enzymatically active IMD recognition protein (PGRP-LB) 12 and 24h post infection. Although Bs is classified as gram-positive bacteria, it produces DAP-type peptidoglycans on its surface and may be better recognized by gram-negative IMD pathway instead of Toll. Additionally, we wanted to compare our previous agar based feeding method to traditional dietary chow lab larvae are accustomed to. At both 12 and 24h post infection it is observed that the level of Cactus remains unchanged no matter what feeding regimen is employed (Figure 4.1.), suggestive of at least no change in the regulation of the Toll pathway during infection experiments. Caudal levels do increase in Bs and
Pa more than 2 fold 12h post infection, but not significantly (Figure 4.1.A), and continue this increase for all feeding regimen at 24h (Figure 4.1.B). We also observe a non-significant increase in PGRP-LB levels for Agar fed insects for control and Bs at 12h, and for all 3 agar feeding regimens at 24h (Figure 4.1.).

4.3.2. Comparison of chow fed, agar fed, and infected individuals for autophagy and immunity transcripts

In continuing to assess differences between the laboratory chow diet and agar based feeding methods we measured transcript levels of ATG1, 6, and 8 to see if there were any changes between feeding conditions that maybe related to nutrition and stress. Additionally, we measured the levels of USP36 and IMD transcripts to see if there may be differences associated with immunity when larvae were fed chow, with its own associate microbial community, as opposed to agar based diets that contain controlled microbial communities. We matched the time points from the first study to see if any significance was observed [79]. With respect to the autophagy transcripts surveyed there were no significantly different changes observed across feeding treatments. However, there is a variable increase in the mean for agar fed regimens versus chow fed for ATG8 at all time points (Figure 4.2.A, C, E). Interestingly, we see a decrease in the expression level of immune transcripts USP36 and IMD that increase with time. At 12h post infection this difference is not significantly noticeable, but at 24h (P<0.001) and 36h (P<0.01) post infection the expression of both USP36 and IMD is diminished significantly for agar fed versus chow fed (Figure 4.2.B, D, F).
4.3.3. RNAi efficiency for Pirk transcript in L3 larvae stage

Our previous study demonstrated that Pirk may play a significant role in suppressing IMD effector molecule production after infection with microbes [79]. Here we injected dsRNA for *GFP, Pirk*, or the larvae remained uninjected. We allowed larvae to recover for 12h and then placed them on their respective chow or agar based feeding regimens and measured *Pirk* transcript level for individually injected insects at 12 and 24h post injection. Our initial measurements for efficiency of knockdown for insects fed on chow indicated a >50% non-significant reduction for the *Pirk* transcript when compared to uninjected or dsGFP injected individuals (Figure 4.3.A & B). However, this particular experiment only included n=5 individuals per time point per treatment. Combining this knockdown efficiency data with later knockdowns including all individuals placed on all feeding regimens, n=20 per time point per treatment, yielded a significant 40% knockdown efficiency of dsPirk individuals compared to uninjected and dsGFP injected individuals by 24h post injection, but not 12h post injection (Figure 4.3.C & D). This evidence suggests a partial knockdown of *Pirk* transcripts in feeding L3 larvae is achievable and can be further investigated.

4.3.4. Effect of dsPirk knockdown on transcript levels related to IMD pathway

Observing that RNAi mediated knockdown of *Pirk* transcript levels was at least partially effective, we moved on to see what effect *Pirk* transcript knockdown may have on other immune related molecules. Following the previously mentioned injection procedure, we expanded our studies to include treated insects fed on control agar, Bs agar, or Pa agar 12 and 24h post injection. Our previous study suggested that Pirk may be suppressing the anti-microbial peptide attacin, and modulating changes in IMPer and USP36 levels [79]. However, at no observed time
point or treatment was any significant difference seen between uninjected, dsGFP, and dsPirk injected individuals fed on control versus infected agar. There was a very high degree of variability associated with a number of samples that is discussed further, and IMPer levels for Pa fed 12h post injection were not determine. The data for Pirk expression represented for each time point and feeding regimen is, however, more stable and is included in the previous injection reports (Figures 4.4.-4.9.).

4.3.5. Supplementary sequencing, amplicon, BLAST, and alignment results

In order to further explain the studies presented here and in chapters 2 & 3 additional information is provided now and discussed hereafter. While specific alignments and phylogenetic trees are provided in chapter 2 to support the sequences of Vein, ATG1, ATG6, and ATG8 further information is provided here for the putative sequence of the transcript, sequence of the amplicons used, and the position of the primers to generate said amplicons (Figures A.1., A.2., A.5., A.6., A.9.-11., A.14., A.17., A.20., A.23., & A.26.). For the sequences used here and in Chapter 2 similar data are also provided, and when appropriate discussed further. All BLAST data used to obtain these sequences unless otherwise stated is summarized (Table 4.3.), and additional PCR primers used are also provided (Tables 2.4., 3.2., and 4.2.). With regard to the position of the Pirk amplicon and dsRNA primer positions data showing a non-overlap of probe and dsRNA targeting region are provided (Figure A.2.). Also, accession numbers are provided (Table 4.3.) for the sequences found in the multiple sequence alignments (Figures A.4., A.8., A.13., A.16., A.19., A.22., A.25., & A.28.), and neighbor joining phylogenetic trees (A.3., A.7., A.12., A.15., A.18., A.21., A.24., A.27.,). Specific details about Blast data and sequence similarity are discussed below.
4.4. Discussion

We had previously sought to understand effector molecule production and IMD regulation in our previous work [79]. Here we further explore other possible transcripts that may play a role in Toll and IMD pathways.

During conditions where fungal and/or Gram-positive microbes are not detected by an insect Cactus generally remains stably bound to dorsal preventing its entry into the nucleus and downstream transcription of Toll immune effector molecules, and also plays a role in early dorso-ventral development [25, 26]. However, when recognition of these organisms and their associated molecules occurs phosphorylation and degradation of Cactus occurs and the transcription factors may enter the nucleus [92]. *B. subtilis* is an interesting microbe to study in this scenario. While it is characterized as a gram-positive bacteria, it presents DAP type peptidoglycans on its surface, making it more putatively recognizable by PGRPs associated with IMD pathway. However, we did not detect any noticeable decrease in Cactus transcript level at any time point or treatment suggesting that there is no change in expression for this negative regulator of immunity and midgut immunity in this case may be dictated by the regulation of IMD pathway.

Caudal is typically associated with developing *D. melanogaster* at the embryonic stage [93]. In the case of immunity, caudal is known for its prohibitory action in allowing Relish to enter the nucleus and initiate transcription of immune effector molecules [94]. While we had previously shown that *Pirk* was capable of providing negative feedback to the IMD pathway we wanted to observe what effect caudal might have in the midguts of sand flies infected with *B. subtilis* or *P. agglomerans*. Although the increase was not statistically significant, it was apparent that infected midguts had a greater mean expression of this molecule. While this is not
conclusive, it does not rule out the role of caudal in inhibiting IMD pathway responses in the midgut when infection with commensal bacteria occurs.

