Toll Signaling Immune Function and Evolution in Anopheline Mosquitoes

by

Victoria L. M. Rhodes

B.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 Biology

College of Arts and Sciences

KANSAS STATE UNIVERSITY

Manhattan, Kansas

2018

Abstract

Malaria remains a major human vector-borne disease, greatly contributing to global human morbidity and mortality. Control of mosquito vectors that transmit malaria continues to be dependent on the widespread application of chemical insecticides through indoor residual spraying and insecticide treated bed nets. However, resistance to these insecticides is spreading within many mosquito populations, adding an ever-increasing urgency to the development of alternative vector control measures. The mosquito immune system is a potential novel target for such alternative measures, as the immune response initiated in these insects during infection with vector-borne disease agents is a key determinant of vector competence and, thus, contributes to a species' vectorial capacity. These immune responses, additionally, interact with and respond to parasitic or symbiotic biocontrol agents employed to kill or manipulate infection outcome with vector-borne disease agents. Entomopathogenic fungi, including *Beauveria bassiana*, have been considered as an alternative vector control measure, functioning as biopesticides. The Toll pathway is a key antifungal immune pathway in insects that impacts an insect's ability to survive fungal infections. A better understanding of Toll signaling immune function and evolution in anophelines, both vector and nonvector, can thus help to improve future biocontrol methods of important vector mosquitoes like *Anopheles gambiae*. In this dissertation, I report the use of *B. bassiana* strain I93-825 in *An. gambiae* to analyze the impact of Toll pathway modulation on mosquito survival. Mosquito survivorship was strongly affected by *B. bassiana* exposure dose by several measured parameters including median survival, longevity, and hazard. Modulation of Toll signaling, by way of knockdown by RNA interference, revealed a dose-dependent trade-off between immune activation state and survivorship in *An. gambiae*. To better determine the full Toll immune signaling repertoire in mosquitoes, I annotated and describe the evolutionary

history of intracellular Toll pathway members and Toll-like receptors (TLRs) within 21 mosquito genomes. The intracellular signaling pathway is conserved with 1:1 orthology, and evolutionary rates across different intracellular pathway members vary widely as compared to the conserved protein core of these mosquito species. In contrast, TLRs evolved largely by duplication events within certain anopheline lineages, most dramatically in the *An. gambiae* complex, where six TOLL1/5 paralogs likely possess different ligand binding specificities. Thus, these TLRs should be prioritized for experimental analyses of TLR immune function in *An. gambiae*.

Taken together, the work in this dissertation identifies Toll pathway modulation as a potential resistance mechanism that could impact malaria biocontrol strategies and provides a foundation for future detailed studies of Toll pathway function in important mosquito vector species.

Toll Signaling Immune Function and Evolution in Anopheline Mosquitoes

by

Victoria L. M. Rhodes

B.S., Kansas State University, 2012

A DISSERTATION

submitted in partial fulfillment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Biology

College of Arts and Sciences

KANSAS STATE UNIVERSITY

Manhattan, Kansas

2018

Approved by:

Major Professor

Dr. Kristin Michel

Copyright

Abstract

Malaria remains a major human vector-borne disease, greatly contributing to global human morbidity and mortality. Control of mosquito vectors that transmit malaria continues to be dependent on the widespread application of chemical insecticides through indoor residual spraying and insecticide treated bed nets. However, resistance to these insecticides is spreading within many mosquito populations, adding an ever-increasing urgency to the development of alternative vector control measures. The mosquito immune system is a potential novel target for such alternative measures, as the immune response initiated in these insects during infection with vector-borne disease agents is a key determinant of vector competence and, thus, contributes to a species' vectorial capacity. These immune responses, additionally, interact with and respond to parasitic or symbiotic biocontrol agents employed to kill or manipulate infection outcome with vector-borne disease agents. Entomopathogenic fungi, including *Beauveria bassiana*, have been considered as an alternative vector control measure, functioning as biopesticides. The Toll pathway is a key antifungal immune pathway in insects that impacts an insect's ability to survive fungal infections. A better understanding of Toll signaling immune function and evolution in anophelines, both vector and nonvector, can thus help to improve future biocontrol methods of important vector mosquitoes like *Anopheles gambiae*. In this dissertation, I report the use of *B. bassiana* strain I93-825 in *An. gambiae* to analyze the impact of Toll pathway modulation on mosquito survival. Mosquito survivorship was strongly affected by *B. bassiana* exposure dose by several measured parameters including median survival, longevity, and hazard. Modulation of Toll signaling, by way of knockdown by RNA interference, revealed a dose-dependent trade-off between immune activation state and survivorship in *An. gambiae*. To better determine the full Toll immune signaling repertoire in mosquitoes, I annotated and describe the evolutionary

history of intracellular Toll pathway members and Toll-like receptors (TLRs) within 21 mosquito genomes. The intracellular signaling pathway is conserved with 1:1 orthology, and evolutionary rates across different intracellular pathway members vary widely as compared to the conserved protein core of these mosquito species. In contrast, TLRs evolved largely by duplication events within certain anopheline lineages, most dramatically in the *An. gambiae* complex, where six TOLL1/5 paralogs likely possess different ligand binding specificities. Thus, these TLRs should be prioritized for experimental analyses of TLR immune function in *An. gambiae*.

Taken together, the work in this dissertation identifies Toll pathway modulation as a potential resistance mechanism that could impact malaria biocontrol strategies and provides a foundation for future detailed studies of Toll pathway function in important mosquito vector species.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

Amino acid ..... aa

Achaete ..... AC

Antimicrobial peptide ..... AMP

Anopheles Plasmodium-responsive Leucine-Rich Repeat 1 ..... APL1

Base pairs ..... bp

Bootstrap ..... BT

Cactus ..... CACT

Cysteine cluster on the C-terminal end of LRRs motif ..... CF motif

Damage-associated molecular pattern ..... DAMP

Deformed epidermal autoregulatory factor-1 ..... DEAF1

Dorsal-related immunity factor ..... Dif

An. gambiae 1000 Genomes Consortium AG1000G

An. gambiae injected with dsRNA specific to *Cactus* ..... ds*Cactus*

An. gambiae injected with dsRNA specific to *GFP* ..... ds*GFP*

An. gambiae injected with dsRNA specific to *REL1* ds*REL1*

FKBP-rapamycin-associated protein ..... FRAP

Gram-negative binding protein ..... GNBP

Gene ontology ..... GO

G protein-coupled receptor kinase 2 ..... GPRK2

Hazard ratio ..... HR

Hepatocyte growth factor-regulated tyrosine kinase substrate ..... HRS

Insulin/insulin-like growth factor signaling ..... IIS

Interleukin-1 receptor.....IL-1R

Immunodeficiency .....IMD

Indoor residual spraying ..... IRS

Insecticide treated bed net.....ITN

Janus kinase ..... JAK

Jun-N-terminal kinase ..... JNK

Jones-Taylor-Thornton ..... JTT

Kilobase .....kb

Knockdown.....kd

Non-LTR retrotransposon..... LINE

Leucine-rich repeat immune .....LRIM

Leucine-rich repeat .....LRR

Leucine rich repeat C-terminal domain ..... LRR-CT

Leucine rich repeat N-terminal domain ..... LRR-NT

Lethal time 50 .....LT50

Microbe-associated molecular pattern ..... MAMP

Mitogen-activated protein kinase.....MAPK

Megabase ..... Mb

Multiple cysteine cluster TLR ..... mccTLR

Miniature inverted-repeat TE.....MITE

Maximum-likelihood .....ML

Modular serine protease.....ModSP

Myopic .....MOP

Nuclear factor-κB ..... NF-κB

Nearest-Neighbor-Interchange ..... NNI

Odorant binding proteins ..... OBP

Odorant receptors ..... OR

Polymerase chain reaction ..... PCR

Pellino ..... PLI

Pannier ..... PNR

Peptidoglycan recognition protein ..... PRGP

PAX transcription activation domain interacting protein ..... PTIP

RNA interference ..... RNAi

Ribosomal protein S7 ..... rpS7

Reverse transcription-quantitative PCR ..... RT-qPCR

Single cysteine cluster TLR ..... sccTLR

Standard error ..... SE

Short intersperse nuclear element ..... SINE

Small interfering RNA ..... siRNA

Supernumerary limbs ..... SLMB

Single nucleotide polymorphism ..... SNP

Serine protease inhibitor ..... SRPN

Signal transducer and activator of transcription ..... STAT

Short tandem repeat ..... STR

Thioester-containing protein 1 ..... TEP1

Transposable elements ..... TEs

Transforming growth factor ..... TGT

Toll/interleukin-1 receptor ..... TIR

Toll-like receptor ..... TLR

Target of rapamycin ..... TOR

TNF-receptor-associated factor 6 ..... TRAF6

U-shaped ..... USH

Acknowledgements

This work could not have been accomplished without the help and expertise of several individuals. These people have aided me over the past five years in a multitude of ways, and as such I would like to thank them here for their various contributions to this body of work.

Thank you to Bart Bryant, Melissa Gulley, and Tinea Graves for showing me the ropes of mosquito rearing, data analysis, and variety of molecular techniques, such as Western blotting, qPCR, and microinjection. Thank you to the several undergraduate students that have assisted in various aspects of lab maintenance and mosquito rearing. I would like to thank Konner Winkley and Samantha Davidson by name, as both aided me greatly in time point collections and recording daily survivorships during their time in the lab. Both of you were a pleasure to mentor, and I hope you go far in life. Thank you to Drs. Mike Kanost and Chris Herren for always being willing to write letters of recommendation for me for various awards and positions. Thank you to my committee members Drs. Mike Herman, Sue Brown, and Mike Kanost for your feedback and questions over the years. I would like to give a special thank you to my advisor, Dr. Kristin Michel. I appreciate everything you have done for me. Your encouragement and guidance has not gone unnoticed.

I would like to thank my family for their support over the years. I know that my inability to make a variety of special occasions and my overall distance and inaccessibility was hard at times. I appreciate their patience, love, and support.

To Chris Rhodes, my husband, thank you for your support in almost every aspect of my academic journey thus far. You were a rock in the stressful times, where I may have been short and snippy, and served as my much-needed support system. Your willingness to help me run

through countless presentations, spend Sunday mornings sorting mosquitoes, and talk me through complex subjects like ‘emotions’ will forever be fond memories.

Additionally, I would like to thank the members of my graduate student cohort, particularly Bram Verheijen, Kirsten Grond, Elizabeth Everman, Ellen Welti, and Kim O’Keefe, with whom I enjoyed many late night talks ranging over many topics, such as the importance of the European honey bee, the vast array of bird species, and movies so terrible that they deserve watching four times. To MK Mills, I look forward to retiring on your cul-de-sac. If ever there was a relationship to make a life-long impression, my friendship with you these past five years has accomplished it.

And lastly, I would like to thank my funding sources, including NIH, Kansas State University Graduate School, and the Biology Graduate Student Association. Their financial support helped make this work possible.

P.S. Radina’s Coffeehouse, I love you.

Dedication

This dissertation is dedicated to all those who inspired me to pursue a career in science. You pushed me, mentored me, and encouraged me in many ways. Your instruction showed me my true passion of teaching others the wonders of biology.

Chapter 1 - Introduction

Victoria Rhodes[1] and Kristin Michel[1*]

[1]Division of Biology, Kansas State University, Manhattan, KS 66506, USA

[*]Corresponding author: Kristin Michel, kmichel@ksu.edu

Portions of this introduction published in

"Modulation of Mosquito Immune Defenses as a Control Strategy"

ARTHROPOD VECTOR: CONTROLLER OF

DISEASE TRANSMISSION

VOLUME 1: VECTOR MICROBIOME AND INNATE

IMMUNITY OF ARTHROPODS

Victoria L.M. Rhodes, Kristin Michel

Kansas State University, Manhattan, KS, United States

Anopheline Mosquitoes and the Malaria Life Cycle

Human malaria, caused by five species of apicomplexan protozoans in the genus *Plasmodium*, requires the parasite to successfully infect and survive within its mosquito vector and mammalian host. Successful completion of the *Plasmodium* life cycle depends on a number of non-genetic determinants, as well the interactions between the genotypes of the parasite, host, and vector. The ability of mosquitoes to acquire, maintain, and transmit human malaria parasite species, also referred to as vector competence, is rare in nature. The majority of mosquito species are unable to sustain the development of human malaria parasites and are refractory to infection.

Currently, there are over 3,500 species of mosquito described and, of these, only the genus *Anopheles* transmits the *Plasmodium* parasite that infect humans (Harbach, 2013). Of the more than 450 formally described species within the *Anopheles* genus, only about 70 can transmit human parasites and only a subset, 41, of these species are considered dominant vector species able to transmit malaria on a scale that threatens public health (Harbach, 2013; Hay et al., 2010; Sinka et al., 2012). Indeed, even in competent vector species, the majority of parasites are eliminated or killed at various stages in their development.

Within the mosquito species with published genome assemblies, the focus has been on those species that serve as successful vectors of *Plasmodium* (Figure 1-1). Comparative genomic analyses of these genomes, along with non-*Plasmodium* vectors from culicine species, revealed that protein families expected to contribute to the ability of *Anopheles* to successfully vector *Plasmodium* were gained shortly after the split between *Culicinae* and *Anophelinae* (Oppenheim et al., 2017). Distribution of nonvectors in phylogenetic analyses of anophelines reveals that vector competence has likely been lost repeatedly over evolution (Oppenheim et al., 2017). A mechanistic insight into the genetic basis of vector competence carries the promise that naturally

susceptible mosquito populations could be rendered refractory as an intervention strategy for malaria.

Susceptible mosquitoes allow *Plasmodium* spp. to undergo sexual reproduction, known as sporogony (Figure 1-2, reviewed in Baton and Ranford-Cartwright, 2005; Beier, 1998), and serve as the definitive host, because sexual reproduction of the parasite occurs strictly within the insect's midgut. The mosquito becomes infected when taking a blood meal from a human containing male and female gametocytes. These undergo gametogenesis within minutes of reaching the midgut lumen. Fertilization takes place within the next hour, and, in the following 2 hours, the zygote undergoes meiosis. Within the first day post blood feeding, the zygote transforms into the motile ookinete, which then traverses the peritrophic matrix surrounding the blood bolus and enters the midgut epithelium. Upon traversal, the ookinete emerges on the basal side of the midgut epithelium, rounds up, and forms the oocyst. In the subsequent 4–5 days, the oocyst undergoes mitosis, ultimately leading to the formation of several thousand sporozoites (Rosenberg and Rungsiwongse, 1991). Sporozoites are released into the hemocoel and distributed passively by hemolymph flow throughout the open circulatory system (Hillyer et al., 2007). Between 10 and 14 days after the initial gametocyte-containing blood meal, the mosquito becomes infective to the next human host, as sporozoites cross the salivary gland epithelium, and reach the mosquito saliva. Malaria parasite losses occur at multiple stages of the sporogonic life cycle (Figure 1-2). The major bottleneck occurs during the first 2–5 days of infection, during the transition of gametocyte to oocyst, leading to elimination more than 99% of parasites (Gouagna et al., 2004). In addition, less than 25% of sporozoites released into the hemocoel eventually reach the salivary gland lumen (Rosenberg and Rungsiwongse, 1991).

Historical Perspective of the Mosquito Immune System

The experiments performed in the late 1800s by Ross, Bignami, Grassi, Dionisi, and Bastianelli not only demonstrated that malaria parasites indeed are transmitted by mosquitoes, but that vector competence for malaria varies widely among mosquito species (for an excellent early summary of this work see Lyon, 1900). During his time in Secunderabad, India, Ronald Ross discovered and visualized the pigmented cells (oocysts) on the midguts of "dapple-winged" mosquitoes, now known as *Anopheles* spp., after the ingestion of gametocytes from a malaria-infected patient (Ross and Smyth, 1897). However, experiments conducted during the previous two years with "gray" mosquitoes, most likely mosquitoes in the genus *Aedes*, fed on malaria-infected humans never produced oocysts. Similarly, experiments performed by Bignami at the same time in Italy, also demonstrated that mosquitoes outside the genus *Anopheles* were not susceptible to human malaria infection (Bastianelli and Bignami, 1900; Ross, 1899). Variation in vector competence is not only seen across but also within species. Huff observed early that *Culex pipiens* shows variation in natural susceptibility to avian malaria parasites (Huff, 1929, 1927). *C. pipiens*, allowed to take repeated blood meals from birds infected with *Plasmodium cathemerium*, *Plasmodium relictum*, or *Plasmodium elongatum* displayed similar infection rates between first and second feedings with the same *Plasmodium* species. However, the susceptibility of *C. pipiens* was parasite species specific, suggesting that vector competence for a given *Plasmodium* spp. is fixed and thus hereditary (Huff, 1930). These early results have been confirmed subsequently across several mosquito species, which range widely in their ability to support development of distinct *Plasmodium* spp., depending on the vector and parasite species combination (Alavi et al., 2003; Huff, 1929, 1927; Vaughan et al., 1994). Similar observations were made in field-derived populations of *Anopheles gambiae* and *Plasmodium falciparum* in

different locations of human malaria transmission in sub-Saharan Africa (Lambrechts et al., 2005; Niaré et al., 2002).

The genetic basis of vector competence was elegantly demonstrated in selection experiments performed by Huff. He produced *C. pipiens* lines by selective mating that were either susceptible or resistant to the avian malaria parasite, *P. cathemerium*. Crosses of these lines yielded F1 progeny that exhibited Mendelian ratios of susceptibility (Huff, 1935, 1931). The heritability of vector competence for bird malaria was confirmed subsequently in a number of *Plasmodium* spp. and mosquito species combinations, demonstrating increased susceptibility after several generations of selective breeding (Micks, 1949; Trager, 1942). Similarly, selection experiments were used to generate two independent refractory lines of *An. gambiae*, a species usually highly susceptible to human malaria parasite infection (Collins et al., 1986; Vernick et al., 1995). For one of these lines, L35, the genetic basis of refractoriness was mapped to one major and two minor quantitative trait loci (Gorman et al., 1997; Zheng et al., 1997).

As early as 1927, Huff had proposed that the immune system of the mosquito kills malaria parasites and thus is an important determinant of vector competence (Huff, 1927). In fact, the earliest description of a mosquito humoral immune response against malaria parasites had been made by Ross, when he documented degenerated and melanized oocysts as “black spores” (Daniels, 1898; Knowles and Basu, 1933; Sinden and Garnham, 1973). Ultimately, the selection experiments performed by Collins and Vernick confirmed Huff’s early proposal. The L35 line kills rodent malaria parasites and allopatric human *P. falciparum* ookinetes through an immune reaction called melanization (Collins et al., 1986). In the SUAF2 line, ookinetes of the bird malaria parasite *Plasmodium gallinaceum* are killed by an immune effector mechanism called lysis (Vernick et al., 1995). These selection experiments hold the promise that targeted

manipulation of the mosquito immune system can render epidemiologically important vector species refractory and prevent malaria transmission. Over the last three decades, this field has made significant advances in our understanding of the mosquito immune system, including the processes that kill different parasite stages within the mosquito vector.

Synopsis of Mosquito Immunity

Much of the basis of our understanding of the molecular makeup and function of immunity in mosquitoes derives from studies in the model organism *Drosophila melanogaster*. Immunity within insects is composed of innate immune responses hardwired to activate in response to molecular patterns common to pathogens such as viruses, bacteria, fungi, protozoa, and nematodes. The response an insect mounts to an infection is largely dependent on which pathogen-associated molecular patterns (PAMPs) are present, allowing for a more targeted response to the pathogen of interest. These responses can be subdivided in three broad categories: cellular, intracellular, and humoral immune responses (Figure 1-3). A brief overview of these immune branches will be summarized in the below sections.

Cellular Immunity

Hemocytes, insect blood cells, constantly circulate and patrol insect hemolymph. Hemocytes can be divided into three populations based on morphological analyses, lectin binding assays, and enzyme activity assays as 1) granulocytes, 2) oenocytoids, and 3) prohemocytes (reviewed in Hillyer and Strand, 2014). These hemocyte types differ in their primary functions. Granulocytes are highly phagocytic (Hillyer et al., 2003), oenocytoids produce the phenoloxidase used to melanize pathogens (Castillo et al., 2006; Hillyer et al., 2003; Hillyer and Christensen, 2002),

and prohemocytes are hypothesized to serve as progenitor cells (Castillo et al., 2006). However, prohemocytes might also be the byproduct of asymmetric granulocyte divisions and do possess phagocytic abilities (King and Hillyer, 2013). Secondly, hemocytes can be divided based on their locations, with hemocytes that freely circulate in the hemolymph termed 'circulating hemocytes' and those that attach to tissues termed 'sessile hemocytes' (King and Hillyer, 2013).

Additionally, these hemocyte populations also have the ability to work together as a team, performing large-scale maneuvers such as encapsulation and nodulation, which are effective immune responses to large pathogens such as filarial worms and parasitoid wasp eggs (reviewed in Satyavathi et al., 2014). These maneuvers involve the irreversible binding of many hemocytes to the surface of these pathogens, forming a large, multicellular capsule that blocks the pathogen from needed nutrients.

Intracellular Immunity

Insects also deploy various tools that combat infections with intracellular pathogens, those that hide within a host's cells away from cellular responses such as melanization (see next section below). One of these effector mechanisms is apoptosis. Apoptosis controls intracellular pathogens as it is a targeted 'suicide' response by infected cells. Apoptosis is characterized by cell membrane blebbing, chromatin condensation, and DNA fragmentation. This mechanism is well-described in mosquito literature as a key mechanism by which midgut cells of *Anopheles* mosquitoes fight *Plasmodium* parasite penetration (Boëte et al., 2004; Vlachou et al., 2004). It is a race at the microscopic level; *Plasmodium* against the apoptosis clock. *Plasmodium* that successfully escape midgut apoptotic responses are marked for destruction by oxidative stress

and nitration of their surface proteins, which aids in lysing these parasites through complement action once they reach the basal lamina (Kumar et al., 2004; Oliveira et al., 2012).

Viral infections are targeted directly through RNA interference (RNAi). This pathway plays important roles in controlling viral infections, including dengue virus, within mosquitoes (Franz et al., 2006; Sánchez-Vargas et al., 2009). This pathway recognizes viral dsRNA, converting this dsRNA into a single-stranded small interfering RNA (siRNA) capable of recognizing complementary viral RNA sequences (Mongelli and Saleh, 2016). Binding of siRNA to complementary viral mRNA transcripts results in dicing of the target, effectively reducing translation of viral transcripts (Kim et al., 2009).

Humoral Immunity

Along with cellular and intracellular immunity, there is a suite of humoral immune responses deployed in the midgut lumen and hemolymph of mosquitoes. Pathogens in these areas must survive melanization and/or lysis mechanisms that are fast and effective. Like all insects, mosquitoes utilize melanin deposition to seal wounds and physically surround pathogens in a thick, dense melanin coat (Gorman et al., 2007; Lai et al., 2002). Melanization occurs in the hemolymph through a series of reactions that convert tyrosine to eumelanin (reviewed in Nappi and Christensen, 2005). Melanization serves to contain or kill pathogens by sequestering them from available nutrients or by directly killing them through the accumulation of toxic byproducts created during the melanization process (Cerenius et al., 2008; Christensen et al., 2005; Nappi and Christensen, 2005). This response is more rarely observed in malaria-susceptible mosquitoes (Riehle et al., 2006; Schwartz and Koella, 2002). Melanization has received considerable

attention, as it is a selectable phenotype for refractoriness in the laboratory (Collins et al., 1986; Hurd et al., 2005).

Insects also have mechanisms that show similarities to vertebrate C3-mediated lysis. Termed the complement-like pathway, its activation results in the deposition of thioester-containing protein 1 (TEP1) that opsonizes the surface of pathogens. TEP1 is structurally similar to vertebrate complement factor C3 (Baxter et al., 2007; Levashina et al., 2001) and, just as C3, acts as an opsonin. TEP1 circulates in the hemolymph, stabilized by proteins APL1C and LRIM1 (Fraiture et al., 2009; Povelones et al., 2009). However, cleaved TEP1 readily binds to pathogen surfaces through covalent binding through a highly reactive thioester binding motif. Opsonization of pathogen surfaces by TEP1 leads to the lysis of the pathogen and this can be readily seen in *Anopheles* infected with *Plasmodium*, where *Plasmodium* ookinetes with TEP1 deposition show lytic symptoms like membrane blebbing and fragmentation (Blandin et al., 2004; Vernick et al., 1995). However, the intermediate steps between TEP1 deposition and lysis are unclear as of yet.

One of the hallmarks of insect innate immunity is the regulation of gene transcription through canonical signal transduction. These signal transduction pathways are activated by the recognition of pathogen-associated molecular patterns (PAMPs) and result in the transcriptional regulation of hundreds of genes (Garver et al., 2009). These include the nuclear factor (NF)-κB-dependent Toll and IMD pathways and the Janus kinase (JAK)-signal transducer and activator of transcription (STAT) pathway. In addition, mitogen-activated protein kinase (MAPK) signaling also regulates mosquito innate immunity through three pathways, including signaling through ERK, Jun- N-terminal kinase (JNK), and p38 MAP kinase. Comparative genomic analyses have shown that key players within these pathways are highly conserved within insects (Christophides

et al., 2002; Neafsey et al., 2015; Waterhouse et al., 2007) and orthologs even extend into higher vertebrates such as humans.

The Toll Pathway Controls Immune Reactions Targeting Broad Classes of Pathogens

The Toll pathway was first described in *D. melanogaster* as critical to establish the dorsoventral axis in the early embryo (Anderson and Nüsslein-Volhard, 1984), and is named after its transmembrane receptor, Toll (Hashimoto et al., 1988). The role of the Toll pathway in immunity was initially hypothesized based on the sequence similarity between the cytoplasmic domains of the *D. melanogaster* Toll and human interleukin-1 (IL-1) receptor (Gay and Keith, 1991), and later experimentally confirmed to control antimicrobial peptide expression in *D. melanogaster* (Lemaitre et al., 1996). This pathway is a major immune signaling pathway in the insect fat body. In mosquitoes, the Toll pathway mediates antibacterial and antifungal immune responses (Dong et al., 2012; Shin et al., 2006, 2005), regulates malaria parasite killing (Frolet et al., 2006) and has been linked to hemocyte activation (Ramirez et al., 2014).

The Toll pathway consists of an extracellular protease cascade and an intracellular signal transduction pathway, culminating in nuclear translocation of the NF-κB transcription factors, Dorsal and the Dorsal-related immunity factor (Dif) (Figure 1-4; reviewed in Valanne et al., 2011). The Toll pathway is activated through recognition of conserved pathogen-associated molecular patterns by extracellular pattern recognition receptors such as Gram-negative binding proteins and peptidoglycan recognition proteins (PGRPs) (Bischoff et al., 2004; Gobert et al., 2003; Gottar et al., 2006; Michel et al., 2001). Alternatively, Toll activation occurs by sensing "danger" signals that include fungal or bacterial virulence factors, such as digestive proteases, as

well as damage-associated molecular patterns that arise from stressed or damaged cells during infection (El Chamy et al., 2008; Gottar et al., 2006; Ming et al., 2014). Detection of either signal culminates in the proteolytic cleavage of the Toll ligand Spätzle, which, upon binding to the Toll receptor, triggers an intracellular signaling cascade, characterized by the death domain proteins Myd88, Tube, and Pelle that phosphorylate the inhibitor of (I)κB, Cactus. Cactus, in its unphosphorylated state, is bound to Dif/Dorsal, preventing it from entering the nucleus. Upon phosphorylation, Cactus releases Dif/Dorsal, resulting in the translocation of these NF-κB transcription factors to the nucleus to activate gene transcription.

The Toll pathway is tightly controlled by inhibitors at several different intervention points. The extracellular protease cascade is inhibited by several serine protease inhibitors of the serpin family (reviewed in Meekins et al., 2017). In *D. melanogaster*, the serpin Necrotic (Spn43Ac, Nec), prevents Toll signaling in the absence of fungal infection (Levashina et al., 1999). Furthermore, serpin-1 (Spn1) potentially regulates Spätzle activation, possibly by inhibition of the modular serine protease ModSP (Fullaondo et al., 2011). In addition to Cactus, the intracellular signal transduction pathway downstream of Toll is downregulated by ubiquitination. The ubiquitin E3 ligase Pellino associates with MyD88, thereby targeting MyD88 for degradation and, in turn, inhibiting Toll signaling (Ji et al., 2014). Finally, Toll pathway activity is not only controlled through increased Cactus phosphorylation, but also through decreased Cactus transcription controlled by the Hippo signal transduction pathway (Liu et al., 2016).

Components of the intracellular signal transduction pathway are conserved by orthology across insects, and orthologs of each protein have been identified in all sequenced mosquito genomes (Arensburger et al., 2010; Christophides et al., 2002; Neafsey et al., 2015; Waterhouse

et al., 2007). In contrast, no members of the extracellular protease cascade that regulate Spätzle cleavage have been identified, largely because these enzymes are members of protein families with numerous clade-specific expansions and losses that defy orthology assignments within and across orders of insects (An et al., 2011; Christophides et al., 2002; Kanost et al., 2016). The impact of the Toll pathway on *Plasmodium* spp. development has been analyzed in a number of mosquito species. Knockdown of *cactus* increases the number of *Plasmodium* spp. oocysts in multiple parasite–vector species combinations (Cirimotich et al., 2010; Mitri et al., 2009; Riehle et al., 2008; Zou et al., 2011). Several *Plasmodium* spp. killing mechanisms are controlled by the Toll pathway, including antimicrobial peptides (AMPs), transcriptional upregulation of components of the complement-like pathway, as well as a negative regulator of the melanization cascade (Frolet et al., 2006). Overexpression of *REL1*, the ortholog of *Drosophila* Dif/Dorsal, or knockdown of *cactus* affects the expression of 264 and 1850 genes (Zou et al., 2011), respectively, and thus likely elicits pleiotropic effects beyond canonical immune effector mechanisms that affect parasite survival.

Toll-Like Receptors (TLRs)

A fascinating aspect of Toll signaling is the diversity in the receptor family that is responsible for transducing the extracellular signaling cascade to the intracellular signal transduction cascade. The Toll pathway is named after the *D. melanogaster* receptor (toll) first characterized by its function in signaling during dorsoventral development of the early embryo. Sequence analysis of the receptor found similarities to vertebrate interleukin-1 receptor (IL-1R), providing the first evidence that what appeared to be an insect developmental pathway may have similarities to a vertebrate immune pathway (Gay and Keith, 1991; Hashimoto et al., 1988). Shortly after, this

insect dorsal-ventral polarizing developmental pathway was identified as a key pathway in the antifungal immune response of *Drosophila* (Lemaitre et al., 1996).

What was once just a single receptor, *Drosophila* Toll now belongs to a large protein family with members present in animals from sponges to higher chordates (reviewed in Leulier and Lemaitre, 2008). In *D. melanogaster*, there are nine Toll receptors; the original toll (Toll-1) and eight additional receptors, termed Toll-like receptors (TLRs) named Toll-2 through Toll-9. All members of this large family are characterized by an extracellular ligand binding region composed of repeated leucine-rich repeat (LRR) domains, a transmembrane domain, and an intracellular Toll/interleukin-1 receptor (TIR) domain (Figure 1-5). Phylogenetic analyses utilizing several protostome and deuterostome amino acid sequences reveal that the TLR family forms two distinct clades, and that TLRs predate the split between protostomes and deuterostomes (Luo and Zheng, 2000; Palmer and Jiggins, 2015; Roach et al., 2005).

In vertebrates, the biological function(s) of each TLR is well described and each plays a distinct role in immunity, as they directly recognize microbe-associated molecular patterns (MAMPs) (Roach et al., 2005). These membrane receptors have been implicated in diverse functions in *D. melanogaster*, including establishing the dorsoventral axis of the developing embryo (Anderson and Nüsslein-Volhard, 1984), transcription of antimicrobial peptides (Lemaitre et al., 1995), and neuronal axon guidance (Ward et al., 2015). TLR signal transduction in early embryo development appears to be a unique evolutionary adaptation of this receptor family primarily found within insects, as TLRs within birds, fish, and mammals are restricted to immunological roles (reviewed in Brownlie and Allan, 2011; Casanova et al., 2011; Rebl et al., 2010). Of the 9 encoded TLRs in *D. melanogaster*, Toll-1 (Lemaitre et al., 1996), Toll-5 (Luo et al., 2001), Toll-7 (Nakamoto et al., 2012), and Toll-8 (Akhouayri et al., 2011) have evidence for

a role in immune signaling. Furthermore, Toll-1 (Anderson et al., 1985), Toll-7 (Mcilroy et al., 2013), and Toll-8 (Paré et al., 2014) have also been implicated in some aspect of *D. melanogaster* development.. The roles these TLRs play in developmental and/or immunity pathways in other insects are currently largely unknown.

Within mosquitoes, very few TLRs have been attributed specific functional roles. Both *Ae. aegypti* and *An. gambiae* TLRs display unique gene expression patterns over the course of development and during infection (MacCallum et al., 2011). Studies utilizing tools such as RNAi knockdown, SNP analysis, and overexpression of TLRs in cell culture have provided a glimpse into the functions of a select few TLRs within these vector species (Harris et al., 2010; Luna et al., 2006, 2002; Shin et al., 2006). However, the function of individual TLRs remains largely undescribed in these vector species and the frequent expansion events observed in insects (Cao et al., 2015; Leulier and Lemaitre, 2008; Levin and Malik, 2017; Palmer and Jiggins, 2015) make it difficult to apply TLR functions found in one species to others through sequence identity alone. Outside of vertebrates such as *Mus musculus* and insects such as *D. melanogaster*, the question of whether specific TLRs serve immunological and/or developmental roles in animals is still a large gap in knowledge (reviewed in Leulier and Lemaitre, 2008). Therefore, it remains important to study and analyze this important signaling pathway in species of interest, such as mosquito vectors, to facilitate the understanding of the biology of these vectors and the potential for development of novel insect control measures.

Entomopathogenic Fungi

Insecticides have been the classic method for vector control, especially among regions of endemic malaria, for decades (reviewed in Alout et al., 2017). However, the heavy and regular

usage of these insecticides has led to the development of insecticide-resistant vector populations (reviewed in Alout et al., 2017). This development has created an urgency to develop novel vector and disease control strategies.

In recent years the development of a vaccine capable of providing protection from *Plasmodium* has been heavily researched, with the only approved vaccine, RTS,S/AS01, completing Phase 3 evaluation in 2015 (World Health Organization, 2016). However, the World Health Organization does not recommend this vaccine for widespread use, as there is a critical issue in the extent of efficacy in children 5-17 months of age. Within this age group, efficacy steadily decreased over relatively short time scales, despite vaccine boosters being administered every 6 months. While research efforts continue to develop long-lasting, effective vaccines, the time required for their development, coupled with the steadily increasing levels of insecticide resistance in vector populations (Ranson and Lissenden, 2016), has accentuated a need to develop novel vector control methods that can be developed quickly and distributed cheaply.

One such method would be to implement biocontrol agents such as entomopathogenic fungi. *B. bassiana* and *Metarhizium anisopliae*, both ascomycetes of the family *Clavicipitaceae*, have gained attention as such potential agents (Butt et al., 2016; Raghavendra et al., 2011; Scholte et al., 2004; Thomas, 2018). For the purposes of this introduction, I will focus my description on *B. bassiana*, as it has received considerable attention as a putative bioinsecticide specifically to combat mosquito-transmitted diseases (Blanford et al., 2012; Heinig and Thomas, 2015; Howard et al., 2010; Kikankie et al., 2010; Scholte et al., 2003; Vogels et al., 2014). *B. bassiana* is a cosmopolitan fungal species that can utilize resources in a range of environments. This fungus can survive as a saprophyte in soil, as an endophyte in plants, and as an entomopathogen in an astonishingly wide range of insect hosts (Boomsma et al., 2014; Inglis et

al., 2001). *B. bassiana* is haploid and reproduces asexually, making it the anamorphic state, while the teleomorphic, sexually reproducing state is better known under the species name *Cordyceps bassiana* (Sung et al., 2006).

The Entomopathogenic Life Cycle

B. bassiana belongs to a large and diverse group of entomopathogenic species, with most large fungal groups containing at least one example of this parasitic life style (Humber, 2008). However, it does not appear that this life style arose through a single evolutionary event, but likely arose independently multiple times (Humber, 2008; James et al., 2006). This adaptation has also led to species that are obligate pathogens, such as lepidopteran-killing *Nomuraea rileyi* and *Isaria tenuipes*, or facultative pathogens, such as *B. bassiana*. (Vega-Aquino et al., 2010). The jump from plant symbiont to insect parasite is hypothesized to have been partly due to the co-opting of genes involved in plant colonization or horizontal gene transfer to pathogenicity (Barelli et al., 2016; Screen and St. Leger, 2000): Acquisition of genes involved in attaching to insect cuticle, cuticle/tissue degradation/penetration, and immune evasion open up a novel source of nutrition to fungal species capable of accessing it.

The entomopathogenic life cycle of *B. bassiana* begins by first actively penetrating through the cuticle (Figure 1-6). This means that this fungus does not need to be ingested to successfully enter a new host, making this species particularly useful in cases where the targeted insect has sucking mouthparts (such as adult mosquitoes or aphids). Passing through the cuticle begins first with adherence of fungal conidia, or blastospores, to insect cuticle, facilitated by hydrophobins present on the surface of these propagules (Zhang et al., 2011). The formation of a specialized penetrating organ, termed an appressoria, along with the secretion of cuticle-

degrading enzymes, weakens and pierces host cuticle (St. Leger et al., 1998; Valero-Jiménez et al., 2016). The appressoria uses turgor pressure to facilitate cuticle penetration and fungal hyphae grow through the insect cuticle and breach into the hemolymph of the host (Wang and St. Leger, 2007). Upon reaching the nutrient-rich hemolymph, hyphae differentiate into blastospores (also known as hyphal bodies), which are single-celled and, at later growth stages, form long hyphae throughout the insect (Güerri-Agulló et al., 2010; Lewis et al., 2009; Wanchoo et al., 2009; Yassine et al., 2012). Fungal growth is facilitated by the nutrition present in host hemolymph, particularly simple disaccharide sugars such as trehalose (Wang et al., 2014). Once hemolymph nutrients are depleted and the host insect is dead, fungal hyphae penetrate the cuticle, and emerge to develop conidiophores. These conidiophores can produce vast amounts of conidia, which give the cadaver and white, powdery appearance (Kumar et al., 1999). This phenotype originally gave way to the term 'calcino disease', which is derived from the Latin '*calc*', for chalk. Later, thanks to the work of Agostino Bassi in the 1860's in identifying the fungal nature of this disease (at the time decimating the silkworm industry), this became known as 'white muscardine disease' and the fungal agent was eventually named in his honor (Porter, 1973).

B. bassiana as an Insect Control Agent

The use of entomopathogenic fungi to control insect pests in terms of crop maintenance is well established and numerous spore formulations have been developed for the mass market (Castrillo et al., 2011; Faria and Wraight, 2007; reviewed in Jaber and Ownley, 2017). The first commercial product containing *B. bassiana* for the control of insect pests was Boverin, developed in the USSR to control populations of Colorado potato beetles and codling moths (Faria and Wraight, 2007). Since then, the development of hundreds of products based on

entomopathogenic fungi have been developed worldwide (Faria and Wraight, 2007). *B. bassiana* is an attractive candidate for control of dipteran disease vectors, as this species effectively shortens the lifespan and feeding propensity of species such as *Anopheles gambiae* (Blanford et al., 2011; Mnyone et al., 2012). Even more promising, insecticide-resistant mosquitoes exposed to *B. bassiana* displayed similarly shortened life span, and co-exposure to both *B. bassiana* and the insecticide permethrin had a synergistic effect on survival (Howard et al., 2011)(Farenhorst et al., 2010; Howard et al., 2011, 2010). Furthermore, exposed male mosquitoes are able to successfully transmit infection to females during mating (García-Munguía et al., 2011).

Entomopathogenic biopesticides are an attractive method to control insect pests. Their expanded use into the realm of vector control promises to be quite helpful in the control of disease transmission. However, as with all vector control methods, close attention should be paid to potential for selection of vector populations resistant to infection and, as such, studies analyzing vector responses to infection will be a welcome addition to the growing body of work utilizing these pathogenic fungi in insect systems.

Dissertation Overview

As summarized in Chapter 1, while chemical insecticides still prove an important method by which vector populations are controlled, there is a push to establish alternative methods of vector control for these insects. *B. bassiana* and related entomopathogenic fungi are attractive alternatives to currently employed chemical methods. However, it remains to be seen how fungal dose and basal immune activation state in mosquitoes affects infection dynamics of these bioinsecticides. There is a need for studies monitoring *B. bassiana* efficacy in mosquitoes with altered immunity at various doses, as these data are required for effective application of this

bioinsectide, as well as accurate planning and modeling of potential resistance development through increased basal immune activation.

Furthermore, the overall immune repertoire of the antifungal Toll pathway and its signal-transducing TLRs, likely to respond to *B. bassiana* within targeted vectors, is not fully described for the selection of sequenced vector and nonvector mosquito genomes currently available. In mosquitoes, the biological functions of individual TLRs are largely unknown and the recent sequencing of 16 mosquito genomes provides a unique opportunity to study the full toll pathway repertoire in vector and nonvector mosquitoes. Understanding the makeup and evolution of this key antifungal pathway within mosquitoes may aid future efforts to a) manipulate mosquito immunity as a vector control method, b) improve planning and implementation of fungal agents for vector control by highlighting potential sites for selection and c) understand the complexity of immune signal transduction within species important to human health.

To fill these gaps in knowledge, this dissertation aimed to analyze the effects to *An. gambiae* survival at key intersections of *B. bassiana* dose and immune activation state through RNAi manipulation of Toll pathway activation (Chapter 2). Additionally, a thorough manual annotation and phylogenetic analysis of intracellular Toll pathway members and TLRS within 21 mosquito genomes was performed in order to better understand what constitutes the full repertoire of Toll signaling within mosquitoes, both vector and nonvector (Chapter 3). These data would fill an important gap in the mosquito vector field and provide much-needed information for the planning and long-term utilization of entomopathogenic fungi as vector-control methods. In addition, this dissertation provides data for future studies in insect immunology by providing a list of intriguing targets for future functional studies of TLRs in mosquitoes.

References

Alavi, Y., Arai, M., Mendoza, J., Tufet-Bayona, M., Sinha, R., Fowler, K., Billker, O., Franke-Fayard, B., Janse, C. J., Waters, A., Sinden, R. E., 2003. The dynamics of interactions between *Plasmodium* and the mosquito: a study of the infectivity of *Plasmodium berghei* and *Plasmodium gallinaceum*, and their transmission by *Anopheles stephensi*, *Anopheles gambiae*, and *Aedes aegypti*. Int. J. Parasitol. 33, 933–943.

Akhouayri, I., Turc, C., Royet, J., Charroux, B., 2011. Toll-8/tollo negatively regulates antimicrobial response in the *Drosophila* respiratory epithelium. PLoS Pathog. 7, e1002319.

Alout, H., Labbé, P., Chandre, F., Cohuet, A., 2017. Malaria vector control still matters despite insecticide resistance. Trends Parasitol.

An, C., Budd, A., Kanost, M. R., Michel, K., 2011. Characterization of a regulatory unit that controls melanization and affects longevity of mosquitoes. Cell. Mol. Life Sci. 68, 1929–1939.

Anderson, K. V., Bokla, L., Nusslein-Volhard, C., 1985. Establishment of dorsal-ventral polarity in the *Drosophila* embryo: The induction of polarity by the *Toll* gene product. Cell 42, 791–798.

Anderson, K. V., Nüsslein-Volhard, C., 1984. Information for the dorsal-ventral pattern of the *Drosophila* embryo is stored as maternal mRNA. Nature 311, 223–227.

Arensburger, P., Megy, K., Waterhouse, R. M., Abrudan, J., Amedeo, P., Antelo, B., Bartholomay, L., Bidwell, S., Caler, E., Camara, F., Campbell, C. L., Campbell, K. S., Casola, C., Castro, M. T., Chandramouliswaran, I., Chapman, S. B., Christley, S., Costas, J., Eisenstadt, E., Feschotte, C., Fraser-Liggett, C., Guigo, R., Haas, B., Hammond, M., Hansson, B. S., Hemingway, J., Hill, S. R., Howarth, C., Ignell, R., Kennedy, R. C., Kodira, C. D., Lobo, N. F., Mao, C., Mayhew, G., Michel, K., Mori, A., Liu, N., Naveira, H., Nene, V., NamNguyen, Pearson, M. D., Pritham, E. J., Puiu, D., Qi, Y., Ranson, H., Ribeiro, J. M. C., Roberston, H. M., Severson, D., Shumway, M., Stanke, M., Strausberg, R. L., Sun, C., Sutton, G., Tu, Z. (Jake), Tubio, J. M. C.,

Unger, M. F., Vanlandingham, D. L., Vilella, A. J., White, O., White, J. R., Wondji, C. S., Wortman, J., Zdobnov, E. M., Birren, B., Christensen, B. M., Collins, F. H., Cornel, A., Dimopoulos, G., Hannick, L. I., Higgs, S., Lanzaro, G. C., Lawson, D., Lee, N. H., Muskavitch, M. A. T., Raikhel, A. S., Atkinson, P. W., 2010. Sequencing of *Culex quinquefasciatus* establishes a platform for mosquito comparative genomics. Science 330, 86–88.

Barelli, L., Moonjely, S., Behie, S. W., Bidochka, M. J., 2016. Fungi with multifunctional lifestyles: endophytic insect pathogenic fungi. Plant Mol. Biol. 90, 657–664.

Bartholomay, Lyric C., and Kristin Michel., 2018. Mosquito Immunobiology: The Intersection of Vector Health and Vector Competence. Annu. Rev. Entomol. 63.

Bastianelli, G., Bignami, A., 1900. Malaria and Mosquitoes. Lancet 155, 79–80.

Baton, L. A., Ranford-Cartwright, L. C., 2005. Spreading the seeds of million-murdering death: metamorphoses of malaria in the mosquito. Trends Parasitol. 21, 573–580.

Baxter, R. H. G., Chang, C. I., Chelliah, Y., Blandin, S., Levashina, E. A., Deisenhofer, J., 2007. Structural basis for conserved complement factor-like function in the antimalarial protein TEP1. Proc. Natl. Acad. Sci. 104, 11615–11620.

Beier, J. C., 1998. Malaria parasite development in mosquitoes. Annu. Rev. Entomol. 43, 519–543.

Bennink, S., Kiesow, M. J., Pradel, G., 2016. The development of malaria parasites in the mosquito midgut. Cell. Microbiol. 18, 905–918.

Bischoff, V., Vignal, C., Boneca, I. G., Michel, T., Hoffmann, J. A., Royet, J., 2004. Function of the *Drosophila* pattern-recognition receptor PGRP-SD in the detection of Gram-positive bacteria. Nat. Immunol. 5, 1175–1180.

Blandin, S., Shiao, S. H., Moita, L. F., Janse, C. J., Waters, A. P., Kafatos, F. C., Levashina, E. A., 2004. Complement-like protein TEP1 is a determinant of vectorial capacity in the malaria vector *Anopheles gambiae*. Cell 116, 661–670.

Blanford, S., Chan, B. H. K., Jenkins, N., Sim, D., Turner, R. J., Read, A. F., Thomas, M.

B., 2005. Fungal pathogen reduces potential for malaria transmission. Science 308, 1638–1641.

Blanford, S., Shi, W., Christian, R., Marden, J. H., Koekemoer, L. L., Brooke, B. D., Coetzee, M., Read, A. F., Thomas, M. B., 2011. Lethal and pre-lethal effects of a fungal biopesticide contribute to substantial and rapid control of malaria vectors. PLoS One 6, e23591.

Blanford, S., Jenkins, N. E., Read, A. F., Thomas, M. B., 2012. Evaluating the lethal and pre-lethal effects of a range of fungi against adult *Anopheles stephensi* mosquitoes. Malar. J. 11, 365.

Boëte, C., Paul, R. E. L., Koella, J. C., 2004. Direct and indirect immunosuppression by a malaria parasite in its mosquito vector. Proc. R. Soc. L. 271, 1611–1615.

Boomsma, J. J., Jensen, A. B., Meyling, N. V, Eilenberg, J., 2014. Evolutionary interaction networks of insect pathogenic fungi. Annu. Rev. Entomol. 59, 467–485.

Brownlie, R., Allan, B., 2011. Avian toll-like receptors. Cell Tissue Res. 343, 121–130.

Butt, T. M., Coates, C. J., Dubovskiy, I. M., Ratcliffe, N. A., 2016. Entomopathogenic fungi: New insights into host-pathogen interactions, Advances in Genetics. 94, 307-364.

Cao, X., He, Y., Hu, Y., Wang, Y., Chen, Y. R., Bryant, B., Clem, R. J., Schwartz, L. M., Blissard, G., Jiang, H., 2015. The immune signaling pathways of *Manduca sexta*. Insect Biochem. Mol. Biol. 62, 64–74.

Casanova, J.-L., Abel, L., Quintana-Murci, L., 2011. Human TLRs and IL-1Rs in host defense: Natural insights from evolutionary, epidemiological, and clinical genetics. Annu. Rev. Immunol 29, 447–91.

Castillo, J. C., Robertson, A. E., Strand, M. R., 2006. Characterization of hemocytes from the mosquitoes *Anopheles gambiae* and *Aedes aegypti*. Insect Biochem. Mol. Biol. 36, 891–903.

Castrillo, L. A., Griggs, M. H., Ranger, C. M., Reding, M. E., Vandenberg, J. D., 2011. Virulence of commercial strains of *Beauveria bassiana* and *Metarhizium brunneum*

(Ascomycota: Hypocreales) against adult *Xylosandrus germanus* (Coleoptera: Curculionidae) and impact on brood. Biol. Control 58, 121–126.

Cerenius, L., Lee, B. L., Söderhäll, K., 2008. The proPO-system: pros and cons for its role in invertebrate immunity. Trends Immunol. 29, 263–271.

Christensen, B. M., Li, J., Chen, C. C., Nappi, A. J., 2005. Melanization immune responses in mosquito vectors. Trends Parasitol. 21, 192–199.

Christophides, G. K., Zdobnov, E., Barillas-Mury, C., Birney, E., Blandin, S., Blass, C., Brey, P. T., Collins, F. H., Danielli, A., Dimopoulos, G., Hetru, C., Hoa, N. T., Hoffmann, J. A., Kanzok, S. M., Letunic, I., Levashina, E. A., Loukeris, T. G., Lycett, G., Meister, S., Michel, K., Moita, L. F., Müller, H. M., Osta, M. A., Paskewitz, S. M., Reichhart, J. M., Rzhetsky, A., Troxler, L., Vernick, K. D., Vlachou, D., Volz, J., von Mering, C., Xu, J., Zheng, L., Bork, P., Kafatos, F. C., 2002. Immunity-related genes and gene families in *Anopheles gambiae*. Science 298, 159–165.

Cirimotich, C. M., Dong, Y., Garver, L. S., Sim, S., Dimopoulos, G., 2010. Mosquito immune defenses against *Plasmodium* infection. Dev. Comp. Immunol. 34, 387–395.

Collins, F. H., Sakai, R. K., Vernick, K. D., Paskewitz, S., Seeley, D. C., Miller, L. H., Collins, W. E., Campbell, C. C., Gwadz, R. W., 1986. Genetic selection of a *Plasmodium*-refractory strain of the malaria vector *Anopheles gambiae*. Science 234, 607–610.

Daniels, C. W., 1898. On transmission of Proteosoma to birds by the mosquito. Proc. R. Soc. London 64, 443–454.

Dong, Y., Morton, J. C. J., Ramirez, J. L., Souza-Neto, J. A., Dimopoulos, G., 2012. The entomopathogenic fungus *Beauveria bassiana* activate Toll and JAK-STAT pathway-controlled effector genes and anti-dengue activity in *Aedes aegypti*. Insect Biochem. Mol. Biol. 42, 126–132.

El Chamy, L., Leclerc, V., Caldelari, I., Reichhart, J. M., 2008. Sensing of “danger signals” and pathogen-associated molecular patterns defines binary signaling pathways

"upstream" of Toll. Nat. Immunol. 9, 1165–1170.

Farenhorst, M., Knols, B. G. J., Thomas, M. B., Howard, A. F. V, Takken, W., Rowland, M., N'Guessan, R., 2010. Synergy in efficacy of fungal entomopathogens and permethrin against west african insecticide-resistant *Anopheles gambiae* mosquitoes. PLoS One 5, e12081.

Faria, M. R. de, Wraight, S. P., 2007. Mycoinsecticides and Mycoacaricides: A comprehensive list with worldwide coverage and international classification of formulation types. Biol. Control 43, 237–256.

Fraiture, M., Baxter, R. H. G., Steinert, S., Chelliah, Y., Frolet, C., Quispe-Tintaya, W., Hoffmann, J. A., Blandin, S. A., Levashina, E. A., 2009. Two mosquito LRR proteins function as complement control factors in the TEP1-mediated killing of *Plasmodium*. Cell Host Microbe 5, 273–284.

Franz, A. W. E., Sanchez-Vargas, I., Adelman, Z. N., Blair, C. D., Beaty, B. J., James, A. A., Olson, K. E., 2006. Engineering RNA interference-based resistance to dengue virus type 2 in genetically modified *Aedes aegypti*. Proc. Natl. Acad. Sci. U. S. A. 103, 4198–4203.

Frolet, C., Thoma, M., Blandin, S., Hoffmann, J. A., Levashina, E. A., 2006. Boosting NF-κB-dependent basal immunity of *Anopheles gambiae* aborts development of *Plasmodium berghei*. Immunity 25, 677–685.

Fullaondo, A., García-Sánchez, S., Sanz-Parra, A., Recio, E., Lee, S. Y., Gubb, D., 2011. Spn1 regulates the GNBP3-dependent Toll signaling pathway in *Drosophila melanogaster*. Mol. Cell. Biol. 31, 2960–2972.

García-Munguía, A. M., Garza-Hernández, J. A., Rebollar-Tellez, E. A., Rodríguez-Pérez, M. A., Reyes-Villanueva, F., 2011. Transmission of *Beauveria bassiana* from male to female *Aedes aegypti* mosquitoes. Parasit. Vectors 4, 24.

Garver, L. S., Dong, Y., Dimopoulos, G., 2009. Caspar controls resistance to *Plasmodium falciparum* in diverse anopheline species. PLoS Pathog. 5, e1000335.

Gay, N. J., Keith, F. J., 1991. *Drosophila* Toll and IL-1 receptor. Nature 351, 355–356.

Gobert, V., Gottar, M., Matskevich, A. A., Rutschmann, S., Royet, J., Belvin, M. P., Hoffmann, J. A., Ferrandon, D., 2003. Dual activation of the *Drosophila* Toll pathway by two pattern recognition receptors. Science 302, 2126–2130.

Gorman, M. J., An, C., Kanost, M. R., 2007. Characterization of tyrosine hydroxylase from *Manduca sexta*. Insect Biochem. Mol. Biol. 37, 1327–1337.

Gorman, M. J., Severson, D. W., Cornel, A. J., Collins, F. H., Paskewitz, S. M., 1997. Mapping a quantitative trait locus involved in melanotic encapsulation of foreign bodies in the malaria vector, *Anopheles gambiae*. Genetics 146, 965–971.

Gottar, M., Gobert, V., Matskevich, A. A., Reichhart, J. M., Wang, C., Butt, T. M., Belvin, M., Hoffmann, J. A., Ferrandon, D., 2006. Dual detection of fungal infections in *Drosophila* via recognition of glucans and sensing of virulence factors. Cell 127, 1425–1437.

Gouagna, L. C., Bonnet, S., Gounoue, R., Verhave, J. P., Eling, W., Sauerwein, R., Boudin, C., 2004. Stage-specific effects of host plasma factors on the early sporogony of autologous *Plasmodium falciparum* isolates within *Anopheles gambiae*. Trop. Med. Int. Heal. 9, 937–948.

Güerri-Agulló, B., Gómez-Vidal, S., Asensio, L., Barranco, P., Lopez-Llorca, L. V., 2010. Infection of the Red Palm Weevil (*Rhynchophorus ferrugineus*) by the entomopathogenic fungus *Beauveria bassiana*: A SEM study. Microsc. Res. Tech. 73, 714–725.

Harbach, R. E., 2013. Mosquito Taxonomic Inventory [WWW Document]. URL http://mosquito-taxonomic-inventory.info/

Harris, C., Lambrechts, L., Rousset, F., Abate, L., Nsango, S. E., Fontenille, D., Morlais, I., Cohuet, A., 2010. Polymorphisms in *Anopheles gambiae* immune genes associated with natural resistance to *Plasmodium falciparum*. PLoS Pathog. 6, e1001112.

Hashimoto, C., Hudson, K. L., Anderson, K. V., 1988. The *Toll* gene of *Drosophila*, required for dorsal-ventral embryonic polarity, appears to encode a transmembrane protein. Cell 52, 269–279.

Hay, S. I., Sinka, M. E., Okara, R. M., Kabaria, C. W., Mbithi, P. M., Tago, C. C., Benz, D.,

Gething, P. W., Howes, R. E., Patil, A. P., Temperley, W. H., Bangs, M. J., Chareonviriyaphap, T., Elyazar, I. R. F., Harbach, R. E., Hemingway, J., Manguin, S., Mbogo, C. M., Rubio-Palis, Y., Godfray, H. C. J., 2010. Developing global maps of the dominant *Anopheles* vectors of human malaria. PLoS Med. 7, e1000209.

Heinig, R. L., Thomas, M. B., 2015. Interactions between a fungal entomopathogen and malaria parasites within a mosquito vector. Malar. J. 14, 22.

Hillyer, J. F., Barreau, C., Vernick, K. D., 2007. Efficiency of salivary gland invasion by malaria sporozoites is controlled by rapid sporozoite destruction in the mosquito haemocoel. Int. J. Parasitol. 37, 673–681.

Hillyer, J. F., Christensen, B. M., 2002. Characterization of hemocytes from the yellow fever mosquito, *Aedes aegypti*. Histochem Cell Biol 117, 431–440.

Hillyer, J. F., Schmidt, S. L., Christensen, B. M., 2003. Hemocyte-mediated phagocytosis and melanization in the mosquito *Armigeres subalbatus* following immune challenge by bacteria. Cell Tissue Res. 313, 117–127.

Hillyer, J. F., Strand, M. R., 2014. Mosquito hemocyte-mediated immune responses. Curr. Opin. Insect Sci. 3, 14–21.

Howard, A. F., N'Guessan, R., Koenraadt, C. J., Asidi, A., Farenhorst, M., Akogbeto, M., Knols, B. G., Takken, W., 2011. First report of the infection of insecticide-resistant malaria vector mosquitoes with an entomopathogenic fungus under field conditions. Malar. J. 10, 24.

Howard, A. F. V, N'Guessan, R., Koenraadt, C. J. M., Asidi, A., Farenhorst, M., Akogbéto, M., Thomas, M. B., Knols, B. G., Takken, W., 2010. The entomopathogenic fungus *Beauveria bassiana* reduces instantaneous blood feeding in wild multi-insecticide-resistant *Culex quinquefasciatus* mosquitoes in Benin, West Africa. Parasit. Vectors 3, 87.

Huff, C. G., 1935. Natural immunity and susceptibility of culicine mosquitoes to avian malaria. Am. J. Trop. Med. 15, 427–434.

Huff, C. G., 1931. The inheritance of natural immunity to *Plasmodium cathemerium* in two

species of *Culex*. J. Prev. Med. 5, 249–259.

Huff, C. G., 1930. Individual immunity and susceptibility to *Culex pipiens* to various species of bird malaria as studied by means of double infectious feedings. Am. J. Epidemiol. 12, 424–441.

Huff, C. G., 1929. The effects of selection upon susceptibility to bird malaria in *Culex pipiens* Linn. Ann. Trop. Med. Parasitol. 23, 427–442.

Huff, C. G., 1927. Studies on the infectivity of plasmodia of birds for mosquitoes, with special reference to the problem of immunity in the mosquito. Am. J. Epidemiol. 7, 706–734.

Humber, R. A., 2008. Evolution of entomopathogenicity in fungi. J. Invertebr. Pathol. 98, 262–266.

Hurd, H., Taylor, P. J., Adams, D., Underhill, A., Eggleston, P., 2005. Evaluating the costs of mosquito resistance to malaria parasites. Evolution (N. Y). 59, 2560–2572.

Inglis, G. D., Goettel, M. S., Butt, T. M., Strasser, H., 2001. Use of hyphomycetous fungi for managing insect pests, in: Fungi as Biocontrol Agents. pp. 23–69.

Jaber, L. R., Ownley, B. H., 2017. Can we use entomopathogenic fungi as endophytes for dual biological control of insect pests and plant pathogens? Biol. Control.

James, T. Y., Kauff, F., Schoch, C. L., Matheny, P. B., Hofstetter, V., Cox, C. J., Celio, G., Gueidan, C., Fraker, E., Miadlikowska, J., Lumbsch, H. T., Rauhut, A., Reeb, V., Arnold, A. E., Amtoft, A., Stajich, J. E., Hosaka, K., Sung, G. H., Johnson, D., O'Rourke, B., Crockett, M., Binder, M., Curtis, J. M., Slot, J. C., Wang, Z., Wilson, A. W., Schüßler, A., Longcore, J. E., O'Donnell, K., Mozley-Standridge, S., Porter, D., Letcher, P. M., Powell, M. J., Taylor, J. W., White, M. M., Griffith, G. W., Davies, D. R., Humber, R. A., Morton, J. B., Sugiyama, J., Rossman, A. Y., Rogers, J. D., Pfister, D. H., Hewitt, D., Hansen, K., Hambleton, S., Shoemaker, R. A., Kohlmeyer, J., Volkmann-Kohlmeyer, B., Spotts, R. A., Serdani, M., Crous, P. W., Hughes, K. W., Matsuura, K., Langer, E., Langer, G., Untereiner, W. A., Lücking, R., Büdel, B., Geiser, D. M., Aptroot, A., Diederich, P., Schmitt, I., Schultz, M., Yahr, R., Hibbett, D. S., Lutzoni, F., McLaughlin, D. J., Spatafora, J. W., Vilgalys,

R., 2006. Reconstructing the early evolution of Fungi using a six-gene phylogeny. Nature 443, 818–822.

Ji, S., Sun, M., Zheng, X., Li, L., Sun, L., Chen, D., Sun, Q., 2014. Cell-surface localization of Pellino antagonizes Toll-mediated innate immune signalling by controlling MyD88 turnover in *Drosophila*. Nat. Commun. 5.

Kanost, M. R., Arrese, E. L., Cao, X., Chen, Y., Chellapilla, S., Goldsmith, M.R., Grosse-Wilde, E., Heckel, D. G., Herndon, N., Jiang, H., Papanicolaou, A., Qu, J., Soulages, J. L., Vogel, V., Walters, J., Waterhouse, R. M., Ahn, S., Almeida, F. C., An, C., Aqrawi, P., Bretschneider, A., Bryant, W. B., Bucks, S., Chao, H., Chevignon, G., Christen, J. M., Clarke, D. F., Dittmer, N. T., Ferguson, L. C. F., Garavelou, S., Gordon, K. H. J., Gunaratna, R. T., Han, Y., Hauser, F., He, Y., Heidel-Fischer, H., Hirsh, A., Hu, Y., Jiang, H., Kalra, D., Klinner, C., König, C., Kovar, C., Kroll, A. R., Kuwar, S. S., Lee, S. L., Lehman, R., Li, K., Li, Z., Liang, H., Lovelace, S., Lu, Z., Mansfield, J. H., McCulloch, K. J., Mathew, T., Morton, B., Muzny, D. M., Neunemann, D., Ongeri, F., Pauchet, Y., Pu, L., Pyrousis, I., Rao, X., Redding, A., Roesel, C., Sanchez-Gracia, A., Schaack, S., Shukla, A., Tetreau, G., Wang, Y., Xiong, G., Traut, W., Walsh, T. K., Worley, K. C., Wu, D., Wu, W., Wu, Y., Zhang, X., Zou, Z., Zucker, H., Briscoe, A. D., Burmester, T., Clem, R. J., Feyereisen, R., Grimmelikhuijzen, C. J. P., Hamodrakas, S. J., Hansson, B. S., Huguet, E., Jermiin, L. S., Lan, Q., Lehman, H. K., Lorenzen, M., Merzendorfer, H., Michalopoulos, I., Morton, D. B., Muthukrishnan, S., Oakeshott, J. G., Palmer, W., Park, Y., Passarelli, A. L., Rozas, J., Schwartz, L. M., Smith, W., Southgate, A., Vilcinskas, A., Vogt, R., Wang, P., Werren, J., Yu, X., Zhou, J., Brown, S. J., Scherer, S. E., Richards, S., and Blissard, G. W., 2016. Multifaceted biological insights from a draft genome sequence of the tobacco hornworm moth, *Manduca sexta*. Insect biochemistry and molecular biology, 76, 118-147.