IMD pathway recognition is a very complex process that includes numerous peptidoglycan recognition proteins that may be secreted or retain; enzymatically active or inactive. For the sake of our experiment we chose PGRP-LB, a protein that has been demonstrated in protecting the obligate symbiont of tsetse flies, *W. glossinidae*. Interestingly, we see an increase for PGRP-LB in all agar feeding treatments used by 24h. While this increase is not statistically significant it does point a possible role of this molecule for both *B. subtilis* and *P. agglomerans* recognition. It is again important to note that recognition of DAP type peptidoglycans associated with gram-negative bacteria is extremely complex and includes a number of: short, long, enzymatically active, and enzymatically inactive PGRPs as discussed in chapter 1 and the introduction herein. It would require much more extensive analysis of expression patterns (RNA-SEQ or microarray) of infected individuals to more readily acquire the data necessary for delineating these associated complexities.

Our previous studies also involved measuring USP36 and IMD transcript levels during the hours post infection, and the levels of ATG1, 6, and 8 during metamorphosis. We sought to investigate how our agar based feeding system may compare to the normal day to day chow diet the insects were acclimated to. While autophagy plays a role in the rearrangement of cells and tissues, it is predominately linked to its role in TOR signaling and nutrient level sensing. Since the agar based feeding method is starkly different than the detritus materials that sand fly larvae are accustomed to, we measured the levels of autophagic transcripts to gain insights into how this pathway may be behaving. Surprisingly, we did not see any significant change between chow and agar fed insects, although there was a mean increase in autophagy markers suggesting that
the agar diet may be less than a sufficient source of nutrients. One important take away from these two feeding methods is the change in expression of immune molecules. USP36 and IMD were both significantly down regulated in the hours post agar feeding. This points to a possible global down regulation immune molecules in the midgut of insects fed a more simplified and directed microbial community.

Understanding the RNAi effect at all life stages of insect development is important in developing putative and practical strategies for controlling insect populations and competence for pathogens. Previously demonstrated by our group was a highly efficient knockdown the transcript encoding chitinase 1 in *P. papatasi* [24]. Also, effective knockdown of the transcript encoding Caspar, a negative regulator of immunity, was achieved by another group in *L. longipalpis* [18]. However, these studies focused on the adult stage of the sand fly and its competence for *Leishmania* parasites. Here we investigated the possibility of suppressing a midgut immune suppressor *Pirk* in an attempt to understand the fundamental basics of sand fly larvae gut immunity. The larval stage was difficult to inject do to the inherent thickness of its integument, and a suitable spot for injection of L3 larvae was only found in the inter-segmental region between the head capsule and the first thoracic segment. We were able to achieve significant suppression of *Pirk* transcripts approaching 50% by 24h post injection when considering all injected insects. This is a promising result for the possible control of larval sand fly populations via RNAi, especially considering its possible use as a ‘platform’ of paratransgenesis [78].

Using this injection method we attempted to see what effects knockdown might have on other immune related transcripts during infection situations. However, this experiment proved to
extremely laborious and expensive, and its current form produced highly variable results that did not provide much valuable expression data.

The discovery, at times, and analysis of putative transcripts within the *L. longipalpis* genome was a very time intensive process. Although there are numerous EST libraries available for *L. longipalpis*, only a few of the genes predicted and RT-qPCR amplicons synthesized for were supported by EST data. I will discuss the implication of the few that were indeed supported. ATG1 and 6 are both transcripts involved in autophagy, and are often associated with nutritional sensing. ATG6 ESTs (AM092571.1, AM092621.1) were found in total EST libraries. Interestingly, an ATG1 EST (EX991142.1) was found during adult female blood digestion, which coincides with the requirement of ATG1 in *A. aegyti* for digestion and egg production [42, 95]. Also, a USP36 EST (AM108362.1) was found in the total EST library, and more interestingly, ESTs for Attacin (EX211749.1) and IMFer (EX211579.2, EX212073.1) were found in adult females infected with *Leishmania* parasites. These evidences further the possibility of these transcripts studied in larvae being of importance in the adult form of the sand fly *L. longipalpis*. 
Table 4.1. The accession numbers, identity, score, and expected value for each query protein (*D. melanogaster*) against each subject protein (*L. longipalpis*).

<table>
<thead>
<tr>
<th>Gene Symbol</th>
<th>Accession # D. melanogaster</th>
<th>Accession # L. longipalpis</th>
<th>Identity</th>
<th>Score</th>
<th>E-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attacin</td>
<td>AAL23662.1</td>
<td>LLOJ005408</td>
<td>40.50%</td>
<td>220</td>
<td>1.00E-14</td>
</tr>
<tr>
<td>Domeless</td>
<td>NP_523412.1</td>
<td>LLOJ002449</td>
<td>27.75</td>
<td>987</td>
<td>3.00E-81</td>
</tr>
<tr>
<td>Duox</td>
<td>NP_608715.2</td>
<td>LLOJ010494</td>
<td>94.30%</td>
<td>4361</td>
<td>0</td>
</tr>
<tr>
<td>IMD</td>
<td>NP_573394.1</td>
<td>LLOJ010482</td>
<td>40.40%</td>
<td>422</td>
<td>2.00E-34</td>
</tr>
<tr>
<td>Pirk</td>
<td>NP_611598.1</td>
<td>LLOJ004926</td>
<td>40%</td>
<td>329</td>
<td>3.00E-26</td>
</tr>
<tr>
<td>USP36</td>
<td>NP_001246639.1</td>
<td>LLOJ008157</td>
<td>53%</td>
<td>1968</td>
<td>0</td>
</tr>
<tr>
<td>NOS</td>
<td>AAF25682.1</td>
<td>LLOJ005465</td>
<td>59.70%</td>
<td>2081</td>
<td>0</td>
</tr>
<tr>
<td>IMPer</td>
<td>NP_996223.1</td>
<td>LLOJ010504</td>
<td>36.30%</td>
<td>1153</td>
<td>1.00E-103</td>
</tr>
</tbody>
</table>
Table 4.2. The sequence of RT-qPCR primers used for each transcript measured in this study.

<table>
<thead>
<tr>
<th>Gene Symbol</th>
<th>Primer 5’-3’ Forward</th>
<th>Primer 5’-3’ Reverse</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nos</td>
<td>CAACTACGAGCCAAATGTTC</td>
<td>AGCCAGCATAGAGGATAAGC</td>
</tr>
<tr>
<td>dsPirk</td>
<td>TAATACGACTCATATAGGGGCATGCAATGC</td>
<td>TAATACGACTCATATAGGGGATATTGGCGGAATTGTT</td>
</tr>
<tr>
<td>dsGFP</td>
<td>TAATACGACTCATATAGGGGATGGTGAAGCAAGGGCGAGGA</td>
<td>TAATACGACTCATATAGGGTagGGTTGTTGTTGTCGGGCAGCAG</td>
</tr>
<tr>
<td>Caudal</td>
<td>ATCGCAGAGTCATTCAAGTG</td>
<td>GTGATGGTGGAGGGATATG</td>
</tr>
<tr>
<td>Cactus</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>PGRP-LB</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>
Table 4.3. The species name, gene symbol, accession number, and database for all sequences used for multiple sequence alignment and phylogenetic analysis.

<table>
<thead>
<tr>
<th>Species Name</th>
<th>Gene</th>
<th>Accession Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>L. longipalpis</td>
<td>Pirk</td>
<td>LLOJ004926</td>
</tr>
<tr>
<td>P. papatasi</td>
<td>Pirk</td>
<td>AJVK01084806.1</td>
</tr>
<tr>
<td>T. castaneum</td>
<td>Pirk</td>
<td>XP_008193927.1</td>
</tr>
<tr>
<td>A. albopictus</td>
<td>Pirk</td>
<td>KXJ77933.1</td>
</tr>
<tr>
<td>L. migratoria</td>
<td>Pirk</td>
<td>AJF38200.1</td>
</tr>
<tr>
<td>D. busckii</td>
<td>Pirk</td>
<td>ALC41356.1</td>
</tr>
<tr>
<td>D. melanogaster</td>
<td>Pirk</td>
<td>NP_611598.1</td>
</tr>
<tr>
<td>L. longipalpis</td>
<td>Attacin</td>
<td>LLOJ005408</td>
</tr>
<tr>
<td>P. papatasi</td>
<td>Attacin</td>
<td>PPAI003791</td>
</tr>
<tr>
<td>D. melanogaster</td>
<td>Attacin</td>
<td>AAL23662.1</td>
</tr>
<tr>
<td>A. aegyti</td>
<td>Attacin</td>
<td>XP_001656768.1</td>
</tr>
<tr>
<td>T. castaneum</td>
<td>Attacin</td>
<td>XP_001809216.1</td>
</tr>
<tr>
<td>M. domestica</td>
<td>Attacin</td>
<td>XP_011296538.1</td>
</tr>
<tr>
<td>L. longipalpis</td>
<td>Duox</td>
<td>LLOJ010494</td>
</tr>
<tr>
<td>P. papatasi</td>
<td>Duox</td>
<td>PPAI003114</td>
</tr>
<tr>
<td>C. quinquefasciatus</td>
<td>Duox</td>
<td>XP_001844503.1</td>
</tr>
<tr>
<td>D. melanogaster</td>
<td>Duox</td>
<td>NP_608715.2</td>
</tr>
<tr>
<td>A. gambiae</td>
<td>Duox</td>
<td>XP_319115.4</td>
</tr>
<tr>
<td>A. aegyti</td>
<td>Duox</td>
<td>XP_001658452.2</td>
</tr>
<tr>
<td>L. longipalpis</td>
<td>USP36</td>
<td>LLOJ008157</td>
</tr>
<tr>
<td>P. papatasi</td>
<td>USP36</td>
<td>PPAI006417</td>
</tr>
<tr>
<td>A. aegyti</td>
<td>USP36</td>
<td>XP_001662174.1</td>
</tr>
<tr>
<td>A. gambiae</td>
<td>USP36</td>
<td>XP_001237068.2</td>
</tr>
<tr>
<td>D. melanogaster</td>
<td>USP36</td>
<td>NP_001246639.1</td>
</tr>
<tr>
<td>C. quinquefasciatus</td>
<td>USP36</td>
<td>XP_001848432.1</td>
</tr>
<tr>
<td>L. longipalpis</td>
<td>IMD</td>
<td>LLOJ010482</td>
</tr>
<tr>
<td>P. papatasi</td>
<td>IMD</td>
<td>PPAI001062</td>
</tr>
<tr>
<td>T. castaneum</td>
<td>IMD</td>
<td>XP_008199405.1</td>
</tr>
<tr>
<td>D. melanogaster</td>
<td>IMD</td>
<td>NP_573394.1</td>
</tr>
<tr>
<td>A. gambiae</td>
<td>IMD</td>
<td>XP_001688860.1</td>
</tr>
<tr>
<td>A. aegyti</td>
<td>IMD</td>
<td>XP_001660624.1</td>
</tr>
<tr>
<td>C. quinquefasciatus</td>
<td>IMD</td>
<td>XP_001861391.1</td>
</tr>
<tr>
<td>Species</td>
<td>Gene</td>
<td>Accession</td>
</tr>
<tr>
<td>--------------------------</td>
<td>-------</td>
<td>-------------</td>
</tr>
<tr>
<td><em>G. morsitans</em></td>
<td>IMD</td>
<td>GMOY013006</td>
</tr>
<tr>
<td><em>L. longipalpis</em></td>
<td>Nos</td>
<td>LLOJ005465</td>
</tr>
<tr>
<td><em>P. papatasi</em></td>
<td>Nos</td>
<td>PPA005982</td>
</tr>
<tr>
<td><em>A. gambiae</em></td>
<td>Nos</td>
<td>XP_555207.2</td>
</tr>
<tr>
<td><em>A. gambiae</em></td>
<td>Nos</td>
<td>XP_317213.1</td>
</tr>
<tr>
<td><em>D. melanogaster</em></td>
<td>Nos</td>
<td>AAF25682.1</td>
</tr>
<tr>
<td><em>A. aegyti</em></td>
<td>Nos</td>
<td>XP_001660328.1</td>
</tr>
<tr>
<td><em>L. longipalpis</em></td>
<td>Domeless</td>
<td>LLOJ002449</td>
</tr>
<tr>
<td><em>P. papatasi</em></td>
<td>Domeless</td>
<td>PPA001065</td>
</tr>
<tr>
<td><em>D. melanogaster</em></td>
<td>Domeless</td>
<td>NP_523412.1</td>
</tr>
<tr>
<td><em>A. gambiae</em></td>
<td>Domeless</td>
<td>XP_319240.5</td>
</tr>
<tr>
<td><em>A. aegyti</em></td>
<td>Domeless</td>
<td>XP_001662596.1</td>
</tr>
<tr>
<td><em>A. pisum</em></td>
<td>Domeless</td>
<td>XP_003245751.1</td>
</tr>
<tr>
<td><em>C. quinquefasciatus</em></td>
<td>Domeless</td>
<td>XP_001867486.1</td>
</tr>
<tr>
<td><em>L. longipalpis</em></td>
<td>IMPer</td>
<td>LLOJO10504</td>
</tr>
<tr>
<td><em>P. papatasi</em></td>
<td>HPX1</td>
<td>PPA009674</td>
</tr>
<tr>
<td><em>A. gambiae</em></td>
<td>HPX4</td>
<td>AGAP007237</td>
</tr>
<tr>
<td><em>A. aegyti</em></td>
<td>HPX4</td>
<td>AAEL018292</td>
</tr>
<tr>
<td><em>A. gambiae</em></td>
<td>HPX15IMPER</td>
<td>XP_003436891.1</td>
</tr>
<tr>
<td><em>C. quinquefasciatus</em></td>
<td>peroxidase</td>
<td>PPA009674</td>
</tr>
<tr>
<td><em>D. melanogaster</em></td>
<td>peroxidase</td>
<td>NP_996223.1</td>
</tr>
</tbody>
</table>
Figure 4.1. mRNA expression levels for L3 larvae fed normal diet (chow), control agar (Con), B. subtilis agar (Bs), and P. agglomerans agar (Pa).

(A) The fold change with standard errors for the transcripts negative regulator of Toll pathway (Cactus), Caudal, and PGRP-LB after 12h of feeding on their respective substrates. (B) The same measurements after 24h feeding. One way ANOVA was performed with a Tukey posttest for multiple comparisons to measure statistical significance.
Figure 4.2. mRNA expression levels for L3 larvae fed normal diet (chow), control agar (Con), *B. subtilis* agar (Bs), and *P. agglomerans* agar (Pa).

(A, C, & E) The fold change with standard errors for the autophagy transcripts: 1, 6, & 8 after: 12, 24, & 36h of feeding on their respective substrates. (B, D, F) The same measurements the immune transcripts USP36 and IMD. One way ANOVA was performed with a Tukey posttest for multiple comparisons to measure statistical significance (** P<0.01 & *** P<0.001).
Figure 4.3. mRNA expression levels for L3 larvae that were uninjected, injected with dsGFP, or injected with dsPirk RNA.