Kikankie, C. K., Brooke, B. D., Knols, B. G. J., Koekemoer, L. L., Farenhorst, M., Hunt, R. H., Thomas, M. B., Coetzee, M., 2010. The infectivity of the entomopathogenic fungus *Beauveria bassiana* to insecticide-resistant and susceptible *Anopheles arabiensis* mosquitoes at two different temperatures. Malar. J. 9, 71.

Kim, V. N., Han, J., Siomi, M. C., 2009. Biogenesis of small RNAs in animals. Nat. Rev. Mol. Cell Biol. 10, 126–139.

King, J. G., Hillyer, J. F., 2013. Spatial and temporal in vivo analysis of circulating and sessile immune cells in mosquitoes: hemocyte mitosis following infection. BMC Biol. 11, 55.

Knowles, R., Basu, B. C., 1933. The nature of the so-called "black spores" of Ross in malaria-transmitting mosquitoes. Indian J. Med. Res. 20, 757–776.

Kumar, S., Gupta, L., Han, Y. S., Barillas-Mury, C., 2004. Inducible peroxidases mediate nitration of *Anopheles* midgut cells undergoing apoptosis in response to *Plasmodium* invasion. J. Biol. Chem. 279, 53475–53482.

Kumar, V., Singh, G. P., Babu, A. M., Ahsan, M. M., Datta, R. K., 1999. Germination, penetration, and invasion of *Beauveria bassiana* on silkworm, *Bombyx mori*, causing white muscardine. Ital. J. Zool. 66, 39–43.

Lai, S. C., Chen, C. C., Hou, R. F., 2002. Immunolocalization of Prophenoloxidase in the Process of Wound Healing in the Mosquito *Armigeres subalbatus* (Diptera: Culicidae). J. Med. Entomol. 39, 266–274.

Lambrechts, L., Halbert, J., Durand, P., Gouagna, L. C., Koella, J. C., 2005. Host genotype by parasite genotype interactions underlying the resistance of anopheline mosquitoes to *Plasmodium falciparum*. Malar. J. 4, 3.

Lemaitre, B., Meister, M., Govind, S., Georgel, P., Steward, R., Reichhart, J. M., Hoffmann, J. A., 1995. Functional analysis and regulation of nuclear import of dorsal during the immune response in *Drosophila*. EMBO J. 14, 536–545.

Lemaitre, B., Nicolas, E., Michaut, L., Reichhart, J. M., Hoffmann, J. A., 1996. The dorsoventral regulatory gene cassette *spätzle/Toll/cactus* controls the potent antifungal response in *Drosophila* adults. Cell 86, 973–983.

Leulier, F., Lemaitre, B., 2008. Toll-like receptors - taking an evolutionary approach. Nat. Rev. Genet. 9, 165–178.

Levashina, E. A., Langley, E., Green, C., Gubb, D., Ashburner, M., Hoffmann, J. A.,

Reichhart, J. M., 1999. Constitutive activation of Toll-mediated antifungal defense in serpin-deficient *Drosophila*. Science 285, 1917–1919.

Levashina, E. A., Moita, L. F., Blandin, S., Vriend, G., Lagueux, M., Kafatos, F. C., 2001. Conserved role of a complement-like protein in phagocytosis revealed by dsRNA knockout in cultured cells of the mosquito, *Anopheles gambiae*. Cell 104, 709–718.

Levin, T. C., Malik, H. S., 2017. Rapidly evolving *Toll-3/4* genes encode male-specific Toll-like receptors in *Drosophila*. Oxford Univ. Press.

Lewis, M. W., Robalino, I. V., Keyhani, N. O., 2009. Uptake of the fluorescent probe FM4-64 by hyphae and haemolymph-derived *in vivo* hyphal bodies of the entomopathogenic fungus *Beauveria bassiana*. Microbiology 155, 3110–3120.

Liu, B., Zheng, Y., Yin, F., Yu, J., Silverman, N., Pan, D., 2016. Toll receptor-mediated Hippo signaling controls innate immunity in *Drosophila*. Cell 164, 406–419.

Luna, C., Hoa, N. T., Lin, H., Zhang, L., Nguyen, H. L. A., Kanzok, S. M., Zheng, L., 2006. Expression of immune responsive genes in cell lines from two different anopheline species. Insect Mol. Biol. 15, 721–729.

Luna, C., Wang, X., Huang, Y., Zhang, J., Zheng, L., 2002. Characterization of four Toll related genes during development and immune responses in *Anopheles gambiae*. Insect Biochem. Mol. Biol. 32, 1171–1179.

Luo, C., Shen, B., Manley, J. L., Zheng, L., 2001. Tehao functions in the Toll pathway in *Drosophila melanogaster*: Possible roles in development and innate immunity. Insect Mol. Biol. 10, 457–464.

Luo, C., Zheng, L., 2000. Independent evolution of *Toll* and related genes in insects and mammals. Immunogenetics 51, 92–98.

Lyon, I. P., 1900. The inoculation of malaria by the mosquito: A review of the literature. Med. Rec. 57.

MacCallum, R. M., Redmond, S. N., Christophides, G. K., 2011. An expression map for *Anopheles gambiae*. BMC Genomics 12, 620.

Mcilroy, G., Foldi, I., Aurikko, J., Wentzell, J. S., Lim, M. A., Fenton, C., Gay, N. J., Hidalgo, A., 2013. Toll-6 and Toll-7 function as neurotrophin receptors in the *Drosophila* central nervous system. Nat. Neurosci. 16, 1248–1256.

Meekins, D. A., Kanost, M. R., Michel, K., 2017. Serpins in arthropod biology. Semin. Cell Dev. Biol. 62, 105–119.

Michel, T., Reichhart, J. M., Hoffmann, J. A., Royet, J., 2001. *Drosophila* Toll is activated by Gram-positive bacteria through a circulating peptidoglycan recognition protein. Nature 414, 756–759.

Micks, D. W., 1949. Investigations on the mosquito transmission of *Plasmodium elongatum*. J. Natl. Malar. Soc. 8, 206–218.

Ming, M., Obata, F., Kuranaga, E., Miura, M., 2014. Persephone/Spätzle pathogen sensors mediate the activation of Toll receptor signaling in response to endogenous danger signals in apoptosis-deficient *Drosophila*. J. Biol. Chem. 289, 7558–7568.

Mitri, C., Jacques, J. C., Thiery, I., Riehle, M. M., Xu, J., Bischoff, E., Morlais, I., Nsango, S. E., Vernick, K. D., Bourgouin, C., 2009. Fine pathogen discrimination within the APL1 gene family protects *Anopheles gambiae* against human and rodent malaria species. PLoS Pathog. 5, e1000576.

Mnyone, L. L., Lyimo, I. N., Lwetoijera, D. W., Mpingwa, M. W., Nchimbi, N., Hancock, P. A., Russell, T. L., Kirby, M. J., Takken, W., Koenraadt, C. J., 2012. Exploiting the behaviour of wild malaria vectors to achieve high infection with fungal biocontrol agents. Malar. J. 11, 87.

Mongelli, V., Saleh, M. C., 2016. Bugs Are Not to Be Silenced: Small RNA Pathways and Antiviral Responses in Insects. Annu. Rev. Virol. 3, 573–589.

Mueller, A. K., Kohlhepp, F., Hammerschmidt, C., Michel, K., 2010. Invasion of mosquito salivary glands by malaria parasites: Prerequisites and defense strategies. Int. J. Parasitol. 40, 1229–1235.

Nakamoto, M., Moy, R. H., Xu, J., Bambina, S., Yasunaga, A., Spencer, S., Gold, B., Cherry, S., 2012. Virus recognition by Toll-7 activates antiviral autophagy in

Drosophila. Immunity 36, 658–667.

Nappi, A. J., Christensen, B. M., 2005. Melanogenesis and associated cytotoxic reactions: Applications to insect innate immunity. Insect Biochem. Mol. Biol. 35, 443–459.

Neafsey, D. E., Waterhouse, R. M., Abai, M. R., Aganezov, S. S., Alekseyev, M. A., Allen, J. E., Amon, J., Arcà, B., Arensburger, P., Artemov, G., Assour, L. A., Basseri, H., Berlin, A., Birren, B. W., Blandin, S. A., Brockman, A. I., Burkot, T. R., Burt, A., Chan, C. S., Chauve, C., Chiu, J. C., Christensen, M., Costantini, C., Davidson, V. L. M., Deligianni, E., Dottorini, T., Dritsou, V., Gabriel, S. B., Guelbeogo, W. M., Hall, A. B., Han, M. V, Hlaing, T., Hughes, D. S. T., Jenkins, A. M., Jiang, X., Jungreis, I., Kakani, E. G., Kamali, M., Kemppainen, P., Kennedy, R. C., Kirmitzoglou, I. K., Koekemoer, L. L., Laban, N., Langridge, N., Lawniczak, M. K. N., Lirakis, M., Lobo, N. F., Lowy, E., MacCallum, R. M., Mao, C., Maslen, G., Mbogo, C., McCarthy, J., Michel, K., Mitchell, S. N., Moore, W., Murphy, K. A., Naumenko, A. N., Nolan, T., Novoa, E. M., O'Loughlin, S., Oringanje, C., Oshaghi, M. A., Pakpour, N., Papathanos, P. A., Peery, A. N., Povelones, M., Prakash, A., Price, D. P., Rajaraman, A., Reimer, L. J., Rinker, D. C., Rokas, A., Russell, T. L., Sagnon, N., Sharakhova, M. V, Shea, T., Simão, F. A., Simard, F., Slotman, M. A., Somboon, P., Stegniy, V., Struchiner, C. J., Thomas, G. W. C., Tojo, M., Topalis, P., Tubio, J. M. C., Unger, M. F., Vontas, J., Walton, C., Wilding, C. S., Willis, J. H., Wu, Y. C., Yan, G., Zdobnov, E. M., Zhou, X., Catteruccia, F., Christophides, G. K., Collins, F. H., Cornman, R. S., Crisanti, A., Donnelly, M. J., Emrich, S. J., Fontaine, M. C., Gelbart, W., Hahn, M. W., Hansen, I. A., Howell, P. I., Kafatos, F. C., Kellis, M., Lawson, D., Louis, C., Luckhart, S., Muskavitch, M. a T., Ribeiro, J. M., Riehle, M. A., Sharakhov, I. V, Tu, Z., Zwiebel, L. J., Besansky, N. J., 2015. Highly evolvable malaria vectors: The genomes of 16 *Anopheles* mosquitoes. Science 347, 1258522-1-1258522–8.

Niaré, O., Markianos, K., Volz, J., Oduol, F., Touré, A., Bagayoko, M., Sangaré, D., Traoré, S. F., Wang, R., Blass, C., Dolo, G., Bouaré, M., Kafatos, F. C. C., Kruglyak, L., Touré, Y. T., Vernick, K. D. D., 2002. Genetic loci affecting resistance to human malaria parasites in a West African mosquito vector population. Science 298, 213–216.

Oliveira, G. de A., Lieberman, J., Barillas-Mury, C., 2012. Epithelial nitration by a peroxidase/NOX5 system mediates mosquito antiplasmodial immunity. Science 335, 856–859.

Oppenheim, S. J., Rosenfeld, J. A., Desalle, R., 2017. Genome content analysis yields new insights into the relationship between the human malaria parasite *Plasmodium falciparum* and its anopheline vectors. BMC Genomics 18, 205.

Palmer, W. J., Jiggins, F. M., 2015. Comparative genomics reveals the origins and diversity of arthropod immune systems. Mol. Biol. Evol. 32, 2111–2129.

Paré, A. C., Vichas, A., Fincher, C. T., Mirman, Z., Farrell, D. L., Mainieri, A., Zallen, J. A., 2014. A positional Toll receptor code directs convergent extension in *Drosophila*. Nature 515, 523–527.

Porter, J. R., 1973. Agostino Bassi bicentennial (1773-1973). Bacteriol. Rev. 37, 284–8.

Povelones, M., Waterhouse, R. M., Kafatos, F. C., Christophides, G. K., 2009. Leucine-rich repeat protein complex activates mosquito complement in defense against *Plasmodium* parasites. Science 324, 258–261.

Raghavendra, K., Barik, T. K., Reddy, B. P. N., Sharma, P., Dash, A. P., 2011. Malaria vector control: From past to future. Parasitol. Res. 108, 757–779.

Ramirez, J. L., Garver, L. S., Brayner, F. A., Alves, L. C., Rodrigues, J., Molina-Cruz, A., Barillas-Mury, C., 2014. The role of hemocytes in *A. gambiae* antiplasmodial immunity. J. Innate Immun. 6, 119–128.

Ranson, H., Lissenden, N., 2016. Insecticide resistance in African *Anopheles* mosquitoes: A worsening situation that needs urgent action to maintain malaria control. Trends Parasitol. 32, 187–196.

Rebl, A., Goldammer, T., Seyfert, H.M., 2010. Toll-like receptor signaling in bony fish. Vet. Immunol. Immunopathol. 134, 139–150.

Rhodes, V. L. M., Michel, K., 2017. Chapter 4 – Modulation of Mosquito Immune Defenses as a Control Strategy, Arthropod Vector: Controller of Disease Transmission, Volume 1.

Riehle, M. M., Markianos, K., Niaré, O., Xu, J., Li, J., Touré, A. M., Podiougou, B., Oduol, F., Diawara, S., Diallo, M., Coulibaly, B., Ouatara, A., Kruglyak, L., Traoré, S. F., Vernick, K. D., 2006. Natural malaria infection in *Anopheles gambiae* is regulated by a single genomic control region. Science 312, 577–579.

Riehle, M. M., Xu, J., Lazzaro, B. P., Rottschaefer, S. M., Coulibaly, B., Sacko, M., Niare, O., Morlais, I., Traore, S. F., Vernick, K. D., 2008. *Anopheles gambiae* APL1 is a family of variable LRR proteins required for Rel1-mediated protection from themalaria parasite, *Plasmodium berghei*. PLoS One 3, e3672.

Roach, J. C., Glusman, G., Rowen, L., Kaur, A., Purcell, M. K., Smith, K. D., Hood, L. E., Aderem, A., 2005. The evolution of vertebrate Toll-like receptors. Proc. Natl. Acad. Sci. 102, 9577–9582.

Rosenberg, R., Rungsiwongse, J., 1991. The number of sporozoites produced by individual malaria oocysts. Am. J. Trop. Med. Hyg. 45, 574–577.

Ross, R., 1899. Inaugural lecture on the possibility of extirpating malaria from certain localities by a new method. Br. Med. J. 2, 1–4.

Ross, R., Smyth, J., 1897. On some peculiar pigmented cells found in two mosquitoes fed on malarial blood. Indian J. Malariol. 34, 47.

Sánchez-Vargas, I., Scott, J. C., Poole-Smith, B. K., Franz, A. W. E., Rie Barbosa-Solomieu, V., Wilusz, J., Olson, K. E., Blair, C. D., 2009. Dengue virus type 2 infections of *Aedes aegypti* are modulated by the mosquito's RNA interference pathway. PLoS Pathog 5, e1000299.

Satyavathi, V. V., Minz, A., Nagaraju, J., 2014. Nodulation: An unexplored cellular defense mechanism in insects. Cell. Signal. 26, 1753–1763.

Scholte, E. J., Knols, B. G. J., Samson, R. A., Takken, W., 2004. Entomopathogenic fungi for mosquito control: A review. J. Insect Sci. 4, 1–24.

Scholte, E. J., Njiru, B. N., Smallegange, R. C., Takken, W., Knols, B. G. J., 2003. Infection of malaria (*Anopheles gambiae* s.s.) and filariasis (*Culex quinquefasciatus*) vectors with the entomopathogenic fungus *Metarhizium anisopliae*. Malar. J. 2, 29.

Schwartz, A., Koella, J. C., 2002. Melanization of *Plasmodium falciparum* and C-25 Sephadex beads by field-caught *Anopheles gambiae* (Diptera: Culicidae) from Southern Tanzania. J. Med. Entomol. 39, 84–88.

Screen, S. E., St. Leger, R. J., 2000. Cloning , expression , and substrate specificity of a fungal chymotrypsin. J. Biol. Chem. 275, 6689–6694.

Shin, S. W., Bian, G. W., Raikhel, A. S., 2006. A toll receptor and a cytokine, Toll5A and Spz1C, are involved in toll antifungal immune signaling in the mosquito *Aedes aegypti*. J. Biol. Chem. 281, 39388–39395.

Shin, S. W., Kokoza, V., Bian, G., Cheon, H. M., Yu, J. K., Raikhel, A. S., 2005. REL1, a homologue of *Drosophila* Dorsal, regulates Toll antifungal immune pathway in the female mosquito *Aedes aegypti*. J. Biol. Chem. 280, 16499–16507.

Sinden, R. E., Garnham, P. C. C., 1973. A comparative study on the ultrastructure of *Plasmodium* sporozoites within the oocyst and salivary glands, with particular reference to the incidence of the micropore. Trans. R. Soc. Trop. Med. Hyg. 67, 631–637.

Sinka, M. E., Bangs, M. J., Manguin, S., Rubio-Palis, Y., Chareonviriyaphap, T., Coetzee, M., Mbogo, C. M., Hemingway, J., Patil, A. P., Temperley, W. H., Gething, P. W., Kabaria, C. W., Burkot, T. R., Harbach, R. E., Hay, S. I., 2012. A global map of dominant malaria vectors. Parasit. Vectors 5, 69.

St. Leger, R. J., Joshi, L., Roberts, D., 1998. Ambient pH is a major determinant in the expression of cuticle-degrading enzymes and hydrophobin by *Metarhizium anisopliae*. Appl. Environ. Microbiol. 64, 709–713.

Sung, J. M., Lee, J. O., Humber, R. A., Sung, G. H., Shrestha, B., 2006. *Cordyceps bassiana* and Production of Stromata *in vitro* Showing *Beauveria* Anamorph in Korea. Mycobiology 34, 1–6.

Thomas, M. B., 2018. Biological control of human disease vectors: a perspective on challenges and opportunities. BioControl 63, 61–69.

Thomas, M. B., Read, A. F., 2007. Can fungal biopesticides control malaria? Nat. Rev. Microbiol. 5, 377–383.

Trager, W., 1942. A strain of the mosquito *Aedes aegypti* selected for susceptibility to the avian malaria parasite *Plasmodium lophurae*. J. Parasitol. 28, 457–465.

Valanne, S., Wang, J. H., Rämet, M., 2011. The *Drosophila* Toll signaling pathway. J. Immunol. 186, 649–656.

Valero-Jiménez, C. A., Wiegers, H., Zwaan, B. J., Koenraadt, C. J. M., van Kan, J. A. L., 2016. Genes involved in virulence of the entomopathogenic fungus *Beauveria bassiana*. J. Invertebr. Pathol. 133, 41–49.

Vaughan, J. A., Hensley, L., Beier, J. C., 1994. Sporogonic development of *Plasmodium yoelii* in five anopheline species. J. Parasitol. 80, 674–681.

Vega-Aquino, P., Sanchez-Peña, S., Blanco, C. A., 2010. Activity of oil-formulated conidia of the fungal entomopathogens *Nomuraea rileyi* and *Isaria tenuipes* against lepidopterous larvae. J. Invertebr. Pathol. 103, 145–149.

Vernick, K. D., Fujioka, H., Seeley, D. C., Tandler, B., Aikawa, M., Miller, L. H., 1995. *Plasmodium gallinaceum*: A refractory mechanism of ookinete killing in the mosquito, *Anopheles gambiae*. Exp. Parasitol. 80, 583–595.

Vlachou, D., Zimmermann, T., Cantera, R., Janse, C. J., Waters, A. P., Kafatos, F. C., 2004. Real-time, *in vivo* analysis of malaria ookinete locomotion and mosquito midgut invasion. Cell. Microbiol. 6, 671–685.

Vogels, C. B. F., Bukhari, T., Koenraadt, C. J. M., 2014. Fitness consequences of larval exposure to *Beauveria bassiana* on adults of the malaria vector *Anopheles stephensi*. J. Invertebr. Pathol. 119, 19–24.

Wanchoo, A., Lewis, M. W., Keyhani, N. O., 2009. Lectin mapping reveals stage-specific display of surface carbohydrates in *in vitro* and haemolymph- derived cells of the entomopathogenic fungus *Beauveria bassiana*. Microbiology 155, 3121–3133.

Wang, C., St. Leger, R. J., 2007. The *Metarhizium anisopliae* perilipin homolog MPL1 regulates lipid metabolism, appressorial turgor pressure, and virulence. J. Biol. Chem. 282, 21110–21115.

Wang, X. X., He, P. H., Feng, M. G., Ying, S. H., 2014. *BbSNF1* contributes to cell differentiation, extracellular acidification, and virulence in *Beauveria bassiana*, a filamentous entomopathogenic fungus. Appl. Microbiol. Biotechnol. 98, 8657–8673.

Ward, A., Hong, W., Favaloro, V., Luo, L., 2015. Toll receptors instruct axon and dendrite targeting and participate in synaptic partner matching in a *Drosophila* olfactory circuit. Neuron 85, 1013–1028.

Waterhouse, R. M., Kriventseva, E. V., Meister, S., Xi, Z., Alvarez, K. S., Bartholomay, L. C., Barillas-Mury, C., Bian, G., Blandin, S., Christensen, B. M., Dong, Y., Jiang, H., Kanost, M. R., Koutsos, A. C., Levashina, E. A., Li, J., Ligoxygakis, P., Maccallum, R. M., Mayhew, G. F., Mendes, A., Michel, K., Osta, M. A., Paskewitz, S., Shin, S. W., Vlachou, D., Wang, L., Wei, W., Zheng, L., Zou, Z., Severson, D. W., Raikhel, A. S., Kafatos, F. C., Dimopoulos, G., Zdobnov, E. M., Christophides, G. K., 2007. Evolutionary dynamics of immune-related genes and pathways in disease-vector mosquitoes. Science 316, 1738–1743.

Whitten, M. M. A., Shiao, S. H., Levashina, E. A., 2006. Mosquito midguts and malaria: Cell biology, compartmentalization and immunology. Parasite Immunol. 28, 121–130.

World Health Organization, 2016. Malaria vaccine: WHO position paper-January 2016, The Weekly Epidemiological Record.

Yassine, H., Kamareddine, L., Osta, M. A., 2012. The mosquito melanization response is implicated in defense against the entomopathogenic fungus *Beauveria bassiana*. PLoS Pathog. 8, e1003029.

Zhang, S., Xia, Y. X., Kim, B., Keyhani, N. O., 2011. Two hydrophobins are involved in fungal spore coat rodlet layer assembly and each play distinct roles in surface interactions, development and pathogenesis in the entomopathogenic fungus, *Beauveria bassiana*. Mol. Microbiol. 80, 811–826.

Zheng, L., Cornel, A. J., Wang, R., Erfle, H., Voss, H., Ansorge, W., Kafatos, F. C., Collins, F. H., 1997. Quantitative trait loci for refractoriness of *Anopheles gambiae* to *Plasmodium cynomolgi* B. Science 276, 425–428.

Zou, Z., Souza-Neto, J., Xi, Z., Kokoza, V., Shin, S. W., Dimopoulos, G., Raikhel, A., 2011. Transcriptome analysis of *Aedes aegypti* transgenic mosquitoes with altered immunity. PLoS Pathog. 7, e1002394.

Figures – Chapter 1

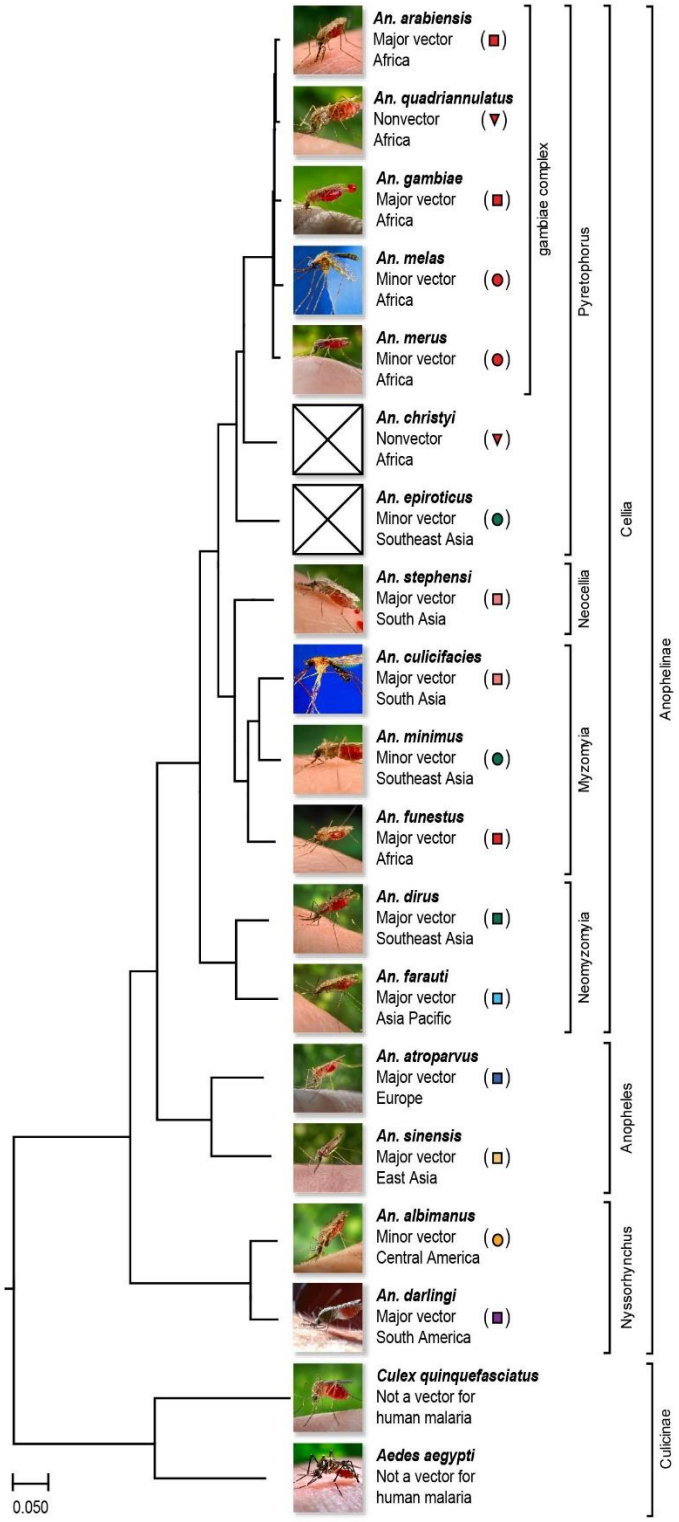

Figure 1-1 Maximum Likelihood Analysis of sequences mosquito genomes

The maximum likelihood molecular phylogeny of all sequenced anophelines and two mosquito outgroups was constructed from the aligned protein sequences of 1085 single-copy orthologs. Branch length values were kindly provided by Dr. Robert Waterhouse. Vector status and geographic distribution are indicated and are reflected in symbols adjacent to each species (rectangles, major vectors; ellipses, minor vectors, triangles, nonvectors) and are colored according to the following geographic ranges: red: Africa, pink: South Asia, green: South-East Asia, light blue: Asia Pacific, dark blue: Europe, light orange: East Asia, dark orange: Central America, purple: South America. Figure adapted from (Neafsey et al., 2015). Photo credits: *An. arabiensis, An. dirus, An.quadriannulatus, An.gambiae, An.merus, An.stephensi, An.minimus, An.funestus, An.farauti, An.atroparvus, An.sinensis, An.albimanus, Culex quinquefasciatus, Aedes aegypti,* Centers for Disease Control, James Gathany. *An. melas* and *An. culicifacies* , Walter J. Stoffer, http://www.faculty.ucr.edu/~legneref/medical/medicalindex.htm..

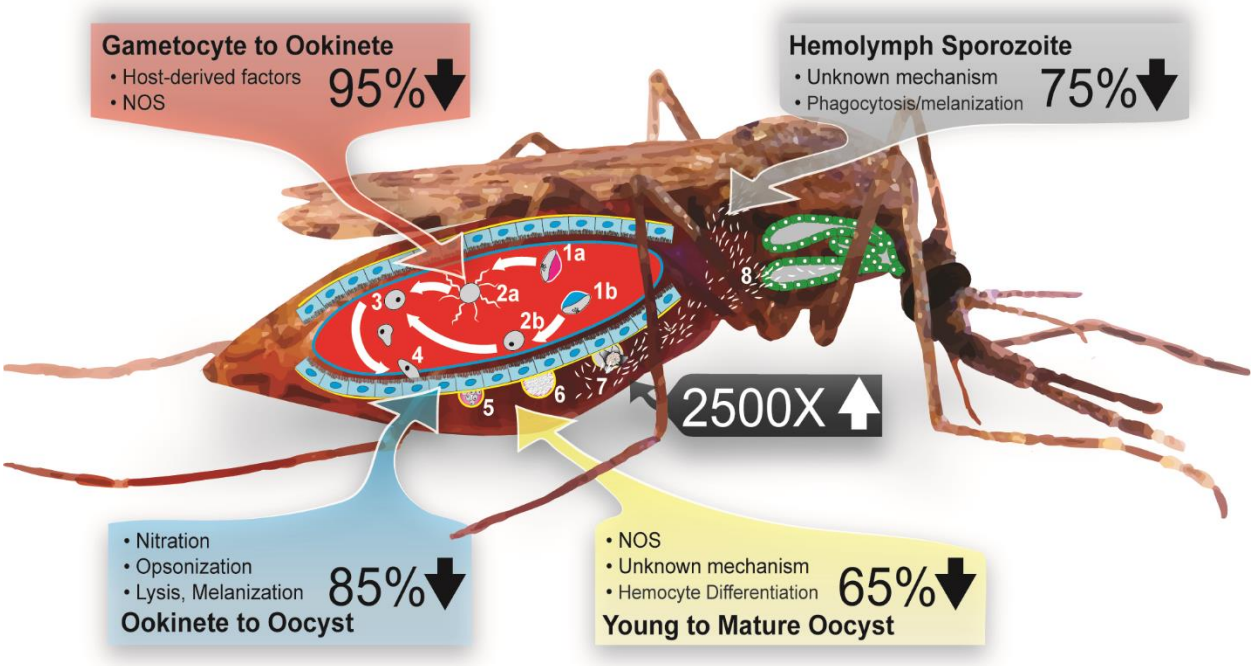

Figure 1-2 The *Plasmodium* life cycle and interactions with host immunity

The sporogonic cycle of Plasmodium parasites begins with the ingestion of male microgametocytes and female macrogametocytes (1a/1b). Exflagellated microgametes (2a) fertilize female macrogametocytes (2b) to produce diploid zygotes (3). Zygotes develop into mobile ookinetes (4), which penetrate through the midgut epithelium and form early oocyst beneath the basal lamina (5). Mature oocysts (6) rupture and release thousands of sporozoites (7) into the hemolymph, and the hemolymph flow carries them to the salivary gland (8). The colored boxes depict major parasite life stage transitions, the immune mechanisms that act upon them, and the percent increase or decrease in parasite numbers that occur at each of these transitions. The figure was adapted from the iconic life cycle depiction by Baton and Ranford-Cartwright, 2005.

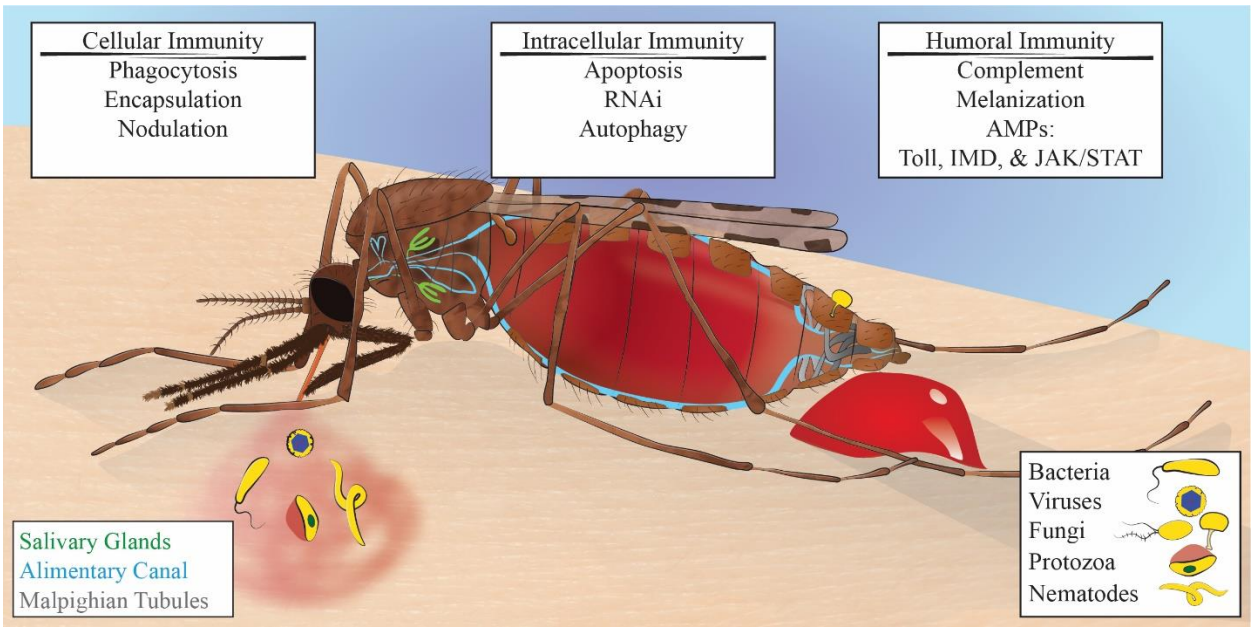

Figure 1-3 The three prongs of insect immunity

Adult mosquitoes must successfully fight infections with a variety of entomopathogens, including bacteria, viruses, fungi, protozoa, and nematodes present in their environment and sugar/blood meals. Mosquito responses to these responses can be broadly divided into three categories: humoral, intracellular, and cellular immune processes. Figure adapted from Bartholomay and Michel 2018.

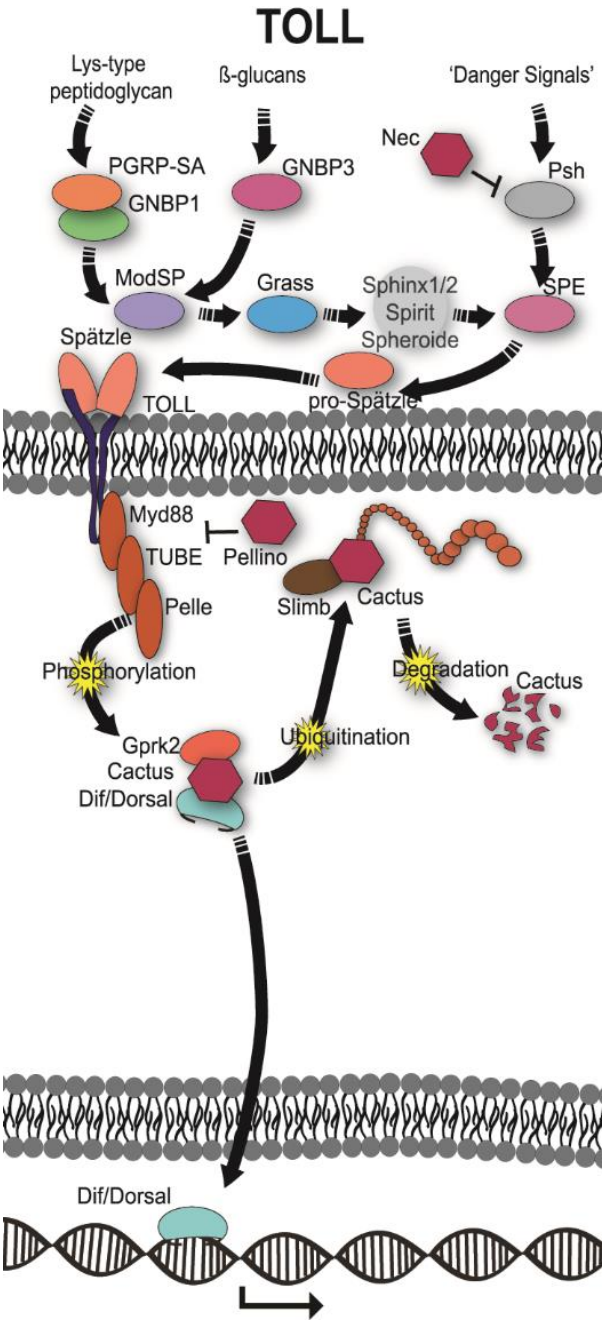

Figure 1-4 The Toll Pathway

The Toll pathway is a major regulator of the immunotranscriptome of insects. Toll is initiated through the binding of the cytokine, spätzle, to the Toll receptor. This NF-κB signaling pathway is characterized by a substantial number of pathway-specific proteins. Their interactions through phosphorylation, ubiquitination, and protease cleavage culminate in the activation of NF-kB transcription factors Dif/Dorsal (REL1 in mosquitoes). Red hexagons, major inhibitors of each pathway, Light blue, transcription factors. Figure adapted from Rhodes and Michel, 2017.

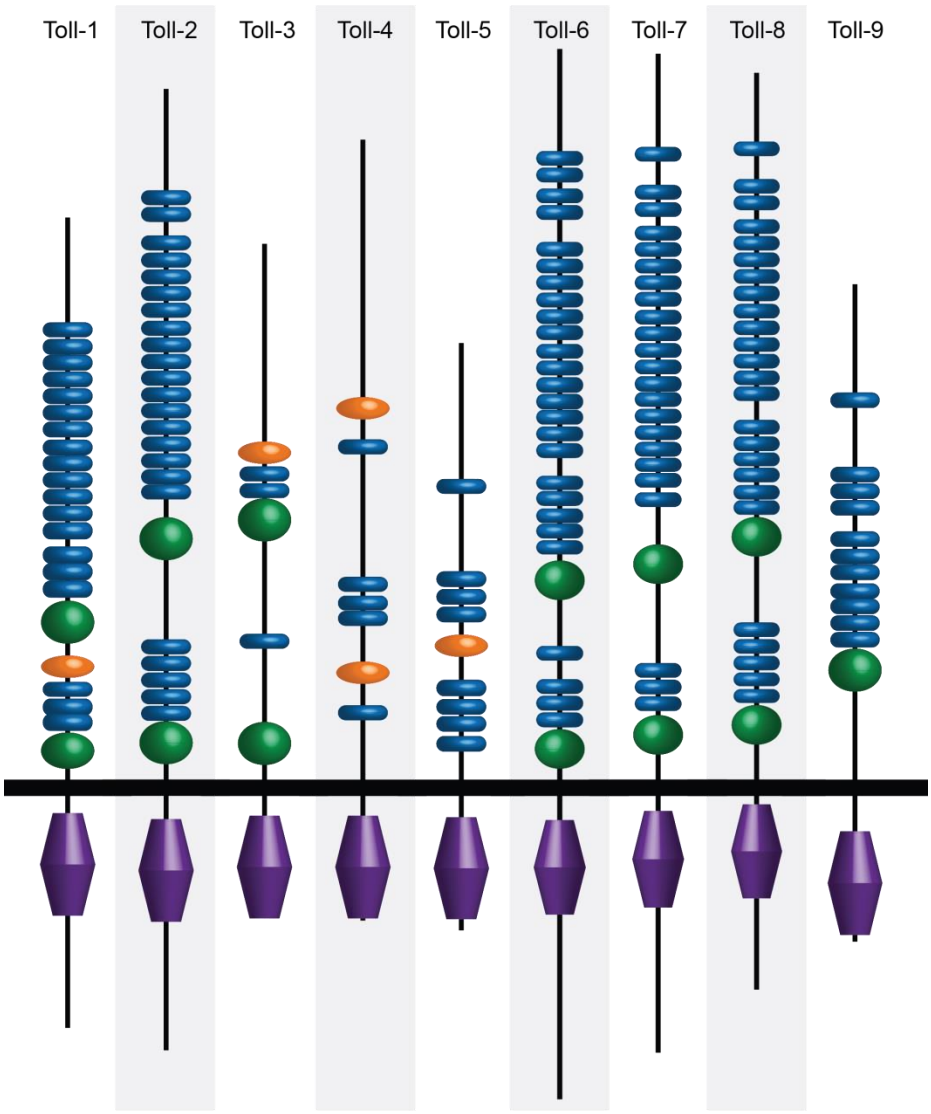

Figure 1-5 *Drosophila* TLR and domain structure

The schematic representations of the predicted domain structure of nine *Drosophila melanogaster* Toll receptors. Domains are drawn to scale and predicted using Pfam, TMHMM Server version 2.0, and LRR finder. LRR, (blue) LRR-CT (green), LRR-NT (orange), and TIR (purple) domains are indicated.

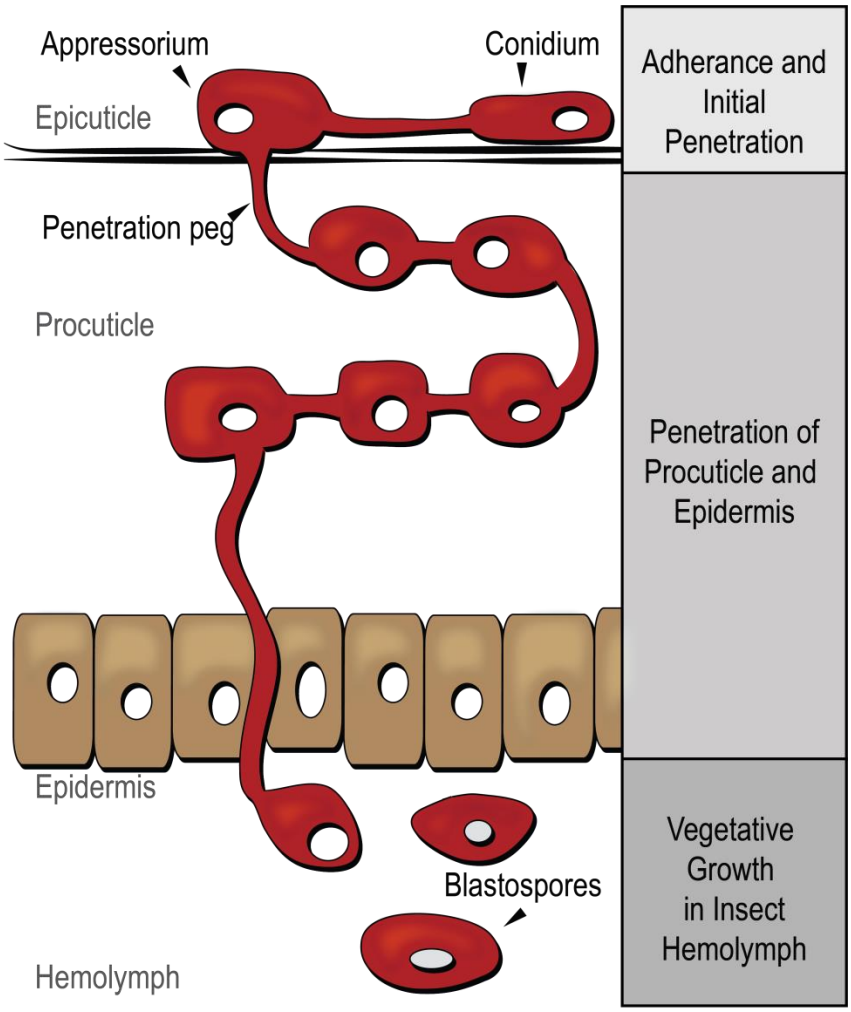

Figure 1-6 Infection of host insect by *B. bassiana*

During the pathogenic life cycle of *B. bassiana*, an infective conidium adheres to the host cuticle and germinates, forming a penetrative structure termed an appressorium. This organ penetrates host cuticle through mechanical pressure and degradative enzymes. Once the fungus penetrates the various layers of cuticle and reaching host hemolymph, it grows vegetatively in the host hemolymph. Figure adapted from Thomas and Read, 2007.

Chapter 2 - The interplay between dose and immune system activation determines fungal infection outcome in the African malaria mosquito, *Anopheles gambiae*

Victoria L. Rhodes[a1], Matthew B. Thomas[b], and Kristin Michel[a*]

[a]Division of Biology, Kansas State University, Manhattan, KS, 66506, USA

[b]Department of Entomology and Center for Infectious Disease Dynamics, The Pennsylvania State University, University Park, PA, 16802, USA

Abstract

The Toll pathway is a central regulator of antifungal immunity in insects. In mosquitoes, the Toll pathway affects infections with the fungal entomopathogen, *Beauveria bassiana,* which is considered a potential mosquito biopesticide. We report here the use of *B. bassiana* strain I93-825 in *Anopheles gambiae* to analyze the impact of Toll pathway modulation on mosquito survival. Exposure to a narrow dose range of conidia by direct contact decreased mosquito longevity and median survival. In addition, fungal exposure dose correlated positively and linearly with hazard ratio. Increased Toll signaling by knockdown of its inhibitor, *cactus*, decreased survivorship of uninfected females, increased mosquito survival after low dose *B. bassiana* exposure, but had little effect following exposure to higher doses. This observed trade-off could have implications for development of *B. bassiana* as a prospective vector control tool. On the one hand, selection for small increases in mosquito immune signaling across a narrow dose range could impair efficacy of *B. bassiana*. On the other hand, costs of immunity and the capacity for higher doses of fungus to overwhelm immune responses could limit evolution of resistance.

Introduction

Control of mosquito vectors continues to be contingent on widespread use of chemical insecticides through insecticide treated bed nets (ITNs) and indoor residual spraying (IRS). However, resistance to widely used public health insecticides is spreading within many vector populations, jeopardizing the effectiveness of these control tools (Benelli and Beier, 2017; Hemingway and Ranson, 2000; Ranson et al., 2011; Ranson and Lissenden, 2016). In light of these circumstances, research into the development of alternative vector control measures has been a topic of continuing interest (Barreaux et al., 2017; Hemingway et al., 2016; Kamareddine, 2012; Thomas et al., 2012). Entomopathogenic fungi have displayed promise as a means to provide such alternative control measures as novel biopesticides (Blanford et al., 2005; Scholte et al., 2005).

Fungal strains of species such as *B. bassiana* and *Metarhizium anisopliae* are virulent to *An. gambiae*, the major vector of malaria in sub-Saharan Africa (Kikankie et al., 2010; Mnyone et al., 2009; Scholte et al., 2003). These fungi are attractive as potential vector control agents as they have been shown to reduce factors related to mosquito physiology that affect disease transmission such as feeding propensity (Blanford et al., 2011; Howard et al., 2010b; Scholte et al., 2006), fecundity (Blanford et al., 2011; Scholte et al., 2006), flight (Blanford et al., 2011), host seeking (George et al., 2011), and vector competence (Blanford et al., 2005), as well as maintaining efficacy against both insecticide-resistant and non-resistant mosquito strains (Blanford et al., 2011; Farenhorst et al., 2009; Howard et al., 2010a). In addition, the use of these agents has the added benefit of horizontal transmission, as seen by transmission through mating from infected males to uninfected females, affecting a population larger than originally exposed (García-Munguía et al., 2011). Formulations of infective conidia can be stable for long periods of

time and remain effective (Blanford et al., 2012a), allowing them to be produced in large batches, shipped, and stored for later use. These infective conidia can be applied in a similar fashion as chemical insecticides through IRS (Heinig et al., 2015; Mnyone et al., 2010), or by application to a variety of substrates that can target host searching mosquitoes (Sternberg et al., 2016, 2014), or resting sites for blood-fed mosquitoes (Farenhorst et al., 2008; Lwetoijera et al., 2010; Mnyone et al., 2012; Scholte et al., 2005).

B. bassiana conidia attach to insect cuticles, where they germinate and penetrate this external barrier and proliferate within the mosquito (reviewed in Mascarin and Jaronski, 2016). This infection can progress in a matter of 3 to 14 days before host death, with both fungal dose and isolate virulence playing an important role in the time of death (Bell et al., 2009; Blanford et al., 2012b, 2011; Farenhorst and Knols, 2010; Heinig and Thomas, 2015; Mnyone et al., 2009; Valero-Jiménez et al., 2014). However, insects have the ability to evolve an increased tolerance and/or resistance to pathogens, including entomopathogenic fungi. For example, laboratory-based selection experiments in the greater wax moth, *Galleria mellonella*, lead to specific resistance to *B. bassiana* through continuous exposure to sublethal doses of conidia in a mere 25 generations (Dubovskiy et al., 2013). The mechanism of enhanced resistance in such a short period of time was linked to augmentations of front line defenses such as cuticle strength, phenoloxidase activity, and antimicrobial peptide (AMP) expression, highlighting the critical role that immune activation state can have on the survivability to fungal infections (Dubovskiy et al., 2013).

The Toll pathway has an important role in antifungal immunity within insects and, thus, alterations in basal activation of this pathway can influence an insect's ability to resist *B. bassiana* (Shin et al., 2005). The Toll pathway is characterized by an extracellular protease

cascade and an intracellular signal transduction pathway (reviewed in Valanne et al., 2011). This pathway culminates in the cleavage and degradation of a major pathway inhibitor, Cactus, whereby the NF-κB transcription factor, REL1, is released to translocate into the nucleus and affect gene transcription. Several pathogen killing mechanisms are controlled by the Toll pathway, including AMPs, transcriptional upregulation of components of the complement-like pathway, as well as negative regulators of the melanization cascade (Frolet et al., 2006).

Changes in Toll pathway activation, indeed, have been shown to affect a mosquito's ability to overcome a *B. bassiana* infection. Previous studies show that increased expression levels of *REL1* boosts basal immunity and positively affected the ability of *Aedes aegypti* to withstand *B. bassiana* infections (Bian et al., 2005; Shin et al., 2006, 2005). Therefore, the Toll pathway could constitute a selection target for resistance to *B. bassiana*. However, the intersection between Toll signaling and *B. bassiana* exposure dose is less clear. In this study, we describe a trade-off between activation of Toll immune signaling and survival in the context of exposure of *An. gambiae* to various doses of *B. bassiana* conidia.

Materials and Methods

Mosquito rearing and maintenance

The *An. gambiae* G3 strain was reared at standard rearing conditions (An et al., 2011). Adult females 2-4 days old (n = 35 per treatment per replicate) were separated, placed in experimental cups (straight-walled paper cans, 1 pint volume, Neptune Paper Products Inc., Fort Lee, NJ, USA), and fed sugar water (8% D-fructose, 2.5 mM 4-aminobenzoic acid, Sigma-Aldrich, St. Louis, MO, USA) *ad libitum* (Beier et al., 1994).

Fungal immune challenge

B. bassiana strain I93-825 conidia were formulated in a mix of mineral oils (80:20 Isopar:Ondina) as described previously (Blanford et al., 2005) with the conidia concentration adjusted to 1.24×10^9 conidia/ml. The resulting conidia formulation was spread on 15 cm filter paper (VWR, Radnor, PA, USA) with volumes adjusted to 4 ml using 80:20 Isopar:Ondina. For all experiments, the following exposure doses were used: (i) low dose = $7.02 \times 10^6/cm^2$, medium dose = $1.4 \times 10^7/cm^2$, high dose = $2.81 \times 10^7/cm^2$. An oil-only formulation was used as the negative control. After drying for 1 day at room temperature, the filter papers were adhered to WHO exposure cones (WHO Collaborating Centre, Universiti Sains Malaysia (USM), Penang, Malaysia), which were modified through the lateral insertion of a 1-inch, mesh-covered straw to allow airflow during mosquito aspirations. Adult females were aspirated into the modified WHO exposure cones through the top of the cone and forced to rest on filter paper for 30 min at room temperature. After exposure, mosquitoes were returned to the experimental cups and kept using the standard rearing conditions cited above. The effects of fungal exposure dose on survival of naive mosquitoes were tested using six biological replicates, with 35 adult female mosquitoes

used for each dose and replicate. The effects of ds*GFP* injection, *REL1*knockdown (kd) and *Cactus*kd on female mosquito survival across the fungal exposure dose range were tested using three biological replicates, with 35 adult mosquitoes used for each dose and replicate.

Mortality analysis

After fungal exposure, survival was monitored daily until mortality reached 100 %. The resulting data were analyzed and graphed using Kaplan−Meier and compared using the log-rank (Mantel−Cox) test and Hazard Ratios (HRs). Lethal time 50 (LT50, the time point after exposure where 50 % of mosquitoes in a given treatment had died) was evaluated statistically using one-way or two-way ANOVA followed by Tukey's multiple comparisons post-test. All statistical analyses were performed using GRAPHPAD PRISM software v.6 (GraphPad Software Inc., La Jolla, CA, USA).

Total RNA extraction

Female mosquitoes were flash-frozen in liquid nitrogen and stored at −80 °C. Samples were homogenized in 200 µl Trizol (Ambion, Life Technologies, Carlsbad, CA, USA), and total RNA was extracted using a final volume of 1 ml Trizol according to the manufacturer's instructions. Pellets were air dried and suspended in 100 µl RNAse-free water (Fisher Scientific, Waltham, MA, USA). RNA was further purified with an RNeasy mini kit (Qiagen, Valencia, CA, USA) using the manufacturer's protocol and eluted in 50 µl RNAse-free water. RNA integrity was verified by agarose gel electrophoresis and concentration determined by Nanodrop (Thermo Fisher Scientific, Waltham, MA, USA).

cDNA synthesis

An. gambiae cDNA was synthesized from 100 ng purified total RNA with an iScript cDNA synthesis kit (Bio-Rad, Hercules, CA, USA), using oligo(dT) and random hexamer primers, in a total reaction volume of 20 µl, following the manufacturer's protocol.

dsRNA synthesis

DNA templates for dsRNA synthesis were generated by two rounds of PCR from cDNA of 3-4 day-old female sugar-fed mosquitoes. The first-round PCR was performed in a 25 µL total reaction volume with 100 ng of cDNA as template, using the following primer sets for REL1, Cactus, and GFP (T7 5' extension is underlined). AGAP009515_REL1_F, 5'-GAATTAATACGACTCACTATAGGGAGAATCA ACAGCACGACGATGAG-3'; AGAP009515_REL1_R, 5'-GAATTAATACGACTCACTATAGGGAGATCGAAAAAGCGCA CCTTAAT-3'; AGAP007938_Cactus_F, 5'-GAATTAATACGACTCACTATAGGGAGAGTCCG CTCTACACATCAGCA-3'; AGAP007938_Cactus_R, 5'-GAATTAATACGACTCACTATAGG GAGACCGTTCGGGTTAATGATGAC-3'; GFP_F, 5'-TAATACGACTCACTATAGGGAGTTC ACCTTGATGCCGTTC-3'; GFP_R,5'-TAATACGACTCACTATAGGGCACATGAAGCAGCA CGACTT-3'. Resulting PCR products were purified by gel extraction (QIAquick Gel Extraction Kit; Qiagen, Hilden, Germany) and eluted in 30 µl Milli-Q® purified water. The second-round PCR was performed using 1 µl of first-round PCR product as template in a 50 µl reaction volume using the following T7 primer: T7_F/R, 5'-TAATACGACTCACTATAGG G-3'. The second-round PCR product was precipitated with 1 volume of isopropanol and the resulting pellet resuspended in 50 µl

Milli-Q purified water. DsRNA synthesis and purification was performed with the AmpliScribe T7-Flash transcription kit (Epicentre, Madison, WI, USA) using 1 μg of second-round purified PCR product following manufacturer's protocol, and the dsRNA after final precipitation was resuspended in RNase-free water at a concentration of 3.0 μg/μL.

dsRNA injection

Adult females (n = 35 per treatment and replicate) were anaesthetized under a constant flow of CO_2 (5 L/min) using a benchtop Flowbuddy^{TM} regulator (Genesee Scientific, San Diego, CA, USA). Mosquitoes were injected with 69 nl of dsRNA solution (210 ng total/mosquito) under the wing base using a nanoinjector (Nanoject II, Drummond Scientific, Broomall, PA, USA). Injected females were kept at standard rearing conditions (see above) until exposed to *B. bassiana* conidia 3 days post injection.

Real Time-quantitative Polymerase Chain Reaction (RT-qPCR)

Total RNA extraction and template cDNA was prepared from female mosquitoes (n = 8 per treatment and replicate) as described in sections 2.4 and 2.5 above. RT-qPCR was performed using iQ SYBR Green Supermix (Bio-Rad, Hercules, CA, USA) according to the manufacturer's protocol with 2 μl of 1:2 diluted cDNA as template for each 20 μL volume reaction. RT-qPCRs were executed on the StepOne Plus RT-PCR System and analyzed with the StepOne software 2.0 (Life Technologies, Carlsbad, CA, USA) with the following amplification protocol: an initial cycle of 5 min at 95 °C, 40 cycles of 15 s at 95 °C, 1 min at 59 °C and 1 min at 72 °C (detection). Primers were as follows: rpS7_F, 5'- CGCTATGGTGTTCGGTT CC-3'; rpS7_R, 5'-TGCTGCAAACTTCGG-3'; REL1_kd_F 5'-TCAACAGATGCCAAAAGAG GAAAT-3';

REL1_ kd_R, 5’-CTGGTTGGAGGGATTGTG-3’; Cactus_ kd_F, 5’-AATCTGGG CCTGATGGACA-3’; Cactus_ kd_R 5’-ACTGCCAGGTGCAGTTGAGT-3’.

To test for potential changes in transcription upon fungal infection, female mosquitoes (n = 8 per treatment and biological replicate) were exposed to fungal conidia (see section 2.2 above), and flash frozen in liquid nitrogen at 1, 2, 4, and 6 days post exposure (dpe). RT-qPCR was used to determine expression of *An. gambiae REL1* (AGAP009515) and *Cactus* (AGAP007938) transcripts relative to the housekeeping gene *40S ribosomal Protein S7* (rp*S7*, AGAP010592). Fold change was assessed using a modification of the delta delta threshold cycle (ΔΔCt) method (Pfaffl, 2001), taking into account primer efficiencies (Figure A1). RpS7 expression was used as reference and unexposed treatments as calibrator conditions. RT-qPCRs were performed with three technical replicates using cDNA templates collected from four (*REL1*) and five (*Cactus*) biological replicates.

To test for gene knockdown after dsRNA treatment, female mosquitoes (n = 8 per treatment and replicate) were collected 3 days post dsRNA injection. Fold change in expression was assessed using a modification of the ΔΔCt method (Pfaffl, 2001) method, taking into account the primer efficiencies (Figure A1). *RpS7* expression was used as reference and ds*GFP* treatments as calibrator conditions. RT-qPCRs were performed in triplicate with three biological replicates.

Results

Time course of fungal-induced mortality

To determine the impact of *B. bassiana* strain I93-825 on mosquito mortality, female *An. gambiae* mosquitoes were exposed to a high dose of *B. bassiana* conidia (2.81 x 10^7/cm^2 filter paper), and the number of dead mosquitoes per treatment was recorded daily until all mosquitoes in the experiment were dead. The resulting survival curves of all treatment groups were sigmoidal (Figure 2-1; individual biological replicates Figure A2). Oil-only control treated mosquitoes had an average maximum life span of 44.1 ± 1.3 (SE) days and a LT50 of 26.5 ± 0.9 (SE) days. Daily percent mortalities were below 5 % for the first three weeks of data collection, resulting in an average of 70 % survival within this treatment at 21 days (Figure 2-1B).

Mosquitoes exposed to the high conidial dose of *B. bassiana* were nine times more likely to die as compared to oil only controls (log-rank test, $P < 0.0001$; HR = 9.035). Their LT50 (9.0 ± 0.8 days) and maximum lifespan (25.5 ± 1.4 days) were significantly shorter as those observed for oil-only controls (Figure 2-1C, D; One-way ANOVA $P < 0.0001$; Tukey's posttest $P < 0.0001$ for both). Daily mortality after high dose fungal exposures was strongly elevated between 6-25 dpe, with an average of 17 % daily mortality across this time interval (Figure 2-1B). In comparison, oil-only control treatments did not reach this percent daily mortality until 40 dpe (Figure 2-1B).

Dose dependence of the mosquito-killing phenotype

To investigate whether the life-shortening phenotype of *B. bassiana* was dose-dependent, we assessed survival after exposure to low (7.02 x 10^6/cm^2 filter paper) and medium (1.4 x 10^7/cm^2 filter paper) doses. Comparison of the survival curves among all four treatment groups revealed

significant differences (log rank test; $df = 3$, $\chi^2 = 314.5$, $P < 0.0001$), with each fungal exposure concentration significantly decreasing mosquito survival over time (Figure 2-1A, individual biological replicates Figure A2). HRs of conidia-exposed mosquitoes as compared to oil-only controls positively correlated with dose (Table 2-1, $R^2 = 0.9973$; $P = 0.0013$). Interestingly, while fungal exposure dose strongly affected the amplitude of increase in daily mortality, the onset of increased mortality rate remained the same, and was first observed between 7-8 days post exposure (dpe) for all fungal treatments (Figure 2-1B). LT50 of fungal treatments decreased with increasing conidial dose (one-way ANOVA, $df = 3$, $F = 26.71$, $P < 0.0001$ with Tukey's multiple comparison test for pairwise analyses; Figure 2-1C). Likewise, maximum mosquito lifespan negatively correlated with *B. bassiana* dose (one-way ANOVA, $df = 3$, $F = 26$, $P < 0.0001$ with Tukey's multiple comparison test for pairwise analyses, Figure 2-1D). Both, LT50 and maximum lifespan correlated negatively with dose, and both relationships fitted a linear regression model, with R^2 of 0.7966 and 0.9343, respectively.

Transient increase in expression of Toll pathway components after fungal infection

To determine whether *B. bassiana* infection in *An. gambiae* affects the Toll pathway, we initially tested for changes in the expression profiles of *REL1* and *Cactus* using RT-qPCR. Transcript levels of *REL1* following exposure to a high dose of *B. bassiana* conidia were determined by RT-qPCR with primer sequences specific to the 5' sequence common to both *REL1* alternative splice isoforms at 1, 2, 4, and 6 dpe.

REL1 expression did not change after *B. bassiana* exposure as compared to oil-only treatments, and in addition remained steady during the time course of the experiment. Thus, neither infection nor age of the sampled mosquitoes impacted *REL1* transcript levels (Figure 2-

2A; individual biological replicates are shown in Figure A-3; Two-way ANOVA, treatment, *df* = 1, *F* = 0.4599, *P* = 0.4434; time, *df* = 3, *F* = 0.9656, *P* = 0.5117; interaction, *df* = 3, *F* = 0.2424, *P* = 0.8650). *Cactus* expression also remained constant over time in oil-only treated mosquitoes. However, *Cactus* expression increased significantly and transiently after fungal exposure, with a maximum two-fold increase observed at 4 dpe (Figure 2-2B; Two-way ANOVA, treatment, *df* = 1, *F* = 11.54, *P* = 0.0043; time, *df* = 3, *F* = 1.363, *P* = 0.2947; interaction, *df* = 3, *F* = 2.87, *P* = 0.0740; Sidak's multiple comparison test, *P* = 0.0099).

*REL1*kd decreases survivorship following fungal challenge in a dose-dependent manner

To investigate whether Toll signaling is activated after *B. bassiana* exposure and does limit *B. bassiana*-induced pathology, we used an RNAi-based silencing approach to inhibit the Toll pathway by depleting *REL1* prior to fungal exposure (Figure 2-3, Figure A-4 for percent knockdown, Figure A-5 for individual replicates). Mosquitoes were injected with ds*GFP* or ds*REL1* and exposed three days post injection to no (oil-only control), low, medium, and high doses of *B. bassiana* conidia.