Larvae were fed on chow, control agar, B. subtilis or P. agglomerans agar. (A) The fold change with standard errors for Pirk expression at 12h post injections feeding on chow, (B) 24h post injection for samples feeding on chow alone. (C) Pirk expression levels 12h post injection for all treated insects feeding on their respective substrates combined, and (D) 24h post injection. One way ANOVA was performed with a Tukey posttest for multiple comparisons to measure statistical significance (* P<0.05).
Figure 4.4. mRNA expression levels for L3 larvae that were uninjected, injected with dsGFP, or injected with dsPirk RNA.

Larvae were fed on control agar. (A) The fold change with standard errors for Attacin expression at 12h post injections, and (B, C, & D) 12h post injection for transcripts IMPer, Pirk, and USP36 respectively. One way ANOVA was performed with a Tukey posttest for multiple comparisons to measure statistical significance (* P<0.05).
Figure 4.5. mRNA expression levels for L3 larvae that were uninjected, injected with dsGFP, or injected with dsPirk RNA.

Larvae were fed on *B. subtilis* agar. (A) The fold change with standard errors for Attacin expression at 12h post injections, and (B, C, & D) 12h post injection for transcripts IMPer, Pirk, and USP36 respectively. One way ANOVA was performed with a Tukey posttest for multiple comparisons to measure statistical significance.
Figure 4.6. mRNA expression levels for L3 larvae that were uninjected, injected with dsGFP, or injected with dsPirk RNA.

Larvae were fed on P. agglomerans agar. (A) The fold change with standard errors for Attacin expression at 12h post injections, and (B, C, & D) 12h post injection for transcripts IMPer (NA = not available), Pirk, and USP36 respectively. One way ANOVA was performed with a Tukey posttest for multiple comparisons to measure statistical significance.
Figure 4.7. mRNA expression levels for L3 larvae that were uninjected, injected with dsGFP, or injected with dsPirk RNA.

Larvae were fed on control agar. (A) The fold change with standard errors for Attacin expression at 24h post injections, and (B, C, & D) 24h post injection for transcripts IMPer, Pirk, and USP36 respectively. One way ANOVA was performed with a Tukey posttest for multiple comparisons to measure statistical significance.
Figure 4.8. mRNA expression levels for L3 larvae that were uninjected, injected with dsGFP, or injected with dsPirk RNA.

Larvae were fed on *B. subtilis* agar. (A) The fold change with standard errors for Attacin expression at 24h post injections, and (B, C, & D) 24h post injection for transcripts IMPer, Pirk, and USP36 respectively. One way ANOVA was performed with a Tukey posttest for multiple comparisons to measure statistical significance.
Figure 4.9. mRNA expression levels for L3 larvae that were uninjected, injected with dsGFP, or injected with dsPirk RNA.

Larvae were fed on *P. agglomerans* agar. (A) The fold change with standard errors for Attacin expression at 24h post injections, and (B, C, & D) 24h post injection for transcripts IMPer, Pirk, and USP36 respectively. One way ANOVA was performed with a Tukey posttest for multiple comparisons to measure statistical significance.
References


Appendix A - Bioinformatic and Phylogenetic Data

>LL_Vein_LLOJ001534
ATGAATGTGCGTGCTCCTGGCAATGATGCTCATGAGAACTCTAATGTACCTGTGGATTGTGGCCTGCA
CAGTACTAATATGATGGTGTATTACGTGAGGCTGCTTCACCATGCAACACAGGAGATTTTATTAGTCC
TGCAAGGAATTCCACCCTCACCACCACCGTGTTGGAAGGCTGCTTTCCGCCATGCACATCACACCGAGTTCTATTAG
GTACACACAAACCAGGACAAACTTTCACGGCAGGACATATTTCACACGACATTTTCTCCGATAGGGTATGGC
TACTGCTGCAACATCATTGCCACAAGAGCAAGAGCAGATGATGAGAACAC
GCGGCCAAG
ACTG
AGTCGGCGTCC
ATGA
TGGAGGCGCTAGTTACCTCATGCGGCCCTATGGG
CGAAATAGGCGCGATGAACTGGAGCAGGGGAAGGTGTACTTTCTTTTTCTCAAGCAATCACCTTGGG
AATTCACAATCATCGCAGCATCCATTGAATTTCTCACGGAAAATCTTCCCGTGCGAGAGGGCAAGAAAG
TCCGTTTGATATGCAGAGTGAAAGGTCAACCGCCACCGAAAATCACATGGTACAAAGATGGGAGTTCCAT
TAAAAGAATCTGATTGATATACTTTTTCTCAGTTAATCCTAATCCAATCTGATGCTGGTA
GGAAGAAAGAACAGACAACGGGACATATGAATGTCGCGCTAAGAATAATTGGCGAAACAACCTGTTACGAA
CAATGTTCTTGAGTGACGTGCCAGCGATCCAAAGTTGGGAATTTGCCCATGGGAGGAGTTACGGA
TTGCTTTGATATGCAAGATGGAAGGCTCAAACCACCAGGAGACATAGTGTCAAAGATGGGAGTCCAT
TTATGGGGAATCGCTGTGAGACGAAAAACCCAACAAATGCCCAAAACGATTACGTACACAAAGACAAT
GGGCACAAAAATATGATTTATGCTGATTTGTTAAAGGGTCTTATCCAGCAATGGGGTGGT
TTATTTTTTTACAAAAAAGCCATCGGAGAAATTCAAGAGACCTCATTTTTTCTAAACATTTTTG
AAAAATCTATAATAATATGCTCTATCATAATAATATATTTTTTTCTACAAATAAAAATTGAGAAAATCAAG
CAAGATTTTTTATATGACAAATAATTTTGTATAATTATTATTTTTAAAAAGGGAAATGAGAAATTTCAGATTTTG
CTTTATGTAAGCTAAAGAAATTCTATTTATTGCTCAATATGCAAGAATAATGTAATAACTTATTTTTTTA
ATTTAAACATTTGACGTTAGAGAATGTGAAAAAAGAGATATTTCTAA

Figure A.1. Putative CDS for mRNA encoding L. longipalpis Vein.

Full length predicted mRNA is depicted in black, while the primers used for RT-qPCR amplicon measurement are shown in pink. Sequences in blue denote agreement between prediction and sequenced amplicons, while nucleotides in red depict a difference at that base pair.
>Ll_Pirk_LLOJ004926
ATGGTGATTAAAGATGAGTGGGAGTGAAGATGCAGCAGCGTGCTGAAACTAATTGGGAATCGCAAGGAGT
TCCGAATTGAGAAGAATGATGCCCATTTGCGTATTGTTGGGAACGATAATCGCATCCTGGTGACACCACAA
CAATGCGTCCGTTGGGACATTGTGTTGGGAATTCTCAACTGTATAATGGGCTCAGAATGCAGGAAGTATCTC
TACACGGGAATTTGTTGGACCGCATCGTCTCCTGGTGAGGATTCATACCAACGTCACACTACACCGGAA
GTGCCGGACGGATGAAGCTGATTGTGGCAGAAGCACAAATGCGCTGTGGGAAGAAAAATCCCGAATGATAAGAA
AGACATCGTAAGACCGTCGGGACAGCAAGCGTGGGAGCAGCCAGCGATCTCCTCTGATGCCACTATGTT
CACATCAACACCAATTCGGCAATATCGTCATCAACAAATACAAATAATTGATGA

Figure A.2. Putative CDS for mRNA encoding *L. longipalpis* Pirk.

Full length predicted mRNA is depicted in black, while the primers used for RT-qPCR amplicon measurement are shown in pink, while primers used for the synthesis of dsPirk RNA molecules are shown in violet. Sequences in blue denote agreement between prediction and sequenced amplicons, while nucleotides in red depict a difference at that base pair.
Figure A.3. Neighbor-joining tree to infer the evolutionary relationship of Pirk sequences.
### Figure A.4. Clustal Omega protein alignment of Pirk sequences from *L. longipalpis*, *P. papatasi*, *D. melanogaster*, *D. buskii*, *A. albopictus*, *T. castaneum*, *S. oryzae*, and *L. migratoria.*
Figure A.5. Putative CDS for mRNA encoding *L. longipalpis* ATG6.

Full length predicted mRNA is depicted in black, while the primers used for RT-qPCR amplicon measurement are shown in pink. Sequences in blue denote agreement between prediction and sequenced amplicons, while nucleotides in red depict a difference at that base pair.
Figure A.6. Putative CDS for mRNA encoding *L. longipalpis* Attacin.

Full length predicted mRNA is depicted in black, while the primers used for RT-qPCR amplicon measurement are shown in pink. Sequences in blue denote agreement between prediction and sequenced amplicons, while nucleotides in red depict a difference at that base pair.
Figure A.7. Neighbor-joining tree to infer the evolutionary relationship of Attacin sequences.
Figure A.8. Clustal Omega protein alignment of Attacin sequences from *L. longipalpis*, *P. papatasi*, *D. melanogaster*, *A. aegyti*, and *T. castaneum*. 
>L1_ATG1_LLOJ007855
ATGATGGAGATTTGCTGTGGGATATTGACATTCCACACAGACAGAGTTCTTGTGGGACTATTGAGTACAACAGCAAAGATCTCATTGGGCACGGGGCATTTGCTGTTGTGTACAA
GGGACGCCACAGGAAGAAGGCAAGCTGCTGACATGAGAAGCGTGTCGTCGCTGCTGCTGAC
TGCAAGGAGATCCTCCACCACTCAATTGCTCTCCTGTCGATGCTGCTGCTGAC
GGTGAATGGGCGGACTAAGTCAAAATCTCTTGGGAGGAGATAAAATACTCAAAGATGCTACGCTGACCCATGAGAAGCGTGTCGTCGCTGCTGCTGAC
TGTTCAACCATGTGACTGCGCCGCCAGGGTTGATAATGCTCCTCCACGATAGACCCCAAGCGAGATTGTGTCGTTGCTGCTGCTGAC
GAACTATTGTTATATCTACCTGCTCGTACGGCGTATCGCATCCTGCTGCTGCTGCTGAC
ACCCATGAGAAGGACCGTGTCGTCGCTGCTGCTGAC
CTCAGGATTCACATGTGTTGACCGAAGATGCTAATAGTCAATCCATGCTGACGGAAGGACCGTGTCGTCGCTGCTGCTGAC
TCGAGTACCTGGCAGGCAGATGCCGATCGCTGCTGCTGCTGAC
GAAATTCGCACGAGAGTCTGGGCAAGAGGGTCTGATGCTGCTGCTGAC
CAGTGGCTCGCCAGTTGAGATCCTGCTCGTACGGCGTATCGCATCCTGCTGCTGCTGAC
TGCAGTGTTTTTGTGGCCAGATGCTGACTTCCACGACCAGGTACGCTGCTGCTGAC
GGAGATGGGACCCGAGGATGGAATTGGCGTAA

Figure A.9. Putative CDS for mRNA encoding *L. longipalpis* ATG1.

Full length predicted mRNA is depicted in black, while the primers used for RT-qPCR amplicon measurement are shown in pink. Sequences in blue denote agreement between prediction and sequenced amplicons, while nucleotides in red depict a difference at that base pair.
Figure A.10. Putative CDS for mRNA encoding *L. longipalpis* ATG8.