Both, the fungal dosage, as well as dsRNA injection treatment significantly affected LT50 (Two-way ANOVA, Table 2-2). In the absence of fungal infection, injection of ds*REL1* did not change mosquito survival as compared to ds*GFP*-injected mosquitoes or uninjected controls (log-rank test, ds*REL1*/uninjected *P* = 0.2145; ds*GFP*/uninjected, *P* = 0.3881; see Table 2-3 for HRs).

As observed previously, in all treatment groups, mosquito survival decreased with increasing dose of conidia. However, neither injection nor the presence of non-specific dsRNA

significantly altered mosquito survival curves with HRs at each dose close to 1 (Figure 2-3, Table 2-3). Daily mortalities of ds*GFP*-injected mosquitoes were comparable to those of uninjected mosquitoes at all fungal exposure doses (Figure A-6). LT50 also remained similar between ds*GFP*-treated and uninjected mosquitoes regardless of conidial dose (Two-way ANOVA, treatment, $df = 1$, $F = 0.4695$, $P = 0.5030$; conidial dose, $df = 3$, $F = 31.77$, $P < 0.0001$; interaction, $df = 3$, $F = 0.3192$, $P = 0.8113$), as did maximum life span. Likewise, after low dose exposure, ds*REL1*-survival curves were not significantly different from ds*GFP*-injected survival curves (log-rank test, $P = 0.1728$) or uninjected control curves (log-rank test, $P = 0.2471$) (Figure 2-3B).

However, after exposure to the medium and high dosages of conidia, ds*REL1*-treated mosquitoes experienced statistically significantly decreased survivorship when compared to ds*GFP*-treated or uninjected mosquitoes (Figure 2-3C, D; log-rank test, all $P < 0.0001$). This decrease was largely a consequence of increased daily mortality beginning at 8 dpe in ds*REL1*-injected mosquitoes. Ds*REL1* injection affected the amplitude of increase in daily mortality, e.g. the percent daily mortality at 8 dpe was four-fold higher in ds*REL1*-injected compared to uninjected mosquitoes (Figure A-6). Again, the onset of this phenotype remained the same, and was first observed between 7-8 dpe for all dsRNA treatments at all fungal doses (Figure 2-3).

Ds*REL1*-injected mosquitoes exposed to higher doses of *B. bassiana* resulted in higher likelihood of mortality when compared to uninjected controls (HR ds*REL1*-injected low dose = 1.061 ± 0.251, HR ds*REL1*-injected medium dose = 1.753 ± 0.167, HR ds*REL1*-injected high dose = 2.848 ± 0.371). After a low dose exposure, LT50 decreased from 14.67 dpe in ds*GFP*-injected control mosquitoes to 12.2 dpe after ds*REL1*-injection (Figure 2-3E). Ds*REL1*-injected mosquitoes had consistently lower LT50s when compared to ds*GFP*-injected controls exposed to

the same amount of fungus (medium dose: ds*GFP*-injected = 12.0 dpe, ds*REL1*-injected = 9.0 dpe; high dose: ds*GFP*-injected = 11.7 dpe, ds*REL1*-injected = 8.0 dpe) (Figure 2-3E).

Cactus knockdown increases survivorship after low dose fungal challenge

To determine if increased basal activity of the Toll pathway can limit *B. bassiana*-induced pathology, we used RNAi to induce the intracellular Toll signaling cascade prior to fungal exposure through the depletion of *cactus* (Figure 2-3, Figure A-5 for individual replicates). Mosquitoes were injected with ds*Cactus* or ds*GFP* as a control treatment and subsequently exposed three days post injection to no (oil-only control), low, medium, and high doses of *B. bassiana* conidia.

Ds*Cactus* injection induced a 33 % decrease in *Cactus* transcripts (Figure A-4). In the absence of infection, *Cactus* knockdown was detrimental to mosquito survival, as ds*Cactus*-depleted mosquitoes had overall significantly decreased survival rates compared to both uninjected and ds*GFP*-injected controls (Figure 2-3A; log-rank test, $P < 0.0001$), with HRs of 1.7 and 2.1, respectively (Table 2-3). *Cactus* knockdown increased daily mortality, with *Cactus*-depleted mosquitoes experiencing a consistently higher percent mortality than all other treatments between 8 to 29 dpe, peaking first at 12 dpe and later at 27 dpe (Figure A-6A).

However, ds*Cactus*-depleted mosquitoes had better survival rates after low dose exposure to *B. bassiana* as compared to both, uninjected and ds*GFP*-injected controls (Figure 2-3D). Ds*Cactus*-depleted mosquitoes consistently displayed lower daily percent mortalities than all other treatments from 5 dpe to 17 dpe, peaking at a later stage of infection at 26 dpe (Figure A-6B). In addition, ds*Cactus*-injected treatments also had significantly increased medium survival compared to ds*GFP*-injected treatments at low dose (Student's t-test, $P = 0.0274$). While we did

see increases in variability between biological replicates at low dose exposures, ds*Cactus*-injected mosquitoes consistently fared better than ds*GFP*-injected mosquitoes, particularly prominent within the third biological replicate (Figure A-5).

The ds*Cactus*-dependent, enhanced survival observed at low dose exposures was lost after mosquitoes were exposed to a medium or high dose of *B. bassiana* (Figure 2-3C, D). Here, ds*Cactus*- and ds*GFP*-injected mosquitoes had overlapping survival curves (Figure 2-3C, D), and the HR of ds*Cactus*-depleted mosquitoes as compared to controls was close to 1 (Table 2-3).

Discussion

This study investigated the activation state of Toll immune signaling and its consequences on survival following exposure of *An. gambiae* to various doses of *B. bassiana* conidia. We find that exposure to *B. bassiana* across a narrow dose range has far reaching effects on mortality in these insects. Additionally, we find that decreasing the basal activation of the Toll signaling pathway through RNAi knockdown of *REL1* leads to severe and significant decreases in survivorship following infection with *B. bassiana*. Conversely, knockdown of a key negative regulator of Toll signaling, *Cactus*, revealed that an increased basal level of Toll activation is beneficial following low dose exposures to this fungus. Interestingly, as fungal dosage increased, any benefits gained through preemptive Toll activation were lost. Together this study highlights the impact fungal exposure dose has on immune system activation and infection outcome in this vector species.

We find that exposure doses over a small range of conidial concentrations can have large effects on the LT50 in *An. gambiae*. While the highest conidial density utilized in this study was merely four fold higher than the lowest dose (2.81×10^7 vs. 7.02×10^6 conidia/cm^2), we observed a strong inverse linear relationship between exposure dose and LT50. Similar relationships between *B. bassiana* dose and LT50, albeit using log-fold changes in dose, have been reported previously in mosquitoes, including *An. gambiae* and *Anopheles stephensi* (Bukhari et al., 2010; Dong et al., 2012; Mnyone et al., 2009), as well as the beetle *Agelastica alni* (Sonmez et al., 2017). However, the relationship between dose and LT50 seems to be highly variable and dependent not only on experimental conditions and the *B. bassiana* strain tested (Bukhari et al., 2010; Heinig et al., 2015; Mnyone et al., 2009), but also on insect species and sex (Bukhari et al., 2010; Taylor et al., 2007). The outcome of *B. bassiana* infection of insects is the result of genotype by genotype interactions, and as well as environmental conditions that likely

affect fungal growth kinetics and virulence factor expression by the fungus. Future studies will have to show whether the strong impact on LT50 of *An. gambiae* over a narrow dose range of *B. bassiana* is influenced by any or all of these factors or is a general feature of *B. bassiana* infection dynamics in *An. gambiae*.

The Toll pathway is a well-known critical negative regulator of fungal infections in mosquitoes (Shin et al., 2005). We were therefore surprised to find that at the lowest dose used in this study this immune signaling pathway does not seem to be activated by *B. bassiana*. This notion is supported by the following two observations: First, at low dose, ds*REL1* injection does not decrease mosquito survival, suggesting that expression of anti-fungal immune factors does not occur or only occurs to levels that do not alter infection outcome. Second, knockdown of *Cactus*, which activates the intracellular Toll signaling cascade, even in absence of immune challenge (Frolet et al., 2006), does increase mosquito survival rates, demonstrating the impact Toll signaling has on limiting *B. bassiana* induced pathology. Possible mechanisms underlying this lack of Toll activation at low dose exposure are mosquito immune system evasion or suppression by *B. bassiana*. Both mechanisms are commonly employed by arthropod pathogens, e.g. *Plasmodium* spp. utilize both to escape nitration, encapsulation, and melanization responses (Boëte et al., 2004; Lambrechts et al., 2007; Michel et al., 2005; Molina-Cruz et al., 2015; Ramphul et al., 2015). *B. bassiana* is no exception, employing a variety of methods to evade host immunity, including utilization of hyphal bodies that evade immune recognition through a lack of antigenic surface compounds (Pendland et al., 1993). *Beauveria* is also capable of modulating insect responses to suppress a host's immune defenses using different molecular classes of proteases and toxins (reviewed in Joop and Vilcinskas, 2016). Given that our data strongly

suggest that immune evasion does occur, future studies might test which, if any, of these mechanisms are employed by *B. bassiana* during infection of *An. gambiae*.

Our data further suggest that this immune evasion/suppression is less effective after exposure to higher doses of *B. bassiana*. The Toll pathway was indeed activated and limited the pathology induced by fungal infection, as evidenced by decreased survivorship after ds*REL1* injection. Intriguingly, at these doses, *Cactus* knockdown did not impact infection outcome, suggesting that the Toll pathway was activated to its maximum level, and the loss of enhanced survival in *Cactus* knockdown treatments is due to overwhelming *B. bassiana* infection in these treatments. One caveat to this explanation is the incomplete *Cactus* knockdown we observed. While the percent knockdown of *Cactus* transcripts to similar levels is observed by other investigators (Frolet et al., 2006; Garver et al., 2009), we cannot rule out that further decrease of *Cactus* transcript levels may have increased survival at higher fungal dose exposures. With technological advances in mosquito genetic manipulation, these technical limitations may be overcome in future studies (Li et al., 2018, 2017; O'Brochta et al., 2011).

Previous studies using the same methodology employed by us have shown that in adult, female *An. stephensi* mosquitoes, *B. bassiana* load slowly increases over time after exposure to conidia increases in mosquitoes (Bell et al., 2009). Using the same strain as employed in this current study, Bell *et al*. observed initial increase in load at the time where survival rates decreased, suggesting that increased pathology is tightly linked to fungal growth. The mechanisms underlying the pathology induced by *B. bassiana* are numerous and include enzymes capable of utilizing and depleting host hemolymph of sugars and toxins such as beauvericin capable of killing host cells (reviewed in Valero-Jiménez et al., 2016).

While our data confirm that even transient genetic manipulation of the Toll pathway has strong impact on pathology of *B. bassiana* infection, it is currently unclear whether these effects are due to altered disease tolerance or pathogen resistance (Medzhitov et al., 2012; Raberg et al., 2009). Tolerance, the ability to survive higher loads of the pathogen, to *B. bassiana* infection has thus far not been observed in insects (reviewed in Lu and St. Leger, 2016). Indeed, *An. gambiae* mosquitoes depleted of the thioester-containing protein 1 (*TEP1*) succumb to the infection more quickly paired with increased fungal loads, favoring the notion that the melanization immune response limits fungal growth and, thus, increases resistance. Interestingly, *TEP1* expression is increased in *Cactus*-depleted *An. gambiae* (Frolet et al., 2006), providing a possible explanation for the increased survivorship we observed in *Cactus*-depleted mosquitoes exposed to low levels of conidia. However, this is likely not the only mechanism underlying the increased resistance to infection. *Cactus* depletion in adult *An. gambiae* changes the expression of 3 % of the protein coding genes (Garver et al., 2009), including the expression of gambicin, one of the two antifungal peptides known in *An. gambiae* (Vizioli et al., 2001, 2000).

An overzealous immune system can have detrimental effects on fitness, due to trade-offs in resource allocation as well as pathology induced by immune byproducts (recently reviewed by Schwenke et al., 2016). This trade-off is exemplified in our experimental system, as *Cactus* knockdown, and thus increased Toll signaling, increases survivorship in the presence of infection, while in the absence of infection leads to higher mortality. Indeed, reduced longevity was previously observed in *An. gambiae* females after *Cactus* knockdown (Garver et al., 2009). In addition, *Cactus* depletion from mosquitoes as well as *D. melanogaster* has additional phenotypic consequences, including melanotic tumor formation, (Frolet et al., 2006; Qiu et al., 1998), proliferation of and shifts in hemocyte subpopulations, (Qiu et al., 1998; Ramirez et al.,

2014), and increased lipid droplet presence in mosquito midguts (Barletta et al., 2016). While increased melanization in a *Cactus*-depleted background is likely to contribute to the mosquito's ability to fight *B. bassiana* infection (Yassine et al., 2012), increased melanization also reduces longevity in absence of infection (An et al., 2011). In addition, constitutive activation of the Toll pathway inhibits Akt signaling and leads to depletion of nutrient stores (DiAngelo et al., 2009), which may contribute to the decreased longevity in *Cactus*-depleted mosquitoes.

Nevertheless, after low fungal exposure dose, we observed increased longevity in adult female mosquitoes, suggesting that under our experimental conditions the fitness cost of a constitutively activated Toll pathway is at least partially rescued by the mosquito's increased ability to fight infection. A more detailed analysis of fitness parameters, including biting frequency and population growth parameters will allow the quantification of the fitness trade-offs between increased basal levels of immunity and the ability to overcome *B. bassiana* infection suggested by the data presented herein. In how far this observed trade-off may have implications for the development of *B. bassiana* as a prospective vector control tool is currently unclear. On the one hand, selection for small increases in mosquito immune signaling across a narrow dose range could impair efficacy of *B. bassiana*. On the other hand, costs of immunity, the capacity for higher doses of fungus to overwhelm immune responses, coupled with the ability to deliver such doses in the field (Heinig et al., 2015), as well as the use of *B. bassiana* in an integrated vector management approach (Sternberg and Thomas, 2017) is likely to limit evolution of resistance. The potential impact of increased basal immunity on *B. bassiana* infection outcome can be experimentally addressed in future studies that recapitulate abiotic field conditions and genetic variation in the mosquito host, along with a spectrum of available fungal strains (Kim et al., 2013; Pava-Ripoll et al., 2008; Peng and Xia, 2015; Xie et al., 2015).

Acknowledgements

We thank all members of the Michel lab for their help with mosquito colony maintenance. The authors would like to thank the College of Veterinary Medicine Core Facilities at Kansas State University for the use of equipment for RT-qPCR-based quantifications. This work is supported by the National Institutes of Health grant number R01AI095842 and through USDA-ARS Specific Cooperative Agreement 58-5430-4-022 (both to K.M.) and by National Institutes of Health grant number R01AI110793 (to M.B.T.). Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the funding agencies. This is contribution 18-202-J from the Kansas Agricultural Experiment Station.

Author Contributions

VR, KM: participated in study design, coordination and wrote the manuscript. MT made specific contributions to the initial phase of the project, including, but not limited to, provision of *B. bassiana* spore formulations and technical assistance with experimental design. All authors read and approved the final manuscript.

References

An, C., Budd, A., Kanost, M. R., Michel, K., 2011. Characterization of a regulatory unit that controls melanization and affects longevity of mosquitoes. Cell. Mol. Life Sci. 68, 1929–1939.

Barletta, A. B. F., Alves, L. R., Nascimento Silva, M. C. L., Sim, S., Dimopoulos, G., Liechocki, S., Maya-Monteiro, C. M., Sorgine, M. H. F., 2016. Emerging role of lipid droplets in *Aedes aegypti* immune response against bacteria and Dengue virus. Sci. Rep. 6, 19928.

Barreaux, P., Barreaux, A. M. G., Sternberg, E. D., Suh, E., Waite, J. L., Whitehead, S. A., Thomas, M. B., 2017. Priorities for broadening the malaria vector control tool kit. Trends Parasitol. 33, 763–774.

Beier, M. S., Pumpuni, C. B., Beier, J. C., Davis, J. R., 1994. Effects of para-aminobenzoic acid, insulin, and gentamicin on *Plasmodium falciparum* development in anopheline mosquitoes (Diptera: Culicidae). J. Med. Entomol. 31, 561–565.

Bell, A. S., Blanford, S., Jenkins, N., Thomas, M. B., Read, A. F., 2009. Real-time quantitative PCR for analysis of candidate fungal biopesticides against malaria: Technique validation and first applications. J. Invertebr. Pathol. 100, 160–168.

Benelli, G., Beier, J. C., 2017. Current vector control challenges in the fight against malaria. Acta Trop. 174, 91–96.

Bian, G., Shin, S. W., Cheon, H. M. M., Kokoza, V., Raikhel, A. S., 2005. Transgenic alteration of Toll immune pathway in the female mosquito *Aedes aegypti*. Proc. Natl. Acad. Sci. 102, 13568–13573.

Blanford, S., Chan, B. H. K., Jenkins, N., Sim, D., Turner, R. J., Read, A. F., Thomas, M. B., 2005. Fungal pathogen reduces potential for malaria transmission. Science 308, 1638–1641.

Blanford, S., Jenkins, N. E., Christian, R., Chan, B. H. K., Nardini, L., Osae, M., Koekemoer, L. L., Coetzee, M., Read, A. F., Thomas, M. B., 2012a. Storage and

persistence of a candidate fungal biopesticide for use against adult malaria vectors. Malar. J. 11, 354.

Blanford, S., Jenkins, N. E., Read, A. F., Thomas, M. B., 2012b. Evaluating the lethal and pre-lethal effects of a range of fungi against adult *Anopheles stephensi* mosquitoes. Malar. J. 11, 365.

Blanford, S., Shi, W., Christian, R., Marden, J. H., Koekemoer, L. L., Brooke, B. D., Coetzee, M., Read, A. F., Thomas, M. B., 2011. Lethal and pre-lethal effects of a fungal biopesticide contribute to substantial and rapid control of malaria vectors. PLoS One 6, e23591.

Boëte, C., Paul, R. E. L., Koella, J. C., 2004. Direct and indirect immunosuppression by a malaria parasite in its mosquito vector. Proc. R. Soc. L. 271, 1611–1615.

Bukhari, T., Middelman, A., Koenraadt, C .J. M., Takken, W., Knols, B. G. J. J., 2010. Factors affecting fungus-induced larval mortality in *Anopheles gambiae* and *Anopheles stephensi*. Malar. J. 9, 22.

DiAngelo, J. R., Bland, M. L., Bambina, S., Cherry, S., Birnbaum, M. J., 2009. The immune response attenuates growth and nutrient storage in *Drosophila* by reducing insulin signaling. Proc. Natl. Acad. Sci. 106, 20853–20858.

Dong, Y., Morton, J. C. J., Ramirez, J. L., Souza-Neto, J. A., Dimopoulos, G., 2012. The entomopathogenic fungus *Beauveria bassiana* activate Toll and JAK-STAT pathway-controlled effector genes and anti-dengue activity in *Aedes aegypti*. Insect Biochem. Mol. Biol. 42, 126–132.

Dubovskiy, I. M., Whitten, M. M. A., Yaroslavtseva, O. N., Greig, C., Kryukov, V. Y., Grizanova, E. V., Mukherjee, K., Vilcinskas, A., Glupov, V. V., Butt, T. M., 2013. Can insects develop resistance to insect pathogenic fungi? PLoS One 8, e60248.

Farenhorst, M., Farina, D., Scholte, E., Takken, W., Hunt, R. H., Coetzee, M., Knols, B. G. J., 2008. African water storage pots for the delivery of the entomopathogenic fungus *Metarhizium anisopliae* to the malaria vectors *Anopheles gambiae s. s.* and *Anopheles funestus*. Am. J. Trop. Med. Hyg. 78, 910–916.

Farenhorst, M., Knols, B. G. J., 2010. A novel method for standardized application of fungal spore coatings for mosquito exposure bioassays. Malar. J. 9, 27.

Farenhorst, M., Mouatcho, J. C., Kikankie, C. K., Brooke, B. D., Hunt, R. H., Thomas, M. B., Koekemoer, L. L., Knols, B. G. J., Coetzee, M., 2009. Fungal infection counters insecticide resistance in African malaria mosquitoes. Proc. Natl. Acad. Sci. 106, 17443–17447.

Frolet, C., Thoma, M., Blandin, S., Hoffmann, J. A., Levashina, E. A., 2006. Boosting NF-κB-dependent basal immunity of *Anopheles gambiae* aborts development of *Plasmodium berghei*. Immunity 25, 677–685.

García-Munguía, A. M., Garza-Hernández, J. A., Rebollar-Tellez, E. A., Rodríguez-Pérez, M. A., Reyes-Villanueva, F., 2011. Transmission of *Beauveria bassiana* from male to female *Aedes aegypti* mosquitoes. Parasit. Vectors 4, 24.

Garver, L. S., Dong, Y., Dimopoulos, G., 2009. Caspar controls resistance to *Plasmodium falciparum* in diverse anopheline species. PLoS Pathog. 5, e1000335.

George, J., Blanford, S., Domingue, M. J., Thomas, M. B., Read, A. F., Baker, T. C., 2011. Reduction in host-finding behaviour in fungus-infected mosquitoes is correlated with reduction in olfactory receptor neuron responsiveness. Malar. J. 10, 219.

Heinig, R. L., Paaijmans, K. P., Hancock, P. A., Thomas, M. B., 2015. The potential for fungal biopesticides to reduce malaria transmission under diverse environmental conditions. J. Appl. Ecol. 52, 1558–1566. doi:10.1111/1365-2664.12522

Heinig, R. L., Thomas, M. B., 2015. Interactions between a fungal entomopathogen and malaria parasites within a mosquito vector. Malar. J. 14, 22.

Hemingway, J., Ranson, H., 2000. Insecticide resistance in insect vectors of human disease. Annu. Rev. Entomol. 45, 371–391.

Hemingway, J., Shretta, R., Wells, T. N. C., Bell, D., Djimdé, A. A., Achee, N., Qi, G., 2016. Tools and Strategies for Malaria Control and Elimination: What Do We Need to Achieve a Grand Convergence in Malaria? PLoS Biol. 14, e1002380.

Howard, A. F. V, Koenraadt, C. J. M., Farenhorst, M., Knols, B. G. J., Takken, W., 2010a. Pyrethroid resistance in *Anopheles gambiae* leads to increased susceptibility to the entomopathogenic fungi *Metarhizium anisopliae* and *Beauveria bassiana*. Malar. J. 9, 168.

Howard, A. F. V, N'Guessan, R., Koenraadt, C. J. M., Asidi, A., Farenhorst, M., Akogbéto, M., Thomas, M. B., Knols, B. G., Takken, W., 2010b. The entomopathogenic fungus *Beauveria bassiana* reduces instantaneous blood feeding in wild multi-insecticide-resistant *Culex quinquefasciatus* mosquitoes in Benin, West Africa. Parasit. Vectors 3, 87.

Joop, G., Vilcinskas, A., 2016. Coevolution of parasitic fungi and insect hosts. Zoology 119, 350–358.

Kamareddine, L., 2012. The biological control of the malaria vector. Toxins (Basel). 4, 748–767.

Kikankie, C. K., Brooke, B. D., Knols, B. G. J., Koekemoer, L. L., Farenhorst, M., Hunt, R. H., Thomas, M. B., Coetzee, M., 2010. The infectivity of the entomopathogenic fungus *Beauveria bassiana* to insecticide-resistant and susceptible *Anopheles arabiensis* mosquitoes at two different temperatures. Malar. J. 9, 71.

Kim, J. S., Choi, J. Y., Lee, J. H., Park, J. Bin, Fu, Z., Liu, Q., Tao, X., Jin, B. R., Skinner, M., Parker, B. L., Je, Y. H., 2013. Bumblebee venom serine protease increases fungal insecticidal virulence by inducing insect melanization. PLoS One 8, e62555.

Lambrechts, L., Morlais, I., Awono-Ambene, P. H., Cohuet, A., Simard, F., Jacques, J. C., Bourgouin, C., Koella, J. C., 2007. Effect of infection by *Plasmodium falciparum* on the melanization immune response of *Anopheles gambiae*. Am. J. Trop. Med. Hyg. 76, 475–480.

Li, M., Akbari, O. S., White, B. J., 2018. Highly efficient site-specific mutagenesis in malaria mosquitoes using CRISPR. G3 Genes, Genomes, Genet. 8, 653–658.

Li, M., Bui, M., Yang, T., Bowman, C. S., White, B. J., Akbari, O. S., 2017. Germline Cas9 expression yields highly efficient genome engineering in a major worldwide disease

vector, *Aedes aegypti*. Proc. Natl. Acad. Sci. 114, E10540–E10549.

Lu, H. L., St. Leger, R. J., 2016. Insect immunity to entomopathogenic fungi. Adv. Genet. 94, 251–285.

Lwetoijera, D. W., Sumaye, R. D., Madumla, E. P., Kavishe, D. R., Mnyone, L. L., Russell, T. L., Okumu, F. O., 2010. An extra-domiciliary method of delivering entomopathogenic fungus, *Metharizium anisopliae* IP 46 for controlling adult populations of the malaria vector, *Anopheles arabiensis*. Parasit. Vectors 3, 18.

Mascarin, G. M., Jaronski, S. T., 2016. The production and uses of *Beauveria bassiana* as a microbial insecticide. World J. Microbiol. Biotechnol. 32, 177.

Medzhitov, R., Schneider, D. S., Soares, M. P., 2012. Disease tolerance as a defense strategy. Science 335, 936–941.

Michel, K., Budd, A., Pinto, S., Gibson, T. J., Kafatos, F. C., 2005. *Anopheles gambiae* SRPN2 facilitates midgut invasion by the malaria parasite *Plasmodium berghei*. EMBO Rep. 6, 891–897.

Mnyone, L. L., Kirby, M. J., Lwetoijera, D. W., Mpingwa, M. W., Knols, B. G. J., Takken, W., Russell, T. L., 2009. Infection of the malaria mosquito, *Anopheles gambiae*, with two species of entomopathogenic fungi: effects of concentration, co-formulation, exposure time and persistence. Malar. J. 8, 1–12.

Mnyone, L. L., Kirby, M. J., Lwetoijera, D. W., Mpingwa, M. W., Simfukwe, E. T., Knols, B. G. J., Takken, W., Russell, T. L., 2010. Tools for delivering entomopathogenic fungi to malaria mosquitoes: effects of delivery surfaces on fungal efficacy and persistence. Malar. J. 9, 246.

Mnyone, L. L., Lyimo, I. N., Lwetoijera, D. W., Mpingwa, M. W., Nchimbi, N., Hancock, P. A., Russell, T. L., Kirby, M. J., Takken, W., Koenraadt, C. J., 2012. Exploiting the behaviour of wild malaria vectors to achieve high infection with fungal biocontrol agents. Malar. J. 11, 87.

Molina-Cruz, A., Canepa, G. E., Kamath, N., Pavlovic, N. V, Mu, J., Ramphul, U. N., Ramirez, J. L., Barillas-Mury, C., 2015. *Plasmodium* evasion of mosquito immunity

and global malaria transmission: The lock-and-key theory. Proc. Natl. Acad. Sci. U. S. A. 112, 15178–15183.

O'Brochta, D. A., Alford, R. T., Pilitt, K. L., Aluvihare, C. U., Harrell, R. A., 2011. *piggyBac* transposon remobilization and enhancer detection in *Anopheles* mosquitoes. Proc. Natl. Acad. Sci. 108, 16339–16344.

Pava-Ripoll, M., Posada, F. J., Momen, B., Wang, C., St. Leger, R., 2008. Increased pathogenicity against coffee berry borer, *Hypothenemus hampei* (Coleoptera: Curculionidae) by *Metarhizium anisopliae* expressing the scorpion toxin (AaIT) gene. J. Invertebr. Pathol. 99, 220–226.

Pendland, J. C., Hung, S. Y., Boucias, D. G., 1993. Evasion of host defense by in vivo-produced protoplast-like cells of the insect mycopathogen *Beauveria bassiana*. J. Bacteriol. 175, 5962–5969.

Peng, G., Xia, Y., 2015. Integration of an insecticidal scorpion toxin (BjαIT) gene into *Metarhizium acridum* enhances fungal virulence towards *Locusta migratoria manilensis*. Pest Manag. Sci. 71, 58–64.

Qiu, P., Pan, P. C., Govind, S., 1998. A role for the *Drosophila* Toll/Cactus pathway in larval hematopoiesis. Development 125, 1909–1920.

Raberg, L., Graham, A. L., Read, A. F., 2009. Decomposing health: tolerance and resistance to parasites in animals. Philos. Trans. R. Soc. B Biol. Sci. 364, 37–49.

Ramirez, J. L., Garver, L. S., Brayner, F. A., Alves, L. C., Rodrigues, J., Molina-Cruz, A., Barillas-Mury, C., 2014. The role of hemocytes in *A. gambiae* antiplasmodial immunity. J. Innate Immun. 6, 119–128.

Ramphul, U. N., Garver, L. S., Molina-Cruz, A., Canepa, G. E., Barillas-Mury, C., 2015. *Plasmodium falciparum* evades mosquito immunity by disrupting JNK-mediated apoptosis of invaded midgut cells. Proc. Natl. Acad. Sci. 112, 1273–1280.

Ranson, H., Lissenden, N., 2016. Insecticide resistance in African *Anopheles* mosquitoes: A eorsening situation that needs urgent action to maintain malaria control. Trends Parasitol. 32, 187–196.

Ranson, H., N'Guessan, R., Lines, J., Moiroux, N., Nkuni, Z., Corbel, V., 2011. Pyrethroid resistance in African anopheline mosquitoes: What are the implications for malaria control? Trends Parasitol. 27, 91–98.

Scholte, E. J., Knols, B. G. J., Takken, W., 2006. Infection of the malaria mosquito *Anopheles gambiae* with the entomopathogenic fungus *Metarhizium anisopliae* reduces blood feeding and fecundity. J. Invertebr. Pathol. 91, 43–49.

Scholte, E. J., Ng'habi, K., Kihonda, J., Takken, W., Paaijmans, K., Abdulla, S., Killeen, G. F., Knols, B. G. J., 2005. An entomopathogenic fungus for control of adult African malaria mosquitoes. Science 308, 1641–1642.

Scholte, E. J., Njiru, B. N., Smallegange, R. C., Takken, W., Knols, B. G. J., 2003. Infection of malaria (*Anopheles gambiae s. s.*) and filariasis (*Culex quinquefasciatus*) vectors with the entomopathogenic fungus *Metarhizium anisopliae*. Malar. J. 2, 29.

Schwenke, R. A., Lazzaro, B. P., Wolfner, M. F., 2016. Reproduction–immunity trade-offs in insects. Annu. Rev. Entomol. 61, 239–256.

Shin, S. W., Bian, G. W., Raikhel, A. S., 2006. A toll receptor and a cytokine, Toll5A and Spz1C, are involved in toll antifungal immune signaling in the mosquito *Aedes aegypti*. J. Biol. Chem. 281, 39388–39395.

Shin, S. W., Kokoza, V., Bian, G., Cheon, H. M., Yu, J. K., Raikhel, A. S., 2005. REL1, a homologue of *Drosophila* Dorsal, regulates Toll antifungal immune pathway in the female mosquito *Aedes aegypti*. J. Biol. Chem. 280, 16499–16507.

Sonmez, E., Kocacevik, S., Sevim, A., Demirbag, Z., Demir, I., 2017. Efficacy of entomopathogenic fungi against the alder leaf beetle *Agelastica alni* (L.) (Coleoptera : Chrysomelidae). Acta Zool. Bulg. 69, 575–581.

Sternberg, E. D., Ng'Habi, K. R., Lyimo, I. N., Kessy, S. T., Farenhorst, M., Thomas, M. B., Knols, B. G. J., Mnyone, L. L., 2016. Eave tubes for malaria control in Africa: Initial development and semi-field evaluations in Tanzania. Malar. J. 15, 447.

Sternberg, E. D., Thomas, M. B., 2017. Insights from agriculture for the management of insecticide resistance in disease vectors. Evol. Appl.

Sternberg, E. D., Waite, J. L., Thomas, M. B., 2014. Evaluating the efficacy of biological and conventional insecticides with the new "MCD bottle" bioassay. Malar. J. 13, 499.

Taylor, K., Kimbrell, D. A., Kiger, J., Natzle, J., Tobin, S., 2007. Host immune response and differential survival of the sexes in *Drosophila*. Fly (Austin). 1, 197–204.

Thomas, M. B., Godfray, H. C. J., Read, A. F., van den Berg, H., Tabashnik, B. E., van Lenteren, J. C., Waage, J. K., Takken, W., 2012. Lessons from agriculture for the sustainable management of malaria vectors. PLoS Med. 9, e1001262.

Valanne, S., Wang, J. H., Rämet, M., 2011. The *Drosophila* Toll signaling pathway. J. Immunol. 186, 649–656.

Valero-Jiménez, C. A., Wiegers, H., Zwaan, B. J., Koenraadt, C. J. M., van Kan, J. A. L., 2016. Genes involved in virulence of the entomopathogenic fungus *Beauveria bassiana*. J. Invertebr. Pathol. 133, 41–49.

Valero-Jiménez, C. A., Debets, A. J. M., van Kan, J. A. L., Schoustra, S. E., Takken, W., Zwaan, B. J., Koenraadt, C. J. M., 2014. Natural variation in virulence of the entomopathogenic fungus *Beauveria bassiana* against malaria mosquitoes. Malar. J. 13, 479.

Vizioli, J., Bulet, P., Charlet, M., Lowenberger, C., Blass, C., Muller, H. M., Dimopoulos, G., Hoffmann, J., Kafatos, F. C., Richman, A., 2000. Cloning and analysis of a cecropin gene from the malaria vector mosquito, *Anopheles gambiae*. Insect Mol Biol 9, 75–84.

Vizioli, J., Bulet, P., Hoffmann, J. A., Kafatos, F. C., Müller, H. M., Dimopoulos, G., 2001. Gambicin: a novel immune responsive antimicrobial peptide from the malaria vector *Anopheles gambiae*. Proc. Natl. Acad. Sci. 98, 12630–12635.

Xie, M., Zhang, Y. J., Zhai, X. M., Zhao, J. J., Peng, D. L., Wu, G., 2015. Expression of a scorpion toxin gene *BmKit* enhances the virulence of *Lecanicillium lecanii* against aphids. J. Pest Sci. (2004). 88, 637–644.

Yassine, H., Kamareddine, L., Osta, M. A., 2012. The Mosquito Melanization Response Is Implicated in Defense against the Entomopathogenic Fungus *Beauveria bassiana*. PLoS Pathog. 8, e1003029.

Tables and Figures – Chapter 2

Table 2-1 Hazard Ratios of mosquitoes exposed to increasing conidial doses

Fungal exposure dose	**Hazard Ratio* (mean ± 1 SE)**
Low	2.183 ± 0.2767
Medium	2.901 ± 0.3525
High	3.702 ± 0.4542

*as compared to oil-only treatment

Table 2-2 Two-way ANOVA: ds*GFP* vs. ds*REL1* vs. ds*Cactus* LT50

ANOVA	*df*	*F*	*P* Value
Interaction	6	6.773	0.0003
Dose	3	85.2	< 0.0001
dsRNA	2	3.549	0.0447
Residual	24		

Table 2-3 Impact of dsRNA injection on hazard ratios with an without fungal exposure

Treatment		Oil-Only*	Low*	Medium*	High*
ds*GFP*	Uninjected	0.729 ± 0.599	1.018 ± 0.224	1.020 ± 0.177	0.873 ± 0.598
ds*REL1*	Uninjected	0.723 ± 0.244	1.061 ± 0.250	1.753 ± 0.167	2.848 ± 0.371
ds*Cactus*	Uninjected	1.733 ± 0.312	1.124 ± 0.405	1.010 ± 0.346	1.349 ± 0.497
ds*REL1*	ds*GFP*	0.911 ± 0.461	1.361 ± 0.208	1.745 ± 0.291	2.641 ± 0.602
ds*Cactus*	ds*GFP*	2.131 ± 1.187	1.071 ± 0.140	1.217 ± 0.302	1.325 ± 0.237
ds*REL1*	ds*Cactus*	2.147 ± 0.705	0.784 ± 0.227	0.586 ± 0.049	0.386 ± 0.046

*median ± ½ range

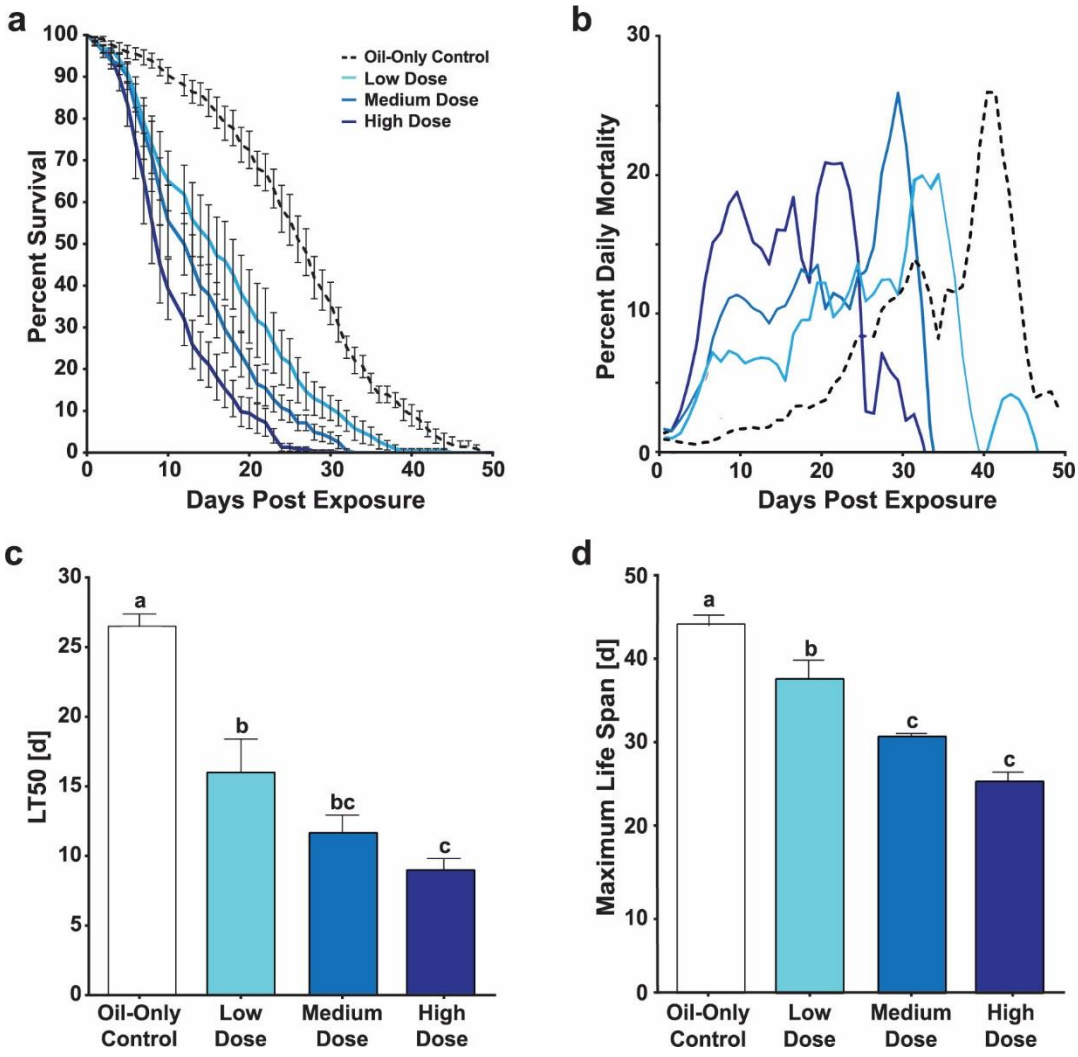

Figure 2-1 Survival following exposure to *B. bassiana* I93-825 in adult, female *An. gambiae*

(**a**) Survival curves and corresponding (**b**) daily mortalities of mosquitoes after exposure to *B. bassiana* strain I93-825 at indicated doses. (**c**) Comparison of LT50 and (**d**) maximum mosquito lifespan after *B. bassiana* exposures. Lettering denotes statistically significant differences between treatments (One-Way ANOVA with Tukey's post test, $P < 0.05$). Data are the combination of six biological replicates (Figure A-2), and are presented as mean ± 1 SEM.

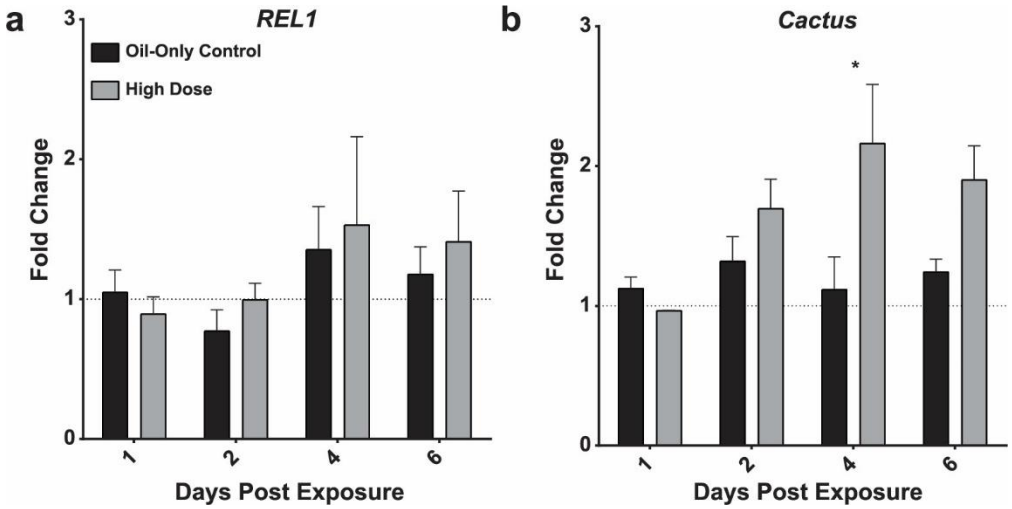

Figure 2-2 Time course of *REL1* and *Cactus* expression following *B. bassiana* exposure

Relative expression of *An. gambiae REL1* and *Cactus* transcripts after exposing female mosquitoes to a high dose of *B. bassiana* strain I93-825. Graphs depict mean transcript levels at 1, 2, 4, and 6 dpe for oil only controls (black) and high dose, *B. bassiana*-exposed mosquitoes (grey). RT-qPCR results were analyzed using *rpS7* as the reference gene and untreated mosquitoes collected at 0 dpe as calibrator condition. Statistically significant differences are indicated by asterisks (Wilkoxon signed-rank test, Sidak's post test, $P < 0.05$). Data are presented as mean ± 1 SEM from four (*REL1*) and five (*Cactus*) biological replicates; individual replicates are shown in Figure A-3.

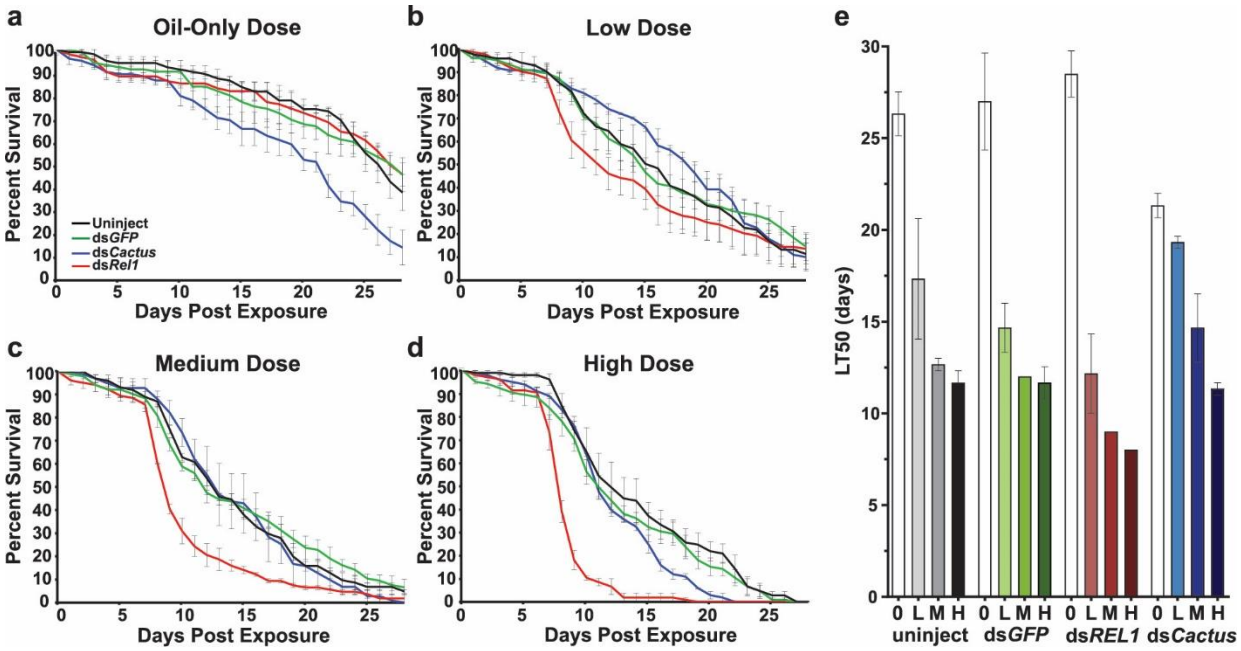

Figure 2-3 Survival of female ds*REL1*- or ds*Cactus*-injected mosquitoes following exposure to *B. bassiana* I93-825

(**a-d**) Survival curves of mosquitoes after exposure to *B. bassiana* strain I93-825 at indicated doses. At each dose, curves represent survival over time of the following treatment groups: black, no injection; green, ds*GFP*-injected; blue, ds*Cactus*-injected; red, ds*REL1*-injected. (**e**) LT50 after *B. bassiana* exposures following no injection (uninject), ds*GFP*-, ds*REL1*-, and ds*Cactus* injection, respectively, after exposure to oil only control (O), low (L), medium (M), and high (H) dose exposures. Data were combined from three biological replicates (Figure A-5) and are presented as mean ± 1 SEM.

Chapter 3 - The Toll pathway immune repertoire in anopheline mosquitoes

Victoria Rhodes[1], and Kristin Michel[1*]

[1]Division of Biology, Kansas State University, Manhattan, KS 66506, USA

[*]Corresponding author: Kristin Michel, kmichel@ksu.edu

Abstract

The Toll signaling pathway plays both developmental and immunological roles in *Drosophila melanogaster*. Toll-like receptors (TLRs) transduce extracellular activating signals to the intracellular toll signaling cascade to control gene expression and have been shown to evolve through duplication and divergence. In mosquitoes, the biological functions of individual TLRs are largely unknown and the recent sequencing of 16 mosquito genomes provides a unique opportunity to study the full toll pathway repertoire in vector and nonvector mosquitoes. Here, we annotate and describe the evolutionary history of intracellular Toll pathway members and TLRs within 21 mosquito genomes. We find the intracellular signaling pathway conserved with 1:1 orthology in all mosquito species included in our analyses. Interestingly, we find variable evolutionary rates across different pathway members, ranging from 0.09 - 4.75-fold amino acid substitution rates as compared to the conserved protein core of these mosquito species. We find that *D. melanogaster* Toll, Toll-5, and Toll-9 orthologs are duplicated in specific anopheline lineages. The most dramatic radiation of TLRs was found in the *Anopheles gambiae* complex, where five consecutive duplication events gave rise to six TOLL1/5 paralogs. These TOLL1/5 paralogs show sequence variation in their ligand binding ectodomains and display unique expression patterns, which likely impact their ligand binding specificities. Thus, these TLRs should be prioritized for experimental analyses of TLR immune function in *An. gambiae*.

Introduction

The Toll pathway, named for the type-I transmembrane *Drosophila melanogaster* protein Toll, was described originally as a signal transduction cascade controlling the development of the dorsoventral axis in the early *D. melanogaster* embryo (Anderson and Nüsslein-Volhard, 1984; Gerttula et al., 1988; Hashimoto et al., 1988). Sequence analysis of *Toll* in *D. melanogaster* quickly found similarities between the insect *Toll* and the immune-functioning human interleukin-1 receptor (IL-1R) in the intracellular regions of *Drosophila toll* and human *IL-1R*, linking what at the time appeared to be solely an insect developmental signal transduction cascade to a vertebrate immune pathway (Gay and Keith, 1991; Hashimoto et al., 1988). Shortly thereafter, Toll signaling was identified as a key pathway in the control of antimicrobial peptide expression and antifungal immunity in *Drosophila* (Lemaitre et al., 1996).

The Toll immune pathway in *D. melanogaster* consists of an extracellular protease cascade and an intracellular signal transduction pathway (reviewed in Valanne et al., 2011). Extracellular pattern recognition receptors such as peptidoglycan recognition proteins (PGRPs) and Gram-negative binding proteins (GNBPs) sense the presence of microbe-associated molecular patterns (MAMPs) (Bischoff et al., 2004; Gobert et al., 2003; Gottar et al., 2006; Michel et al., 2001). Additionally, secreted pathogen factors (El Chamy et al., 2008; Gottar et al., 2006) as well as endogenous factors released by stressed or damaged host cells, termed damage-associated molecular patterns (DAMPs), can trigger Toll signaling through activation of the protease Persephone (Issa et al., 2018; Ming et al., 2014). MAMPs, DAMPs, or pathogen-produced proteases all lead to the activation of a proteolytic cascade which amplifies the signal and culminates in the activation by cleavage of the Toll receptor ligand, Spätzle (Gottar et al., 2006; Ming et al., 2014).

Binding of spätzle to the Toll receptor (Gangloff et al., 2008) triggers receptor dimerization and intracellular signaling by recruiting the death-domain protein adaptors Myd88, Tube, and Pelle to the intracellular Toll/interleukin-1 receptor (TIR) domain of Toll (Moncrieffe et al., 2008). Pelle, functioning as a kinase, is autoactivated by this association(Shen and Manley, 2002). Activation of Pelle leads to the subsequent phosphorylation and degradation of a key inhibitor of Toll signaling, Cactus (Belvin and Anderson, 1996; Grosshans et al., 1994).

Unphosphorylated Cactus is bound to the NF-kB transcription factor Dif, preventing it from entering the nucleus. Upon phosphorylation, Cactus releases Dif, resulting in the translocation of Dif to the nucleus to initiate gene transcription (Wu and Anderson, 1998). In addition to these core members, several other proteins were identified in RNAi screens to impact the Toll signal transduction cascade. However, their placement is not yet defined. These include the proteins Myopic (Mop), Hepatocyte growth factor-regulated tyrosine kinase substrate (Hrs), Deformed epidermal autoregulatory factor 1 (Deaf1), G protein-coupled receptor kinase 2 (Gprk2), U-shaped (Ush), and Toll activation mediating protein (wispy) (Huang et al., 2010; Kuttenkeuler et al., 2010; Valanne et al., 2010).

The intracellular Toll signaling pathway is conserved within insects, and orthologs of each protein have been identified in all currently sequenced mosquito genomes (Bartholomay et al., 2010; Chen et al., 2015; Christophides et al., 2002; Neafsey et al., 2015; Waterhouse et al., 2007). The immunological role this pathway plays in *D. melanogaster* appears conserved in mosquitoes, as knockdown or overexpression of pathway members *Cactus* and the mosquito equivalent of *Dif*, *REL1*, affects survival to fungal and bacterial infections (Bian et al., 2005; Shin et al., 2006, 2005), *Plasmodium* development (Frolet et al., 2006; Garver et al., 2009; Mitri

et al., 2009; Ramirez et al., 2014; Riehle et al., 2008; Zou et al., 2011), and immunity-related gene expression (Bian et al., 2005; Garver et al., 2009; Shin et al., 2005; Zou et al., 2011).

Originally identified as a single receptor of a developmental pathway, *Drosophila Toll* is the founding member of a large gene family of Toll-like receptors (TLRs) extending throughout the Animalia kingdom from sponges to higher chordates (Leulier and Lemaitre, 2008). All members of this large receptor family are characterized by an intracellular TIR domain, a transmembrane domain, and an extracellular ligand binding region abundant in leucine-rich repeat (LRR) domains. TLRs are classified into two major types, based on the number of CF motifs (cysteine cluster on the C-terminal end of LRRs) in the extracellular TLR domain (Leulier and Lemaitre, 2008). All vertebrate TLRs described to date are single cysteine cluster (scc) TLRs, while most insect TLRs belong to the multiple cysteine cluster (mcc) TLRs. In *D. melanogaster*, there are nine Toll receptors; the originally identified *Toll* (Toll-1) plus eight additional TLR receptors, named Toll-2 through Toll-9, of which Toll-1 to 8 are mccTLRs, while Toll-9 is the only sccTLR identified in this species (Imler and Zheng, 2004). The genomes of coleopteran, dipteran, hymenopteran, and lepidopteran insects encode between 5 and 16 TLRs (Cao et al., 2015; Christophides et al., 2002; Evans et al., 2006; Waterhouse et al., 2007; Zou et al., 2007), with species-specific expansions of TLRs commonly observed in insects (Cao et al., 2015; Leulier and Lemaitre, 2008; Levin and Malik, 2017; Palmer and Jiggins, 2015). Previous genome analyses has revealed that mosquitoes possess species-specific Toll expansions, with genes TOLL1A, TOLL1B, TOLL5A, and TOLL5B corresponding to *D. melanogaster* Toll-1 and Toll-5 in *An. gambiae* and *Ae. aegypti*, but clear orthology cannot be determined phylogenetically (Waterhouse et al., 2007). Additionally, Toll-2, Toll-3, and Toll-4 orthologs are

absent mosquitoes, as well as two TLRs present in mosquitoes that are absent in *D. melanogaster*, TOLL10 and TOLL11 (Waterhouse et al., 2007).

For mammals, the biological function(s) of each TLR is well described and each plays a distinct role in immunity, as they directly recognize MAMPs (Roach et al., 2005). However, in insects, the recognition of MAMPs occurs further upstream. Very few TLRs in *D. melanogaster* have been shown to be functional in development and/or immunity. These membrane receptors have been implicated in diverse functions in *D. melanogaster*, including establishing the dorsoventral axis of the developing embryo (Anderson and Nüsslein-Volhard, 1984), transcription of antimicrobial peptides (Lemaitre et al., 1995), and axon guidance(Ward et al., 2015). Of the 9 encoded TLRs in *D. melanogaster*, Toll-1 (Lemaitre et al., 1996), Toll-2 (Eldon et al., 1994), Toll-5 (Luo et al., 2001), Toll-7 (Nakamoto et al., 2012), and Toll-8 (Akhouayri et al., 2011) have been implicated in regulation of immune signaling. Furthermore, Toll-1 (Anderson et al., 1985), Toll-2 (Williams et al., 1997), Toll-7 (Mcilroy et al., 2013), and Toll-8 (Paré et al., 2014) have also been implicated in some aspect of *D. melanogaster* development, highlighting the functional diversity of TLRs within insects.

Even less is known about the function of TLRs in mosquitoes. Both *Ae. aegypti* and *An. gambiae* TLRs display unique gene expression patterns, showing expression differences over the course of development, after blood meals, and following infection (MacCallum et al., 2011). Notably, RNAi knockdown of *Ae. aegypti TOLL5A* increases susceptibility to the entomopathogenic fungus *B. bassiana* (Shin et al., 2006) and single nucleotide polymorphisms (SNPs) within *TOLL6* increases *Plasmodium falciparum* prevalence of infection in *An. gambiae* (Harris et al., 2010). Expression of *An. gambiae* TOLL1A and TOLL5A in *D. melanogaster* cell culture can activate the expression of a firefly luciferase gene under the control of the

antimicrobial peptide Drosomycin promoter (Luna et al., 2006, 2002). However, the function of individual TLRs remains largely undescribed in these vector species and the frequent expansion events observed in insects (Cao et al., 2015; Leulier and Lemaitre, 2008; Levin and Malik, 2017; Palmer and Jiggins, 2015) make it difficult to apply TLR functions found in one species to others through sequence identity alone. Therefore, it remains important to study and analyze this important signaling pathway in species of interest, such as mosquito vectors, to facilitate the understanding of the biology of these vectors and the potential for development of novel insect control measures.

The recent sequencing and publishing of 16 anopheline genomes provides a powerful opportunity to systematically analyze the immune repertoire of TLRs and intracellular Toll pathway members over a range of vector and nonvector mosquito species (Neafsey et al., 2015). In an attempt to further our understanding of this intriguing and multifunctional signaling pathway, we present data resulting from the comprehensive manual annotation and phylogenetic analysis of the coding sequences of intracellular Toll pathway members and TLRs across 21 total mosquito species. Our results show strong 1:1 orthology within the intracellular Toll signaling cascade within mosquitoes. However, several pathway members, including the adaptor proteins Pelle, Tube, and Myd88, are accumulating amino acid substitutions at much higher rates higher than observed for conserved proteins cores identified in Neafsey et al., 2015. Our analyses of TLRs reveals gene expansions within TOLL1 and TOLL5 subfamilies for a subset of anophelines, while TLR subfamilies TOLL6-TOLL11 display strong 1:1 conservation among all analyzed mosquito species.

Materials and Methods

Obtaining Sequences

Sequences of genes orthologous to *Drosophila melanogaster* Toll-like receptors and intracellular components of the Toll signaling pathway were acquired from publically available genome assemblies obtainable through VectorBase (www.vectorbase.org) (Giraldo-Calderon et al., 2015). Genomes included in this study encompass the recently published 16 *Anopheles* genomes, previously published genomes of *An. gambiae*, *Anopheles darlingi*, and *Anopheles coluzzii*, and the culicine species *Aedes aegypti* and *Culex quinquefasciatus* (Arensburger et al., 2010; Holt et al., 2002; Lawniczak et al., 2010; Marinotti et al., 2013; Neafsey et al., 2015; Nene et al., 2007). The following species (gene nomenclature) genome assembly.gene sets were used: *Anopheles albimanus* (AALB) STECLA AaalbS2.2, *Anopheles arabiensis* (AARA) Dongola AaraD1.5, *Anopheles atroparvus* (AATE) EBRO AatrE1.4, *Anopheles christyi* (ACHR) ACHKN1017 AchrA1.4, *Anopheles coluzzii* (ACOM) Mali-NIH AcolM1.4, *Anopheles culicifacies* (ACUA) A-37 AculA1.4, *Anopheles darlingi* (ADAC) AdarC3 AdarC3.5, *Anopheles dirus* (ADIR) WRAIR2 AdirW1.4, *Anopheles epiroticus* (AEPI) Epiroticus2 AepiE1.4, *Anopheles farauti* (AFAF) FAR1 AfarF2.2, *Anopheles funestus* (AFUN) FUMOZ AfunF1.5, *An. gambiae* (AGAP) PEST AgamP4.5, *Anopheles melas* (AMEC) CM1001059_A AmelC2.3, *Anopheles merus* (AMEM) MAF AmerM2.3, *Anopheles minimus* (AMIN) MINIMUS1 AminM1.4, *Anopheles quadriannulatus* (AQUA) SANGWE AquaS1.5, *Anopheles sinensis* (ASIS) SINENSIS AsinS2.2, *Anopheles stephensi* (ASTE) SDA-500 AsteS1.4, *Ae. aegypti* (AAEL) Liverpool AaegL3.4, and *C. quinquefasciatus* (CPIJ) Johannesburg CpipJ2.3. To confirm annotated and identify additional non-annotated intracellular Toll signaling pathway members, all genomes were searched by tBLASTn using amino acid (aa) sequences of all known components from *D.*

melanogaster and *An. gambiae* as queries. Genomes were also searched by tBLASTn using the aa sequences of the TIR domains of *An. gambiae* and *D. melanogaster* TLRs to obtain a comprehensive gene list of putative TLRs across the mosquito genomes.

Manual Annotation

Manual annotation of the resulting gene lists was completed using the web-based genomic annotation editing platform, Apollo (Lee et al., 2013). Genes from species with RNA-seq data support (*An. albimanus, An. arabiensis, An. atroparvus, An. dirus, An. funestus, An. gambiae, An. minimus, An. quadriannulatus, An. stephensi,* and, *Ae. aegypti*) were annotated first. The resulting coding exons were then used as template to annotate gene models in mosquito genomes lacking transcriptional data support. The highly fragmented genome assemblies of *An. christyi* and *An. maculatus* (Neafsey et al., 2015) made it impossible to fully annotate orthologs of several Toll signaling components and TLRs. Thus, all components of the Toll signaling pathway from *An. christyi* and *An. maculatus* as well as TLRs from *An. maculatus* were removed from further analyses. All annotations were submitted to VectorBase for publication in updated gene sets. A summary of the final gene models, including nucleotide and amino acid sequences are listed in Supplemental Table 1 (intracellular Toll signaling pathway members) and Supplemental Table 2 (TLRs).

Alignments and Phylogenetic Analysis

Aa sequences of all gene models were aligned using MUSCLE (Edgar, 2004) within the MEGA7 program (Kumar et al., 2016) using default parameters. Aa sequences of the conserved intracellular TIR domains were used to reconstruct the phylogeny of all annotated mosquito

TLRs, due to the highly variable nature of the extracellular protein regions across the different TLR families. Boundaries of TIR domains were identified using Pfam (Bateman et al., 2002). Alignments and phylogenies of Toll signaling pathway members and TLRs were executed using full length protein coding sequences.

All phylogenetic analyses were performed using maximum-likelihood (ML) methodology in the MEGA7 program (Kumar et al., 2016) using the substitution models and settings outlined in Supplemental Table 3. All trees were run with a Nearest-Neighbor-Interchange (NNI) ML heuristic method, with initial trees made automatically using the NJ/BioNJ algorithm. All positions in the alignments that had less than 95% site coverage were excluded from the phylogenetic analyses. Branch support was calculated by bootstrap using 1,000 replications and is presented as percentages.

TLR Protein Motif Prediction

TLR protein motif prediction was accomplished using Pfam and LRRfinder (Bateman et al., 2002; Offord et al., 2010) to estimate LRR and TIR domain locations within revised sequences. Transmembrane domains were predicted with the TMHMM Server version 2.0 (http://www.cbs.dtu.dk/services/TMHMM/). The resulting domain locations were then translated into a visual format to scale in Adobe Illustrator (individual graphics found in Figure B-25 to B-31) and overlaid to find the consensus motif structure of TLR subclasses.

Pairwise Comparisons

Pairwise distance comparisons of alignments were performed using MEGA7 (Kumar et al., 2016) by way of the Jones-Taylor-Thornton (JTT) amino acid substitution model with Gamma

rate distribution (G). All positions in the alignments that had less than 95% site coverage were excluded from the phylogenetic analyses. Data was normalized to existing species distances previously reported (Neafsey et al., 2015) using conserved protein cores by dividing phylogenetic species distances from maximum likelihood gene trees and dividing these values by the species distances as reported in Neafsey et al., 2015. Data output was used to construct color heat maps using Heatmapper (Babicki et al., 2016).