Full length predicted mRNA is depicted in black, while the primers used for RT-qPCR amplicon measurement are shown in pink. Sequences in blue denote agreement between prediction and sequenced amplicons, while nucleotides in red depict a difference at that base pair.
>Duox
ATGACTGTGATGAGAACAAGCTTGTCAAGGCTGATTTGCAATCGGCTCTTGGAATGAAATCGGACGATATGTTTGT
GCGGAAAATGTCAAATATCTGCGACAAAGGATGCTGAATTTGCATCGGCTCTGGGAATGAAATCGGACGATATGTTTGT
TAGGCTCAGCATGCAAGGACCAAGAAGCTCGGAGATTTGATATGTTGCTGCAAGACGGAAGACAAG
CTGCGGATTATCTTTGATATGTGCGACAACGATAGGAATGGTGTGATTGATAAGGGGGAACTGAGCGA
GATGATGCGTTCCCTTGTTGA
GATTGCGCGC
ACAACAAGCCTGGGGGATGATCAGGTA
GACGGAACTAATTGTGAGGATGTTCCTGCAAGCACAAGAATCTCCTACCTACCGTAA
GTTGGCCTAGCCATTACACGTGCTAGCAGCATGTTGCTCTCTGCTATTCTTCCTTCTTCTTTGCTTGACGGA
CAGTGGGACAGCTAACAGCAGATCGAGATGGTTGATTTGCTTGATTTTGCTTTGAGCTAATCTGAGT
GCTCTTCCACAGCCTAGGGGTAAGCCTTCCTTCTCTTTTCTCAAGTGATTGATATGTTGAAAGAAGG
ATGTAACGAATGTCCTCGAGATTCATATTTTCATCACACAATTCTTC
CTATCTCCTCTTGAGGATGTGCTGCGCGATGTGGAGAAGAAGG
Figure A.11. Putative CDS for mRNA encoding *L. longipalpis* Duox.
Full length predicted mRNA is depicted in black, while the primers used for RT-qPCR amplicon measurement are shown in pink. Sequences in blue denote agreement between prediction and sequenced amplicons, while nucleotides in red depict a difference at that base pair.
Figure A.12. Neighbor-joining tree to infer the evolutionary relationship of Duox sequences.
<table>
<thead>
<tr>
<th>Protein</th>
<th>Alignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>DmDuox</td>
<td>1 MVSAPSAPQRAESKNRVPQPGQKNRKFLKLRHLPQGAYGGALLLLISYGLELGTVHCY</td>
</tr>
<tr>
<td>LLDuox</td>
<td>1 -----------------------------------------------</td>
</tr>
<tr>
<td>PpDuox</td>
<td>1 -----------------------------------------------</td>
</tr>
<tr>
<td>AgAGAP009978</td>
<td>1 -----------------------------------------------</td>
</tr>
<tr>
<td>CqDuox</td>
<td>1 -----------------------------------------------</td>
</tr>
<tr>
<td>AeAEL007563Par</td>
<td>1 -----------------------------------------------</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Protein</th>
<th>Alignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>DmDuox</td>
<td>61 EKMYGQTKEQRYDGWNLAPDG LS:PTPAKAPSYSGVTVNCN:MTPTTLSG</td>
</tr>
<tr>
<td>LLDuox</td>
<td>1 -----------------------------------------------</td>
</tr>
<tr>
<td>PpDuox</td>
<td>1 -----------------------------------------------</td>
</tr>
<tr>
<td>AgAGAP009978</td>
<td>1 -----------------------------------------------</td>
</tr>
<tr>
<td>CqDuox</td>
<td>5 ERLESHEQRYDGWNLAPDG YCNLTKAPPYSGVTVLACGNRPPSTLSR</td>
</tr>
<tr>
<td>AeAEL007563Par</td>
<td>1 -----------------------------------------------</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Protein</th>
<th>Alignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>DmDuox</td>
<td>121 FMRKDGFL-KRNRTALLAFFGQVNCIVMASESGCPIEMHRIEEKCDMYDRCRG</td>
</tr>
<tr>
<td>LLDuox</td>
<td>1 -----------------------------------------------</td>
</tr>
<tr>
<td>PpDuox</td>
<td>45 FMRKDGFL-KRNRTALFFQVTVNEIVMASESGCPIEMHRIEEKCDMYDRCRG</td>
</tr>
<tr>
<td>AgAGAP009978</td>
<td>58 FMRKDGFL-PKRNRTALLAFFGQVTVNEIVMASESGCPIEMHRIEEKCDMYDRCRG</td>
</tr>
<tr>
<td>CqDuox</td>
<td>65 FMRKDGFL-PKRNRTALLAFFGQVTVNEIVMASESGCPIEMHRIEEKCDMYDRCRG</td>
</tr>
<tr>
<td>AeAEL007563Par</td>
<td>35 FMRKDGFL-PKRNRTALFFQVTVNEIVMASESGCPIEMHRIEEKCDMYDRCRG</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Protein</th>
<th>Alignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>DmDuox</td>
<td>181 [YIFHRAAYDR]TEQSPNAPREQINQMTANIDGSFIYSTSEAWLNAMRESSHC:LLIP</td>
</tr>
<tr>
<td>LLDuox</td>
<td>1 -----------------------------------------------</td>
</tr>
<tr>
<td>PpDuox</td>
<td>105 [YIFHRAAYDR]TEQSPNAPREQINQMTANIDGSFIYSTSEAWLNAMRESSHC:LLIP</td>
</tr>
<tr>
<td>AgAGAP009978</td>
<td>118 [YIFHRAAYDR]TEQSPNAPREQINQMTANIDGSFIYSTSEAWLNAMRESSHC:LLIP</td>
</tr>
<tr>
<td>CqDuox</td>
<td>125 [YIFHRAAYDR]TEQSPNAPREQINQMTANIDGSFIYSTSEAWLNAMRESSHC:LLIP</td>
</tr>
<tr>
<td>AeAEL007563Par</td>
<td>95 [YIFHRAAYDR]TEQSPNAPREQINQMTANIDGSFIYSTSEAWLNAMRESSHC:LLIP</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Protein</th>
<th>Alignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>DmDuox</td>
<td>241 D[SPY]NTNRVPLFNFPVR-YA MLSPERE LLGDPRTQNFAPFSAILFRMN</td>
</tr>
<tr>
<td>LLDuox</td>
<td>1 -----------------------------------------------</td>
</tr>
<tr>
<td>PpDuox</td>
<td>165 [S:MPY]NTNRVPLFNFPVR-YA MLSPERE LLGDPRTQNFAPFSAILFRMN</td>
</tr>
<tr>
<td>AgAGAP009978</td>
<td>178 [S:MPY]NTNRVPLFNFPVR-YA MLSPERE LLGDPRTQNFAPFSAILFRMN</td>
</tr>
<tr>
<td>CqDuox</td>
<td>185 [S:MPY]NTNRVPLFNFPVR-YA MLSPERE LLGDPRTQNFAPFSAILFRMN</td>
</tr>
<tr>
<td>AeAEL007563Par</td>
<td>155 [S:MPY]NTNRVPLFNFPVR-YA MLSPERE LLGDPRTQNFAPFSAILFRMN</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Protein</th>
<th>Alignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>LLDuox</td>
<td>1 -----------------------------------------------</td>
</tr>
<tr>
<td>Protein</td>
<td>Length</td>
</tr>
<tr>
<td>-------------</td>
<td>--------</td>
</tr>
<tr>
<td>DmDuoxB</td>
<td>900</td>
</tr>
<tr>
<td>LlDuox</td>
<td>66</td>
</tr>
<tr>
<td>PpDuox</td>
<td>825</td>
</tr>
<tr>
<td>AgAGAP009978</td>
<td>838</td>
</tr>
<tr>
<td>CqDuox</td>
<td>845</td>
</tr>
<tr>
<td>AeAAE007563Par</td>
<td>815</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Protein</th>
<th>Length</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>DmDuoxB</td>
<td>960</td>
<td>DFKLMMKEYGDFVAIGLDCKGAKQNFDTSTNVARMTSFHIEPMPKDRHWLAKWDY</td>
</tr>
<tr>
<td>LlDuox</td>
<td>126</td>
<td>DFKLMMKEYGDFVAIGLDCKGAKQNFDTSTNVARMTSFHIEPMPKDRHWLAKWDY</td>
</tr>
<tr>
<td>PpDuox</td>
<td>885</td>
<td>DFKLMMKEYGDFVAIGLDCKGAKQNFDTSTNVARMTSFHIEPMPKDRHWLAKWDY</td>
</tr>
<tr>
<td>AgAGAP009978</td>
<td>898</td>
<td>DFKLMMKEYGDFVAIGLDCKGAKQNFDTSTNVARMTSFHIEPMPKDRHWLAKWDY</td>
</tr>
<tr>
<td>CqDuox</td>
<td>905</td>
<td>DFKLMMKEYGDFVAIGLDCKGAKQNFDTSTNVARMTSFHIEPMPKDRHWLAKWDY</td>
</tr>
<tr>
<td>AeAAE007563Par</td>
<td>875</td>
<td>DFKLMMKEYGDFVAIGLDCKGAKQNFDTSTNVARMTSFHIEPMPKDRHWLAKWDY</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Protein</th>
<th>Length</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>DmDuoxB</td>
<td>1020</td>
<td>TTFLEENQRQNYLFLFYVITILFVERFIHYSFNAETHDLRHIMGVGIAITRGSAAALS</td>
</tr>
<tr>
<td>LlDuox</td>
<td>186</td>
<td>TTFLEENQRQNYLFLFYVITILFVERFIHYSFNAETHDLRHIMGVGIAITRGSAAALS</td>
</tr>
<tr>
<td>PpDuox</td>
<td>945</td>
<td>TTFLEENQRQNYLFLFYVITILFVERFIHYSFNAETHDLRHIMGVGIAITRGSAAALS</td>
</tr>
<tr>
<td>AgAGAP009978</td>
<td>958</td>
<td>TTFLEENQRQNYLFLFYVITILFVERFIHYSFNAETHDLRHIMGVGIAITRGSAAALS</td>
</tr>
<tr>
<td>CqDuox</td>
<td>965</td>
<td>TTFLEENQRQNYLFLFYVITILFVERFIHYSFNAETHDLRHIMGVGIAITRGSAAALS</td>
</tr>
<tr>
<td>AeAAE007563Par</td>
<td>935</td>
<td>TTFLEENQRQNYLFLFYVITILFVERFIHYSFNAETHDLRHIMGVGIAITRGSAAALS</td>
</tr>
<tr>
<td>Protein</td>
<td>Length</td>
<td>Sequence</td>
</tr>
<tr>
<td>--------------</td>
<td>--------</td>
<td>--------------------------------------------------------------------------</td>
</tr>
<tr>
<td>DmDuoxB</td>
<td>1440</td>
<td>QFFHKFDLRTTMLYICENHFQRLSKTSFTGLKAVNHFGRPDMSFLKTVKKHSVYVSFKI</td>
</tr>
<tr>
<td>LiDuox</td>
<td>604</td>
<td>QFFHKFDLRTTMLYICENHFQRLSKTSFTGLKAVNHFGRPDMSFLKTVKKHSVYVSFKI</td>
</tr>
<tr>
<td>PpDuox</td>
<td>1365</td>
<td>QFFHKFDLRTTMLYICENHFQRLSKTSFTGLKAVNHFGRPDMSFLKTVKKHSVYVSFKI</td>
</tr>
<tr>
<td>AgAGAP009978</td>
<td>1378</td>
<td>QFFHKFDLRTTMLYICENHFQRLSKTSFTGLKAVNHFGRPDMSFLKTVKKHSVYVSFKI</td>
</tr>
<tr>
<td>CqDuox</td>
<td>1385</td>
<td>QFFHKFDLRTTMLYICENHFQRLSKTSFTGLKAVNHFGRPDMSFLKTVKKHSVYVSFKI</td>
</tr>
<tr>
<td>AeAAEL007563Par</td>
<td>1355</td>
<td>QFFHKFDLRTTMLYICENHFQRLSKTSFTGLKAVNHFGRPDMSFLKTVKKHSVYVSFKI</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Protein</th>
<th>Length</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>DmDuoxB</td>
<td>1500</td>
<td>GVFSCGPRPLTKSVMSACDEVNKGRKLPYFIHHFENFC</td>
</tr>
<tr>
<td>LiDuox</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PpDuox</td>
<td>1425</td>
<td>GVFSCGPRPLTKSVMSACDEVNKGRKLPYFIHHFENFC</td>
</tr>
<tr>
<td>AgAGAP009978</td>
<td>1438</td>
<td>GVFSCGPRPLTKSVMSACDEVNKGRKLPYFIHHFENFC</td>
</tr>
<tr>
<td>CqDuox</td>
<td>1445</td>
<td>GVFSCGPRPLTKSVMSACDEVNKGRKLPYFIHHFENFC</td>
</tr>
<tr>
<td>AeAAEL007563Par</td>
<td>1415</td>
<td>GVFSCGPRPLTKSVMSACDEVNKGRKLPYFIHHFENFC</td>
</tr>
</tbody>
</table>