Results

Gene model refinement for Toll pathway and TLRs

To compile a detailed list of all known *D. melanogaster* Toll pathway members, we first utilized a literature search through PubMed to mine and compile a list of identified *D. melanogaster* intracellular pathway members. Upon synthesis of the current literature, we categorized and assembled a list of 18 pathway members; Achaete (AC) (Valanne et al., 2010), Cactus (CACT) (Belvin and Anderson, 1996; Grosshans et al., 1994), Deformed epidermal autoregulatory factor-1 (DEAF1) (Kuttenkeuler et al., 2010), Dorsal-related immunity factor (DIF) (Wu and Anderson, 1998), G protein-coupled receptor kinase 2 (GPRK2) (Valanne et al., 2010), Hepatocyte growth factor regulated tyrosine kinase substrate (HRS) (Huang et al., 2010), Myd88 (MYD88) (Moncrieffe et al., 2008), Myopic (MOP) (Huang et al., 2010), Pannier (PNR) (Valanne et al., 2010), PAX transcription activation domain interacting protein (PTIP) (Valanne et al., 2010), Pelle (PELLE) (Moncrieffe et al., 2008), Pellino (PLI) (Ji et al., 2014), Supernumerary Limbs (SLMB) (Spencer et al., 1999), Spt6 (SPT6) (Valanne et al., 2010), TNF-receptor-associated factor 6 (TRAF6) (Cha et al., 2003), Tube (TUBE) (Moncrieffe et al., 2008), U-shaped (USH) (Valanne et al., 2010), and Wispy (WISP) (Valanne et al., 2010).From this list, we assembled an inventory of identified previously orthologous genes (Neafsey et al., 2015). We then mined the published annotated genomes available on VectorBase (Giraldo-Calderon et al., 2015) for 21 mosquito species using blastn and tblastn searches using the *D. melanogaster* CDS sequences as queries. Genomes in this study encompassed the recently published 16 *Anopheles* genomes (Neafsey et al., 2015), previously published genomes of *An. gambiae* (Zdobnov et al., 2002), *Anopheles coluzzii* (Coetzee et al., 2013), and *Anopheles darlingi* (Marinotti et al., 2013), as well as the genomes of the culicine species *Ae. aegypti* (Nene et al., 2007) and *Culex quinquefasciatus*

(Arensburger et al., 2010). The following species were used: *Anopheles albimanus*, *Anopheles arabiensis*, *Anopheles atroparvus*, *Anopheles christyi*, *An. coluzzii*, *Anopheles culicifacies*, *Anopheles darlingi*, *Anopheles dirus*, *Anopheles epiroticus*, *Anopheles farauti*, *Anopheles funestus*, *An. gambiae*, *Anopheles maculatus*, *Anopheles melas*, *Anopheles merus*, *Anopheles minimus*, *Anopheles quadriannulatus*, *Anopheles sinensis*, *Anopheles stephensi*, *Ae. aegypti*, and *C. quinquefasciatus*.

Using this compiled inventory, current gene models available through VectorBase (Giraldo-Calderon et al., 2015) were then manually refined for all sequenced mosquito genomes using expression data, sequence alignment, and genome comparison to identify missing exons and misannotated intron/exon boundaries. The genome assemblies of *An. christyi* and *An. maculatus* are fragmented (scaffold count/N50 values of 30,369/9,057 and 47,797/3,841, respectively), which prevented adequate annotation of full length gene models of Toll pathway members in these species. Thus, we removed these gene models from further analyses.

Of the 342 gene model coding sequences compiled across the 19 mosquito genomes, 150 required no changes to the currently published coding sequence, 140 required annotation refinements such as exon/intron boundary adjustments or removal/addition of exons, and 44 could not be fully annotated due to genome constraints of the assembled genomes, such as sequence gaps and scaffold locations (Supplemental Table 1). Additionally, sequences for eight gene models within various species could not be identified within their corresponding assembled genomes (Figure 3-1).

The same analysis of TLRs, using *D. melanogaster* Toll-1 through Toll-9 and *An. gambiae* TOLL1A, TOLL1B, TOLL5A, TOLL5B, and TOLL6 through TOLL11 sequences as blast queries, led to the documentation of 197 putative TLRs encoded in the 20 mosquito

genomes analyzed in this study. For this analysis, TLR genes encoded in the *An. christyi* genome yielded complete TLR gene models, and thus this species was included in this portion of our analyses. However, the fragmentation of the *An. maculatus* assembled genome sequence still proved difficult, preventing annotation of complete TLR gene models for this species. Of the 197 TLR gene models, 102 necessitated no changes to the currently published annotations, 68 required annotation refinements, and 25 were partial gene predictions and could not be fully annotated due to their locations at the edges of contigs within assembled genome sequences (Supplemental Table 2). Additionally, two TLR gene models, AMEM bae36a1e and AMEM 3eac684b, both belonging to *Anopheles merus* were novel predictions.

The gene models that required annotation edits are not evenly distributed across the gen families. Of the orthologous groups within our analyses, eleven (*DEAF1, HRS, SPT6, PELLE, MYD88, TUBE, CACT, TOLL6, TOLL7, TOLL8,* and *TOLL10*) required little refinement from publicly available gene models, with more than 70% of published gene models unchanged within a orthologous gene set across species (Figure 3-1, Supplemental Table 2). However, multiple orthologous gene groups consistently required refinement (> 50% of gene models), by editing of intron/exon boundaries and addition or removal of coding exons, including *GPRK2*, *TOLL11*, *MOP*, *PNR*, *PTIP*, *REL1*, *SLMB*, and *WISP*. This observation may be the result of exon number, as those gene families regularly requiring changes consistently possessed higher exon numbers (average exon number 6.25 versus 3.0 in those orthologous groups with few manual edits). Additionally, gene models for the Toll pathway transcription factor *REL1* often required editing due to the presence of alternative splicing. The complete list of all genes used in this study, including gene names, VectorBase identifiers, nucleotide sequences, and amino acid sequences, is available as supporting material (Supplemental Table 1 and 2).

Phylogenetic analysis of Toll pathway members

To identify the phylogenetic relationships among all 18 putative Toll pathway members in mosquitoes, we performed a detailed maximum likelihood analysis using the alignments of their full-length amino acid sequences. All 18 pathway members were conserved across the species included in the analyses with 1:1 orthology, which included *An. albimanus, An. arabiensis, An. atroparvus, An. coluzzii, An. culicifacies, An. darlingi, An. dirus, An. epiroticus, An. farauti, An. funestus, An. gambiae, An. melas, An. merus, An. minimus, An. quadriannulatus, An. sinensis, An. stephensi, Ae. aegypti,* and *C. quinquefasciatus*. Phylogeny of these pathway members typically followed published species tree topology (Figures B-1 to B-19) (Neafsey et al., 2015). Trees that had discrepancies between the phylogenetic relationships of proteins vs. published species relationships (Neafsey et al., 2015) were restricted to the species *An. sinensis, An. atroparvus, An. farauti*, and *An. dirus* belonging to the subgenera *Anopheles* and *Nyssorhynchus*. In these instances, encompassing the phylogenetic analyses of *AC, CACT, GPRK2, MOP, PELLE, PTIP, SPT6,* these subgenera were placed as sister groups, while published species topology shows the subgenus *Nyssorhynchus* basal to *Anopheles*. However, in every instance, this placement lacked sufficient support, with bootstrap confidence values under 75, ranging from 42-69 (Figures B-1 to B-19). In the phylogenetic analysis of *DEAF1* and *HRS*, species belonging to the subseries *Pyretophorus* were split, but this split was unsupported by bootstrap values over 75 (69 and 66, respectively).

Evolutionary distances among the orthologs of these genes varied. To analyze this variation, we performed a pairwise comparison of the evolutionary distances for all genes, normalizing these distances to the previously published evolutionary distances of these species

(Neafsey et al., 2015). This allows one to observe genes acquiring substitutions at a higher rate than the overall evolution of these species, with values of 1.0 signifying equal gene tree and species tree distances. Visualizing these pairwise comparisons across each protein of the intracellular Toll signaling transduction cascade by heat map revealed that evolutionary distances ranged from slow-evolving, with very low phylogenetic distances, to quick-evolving, with long phylogenetic distances. *PLI*, *GPRK2*, *SLMB*, *DEAF1*, *SPT6*, *PNR* and *HRS* were highly conserved among mosquitoes, with short normalized branch lengths indicative of a low rate of site substitution (mean normalized branch lengths of each orthologous group between 0.09 and 0.39) (Figure 3-2). The genes *AC*, *REL1-B*, *MOP*, *PTIP*, *USH*, and *REL1-A* were also were conserved (mean normalized branch lengths of each gene between 0.48 and 0.99) as compared to the orthologous protein core of these species. Interestingly, those genes with well-established functions and placement within the Toll pathway, including *PELLE*, *CACT*, *MYD88*, *TUBE*, and *WISP* exhibited the high average evolutionary distances in our analyses (mean normalized branch lengths of each orthologous group between 1.20 and 2.50), with the ubiquitin ligase, *TRAF6*, exhibiting the highest evolutionary distances (mean normalized branch lengths = 4.75)(Figure 3-2).

Phylogenetic analysis of mosquito TLRs

To investigate the evolutionary relationships among all TLRs included in this study, we compiled a detailed inventory of TLR paralogs and performed a phylogenetic analysis utilizing the amino acid sequences of the highly conserved intracellular TIR domains (Figure 3-3). As expected from previous analyses of mosquito TLRs (Waterhouse et al., 2007), these receptors formed well-supported clades (bootstrap values 78-99), with the majority of TLR subfamilies

(TOLL6-11) grouped at the ends of long branches (Figure 3-3). TOLL6, TOLL7, TOLL10, and TOLL11 are close paralogs of each other based on tree topology, with duplication events giving rise to TOLL6, then TOLL7, and finally the closely-related and TOLL10 and TOLL11. However, TOLL1A/1B and TOLL5A/5B, along with their closest paralogs, did not cluster into distinct, well-supported clades, but instead formed a single, large cluster that we termed the "TOLL1/5 clade". Within the TOLL1/5 clade, TOLL1A sequences segregated from all other anopheline TOLL1/5 sequences. This subclade reveals expansions for Toll-1 and Toll-5 paralogs, with several duplications of TLRs observed in species belonging to the gambiae complex anophelines (Figure 3-3) including five *An. gambiae* TLRs (TOLL1B, TOLL5A, TOLL5B, TOLLX, and TOLLY).

A lack of bootstrap support within these clades prevented further interpretation of the evolutionary relationships of these receptors within each TLR orthologous group (Figure 3-3). The low bootstrap values are likely due to the relatively conserved nature of TIR domains within, but not between, each of these TLR orthologous groups, leading to a lack of informative positions (66 sites) in the final alignment.

To better understand whether the tight phylogenetic clustering of TLRs we observed based on TIR sequence analysis reflected overall protein conservation, we analyzed the variation in protein domain structure among and within the TLR subfamilies using Pfam (Bateman et al., 2002) and LRRfinder (Offord et al., 2010). The characteristic structure of Toll receptors is preserved in all TLRs analyzed in this study, with an intracellular TIR domain and an extracellular domain containing multiple LRRs separated by a single transmembrane helix (Figure 3-4). The overall number and location of domain structures within TLR clade gene predictions for members of the TOLL6, 7, 8, 10, and 11 subfamilies was highly conserved, with

protein motif numbers and locations similar throughout each of these TLR clades (Figures B-25 to B-31). Within the TOLL1/5 clade, we found that domain architecture varied depending on subclade. Members of the TOLL1A subclade had similar domain architecture, possessing two LRR-NT domains absent from other TOLL1/5 clade anopheline TLRs (Figure 3-4, Figure B-25). Phylogenetic relationships (Figures B-20 to B-24) were mirrored in domain architecture similarities (Figures B-25 to B-31).

To resolve the phylogenetic history of each orthologous TLR clade, we performed separate phylogenetic analyses for each TLR subfamily using full-length amino acid sequences. For the majority of TLR subfamilies, we found a 1:1 orthology among all 21 mosquito species included in our analyses, with three key exceptions: TOLL8, TOLL9, and the TOLL1/5 clade (Figures B-20 to B-24). These are discussed further in the following sections.

TOLL8 Phylogeny

In an effort to gain additional informative residues in the analysis, the phylogeny of individual TLR subfamilies was also determined on its own, without the inclusion of other subfamilies. In our analysis of TOLL8 phylogeny, for example, the number of informative sites increased from 66 to 1229, improving the ability to determine phylogenetic relationships of orthologs within individual TLR subfamilies across the analyzed mosquito genomes.

Maximum likelihood analysis of the TOLL8 subfamily revealed that the *Neomyzomya* series (*An. dirus* and *An. farauti*) are placed within series *Myzomyia*, differing from its species placement as the basal series within the subgenus *Cellia*. While the grouping of these two species is corroborated by a strong bootstrap value of 99, placement within *Myzomyia* is not, and thus we cannot make determination on the true nature of this phylogeny. Additionally, *C.*

quinquefasciatus was found to encode two copies of *TOLL8* (Figures B-22, B-28). Both *TOLL8* genes, CPIJ018010 and CPIJ019764 are single exon gene sequences, with highly similar amino acid (99.92% identity) and nucleotide sequences (99.21% identity). Both genes are located on relatively short contigs (34.29 Kb and 129.18 Kb, respectively), located immediately adjacent to genomic gaps. Additionally, both gene models share almost identical 3'-UTR sequences (98.31% sequence identity). Together, these data suggest that this observed duplication of *TOLL8* in *C. quinquefaciatus* is artificial and these two genes represent haplotypes of the same gene.

TOLL9 Phylogeny

Previous studies have shown that *Ae. aegypti* and *C. quinquefasciatus* possess two copies of *TOLL9*, termed *TOLL9A* and *9B*, with this duplication absent from the anophelines (Arensburger et al., 2010; Bartholomay et al., 2010; Waterhouse et al., 2008, 2007). However, in our annotations, we identified additional TOLL9 duplications in *An. albimanus* (AALB007527) and *An. darlingi* (ADAC000052). Maximum likelihood analysis of the TOLL9 subfamily revealed that these duplications in *Neomyzomya* species *An. albimanus* and *An. darlingi* cluster with the previously described TOLL9B genes AAEL011734 and CPIJ006150. This clustering is strongly supported by a bootstrap value of 100. Likewise, the additional copies of TOLL9 paralogs AALB005549 and ADAC008087 cluster with strong bootstrap support (100) with the existing TOLL9 anopheline sequences as well as the TOLL9 culicine paralogs. Within the TOLL9A subclade, the Nyssorhynchus subseries is placed as a sister group to the Anopheles subseries, but this grouping is poorly supported (bootstrap value of 65), making determination of true TOLL9A topology difficult in this region. Our analyses revealed that a *TOLL9* duplication event extends into the anopheline *Nyssorhynchus* subgenus and suggests an ancient duplication occurred after

the separation of Drosophilidae and Culicidae. This duplication event was following by a subsequent gene loss after the divergence of *Nyssorhynchus* from other anopheline species (Figure 3-5).

The amino acid sequence of these two *Nyssorhynchus* TOLL9B proteins, AALB007527 and ADAC000052, are divergent from the anopheline *TOLL9A* gene models with an average percent sequence identity of 36.33% and 40.97%, respectively. In comparison, the TOLL9A paralogs AALB005549 and ADAC008087 share 72.38% and 71.71% sequence identity with other anopheline TOLL9A sequences. Additionally, protein domain predictions of the TOLL9B paralogs revealed differences in their extracellular LRR locations when compared to those of the TOLL9A domain architectures (Figures B-29). However, phylogenetic analyses of the TIR domain of these genes (Figure 3-3) cluster these duplications with high confidence with established mosquito *TOLL9* orthologs, establishing that TOLL9A and 9B proteins belong to the same TLR subfamily and arose by duplication of the ancestral *TOLL9* gene.

TOLL1/5 Phylogeny

To better elucidate the phylogenetic histories of the TLRs that clustered within the TOLL1/5 clade based on TIR sequence comparison (Figure 3-2), we performed an additional maximum analysis focused on this clade utilizing their full amino acid sequences. Focusing our analysis increased the number of informative sites from 66 to 752 (Supplemental Table 3) and further emphasized that TOLL1/5 anopheline TLRs do not cluster into subclades corresponding to TOLL1A, TOLL1B, TOLL5A, and TOLL5B *An. gambiae* TLRs. Indeed, these anopheline TLRs cluster together into one large clade, supported by a bootstrap value of 94 (Figure 3-6). Within this clade, 1:1 orthologs to *An. gambiae* TOLL1A formed a distinct subclade (bootstrap value

100), and all other anopheline sequences clustered into a single secondary subclade (bootstrap value 74) (Figure 3-3, Figure 3-6). Within each subclade, gene topology matched species tree topologies for species belonging to the Neocellia, Myzomyia, Neomyzymia series as well as the Anopheles and Nyssorhynchus subgenera. However, outside of the TOLL1A subclade, we observed repeated duplications of TLRs within the *gambiae* complex (Figure 3-6). Based on species tree topology, we can infer five independent gene duplications have given rise to six TOLL1/5 paralogs in *An. gambiae* (Figure 3-6).

Within *An. gambiae*, the six gene models are located on two chromosomes, X and 3L (Figure 3-7). All six gene models have similar gene structure, with three exons and two introns of similar length (Figure 3-7). The four duplications of TLRs on the X chromosome in *An. gambiae* are located near each other (within 25 thousand bp) and are oriented in the same direction, while the two coding sequences on the 3L chromosome are separated by 404 thousand bp and lie in opposing directions (Figure 3-7). Given the phylogenic analyses of these sequences (Figure 3-6), coupled with their genomic locations (Figure 3-7), it is likely that the TOLLX and TOLLY genes described in this study arose through two separate duplication events that led to the genes TOLL5A, TOLLX, and TOLLY tandem on chromosome X.

Discussion

Molecular mechanisms of mosquito immune responses like the Toll pathway are important for our understanding of vector biology, including aspects of vector control and disease transmission. In this study, we identified and manually annotated the coding sequences of intracellular Toll pathway members and TLRs to identify and characterize the complete immune repertoire of Toll signaling within 21 mosquito species.

Here we show, through a combination of manual annotation and maximum likelihood phylogenetic analysis, that the intracellular Toll signaling cascade is maintained throughout the anophelines, with 1:1 orthology found for all members. While these pathway members are conserved in gene number, we find that the evolutionary distances between anopheline species varies between pathway members. There does appear to be a higher rate of amino acid substitutions within pathway members that are central to Toll pathway signaling, including the PELLE, MYD88, and TUBE adaptor proteins, indicating that the coding sequences of these pathway members are diversifying within the anophelines. This is very similar to what can be observed in related evolutionary studies, where immune modulators evolve more quickly (Neafsey et al., 2015; Sackton et al., 2007).

Interestingly, we found differences in the phylogenetic distances between the two splice isoforms of REL1 in our analyses. REL1 is alternatively spliced to create a short (REL1-A) and a long (REL1-B) mRNA transcript. Phylogenetic analysis of the amino acid sequences of these genes revealed that REL1-A displayed more divergent sequences between anophelines (mean normalized phylogenetic distance = 0.99 substitutions/site) compared to that of REL1-B (mean normalized phylogenetic distance = 0.48 substitutions/site). Previously published work analyzing the function of these splice isoforms in *Ae. aegypti* stated that REL1-B does not display binding

affinity for κB motifs, indicating that it may not serve as a transcription factor (Shin et al., 2005). However, the same study also observed that REL1-B works cooperatively with REL1-A to initiate a higher level of transcription of immune genes (Shin et al., 2005) This is similar to results obtained in *D. melanogaster* for the developmental Toll pathway NF-κB transcription factors Dorsal-A and Dorsal-B (Gross et al., 1999). The higher rates of amino acid substitution that we observed in REL1-A compared to REL1-B may be reflective of the evolutionary pressures placed on the transcriptional regulator, REL1-A that are not shared by the cooperative activator, REL1-B.

Additionally, we performed a manual annotation of TLR genes encoded within 20 sequenced mosquito genomes and found that the overall repertoire of TLR genes was consistent with previous descriptions in *Ae. aegypti* and *An. gambiae* (Waterhouse et al., 2007). 1:1 orthology was observed within the orthologous TLR groups TOLL6, TOLL 7, TOLL 8, TOLL 10, and TOLL 11 by phylogenetic analysis and each TLR subfamily possessed unique ectodomain architecture, providing support that these TLRs are conserved in mosquitoes belonging to both culicine and anopheline genera in terms amino acid sequence and protein architecture. In *D. melanogaster*, these TLRs have been implicated in convergent extension of developing embryos (TOLL-6/TOLL-8) (Paré et al., 2014), regulation of autophagy and recognition of viral infections (TOLL-7) (Nakamoto et al., 2012), and negative regulation of antimicrobial responses (TOLL-8) (Akhouayri et al., 2011). However, whether these biological functions of these genes are conserved in mosquitoes remains to be experimentally confirmed. Additionally, TOLL10 and TOLL11 are absent in *D. melanogaster* and have been functionally described in other insect species, including the mosquito species included in this study.

However, the conservation observed in TOLL6-TOLL11 does not mean that no duplication and diversification was observed in our analysis of these receptor subfamilies. We observed duplication of *TOLL9* within *An. albimanus* and *An. darlingi* in our analysis. These duplicated TOLL9 amino acid coding sequences cluster with other mosquito TOLL9 sequenced by phylogenetic analysis of TIR domains, and cluster more closely to culicine TOLL9B genes than the single-copy TOLL9 representatives found in other anophelines. This topology provides evidence for a TOLL9 duplication occurrence prior to the culicine/anopheline speciation event. Following this event, it is likely that TOLL9B orthologs were loss prior to the split of *Nyssorhynchus* from other anophelines. These duplicated *TOLL9* genes within *An. albimanus* and *An. darlingi* could possibly have functions separate to their paralogs within these species. This is largely evidenced by the difference in sequence identity of these duplications, AALB007527 and ADAC000052, as these genes possess an average percent sequence identity of 36.33% and 40.97%, respectively, compared to other anopheline TOLL9 coding sequences. These percent identities are over 30% more divergent in sequence than their paralogs AALB005549 and ADAC008087 with other anopheline TOLL9 sequences. *D. melanogaster* TOLL-9 has been linked to immunity (Ooi et al., 2002) and gene duplication of immune genes can lead to novel ligand specificities or function (Conant and Wolfe, 2008; Hughes, 1994; Zhang, 2003), making these novel TLRs intriguing candidates for study of novel TLR binding affinity within mosquitoes.

Lastly, our data show that *TOLL1/5* genes in the *gambiae* complex lineage have experienced repeated gene duplications, leading to an expansion of *TOLL1/5* genes. Our comparative approach allowed us to characterize these duplications and phylogenetic analysis of the manually annotated sequences reveal that these genes encode complete TLRs, with

extracellular repeated LRR domains, a single transmembrane domain, and a TIR domain. Interestingly, within the TOLL1/5 clade, there is a division between genes orthologous to TOLL1A and all other TOLL1/5 coding sequences. As this TOLL1/5 clade contains members from *D. melanogaster* (*Toll* and *Toll-5*) and *Ae. aegypti* (*TOLL5A*) that serve in development (Anderson et al., 1985) and immunity (Lemaitre et al., 1996; Shin et al., 2006), it is tempting to speculate that these expansions may impact development and immune responses of vector mosquito species such as *An. gambiae*, *An. arabiensis*, *An. merus*, and *An. melas*.

The influences that drive the evolution and diversification of TLRs within insects remain largely unknown. This is, at least in part, due to the fact that there is still little understanding on the functional role these genes play in insect biology. Even within *D. melanogaster*, knockdown or overexpression of many TLRs does not lead to visible phenotypes in survivorship, morphology, or expression of antimicrobials (Nakamoto et al., 2012; Ooi et al., 2002; Samaraweera et al., 2013; Yagi et al., 2010). This may be due to the heterodimerization of these receptors leading to an array of possible receptor combinations. Within TLRs, this is not without precedent, as heterodimerization of TLRs within humans can have drastic effects on Toll signaling outcomes (De Nardo, 2015). However, a thorough understanding of the complete repertoire of this gene family will aid future studies of TLR function within these mosquito vectors and nonvectors by improving on our understanding on the possible heterodimeric combinations. In summary, this study provides a thorough description of the complete repertoire of TLRs and intracellular Toll pathway members for all anophelines sequenced to date. This research provides a much needed description of the phylogenetic relationships and conservation of a signaling pathway that is not only diverse in sequence but diverse in function. We show that the intracellular Toll signaling cascade is conserved, with 1:1 orthologs found in all species

included in this study. In addition, we have provided a complete annotation of the TLR family in these anophelines, enabling future studies on their biological functions.

Acknowledgements

This work is supported by the National Institutes of Health grant number R01AI095842 and through USDA-ARS Specific Cooperative Agreement 58-5430-4-022 (both to K.M.). Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the funding agencies.

Authors' Contributions

VLMR and KM conceived and designed the study; KM provided funding acquisition; VLMR wrote the initial draft of the manuscript, and VLMR and KM wrote and edited the final manuscript version. Both authors read and approved the final manuscript.

References

Akhouayri, I., Turc, C., Royet, J., Charroux, B., 2011. Toll-8/tollo negatively regulates antimicrobial response in the *Drosophila* respiratory epithelium. PLoS Pathog. 7, e1002319.

Anderson, K. V., Bokla, L., Nusslein-Volhard, C., 1985. Establishment of dorsal-ventral polarity in the *Drosophila* embryo: The induction of polarity by the *Toll* gene product. Cell 42, 791–798.

Anderson, K. V., Nüsslein-Volhard, C., 1984. Information for the dorsal-ventral pattern of the *Drosophila* embryo is stored as maternal mRNA. Nature 311, 223–227.

Arensburger, P., Megy, K., Waterhouse, R. M., Abrudan, J., Amedeo, P., Antelo, B., Bartholomay, L., Bidwell, S., Caler, E., Camara, F., Campbell, C. L., Campbell, K. S., Casola, C., Castro, M. T., Chandramouliswaran, I., Chapman, S. B., Christley, S., Costas, J., Eisenstadt, E., Feschotte, C., Fraser-Liggett, C., Guigo, R., Haas, B., Hammond, M., Hansson, B. S., Hemingway, J., Hill, S. R., Howarth, C., Ignell, R., Kennedy, R. C., Kodira, C. D., Lobo, N. F., Mao, C., Mayhew, G., Michel, K., Mori, A., Liu, N., Naveira, H., Nene, V., NamNguyen, Pearson, M. D., Pritham, E. J., Puiu, D., Qi, Y., Ranson, H., Ribeiro, J. M. C., Roberston, H. M., Severson, D., Shumway, M., Stanke, M., Strausberg, R. L., Sun, C., Sutton, G., Tu, Z. (Jake), Tubio, J. M. C., Unger, M. F., Vanlandingham, D. L., Vilella, A. J., White, O., White, J. R., Wondji, C. S., Wortman, J., Zdobnov, E. M., Birren, B., Christensen, B. M., Collins, F. H., Cornel, A., Dimopoulos, G., Hannick, L. I., Higgs, S., Lanzaro, G. C., Lawson, D., Lee, N. H., Muskavitch, M. A. T., Raikhel, A. S., Atkinson, P. W., 2010. Sequencing of *Culex quinquefasciatus* establishes a platform for mosquito comparative genomics. Science 330, 86–88.

Babicki, S., Arndt, D., Marcu, A., Liang, Y., Grant, J. R., Maciejewski, A., Wishart, D. S., 2016. Heatmapper: web-enabled heat mapping for all. Nucleic Acids Res. 44, W147–W153.

Barillas-Mury, C., Charlesworth, A., Gross, I., Richman, A., Hoffmann, J. A., Kafatos, F. C., 1996. Immune factor Gambif1, a new rel family member from the human malaria

vector, *Anopheles gambiae*. EMBO J. 15, 4691–4701.

Barreaux, P., Barreaux, A. M. G., Sternberg, E. D., Suh, E., Waite, J. L., Whitehead, S. A., Thomas, M. B., 2017. Priorities for broadening the malaria vector control tool kit. Trends Parasitol. 33, 763–774.

Bartholomay, L. C., Waterhouse, R. M., Mayhew, G. F., Campbell, C. L., Michel, K., Zou, Z., Ramirez, J. L., Das, S., Alvarez, K. S., Arensburger, P., Bryant, B., Chapman, S. B., Dong, Y., Erickson, S. M., Karunaratne, S. H., Kokoza, V., Kodira, C. D., Pignatelli, P., Shin, S. W., Vanlandingham, D. L., Atkinson, P. W., Birren, B., Christophides, G. K., Clem, R. J., Hemingway, J., Higgs, S., Megy, K., Ranson, H., Zdobnov, E. M., Raikhel, A. S., Christensen, B. M., Dimopoulos, G., Muskavitch, M. A. T., 2010. Pathogenomics of *Culex quinquefasciatus* and meta-analysis of infection responses to diverse pathogens. Science 330, 88–90.

Bateman, A., Birney, E., Cerruti, L., Durbin, R., Etwiller, L., Eddy, S. R., Griffiths-Jones, S., Howe, K. L., Marshall, M., Sonnhammer, E. L. L., 2002. The Pfam protein families database. Nucleic Acids Res. 30, 276–280.

Belvin, M. P., Anderson, K. V., 1996. A conserved signaling pathway: the *Drosophila* toll-dorsal pathway. Annu. Rev. Cell Dev. Biol. 12, 393–416.

Benelli, G., Beier, J. C., 2017. Current vector control challenges in the fight against malaria. Acta Trop. 174, 91–96.

Bian, G., Shin, S. W., Cheon, H. M., Kokoza, V., Raikhel, A. S., 2005. Transgenic alteration of Toll immune pathway in the female mosquito *Aedes aegypti*. Proc. Natl. Acad. Sci. U. S. A. 102, 13568–13573.

Bischoff, V., Vignal, C., Boneca, I. G., Michel, T., Hoffmann, J. A., Royet, J., 2004. Function of the *Drosophila* pattern-recognition receptor PGRP-SD in the detection of Gram-positive bacteria. Nat. Immunol. 5, 1175–1180.

Blanford, S., Chan, B. H. K., Jenkins, N., Sim, D., Turner, R. J., Read, A. F., Thomas, M. B., 2005. Fungal pathogen reduces potential for malaria transmission. Science 308, 1638–1641.

Boëte, C., Paul, R. E. L., Koella, J. C., 2004. Direct and indirect immunosuppression by a malaria parasite in its mosquito vector. Proc. R. Soc. L. 271, 1611–1615.

Brownlie, R., Allan, B., 2011. Avian toll-like receptors. Cell Tissue Res. 343, 121–130.

Buckley, K. M., Rast, J. P., 2012. Dynamic evolution of toll-like receptor multigene families in echinoderms. Front. Immunol. 3.

Cao, X., He, Y., Hu, Y., Wang, Y., Chen, Y. R., Bryant, B., Clem, R. J., Schwartz, L. M., Blissard, G., Jiang, H., 2015. The immune signaling pathways of *Manduca sexta*. Insect Biochem. Mol. Biol. 62, 64–74.

Casanova, J. L., Abel, L., Quintana-Murci, L., 2011. Human TLRs and IL-1Rs in Host Defense: Natural Insights from Evolutionary, Epidemiological, and Clinical Genetics. Annu. Rev. Immunol 29, 447–91.

Cha, G., Cho, K. S., Lee, J. H., Kim, E., Park, J., Lee, S. B., Kim, M., Chung, J., 2003. Discrete functions of TRAF1 and TRAF2 in *Drosophila melanogaster* mediated by c-Jun N-terminal kinase and NF-κB-dependent signaling pathways. Mol. Cell. Biol. 23, 7982–7991.

Chen, X. G., Jiang, X., Gu, J., Xu, M., Wu, Y., Deng, Y., Zhang, C., Bonizzoni, M., Dermauw, W., Vontas, J., Armbruster, P., Huang, X., Yang, Y., Zhang, H., He, W., Peng, H., Liu, Y., Wu, K., Chen, J., Lirakis, M., Topalis, P., Van Leeuwen, T., Hall, A. B., Jiang, X., Thorpe, C., Mueller, R. L., Sun, C., Waterhouse, R. M., Yan, G., Tu, Z. J., Fang, X., James, A. A., 2015. Genome sequence of the Asian Tiger mosquito, *Aedes albopictus*, reveals insights into its biology, genetics, and evolution. Proc. Natl. Acad. Sci. U. S. A. 112, E5907-5915.

Christophides, G. K., Zdobnov, E., Barillas-Mury, C., Birney, E., Blandin, S., Blass, C., Brey, P. T., Collins, F. H., Danielli, A., Dimopoulos, G., Hetru, C., Hoa, N. T., Hoffmann, J. A., Kanzok, S. M., Letunic, I., Levashina, E. A., Loukeris, T. G., Lycett, G., Meister, S., Michel, K., Moita, L. F., Müller, H. M., Osta, M. A., Paskewitz, S. M., Reichhart, J. M., Rzhetsky, A., Troxler, L., Vernick, K. D., Vlachou, D., Volz, J., von Mering, C., Xu, J., Zheng, L., Bork, P., Kafatos, F. C.,

2002. Immunity-related genes and gene families in *Anopheles gambiae*. Science 298, 159–165.

Coetzee, M., Hunt, R. H., Wilkerson, R., Torre, A. Della, Coulibaly, M. B., Besansky, N. J., 2013. *Anopheles coluzzii* and *Anopheles amharicus*, new members of the *Anopheles gambiae* complex. Zootaxa 3619, 246–274.

Conant, G. C., Wolfe, K. H., 2008. Turning a hobby into a job: How duplicated genes find new functions. Nat. Rev. Genet. 9, 938.

De Nardo, D., 2015. Toll-like receptors: Activation, signalling and transcriptional modulation. Cytokine 74, 181–189.

Edgar, R. C., 2004. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1792–1797.

Eldon, E., Kooyer, S., D'Evelyn, D., Duman, M., Lawinger, P., Botas, J., Bellen, H. (1994). The *Drosophila* 18 wheeler is required for morphogenesis and has striking similarities to Toll. Development, 120, 885-899.

El Chamy, L., Leclerc, V., Caldelari, I., Reichhart, J. M., 2008. Sensing of "danger signals" and pathogen-associated molecular patterns defines binary signaling pathways "upstream" of Toll. Nat. Immunol. 9, 1165–1170.

Evans, J. D., Aronstein, K., Chen, Y. P., Hetru, C., Imler, J. L., Jiang, H., Kanost, M. R., Thompson, G. J., Zou, Z., Hultmark, D., 2006. Immune pathways and defence mechanisms in honey bees *Apis mellifera*. Insect Mol. Biol. 15, 645–656.

Frolet, C., Thoma, M., Blandin, S., Hoffmann, J. A., Levashina, E. A., 2006. Boosting NF-κB-dependent basal immunity of *Anopheles gambiae* aborts development of *Plasmodium berghei*. Immunity 25, 677–685.

Gangloff, M., Murali, A., Xiong, J., Arnot, C. J., Weber, A. N., Sandercock, A. M., Robinson, C. V., Sarisky, R., Holzenburg, A., Kao, C., Gay, N. J., 2008. Structural insight into the mechanism of activation of the Toll receptor by the dimeric ligand Spätzle. J. Biol. Chem. 283, 14629–14635.

Garver, L. S., Dong, Y., Dimopoulos, G., 2009. Caspar controls resistance to *Plasmodium falciparum* in diverse anopheline species. PLoS Pathog. 5, e1000335.

Gay, N. J., Keith, F. J., 1991. *Drosophila* Toll and IL-1 receptor. Nature 351, 355–356.

Gerttula, S., Jin, Y. S., Anderson, K. V., 1988. Zygotic expression and activity of the *Drosophila Toll* gene, a gene required maternally for embryonic dorsal-ventral pattern formation. Genetics 119, 123–133.

Giraldo-Calderon, G. I., Emrich, S. J., MacCallum, R. M., Maslen, G., Emrich, S. J., Collins, F. H., Dialynas, E., Topalis, P., Ho, N., Gesing, S., Madey, G., Collins, F. H., Lawson, D., Kersey, P., Allen, J., Christensen, M., Hughes, D., Koscielny, G., Langridge, N., Gallego, E. L., Megy, K., Wilson, D., Gelbart, B., Emmert, D., Russo, S., Zhou, P., Christophides, G., Brockman, A., Kirmitzoglou, I., MacCallum, B., Tiirikka, T., Louis, K., Dritsou, V., Mitraka, E., Werner-Washburn, M., Baker, P., Platero, H., Aguilar, A., Bogol, S., Campbell, D., Carmichael, R., Cieslak, D., Davis, G., Konopinski, N., Nabrzyski, J., Reinking, C., Sheehan, A., Szakonyi, S., Wieck, R., Giraldo-Calde On, G. I., Emrich, S. J., MacCallum, R. M., Maslen, G., Dialynas, E., Topalis, P., Ho, N., Gesing, S., Consortium, V., Madey, G., Collins, F. H., Lawson, D., Giraldo-Calderon, G. I., Emrich, S. J., MacCallum, R. M., Maslen, G., Dialynas, E., Topalis, P., Ho, N., Gesing, S., Madey, G., Collins, F. H., Lawson, D., Lawson, D., 2015. VectorBase: an updated bioinformatics resource for invertebrate vectors and other organisms related with human diseases. Nucleic Acids Res. 43, D707–D713.

Gobert, V., Gottar, M., Matskevich, A. A., Rutschmann, S., Royet, J., Belvin, M. P., Hoffmann, J. A., Ferrandon, D., 2003. Dual activation of the *Drosophila* Toll pathway by two pattern recognition receptors. Science 302, 2126–2130.

Goldman, N., 1998. Phylogenetic information and experimental design in molecular systematics. Proc Biol Sci 265, 1779–1786.

Gottar, M., Gobert, V., Matskevich, A. A., Reichhart, J. M., Wang, C., Butt, T. M., Belvin, M., Hoffmann, J. A., Ferrandon, D., 2006. Dual detection of fungal infections in *Drosophila* via recognition of glucans and sensing of virulence factors. Cell 127, 1425–

1437.

Gross, I., Georgel, P., Oertel-Buchheit, P., Schnarr, M., Reichhart, J. M., 1999. Dorsal-B, a splice variant of the *Drosophila* factor Dorsal, is a novel REL/NF-κB transcriptional activator. Gene 228, 233–242.

Grosshans, J., Bergmann, A., Haffter, P., Nüsslein-Volhard, C., 1994. Activation of the kinase Pelle by Tube in the dorsoventral signal transduction pathway of *Drosophila* embryo. Nature.

Haghayeghi, A., Sarac, A., Czerniecki, S., Grosshans, J., Schöck, F., 2010. Pellino enhances innate immunity in *Drosophila*. Mech. Dev. 127, 301–307.

Harris, C., Lambrechts, L., Rousset, F., Abate, L., Nsango, S. E., Fontenille, D., Morlais, I., Cohuet, A., 2010. Polymorphisms in *Anopheles gambiae* immune genes associated with natural resistance to *Plasmodium falciparum*. PLoS Pathog. 6, e1001112.

Hashimoto, C., Hudson, K. L., Anderson, K. V., 1988. The *Toll* gene of *Drosophila*, required for dorsal-ventral embryonic polarity, appears to encode a transmembrane protein. Cell 52, 269–279.

Hemingway, J., Ranson, H., 2000. Insecticide resistance in insect vectors of human disease. Annu. Rev. Entomol. 45, 371–391.

Hemingway, J., Shretta, R., Wells, T. N. C., Bell, D., Djimdé, A. A., Achee, N., Qi, G., 2016. Tools and Strategies for Malaria Control and Elimination: What Do We Need to Achieve a Grand Convergence in Malaria? PLoS Biol. 14, e1002380.

Holt, R. A., Subramanian, G. M., Halpern, A., Sutton, G. G., Charlab, R., Nusskern, D. R., Wincker, P., Clark, A. G., Ribeiro, J. M. C., Wides, R., Salzberg, S. L., Loftus, B., Yandell, M., Majoros, W. H., Rusch, D. B., Lai, Z., Kraft, C. L., Abril, J. F., Anthouard, V., Arensburger, P., Atkinson, P. W., Baden, H., Berardinis, V. de, Baldwin, D., Benes, V., Biedler, J., Blass, C., Bolanos, R., Boscus, D., Barnstead, M., Cai, S., Center, A., Chatuverdi, K., Christophides, G. K., Chrystal, M. A., Clamp, M., Cravchik, A., Curwen, V., Dana, A., Delcher, A., Dew, I., Evans, C. A., Flanigan, M., Grundschober-Freimoser, A., Friedli, L., Gu, Z., Guan, P., Guigo, R.,

Hillenmeyer, M. E., Hladun, S. L., Hogan, J. R., Hong, Y. S., Hoover, J., Jaillon, O., Ke, Z., Kodira, C., Kokoza, E., Koutsos, A., Letunic, I., Levitsky, A., Liang, Y., Lin, J. J., Lobo, N. F., Lopez, J. R., Malek, J. A., McIntosh, T. C., Meister, S., Miller, J., Mobarry, C., Mongin, E., Murphy, S. D., O'Brochta, D. A., Pfannkoch, C., Qi, R., Regier, M. A., Remington, K., Shao, H., Sharakhova, M. V., Sitter, C. D., Shetty, J., Smith, T. J., Strong, R., Sun, J., Thomasova, D., Ton, L. Q., Topalis, P., Tu, Z., Unger, M. F., Walenz, B., Wang, A., Wang, J., Wang, M., Wang, X., Woodford, K. J., Wortman, J. R., Wu, M., Yao, A., Zdobnov, E. M., Zhang, H., Zhao, Q., Zhao, S., Zhu, S. C., Zhimulev, I., Coluzzi, M., Torre, A. della, Roth, C. W., Louis, C., Kalush, F., Mural, R. J., Myers, E. W., Adams, M. D., Smith, H. O., Broder, S., Gardner, M. J., Fraser, C. M., Birney, E., Bork, P., Brey, P. T., Venter, J. C., Weissenbach, J., Kafatos, F. C., Collins, F. H., Hoffman, S. L., 2002. The genome sequence of the malaria mosquito *Anopheles gambiae*. Science 298, 129–149.

Huang, H. R., Chen, Z. J., Kunes, S., Chang, G. D., Maniatis, T., 2010. Endocytic pathway is required for *Drosophila* Toll innate immune signaling. Proc. Natl. Acad. Sci. U. S. A. 107, 8322–8327.

Hughes, A. L., 1994. The Evolution of Functionally Novel Proteins after Gene Duplication. Proc. Biol. Sci. 256, 119–124.

Imler, J. L., Zheng, L., 2004. Biology of Toll receptors: lessons from insects and mammals. J. Leukoc. Biol. 75, 18–26.

Issa, N., Guillaumot, N., Lauret, E., Matt, N., Schaeffer-Reiss, C., Van Dorsselaer, A., Reichhart, J. M., Veillard, F., 2018. The Circulating Protease Persephone is an Immune Sensor for Micobial Proteolytic Activities Upstream of the *Drosophila* Toll Pathway. Mol. Cell 69, 539–550.

Ji, S., Sun, M., Zheng, X., Li, L., Sun, L., Chen, D., Sun, Q., 2014. Cell-surface localization of Pellino antagonizes Toll-mediated innate immune signalling by controlling MyD88 turnover in *Drosophila*. Nat. Commun. 5.

Joop, G., Vilcinskas, A., 2016. Coevolution of parasitic fungi and insect hosts. Zoology 119, 350–358.

Kamareddine, L., 2012. The biological control of the malaria vector. Toxins (Basel). 4, 748–767.

Kawai, T., Akira, S., 2006. TLR signaling. Cell Death Differ. 13, 816–825.

Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 33, 1870–1874.

Kuttenkeuler, D., Pelte, N., Ragab, A., Gesellchen, V., Schneider, L., Blass, C., Axelsson, E., Huber, W., Boutros, M., 2010. A large-scale RNAi screen identifies *Deaf1* as a regulator of innate immune responses in *Drosophila*. J. Innate Immun. 2, 181–194.

Lambrechts, L., Morlais, I., Awono-Ambene, P. H., Cohuet, A., Simard, F., Jacques, J. C., Bourgouin, C., Koella, J. C., 2007. Effect of infection by *Plasmodium falciparum* on the melanization immune response of *Anopheles gambiae*. Am. J. Trop. Med. Hyg. 76, 475–480.

Lawniczak, M. K. N., Emrich, S., Holloway, A. K., Regier, A. P., Olson, M., White, B., Redmond, S., Fulton, L., Appelbaum, E., Godfrey, J., Farmer, C., Chinwalla, A., Yang, S. P., Minx, P., Nelson, J., Kyung, K., Walenz, B. P., Garcia-Hernandez, E., Aguiar, M., Viswanathan, L. D., Rogers, Y. H., Strausberg, R. L., Saski, C. A., Lawson, D., Collins, F. H., Kafatos, F. C., Christophides, G. K., Clifton, S. W., Kirkness, E. F., Besansky, N. J., 2010. Widespread divergence between incipient *Anopheles gambiae* species revealed by whole genome sequences. Science 330, 512–514.

Lee, E., Helt, G. A., Reese, J. T., Munoz-Torres, M. C., Childers, C. P., Buels, R. M., Stein, L., Holmes, I. H., Elsik, C. G., Lewis, S. E., 2013. Web Apollo: a web-based genomic annotation editing platform. Genome Biol. 14, R93.

Lemaitre, B., Meister, M., Govind, S., Georgel, P., Steward, R., Reichhart, J. M., Hoffmann, J. A., 1995. Functional analysis and regulation of nuclear import of dorsal during the immune response in *Drosophila*. EMBO J. 14, 536–545.

Lemaitre, B., Nicolas, E., Michaut, L., Reichhart, J. M., Hoffmann, J. A., 1996. The dorsoventral regulatory gene cassette *spätzle/Toll/cactus* controls the potent antifungal response in *Drosophila* adults. Cell 86, 973–983.

Leulier, F., Lemaitre, B., 2008. Toll-like receptors - taking an evolutionary approach. Nat. Rev. Genet. 9, 165–178.

Levin, T. C., Malik, H. S., 2017. Rapidly evolving *Toll-3/4* genes encode male-specific Toll-like receptors in *Drosophila*. Oxford Univ. Press.

Luna, C., Hoa, N. T., Lin, H., Zhang, L., Nguyen, H. L. A., Kanzok, S. M., Zheng, L., 2006. Expression of immune responsive genes in cell lines from two different anopheline species. Insect Mol. Biol. 15, 721–729.

Luna, C., Hoa, N. T., Zhang, J., Kanzok, S. M., Brown, S. E., Imler, J. L., Knudson, D. L., Zheng, L., 2003. Characterization of three Toll-like genes from mosquito *Aedes aegypti*. Insect Mol. Biol. 12, 67–74.

Luna, C., Wang, X., Huang, Y., Zhang, J., Zheng, L., 2002. Characterization of four Toll related genes during development and immune responses in *Anopheles gambiae*. Insect Biochem. Mol. Biol. 32, 1171–1179.

Luo, C., Shen, B., Manley, J. L., Zheng, L., 2001. Tehao functions in the Toll pathway in *Drosophila melanogaster*: Possible roles in development and innate immunity. Insect Mol. Biol. 10, 457–464.

MacCallum, R. M., Redmond, S. N., Christophides, G. K., 2011. An expression map for *Anopheles gambiae*. BMC Genomics 12, 620.

Marinotti, O., Cerqueira, G. C., De Almeida, L. G. P., Ferro, M. I. T., Da Silva Loreto, E. L., Zaha, A., Teixeira, S. M. R., Wespiser, A. R., E Silva, A. A., Schlindwein, A. D., Pacheco, A. C. L., Da Costa Da Silva, A. L., Graveley, B. R., Walenz, B. P., De Araujo Lima, B., Ribeiro, C. A. G., Nunes-Silva, C. G., De Carvalho, C. R., De Almeida Soares, C. M., De Menezes, C. B. A., Matiolli, C., Caffrey, D., Araújo, D. A. M., De Oliveira, D. M., Golenbock, D., Grisard, E. C., Fantinatti-Garboggini, F., De Carvalho, F. M., Barcellos, F. G., Prosdocimi, F., May, G., De Azevedo Junior, G. M., Guimarães, G. M., Goldman, G. H., Padilha, I. Q. M., Da Silva Batista, J., Ferro, J. A., Ribeiro, J. M. C., Fietto, J. L. R., Dabbas, K. M., Cerdeira, L., Agnez-Lima, L. F., Brocchi, M., De Carvalho, M. O., De Melo Teixeira, M., De Mascena

Diniz Maia, M., Goldman, M. H. S., Schneider, M. P. C., Felipe, M. S. S., Hungria, M., Nicolás, M. F., Pereira, M., Montes, M. A., Cantão, M. E., Vincentz, M., Rafael, M. S., Silverman, N., Stoco, P. H., Souza, R. C., Vicentini, R., Gazzinelli, R. T., De Oliveira Neves, R., Silva, R., Astolfi-Filho, S., Maciel, T. E. F., Ürményi, T. P., Tadei, W. P., Camargo, E. P., De Vasconcelos, A. T. R., 2013. The genome of *Anopheles darlingi*, the main neotropical malaria vector. Nucleic Acids Res. 41, 7387–7400.

Mcilroy, G., Foldi, I., Aurikko, J., Wentzell, J. S., Lim, M. A., Fenton, C., Gay, N. J., Hidalgo, A., 2013. Toll-6 and Toll-7 function as neurotrophin receptors in the *Drosophila* central nervous system. Nat. Neurosci. 16, 1248–1256.

Michel, K., Budd, A., Pinto, S., Gibson, T. J., Kafatos, F. C., 2005. *Anopheles gambiae* SRPN2 facilitates midgut invasion by the malaria parasite *Plasmodium berghei*. EMBO Rep. 6, 891–897.

Michel, T., Reichhart, J. M., Hoffmann, J. A., Royet, J., 2001. *Drosophila* Toll is activated by Gram-positive bacteria through a circulating peptidoglycan recognition protein. Nature 414, 756–759.

Ming, M., Obata, F., Kuranaga, E., Miura, M., 2014. Persephone/Spätzle pathogen sensors mediate the activation of toll receptor signaling in response to endogenous danger signals in apoptosis-deficient *Drosophila*. J. Biol. Chem. 289, 7558–7568.

Mitri, C., Jacques, J. C., Thiery, I., Riehle, M. M., Xu, J., Bischoff, E., Morlais, I., Nsango, S. E., Vernick, K. D., Bourgouin, C., 2009. Fine pathogen discrimination within the APL1 gene family protects *Anopheles gambiae* against human and rodent malaria species. PLoS Pathog. 5, e1000576.

Molina-Cruz, A., Canepa, G. E., Kamath, N., Pavlovic, N. V, Mu, J., Ramphul, U. N., Ramirez, J. L., Barillas-Mury, C., 2015. *Plasmodium* evasion of mosquito immunity and global malaria transmission: The lock-and-key theory. Proc. Natl. Acad. Sci. U. S. A. 112, 15178–15183.

Moncrieffe, M. C., Grossmann, J. G., Gay, N. J., 2008. Assembly of oligomeric death domain

complexes during Toll receptor signaling. J. Biol. Chem. 283, 33447–33454.

Nakamoto, M., Moy, R. H., Xu, J., Bambina, S., Yasunaga, A., Spencer, S., Gold, B., Cherry, S., 2012. Virus recognition by Toll-7 activates antiviral autophagy in *Drosophila*. Immunity 36, 658–667.

Neafsey, D. E., Waterhouse, R. M., Abai, M. R., Aganezov, S. S., Alekseyev, M. A., Allen, J. E., Amon, J., Arcà, B., Arensburger, P., Artemov, G., Assour, L. A., Basseri, H., Berlin, A., Birren, B. W., Blandin, S. A., Brockman, A. I., Burkot, T. R., Burt, A., Chan, C. S., Chauve, C., Chiu, J. C., Christensen, M., Costantini, C., Davidson, V. L. M., Deligianni, E., Dottorini, T., Dritsou, V., Gabriel, S. B., Guelbeogo, W. M., Hall, A. B., Han, M. V, Hlaing, T., Hughes, D. S. T., Jenkins, A. M., Jiang, X., Jungreis, I., Kakani, E. G., Kamali, M., Kemppainen, P., Kennedy, R. C., Kirmitzoglou, I. K., Koekemoer, L. L., Laban, N., Langridge, N., Lawniczak, M. K. N., Lirakis, M., Lobo, N. F., Lowy, E., MacCallum, R. M., Mao, C., Maslen, G., Mbogo, C., McCarthy, J., Michel, K., Mitchell, S. N., Moore, W., Murphy, K. A., Naumenko, A. N., Nolan, T., Novoa, E. M., O'Loughlin, S., Oringanje, C., Oshaghi, M. A., Pakpour, N., Papathanos, P. A., Peery, A. N., Povelones, M., Prakash, A., Price, D. P., Rajaraman, A., Reimer, L. J., Rinker, D. C., Rokas, A., Russell, T. L., Sagnon, N., Sharakhova, M. V., Shea, T., Simão, F. A., Simard, F., Slotman, M. A., Somboon, P., Stegniy, V., Struchiner, C. J., Thomas, G. W. C., Tojo, M., Topalis, P., Tubio, J. M. C., Unger, M. F., Vontas, J., Walton, C., Wilding, C. S., Willis, J. H., Wu, Y. C., Yan, G., Zdobnov, E. M., Zhou, X., Catteruccia, F., Christophides, G. K., Collins, F. H., Cornman, R. S., Crisanti, A., Donnelly, M. J., Emrich, S. J., Fontaine, M. C., Gelbart, W., Hahn, M. W., Hansen, I. A., Howell, P. I., Kafatos, F. C., Kellis, M., Lawson, D., Louis, C., Luckhart, S., Muskavitch, M. A. T., Ribeiro, J. M., Riehle, M. A., Sharakhov, I. V, Tu, Z., Zwiebel, L. J., Besansky, N. J., 2015. Highly evolvable malaria vectors: The genomes of 16 *Anopheles* mosquitoes. Science 347, 1258522-1-1258522–8.

Nene, V., Wortman, J. R., Lawson, D., Haas, B., Kodira, C., Tu, Z., Loftus, B., Xi, Z., Megy, K., Grabherr, M., Ren, Q., Zdobnov, E. M., Lobo, N. F., Campbell, K. S., Brown, S. E., Bonaldo, M. F., Zhu, J., Sinkins, S. P., Hogenkamp, D. G., Amedeo, P.,

Arensburger, P., Atkinson, P. W., Bidwell, S., Biedler, J., Birney, E., Bruggner, R. V, Costas, J., Coy, M. R., Crabtree, J., Crawford, M., DeBruyn, B., DeCaprio, D., Eiglmeier, K., Eisenstadt, E., El-Dorry, H., Gelbart, W. M., Gomes, S. L., Hammond, M., Hannick, L. I., Hogan, J. R., Holmes, M. H., Jaffe, D., Spencer Johnston, J., Kennedy, R. C., Koo, H., Kravitz, S., Kriventseva, E. V, Kulp, D., LaButti, K., Lee, E., Li, S., Lovin, D. D., Mao, C., Mauceli, E., M Menck, C. F., Miller, J. R., Montgomery, P., Mori, A., Nascimento, A. L., Naveira, H. F., Nusbaum, C., Orvis, J., Pertea, M., Quesneville, H., Reidenbach, K. R., Rogers, Y. H., Roth, C. W., Schneider, J. R., Schatz, M., Shumway, M., Stanke, M., Stinson, E. O., C Tubio, J. M., VanZee, J. P., Verjovski-Almeida, S., Werner, D., White, O., Wyder, S., Zeng, Q., Zhao, Q., Zhao, Y., Hill, C. A., Raikhel, A. S., Soares, M. B., Knudson, D. L., Lee, N. H., Galagan, J., Salzberg, S. L., Paulsen, I. T., Dimopoulos, G., Collins, F. H., Birren, B., Fraser-Liggett, C. M., Severson, D. W., 2007. Genome sequence of *Aedes aegypti*, a major arbovirus vector. Science 316, 1718–1723.

Offord, V., Coffey, T. J., Werling, D., 2010. LRRfinder: A web application for the identification of leucine-rich repeats and an integrative Toll-like receptor database. Dev. Comp. Immunol. 34, 1035–1041.

Ooi, J. Y., Yagi, Y., Hu, X., Ip, Y. T., 2002. The *Drosophila* Toll-9 activates a constitutive antimicrobial defense. EMBO Rep. 3, 82–87.

Palmer, W. J., Jiggins, F. M., 2015. Comparative genomics reveals the origins and diversity of arthropod immune systems. Mol. Biol. Evol. 32, 2111–2129.

Paré, A. C., Vichas, A., Fincher, C. T., Mirman, Z., Farrell, D. L., Mainieri, A., Zallen, J. A., 2014. A positional Toll receptor code directs convergent extension in *Drosophila*. Nature 515, 523–527.

Pendland, J. C., Hung, S. Y., Boucias, D. G., 1993. Evasion of host defense by in vivo-produced protoplast-like cells of the insect mycopathogen *Beauveria bassiana*. J. Bacteriol. 175, 5962–5969.

Ramirez, J. L., Garver, L. S., Brayner, F. A., Alves, L. C., Rodrigues, J., Molina-Cruz, A., Barillas-Mury, C., 2014. The role of hemocytes in *A. gambiae* antiplasmodial immunity.

J. Innate Immun. 6, 119–128.

Ranson, H., Lissenden, N., 2016. Insecticide resistance in African *Anopheles* mosquitoes: A worsening situation that needs urgent action to maintain malaria control. Trends Parasitol. 32, 187–196.

Ranson, H., N'Guessan, R., Lines, J., Moiroux, N., Nkuni, Z., Corbel, V., 2011. Pyrethroid resistance in African anopheline mosquitoes: What are the implications for malaria control? Trends Parasitol. 27, 91–98.

Rebl, A., Goldammer, T., Seyfert, H. M., 2010. Toll-like receptor signaling in bony fish. Vet. Immunol. Immunopathol. 134, 139–150.

Riehle, M. M., Xu, J., Lazzaro, B. P., Rottschaefer, S. M., Coulibaly, B., Sacko, M., Niare, O., Morlais, I., Traore, S. F., Vernick, K. D., 2008. *Anopheles gambiae* APL1 is a family of variable LRR proteins required for Rel1-mediated protection from themalaria parasite, *Plasmodium berghei*. PLoS One 3, e3672.

Roach, J. C., Glusman, G., Rowen, L., Kaur, A., Purcell, M. K., Smith, K. D., Hood, L. E., Aderem, A., 2005. The evolution of vertebrate Toll-like receptors. Proc. Natl. Acad. Sci. 102, 9577–9582.

Sackton, T. B., Lazzaro, B. P., Schlenke, T. A., Evans, J. D., Hultmark, D., Clark, A. G., 2007. Dynamic evolution of the innate immune system in *Drosophila*. Nat. Genet. 39, 1461–1468.

Samaraweera, S. E., O'Keefe, L. V., Price, G. R., Venter, D. J., Richards, R. I., 2013. Distinct roles for *Toll* and autophagy pathways in double-stranded RNA toxicity in a *Drosophila* model of expanded repeat neurodegenerative diseases. Hum. Mol. Genet. 22, 2811–2819.

Scholte, E. J., Ng'habi, K., Kihonda, J., Takken, W., Paaijmans, K., Abdulla, S., Killeen, G. F., Knols, B. G. J., 2005. An entomopathogenic fungus for control of adult African malaria mosquitoes. Science 308, 1641–1642.

Schwenke, R. A., Lazzaro, B. P., Wolfner, M. F., 2016. Reproduction–immunity trade-offs in insects. Annu. Rev. Entomol. 61, 239–256.

Shen, B., Manley, J. L., 2002. Pelle kinase is activated by autophosphorylation during Toll signaling in *Drosophila*. Development 129, 1925–1933.

Shin, S. W., Bian, G. W., Raikhel, **A. S., 2006**. A toll receptor and a cytokine, Toll5A and Spz1C, are involved in toll antifungal immune signaling in the mosquito *Aedes aegypti*. J. Biol. Chem. 281, 39388–39395.

Shin, S. W., Kokoza, V., Bian, G., Cheon, H. M., Yu, J. K., Raikhel, A. S., 2005. REL1, a homologue of *Drosophila* Dorsal, regulates Toll antifungal immune pathway in the female mosquito *Aedes aegypti*. J. Biol. Chem. 280, 16499–16507.

Spencer, E., Jiang, J., Chen, Z. J., 1999. Signal-induced ubiquitination of IκBα by the F-box protein Slimb/β-TrCP. Genes Dev. 13, 284–294.

Thomas, M. B., Godfray, H. C. J., Read, A. F., van den Berg, H., Tabashnik, B. E., van Lenteren, J. C., Waage, J. K., Takken, W., 2012. Lessons from agriculture for the sustainable management of malaria vectors. PLoS Med. 9, e1001262.

Valanne, S., Myllymäki, H., Kallio, J., Schmid, M. R., Kleino, A., Murumägi, A., Airaksinen, L., Kotipelto, T., Kaustio, M., Ulvila, J., Esfahani, S. S., Engström, Y., Silvennoinen, O., Hultmark, D., Parikka, M., Rämet, M., 2010. Genome-wide RNA interference in *Drosophila* cells identifies G protein-coupled receptor kinase 2 as a conserved regulator of NF-κB signaling. J. Immunol. 184, 6188–6198.

Valanne, S., Wang, J. H., Rämet, M., 2011. The *Drosophila* Toll signaling pathway. J. Immunol. 186, 649–656.

Ward, A., Hong, W., Favaloro, V., Luo, L., 2015. Toll receptors instruct axon and dendrite targeting and participate in synaptic partner matching in a *Drosophila* olfactory circuit. Neuron 85, 1013–1028.

Waterhouse, R. M., Kriventseva, E. V., Meister, S., Xi, Z., Alvarez, K. S., Bartholomay, L. C., Barillas-Mury, C., Bian, G., Blandin, S., Christensen, B. M., Dong, Y., Jiang, H., Kanost, M. R., Koutsos, A. C., Levashina, E. A., Li, J., Ligoxygakis, P., Maccallum, R. M., Mayhew, G. F., Mendes, A., Michel, K., Osta, M. A., Paskewitz, S., Shin, S. W., Vlachou, D., Wang, L., Wei, W., Zheng, L., Zou, Z., Severson, D. W., Raikhel,

A. S., Kafatos, F. C., Dimopoulos, G., Zdobnov, E. M., Christophides, G. K., 2007. Evolutionary dynamics of immune-related genes and pathways in disease-vector mosquitoes. Science 316, 1738–1743.

Waterhouse, R. M., Wyder, S., Zdobnov, E. M., 2008. The *Aedes aegypti* genome: A comparative perspective. Insect Mol. Biol. 17, 1–8.

Williams, M.J., Rodriguez, A., Kimbrell, D. a., Eldon, E.D., 1997. The 18-wheeler mutation reveals complex antibacterial gene regulation in *Drosophila* host defense. EMBO J. 16, 6120–6130.

Wu, L. P., Anderson, K. V., 1998. Regulated nuclear import of Rel proteins in the *Drosophila* immune response. Nature 392, 93–97.

Yagi, Y., Nishida, Y., Ip, Y. T., 2010. Functional analysis of *Toll*-related genes in *Drosophila*. Dev. Growth, Differ. 52, 771–783.

Yang, Z., Jiang, H., Zhao, X., Lu, Z., Luo, Z., Li, X., Zhao, J., Zhang, Y., 2017. Correlation of cell surface proteins of distinct *Beauveria bassiana* cell types and adaption to varied environment and interaction with the host insect. Fungal Genet. Biol. 99, 13–25.

Zdobnov, E. M., von Mering, C., Letunic, I., Torrents, D., Suyama, M., Copley, R. R., Christophides, G. K., Thomasova, D., Holt, R. A., Subramanian, G. M., Mueller, H. M., Dimopoulos, G., Law, J. H., Wells, M. A., Birney, E., Charlab, R., Halpern, A. L., Kokoza, E., Kraft, C. L., Lai, Z., Lewis, S., Louis, C., Barillas-Mury, C., Nusskern, D., Rubin, G. M., Salzberg, S. L., Sutton, G. G., Topalis, P., Wides, R., Wincker, P., Yandell, M., Collins, F. H., Ribeiro, J., Gelbart, W. M., Kafatos, F. C., Bork, P., 2002. Comparative genome and proteome analysis of *Anopheles gambiae* and *Drosophila melanogaster*. Science 298, 149–159.

Zhang, J., 2003. Evolution by gene duplication: an update. Trends Ecol. Evol. 18, 292–298.

Zou, Z., Evans, J. D., Lu, Z., Zhao, P., Williams, M., Sumathipala, N., Hetru, C., Hultmark, D., Jiang, H., 2007. Comparative genomic analysis of the *Tribolium* immune system. Genome Biol. 8, R177.

Zou, Z., Souza-Neto, J., Xi, Z., Kokoza, V., Shin, S. W., Dimopoulos, G., Raikhel, A., 2011. Transcriptome analysis of *Aedes aegypti* transgenic mosquitoes with altered immunity. PLoS Pathog. 7, e1002394.