**Figure A.13.** Clustal Omega protein alignment of Duox sequences from *L. longipalpis*, *P. papatasi*, *D. melanogaster*, *A. gambiae*, *A. aegyti*, and *C. quinquefasciatus.*
USP36
ATGGTGTGCGACATCGTCAGTGCGGCACTCCGGCAGTGACTTCCTCATCCGCCAAGGCG
GCAGGACGGGCTGACCGTAATCTCAGACGACGCTCTCTGTTAGTACGACAAAGCGGAAT
TTTTCTCACAAAAATTTGAATACGAAGACGTGGACAGTACCCACAGCACTTTGCTGAGCG
CCTCAAATTTATATATTGTCTTTCAAACCTCCTCGGACATTATTTCGCTGTAACCACG
GCTGAGGATAAAATGCTCTGTGTGTCGACGACGCGCTCTCTGCTGACACAGAGATGG
CACAACCTGCACTGTAATGCGACAGCAGCAAGATGCTCAGTTGCCAGCACCACCAAC
GTGTCCTTTCCTCCCGCAGACAGCTCAGGAGTTGAGCTGGGCTGCGGGCGAGTGGACGAAAAT
GCCAGGTGCGCGGCTGATGTAACACGTGCTACCTCAAATTGGCAGCGGAGCGATGCG
CACCAGGAAATCATCGCAAGATAGTCGGCGGCTGCGGGGACGGGTGCTTGTTGATAAATTCGCG
AATGGGCACGCCTACTCATCTCGTGCACACAAAAAACAACAAATTCACCCCCGTAAGCCGTGCT
GATTCACTCABAACACTCCCGCTGGTGTTGTAAGCAGACTTACTTCCCCGGGCGCTCAGAGGA
TGCCGATGAGTCTCTGCGTGACCTGCGTGACCTCGGAGCAAGCATACCTCGCAC
GCCGTGAAGAATGCGTATGTTGGATATAGTGACAGCAGGAAGACGACGCGCTCTCAAT
CAAATTTCCGTGCAGCTACGTACCTACGGCTCCACTGCAGCACGCTCTCACTACGTC
TGGGAGAGCCTGCTCCGGCAGCTCGGACGCTCAGTTGGGAGACTTACGACCGGTGCTCAGG
GCCGTGCTGGAGCGCCCTCTCTGTTTACCGCTCTGGTGCAATGGTGACGCATTTGGGTGC
TTCGCAGAATTGTGGGGCACTACACGGCTATTGGGCTGACGGAAACGGGCAATTACT
ACACTTTAGCAGATCGCGTGACCAAATTAGTGAGCAAAATGCTTCGACAGACAA
ATGCTCCTACATTATCTTCTACGGAATCAGATGGTGCACCTGCAAAAAGCCCCCCATCAAC
CACCACCCCAGCAGATCGACGAGCAAGACCCCTACAGCCCTTCAGAGCTTGAGCAAT
GGGGAGATGTGAATGCACTGGCTGGTGTATTCAATACGTACGCTCTCCAGAAT
GAGTTACCTTCGACCTACCCAGGACGATGAGTCGTAATTCCTGACCCGCAAAATCGTTCTCAGG
GGCTGAGAGAAGAGCGCGGGAAGAGGTAACCTCAATGGGGGATGCAAATCG
AAGATTACCTTCGACCTACCCAGGAGCTGAGTGAGGCGCGAGCAGTCCCTTCAGGCACA
GAGCACCAGGGGCTAGCAGGCTCCCATTTCTCCAGTAAACTCGTCACAGTGCGGCAA

128
GAAGATGACCTTCAATGGTAGCTCAAATGGTGTTGACGCAGCCCCAAAGAGATTCA
CCCACTTATCCTCACGAATAAAGAGCAACAATTACGCAGGATAAGGGGCAAGGGGA
AAGTCCGTGGAGCAGGCAATGATTACCCAGCCGGAGAAAAAGATGAAAATCGGGCC
ATCGCTTACGCCCAACCCCAAAAGTCTCGTCCCTATGAAATCGGACGATGAGTCGAA
TCCATCGGAAAAATGATGAACCGTCATACTGAAGACACAAACAGGTGTTGTTGGCA
TTACACCAATGGATGATCCAAAATGTAGTCAAGAACACGCTACCCAACATCGTCGTC
AATGGGGGATGGATGATCCTATGGAATGTTCTCGTCCACGACCTGAG
AATGGGGGAGGAAATGAATTATACGTAAACTGGCCACAGTTAAATCATTTTGCTCAA
GACAGGTGGCATCGAGGTATGGGATGGCATGAAATGTCTATCGTGGAATGGCCAAA
AGACATCCCTCGACAGAGAGGTGCAAGAAGGACAAGAAGGAAGGGGAAAGCTGA
ATTGGGACATCAAATGGAGAATGGTGCATGGACAGGGGAGGAGAGCAGAAGGGA
AAGGTGCACCATCCACCAGGAAGGACAATCGGGCTATAATCTCCCTCCGAGATA
CCAGATGCGCAATTGGACCTGGCTCAAATAATGGTGTCGTCAGGGCTACTCAGGGCA
TCTACCAGCCAAAAAGTCTCACAGCAACTACTACAACAAACATACCCGGGCAAGGA
GGGGGTGTTGTTGGTGTTGCCCTCCGGGCGGAGCCATCGGGGTTGTCTTTTAATAGCA
ATTCGTTTCCACCATGGGACGTCGACAAATGA