Figures – Chapter 3

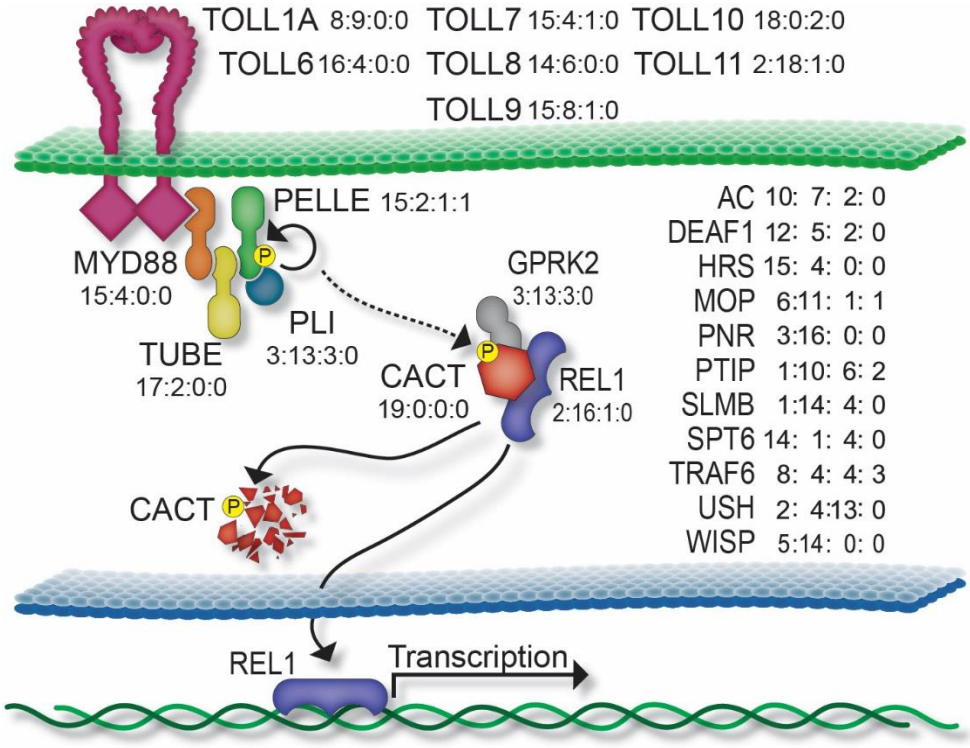

Figure 3-1 Schematic representation of the Toll signaling pathway and annotation summary

Members of the pathway that have defined placement in the cascade (through synthesis of *D. melanogaster* literature) are indicated with shapes. Solid arrows indicate confirmed molecular interactions, while the dotted arrow indicates interaction that may or may not be direct. Pathway members that have been implicated in Toll signaling, but whose placement in the pathway is unknown, are listed on the right of the schematic. TLRs, annotated across 18 anopheline genomes, and those of *C. quinquefaciatus,* and *Ae. aegypti* are listed along the top of the figure. TLR coding sequences were annotated in 20 mosquito genomes, as reliable annotation of TLRs in *An. maculatus* was hindered by its fragmented genome assembly. We excluded TOLL1B, TOLL5A and TOLL5B from this figure, as 1:1 orthology across the 20 mosquito genomes could

not be assigned (see Figure 3). Toll pathway member coding sequences were annotated in 19 mosquito genomes, as reliable annotation of pathway members in *An. maculatus* and *An. christyi* was hindered by its fragmented genome assembly. Numbers adjacent to each pathway member indicate the different types of changes made to the annotation of gene models across the mosquito genomes (no changes to existing gene model: improved annotation: incomplete coding sequence: gene not identified; see Supplemental Tables 1 and 2for details).

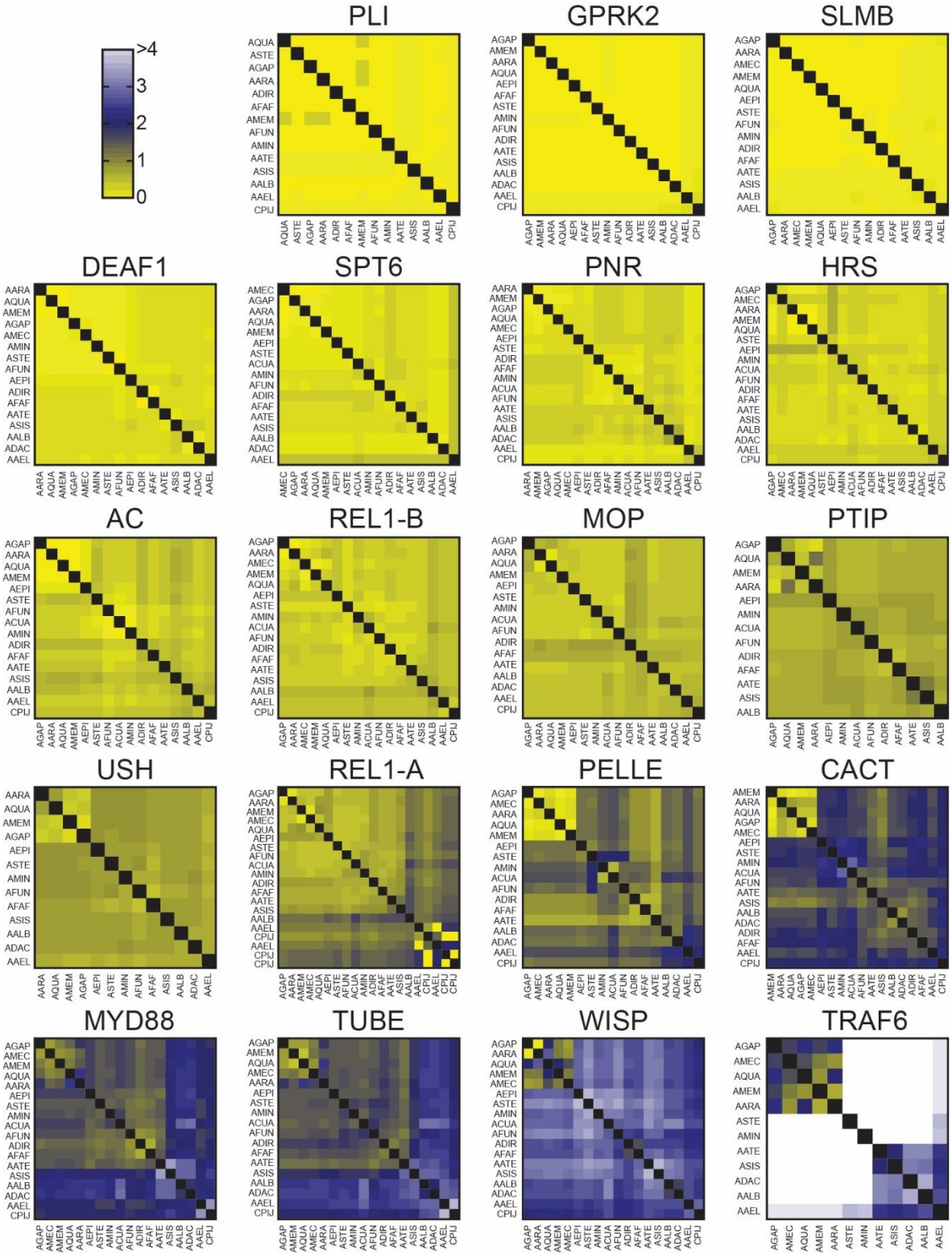

Figure 3-2 Heat map representations of phylogenetic distances of annotated Toll pathway members

Heat maps indicate the pairwise comparisons of phylogenetic distances in substitutions/site within each representative toll pathway member. Compared species for each gene model set are listed along the y- and x-axis. Scale indicated in top left, with yellow indicated highly similar sequences (substitutions/site, normalized to the phylogenetic distances for each corresponding species comparison as published in Neafsey et al., 2015 and blue and white indicating higher levels of sequence divergence. Pathway members are ordered from least (PLI) to greatest (TRAF6) average normalized pairwise distances.

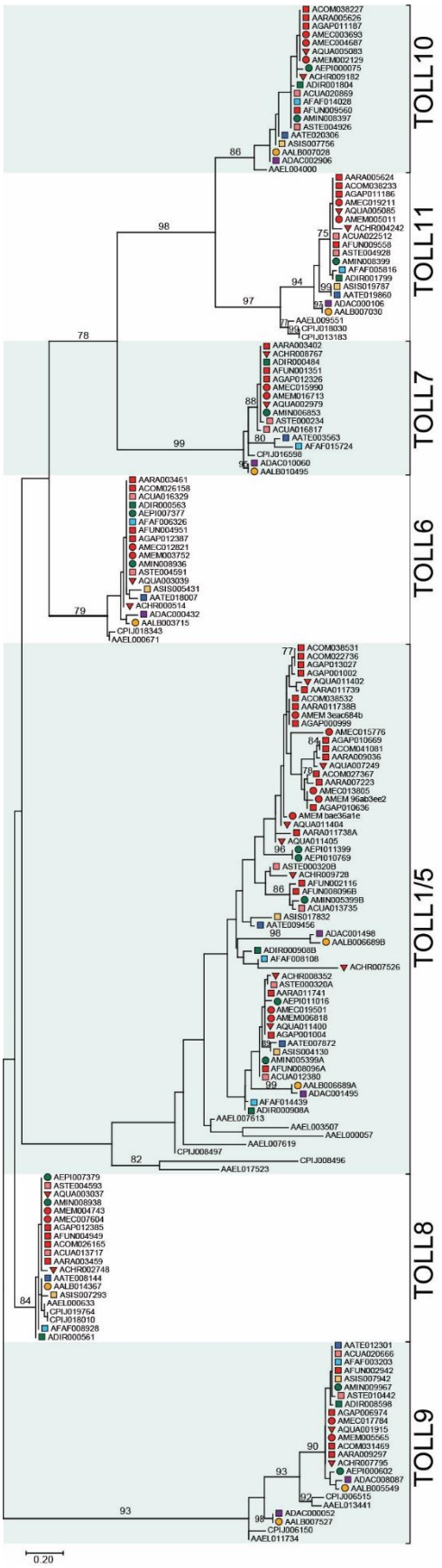

Figure 3-3 Phylogenetic relationships of Toll-like receptors from 20 mosquito species

Maximum likelihood phylogeny of the TIR domain of the TLR family shows strong support for branches supporting clades of TOLL6, TOLL7, TOLL8, TOLL9, TOLL10, and TOLL11 TLR subfamilies across the examined mosquito genomes. The tree is drawn to scale, with the scale bar indicating substitutions per site per unit of branch length and the number at each branch reflects bootstrap support in percent (1000 replications). Only branches with 75% support or higher have values listed. Branch labels are coded according to Neafsey et al., 2015 and indicate vector status and geographic distribution of species (square, major vector; circle, minor vector; triangle, nonvector; red, Africa; pink, South Asia; green, South-East Asia; light blue, Asia Pacific; dark blue, Europe; light orange, East Asia; dark orange, Central America; purple, South America). The analysis involved 183 amino acid sequences. All positions with less than 95% site coverage were eliminated. That is, fewer than 5% alignment gaps, missing data, and ambiguous bases were allowed at any position. There were a total of 66 positions in the final dataset.

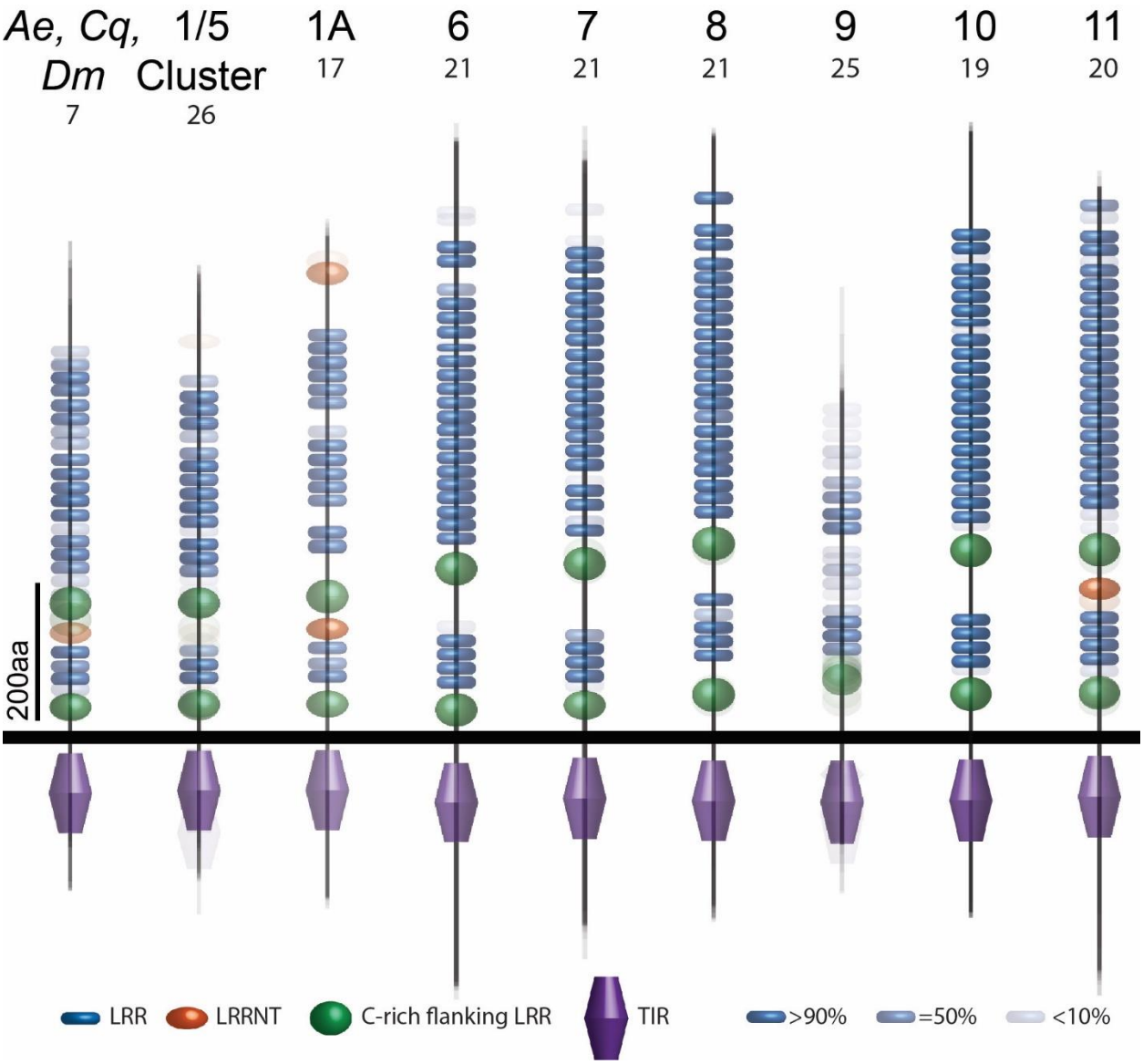

Figure 3-4 Schematic representations of predicted domains within mosquito TLR subfamilies

Domains are drawn to scale and predicted using Pfam, TMHMM Server version 2.0, and LRR finder. LRR, (blue) LRR-CT (green), LRR-NT (orange), and TIR (purple) domains are indicated. Black rectangle is a transmembrane domain. Each subfamily depicts a protein motif prediction overlay, with more opaque motifs indicating highly conserved motifs within a subfamily. Subfamilies listed (from left to right: *Ae. aegypti, C. quinquefasciatus*, and *D. melanogaster*

TOLL1/5, anopheline TOLL1/5 cluster, anopheline TOLL1A, TOLL6, TOLL7, TOLL8, TOLL9, TOLL10, and TOLL11) and the corresponding gene models included (listed above) are displayed individually in supplemental files (Figures B-25 to B-31).

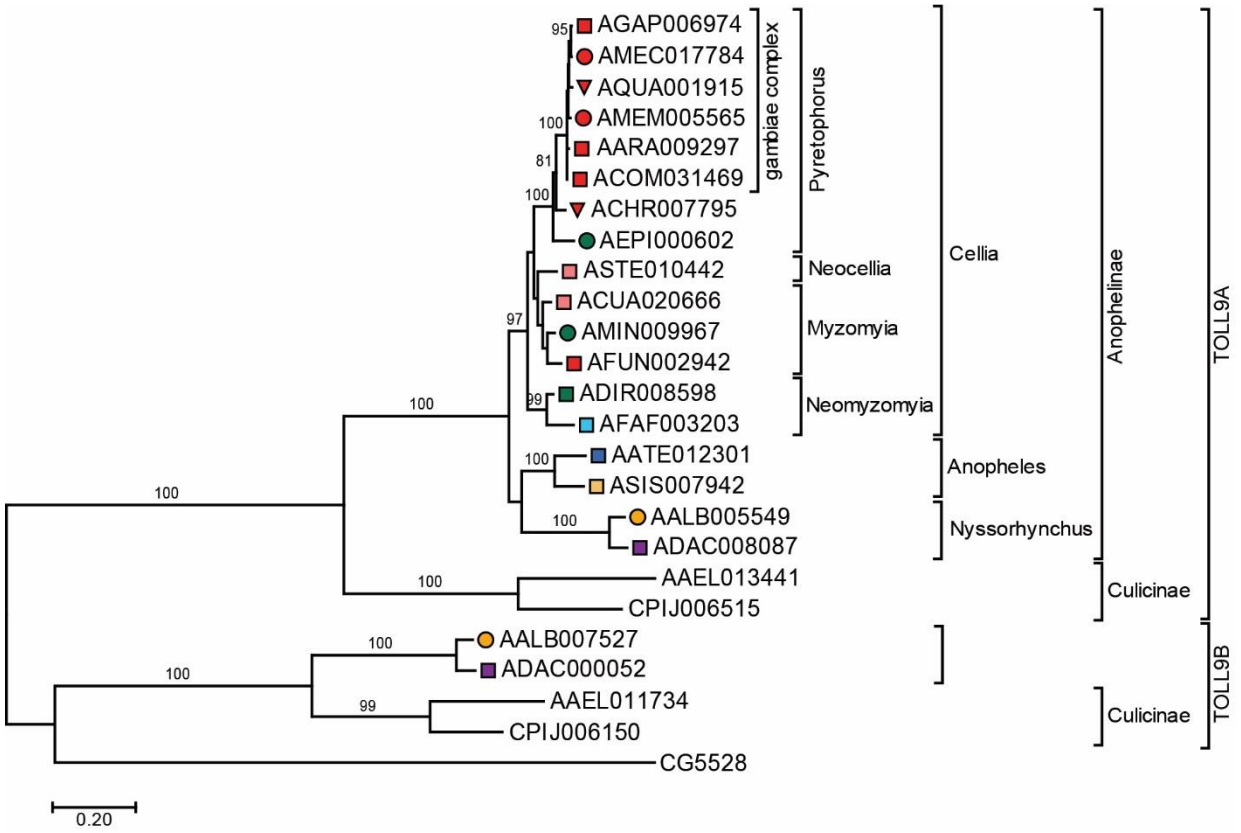

Figure 3-5 Phylogenetic relationships of TOLL9 from 20 mosquito species

Maximum likelihood phylogeny of the full protein sequence of the *TOLL9* subfamily indicates strong support for a duplication of *TOLL9* within *An. albimanus* (AALB007527) and *An. darlingi* (ADAC00052). Scale bar indicates substitutions per site per unit of branch length and the number at each branch reflects bootstrap percentages (1000 replications). Only branches with 75% support have values listed. Branch labels are coded according to Neafsey et al., 2015 and indicate vector status and geographic distribution of species (square, major vector; circle, minor vector; triangle, nonvector; red, Africa; pink, South Asia; green, South-East Asia; light blue, Asia Pacific; dark blue, Europe; light orange, East Asia; dark orange, Central America; purple, South America).

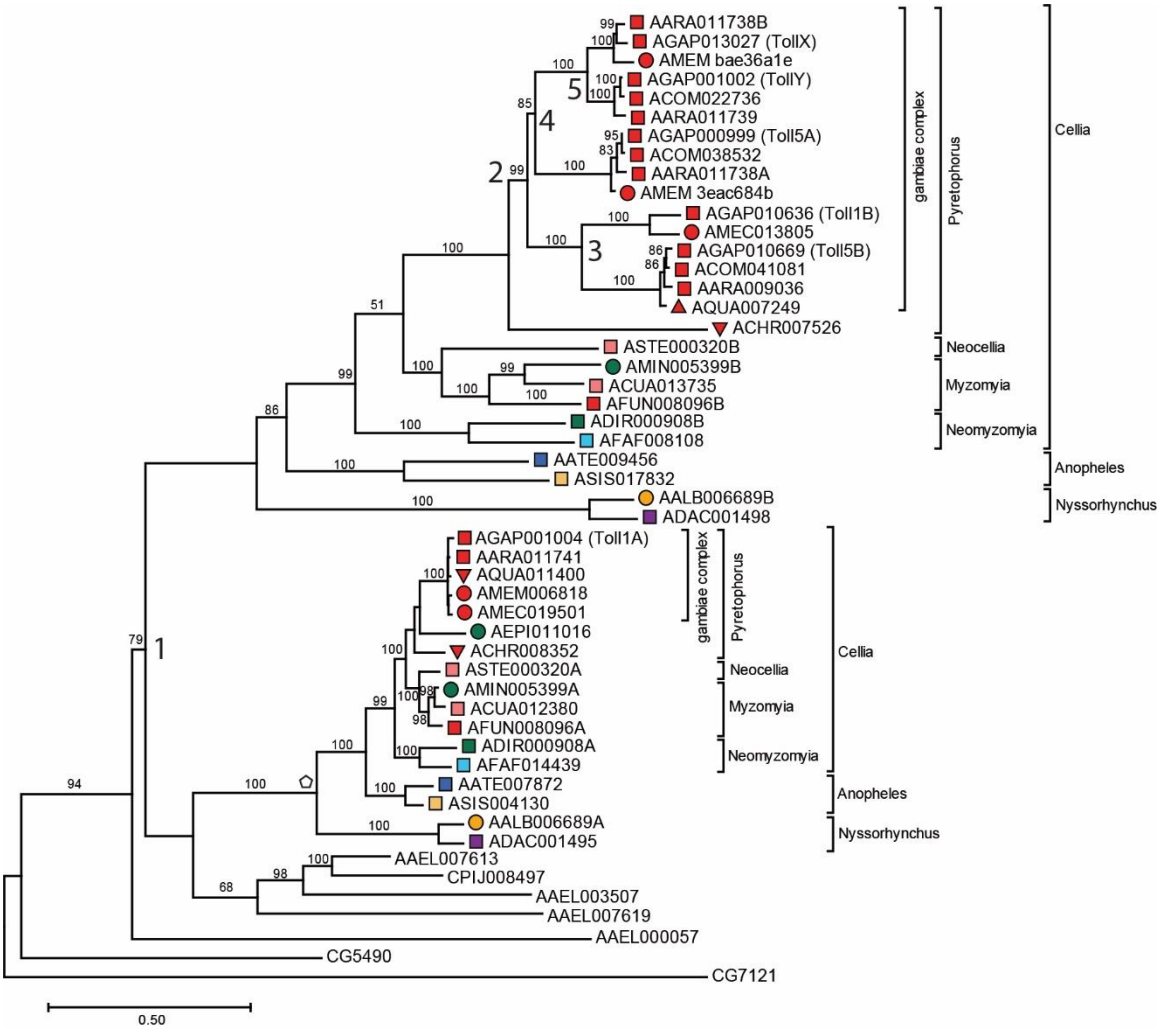

Figure 3-6 Phylogenetic relationships of TOLL1/5 expansion cluster from 20 mosquito species

Maximum likelihood phylogeny of the full protein sequence of TOLL1/5 subfamily indicates strong support for multiple duplication events within anophelines (numbered 1-5). 1:1 orthology observed for protein sequences corresponding to TOLL1A (pentagon). Scale bar indicates substitutions per site per unit of branch length and the number at each branch reflects bootstrap percentages (1000 replications). Only branches with 75% support have values listed. Branch labels are coded according to Neafsey et al., 2015and indicate vector status and geographic distribution of species (square, major vector; circle, minor vector; triangle, nonvector; red,

Africa; pink, South Asia; green, South-East Asia; light blue, Asia Pacific; dark blue, Europe; light orange, East Asia; dark orange, Central America; purple, South America).

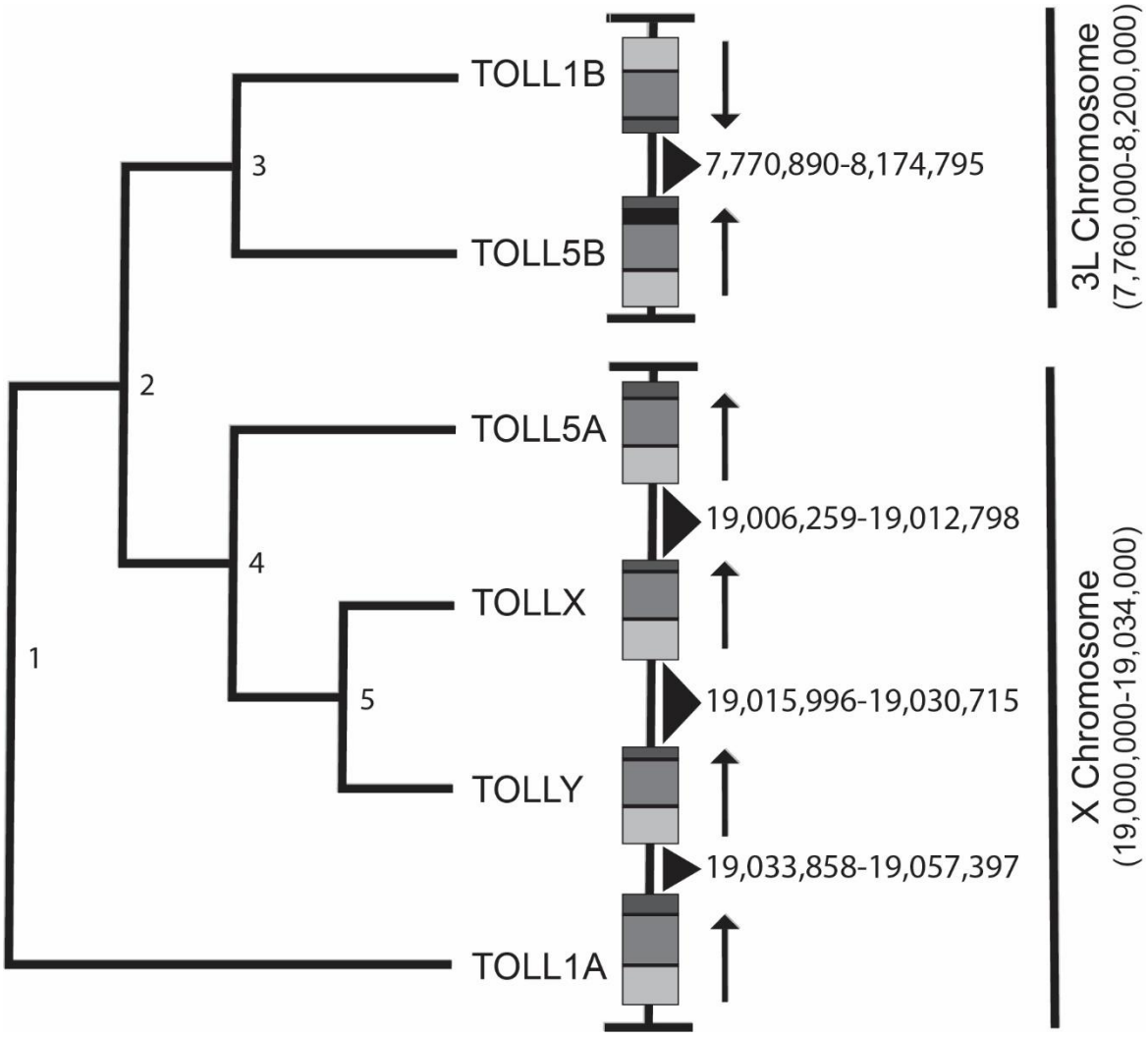

Figure 3-7 Genomic locations of *TOLL1/5* cluster genes within *An. gambiae*

Schematic depiction of TLR locations belonging to the TOLL1/5 expansion cluster within *An. gambiae*, with phylogenetic relationships depicted on the left and duplication events numbered (1-5). Orientations depicted by arrows. All gene models are drawn to scale and contain three exons, with first exon in light gray and third exon in dark gray. Intronic spaces indicated in black. Lengths of separating genomic sequences provided (indicated by black triangles) and the overall genomic sequence of gene models is provided on the right.

Chapter 4 - Conclusions

The research presented in this dissertation was completed to (A) establish the interaction between pathogen dose, immune activation state, and survival in the most important malaria vector species and (B) determine the Toll pathway immune repertoire in several recently sequenced mosquito species, that are both vectors and nonvectors for human malaria. This chapter provides a summary of the major findings towards these two objectives, which contributed critical information to our understanding of the immune repertoire of mosquitoes and the interaction between basal immune activation and *An. gambiae* survival. The following sections highlight these major findings in the context of current literature and give a brief outlook on possible future studies focusing on the interactions of immunity and disease outcome in vector mosquitoes.

A. Establish the interaction between pathogen dose, immune activation state, and survival in an influential vector species

Control of mosquito vectors continues to be dependent on the widespread application of chemical insecticides through indoor residual spraying (IRS) and insecticide treated bed nets (ITNs). However, resistance to these insecticides is spreading within many vector populations (Benelli and Beier, 2017; Hemingway and Ranson, 2000; Ranson et al., 2011; Ranson and Lissenden, 2016), adding an ever-increasing urgency to the development of alternative vector control measures (reviewed in Barreaux et al., 2017; Hemingway et al., 2016; Kamareddine, 2012; Thomas et al., 2012). Entomopathogenic fungi, including *B. bassiana*, have been considered as an alternative control measure due to their use as biopesticides (Blanford et al., 2005; Scholte et al., 2005).

Biopesticides, such as *B. bassiana*, kill their hosts over the course of an invasive infection. This invasive infection results in the activation of immune pathways, including the antifungal Toll signaling pathway, which can affect an insect's ability to withstand the infection. Indeed, previous studies demonstrate that increased expression of the *REL1* Toll transcription factor boosts basal immunity and positively affects the ability of *Aedes aegypti* to withstand *B. bassiana* infection (Bian et al., 2005; Shin et al., 2006, 2005). While these studies suggest the Toll pathway could constitute as a selection target for resistance against *B. bassiana* infection, information concerning the relationship between Toll signaling activation and *B. bassiana* exposure dose is unclear.

To this end, the work presented in Chapter 2 hypothesized that changes to the basal activation state of the Toll pathway would affect mosquito survival following exposure to varying doses of *B. bassiana*. We confirmed that *B. bassiana* exposure dose directly affects mosquito survival, as increasing *B. bassiana* infection dose resulted in increased percent daily mortality and decreased median survival and longevity in infected female mosquitoes. These dose-dependent infection dynamics were reflected in the observed hazard ratios, with positively correlated doses and hazard ratios for all treatments. Furthermore, RNA interference (RNAi) knockdown of target transcripts *REL1* and *Cactus*, in the context of these doses, also affected the survival of mosquitoes. These transcripts were targeted for RNAi knockdown due to their central roles in Toll signaling. *REL1* acts as the terminal transcription factor for the Toll pathway, while *Cactus* serves as the key inhibitor of *REL1* (Barillas-Mury et al., 1996; Christophides et al., 2002). In addition, knockdown of these transcripts was also shown to strongly affect Toll signaling in mosquitoes (Bian et al., 2005; Frolet et al., 2006; Garver et al., 2009).

To determine if inhibition of the Toll pathway activation, by removal of the *REL1* terminal transcription factor, during *B. bassiana* infection would negatively affect mosquito survival, mosquitoes were injected with dsRNA specific to *REL1* (ds*REL1*) and subsequently exposed to varying doses of *B. bassiana*. After RNAi knockdown of *REL1* transcripts following *B. bassiana* infection, we observed increased mortality in infected mosquitoes compared to ds*GFP* controls, which was in agreement with results from a previous study in *Ae. aegypti* mosquitoes (Shin et al., 2005). Furthermore, we also observed that ds*REL1* injection resulted in significantly decreased median survival and longevity compared to control *dsGFP*-injected mosquitoes infected with the same medium and high *B. bassiana* doses.

To test if increased basal activation of Toll signaling had the inverse effect on mosquito survival following *B. bassiana* exposure, we performed similar studies in which dsRNA specific to *Cactus* (ds*Cactus*) was injected into female mosquitoes before exposure to varying doses of *B. bassiana.* We found ds*Cactus*-injection was detrimental to mosquito survival in the absence of infection, significantly decreasing survivorship compared to controls. However, ds*Cactus*-depleted mosquitoes had better survival rates and higher median survival after low dose *B. bassiana* exposure compared to similarly infected dsRNA-treated control mosquitoes. Interestingly, as fungal dosage increased, any benefits gained through preemptive Toll activation by *Cactus* depletion were lost.

The work presented in Chapter 2 suggests that Toll immune signaling does not appear to be activated after exposure to the lowest dose of *B. bassiana* used in these analyses. We observed that knockdown of *REL1* by ds*REL1* injection does not significantly decrease mosquito survival after exposure to a low conidial dose of *B. bassiana.* This implies that expression of antifungal immune factors at levels high enough to alter infection outcome is unlikely. Due to the observed

increase in mosquito survival after knockdown of *Cactus* prior to low dose exposure, these findings also reflect the positive impact that activated Toll signaling has on limiting *B. bassiana*-induced pathology.

Possible mechanisms underlying this lack of Toll activation at low dose *B. bassiana* exposure are mosquito immune system evasion or suppression by the infecting fungus. Both mechanisms are commonly employed by arthropod pathogens, e.g. *Plasmodium* spp. utilize both of these mechanisms to escape mosquito nitration, encapsulation, and melanization responses (Boëte et al., 2004; Lambrechts et al., 2007; Michel et al., 2005; Molina-Cruz et al., 2015). *B. bassiana* also employs a variety of methods to evade host immunity, such as the use of nonimmunogenic hyphal bodies during infection that evade immune recognition through a lack of antigenic surface compounds (Pendland et al., 1993). *Beauveria* is also capable of modulating insect responses to suppress a host's immune defenses by using different molecular classes of proteases and toxins (reviewed in Joop and Vilcinskas, 2016). Future testing utilizing purified proteins from the *B. bassiana* secretome injected into vector mosquitoes would shed light on the immunomodulatory effects these proteases and toxins may possess.

Overall, these data presented in Chapter 2 demonstrate that *An. gambiae* immune activation state has the potential to enhance survival against biopesticides such as *B. bassiana*. Additional studies are needed to determine if the observed impact on lethal time 50 (LT50) is influenced by experimental conditions and/or *B. bassiana* strain tested. Given that the data strongly suggest immune evasion does occur, future studies should focus on understanding the evasion mechanisms employed by *B. bassiana* during infection within *An. gambiae* as well as testing additional strains of *B. bassiana* to gauge whether similar effects are observed.

B. Determining the Toll pathway immune repertoire in several recently sequenced anopheline mosquito species

The sequencing of several species within a single genus, like *Anopheles*, provides a unique opportunity to compare and understand the genetic basis underlying vector vs. nonvector status in insects (Neafsey et al., 2013). Comparative genomics also allows for a better understanding of coding and/or regulatory DNA sequence differences between vector and nonvector species. Data from these studies can facilitate the identification and targeting of specific gene products that predispose species to be successful vector pathogens. For example, when sequencing 16 anophelines from Africa, Europe, Asia, and Latin America, chemosensory genes, important in vector host selection, were reported to diversify across species through changes in protein sequence (Neafsey et al., 2015). These protein sequence changes on mosquito host-selection receptors have the potential to change host preference; thus, predisposing one species to vector disease in humans, while a related species is considered a nonvector. Additionally, genes that control immunity to *Plasmodium* malaria parasites also varied among mosquito species, with signal transducer and activator of transcription (*STAT*) genes undergoing duplications after the divergence of *An. dirus* and *An. farauti* from the *Cellia* subgenus, possibly affect the immune responses elicited in the species belonging to the gambiae complex (Neafsey et al., 2015). Overall, the vector community has the potential to develop novel methods of vector control through the utilization of these genome level differences between vector and nonvector mosquito species.

The Toll pathway repertoire outlined in Chapter 3 provides a thorough identification and manual annotation of 18 intracellular Toll pathway members, first described in *D. melanogaster* (Haghayeghi et al., 2010; Huang et al., 2010; Kuttenkeuler et al., 2010; Valanne et al., 2011,

2010), in the assembled genomes of 20 different mosquito species (Arensburger et al., 2010; Holt et al., 2002; Lawniczak et al., 2010; Neafsey et al., 2015; Nene et al., 2007). Of the 342 gene model coding sequences included in this work, 150 gene models required no changes to the currently published coding sequence while 140 required annotation refinements. Additionally, 44 CDS gene models could not be fully annotated due to sequence gaps or their location on scaffold edges within the assembled genome sequences. Lastly, eight gene models could not be located in their corresponding genome published sequences. Similar to previous studies focusing on immune signaling pathways in mosquitoes (Neafsey et al., 2015; Waterhouse et al., 2007), I found that the intracellular Toll pathway is conserved among 19 mosquito species, with 1:1 orthology, as well as *D. melanogaster*. I also observed that a subset of these pathway members are acquiring amino acid substitutions at higher rates than observed for the universal single-copy ortholog core (identified in Neafsey et al., 2015), indicating these pathway members may be diversifying within the anophelines. Interestingly, the pathway members acquiring these substitutions were often pathway members that have defined placement within the Toll signaling pathway (reviewed in Valanne et al., 2011), such as PELLE, MYD88, TUBE, and CACT.

Chapter 3 also details the annotative and phylogenetic analyses of the Toll-like receptor (TLR) family within anopheline species (Arensburger et al., 2010; Holt et al., 2002; Lawniczak et al., 2010; Neafsey et al., 2015; Nene et al., 2007). TLR repertoires were observed to vary drastically between vertebrate species through comparative genomics, showing expanded TLR repertoires in multiple lineages such as vertebrates (Roach et al., 2005), echinoderms (Buckley and Rast, 2012), and arthropods (Palmer and Jiggins, 2015). Duplication events in these expanded TLR repertoires make it difficult to relate TLR function observed in one genus to other genera based on sequence identity alone. Therefore, it is essential to study and analyze this

important receptor family within mosquito vectors important to human health, to promote the understanding of the biology of these vectors and further increase the possibility for development of novel vector control measures.

In the analyses detailed in Chapter 3, I observed that the overall TLR repertoire among the 20 *Anopheles* species was consistent with previous descriptions in *Ae. aegypti* and *An. gambiae* (Waterhouse et al., 2007). The 1:1 orthology observed within TLR clades containing TOLL6, TOLL7, TOLL8, TOLL10, and TOLL11 Toll/interleukin-1 receptor (TIR) domain amino acid sequences, coupled with short phylogenetic distances between the orthologs from the different anopheline species (mean phylogenetic distance < 0.1164 substitutions/site), provides support that these TLRs are highly conserved on a sequence level in mosquitoes belonging to both culicine and anopheline genera. In *D. melanogaster*, these TLRs have been implicated in dendrite guidance (TOLL-6/TOLL-7; Ward et al., 2015), convergent extension of developing embryos (TOLL-6/TOLL-8; Paré et al., 2014), regulation of autophagy (TOLL-7; Nakamoto et al., 2012), defense to viral infection (TOLL-7; Nakamoto et al., 2012), and negative regulation of antimicrobial responses (TOLL-8; Akhouayri et al., 2011). While the biological functions of these TLRs in mosquitoes are currently unknown, their strong sequence conservation (> 80%) suggests identical functions of these proteins across mosquitoes that are likely similar to those found in *D. melanogaster*, where percent sequence identities of anopheline TOLL proteins to *D. melanogaster* TLRs remains above 60%. Additionally, TOLL10 and TOLL11 are absent in *D. melanogaster* and have not yet been functionally characterized in other insect species, including those species included within Chapter 3.

While we reported 1:1 orthology within TLR clades, duplication and diversification were also observed in our analysis of this receptor family. We observed duplication of *TOLL9* within

An. albimanus and *An. darlingi*, with both species belonging to the subgenus *Nyssorhynchus*. The presence of these duplications in *Nyssorhynchus* implies there was an ancient duplication of *TOLL9* before the divergence of the culicine and anopheline subfamilies, which was subsequently lost before the split between *Nyssorhynchus* and the rest of the anopheline mosquitoes. The functional role *D. melanogaster* Toll-9 performs has previously been linked to immunity (Bettencourt et al., 2004, Ooi et al., 2002), but conflicting evidence suggests that this receptor may not be critical for the initiation of immune responses (Narbonne-Reveau, Charroux, and Royet, 2011). Within the lepidopteran *Bombyx mori*, BmToll9 is strongly and constitutively expressed in different parts of the gut and displays inducible expression in epidermis, hemocytes, trachea, and Malpighian tubules in response to *Staphylococcus aureus*, *Escherichia coli*, and *B. bassiana* injected or ingestion (Wu et al., 2010). The sequence divergence between these duplicated genes and their paralogs within *An. albimanus* (36.2% sequence identity) and *An. darlingi* (36.3% sequence identity) supports that these TLR paralogs have widely distinct sequences and, therefore, may bind different ligands.

Lastly, the data in Chapter 3 show that *TOLL1/5* genes in the *gambiae* complex lineage experienced repeated gene duplications, leading to an expansion of *TOLL1/5* genes in these vector species. Phylogenetic analysis of the manually annotated TOLL1/5 sequences reveals that these genes encode complete TLRs, with extracellular repeated LRR domains, a single transmembrane domain, and a TIR domain. Interestingly, within the TOLL1/5 clade, there is a division between genes orthologous to *TOLL1A* and all other *TOLL1/5* coding sequences. As this TOLL1/5 clade contains members from *D. melanogaster* (*Toll* and *Toll-5*) and *Ae. aegypti* (*TOLL5A*) that serve in development (Anderson et al., 1985) and immunity (Lemaitre et al., 1996; Shin et al., 2006), it is tempting to speculate that these expansions may impact

development and immune responses of vector mosquito species such as *An. gambiae*, *An. arabiensis, An. merus,* and *An. melas*.

In summary, the data presented in Chapter 3 provides a thorough description of the complete repertoire of TLRs and intracellular Toll pathway members for all anophelines sequenced to date. This catalog enabled the analysis of the phylogenetic relationships and conservation of a signaling pathway that is not only diverse in sequence, but also diverse in function. We show that the intracellular Toll signaling cascade is conserved, with 1:1 orthologs found in all species included in this study. In addition, we have provided a complete annotation of the TLR family in anophelines, enabling future studies on their biological functions.

C. Future Directions

The work in this dissertation provides a foundation for the study of Toll pathway function within several mosquito species. By utilizing the description of mosquito immunity provided in Chapter 1, the interaction between the fungal entomopathogen *B. bassiana* and the malaria vector *An. gambiae* illustrated in Chapter 2, and the putative orthologs of the Toll pathway described in Chapter 3, future studies can probe this specific immune pathway to determine its role in both vector and nonvector mosquitoes.

Future experiments should determine the role of the specific TLRs in Toll pathway activation as an antifungal defense against *B. bassiana* infection within *An. gambiae*. The work presented in Chapter 2 demonstrates that the Toll pathway does, in fact, respond to infection with this entomopathogen at easily applied doses of infective conidial spores. Furthermore, we provide data that knockdown of expression of pathway components (*REL1* and *CACT*) by RNAi has observable phenotypes shortly after exposure to these conidia. In addition, previous

studies have confirmed the antifungal function of the Toll pathway *in vivo* through the decreased expression of Toll pathway members (Lemaitre et al., 1996; Shin et al., 2006, 2005). Future *in vivo* studies should be completed to determine if the TLRs described in Chapter 3 recognize and limit *B. bassiana* infection within *An. gambiae* through the knockdown of these receptors by RNAi.

If TLRs act in the antifungal defense against *B. bassiana*, RNAi knockdown of these receptors would result in decreased survival following infection; likely a result of increased replicative growth of the fungus in mosquito hemocoel. In contrast, if the knockdown of TLRs led to no change in hazard compared to controls, these TLRs are likely not activated during *B. bassiana* infection. Poor knockdown efficiency or low basal expression of these receptors may limit the effectiveness of an RNAi-based functional assay. Under these circumstances, utilizing a gene knockout system such as CRISPR/cas9 would prove an excellent alternative, as CRISPR is established and has been successfully utilized in *Anopheles* (Gantz et al., 2015, Hammond et al., 2016). These experiments would be of great interest, as the identification of TLR function in mosquitoes is still largely undescribed (Luna et al., 2003, 2002; Shin et al., 2006). Furthermore, future studies can analyze the effect that TLR knockdown has on gene expression in vector anophelines, both under infection and control conditions, allowing for a detailed look into the expression profiles controlled by this fascinating receptor family.

The work presented in Chapter 2 demonstrates that the basal immune activation state in *An. gambiae* does have the potential to enhance survival to a potential biopesticide like *B. bassiana*. However, further studies will be required to show whether this strong impact on LT50 is influenced by experimental conditions and/or the *B. bassiana* strain tested. Given that the data strongly suggest that immune evasion does occur, future studies might test which, if any, of these

mechanisms are employed by *B. bassiana* during infection of *An. gambiae*. Additional studies are also needed to determine if alterations in the immune activation state lowers the reproductive output of *An. gambiae* through immune/reproduction trade-offs, as previously seen in other model systems (Schwenke et al., 2016). Such studies should monitor the success of eggs laid and hatched, as these data could provide support for the description of a fitness trade-off in this system and should be taken into consideration when determining the long-term efficacy of *B. bassiana* as a late-acting bioinsecticide.

The work presented in Chapter 3 details the annotation and description of the complete catalog of Toll pathway members and TLRs in several anopheline species. Within species where immunity has hardly been studied or analyzed, this catalog will prove useful, as detailed annotation is a crucial first step in any molecular analysis. Using the work presented in Chapter 3, will enable the analysis of the evolutionary history of TLRs and Toll pathway members in greater detail. For example, the 1000 Genomes Consortium (AG1000G) recently published the genome sequences of 765 individual *An. gambiae* and *An. coluzzii* specimens from locations across Africa (Miles et al., 2017). This data, coupled with the detailed annotations described in this dissertation, makes it possible to test whether TLRs and Toll pathway members in *An. gambiae* and *An. coluzzii* contain signatures of purifying and/or positive selection .

Additionally, the full catalog of TLRs described in Chapter 3 provide an opportunity to perform functional ligand binding assays to analyze the binding affinity of these receptors. Such assays have provided insights into the function and activation of human TLRs (Latz et al., 2007) and similar studies analyzing the affinity of *An. gambiae* TLRs to the six putative spätzle TLR ligands encoded within the genome would provide the first description of TLR

binding specificity within an anopheline. Particularly, the TLR duplication of TOLL9 in *An. albimanus* and *An. darlingi* possess unique ligand-binding ectodomain sequences and may reveal TLRs with new ligand specificity within these vectors.

Together, the work in this dissertation has greatly enhanced our understanding of the Toll pathway in mosquito immunity. It has established a natural infection system that will allow the description of TLR function in mosquito immunity as well as the quantification of putative trade-offs between immunity and longevity. This dissertation also provides a comprehensive catalog of TLRs and Toll pathway members in mosquito species where mosquito immunity thus far has not been explored. The descriptions of the Toll pathway and TLR repertoire in 21 mosquito species provide a means to assess the importance of subfamilies in both mosquito development and immunity. Finally, this dissertation can also offer information useful for novel malaria control strategies that are being implemented in the field. The description of the interaction between dose and Toll pathway activation within the malaria vector *An. gambiae* can contribute to the successful application of this potential biopesticide to reduce malaria transmission in the field.

References

Akhouayri, I., Turc, C., Royet, J., Charroux, B., 2011. Toll-8/tollo negatively regulates antimicrobial response in the *Drosophila* respiratory epithelium. PLoS Pathog. 7, e1002319.

Anderson, K. V., Bokla, L., Nusslein-Volhard, C., 1985. Establishment of dorsal-ventral polarity in the *Drosophila* embryo: The induction of polarity by the *Toll* gene product. Cell 42, 791–798.

Anderson, K. V., Nüsslein-Volhard, C., 1984. Information for the dorsal-ventral pattern of the *Drosophila* embryo is stored as maternal mRNA. Nature 311, 223–227.

Arensburger, P., Megy, K., Waterhouse, R. M., Abrudan, J., Amedeo, P., Antelo, B., Bartholomay, L., Bidwell, S., Caler, E., Camara, F., Campbell, C. L., Campbell, K. S., Casola, C., Castro, M. T., Chandramouliswaran, I., Chapman, S. B., Christley, S., Costas, J., Eisenstadt, E., Feschotte, C., Fraser-Liggett, C., Guigo, R., Haas, B., Hammond, M., Hansson, B. S., Hemingway, J., Hill, S. R., Howarth, C., Ignell, R., Kennedy, R. C., Kodira, C. D., Lobo, N. F., Mao, C., Mayhew, G., Michel, K., Mori, A., Liu, N., Naveira, H., Nene, V., NamNguyen, Pearson, M. D., Pritham, E. J., Puiu, D., Qi, Y., Ranson, H., Ribeiro, J. M. C., Roberston, H. M., Severson, D., Shumway, M., Stanke, M., Strausberg, R. L., Sun, C., Sutton, G., Tu, Z. (Jake), Tubio, J. M. C., Unger, M. F., Vanlandingham, D. L., Vilella, A. J., White, O., White, J. R., Wondji, C. S., Wortman, J., Zdobnov, E. M., Birren, B., Christensen, B. M., Collins, F. H., Cornel, A., Dimopoulos, G., Hannick, L. I., Higgs, S., Lanzaro, G. C., Lawson, D., Lee, N. H., Muskavitch, M. A. T., Raikhel, A. S., Atkinson, P. W., 2010. Sequencing of *Culex quinquefasciatus* establishes a platform for mosquito comparative genomics. Science 330, 86–88.

Barillas-Mury, C., Charlesworth, A., Gross, I., Richman, A., Hoffmann, J. A., Kafatos, F. C., 1996. Immune factor Gambif1, a new rel family member from the human malaria vector, *Anopheles gambiae*. EMBO J. 15, 4691–4701.

Barreaux, P., Barreaux, A. M. G., Sternberg, E. D., Suh, E., Waite, J. L., Whitehead, S. A.,

Thomas, M. B., 2017. Priorities for broadening the malaria vector control tool kit. Trends Parasitol. 33, 763–774.

Benelli, G., Beier, J. C., 2017. Current vector control challenges in the fight against malaria. Acta Trop. 174, 91–96.

Bettencourt, R., Tanji, T., Yagi, Y., & Ip, Y. T. (2004). Toll and Toll-9 in *Drosophila* innate immune response. Journal of endotoxin research, 10, 261-268.

Bian, G., Shin, S. W., Cheon, H.-M., Kokoza, V., Raikhel, A. S., 2005. Transgenic alteration of Toll immune pathway in the female mosquito *Aedes aegypti*. Proc. Natl. Acad. Sci. U. S. A. 102, 13568–13573.

Blanford, S., Chan, B. H. K., Jenkins, N., Sim, D., Turner, R. J., Read, A. F., Thomas, M. B., 2005. Fungal pathogen reduces potential for malaria transmission. Science 308, 1638–1641.

Boëte, C., Paul, R. E. L., Koella, J. C., 2004. Direct and indirect immunosuppression by a malaria parasite in its mosquito vector. Proc. R. Soc. L. 271, 1611–1615.

Brownlie, R., Allan, B., 2011. Avian toll-like receptors. Cell Tissue Res. 343, 121–130.

Buckley, K. M., Rast, J. P., 2012. Dynamic evolution of toll-like receptor multigene families in echinoderms. Front. Immunol. 3.

Casanova, J. L., Abel, L., Quintana-Murci, L., 2011. Human TLRs and IL-1Rs in Host Defense: Natural Insights from Evolutionary, Epidemiological, and Clinical Genetics. Annu. Rev. Immunol 29, 447–91.

Christophides, G. K., Zdobnov, E., Barillas-Mury, C., Birney, E., Blandin, S., Blass, C., Brey, P. T., Collins, F. H., Danielli, A., Dimopoulos, G., Hetru, C., Hoa, N. T., Hoffmann, J. A., Kanzok, S. M., Letunic, I., Levashina, E. A., Loukeris, T. G., Lycett, G., Meister, S., Michel, K., Moita, L. F., Müller, H. M., Osta, M. A., Paskewitz, S. M., Reichhart, J. M., Rzhetsky, A., Troxler, L., Vernick, K. D., Vlachou, D., Volz, J., von Mering, C., Xu, J., Zheng, L., Bork, P., Kafatos, F. C., 2002. Immunity-related genes and gene families in *Anopheles gambiae*. Science 298, 159–165.

Frolet, C., Thoma, M., Blandin, S., Hoffmann, J. A., Levashina, E. A., 2006. Boosting NF-κB-dependent basal immunity of *Anopheles gambiae* aborts development of *Plasmodium berghei*. Immunity 25, 677–685.

Gantz, V. M., Jasinskiene, N., Tatarenkova, O., Fazekas, A., Macias, V. M., Bier, E., & James, A. A., 2015. Highly efficient Cas9-mediated gene drive for population modification of the malaria vector mosquito *Anopheles stephensi*. Proceedings of the National Academy of Sciences. 112, E6736-E6743.

Garver, L. S., Dong, Y., Dimopoulos, G., 2009. Caspar controls resistance to *Plasmodium falciparum* in diverse anopheline species. PLoS Pathog. 5, e1000335.

Haghayeghi, A., Sarac, A., Czerniecki, S., Grosshans, J., Schöck, F., 2010. Pellino enhances innate immunity in *Drosophila*. Mech. Dev. 127, 301–307.

Hammond, A., Galizi, R., Kyrou, K., Simoni, A., Siniscalchi, C., Katsanos, D., Gribble, M., Baker, D., Marois, E., Russell, S. and Burt, A., 2016. A CRISPR-Cas9 gene drive system targeting female reproduction in the malaria mosquito vector *Anopheles gambiae*. Nature biotechnology, 34, 78.

Hemingway, J., Ranson, H., 2000. Insecticide resistance in insect vectors of human disease. Annu. Rev. Entomol. 45, 371–391.

Hemingway, J., Shretta, R., Wells, T. N. C., Bell, D., Djimdé, A. A., Achee, N., Qi, G., 2016. Tools and Strategies for Malaria Control and Elimination: What Do We Need to Achieve a Grand Convergence in Malaria? PLoS Biol. 14, e1002380.

Holt, R. A., Subramanian, G. M., Halpern, A., Sutton, G. G., Charlab, R., Nusskern, D. R., Wincker, P., Clark, A. G., Ribeiro, J. M. C., Wides, R., Salzberg, S. L., Loftus, B., Yandell, M., Majoros, W. H., Rusch, D. B., Lai, Z., Kraft, C. L., Abril, J. F., Anthouard, V., Arensburger, P., Atkinson, P. W., Baden, H., Berardinis, V. de, Baldwin, D., Benes, V., Biedler, J., Blass, C., Bolanos, R., Boscus, D., Barnstead, M., Cai, S., Center, A., Chatuverdi, K., Christophides, G. K., Chrystal, M. A., Clamp, M., Cravchik, A., Curwen, V., Dana, A., Delcher, A., Dew, I., Evans, C. A., Flanigan, M., Grundschober-Freimoser, A., Friedli, L., Gu, Z., Guan, P., Guigo, R.,

Hillenmeyer, M. E., Hladun, S. L., Hogan, J. R., Hong, Y. S., Hoover, J., Jaillon, O., Ke, Z., Kodira, C., Kokoza, E., Koutsos, A., Letunic, I., Levitsky, A., Liang, Y., Lin, J. J., Lobo, N. F., Lopez, J. R., Malek, J. A., McIntosh, T. C., Meister, S., Miller, J., Mobarry, C., Mongin, E., Murphy, S. D., O'Brochta, D. A., Pfannkoch, C., Qi, R., Regier, M. A., Remington, K., Shao, H., Sharakhova, M. V., Sitter, C. D., Shetty, J., Smith, T. J., Strong, R., Sun, J., Thomasova, D., Ton, L. Q., Topalis, P., Tu, Z., Unger, M. F., Walenz, B., Wang, A., Wang, J., Wang, M., Wang, X., Woodford, K. J., Wortman, J. R., Wu, M., Yao, A., Zdobnov, E. M., Zhang, H., Zhao, Q., Zhao, S., Zhu, S. C., Zhimulev, I., Coluzzi, M., Torre, A. della, Roth, C. W., Louis, C., Kalush, F., Mural, R. J., Myers, E. W., Adams, M. D., Smith, H. O., Broder, S., Gardner, M. J., Fraser, C. M., Birney, E., Bork, P., Brey, P. T., Venter, J. C., Weissenbach, J., Kafatos, F. C., Collins, F. H., Hoffman, S. L., 2002. The genome sequence of the malaria mosquito *Anopheles gambiae*. Science 298, 129–149.

Huang, H. R., Chen, Z. J., Kunes, S., Chang, G. D., Maniatis, T., 2010. Endocytic pathway is required for *Drosophila* Toll innate immune signaling. Proc. Natl. Acad. Sci. U. S. A. 107, 8322–8327.

Joop, G., Vilcinskas, A., 2016. Coevolution of parasitic fungi and insect hosts. Zoology 119, 350–358.

Kamareddine, L., 2012. The biological control of the malaria vector. Toxins (Basel). 4, 748–767.

Kawai, T., Akira, S., 2006. TLR signaling. Cell Death Differ. 13, 816–825.

Kuttenkeuler, D., Pelte, N., Ragab, A., Gesellchen, V., Schneider, L., Blass, C., Axelsson, E., Huber, W., Boutros, M., 2010. A large-scale RNAi screen identifies *Deaf1* as a regulator of innate immune responses in *Drosophila*. J. Innate Immun. 2, 181–194.

Lambrechts, L., Morlais, I., Awono-Ambene, P. H., Cohuet, A., Simard, F., Jacques, J. C., Bourgouin, C., Koella, J. C., 2007. Effect of infection by *Plasmodium falciparum* on the melanization immune response of *Anopheles gambiae*. Am. J. Trop. Med. Hyg. 76, 475–480.

Latz, E., Verma, A., Visintin, A., Gong, M., Sirois, C.M., Klein, D.C., Monks, B.G., McKnight, C.J., Lamphier, M.S., Duprex, W.P. and Espevik, T., 2007. Ligand-induced conformational changes allosterically activate Toll-like receptor 9. Nature Immunology, 8, 772.

Lawniczak, M. K. N., Emrich, S., Holloway, A. K., Regier, A. P., Olson, M., White, B., Redmond, S., Fulton, L., Appelbaum, E., Godfrey, J., Farmer, C., Chinwalla, A., Yang, S. P., Minx, P., Nelson, J., Kyung, K., Walenz, B. P., Garcia-Hernandez, E., Aguiar, M., Viswanathan, L. D., Rogers, Y. H., Strausberg, R. L., Saski, C. A., Lawson, D., Collins, F. H., Kafatos, F. C., Christophides, G. K., Clifton, S. W., Kirkness, E. F., Besansky, N. J., 2010. Widespread divergence between incipient *Anopheles gambiae* species revealed by whole genome sequences. Science 330, 512–514.

Lemaitre, B., Meister, M., Govind, S., Georgel, P., Steward, R., Reichhart, J. M., Hoffmann, J. A., 1995. Functional analysis and regulation of nuclear import of dorsal during the immune response in Drosophila. EMBO J. 14, 536–545.

Lemaitre, B., Nicolas, E., Michaut, L., Reichhart, J. M., Hoffmann, J. A., 1996. The dorsoventral regulatory gene cassette *spätzle/Toll/cactus* controls the potent antifungal response in *Drosophila* adults. Cell 86, 973–983.

Luna, C., Hoa, N. T., Zhang, J., Kanzok, S. M., Brown, S. E., Imler, J. L., Knudson, D. L., Zheng, L., 2003. Characterization of three Toll-like genes from mosquito *Aedes aegypti*. Insect Mol. Biol. 12, 67–74.

Luna, C., Wang, X., Huang, Y., Zhang, J., Zheng, L., 2002. Characterization of four Toll related genes during development and immune responses in *Anopheles gambiae*. Insect Biochem. Mol. Biol. 32, 1171–1179.

Michel, K., Budd, A., Pinto, S., Gibson, T. J., Kafatos, F. C., 2005. *Anopheles gambiae* SRPN2 facilitates midgut invasion by the malaria parasite *Plasmodium berghei*. EMBO Rep. 6, 891–897.

Miles, A., Harding, N. J., Bottà, G., Clarkson, C. S., Antão, T., Kozak, K., Schrider, D. R., Kern, A. D., Redmond, S., Sharakhov, I., Pearson, R. D., Bergey, C., Fontaine, M.

C., Donnelly, M. J., Lawniczak, M. K. N., Kwiatkowski, D. P., Donnelly, M. J., Ayala, D., Besansky, N. J., Burt, A., Caputo, B., della Torre, A., Fontaine, M. C., Godfray, H. C. J., Hahn, M. W., Kern, A. D., Kwiatkowski, D. P., Lawniczak, M. K. N., Midega, J., Neafsey, D. E., O'Loughlin, S., Pinto, J., Riehle, M. M., Sharakhov, I., Vernick, K. D., Weetman, D., Wilding, C. S., White, B. J., Troco, A. D., Pinto, J., Diabaté, A., O'Loughlin, S., Burt, A., Costantini, C., Rohatgi, K. R., Besansky, N. J., Elissa, N., Pinto, J., Coulibaly, B., Riehle, M. M., Vernick, K. D., Pinto, J., Dinis, J., Midega, J., Mbogo, C., Bejon, P., Wilding, C. S., Weetman, D., Mawejje, H. D., Donnelly, M. J., Weetman, D., Wilding, C. S., Donnelly, M. J., Stalker, J., Rockett, K., Drury, E., Mead, D., Jeffreys, A., Hubbart, C., Rowlands, K., Isaacs, A. T., Jyothi, D., Malangone, C., Vauterin, P., Jeffery, B., Wright, I., Hart, L., Kluczyński, K., Cornelius, V., MacInnis, B., Henrichs, C., Giacomantonio, R., Kwiatkowski, D. P., 2017. Genetic diversity of the African malaria vector *Anopheles gambiae*. Nature 552, 96.

Molina-Cruz, A., Canepa, G. E., Kamath, N., Pavlovic, N. V, Mu, J., Ramphul, U. N., Ramirez, J. L., Barillas-Mury, C., 2015. *Plasmodium* evasion of mosquito immunity and global malaria transmission: The lock-and-key theory. Proc. Natl. Acad. Sci. U. S. A. 112, 15178–15183.

Nakamoto, M., Moy, R. H., Xu, J., Bambina, S., Yasunaga, A., Spencer, S., Gold, B., Cherry, S., 2012. Virus recognition by Toll-7 activates antiviral autophagy in *Drosophila*. Immunity 36, 658–667.

Narbonne-Reveau, K., Charroux, B., Royet, J., 2011. Lack of an antibacterial response defect in *Drosophila Toll-9* mutant. PLoS One 6, 1–11.