**Figure A.14. Putative CDS for mRNA encoding *L. longipalpis* USP36.**

Full length predicted mRNA is depicted in black, while the primers used for RT-qPCR amplicon measurement are shown in pink. Sequences in blue denote agreement between prediction and sequenced amplicons, while nucleotides in red depict a difference at that base pair.
Figure A.15. Neighbor-joining tree to infer the evolutionary relationship of USP36 sequences.
Figure A.16. Clustal Omega protein alignment of USP36 sequences from *L. longipalpis*, *P. papatasi*, *D. melanogaster*, *A. gambiae*, *A. aegyti*, and *C. quinquefasciatus.*
Figure A.17. Putative CDS for mRNA encoding *L. longipalpis* IMD.

Full length predicted mRNA is depicted in black, while the primers used for RT-qPCR amplicon measurement are shown in pink. Sequences in blue denote agreement between prediction and sequenced amplicons, while nucleotides in red depict a difference at that base pair.
Figure A.18. Neighbor-joining tree to infer the evolutionary relationship of IMD sequences.
<table>
<thead>
<tr>
<th>Gene</th>
<th>Sequence</th>
</tr>
</thead>
</table>
| TcLOC660509  | GmIMD            47
|              | DmIMD            70
|              | LlIMD            64
|              | CqConserved_Hyp  60
|              | LiIMD            109
|              | DmIMD            80
|              | GmIMD            62 |
|              | GmIMD            131
|              | DmIMD            166
|              | LlIMD            192
|              | CqConserved_Hyp  172
|              | LiIMD            109
|              | DmIMD            80
|              | GmIMD            62 |
|              | PpIMD            479
|              | TcLOC660509      47
|              | AgAGAP004959     70
|              | AaAAEL010083     64
|              | CqConserved_Hyp  60
|              | LiIMD            109
|              | DmIMD            80
|              | GmIMD            62 |
|              | PpIMD            539
|              | TcLOC660509      47
|              | AgAGAP004959     70
|              | AaAAEL010083     64
|              | CqConserved_Hyp  60
|              | LiIMD            109
|              | DmIMD            80
|              | GmIMD            62 |
|              | PpIMD            599
|              | TcLOC660509      47
|              | AgAGAP004959     70
|              | AaAAEL010083     64
|              | CqConserved_Hyp  60
|              | LiIMD            109
|              | DmIMD            80
|              | GmIMD            62 |
|              | PpIMD            644
|              | TcLOC660509      47
|              | AgAGAP004959     70
|              | AaAAEL010083     64
|              | CqConserved_Hyp  60
|              | LiIMD            109
|              | DmIMD            80
|              | GmIMD            62 |
|              | PpIMD            704
|              | TcLOC660509      47
|              | AgAGAP004959     70
|              | AaAAEL010083     64
|              | CqConserved_Hyp  60
|              | LiIMD            109
|              | DmIMD            80
|              | GmIMD            62 |
|              | PpIMD            764
|              | TcLOC660509      47
|              | AgAGAP004959     70
|              | AaAAEL010083     64
|              | CqConserved_Hyp  60
|              | LiIMD            109
|              | DmIMD            80
|              | GmIMD            62 |
Figure A.19. Clustal Omega protein alignment of IMD sequences from *L. longipalpis*, *P. papatasii*, *D. melanogaster*, *A. gambiae*, *A. aegyti*, *C. quinquefasciatus*, *T. Castaneum*, and *G. morsitans*.
Figure A.20. Putative CDS for mRNA encoding *L. longipalpis* Nos.

Full length predicted mRNA is depicted in black, while the primers used for RT-qPCR amplicon measurement are shown in pink. Sequences in blue denote agreement between prediction and sequenced amplicons, while nucleotides in red depict a difference at that base pair.
Figure A.21. Neighbor-joining tree to infer the evolutionary relationship of Nos sequences.
Figure A.22. Clustal Omega protein alignment of Nos sequences from *L. longipalpis*, *P. papatasi*, *D. melanogaster*, *A. gambiae*, and *A. aegyti*. 
GGATGCATTGAATGACATTAATAAAGATTCAAAATGCCCCAAAATGGGTGAATGATCGATGGGGGTGTACTGCGGA
ATGTTCAAGATACATCGTGGGGAGATGCGTGTGACGAGCAGGATGAGCAGGAGAGGAGTCTCCTGAGGAATCATATGGAG
TCAAGTCTAGTAGTAGTAGTAGTAACGAGTAGTAGACGGCAAAATGTTGACAAACCAGAGTCAGTGTAATCTACAGATGGACCTCC
GGAGGAGATGATGGACTCCATGGAGGAGCATACGCATCATCAACCAGAGGAGGATGAGGATAAGCATTCAGTTGAAAT
CTCAACCCGGAGATCCGGCTCACAGGAGATCTCCTTTGATGTGCTGCAGCCAGAAAAATTGGACGAAATTATGATTCGT
CCACAAATTGAGCCACAGCAAACGCATCAAGTTGCTCTCTGGGAGTCAATCAATACGTCCACTTTGCTAAGCC
CTCACAGGGTTATACACAACGTGCACTCCCCTTAAGGTGCCCATTTAACGCCCCATGGGGGCTGAGACTGATGAAACGG
GAATTTCCAGGATACGTTACACGAAAAACTCTCTGATTGTGGGCGAACGTATGTA

**Figure A.23. Putative CDS for mRNA encoding *L. longipalpis* Domeless.**

Full length predicted mRNA is depicted in black, while the primers used for RT-qPCR amplicon measurement are shown in pink. Sequences in blue denote agreement between prediction and sequenced amplicons, while nucleotides in red depict a difference at that base pair.
Figure A.24. Neighbor-joining tree to infer the evolutionary relationship of Domeless sequences.
Figure A.25. Clustal Omega protein alignment of Domeless sequences from *L. longipalpis*, *P. papatasi*, *D. melanogaster*, *A. gambiae*, *A. aegyti*, and *C. quinquefasciatus*. 
Figure A.26. Putative CDS for mRNA encoding *L. longipalpis* IMPer.

Full length predicted mRNA is depicted in black, while the primers used for RT-qPCR amplicon measurement are shown in pink. Sequences in blue denote agreement between prediction and sequenced amplicons, while nucleotides in red depict a difference at that base pair.
Figure A.27. Neighbor-joining tree to infer the evolutionary relationship of IMPer sequences.
<table>
<thead>
<tr>
<th>Protein</th>
<th>Accession</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>PpHPX1_PPAA0096</td>
<td>1</td>
<td>FFPPCSSK --FRSTGECNNINHRD-- WGA</td>
</tr>
</tbody>
</table>
Figure A.28. Clustal Omega protein alignment of IMPer sequences from *L. longipalpis*, *P. papatas*, *D. melanogaster*, *A. gambiae*, *A. aegyti*, and *C. quinequefasciatus*. 