Neafsey, D. E., Waterhouse, R. M., Abai, M. R., Aganezov, S. S., Alekseyev, M. A., Allen, J. E., Amon, J., Arcà, B., Arensburger, P., Artemov, G., Assour, L. A., Basseri, H., Berlin, A., Birren, B. W., Blandin, S. A., Brockman, A. I., Burkot, T. R., Burt, A., Chan, C. S., Chauve, C., Chiu, J. C., Christensen, M., Costantini, C., Davidson, V. L. M., Deligianni, E., Dottorini, T., Dritsou, V., Gabriel, S. B., Guelbeogo, W. M., Hall, A. B., Han, M. V, Hlaing, T., Hughes, D. S. T., Jenkins, A. M., Jiang, X., Jungreis, I., Kakani, E. G., Kamali, M., Kemppainen, P., Kennedy, R. C.,

Kirmitzoglou, I. K., Koekemoer, L. L., Laban, N., Langridge, N., Lawniczak, M. K. N., Lirakis, M., Lobo, N. F., Lowy, E., MacCallum, R. M., Mao, C., Maslen, G., Mbogo, C., McCarthy, J., Michel, K., Mitchell, S. N., Moore, W., Murphy, K. A., Naumenko, A. N., Nolan, T., Novoa, E. M., O'Loughlin, S., Oringanje, C., Oshaghi, M. A., Pakpour, N., Papathanos, P. A., Peery, A. N., Povelones, M., Prakash, A., Price, D. P., Rajaraman, A., Reimer, L. J., Rinker, D. C., Rokas, A., Russell, T. L., Sagnon, N., Sharakhova, M. V, Shea, T., Simão, F. A., Simard, F., Slotman, M. A., Somboon, P., Stegniy, V., Struchiner, C. J., Thomas, G. W. C., Tojo, M., Topalis, P., Tubio, J. M. C., Unger, M. F., Vontas, J., Walton, C., Wilding, C. S., Willis, J. H., Wu, Y. C., Yan, G., Zdobnov, E. M., Zhou, X., Catteruccia, F., Christophides, G. K., Collins, F. H., Cornman, R. S., Crisanti, A., Donnelly, M. J., Emrich, S. J., Fontaine, M. C., Gelbart, W., Hahn, M. W., Hansen, I. A., Howell, P. I., Kafatos, F. C., Kellis, M., Lawson, D., Louis, C., Luckhart, S., Muskavitch, M. a T., Ribeiro, J. M., Riehle, M. A., Sharakhov, I. V, Tu, Z., Zwiebel, L. J., Besansky, N. J., 2015. Highly evolvable malaria vectors: The genomes of 16 *Anopheles* mosquitoes. Science 347, 1258522-1-1258522–8.

Nene, V., Wortman, J. R., Lawson, D., Haas, B., Kodira, C., Tu, Z., Loftus, B., Xi, Z., Megy, K., Grabherr, M., Ren, Q., Zdobnov, E. M., Lobo, N. F., Campbell, K. S., Brown, S. E., Bonaldo, M. F., Zhu, J., Sinkins, S. P., Hogenkamp, D. G., Amedeo, P., Arensburger, P., Atkinson, P. W., Bidwell, S., Biedler, J., Birney, E., Bruggner, R. V, Costas, J., Coy, M. R., Crabtree, J., Crawford, M., DeBruyn, B., DeCaprio, D., Eiglmeier, K., Eisenstadt, E., El-Dorry, H., Gelbart, W. M., Gomes, S. L., Hammond, M., Hannick, L. I., Hogan, J. R., Holmes, M. H., Jaffe, D., Spencer Johnston, J., Kennedy, R. C., Koo, H., Kravitz, S., Kriventseva, E. V, Kulp, D., LaButti, K., Lee, E., Li, S., Lovin, D. D., Mao, C., Mauceli, E., M Menck, C. F., Miller, J. R., Montgomery, P., Mori, A., Nascimento, A. L., Naveira, H. F., Nusbaum, C., Orvis, J., Pertea, M., Quesneville, H., Reidenbach, K. R., Rogers, Y. H., Roth, C. W., Schneider, J. R., Schatz, M., Shumway, M., Stanke, M., Stinson, E. O., C Tubio, J. M., VanZee, J. P., Verjovski-Almeida, S., Werner, D., White, O., Wyder, S., Zeng, Q., Zhao, Q., Zhao, Y., Hill, C. A., Raikhel, A. S., Soares, M. B., Knudson, D. L., Lee, N. H., Galagan, J., Salzberg, S. L., Paulsen, I. T., Dimopoulos,

G., Collins, F. H., Birren, B., Fraser-Liggett, C. M., Severson, D. W., 2007. Genome sequence of *Aedes aegypti*, a major arbovirus vector. Science 316, 1718–1723.

Ooi, J. Y., Yagi, Y., Hu, X., Ip, Y. T., 2002. The *Drosophila* Toll-9 activates a constitutive antimicrobial defense. EMBO Rep. 3, 82–87.

Palmer, W. J., Jiggins, F. M., 2015. Comparative genomics reveals the origins and diversity of arthropod immune systems. Mol. Biol. Evol. 32, 2111–2129.

Paré, A. C., Vichas, A., Fincher, C. T., Mirman, Z., Farrell, D. L., Mainieri, A., Zallen, J. A., 2014. A positional Toll receptor code directs convergent extension in *Drosophila*. Nature 515, 523–527.

Pendland, J. C., Hung, S. Y., Boucias, D. G., 1993. Evasion of host defense by in vivo-produced protoplast-like cells of the insect mycopathogen *Beauveria bassiana*. J. Bacteriol. 175, 5962–5969.

Ranson, H., Lissenden, N., 2016. Insecticide resistance in African *Anopheles* mosquitoes: A worsening situation that needs urgent action to maintain malaria control. Trends Parasitol. 32, 187–196.

Ranson, H., N'Guessan, R., Lines, J., Moiroux, N., Nkuni, Z., Corbel, V., 2011. Pyrethroid resistance in African anopheline mosquitoes: What are the implications for malaria control? Trends Parasitol. 27, 91–98.

Rebl, A., Goldammer, T., Seyfert, H. M., 2010. Toll-like receptor signaling in bony fish. Vet. Immunol. Immunopathol. 134, 139–150.

Roach, J. C., Glusman, G., Rowen, L., Kaur, A., Purcell, M. K., Smith, K. D., Hood, L. E., Aderem, A., 2005. The evolution of vertebrate Toll-like receptors. Proc. Natl. Acad. Sci. 102, 9577–9582.

Samaraweera, S. E., O'Keefe, L. V., Price, G. R., Venter, D. J., Richards, R. I., 2013. Distinct roles for *Toll* and autophagy pathways in double-stranded RNA toxicity in a *Drosophila* model of expanded repeat neurodegenerative diseases. Hum. Mol. Genet. 22, 2811–2819.

Scholte, E.-J., Ng'habi, K., Kihonda, J., Takken, W., Paaijmans, K., Abdulla, S., Killeen, G. F., Knols, B. G. J., 2005. An entomopathogenic fungus for control of adult African malaria mosquitoes. Science 308, 1641–1642.

Schwenke, R. A., Lazzaro, B. P., Wolfner, M. F., 2016. Reproduction–immunity trade-offs in insects. Annu. Rev. Entomol. 61, 239–256.

Shin, S. W., Bian, G. W., Raikhel, A. S., 2006. A toll receptor and a cytokine, Toll5A and Spz1C, are involved in toll antifungal immune signaling in the mosquito *Aedes aegypti*. J. Biol. Chem. 281, 39388–39395.

Shin, S. W., Kokoza, V., Bian, G., Cheon, H. M., Yu, J. K., Raikhel, A. S., 2005. REL1, a homologue of *Drosophila* Dorsal, regulates Toll antifungal immune pathway in the female mosquito *Aedes aegypti*. J. Biol. Chem. 280, 16499–16507.

Thomas, M. B., Godfray, H. C. J., Read, A. F., van den Berg, H., Tabashnik, B. E., van Lenteren, J. C., Waage, J. K., Takken, W., 2012. Lessons from agriculture for the sustainable management of malaria vectors. PLoS Med. 9, e1001262.

Valanne, S., Myllymäki, H., Kallio, J., Schmid, M. R., Kleino, A., Murumägi, A., Airaksinen, L., Kotipelto, T., Kaustio, M., Ulvila, J., Esfahani, S. S., Engström, Y., Silvennoinen, O., Hultmark, D., Parikka, M., Rämet, M., 2010. Genome-wide RNA interference in *Drosophila* cells identifies G protein-coupled receptor kinase 2 as a conserved regulator of NF-κB signaling. J. Immunol. 184, 6188–6198.

Valanne, S., Wang, J. H., Rämet, M., 2011. The *Drosophila* Toll signaling pathway. J. Immunol. 186, 649–656.

Ward, A., Hong, W., Favaloro, V., Luo, L., 2015. Toll receptors instruct axon and dendrite targeting and participate in synaptic partner matching in a *Drosophila* olfactory circuit. Neuron 85, 1013–1028.

Waterhouse, R. M., Kriventseva, E. V., Meister, S., Xi, Z., Alvarez, K. S., Bartholomay, L. C., Barillas-Mury, C., Bian, G., Blandin, S., Christensen, B. M., Dong, Y., Jiang, H., Kanost, M. R., Koutsos, A. C., Levashina, E. A., Li, J., Ligoxygakis, P., Maccallum, R. M., Mayhew, G. F ., Mendes, A., Michel, K., Osta, M. A., Paskewitz, S., Shin, S.

W., Vlachou, D., Wang, L., Wei, W., Zheng, L., Zou, Z., Severson, D. W., Raikhel, A. S., Kafatos, F. C., Dimopoulos, G., Zdobnov, E. M., Christophides, G. K., 2007. Evolutionary dynamics of immune-related genes and pathways in disease-vector mosquitoes. Science 316, 1738–1743.

Wu, S., Zhang, X., Chen, X., Cao, P., Beerntsen, B. T., & Ling, E., 2010. BmToll9, an arthropod conservative Toll, is likely involved in the local gut immune response in the silkworm, *Bombyx mori*. Developmental & Comparative Immunology. 34, 93-96.

Yagi, Y., Nishida, Y., Ip, Y. T., 2010. Functional analysis of *Toll*-related genes in *Drosophila*. Dev. Growth, Differ. 52, 771–783.

Yang, Z., Jiang, H., Zhao, X., Lu, Z., Luo, Z., Li, X., Zhao, J., Zhang, Y., 2017. Correlation of cell surface proteins of distinct *Beauveria bassiana* cell types and adaption to varied environment and interaction with the host insect. Fungal Genet. Biol. 99, 13–25.

Appendix A - Chapter 2 Supplement

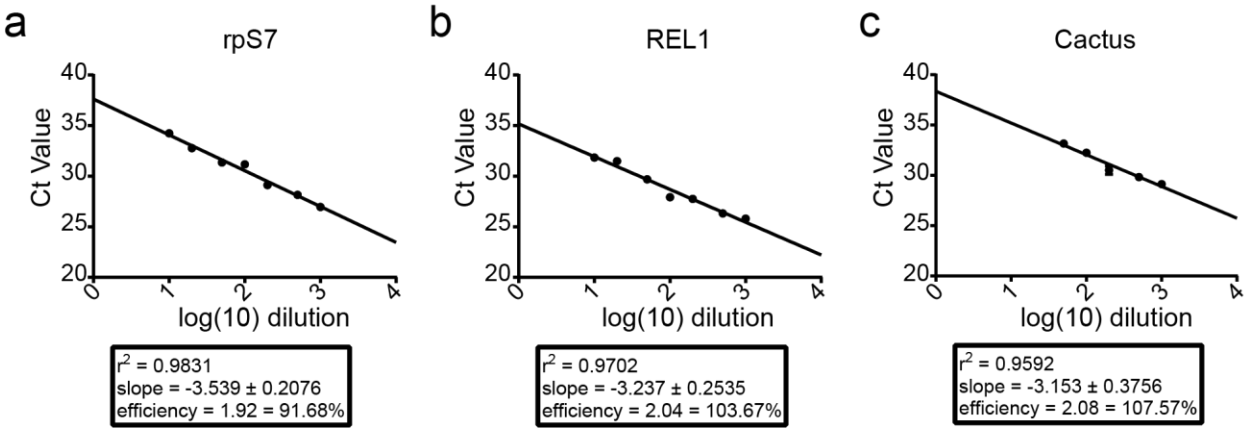

Figure A-1 Primer efficiencies for RT-qPCR analysis

Primer efficiencies were measured using dilutions of the cDNA of untreated 2-3 day-old female *An. gambiae* for RT-qPCR analysis. Graphs show Ct values plotted against log values of the dilutions for ribosomal protein S7 *(rpS7*) (a), and genes of interest, *REL1* (b) and *Cactus* (c). r^2 = goodness of fit.

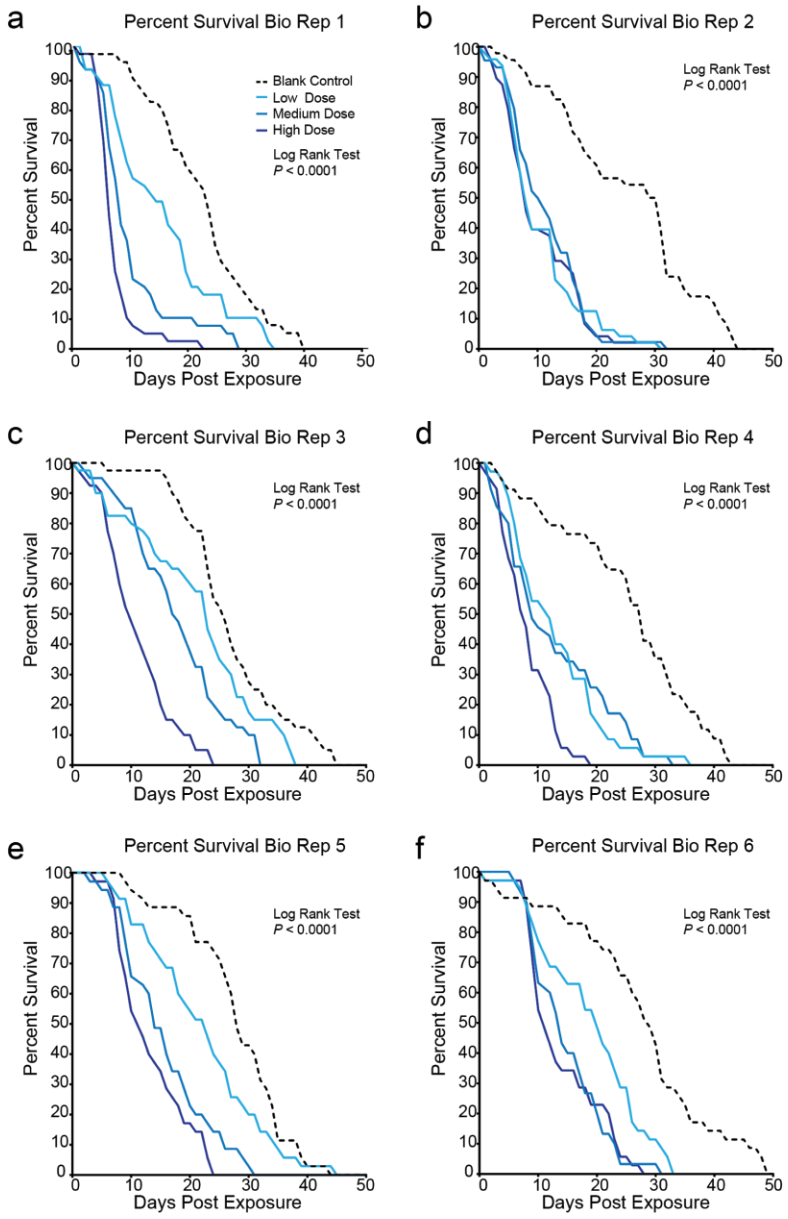

Figure A-2 Individual biological replicates of survival analyses of *An. gambiae* after exposure to *B. bassiana*

(a-f) Individual biological replicates of survival curves after exposure to 0, low, medium, and high *B. bassiana* conidial doses corresponding to data shown in Figure 2-1. Survival curves are presented as percent survival and each biological replicate utilized 35 adult, female mosquitoes

per treatment. Statistical significant differences between survival curves were assessed by Log Rank Test, with resulting *P* values shown in the figure.

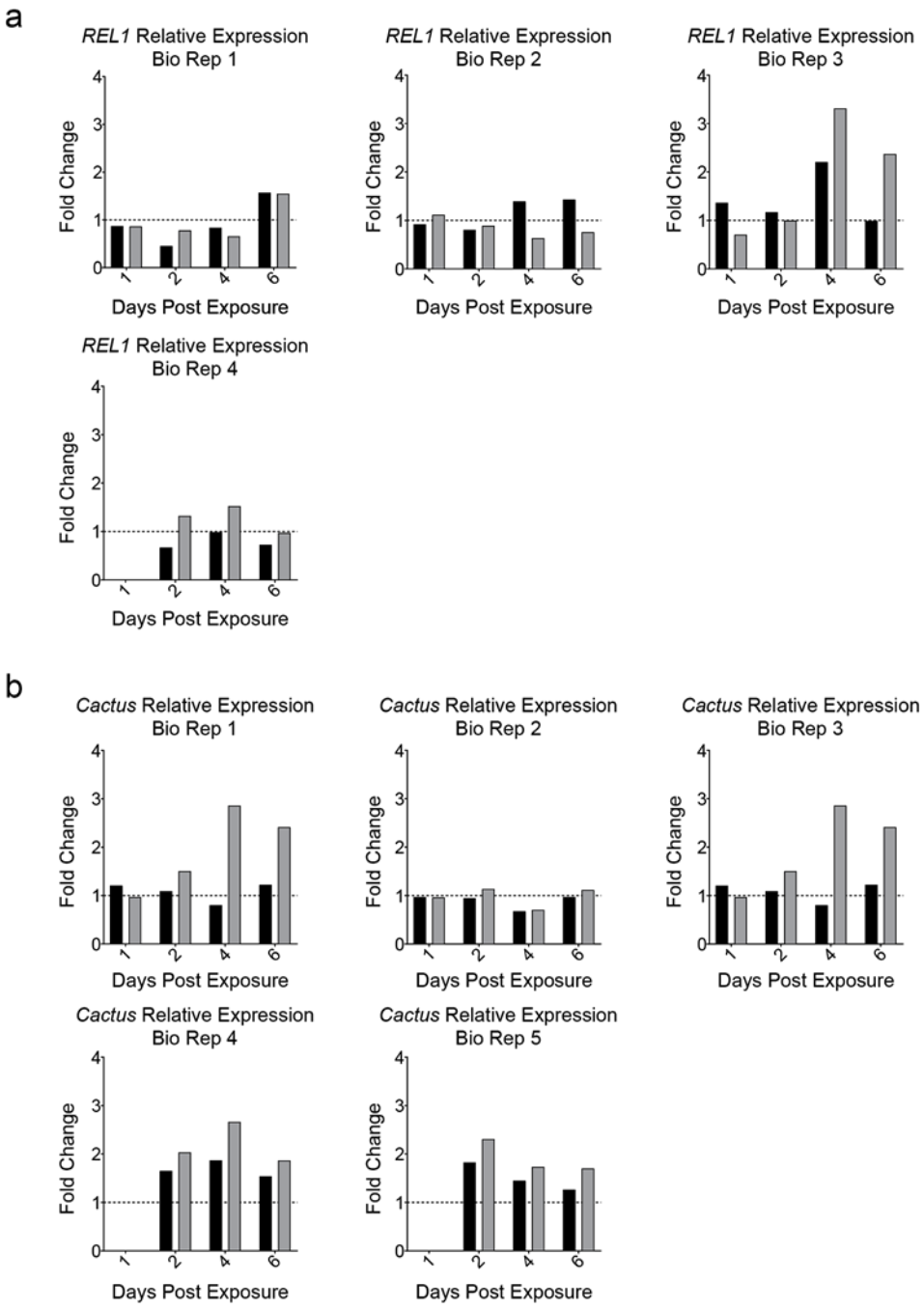

Figure A-3 Individual biological replicates of the *REL1* and *Cactus* relative expression time course following exposure to *B. bassiana*

Individual biological replicates of the relative expression of *An. gambiae* **(a)** *REL1* and **(b)** *Cactus* transcripts after exposure to the high dose of *B. bassiana* conidia in adult, female mosquitoes. Graphs depict transcript levels at 1, 2, 4, and 6 days post exposure for oil-only and

B. bassiana-exposed mosquitoes relative to untreated, day 0 controls. Quantitative RT-qPCR results were analyzed using *rpS7* as the reference gene and untreated mosquitoes as calibrator condition.

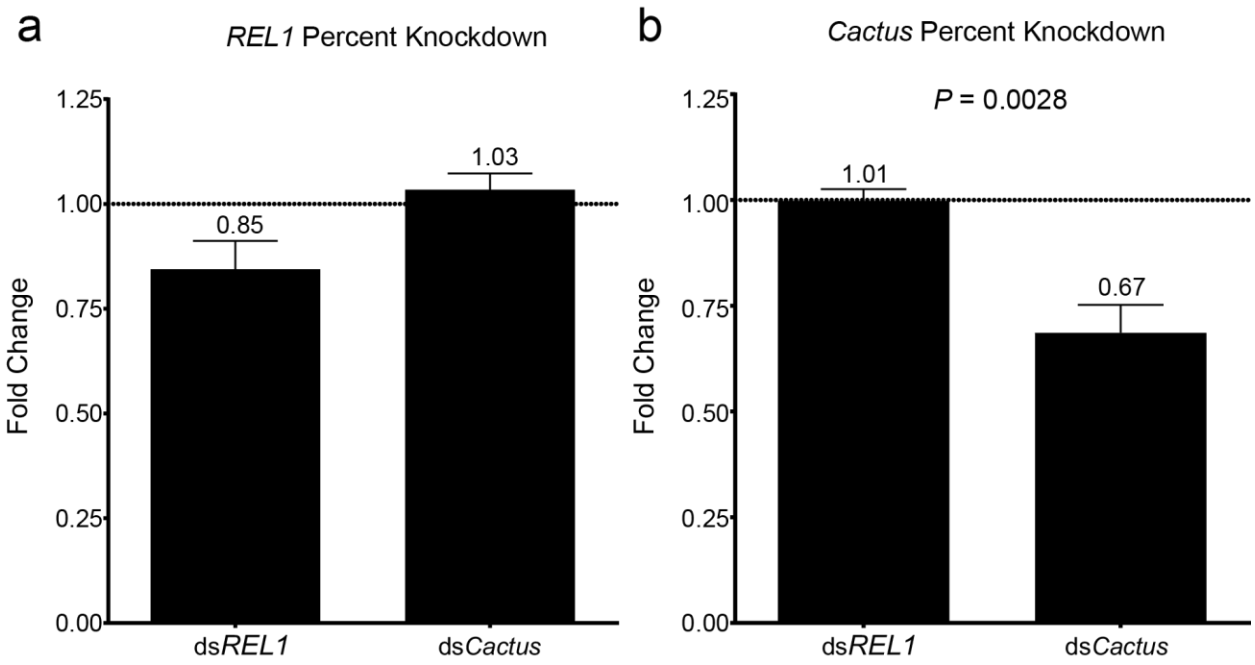

Figure A-4 Percent knockdown of *REL1* and *Cactus* transcripts by RNAi

Mosquitoes were analyzed for knockdown efficiency by quantitative RT-qPCR. Relative expression of *An. gambiae* (**a**) *REL1* and (**b**) *Cactus* transcripts three days after injection of dsRNA [3.0 µg/µL] specific to each transcript in adult, female mosquitoes. Graphs depict mean transcript levels relative to ds*GFP*-injected controls. Quantitative RT-qPCR results were analyzed using *rpS7* as the reference gene and untreated mosquitoes as calibrator condition. Data are presented as mean ± 1 SEM from three biological replicates.

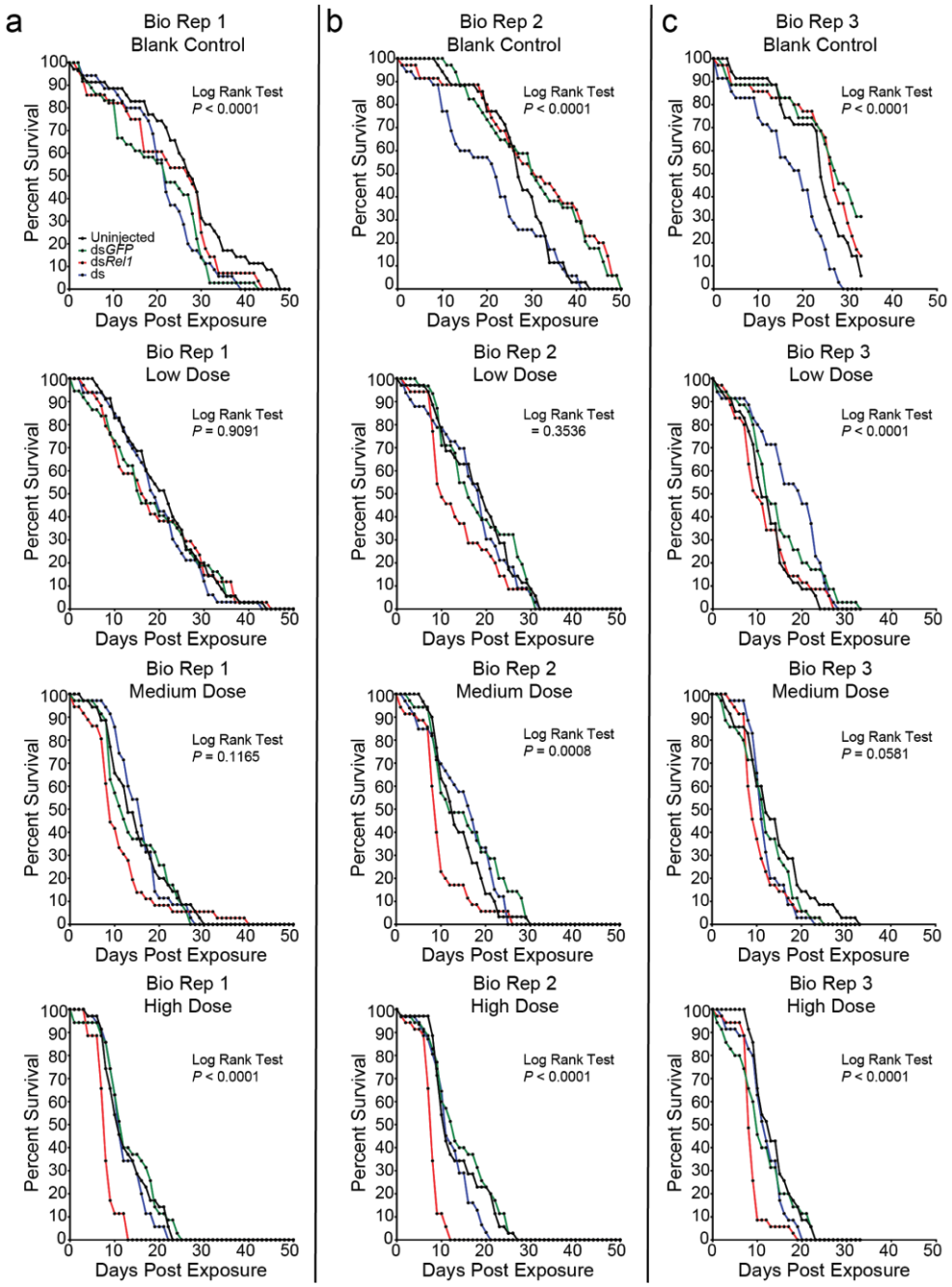

Figure A-5 Individual biological replicates of survival analyses of *An. gambiae* following exposure to *B. bassiana*

(a-c) Figure shows the individual biological replicates of survival curves after RNAi treatments (uninjected, ds*GFP*, ds*REL1*, ds*Cactus*) and exposure to oil-only, low, medium, and high *B. bassiana* conidial doses, corresponding to data shown in Figure 2-3. Survival curves are

presented as percent survival and each biological replicate utilizing 35 female mosquitoes per treatment. Statistically significant differences between survival curves were assessed by Log Rank Test, with resulting *P* values shown in the figure.

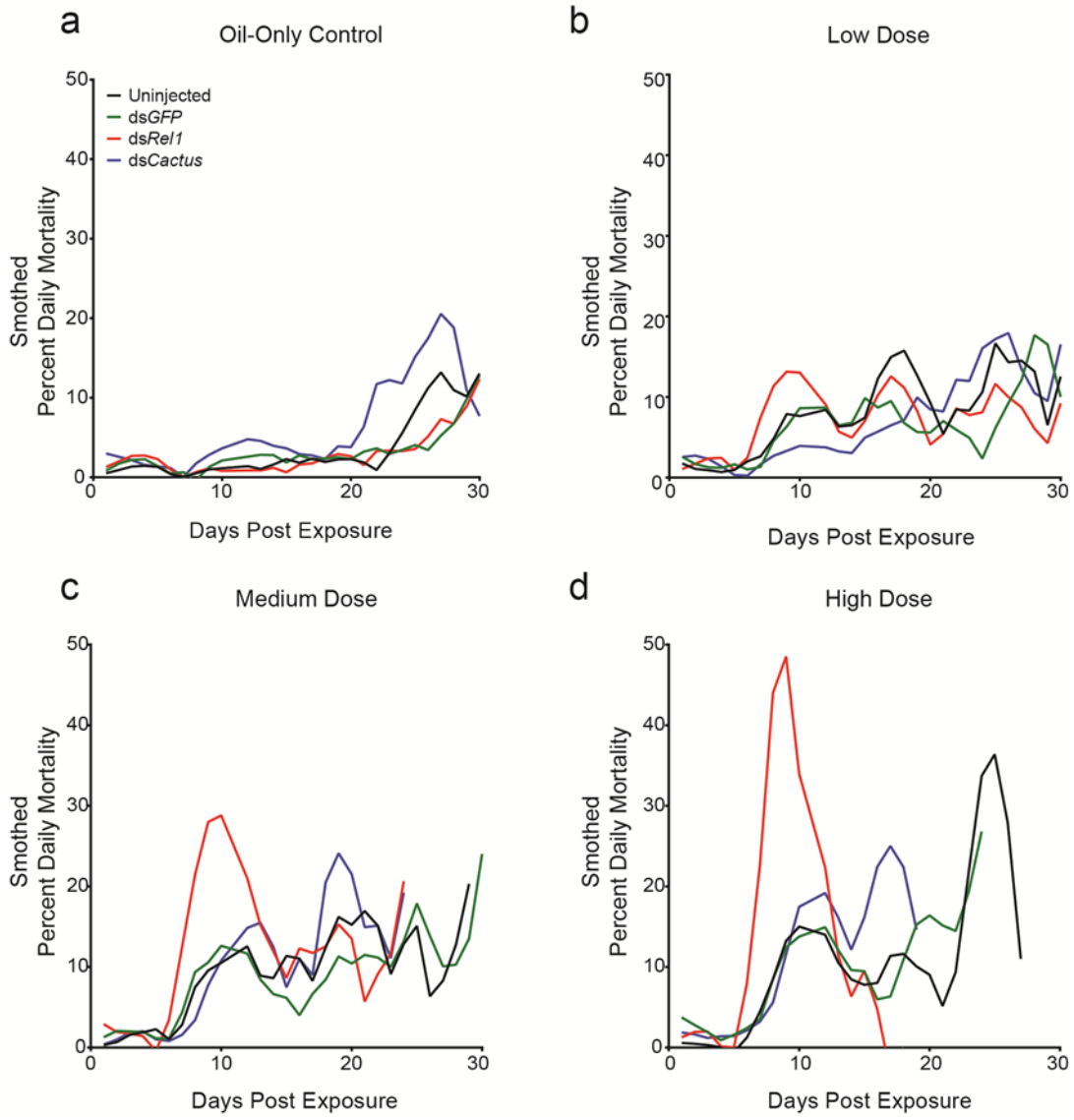

Figure A-6 Smoothed percent daily mortality curves of *An. gambiae* following exposure to *B. bassiana*

(**a-d**) Combined percent daily mortality rate from proportional survivals shown in Figure 3 of the main manuscript. Curves represent the average of three biological replicates. Peaks within ds*REL1*- and ds*Cactus*-injected treatments indicated with an asterisk (first peak) or a square (second peak).

Appendix B - Chapter 3 Supplement

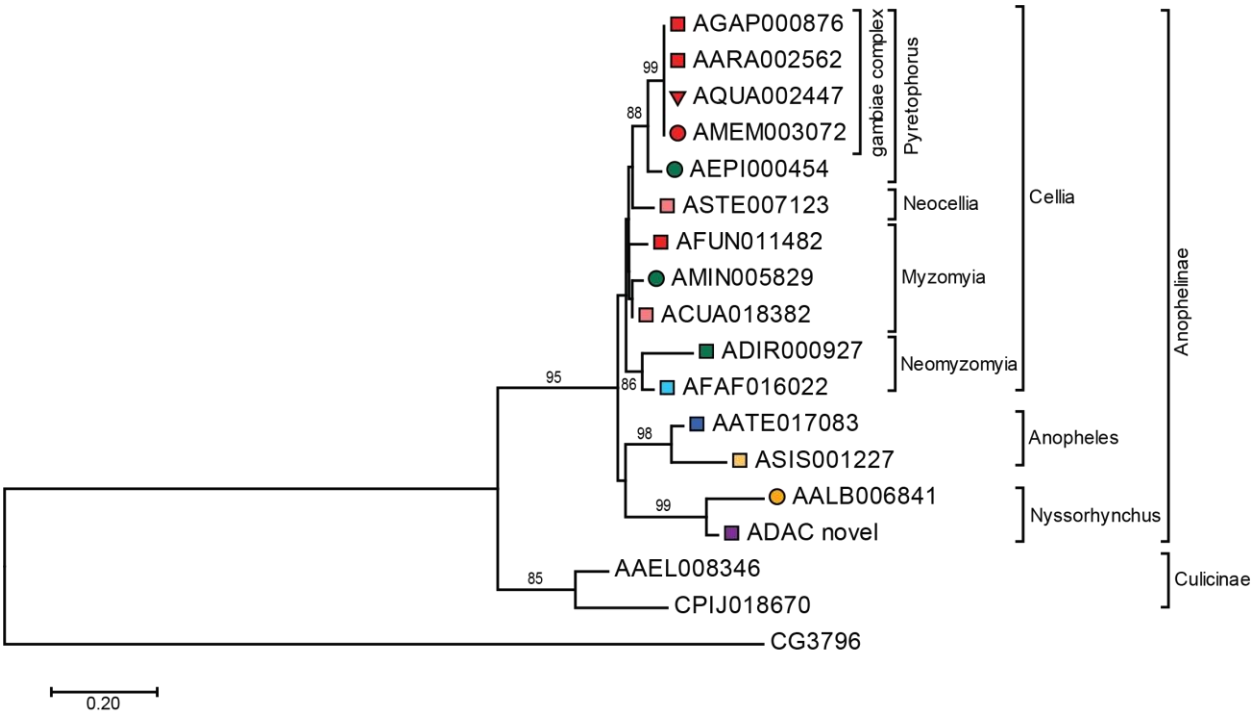

Figure B-1 AC Phylogeny

Phylogenetic tree obtained by maximum-likelihood analysis, reflecting the relationships of Achaete amino acid sequences. Scale bars indicate substitutions per site per unit of branch length and the number at each branch reflects bootstrap percentages (1000 replications). Only branches with support over 75% have values listed. Branch labels are coded according to Neafsey et al. and indicate vector status and geographic distribution of species (square, major vector; circle, minor vector; triangle, nonvector; red, Africa; pink, South Asia; green, South-East Asia; light blue, Asia Pacific; dark blue, Europe; light orange, East Asia; dark orange, Central America; purple, South America).

Tree follows species topology as published in Neafsey et al., 2015 with an exception in the placement of the subgenus *Neomyzomyia* as the sister group to the subgenus *Myzomyia* rather

than its normal placement basally within *Anophelinae*, although low support for this node leaves the true topology unclear. Further details on tree analysis can be found in Supplemental Table 3.

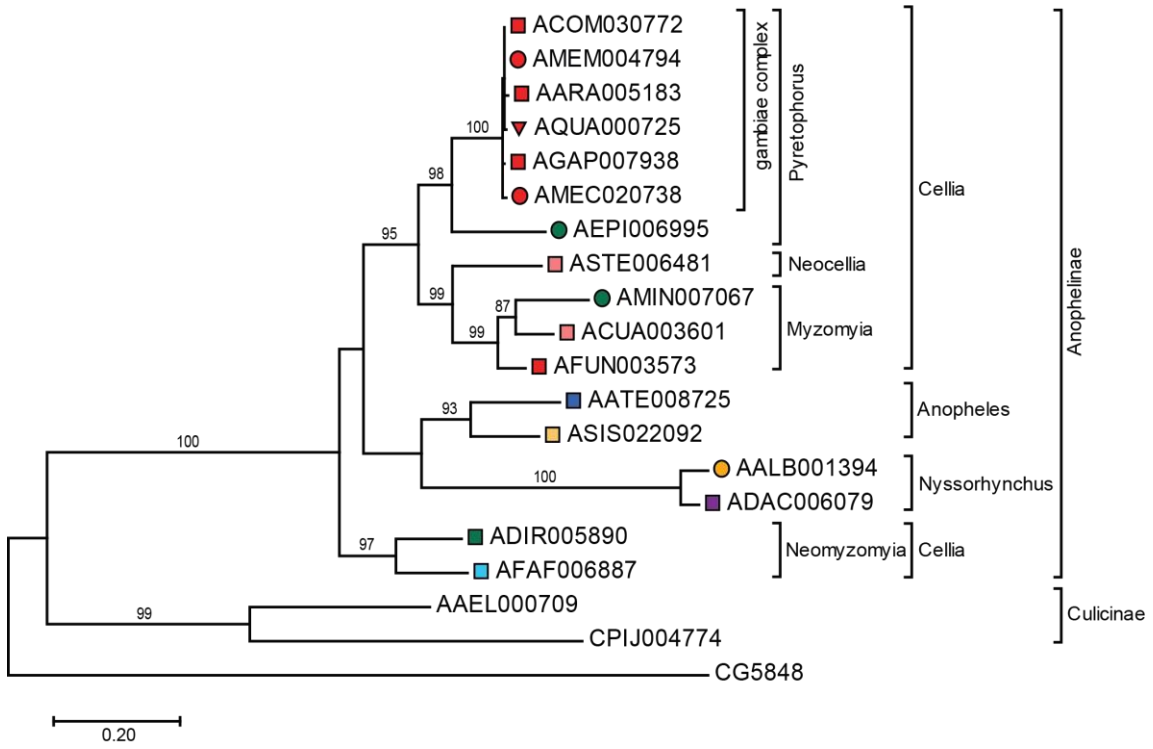

Figure B-2 CACT Phylogeny

Phylogenetic tree obtained by maximum-likelihood analysis, reflecting the relationships of Cactus amino acid sequences. Scale bars indicate substitutions per site per unit of branch length and the number at each branch reflects bootstrap percentages (1000 replications). Only branches with support over 75% have values listed. Branch labels are coded according to Neafsey et al. 2015 and indicate vector status and geographic distribution of species (square, major vector; circle, minor vector; triangle, nonvector; red, Africa; pink, South Asia; green, South-East Asia; light blue, Asia Pacific; dark blue, Europe; light orange, East Asia; dark orange, Central America; purple, South America).

Subgenus *Anopheles* is the sister group to series *Nyssorhynchus* rather than subgenus *Cellia*, although low support for this branch leaves the true topology unclear. Additionally, *An. dirus* and *An. farauti*, while remaining closely related to each other, branch basally within

Anophelinae rather than being placed basally within *Cellia*. Further details on tree analysis can be found in Supplemental Table 3.

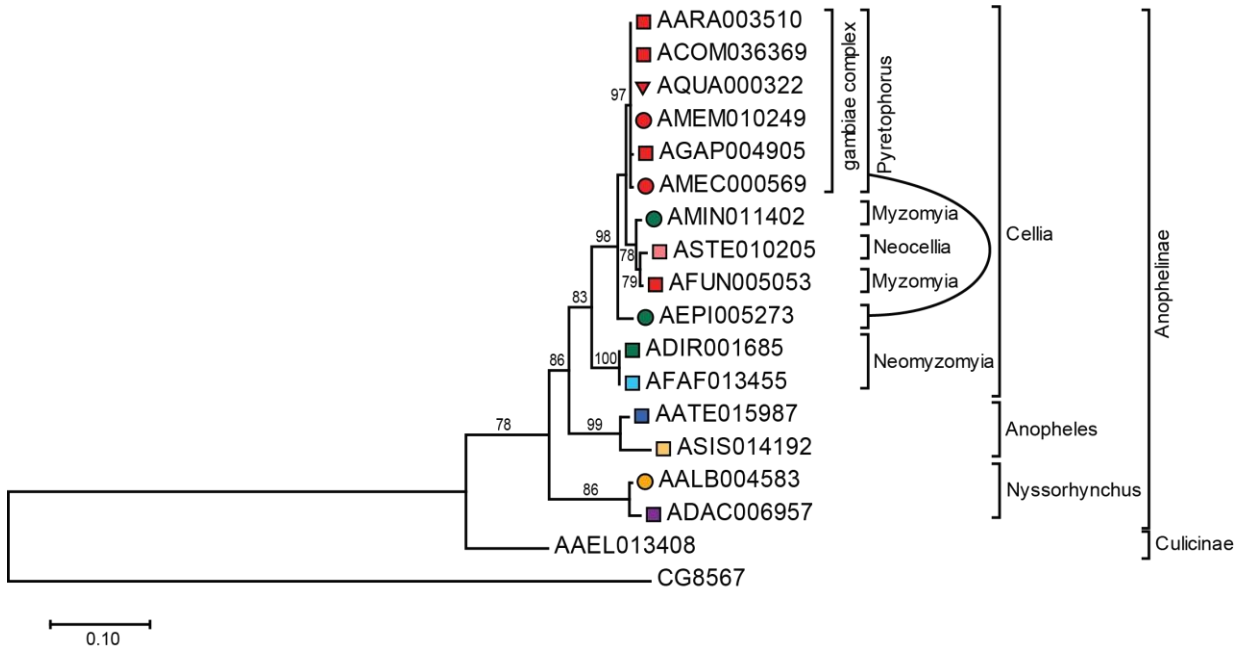

Figure B-3 DEAF1 Phylogeny

Phylogenetic tree obtained by maximum-likelihood analysis, reflecting the relationships of Deaf1 amino acid sequences. Scale bars indicate substitutions per site per unit of branch length and the number at each branch reflects bootstrap percentages (1000 replications). Only branches with support over 75% have values listed. Branch labels are coded according to Neafsey et al. 2015 and indicate vector status and geographic distribution of species (square, major vector; circle, minor vector; triangle, nonvector; red, Africa; pink, South Asia; green, South-East Asia; light blue, Asia Pacific; dark blue, Europe; light orange, East Asia; dark orange, Central America; purple, South America).

Series *Pyretophorus* and *Myzomyia* are not supported, as *An. epiroticus* (AEPI005273) is basal to the *Cellia*, differing from the species topology placement of *An. epiroticus* basal to the *gambiae* complex. In addition, *An. stephensi* is placed as the within the series *Myzomyia*, rather than its normal placement as a sister group to this group. Further details on tree analysis can be found in Supplemental Table 3.

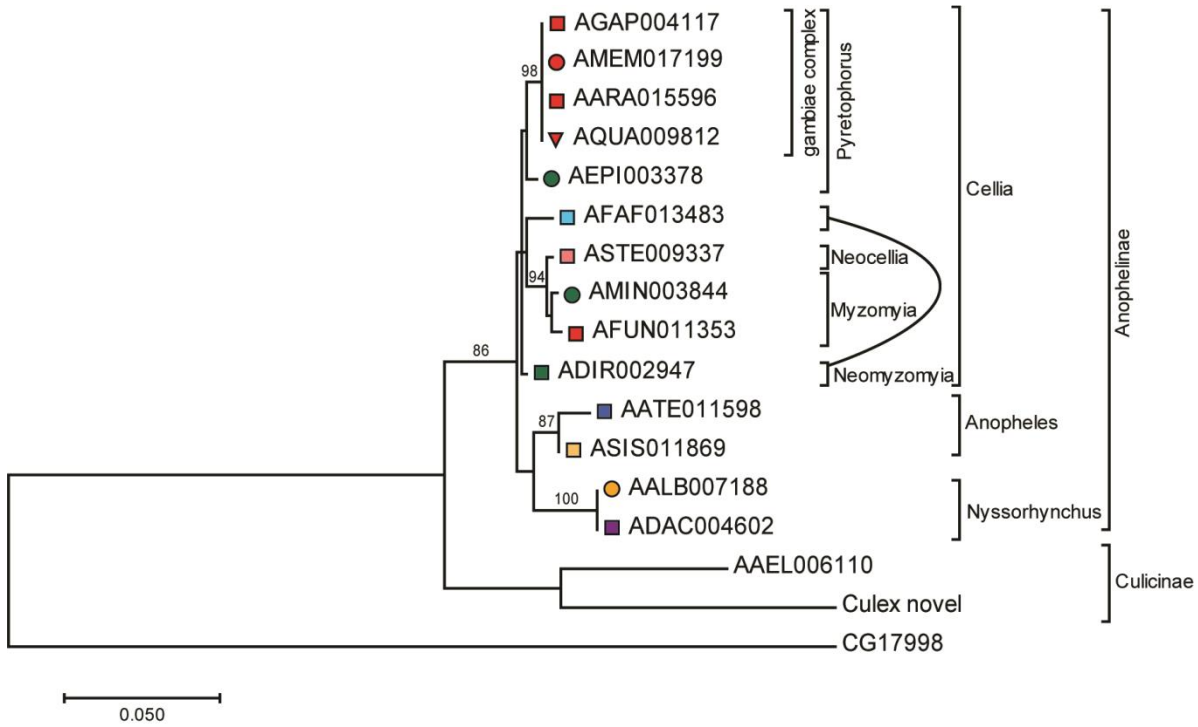

Figure B-4 GPRK2 Phylogeny

Phylogenetic tree obtained by maximum-likelihood analysis, reflecting the relationships of Gprk2 amino acid sequences. Scale bars indicate substitutions per site per unit of branch length and the number at each branch reflects bootstrap percentages (1000 replications). Only branches with support over 75% have values . Branch labels are coded according to Neafsey et al. 2015 and indicate vector status and geographic distribution of species (square, major vector; circle, minor vector; triangle, nonvector; red, Africa; pink, South Asia; green, South-East Asia; light blue, Asia Pacific; dark blue, Europe; light orange, East Asia; dark orange, Central America; purple, South America).

The overall amino acid identity of these sequences within the anophelines (97.83%) prevents accurate prediction of sequence evolution. Further details on tree analysis can be found in Supplemental Table 3.

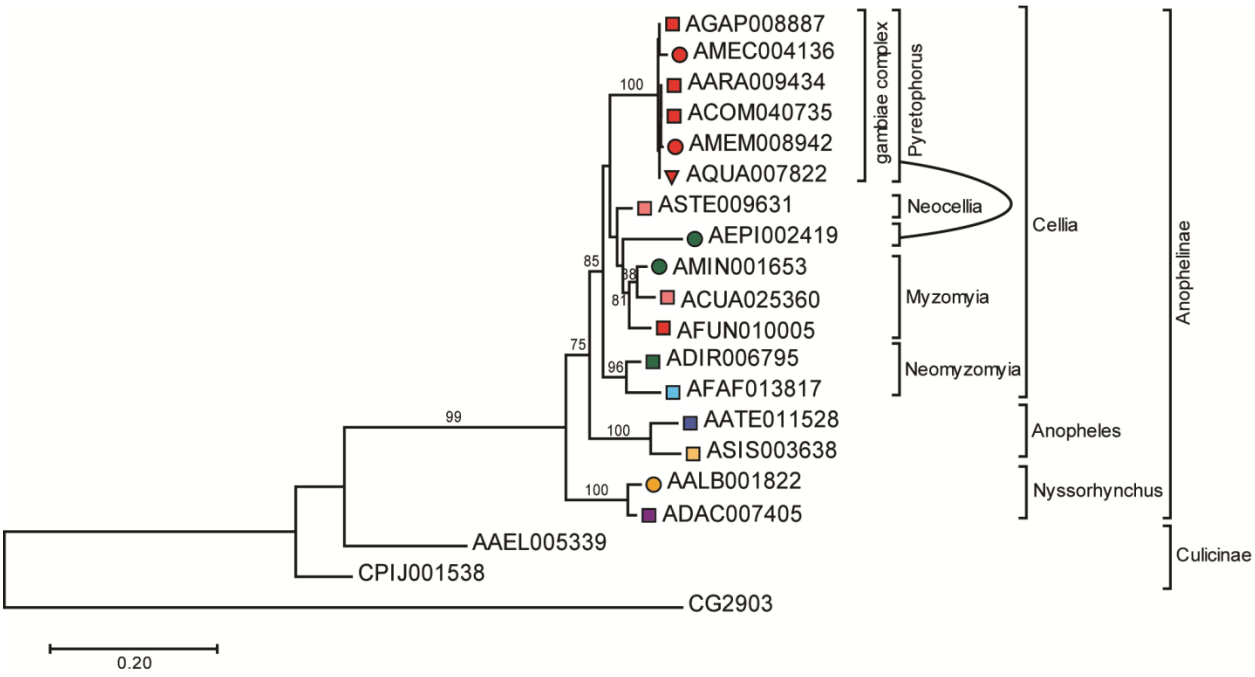

Figure B-5 HRS Phylogeny

Phylogenetic tree obtained by maximum-likelihood analysis, reflecting the relationships of Hrs amino acid sequences. Scale bars indicate substitutions per site per unit of branch length and the number at each branch reflects bootstrap percentages (1000 replications). Only branches with support over 75% have values listed Branch labels are coded according to Neafsey et al. 2015 and indicate vector status and geographic distribution of species (square, major vector; circle, minor vector; triangle, nonvector; red, Africa; pink, South Asia; green, South-East Asia; light blue, Asia Pacific; dark blue, Europe; light orange, East Asia; dark orange, Central America; purple, South America).

Series *Pyretophorus* is not supported, as *An. epiroticus* (AEPI002419) is placed between series *Neocellia* and *Myzomyia*. This placement, however, is not bootstrap supported and as such leaves true topology unclear. Further details on tree analysis can be found in Supplemental Table 3.

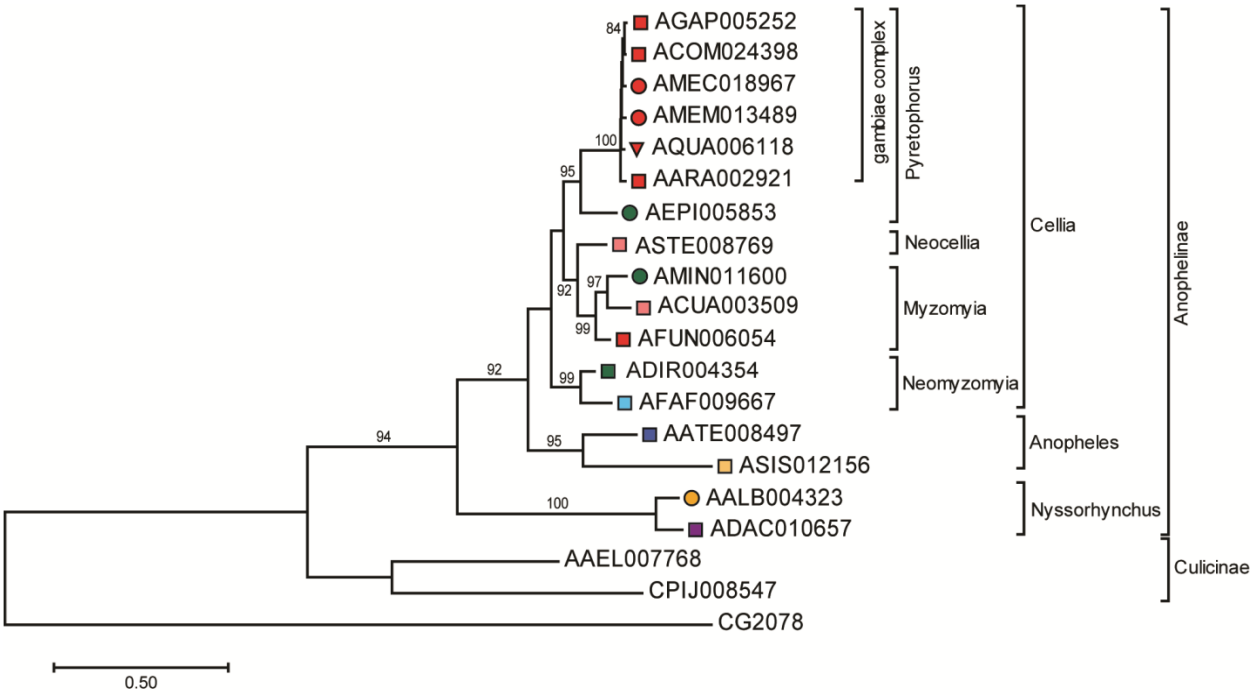

Figure B-6 MYD88 Phylogeny

Phylogenetic tree obtained by maximum-likelihood analysis, reflecting the relationships of Myd88 amino acid sequences. Scale bars indicate substitutions per site per unit of branch length and the number at each branch reflects bootstrap percentages (1000 replications). Only branches with support over 75% have values listed. Branch labels are coded according to Neafsey et al. 2015 and indicate vector status and geographic distribution of species (square, major vector; circle, minor vector; triangle, nonvector; red, Africa; pink, South Asia; green, South-East Asia; light blue, Asia Pacific; dark blue, Europe; light orange, East Asia; dark orange, Central America; purple, South America).

Tree topology fully follows published species phylogeny, with strong bootstrap support (>75%) indicated at branch points. Further details on tree analysis can be found in Supplemental Table 3.

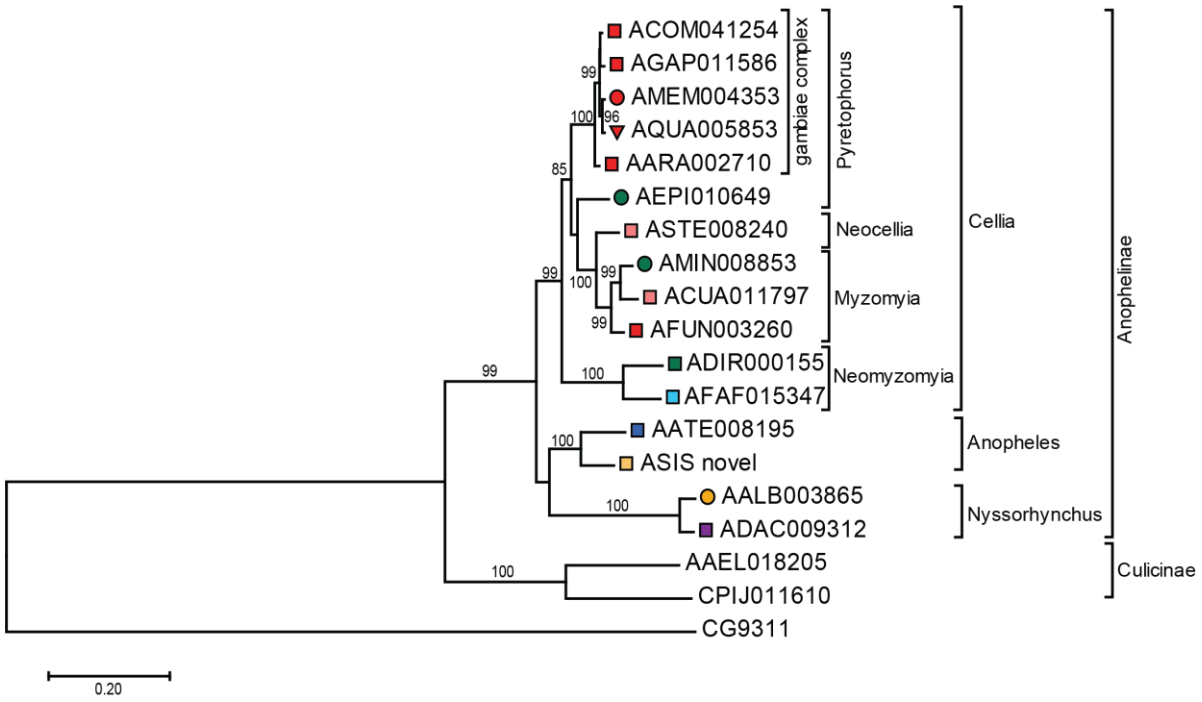

Figure B-7 MOP Phylogeny

Phylogenetic tree obtained by maximum-likelihood analysis, reflecting the relationships of Myopic amino acid sequences. Scale bars indicate substitutions per site per unit of branch length and the number at each branch reflects bootstrap percentages (1000 replications). Only branches with support over 75% have values listed. Branch labels are coded according to Neafsey et al. 2015 and indicate vector status and geographic distribution of species (square, major vector; circle, minor vector; triangle, nonvector; red, Africa; pink, South Asia; green, South-East Asia; light blue, Asia Pacific; dark blue, Europe; light orange, East Asia; dark orange, Central America; purple, South America).

Tree topology fully follows published species phylogeny, with strong bootstrap support (>75%) indicated at branch points. Further details on tree analysis can be found in Supplemental Table 3.

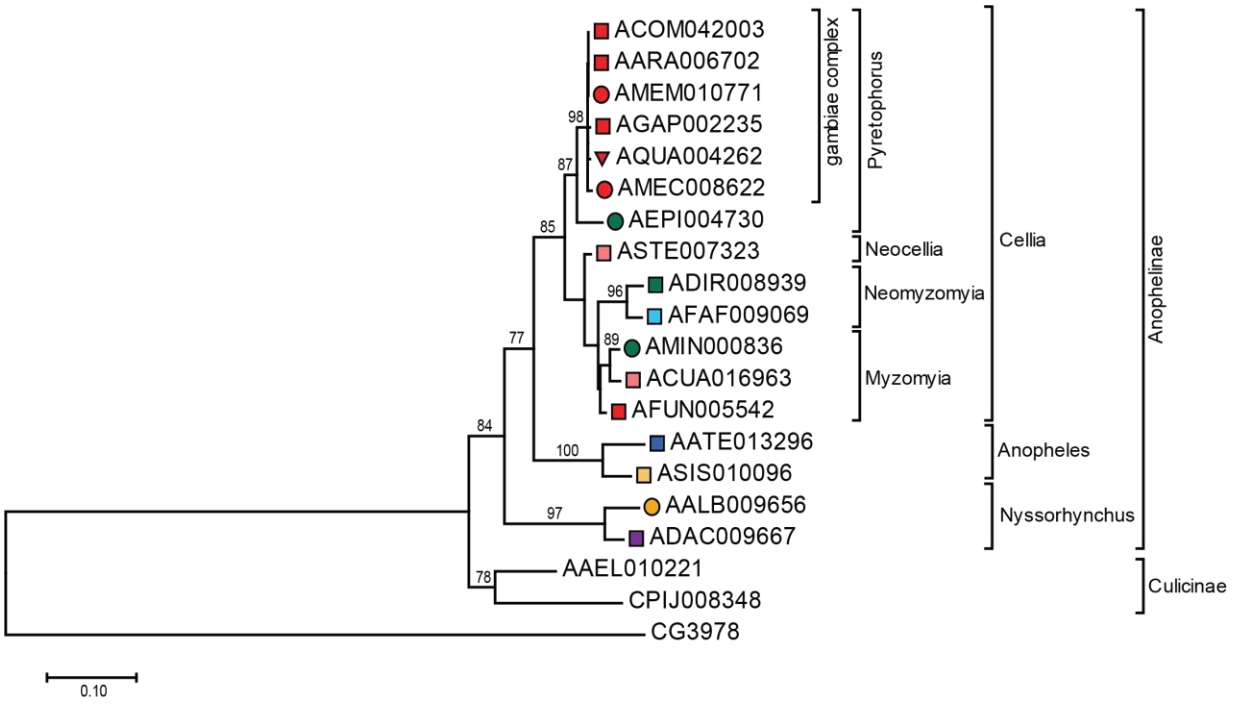

Figure B-8 PNR Phylogeny

Phylogenetic tree obtained by maximum-likelihood analysis, reflecting the relationships of Pannier amino acid sequences. Scale bars indicate substitutions per site per unit of branch length and the number at each branch reflects bootstrap percentages (1000 replications). Only branches with support over 75% have values listed. Branch labels are coded according to Neafsey et al. 2015 and indicate vector status and geographic distribution of species (square, major vector; circle, minor vector; triangle, nonvector; red, Africa; pink, South Asia; green, South-East Asia; light blue, Asia Pacific; dark blue, Europe; light orange, East Asia; dark orange, Central America; purple, South America).

Series *Neomyzomyia* is placed as the sister group to series *Myzomyia*, disrupting the usual species topology. However, this placement is not bootstrap supported, and as such true topology remains unclear. Further details on tree analysis can be found in Supplemental Table 3.

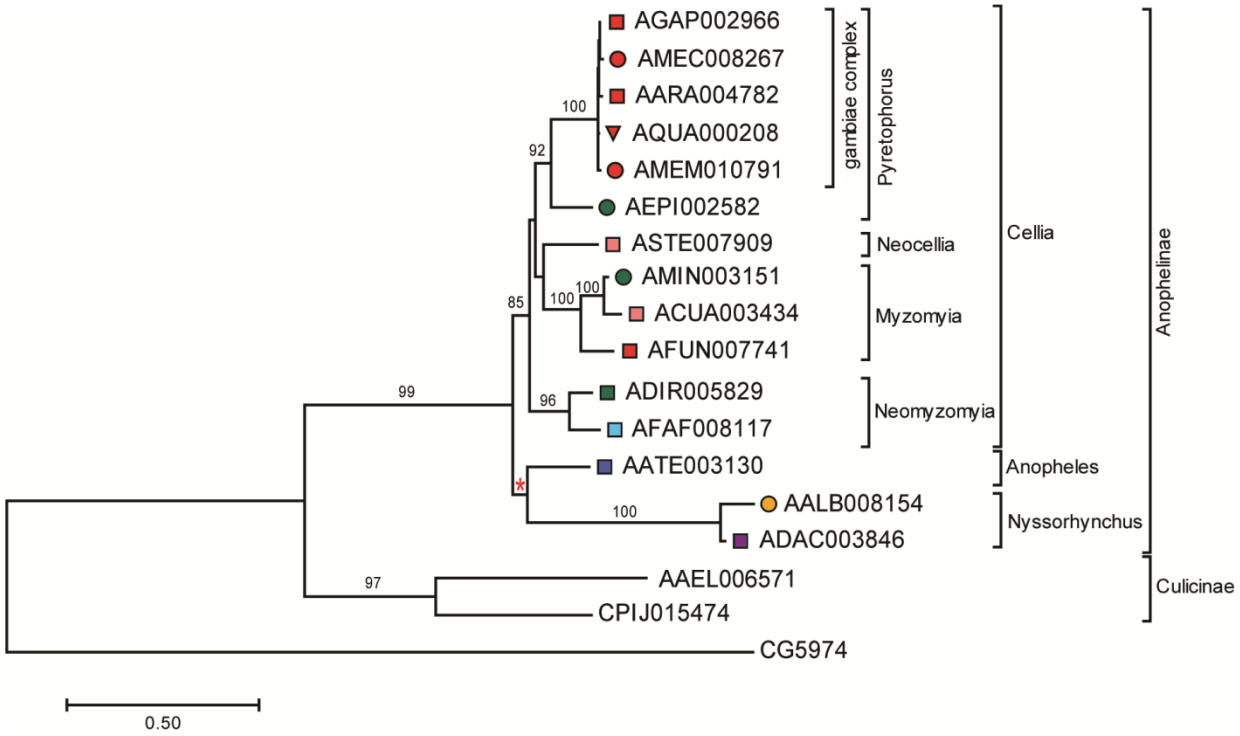

Figure B-9 PELLE Phylogeny

Phylogenetic tree obtained by maximum-likelihood analysis, reflecting the relationships of Pelle amino acid sequences. Scale bars indicate substitutions per site per unit of branch length and the number at each branch reflects bootstrap percentages (1000 replications). Only branches with support over 75% have values listed. Branch labels are coded according to Neafsey et al. 2015 and indicate vector status and geographic distribution of species (square, major vector; circle, minor vector; triangle, nonvector; red, Africa; pink, South Asia; green, South-East Asia; light blue, Asia Pacific; dark blue, Europe; light orange, East Asia; dark orange, Central America; purple, South America).

An. atroparvus in the subgenus *Anopheles* is places as the sister group to *Nyssorhynchus* as opposed to its normal placement basally to *Cellia.* However, low bootstrap support at this node (red asterisk) makes true topology unclear. Further details on tree analysis can be found in Supplemental Table 3.

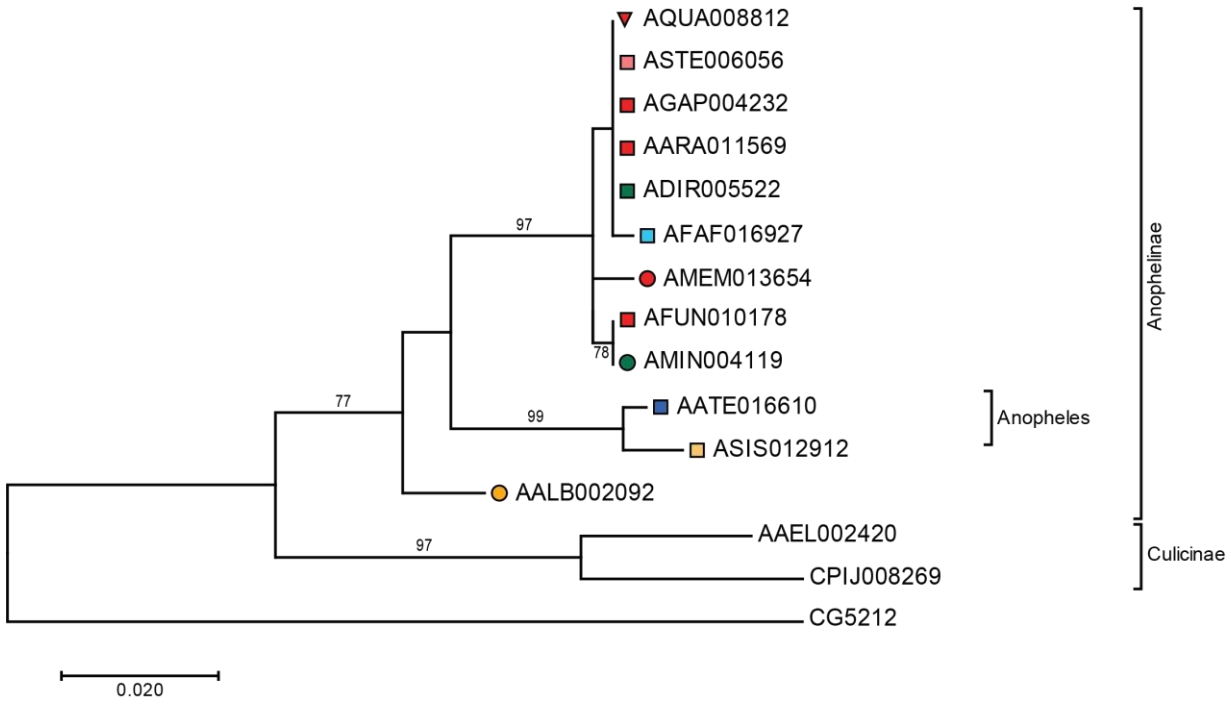

Figure B-10 PLI Phylogeny

Phylogenetic tree obtained by maximum-likelihood analysis, reflecting the relationships of Pellino amino acid sequences. Scale bars indicate substitutions per site per unit of branch length and the number at each branch reflects bootstrap percentages (1000 replications). Only branches with support over 75% have values listed. Branch labels are coded according to Neafsey et al. 2015 and indicate vector status and geographic distribution of species (square, major vector; circle, minor vector; triangle, nonvector; red, Africa; pink, South Asia; green, South-East Asia; light blue, Asia Pacific; dark blue, Europe; light orange, East Asia; dark orange, Central America; purple, South America).

Due to the highly conserved nature of Pellino within anophelines (97.95% amino acid identity), tree topology is not resolved. Further details on tree analysis can be found in Supplemental Table 3.

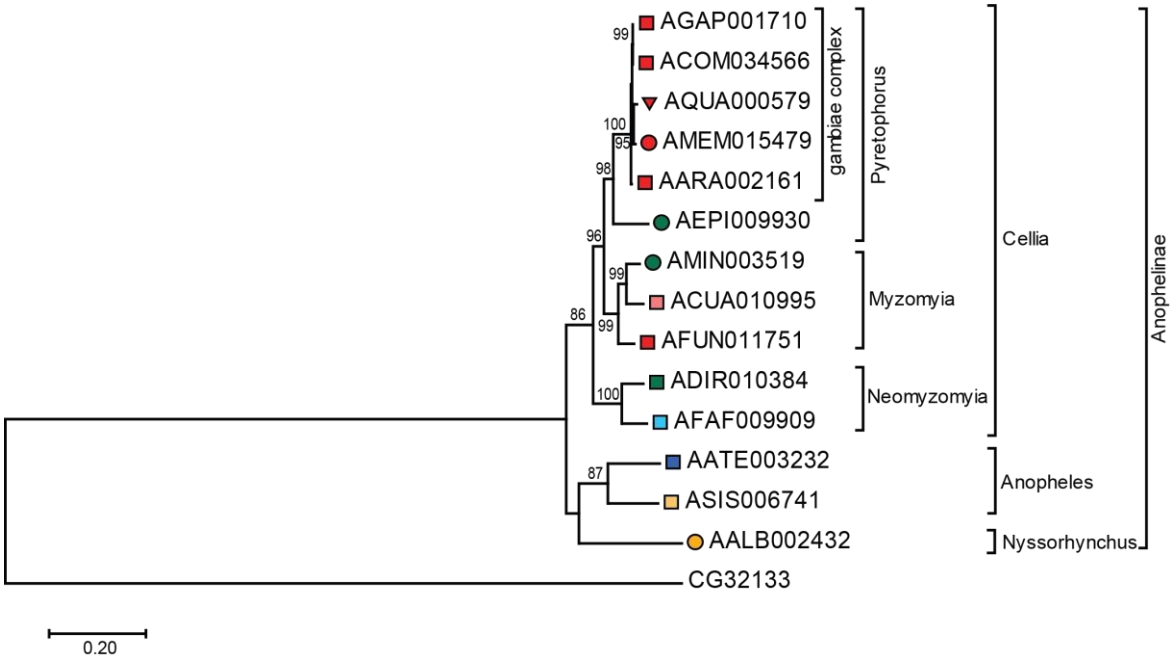

Figure B-11 PTIP Phylogeny

Phylogenetic tree obtained by maximum-likelihood analysis, reflecting the relationships of Ptip amino acid sequences. Scale bars indicate substitutions per site per unit of branch length and the number at each branch reflects bootstrap percentages (1000 replications). Only branches with support over 75% have values listed. Branch labels are coded according to Neafsey et al. 2015 and indicate vector status and geographic distribution of species (square, major vector; circle, minor vector; triangle, nonvector; red, Africa; pink, South Asia; green, South-East Asia; light blue, Asia Pacific; dark blue, Europe; light orange, East Asia; dark orange, Central America; purple, South America).

Subgenus *Anopheles* is placed as the sister group to *Nyssorhynchus*, rather than to the subgenus *Cellia*, however low bootstrap support makes true placement of this clade unclear. Further details on tree analysis can be found in Supplemental Table 3.

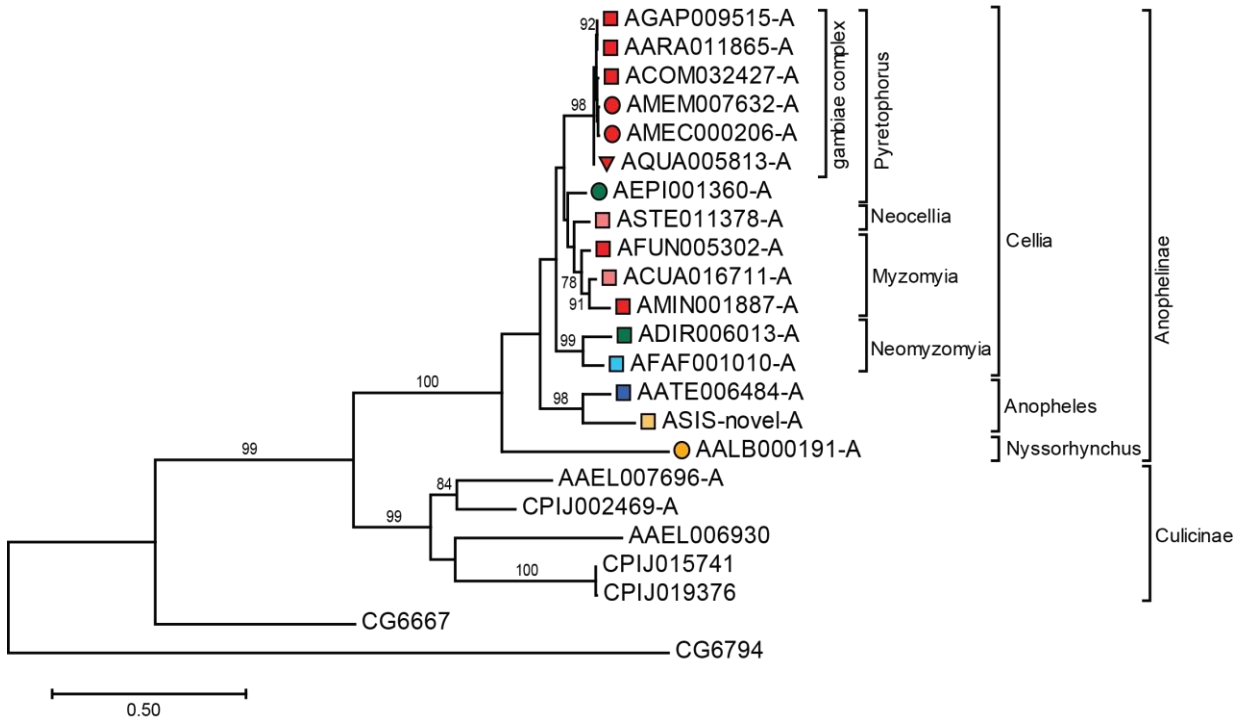

Figure B-12 REL1-A Phylogeny

Phylogenetic tree obtained by maximum-likelihood analysis, reflecting the relationships of REL1-A amino acid sequences. Scale bars indicate substitutions per site per unit of branch length and the number at each branch reflects bootstrap percentages (1000 replications). Only branches with support over 75% have values listed. Branch labels are coded according to Neafsey et al. 2015 and indicate vector status and geographic distribution of species (square, major vector; circle, minor vector; triangle, nonvector; red, Africa; pink, South Asia; green, South-East Asia; light blue, Asia Pacific; dark blue, Europe; light orange, East Asia; dark orange, Central America; purple, South America).

Tree topology fully follows published species phylogeny, with strong bootstrap support (>75%) indicated at branch points. Additionally, it appears that ancient duplications in REL1 within Culicinae has led to single exon genes AAEL006930, CPIJ015741, and CPIJ0193976,

which remain similar in sequence to the A splice isoform of REL1 within mosquitoes. Further details on tree analysis can be found in Supplemental Table 3.

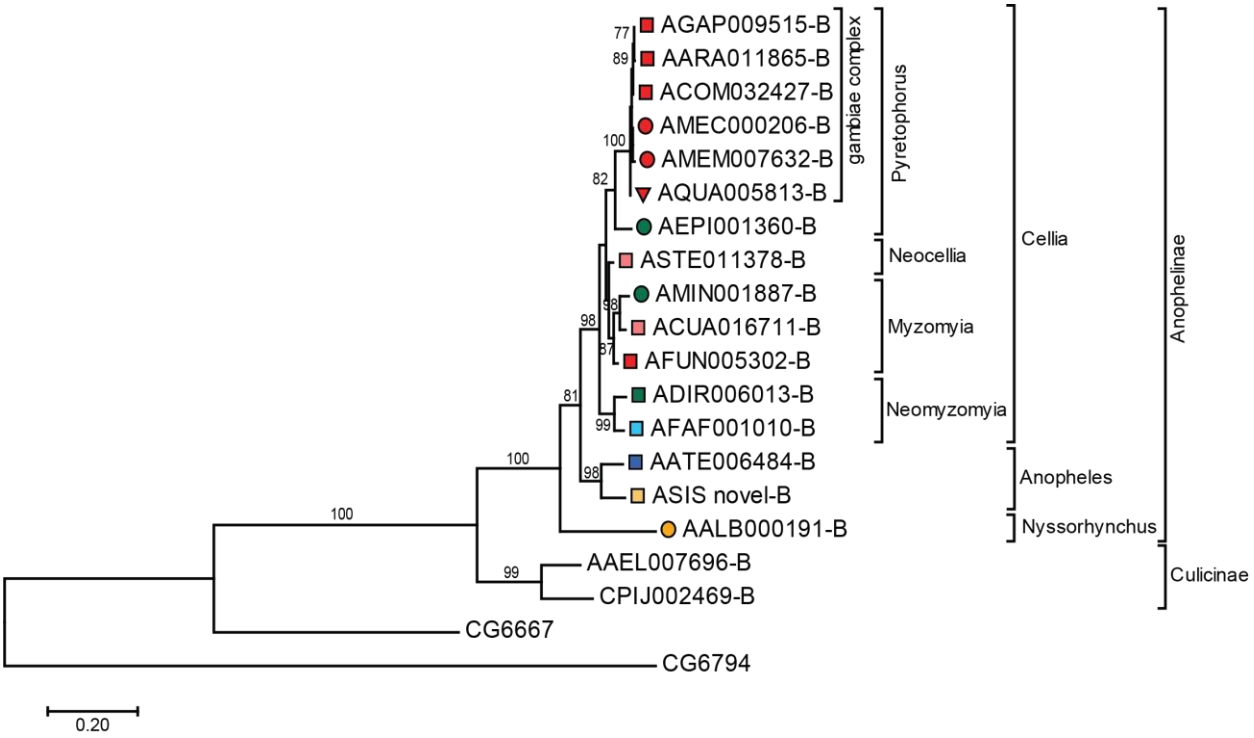

Figure B-13 REL1-B Phylogeny

Phylogenetic tree obtained by maximum-likelihood analysis, reflecting the relationships of REL1-B amino acid sequences. Scale bars indicate substitutions per site per unit of branch length and the number at each branch reflects bootstrap percentages (1000 replications). Only branches with support over 75% have values listed. Branch labels are coded according to Neafsey et al. 2015 and indicate vector status and geographic distribution of species (square, major vector; circle, minor vector; triangle, nonvector; red, Africa; pink, South Asia; green, South-East Asia; light blue, Asia Pacific; dark blue, Europe; light orange, East Asia; dark orange, Central America; purple, South America).

Tree topology fully follows published species phylogeny, with strong bootstrap support (>75%) indicated at branch points. Further details on tree analysis can be found in Supplemental Table 3.

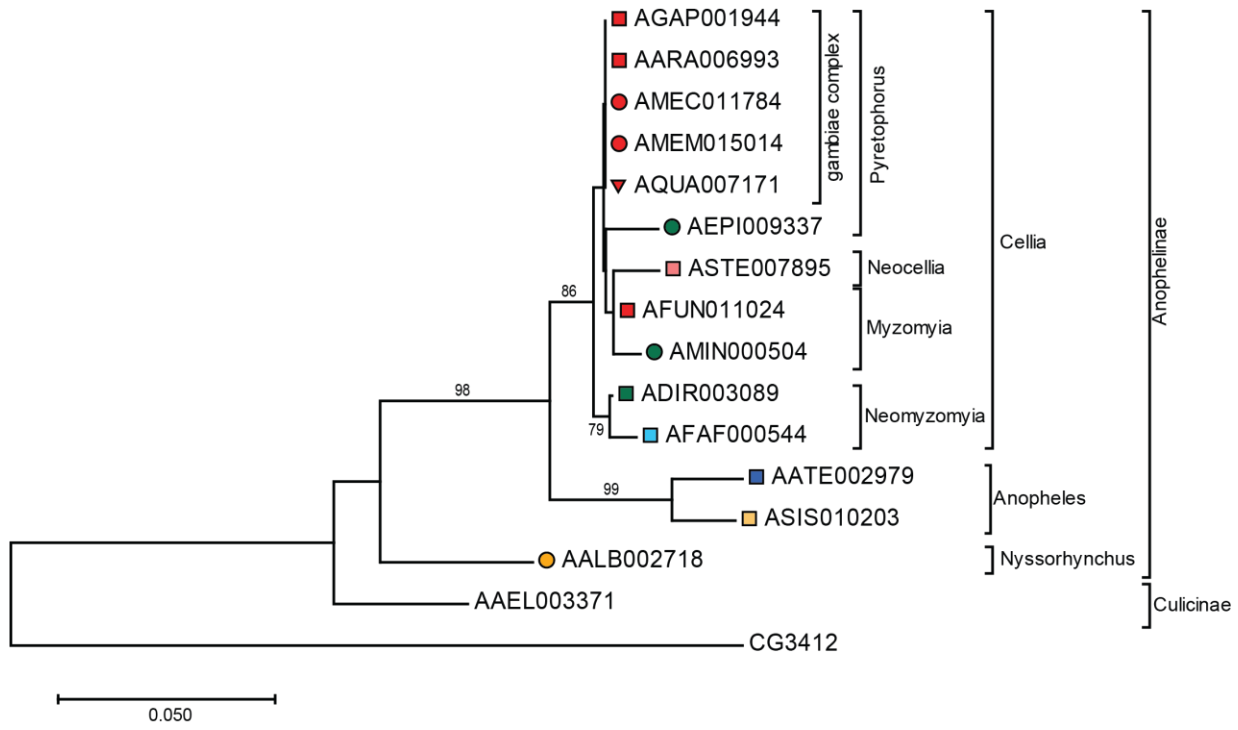

Figure B-14 SLMB Phylogeny

Phylogenetic tree obtained by maximum-likelihood analysis, reflecting the relationships of Slimb amino acid sequences. Scale bars indicate substitutions per site per unit of branch length and the number at each branch reflects bootstrap percentages (1000 replications). Only branches with support over 75% have values listed. Branch labels are coded according to Neafsey et al. 2015 and indicate vector status and geographic distribution of species (square, major vector; circle, minor vector; triangle, nonvector; red, Africa; pink, South Asia; green, South-East Asia; light blue, Asia Pacific; dark blue, Europe; light orange, East Asia; dark orange, Central America; purple, South America).

The overall amino acid identity of these sequences within the anophelines (94.2%) prevents accurate prediction of sequence evolution. Further details on tree analysis can be found in Supplemental Table 3.

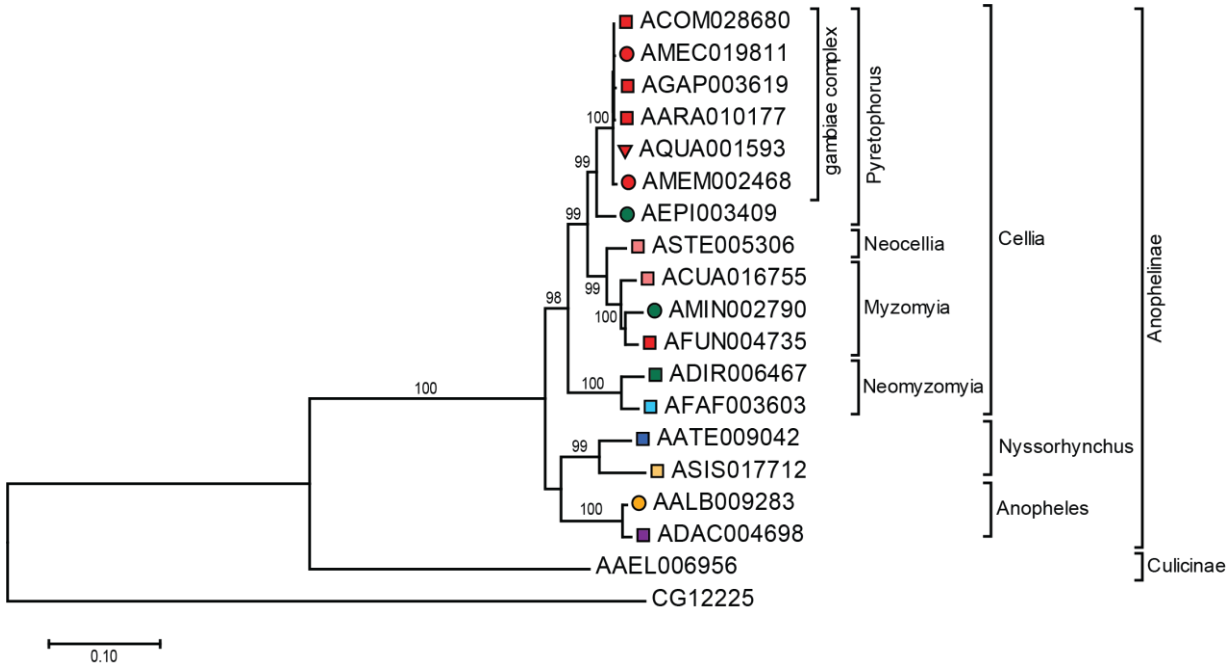

Figure B-15 SPT6 Phylogeny

Phylogenetic tree obtained by maximum-likelihood analysis, reflecting the relationships of Spt6 amino acid sequences. Scale bars indicate substitutions per site per unit of branch length and the number at each branch reflects bootstrap percentages (1000 replications). Only branches with support over 75% have values listed. Branch labels are coded according to Neafsey et al. 2015 and indicate vector status and geographic distribution of species (square, major vector; circle, minor vector; triangle, nonvector; red, Africa; pink, South Asia; green, South-East Asia; light blue, Asia Pacific; dark blue, Europe; light orange, East Asia; dark orange, Central America; purple, South America).

Tree topology fully follows published species phylogeny, with strong bootstrap support (>75%) indicated at branch points. Further details on tree analysis can be found in Supplemental Table 3.

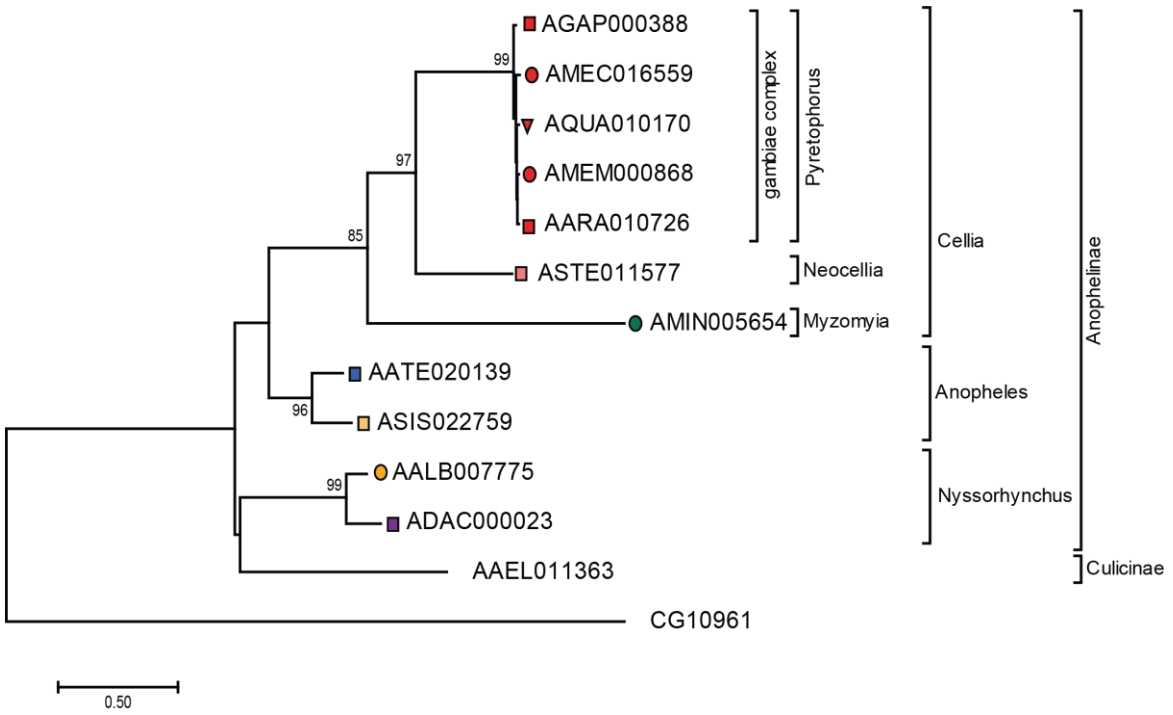

Figure B-16 TRAF6 Phylogeny

Phylogenetic tree obtained by maximum-likelihood analysis, reflecting the relationships of TRAF6 amino acid sequences. Scale bars indicate substitutions per site per unit of branch length and the number at each branch reflects bootstrap percentages (1000 replications). Only branches with support over 75% have values listed. Branch labels are coded according to Neafsey et al. 2015 and indicate vector status and geographic distribution of species (square, major vector; circle, minor vector; triangle, nonvector; red, Africa; pink, South Asia; green, South-East Asia; light blue, Asia Pacific; dark blue, Europe; light orange, East Asia; dark orange, Central America; purple, South America).

Ae. aegypti is placed as the sister group to subgenus *Nyssorhynchus* in this tree topology. However, this placement is not corroborated by strong bootstrap support, and thus a determination on true topology cannot be made. Further details on tree analysis can be found in Supplemental Table 3.

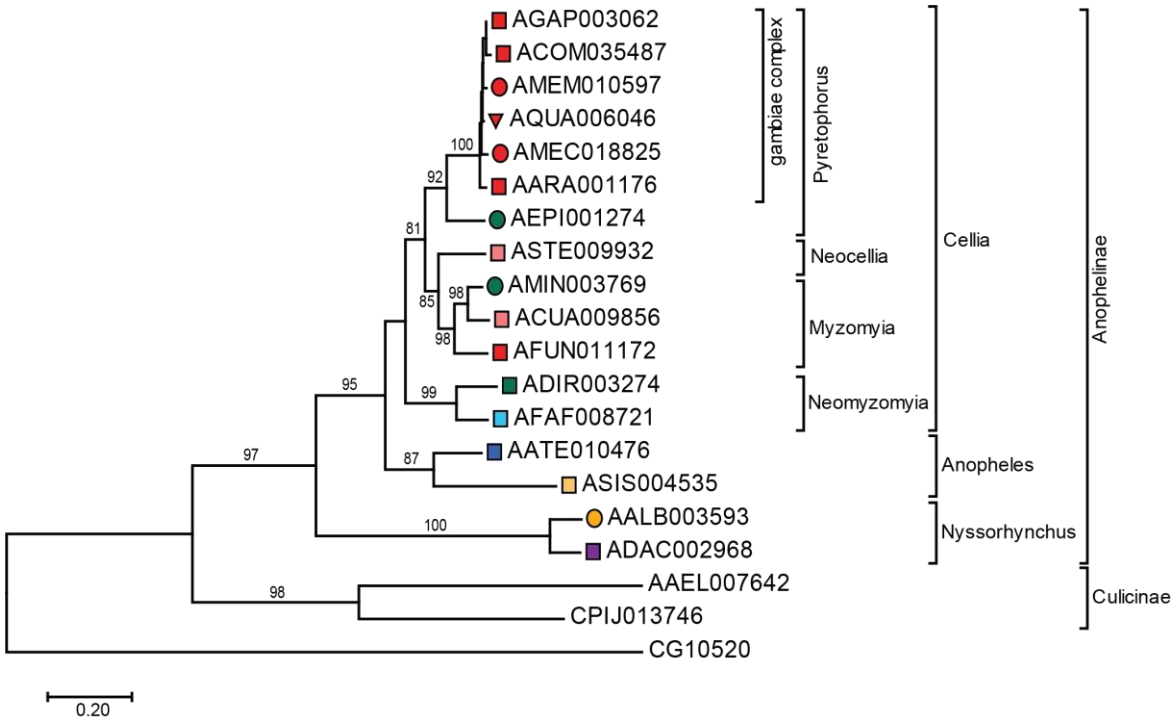

Figure B-17 TUBE Phylogeny

Phylogenetic tree obtained by maximum-likelihood analysis, reflecting the relationships of TUBE amino acid sequences. Scale bars indicate substitutions per site per unit of branch length and the number at each branch reflects bootstrap percentages (1000 replications). Only branches with support over 75% have values listed. Branch labels are coded according to Neafsey et al. 2015 and indicate vector status and geographic distribution of species (square, major vector; circle, minor vector; triangle, nonvector; red, Africa; pink, South Asia; green, South-East Asia; light blue, Asia Pacific; dark blue, Europe; light orange, East Asia; dark orange, Central America; purple, South America).

Tree topology fully follows published species phylogeny, with strong bootstrap support (>75%) indicated at branch points. Further details on tree analysis can be found in Supplemental Table 3.

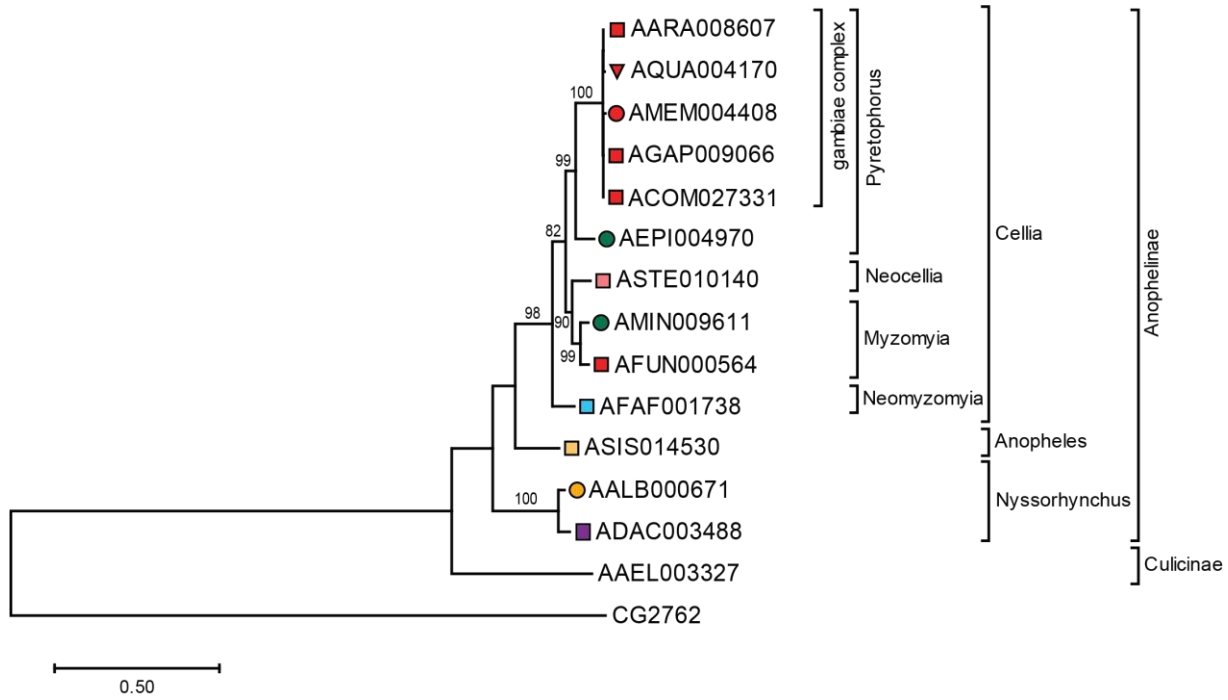

Figure B-18 USH Phylogeny

Phylogenetic tree obtained by maximum-likelihood analysis, reflecting the relationships of U-shaped amino acid sequences. Scale bars indicate substitutions per site per unit of branch length and the number at each branch reflects bootstrap percentages (1000 replications). Only branches with support over 75% have values listed. Branch labels are coded according to Neafsey et al. 2015 and indicate vector status and geographic distribution of species (square, major vector; circle, minor vector; triangle, nonvector; red, Africa; pink, South Asia; green, South-East Asia; light blue, Asia Pacific; dark blue, Europe; light orange, East Asia; dark orange, Central America; purple, South America).

Tree topology fully follows published species phylogeny, with strong bootstrap support (>75%) indicated at branch points. Further details on tree analysis can be found in Supplemental Table 3.

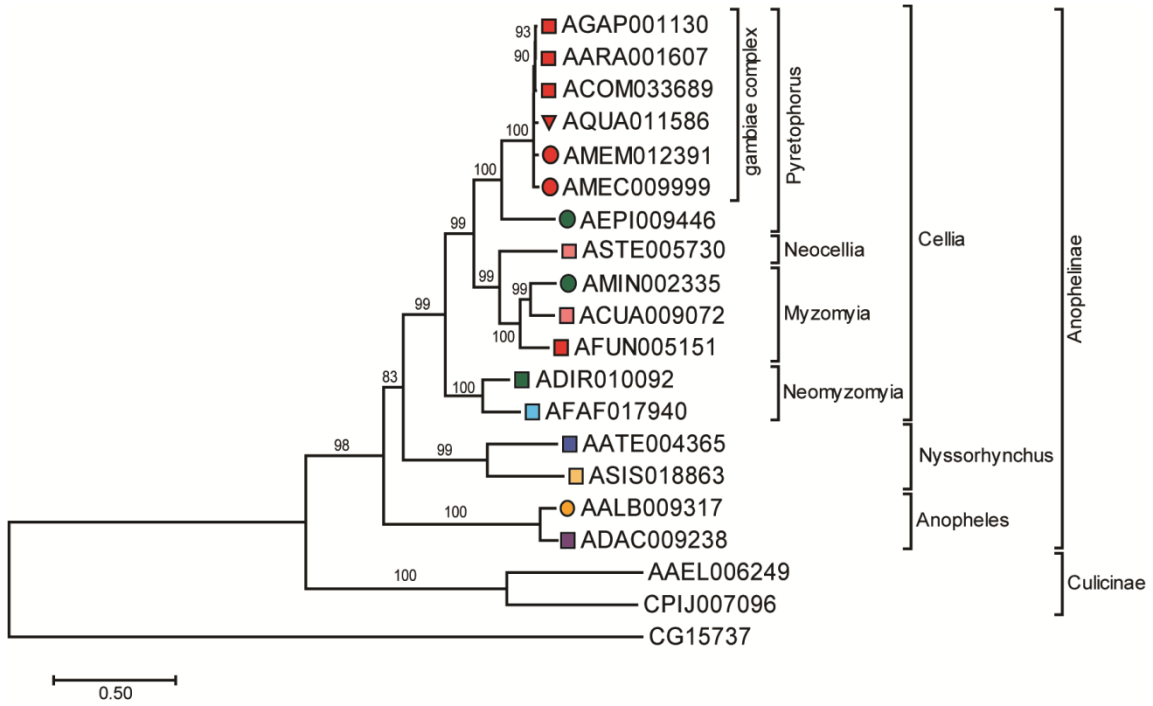

Figure B-19 WISP Phylogeny

Phylogenetic tree obtained by maximum-likelihood analysis, reflecting the relationships of Wispy amino acid sequences. Scale bars indicate substitutions per site per unit of branch length and the number at each branch reflects bootstrap percentages (1000 replications). Only branches with support over 75% have values listed. Branch labels are coded according to Neafsey et al. 2015 and indicate vector status and geographic distribution of species (square, major vector; circle, minor vector; triangle, nonvector; red, Africa; pink, South Asia; green, South-East Asia; light blue, Asia Pacific; dark blue, Europe; light orange, East Asia; dark orange, Central America; purple, South America).

Tree topology fully follows published species phylogeny, with strong bootstrap support (>75%) indicated at branch points. Further details on tree analysis can be found in Supplemental Table 3.

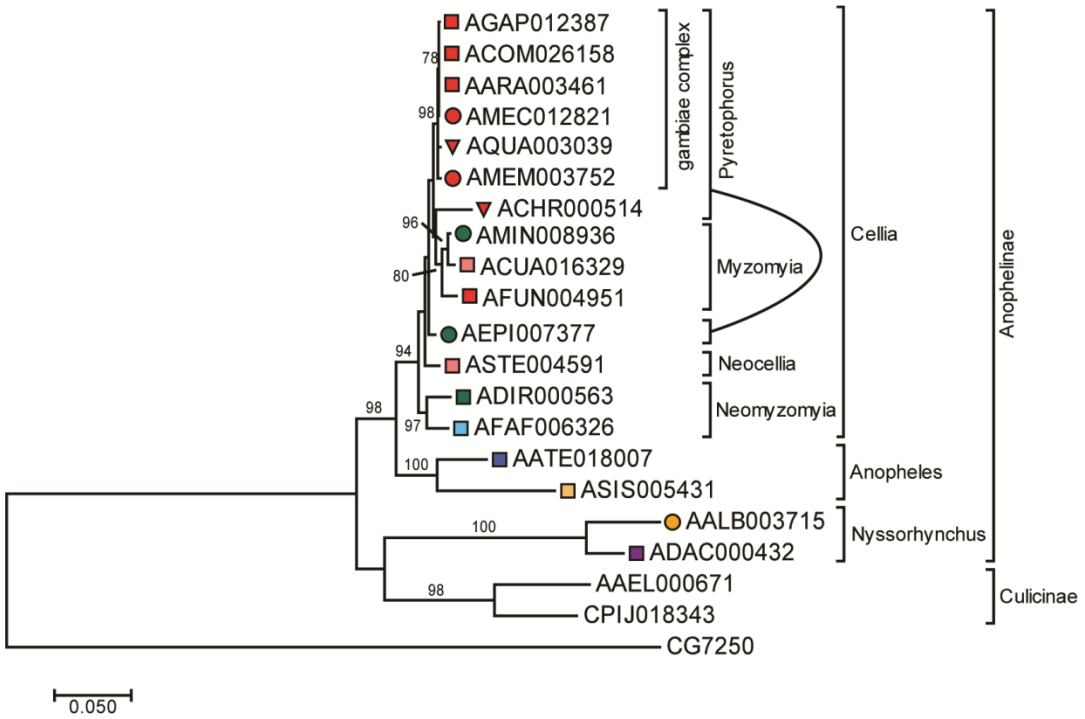

Figure B-20 TOLL6 Phylogeny

Phylogenetic tree obtained by maximum-likelihood analysis, reflecting the relationships of Toll6 amino acid sequences. Scale bars indicate substitutions per site per unit of branch length and the number at each branch reflects bootstrap percentages (1000 replications). Only branches with support over 75% have values listed. Branch labels are coded according to Neafsey et al. 2015 and indicate vector status and geographic distribution of species (square, major vector; circle, minor vector; triangle, nonvector; red, Africa; pink, South Asia; green, South-East Asia; light blue, Asia Pacific; dark blue, Europe; light orange, East Asia; dark orange, Central America; purple, South America).

Topology of *Neocellia* series and *Myzomyia* series is disrupted by the placement of *An. christyi* and *An. epiroticus* sequences. However, low bootstrap values at these branch points (<75%) makes true placement of these species undetermined. Further details on tree analysis can be found in Supplemental Table 3.

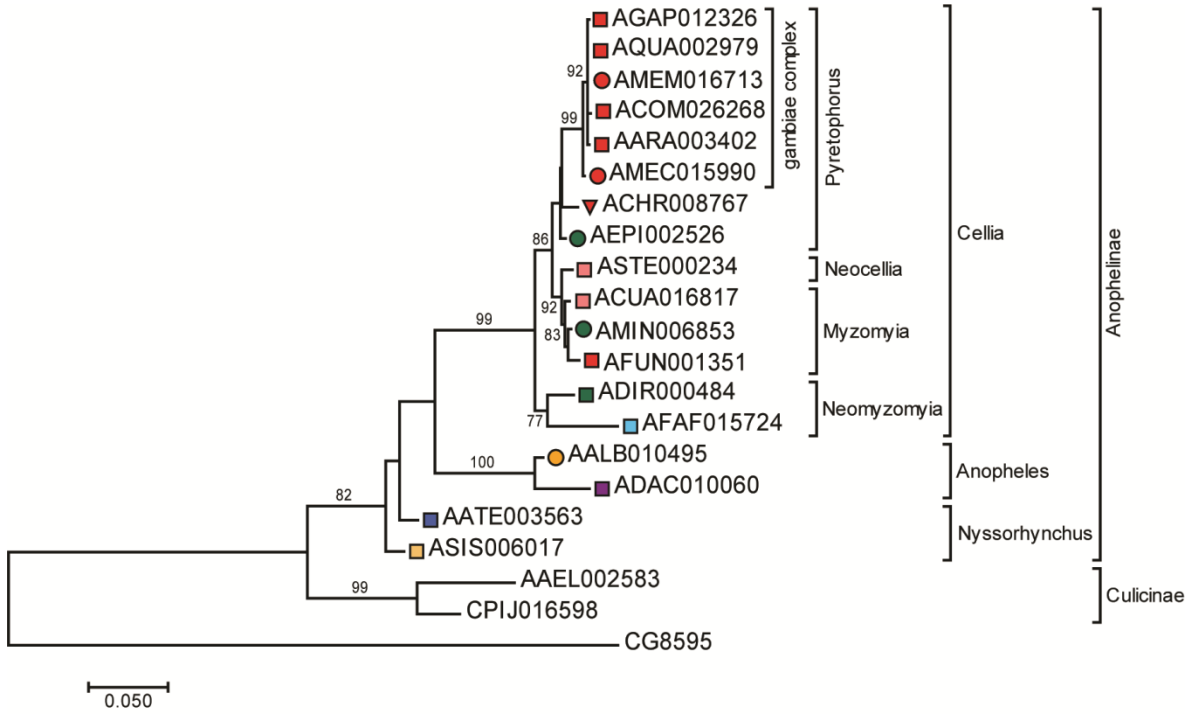

Figure B-21 TOLL7 Phylogeny

Phylogenetic tree obtained by maximum-likelihood analysis, reflecting the relationships of Toll7 amino acid sequences. Scale bars indicate substitutions per site per unit of branch length and the number at each branch reflects bootstrap percentages (1000 replications). Only branches with support over 75% have values listed. Branch labels are coded according to Neafsey et al. 2015 and indicate vector status and geographic distribution of species (square, major vector; circle, minor vector; triangle, nonvector; red, Africa; pink, South Asia; green, South-East Asia; light blue, Asia Pacific; dark blue, Europe; light orange, East Asia; dark orange, Central America; purple, South America).

Tree topology fully follows published species phylogeny, with strong bootstrap support (>75%) indicated at branch points. Further details on tree analysis can be found in Supplemental Table 3.

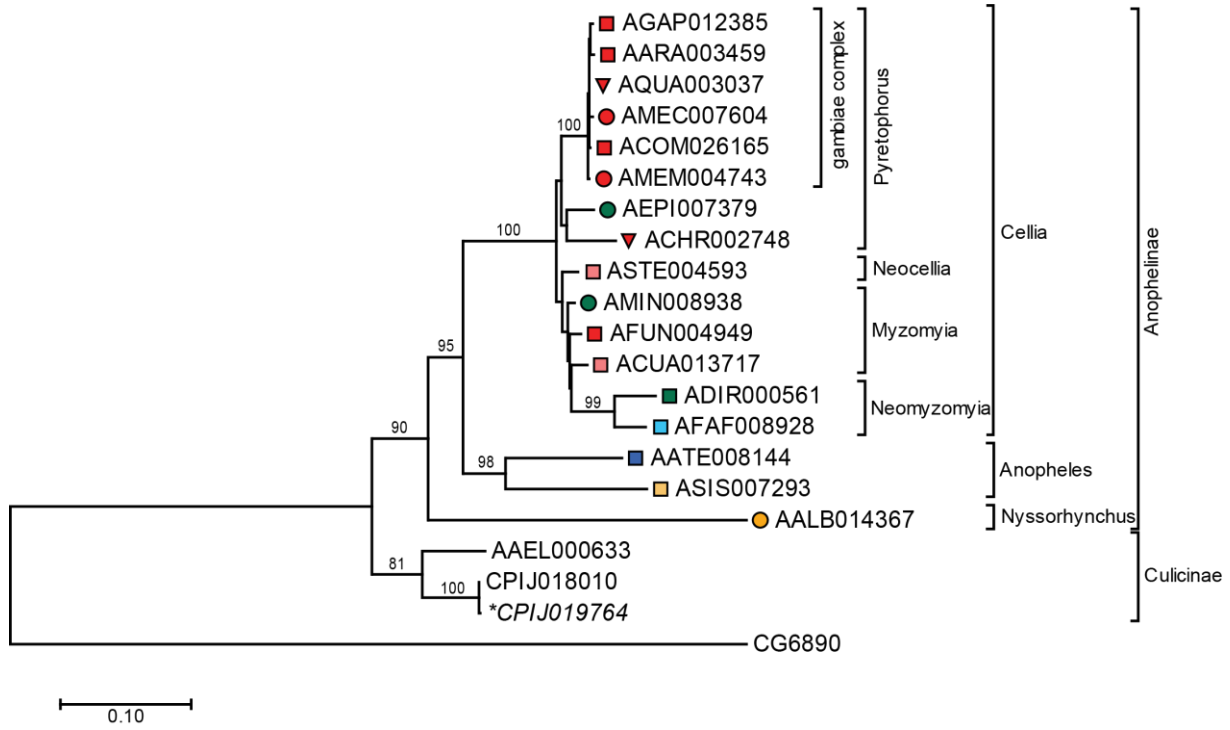

Figure B-22 TOLL8 Phylogeny

Phylogenetic tree obtained by maximum-likelihood analysis, reflecting the relationships of Toll8 amino acid sequences. Scale bars indicate substitutions per site per unit of branch length and the number at each branch reflects bootstrap percentages (1000 replications). Only branches with support over 75% have values listed. Branch labels are coded according to Neafsey et al. 2015 and indicate vector status and geographic distribution of species (square, major vector; circle, minor vector; triangle, nonvector; red, Africa; pink, South Asia; green, South-East Asia; light blue, Asia Pacific; dark blue, Europe; light orange, East Asia; dark orange, Central America; purple, South America).

Neomyzomya series (*An. dirus* and *An. farauti*) are placed within series *Myzomyia*, differing from its usual placement as the basal series within the subgenus *Cellia*. While the grouping of these two species is corroborated by a strong bootstrap value of 99, placement within *Myzomyia* is not, and thus we cannot make determination on the true nature of this phylogeny.

Additionally, there is a duplication of Toll8 within *C. quinquefasciatus*, but these genes are likely not reflective a true duplication and, as such, CPIJ019764 has been excluded from our final data set (asterisk). Further details on tree analysis can be found in Supplemental Table 3.

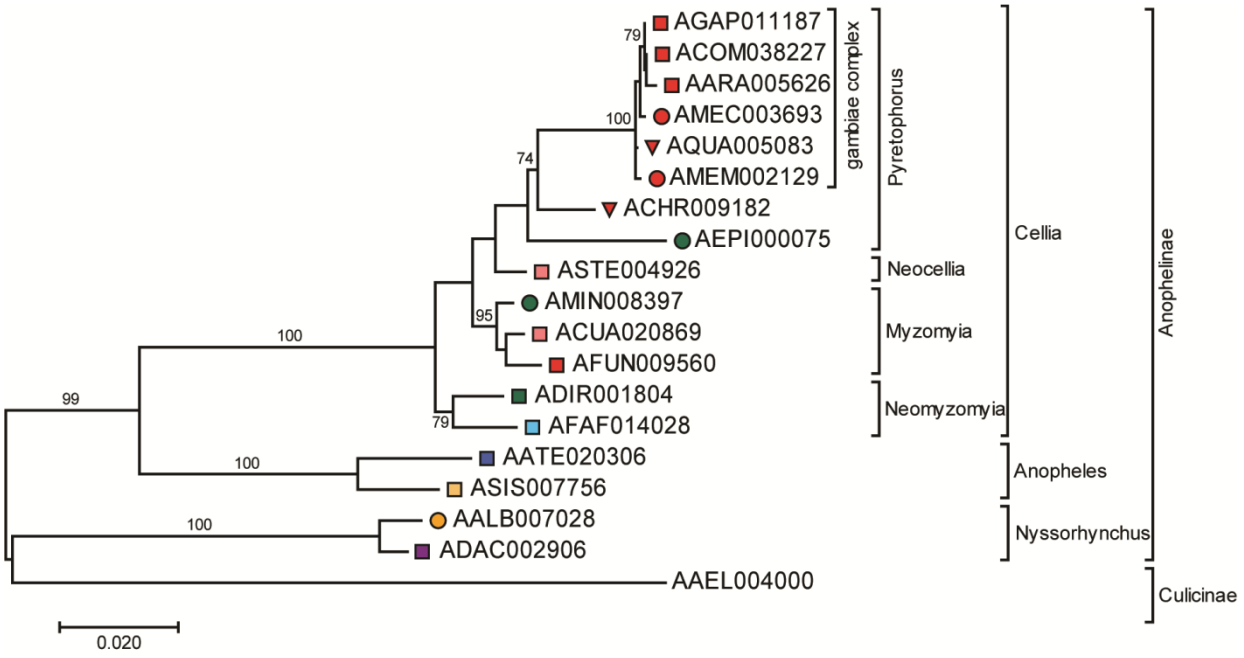

Figure B-23 TOLL10 Phylogeny

Phylogenetic tree obtained by maximum-likelihood analysis, reflecting the relationships of Toll10 amino acid sequences. Scale bars indicate substitutions per site per unit of branch length and the number at each branch reflects bootstrap percentages (1000 replications). Only branches with support over 75% have values listed. Branch labels are coded according to Neafsey et al. 2015 and indicate vector status and geographic distribution of species (square, major vector; circle, minor vector; triangle, nonvector; red, Africa; pink, South Asia; green, South-East Asia; light blue, Asia Pacific; dark blue, Europe; light orange, East Asia; dark orange, Central America; purple, South America).

Tree topology fully follows published species phylogeny, with strong bootstrap support (>75%) indicated at branch points. Further details on tree analysis can be found in Supplemental Table 3.

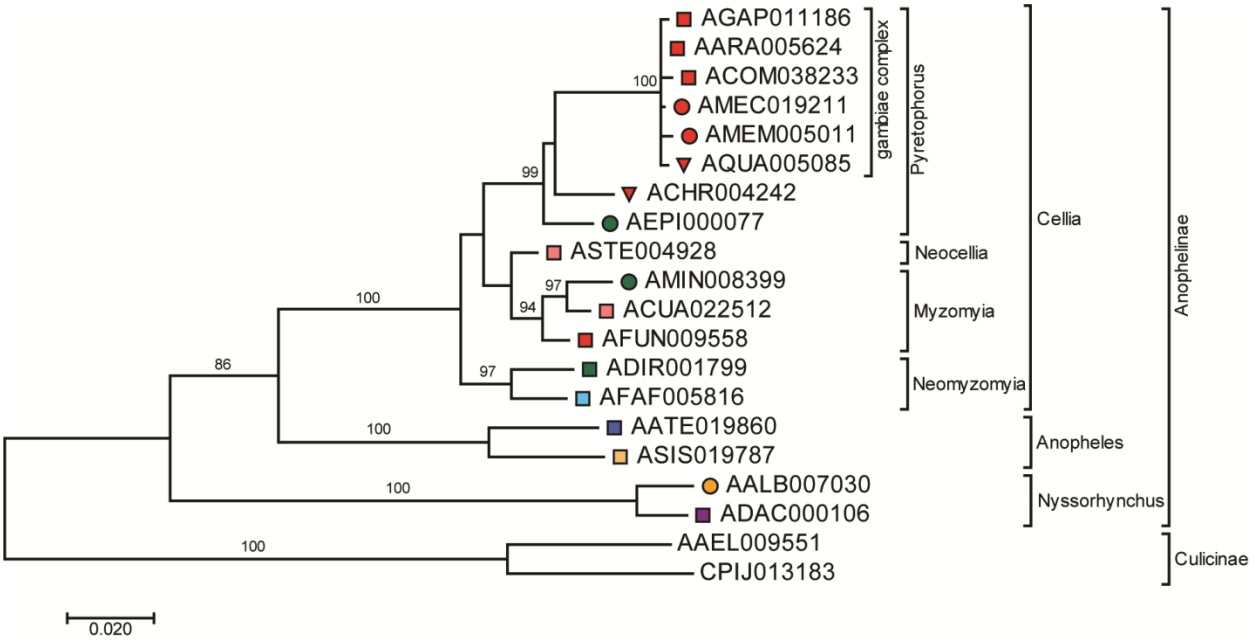

Figure B-24 TOLL11 Phylogeny

Phylogenetic tree obtained by maximum-likelihood analysis, reflecting the relationships of Toll amino acid sequences. Scale bars indicate substitutions per site per unit of branch length and the number at each branch reflects bootstrap percentages (1000 replications). Only branches with support over 75% have values listed. Branch labels are coded according to Neafsey et al. 2015 and indicate vector status and geographic distribution of species (square, major vector; circle, minor vector; triangle, nonvector; red, Africa; pink, South Asia; green, South-East Asia; light blue, Asia Pacific; dark blue, Europe; light orange, East Asia; dark orange, Central America; purple, South America).

Tree topology fully follows published species phylogeny, with strong bootstrap support (>75%) indicated at branch points. Further details on tree analysis can be found in Supplemental Table 3.

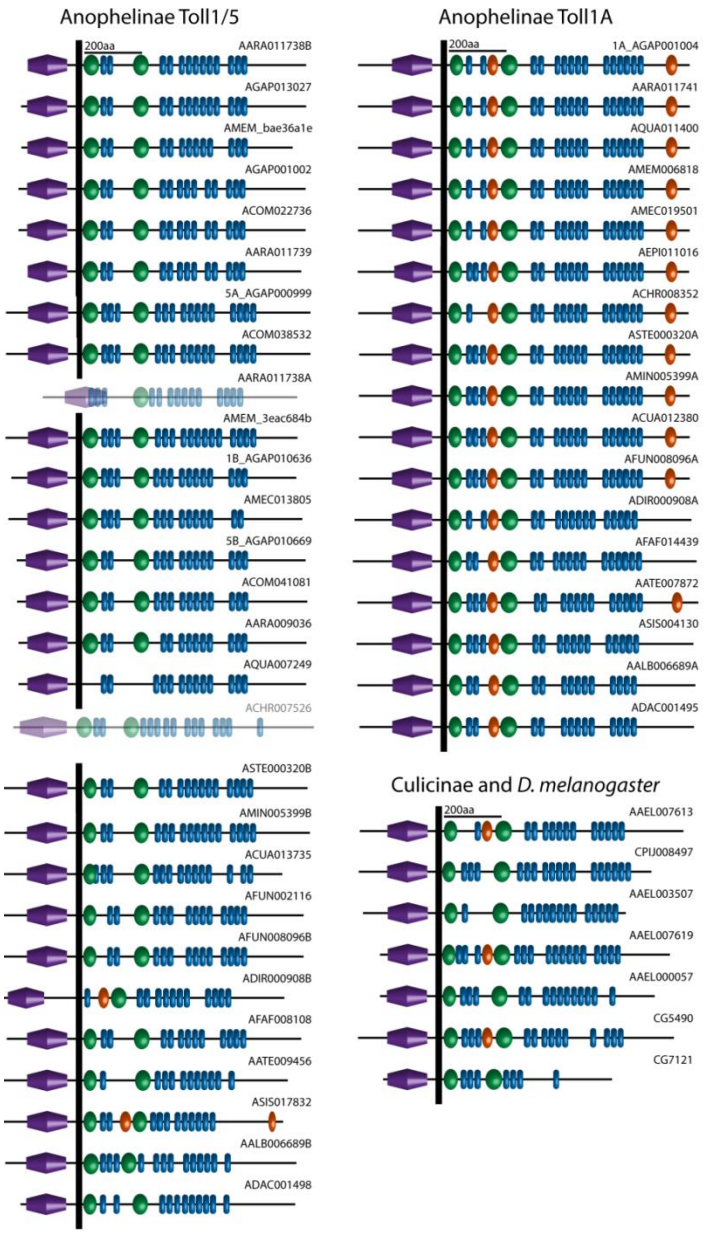

Figure B-25 TOLL1/5 Cluster Individual Protein Motifs

Schematic representation of predicted domains within mosquito and *Drosophila melanogaster* TLRs within the TOLL1/5 expansion cluster. Domains are drawn to scale and predicted using Pfam, TMHMM Server version 2.0, and LRR finder. LRR, (blue) LRR-CT (green), LRR-NT (orange), and TIR (purple) domains are indicated. Black rectangle is a transmembrane domain.

Two sequences, AARA011738A and ACHR007526, are lightened to indicate the lack of a predicted transmembrane domain.

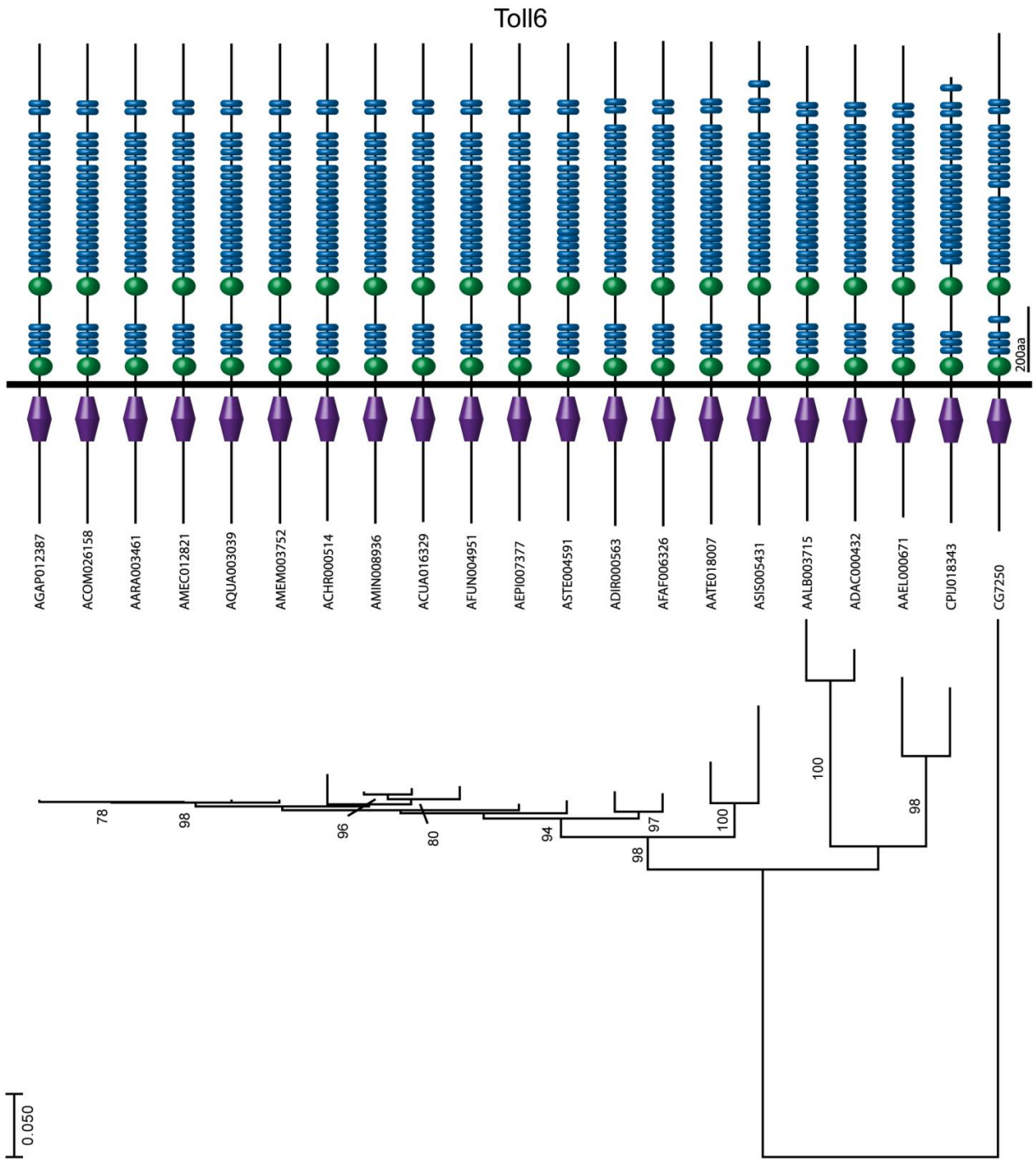

Figure B-26 TOLL6 Individual Protein Motifs

Schematic representation of predicted domains within mosquito and *Drosophila melanogaster* TLR6 orthologs. Domains are drawn to scale and predicted using Pfam, TMHMM Server version 2.0, and LRR finder. LRR, (blue) LRR-CT (green), LRR-NT (orange), and TIR (purple)

domains are indicated. Black rectangle is a transmembrane domain. Phylogenetic relationships are indicated below motif schematics and are to scale (bottom left).

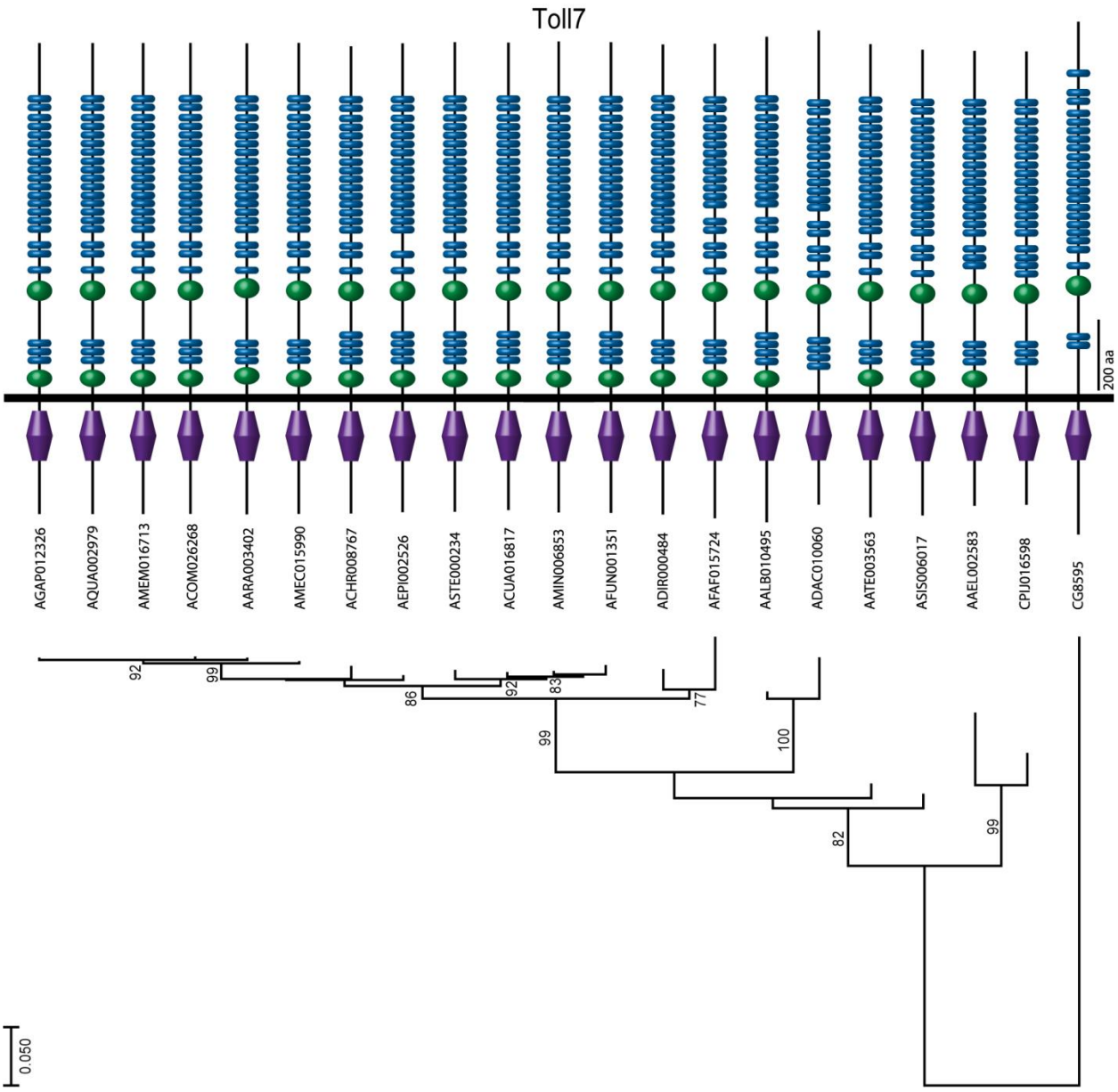

Figure B-27 TOLL7 Individual Proteins Motifs

Schematic representation of predicted domains within mosquito and *Drosophila melanogaster* TLR7 orthologs. Domains are drawn to scale and predicted using Pfam, TMHMM Server version 2.0, and LRR finder. LRR, (blue) LRR-CT (green), LRR-NT (orange), and TIR (purple) domains are indicated. Black rectangle is a transmembrane domain. Phylogenetic relationships are indicated below motif schematics and are to scale (bottom left).

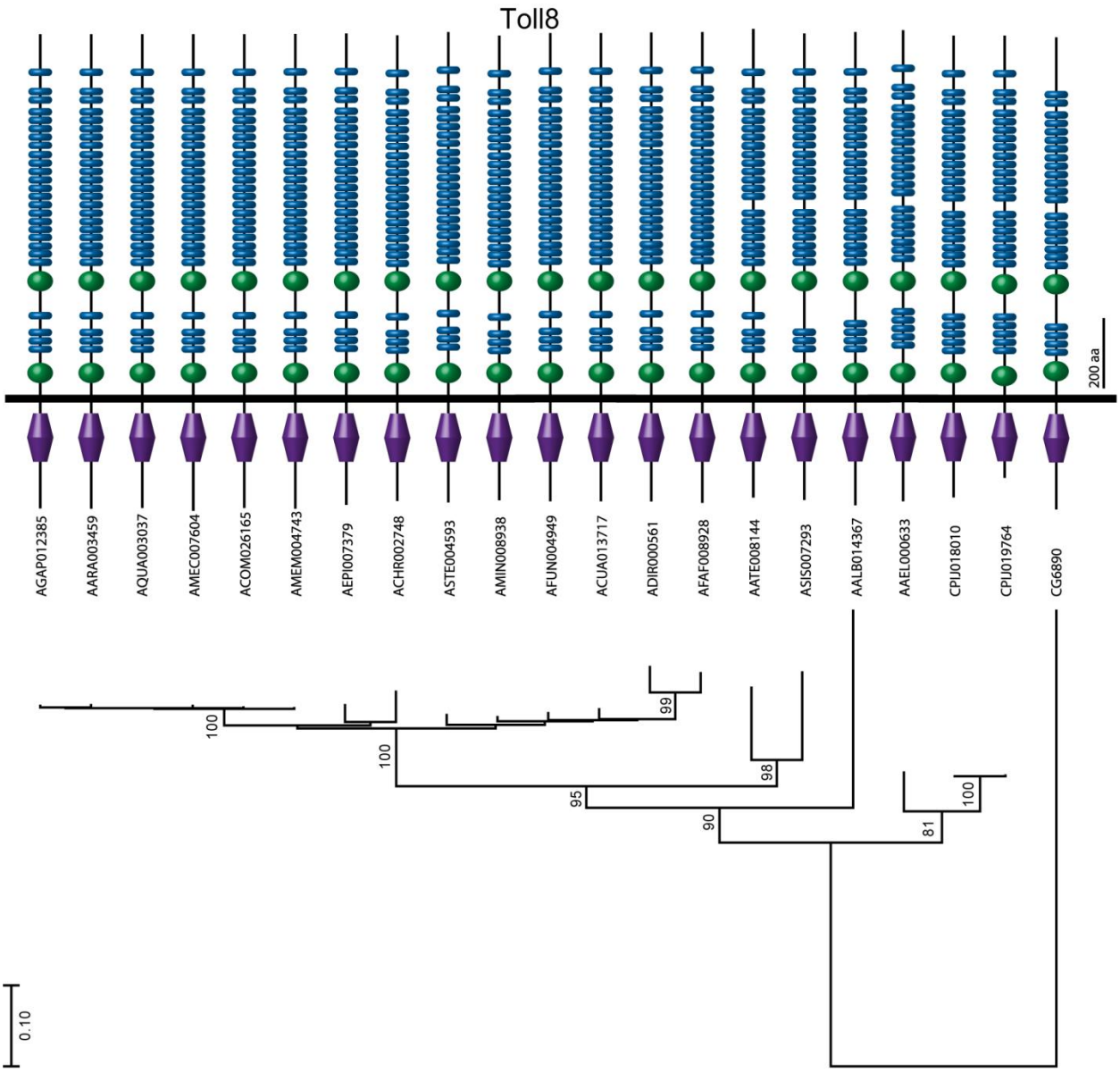

Figure B-28 TOLL8 Individual Protein Motifs

Schematic representation of predicted domains within mosquito and *Drosophila melanogaster* TLR8 orthologs. Domains are drawn to scale and predicted using Pfam, TMHMM Server version 2.0, and LRR finder. LRR, (blue) LRR-CT (green), LRR-NT (orange), and TIR (purple) domains are indicated. Black rectangle is a transmembrane domain. Phylogenetic relationships are indicated below motif schematics and are to scale (bottom left).

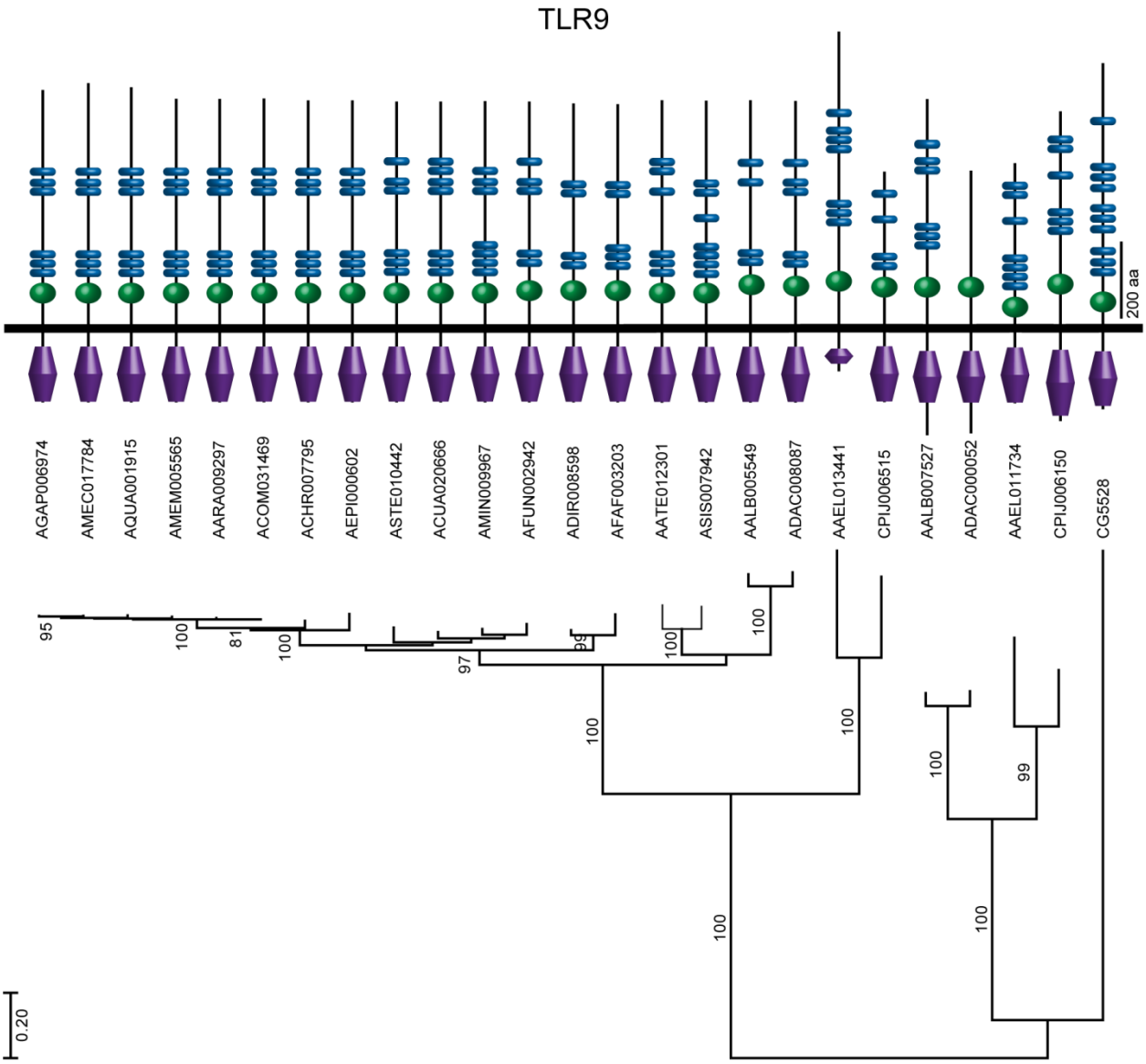

Figure B-29 TOLL9 Individual Protein Motifs

Schematic representation of predicted domains within mosquito and *Drosophila melanogaster* TLR9 orthologs. Domains are drawn to scale and predicted using Pfam, TMHMM Server version 2.0, and LRR finder. LRR, (blue) LRR-CT (green), LRR-NT (orange), and TIR (purple) domains are indicated. Black rectangle is a transmembrane domain. Phylogenetic relationships are indicated below motif schematics and are to scale (bottom left).

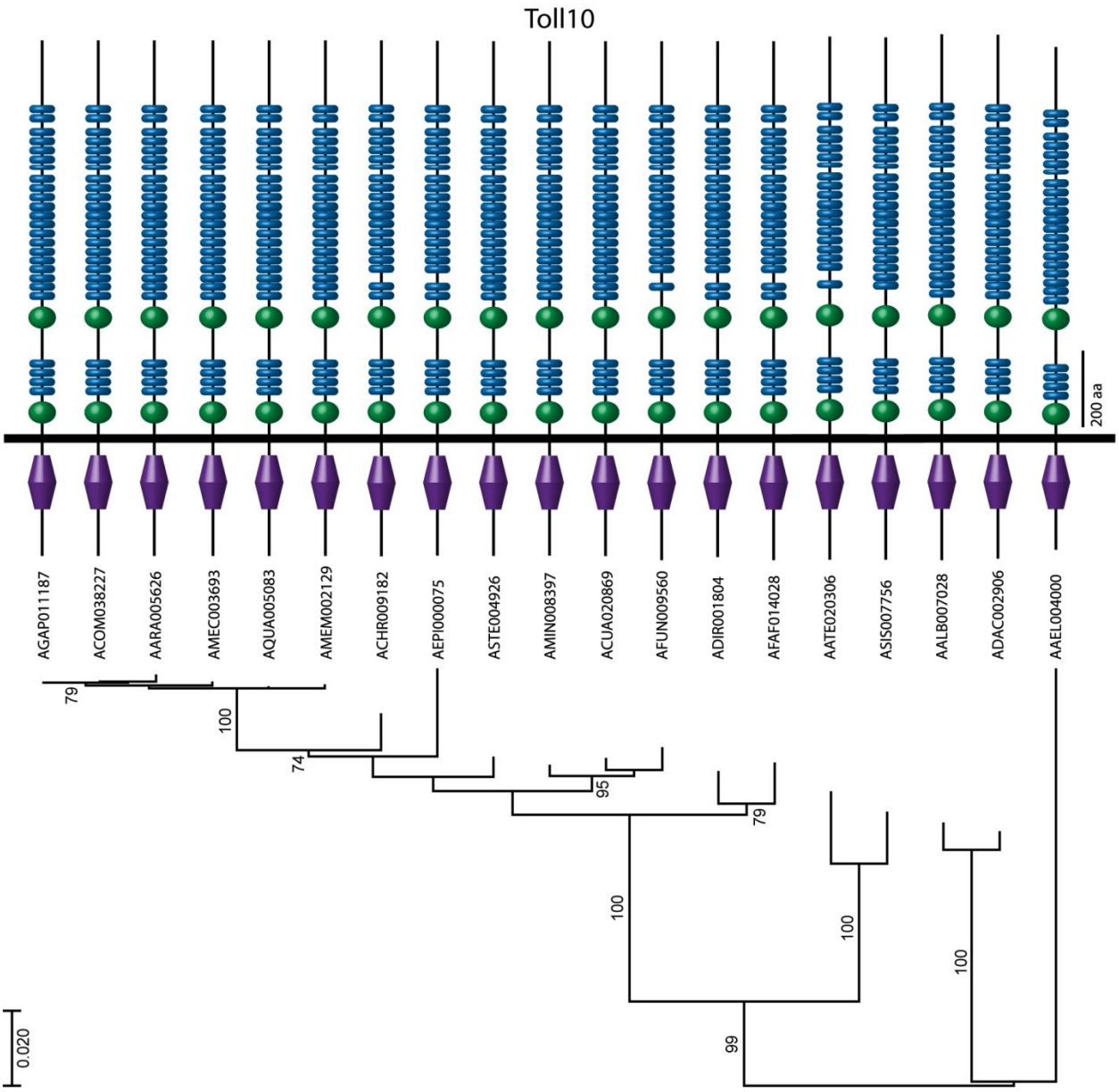

Figure B-30 TOLL10 Individual Protein Motifs

Schematic representation of predicted domains within mosquito and *Drosophila melanogaster* TLR10 orthologs. Domains are drawn to scale and predicted using Pfam, TMHMM Server version 2.0, and LRR finder. LRR, (blue) LRR-CT (green), LRR-NT (orange), and TIR (purple) domains are indicated. Black rectangle is a transmembrane domain. Phylogenetic relationships are indicated below motif schematics and are to scale (bottom left).

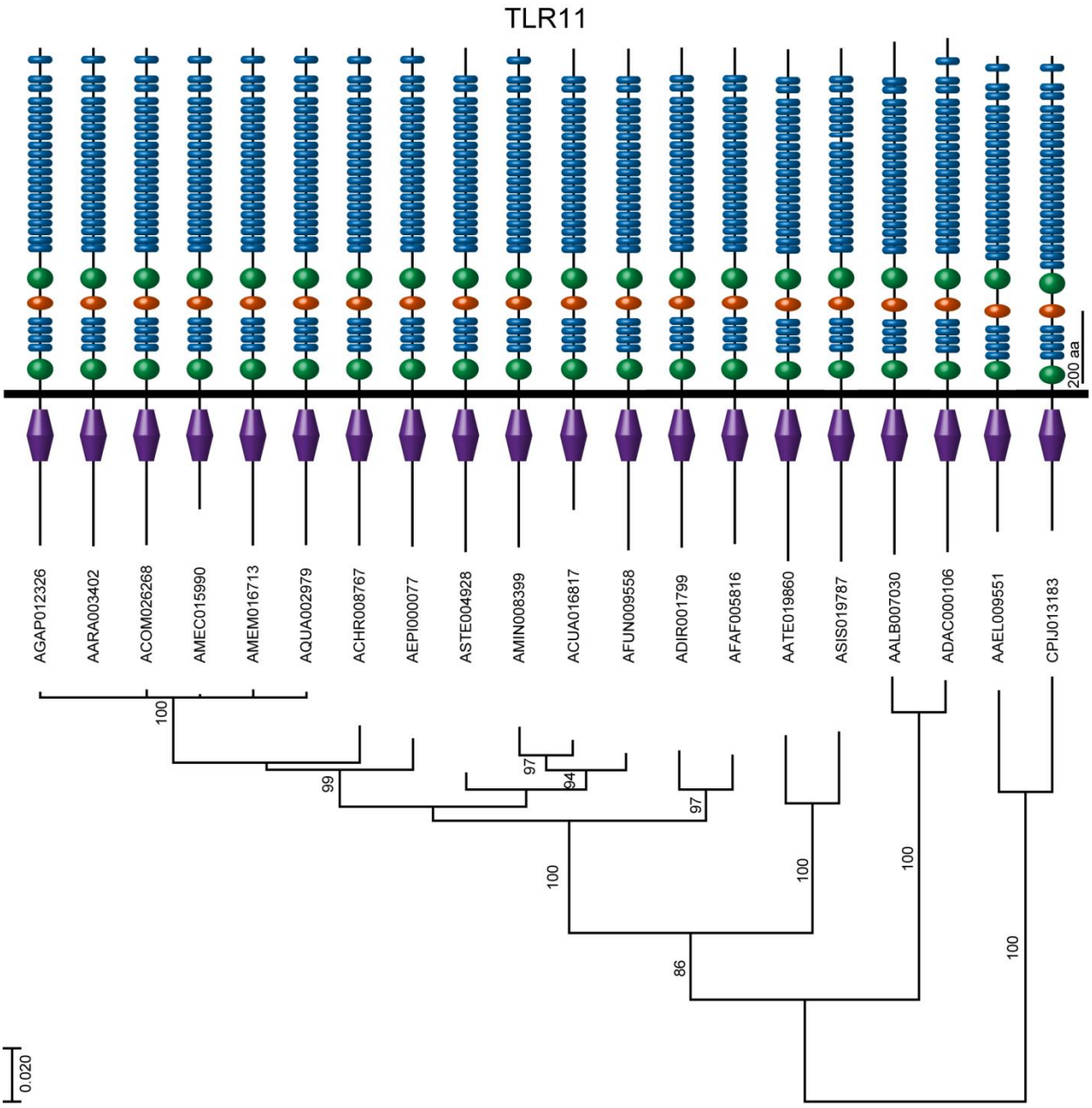

Figure B-31 TOLL11 Individual Protein Motifs

Schematic representation of predicted domains within mosquito and *Drosophila melanogaster* TLR11 orthologs. Domains are drawn to scale and predicted using Pfam, TMHMM Server version 2.0, and LRR finder. LRR, (blue) LRR-CT (green), LRR-NT (orange), and TIR (purple) domains are indicated. Black rectangle is a transmembrane domain. Phylogenetic relationships are indicated below motif schematics and are to scale (bottom left).

Appendix C - Genome analysis of a major urban malaria vector mosquito, *Anopheles stephensi*

Xiaofang Jiang[1,2†], Ashley Peery[3†], A Brantley Hall[1,2], Atashi Sharma[3], Xiao-Guang Chen[4], Robert M Waterhouse[5,6,7,8], Aleksey Komissarov[9,10], Michelle M Riehle[11], Yogesh Shouche[12], Maria V Sharakhova[3], Dan Lawson[13], Nazzy Pakpour[14], Peter Arensburger[15], Victoria L M Davidson[16], Karin Eiglmeier[17,18], Scott Emrich[19], Phillip George[3], Ryan C Kennedy[20], Shrinivasrao P Mane[21], Gareth Maslen[12], Chioma Oringanje[22], Yumin Qi[2], Robert Settlage[23], Marta Tojo[22], Jose M C Tubio[24], Maria F Unger[25], Bo Wang[14], Kenneth D Vernick[17,18], Jose M C Ribeiro[26], Anthony A James[27], Kristin Michel[17,18], Michael A Riehle[22], Shirley Luckhart[14], Igor V Sharakhov[1,3*] and Zhijian Tu[1,2*]

* Correspondence: igor@vt.edu; jaketu@vt.edu

†Equal contributors

Author details listed at the end of this chapter

Abstract

Background: Anopheles *stephensi* is the key vector of malaria throughout the Indian subcontinent and Middle East and an emerging model for molecular and genetic studies of mosquito-parasite interactions. The type form of the species is responsible for the majority of urban malaria transmission across its range.

Results: Here, we report the genome sequence and annotation of the Indian strain of the type form of *An. stephensi*. The 221 Mb genome assembly represents more than 92% of the entire genome and was produced using a combination of 454, Illumina, and PacBio sequencing. Physical mapping assigned 62% of the genome onto chromosomes, enabling chromosome-based analysis. Comparisons between *An. stephensi* and *An. gambiae* reveal that the rate of gene order reshuffling on the X chromosome was three times higher than that on the autosomes. *An. stephensi* has more heterochromatin in pericentric regions but less repetitive DNA in chromosome arms than *An. gambiae*. We also identify a number of Y-chromosome contigs and BACs. Interspersed repeats constitute 7.1% of the assembled genome while LTR retrotransposons alone comprise more than 49% of the Y contigs. RNA-seq analyses provide new insights into mosquito innate immunity, development, and sexual dimorphism.

Conclusions: The genome analysis described in this manuscript provides a resource and platform for fundamental and translational research into a major urban malaria vector. Chromosome-based investigations provide unique perspectives on *Anopheles* chromosome evolution. RNA-seq analysis and studies of immunity genes offer new insights into mosquito biology and mosquito-parasite interactions.

Introduction

Mosquitoes in the genus *Anopheles* are the primary vectors of human malaria parasites and the resulting disease is one of the most deadly and costly in history (Feachem et al., 2010; White et al., 2011). Publication and availability of the *An. gambiae* genome sequence accelerated research that has not only enhanced our basic understanding of vector genetics, behavior, and physiology and roles in transmission, but also contributed to new strategies for combating malaria (Holt et al., 2002). Recent application of next-generation sequencing technologies to mosquito genomics offers exciting opportunities to expand our understanding of mosquito biology in many important vector species and harness the power of comparative genomics. Such information will further facilitate the development of new strategies to combat malaria and other mosquito-borne diseases. *An. stephensi* is among approximately 60 species considered important in malaria transmission and is the key vector of urban malaria on the Indian subcontinent and the Middle East (Rafinejad et al., 2008; Sharma, 1999). The fact that a recent resurgence of human malaria in Africa could have been caused by the sudden appearance of *An. stephensi* indicates that *An. stephensi* may pose an even greater risk to human health in the future (Faulde, Rueda, and Khaireh, 2014). Of the three forms, type, *mysorensis*, and intermediate, the former is responsible for the majority, if not all, of urban malaria transmission across its range and accounts for approximately 12% of all transmission in India (Gaskar, Sharma, and Sharma 2013). Thus efforts to control it can be expected to contribute significantly to the malaria eradication agenda (Alonso et al., 2011; Murray et al., 2012). *An. stephensi* is amenable to genetic manipulations such as transposon-based germline transformation (Nolan et al., 2002), genome-wide mutagenesis

(O'Brochta et al., 2011), site-specific integration (Isaacs et al., 2012), genome-editing (Smidler et al., 2013), and RNAi-based functional genomics analysis (Brown et al., 2003). Our understanding of the interactions between *An. stephensi* and the malaria parasites is rapidly improving (Bian et al., 2013; Dong et al., 2011; Garver, Dong, and Dimopoulos, 2009; Luckhart et al., 2013; Mitri, Theiry, and Bourgouin, 2009; Pakpour et al., 2012). Thus *An. stephensi* is emerging as a model species for genetic and molecular studies. We report the draft genome sequence of the Indian strain of the type form of *An. stephensi* as a resource and platform for fundamental and translational research. We also provide unique perspectives on *Anopheles* chromosome evolution and offer new insights into mosquito biology and mosquito-parasite interactions.

Results and Discussion

Draft genome sequence of *An. stephensi*: Assembly and verification

The *An. stephensi* genome was sequenced using 454 GS FLX, Illumina HiSeq, and PacBio RS technologies. The 454 reads comprised 19.4× coverage: 12.2× from single-end reads, 2.2× from 3 kilobase (kb) paired-end reads, 3.4× from 8 kb paired-end reads, and 1.7× from 20 kb paired-end reads. The majority of 454 reads was in the range of 194 to 395 base-pairs (bp) in length. A single lane of Illumina sequencing of male genomic DNA resulted in 86.4× coverage of 101 bp paired-end reads with an average insert size of approximately 200 bp. Ten cells of PacBio RS sequencing of male genomic DNA produced 5.2× coverage with a median length of 1,295 bp. A hybrid assembly combining 454 and Illumina data produced a better overall result than using 454 data alone (Materials and methods). The resulting assembly was further improved by filling gaps with errorcorrected PacBio reads and scaffolding with BAC-ends. The current assembly, verified using various methods, contains 23,371 scaffolds spanning 221 Mb. The assembly includes 11.8 Mb (5.3%) of gaps filled with Ns (Table C-1), which is slightly lower than the size of gaps in the *An. gambiae* assembly (20.7 Mb, 7.6%). The N50 scaffold size is 1.59 Mb and the longest scaffold is 5.9 Mb. The number of scaffolds is inflated because we choose to set the minimum scaffold length to 500 bp to include repeat-rich short scaffolds. The assembled size of 221 Mb is consistent with the previous estimate of the *An. stephensi* genome size of approximately 235 Mb (Rai and Black, 1999).

Table C-1 Assembly Statistics

Statistic	Value
Scaffolds (n)	23,371
Scaffold N50 size	1,591,355
Maximum scaffold length	5,975,090
Minimum scaffold length	486
Percent Ns	5.35%
Contigs (n)	31,761
Contig N50 size	36,511
Maximum contig length	475,937
Minimim contig length	347
Total length of contigs	209,483,518
GC percent	44.80%

Physical mapping

Mapping of 227 probes was sufficient to assign 86 scaffolds to unique positions on the *An. stephensi* polytene chromosomes (Figure C-1; Table C-2). These 86 scaffolds comprise 137.14 Mb or 62% of the assembled genome. Our physical map includes 28 of the 30 largest scaffolds and we were able to determine the orientation of 32 of the 86 scaffolds. We expect that relatively little of the heterochromatin was captured in our chromosomal assembly based on the morphology of the chromosomes in regions to which the scaffolds mapped. For this reason, subsequent comparisons with *An. gambiae* on molecular features of the genome landscape exclude regions of known

heterochromatin from the *An. gambiae* dataset. *An. stephensi* and *An. gambiae* have different chromosome arm associations with 2L of *An. gambiae* homologous to 3L of *An. stephensi* (Sharakhova et al., 2011). Therefore, all ensuing discussion of synteny between the two species refers to *An. stephensi* chromosome arms listed in homologous order to those of *An. gambiae*: X, 2R, 3L, 3R, and 2L. While draft genomes also are available for *An. darlingi* and *An. sinensis* (Marinotti et al., 2013; Zhou et al., 2014), we focused our comparative analysis on *An. stephensi* and *An. gambiae*, the only two species that have chromosome-based assembly.

Gene annotation

A total of 11,789 protein-encoding genes were annotated using a combination of homology and *de novo* prediction. These gene models have been submitted to the NCBI (GCA_000300775.2) and are hosted in VectorBase (Giraldo-Calderon et al., 2015). The average transcript length was 3,666 bp and the average number of exons per transcript was 4.18. Evolutionary relationships among *An. stephensi* and other *dipteran* insects were evaluated by constructing a maximum likelihood molecular species phylogeny using universal single-copy orthologs (Figure C-2A). *An. stephensi* and *An. gambiae* form a well-supported clade representing the subgenus *Cellia* within the genus *Anopheles*. This phylogeny provides the evolutionary context for current and future comparative genomics analysis. A total 10,492 (89.0%) of the 11,789 predicted *An. stephensi* protein-encoding genes had orthologs in *An. gambiae*, *Aedes aegypti*, and *Drosophila melanogaster* (Figure C-2).

Global transcriptome analysis

Eleven RNA-seq samples were prepared from 0 to 1, 2 to 4, 4 to 8, and 8 to 12 h post-egg deposition embryos, larvae, pupae, adult males, adult females, non-blood-fed ovaries, blood-fed ovaries, and 24 h post-blood-fed female carcasses without ovaries (Criscione et al., 2013). The corresponding genes were clustered into 20 distinct groups in sizes in the range of 8 to 2,106 genes per group on the basis of similar expression patterns (Figure C-3). Many of the clusters correspond to either a specific developmental stage or sex (Additional file 2). A search for over-represented gene ontology (GO) terms in the 20 clusters found that many of the co-regulated genes have similar inferred functions or roles. Adult females require a protein-rich blood-meal for oogenesis and thus are the most interesting sex from a health perspective.

Table C-2 Physical map information

Arm	Scaffolds per arm (n)	Length (Mb)	Mapped genome (%)	Total genome (%)
X	9	14.95	10.90	6.77
2R	21	39.50	28.80	17.87
2L	15	22.40	16.33	10.14
3R	24	37.83	27.59	17.12
3L	17	22.45	16.37	10.16
Total	86	137.14	100	62.05

Scaffolds mapped to each chromosome, total bp to each chromosome, percent of the predicted genome covered.

Genes in clusters 1, 10, and 17 are induced in the female soma after blood-feeding. These clusters are enriched for genes encoding proteins with proteolytic

activity, including serine peptidases, and involved in blood-meal digestion. Mosquitoes have undergone lineage-specific amplification of serine peptidases when compared to *Drosophila*, many of which are found in the three clusters described above. Cluster 9 contains 258 genes that showed peak expression in the pupal stage and it is enriched for genes whose products are involved in exoskeleton development. GO analyses of other clusters are described in the Additional file 1: Text. We identified 241 and 313 genes with female or male biased expression, respectively (Additional file 2: Sex-biased genes list and GO terms). The male-biased genes are enriched for those whose products are involved in spermatogenesis and the auditory perception. Male mosquitoes detect potential mates using their Johnston's organ, which has twice the number of sensory neurons as that of the females (Göpfert and Robert, 2011; Gibsen, Warren, and Russell, 2010). The female-biased genes are enriched for those whose products are involved in proteolysis and other metabolic processes likely relevant to blood digestion.

Immunity genes

Manual annotation was performed on genes involved in innate immunity including those that encode the LRR immune (LRIM) and the *Anopheles Plasmodium*-responsive leucine-rich repeat 1 (APL1) proteins, and the genes of the Toll, immune deficiency (IMD), insulin/insulin-like growth factor signaling (IIS), mitogen-activated protein kinase (MAPK), and TGF-β signaling pathways. A number of studies have demonstrated the importance of these genes or pathways in mosquito defense against parasites or viruses (Dong et al., 2011; Garver, Dong, and Dimopoulos, 2009; Horton et al., 2011; Luckhart et al., 2013; Mitri, Thiery and Bourgouin, 2009; Pakpour et al., 2012; Price et al., 2013; Xi, Ramirez, and

Dimopoulos, 2008). Manual analysis showed overall agreement with the automated annotation and improved the gene models in some cases (Additional file 2). A high level of orthology is generally observed between *An. stephensi* and *An. gambiae* and we highlight here a few potentially interesting exceptions. *An. stephensi* may have only one APL1 gene (ASTEI02571) instead of the three APL1 gene cluster found in *An. gambiae*. We also observed the apparent lack of TOLL1B and 5B sequences in *An. stephensi*, which in *An. gambiae* are recent duplications of TOLL1A and 5A, respectively.

Expression profiles of all immunity genes were analyzed using the 11 RNA-seq samples to provide insights into their biological functions (Additional file 2: RNA-seq expression profile of immunity-related genes). For example, FKBP12, a protein known to regulate both transforming growth factor (TGF)-β and target of rapamycin (TOR) signaling, showed abundant transcript levels across immature stages and adult tissues. The high expression levels of AsteFKBP12 in all examined stages and tissues were unexpected. Examination of existing publicly-available microarray data confirmed these expression levels and patterns (Choi et al., 1996). FKBP12 in mammals forms a complex with rapamycin and FKBP-rapamycin-associated protein (FRAP) to inhibit TOR (Choi et al.,). Given that TOR signaling is fundamental to many biological functions in mammals (Laplante and Sabatini, 2012) and cumulative data support the same for *D. melanogaster* (Grewal, 2009), a high level of FBKP12 expression may be critical for tight regulation of TOR activity in *An. stephensi* and perhaps *An. gambiae* (Arsic and Guerin, 2008). Expression patterns of the *An. gambiae* FKBP12 ortholog, AGAP012184, from microarray datasets (Giraldo-Calderon, 2015) support the hypothesis that this protein is involved in a broad array of *Anopheline* physiologies including: development, blood-feeding, molecular

form-specific insecticide resistance, circadian rhythms, desiccation resistance, mating status, and possibly also broad regulation of infection based on studies with murine (*Plasmodium berghei*) and human (*Plasmodium falciparum*) malaria parasites. Whether these same physiologies and others are regulated by FKBP12 in *An. stephensi* will require experimental confirmation. Given that signaling pathways regulating embryonic pattern formation in *Drosophila* (for example, the Toll pathway (Anderson, Bokla, and Nüsslein-Volhard, 1985) have been co-opted in the adult fly for regulation of various physiologies including metabolism and immune defense, the data presented here support the hypothesis that pathways integral to adult biology in adult *Anophelines* also have been similarly co-opted from important developmental roles.

Salivary genes

Saliva of blood-feeding arthropods contains a cocktail of pharmacologically active components that disarm vertebrate host's blood clotting and platelet aggregation, induce vasodilation, and affect inflammation and immunity. These salivary proteins are under accelerated evolution due most likely to their host's immune pressure. A previous salivary gland transcriptome study identified 37 corresponding salivary proteins in *An. stephensi*, most of which are shared with *An. gambiae*, including mosquito and *Anopheles*-specific protein families (Valenzuela et al., 2003). A more extensive sialotranscriptome based on approximately 3,000 ESTs identified the templates for 71 putative secreted proteins for *An. gambiae* (Arca et al., 2005). The combined data verify the identity of 71 putative salivary secreted proteins for *An. stephensi*, seven of which have no similarities to *An. gambiae* proteins. The current assembly of the *An. stephensi* genome shows that many salivary gland

genes are present as tandem repeated genes and represent families that arose by gene duplication events. Tandem repeated gene families often are poorly annotated by automated approaches, therefore, manual annotation was necessary to improve the salivary gland gene models. In particular, *An. gambiae* has eight genes of the D7 family, which has modified odorant binding domains (OBD) that strongly bind agonists of platelet aggregation and vasoconstriction (histamine, serotonin, epinephrine, and norepinephrine) (Ribeiro, Mans, and Arcà, 2010). Three of these genes have two OBDs while the remaining five have only one domain each. As in *An. gambiae*, the short forms are oriented in tandem and in the opposite orientation of the long-form genes. However, *An. stephensi* has apparently collapsed the second long form to create a sixth short form.

Comparative analysis of additional gene families

Functional annotations of a number of gene families in *An. stephensi* were obtained based on their InterPro ID (Hunter et al., 2011). We also compared gene numbers in these gene families across several species. *An. stephensi* and *An. gambiae* showed similar gene numbers in most of the gene families (Holt et al., 2002) and this is consistent with the close phylogenetic relationship between the two species. As observed with manually annotated immunity-related genes, strong one-to-one relationship was observed between *An. stephensi* and *An. gambiae* genes in odorant binding proteins (OBPs) and other gene families studied. There are a few gene families that showed obvious difference in numbers between *An. stephensi* and *An. gambiae*. We performed phylogenetic analysis of these gene families. The results indicate gene expansion in the odorant receptors (OR) and fibrinogen-related proteins in *An. gambiae*. Interestingly, a plurality of expanded genes is physically clustered

in *An. gambiae*, suggesting that the gene expansions in *An. gambiae* may have arisen from local duplications. For example, the *An. stephensi* single-copy OR gene ASTEI08685 has four orthologs in *An. gambiae* (AGAP004354, AGAP004355, AGAP004356, and AGAP004357). The putative orthologs of these 'expanded' genes tend to be single or low-copy in *An. stephensi* and other related species in VectorBase, supporting the interpretation that the lack of duplicated copies in *An. stephensi* is not due to assembly or annotation error. Further analysis that includes all species in the ongoing 16 *Anopheles* genomes project (Neafsey et al., 2013) will facilitate future comparative analysis of gene family expansions and gene losses.

Repeat content

Transposable elements (TEs) and other unclassified interspersed repeats constitute 7.1% of the assembled *An. stephensi* genome (Table C-3). TE occupancy of the euchromatic genome in *D. melanogaster* and *An. gambiae* is 2% and 16%, respectively (Holt et al., 2002). Thus variations in the size of the genomes correlate with different amounts of repetitive DNA in these three species. More than 200 TEs have been annotated. DNA transposons and miniature inverted-repeat TEs (MITEs) comprise 0.44% of the genome. Non-LTR retrotransposons (or LINEs) comprise 2.36% of the genome. Short intersperse nuclear elements (SINEs), although less than 300 bp in length, are highly repetitive and comprise 1.7% of the genome. There is considerable diversity among the LTR-retrotransposons although they occupy only 0.7% of the genome. Approximately 2% of the genome consists of interspersed repeats that remain to be classified.

Genome landscape: a chromosomal arm perspective

The density of genes, TEs, and short tandem repeats (STRs) for each chromosome were determined based on the physical map (Figure C-4). The average numbers of genes for each chromosome arm are consistent with those in *An. gambiae*. The X had the lowest number of genes per 100 kb, and the highest densities of genes per 100 kb were seen on 2R and 3 L (Figure C-5). Chromosomes 2R and 3 L also contain the greatest numbers of polymorphic inversions (Mahmood and Sakai, 1984). Genes functioning as drivers of adaptation could be expected to occur in greater densities on chromosome arms with higher numbers of polymorphic inversions (Hoffmann and Sgrò, 2004). *An. stephensi* has a lower density of transposable elements across all chromosome arms than *An. gambiae* (Figure C-5). The density of transposable elements on the *An. stephensi* X is more than twice that of the autosomes. A comparison of the *An. stephensi* simple repeats with those in *An. gambiae* euchromatin showed that densities in the latter were approximately 2-2.5× higher (Figure C-5). The greatest densities of simple repeats were found on the X chromosome and this is consistent with a previous study in *An. gambiae* (Xia et al., 2010). Although *An. stephensi* shows lower densities of simple repeats across all arms compared to *An. gambiae*, its X appears to harbor an over-representation of simple repeats compared to its autosomes. Scaffold/Matrix-associated regions (S/MARs) can potentially affect chromosome mobility in the cell nucleus and rearrangements during evolution (Baricheva et al., 1996; Dechat et al., 2008) and these were found to be enriched in the 2 L and 3R arms (Figure C-5).

Table C-3 Transposable elements and other interspersed repeats

Type	Elements (n)	Length occupied (bp)	Genome (%)
SINEs	30,514	3,739,253	1.69

LINEs	22,022	5,231,240	2.36
LTR elements	4,359	1,499,282	0.68
DNA elements	4,611	966,667	0.44
Unclassified	30,611	4,322,468	1.95
Total	92,117	15,758,910	7.12

Molecular organization of pericentric heterochromatin We observed clear differences in heterochromatin staining patterns when comparing mitotic chromosome squashes prepared from imaginal discs of *An. gambiae* and *An. stephensi. An. stephensi* appears to have more pericentric heterochromatin than *An. gambiae*. This is particularly evident in the sex chromosomes. Mitotic X chromosomes in *An. stephensi* possess much more pericentric heterochromatin compared with X chromosomes from several different strains of *An. gambiae*. Finally, the Y chromosome in *An. stephensi* has a large block of heterochromatin. We further investigated whether particular tandem repeats are concentrated in heterochromatin. Aste72A and Aste190A, the two repeats with highest coverage in raw genomic data reads, were selected as probes for FISH analysis. Aste72A, which comprises approximately 1% of the raw genomic reads, was mapped to the pericentric heterochromatin of X and Y chromosomes (Figure C-6). Aste190A, which comprises approximately 2% of the raw genomic reads, was mapped to centromere of both autosomes. The Aste72A tandem repeat has a 26.7% mean GC content and contributes significantly to the AT-rich peak in the plot of GC distribution of raw genomic reads.

Y chromosome

Anopheles mosquitoes have heteromorphic sex-chromosomes where males are heterogametic (XY) and females homogametic (XX) (Baker and Sakai, 1979). The high repetitive DNA content of Y chromosomes makes them difficult to assemble and they often are ignored in genome projects. An approach called the chromosome quotient (Hall et al., 2013) was used to identify 57 putative Y sequences spanning 50,375 bp . All of these sequences are less than 4,000 bp in length and appear to be highly repetitive. Five BACs that appeared to be Y-linked based on the CQs of their end sequences were analyzed by sequencing and their raw PacBio reads were assembled with the HGAP assembler (Chin et al., 2013). Eleven contigs spanning 196,498 bp of predicted Y-linked sequences were obtained. The 57 Y-linked sequences and 11 contigs from the Y-linked BACs represent currently the most abundant set of Y sequences in any *Anopheles* species. RepeatMasker analysis using the annotated *An. stephensi* interspersed repeats showed that approximately 65% of the *An. stephensi* Y sequences are interspersed repeats. LTR retrotransposons alone occupy approximately 49% of the annotated Y.

Synteny and gene order evolution

We used the chromosomal location and orientation of 6,448 one-to-one orthologs from *An. gambiae* and *An. stephensi* to examine synteny and estimate the number of chromosomal inversions between these two species (Figure C-7). Syntenic blocks were defined as those that had at least two genes and all genes within the block had the same order and orientation with respect to one another in both genomes. The X chromosome has markedly more inversions than the autosomes. The number of chromosomal inversions that might have happened since *An. stephensi* and *An. gambiae* last shared a common ancestor was

determined with GRIMM (Tesler, 2002). We calculated the density of inversions per chromosome arm ignoring breakpoint reuse and assuming two breakpoints per inversion. The length of *An. stephensi* assembly was used as a proxy for the size of the *An. stephensi* chromosomes. The density of inversions per megabase on the X chromosome supports the conclusion that it is much more prone to rearrangement than the autosomes. Genomic segments on the X are approximately three-fold more likely to change order than those on the autosomes (Figure C-8A). The fast rate of X chromosome rearrangements contrasts with the lack of polymorphic inversions in *An. stephensi* and *An. gambiae*. Interestingly, a recent comparative genomic study between *An. gambiae* and *Ae. aegypti* revealed that the homomorphic sex-determining chromosome in *Ae. aegypti* has a higher rate of genome rearrangements than autosomes (Timoshevskiy et al., 2014).

Rates of chromosome evolution in *Drosophila* and *Anopheles* Recent studies have established that both *Anopheles* and *Drosophila* species have high rates of chromosomal evolution as compared with mammalian species (Bhutkar et al., 2008; Bourque and Pevzner, 2002; Chaisson, Raphael, and Pevzner, 2006; Peng, Pevzner, and Tesler, 2006; Ranz et al., 2007; Ranz and Ruiz, 2001; Shaeffer et al., 2008; Sharakhov et al., 2002; Xia et al., 2010). We compared the number of breaks per megabase for the X chromosome and all chromosomes to understand the differences in the dynamics of chromosome evolution between *Drosophila* and *Anopheles*. These results reveal a higher ratio of the rates of evolution of sex chromosome to all chromosomes in *Anopheles* than *Drosophila*, with means of 2.116 and 1.197, respectively (Figure C-8B). We correlated densities of different molecular features including simple repeats, TEs, genes, and S/MARs with the rates of rearrangement calculated for each arm. The strongest correlations were found among the

rates of evolution across all chromosome arms and the densities of microsatellites, minisatellites, and satellites in both *An. gambiae* and *An. stephensi*. The highly-positive correlations between rates of inversion across all chromosome arms and satellites of different sizes are due most likely to the co-occurring abundance of satellites and inversions on the X chromosome. Rates of inversions and satellite densities are much lower on the autosomes. S/MARs in autosomes were correlated negatively and genes correlated positively with polymorphic inversions.

Genetic diversity of the genome

The genome sequencing effort reported in the current study is based on an inbred laboratory strain to ensure good assembly. Nonetheless, we performed genome-wide SNP analysis based on the available data. A total of 530,997 SNPs were detected. A total of 319,751 SNPs were assigned to chromosomes based on mapping information. The SNP calls were assessed for their effect on the primary sequence of transcripts. These analyses will help future population genomic studies and facilitate association studies. We found that the X chromosome has a markedly lower frequency of SNPs than the autosomes in agreement with the similar observation in *An. gambiae* (Holt et al., 2002). The observed pattern may be explained by a smaller effective population size of the X chromosome due to male hemizygosity and lower sequence coverage of the X chromosome (Lawniczak et al., 2010).

Conclusions

The genome assembly of the type-form of the Indian strain of *An. stephensi* was produced using a combination of 454, Illumina, and PacBio sequencing and verified by analysis of BAC clones and ESTs. Physical mapping was in complete agreement with the genome assembly and resulted in a chromosome-based assembly that includes 62% of the genome. Such an assembly enabled analysis of chromosome arm-specific differences that are seldom feasible in next-gen genome projects.

Comparative analyses between *An. stephensi* and *An. gambiae* showed that the *Anopheles* X has a high rate of chromosomal rearrangement when compared with autosomes, despite the lack of polymorphic inversions in the X chromosomes in both species. Additionally, the difference between the rates of X chromosome and all chromosome evolution is much more striking in *Anopheles* than in *Drosophila.* The high rate of evolution on the X correlates well with the density of simple repeats. Our data indicate that overall high rates of chromosomal evolution are not restricted to *Drosophila* but may be a feature common to *Diptera*.

The genome landscape of *An. stephensi* is characterized by relatively low repeat content compared to *An. gambiae*. *An. stephensi* appears to have larger amount of repeat-rich heterochromatin in pericentric regions but far less repetitive sequences in chromosomal arms as compared with *An. gambiae*. Using a newly developed chromosome quotient method, we identified a number of Y-chromosome contigs and BACs, which together represent currently the most abundant set of Y sequences in any *Anopheles* species.

The current assembly contains 11,789 predicted protein coding genes, 127 miRNA genes, 434 tRNA genes, and 53 fragments of rRNA genes. *An. stephensi* appears to have fewer gene duplications than *An. gambiae* according to orthology analysis, which may explain the slightly lower number of gene models.

This genome project is accompanied by the first comprehensive RNA-seq-based transcriptomic analysis of an *Anopheles* mosquito. Twenty gene clusters were identified according to gene expression profiles, many of which are stage or sex-specific. GO term analysis of these gene clusters provided biological insights and leads for important research. For example, male-biased genes were enriched for genes involved in spermatogenesis and the auditory perception.

Close attention was paid to genes involved in innate immunity including LRIMs, APL1, and proteins in the Toll, IMD, insulin, and TGF-β signaling pathways. A high level of orthology is generally observed between *An. stephensi* and *An. gambiae*. RNA-seq analysis, which was corroborated by other expression analysis methods, provided novel insights. For example, a protein known to interact with both TOR and TGF-β signaling pathways showed abundant mRNA expression in a wide range of tissues, providing new leads for insights into both TOR and TGF-β signaling in mosquitoes.

Material and Methods

Strain selection

The Indian strain of *An. stephensi*, a representative of the type form was sequenced. The lab colony from which we selected mosquitoes for sequencing was originally established from wild mosquitoes collected in India. The lab colony has been maintained continuously for many generations so we did not attempt to inbreed it.

Sample collection

DNA was isolated from more than 50 adult male and female *An. stephensi* using the Qiagen (Hilden, Germany) DNeasy Blood and tissue kit following the suggested protocol. The integrity of the DNA was verified by running an aliquot on a 1% agarose gel to visualize any degradation. Total RNA was isolated using the standard protocol of the mirVana RNA isolation kit (Life Technologies, Carlsbad, CA, USA) and quality was verified using Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA).

Sequencing

The *An. stephensi* genome was sequenced to 19.4× coverage using 454 FLX Titanium sequencing performed by the Virginia Bioinformatics Institute (VBI) core laboratory. Sequencing was performed on four different libraries: a single-end shotgun library, and 3 kb, 8 kb, and 20 kb matepair libraries. A 200 bp insert size library produced from male *An. stephensi* genomic DNA was prepared and subjected to a single lane of Illumina HiSeq. Genomic DNA from male *An.* sequence was subjected to 10 SMRT cells of Pacific Biosciences (PacBio) v1 sequencing. Only males were sequenced with PacBio because we

are interested in increasing the probability of finding Y chromosome sequences. Sanger sequencing performed by Amplicon Express was used to sequence 7,263 BAC-ends.

Genome assembly

We used several approaches to combine the Illumina and 454 data to generate a better assembly. Newbler can take raw Illumina data as input, so we tried a Newbler assembly with the 454 and Illumina data. However, this resulted in a worse assembly than 454 alone. We had much more success with the strategy used to assemble the *Solenopsis invicta* genome (Wurm et al., 2011). We assembled the Illumina data first, and then cut the assembly into pseudo-454 reads. These reads were then used along with the real 454 data as input to Newbler (Kumar and Blaxter, 2010).

De novo Illumina assembly with Celera

We assembled the paired-end Illumina reads using the Celera assembler (Denisov et al., 2008) with the parameters: 'overlapper = ovl; unitigger = bogart; utgBubblePopping = 1; kickOutNonOvlContigs = 1; cgwDemoteRBP = 0; cgwMergeMissingThreshold = 0.5; merSize = 14'. The Celera assembler output comprised 41,213 contigs spanning 212.8 Mb. The N50 contig size of this assembly was 16.8 kb.

De novo 454 and Illumina pseudo-454 reads assembly with Newbler 2.8

The contigs of the aforementioned Illumina assembly were shredded informatically into 400 bp pieces with overlapping 200 bp to approximate 454 reads. To artificially simulate coverage depth, we started the shredding at offsets with the values of 0, 10, and 20.

Shredding the Illumina assembly resulted in 2,452,038 pseudo-454 reads simulating 4.17 × coverage.

We generated an assembly of the 454 and pseudo-454 reads with Newbler 2.8 using the '-het -scaffold -large -s 500' parameters. The resulting assembly contained 23,595 scaffolds spanned 221 Mb. The scaffold N50 size was 1.34 Mb. Mitochondrial DNA (1 scaffold), and other contamination (87 scaffolds) were identified by blastn and removed from the assembly.

Gap-filling with PacBio reads

PacBio data was used to fill gaps in the scaffolds to further improve the genome assembly. We error-corrected raw PacBio reads using the 454 sequencing data with the Celera pacBioToCa pipeline. pacBioToCa produced 0.88 Gb of error-corrected PacBio reads. Using the error-corrected PacBio data as input, Pbjelly (English et al., 2012) was used to fill gaps with parameters: '-minMatch 30 -minPctIdentity 98 -bestn 10 –n Candidates 5 -maxScore -500 -nproc 36-noSplitSubreads'. Pbjelly filled 1,310 gaps spanning 5.4 Mb.

Further scaffolding with BAC-ends

The scaffolds of the assembly were improved subsequently through the integration of 3,527 BAC-end pairs (120 kb ± 70 kb) using the Bambus scaffolder (Pop, Kosack, and Salzberg, 2004). The BAC-end sequences were mapped to the scaffolds using Nucmer (Kurtz et al., 2004). The output files were used to generate the '.contig' format files required for Bambus. In total, 275 links between scaffolds were detected. Of these, 169 were retained as potential

valid links, which are links connected by uniquely mapped BAC-ends. Links confirmed by less than two BAC-ends were rejected. A total of 46 links were retained that together connected 22 scaffolds, increasing the N50 scaffold size from 1,378 kb to 1,572 kb.

Assembly validation

CEGMA (Core Eukaryotic Genes) We used CEGMA (Parra, Bradnam, and Korf, 2007) to search for the number of core eukaryotic genes to test the completeness and correctness of the genome assembly. CEGMA provides additional information as to whether the entire core eukaryotic genes are present (>70%) or only partially present (>20% and <70%). In total, CEGMA found 96.37% of the 248 core eukaryotic genes to be present, and 97.89% of the core eukaryotic genes to be partially present.

BAC-ends

We checked whether BAC-ends align concordantly to the genome to study the structural correctness of the *de novo* assembly. BAC-ends were aligned to the scaffolds using NUCMER. In order to ensure unambiguous mapping, only sequences that aligned to a unique location with >95% coverage and 99% identity were used. In total, 21.6% of the BAC-end sequence pairs could be aligned to a unique position in the *An. stephensi* genome with these stringent criteria. Pairs of BAC-end sequence that aligned discordantly to a single scaffold were considered indicative of potential misassembly. Only four of 717 aligned BAC-end pairs aligned discordantly with the assembly confirming overall structural correctness.

ESTs

An. stephensi EST sequences were downloaded from both the NCBI and VectorBase. We screened the EST sequences to remove any residual vector sequence. The screened ESTs were aligned to the assembly with GMAP (Wu and Watanabe, 2005). In total, 35,367 of 36,064 ESTs aligned to the assembly. Of these, 26,638 aligned over at least 95% of their length with an identity of >98%. The high percentage of aligned ESTs demonstrates the near-completeness of the *An. stephensi* genome assembly.

Fluorescent *in situ* hybridization (FISH): Slides were prepared from ovaries of lab reared, half-gravid females of the *An. stephensi* Indian wild-type strain. Slide preparation and hybridization experiments followed the techniques described in Sharakhova *et al.* (Sharakhova et al., 2010). Fluorescent microscope images were converted to black and white and inverted in Adobe Photoshop. FISH signals were mapped to specific bands or interbands on the physical map for *An. stephensi* presented by Sharakhova *et al.* (Sharakhova et al., 2010).

Constructing the physical map

For the chromosomal based genome assembly, all probes mapped by *in situ* hybridization by Sharakhova (Sharakhova et al., 2010) and this study were aligned to the final version of the *An. stephensi* genome using NCBI blast + blastn. Different blastn parameters were used for probes from different sources to determine if the probe was kept in the final assembly. An e-value of 1e-40 and an identity of >95% was required for probes from *An. stephensi*. An e-value of 1e-5 was required for probes from species other than *An. stephensi*. Probes that mapped to more than one location in the genome were discarded. The work by

Sharakhova *et al.* (Sharakhova et al., 2010) hybridized 345 probes however, only approximately 200 probes from that study were maintained in the final chromosomal assembly. An additional 27 PCR products and BAC clones were hybridized to increase the coverage of our chromosomal assembly.

Annotation

The genome assembly was annotated initially using the MAKER pipeline (Cantarel et al., 2008). This software synthesizes the results from *ab initio* gene prediction with experimental gene evidence to produce final annotations. Within the MAKER framework, RepeatMasker (Tempel et al., 2012) was used to mask low-complexity genomic sequence based on the repeat library from previous prediction. First, ESTs and proteins were aligned to the genome by MAKER using BLASTn and BLASTx, respectively. MAKER uses the program Exonerate to polish BLAST hits. Next, within the MAKER framework, SNAP (Korf, 2004) and AUGUSTUS (Stanke et al., 2004) were run to produce *ab initio* gene predictions based on the initial training data. SNAP and AUGUSTUS were run once again inside of MAKER using the initial training obtained from the ESTs and protein alignments to produce the final annotations.

Orthology and molecular species phylogeny

Orthologs of predicted *An. stephensi* genes were assigned by OrthoDB (Waterhouse et al., 2013). Information about orthologous genes for *An. gambiae*, *Ae. aegypti*, and *D. melanogaster* also were downloaded from OrthoDB. Enrichment analysis was performed for categories of orthologs using the methods provided in the ontology section. The

molecular phylogeny of the 10 selected species was determined from the concatenated protein sequence alignments using MUSCLE (Edgar, 2004) (default parameters) followed by alignment trimming with trimAl (Capella-Gutierrez et al., 2009) (automated1 parameters) of 3,695 relaxed single-copy orthologs (a maximum of three paralogs allowed in no more than two species, longest protein selected) from OrthoDB (Waterhouse et al., 2013).. The resulting 2,246,060 amino acid columns with 932,504 distinct alignment patterns was analyzed with RAxML (Stamatakis, 2014) with the PROTGAMMAJTT model to estimate the maximum likelihood species phylogeny with 100 bootstrap samples.

Transcriptomics

RNA-seq from 11 samples including: 0 to 1, 2 to 4, 4 to 8, and 8 to 12 h embryos, larva, pupa, adult males, adult females, non-blood-fed ovaries, blood-fed ovaries, and female carcasses without ovaries as described (Criscione et al., 2013) were used for transcriptome analysis. These RNA-seq samples are available from the NCBI SRA (SRP013839). Tophat (Trapnell, Pachter, and Salzberg, 2009) was used to align these RNA-seq reads to the *An. stephensi* genome and HTSeq-count (Anders, Pyl, and Huber, 2014) was used to generate an occurrence table for each gene in each sample. The numbers of alignments to each gene in each sample then were clustered using MBCluster.Seq (Si et al., 2014), an R package designed to cluster genes by expression profile based on Poisson or Negative-Binomial models. MBCluster.Seq generated 20 clusters. To visualize these results we performed regularized log transformation to the original occurrence tables for all 20 clusters using DESeq2 (Love, Huber, and Anders, 2014). The results were plotted using ggplot2 (Wickham, 2011).

Ontology

Gene ontology (GO) terms were assigned for the 20 clusters of predicted *An. stephensi* genes. GO terms were assigned using Blast2Go (Conesa et al. 2005). The predicted proteins are blasted against the NCBI non-redundant protein database and scanned with InterProScan (Quevillon et al., 2005) against InterPro's signatures. After GO terms were assigned, GO-slim results were generated for the available annotation based on the Generic GO slim mapping. The GO terms assigned by Blast2GO were subject to GO term enrichment. Overrepresented GO terms were identified using a hypergeometric test using the GOstats package in R (Falcon and Gentleman, 2007).

Functional annotation of key gene families

We obtained the InterPro ID information for proteins in *An. stephensi* from the ontology analysis. We functionally annotated gene families based on the assigned InterPro ID. The gene families, including genes involved in immunity, chemosensation, and detoxification were studied. For comparative genome analysis, we retrieved the InterPro ID for seven other species (*An. gambiae*, *An. darlingi*, *Ae. aegypti*, *Culex quinquefasciatus*, *D. melanogaster*, *Bombyx mori*, and *Tribolium castaneum*) using Biomart (Kasprzyk, 2011) from vectorbase (VectorBase) and Ensembl Metazoa (Ensemble). We compared gene numbers in gene families of interest. For gene families with obvious differences in numbers between *An. stephensi* and *An. gambiae*, we preformed phylogenetic analysis of these genes. First we aligned these genes from *Anopheles* species using MUSCLE (Edgar, 2004).

Then, we constructed phylogenetic tree using Neighbor-joining method with 1,000 bootstrap replicates by CLC Genomics Workbench 4 (CLC bio).

Non-coding RNA

We used tRNAScan-SE (Lowe and Eddy, 1997) with the default eukaryotic mode to predict 434 tRNAs in the *An. stephensi* genome. Other non-coding RNAs were predicted with INFERNAL (Nawrocki, Kolbe, and Eddy, 2009) by searching against Rfam database version 11.0 (Griffiths-Jones et al., 2003). A total of 53 fragmental ribosomal RNA, 34 snRNA, 7 snoRNA, 127 miRNA, and 148 sequences with homology to the *An. gambiae* self-cleaving riboswitch were predicted with an e-value cutoff of 1e-5.

Transposable elements and other interspersed repeats Transposable element discovery and classification were performed on the *An. stephensi* scaffold sequences using previously-described pipelines for LTR-retrotransposons, non-LTR-retrotransposons, SINEs, DNA-transposons, and MITEs, followed by manual inspection (Nene et al., 2007). The manually-annotated TE libraries then were compared with the RepeatModeler output to remove redundancy and to correct mis-classification by RepeatModler. A repeat library was produced that contains all manually-annotated TEs and non-redundant sequences from RepeatModeler. The repeat library was used to run RepeatMasker at default settings on the *An. stephensi* assembly to calculate TE copy number and genome occupancy.

Simple repeats

The number of microsatellites, minisatellites, and satellites present in the mapped scaffolds for each chromosome were derived by dividing the scaffolds into strings of 100,000 bp and then concatenating them into a multiFASTA file to represent an *An. stephensi* pseudo chromosome. Scaffolds were oriented when possible, and all unoriented scaffolds were given the default positive orientation for that chromosome. The multiFASTA file for each pseudo-chromosome was analyzed using a local copy of TandemRepeatsFinder v 4.07b (Benson, 1999). Parameters for the analysis followed those used by Xia *et al.* (Xia et al., 2010): microsatellites were those of period size 2 to 6 with copy number of >8. Minisatellites had period size 7 to 99 while repeats were considered satellites if they had a period size of >100. Both satellites and minisatellites were considered only if they had a copy number of >2. Simple repeats were recorded only if they had at least 80% identity.

Identification of S/MARs

Scaffold/matrix associated regions were identified using the SMARTest bioinformatic tool provided by Genomatix (Frisch et al., 2002). Densities of genes and TEs per 100 kb window were calculated using Bedtools coverage based on the genome annotation and TE annotation, respectively.

Synteny, gene order evolution, and inversions

One-to-one orthologs from *An. gambiae* and *An. stephensi* were identified using OrthoDB (Waterhouse et al., 2013) and their locations on the *An. gambiae* and *An. stephensi* scaffolds determined. Comparative positions of the genes on the scaffolds based on ontology relationships were plotted using genoPlotR (Guy, Kultima, and Andersson, 2010). Scaffolds

that mapped using two or more probes were oriented properly, but those anchored by only one probe were used in their default orientation. The number of synteny blocks for each pair of homologous chromosome arms between *An. stephensi* and *An. gambiae* was determined from the images output from genoPlotR. Two criteria were imposed to determine the number of synteny blocks: the orientation of two or more orthologous genes, and whether the genes remained in the same order on the chromosome of *An. stephensi* as in *An. gambiae*. Thus, a group of two or more genes is assigned to the same synteny block if it has the same orientation and order in both species. Synteny blocks were numbered 1, 2, 3, 4, and so on along the chromosome by assigning *An. gambiae* as the default gene order. *An. stephensi* was considered rearranged compared to *An. gambiae* when the numbering of synteny blocks was the same in both species but the order was rearranged in *An. stephensi*. After quantifying the number of synteny blocks and the amount of gene rearrangement between the two species, we estimated the number of chromosomal inversions between them using the programs Genome Rearrangements in Mouse and Man (GRIMM (Tesler, 2002)).

SNP analysis

We used CLC Genomics Workbench 4 (CLC bio) to identify SNPs using a combination of the male and female Illumina data (Accession number: SRP013838). The required coverage was 20 and minimum variant frequency was 35. SNP calls made on the assembly were assessed for their effect on transcripts from the gene build using the Ensembl e-hive, variation database, and variation consequence pipeline (available from github (EnsEMBL) and (The Ensemble Variation Perl)). The Ensembl variation consequence pipeline uses the

Ensembl API in the same manner as the Variant Effect Predictor (McLaren et al., 2010) and produces equivalent output. The variation consequence pipeline directly loaded the analysis results into an Ensembl MySQL variation database which was used to generate summary statistics of transcript consequences classified using Sequence Ontologs (Eilbeck et al., 2005).

Data access

The *An. stephensi* genome assembly has been deposited in GenBank under the accession number ALPR00000000 and is available at (VectorBase). The raw sequence data used for genome assembly are available in the NCBI SRA: 454 SRP037783, Illumina SRP037783, and PacBio SRP037783. The BAC-ends used for scaffolding are available from the NCBI dbGSS accession numbers: KG772729 KG777469. RNA-Seq data can be accessed at the NCBI SRA with ID SRP013839.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

Conceived and designed experiments: ZT and IVS; Data generation, analysis, and presentation: XJ, AP, AS, ABH, MK, MVS, AK, BW, CO, DL, KE, KM, JMCT, JMCR, MAR, MRR, MU, NP, PA, PG, RK, RS, RMW, SL, SM, VLMD, YQ, and ZT; Writing of the manuscript: XJ, ABH, AAJ, AP, AS, JMCR, KDV, KM, KP, MK, MAR, MMR, SL, IVS, and ZT; Provided resources and tools and critical reviewed manuscript: XC and YS. All authors read and approved the final manuscript.

Acknowledgements

This work is supported by the Fralin Life Science Institute and the Virginia Experimental Station, and by NIH grants AI77680 and AI105575 to ZT, AI094289 and AI099528 to IVS, AI29746 to AAJ, AI095842 to KM, AI073745 to MAR, AI080799 and AI078183 to SL, and AI042361 and AI073685 to KDV. AP and IVS are supported in part by the Institute for Critical Technology and Applied Science and the NSF award 0850198. RMW is supported by Marie Curie International Outgoing Fellowship PIOF-GA-2011-303312. XC is supported by GDUPS (2009). This work was also supported in part by NSF grant CNS-0960081 and the HokieSpeed and BlueRidge supercomputers at Virginia Tech. YS thanks the Department of Biotechnology, Government of India for the financial support. Drew Cocrane provided assistance with *in situ* hybridization.

Author details

[1]Program of Genetics, Bioinformatics, and Computational Biology, Virginia Tech, Blacksburg, VA, USA.

[2]Department of Biochemistry, Virginia Tech, Blacksburg, VA, USA.

[3]Department of Entomology, Virginia Tech, Blacksburg, VA, USA.

[4]Department of Pathogen Biology, Southern Medical University, Guangzhou, Guangdong, China.

[5]Department of Genetic Medicine and Development, University of Geneva Medical School, rue Michel-Servet 1,1211 Geneva, Switzerland.

[6]Swiss Institute of Bioinformatics, rue Michel-Servet 1, 1211 Geneva, Switzerland.

[7]Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, 32 Vassar Street, Cambridge, MA, USA.

[8]The Broad Institute of MIT and Harvard, 7 Cambridge Center, Cambridge, MA, USA.

[9]Theodosius Dobzhansky Center for Genome Bioinformatics, St. Petersburg State University, St. Petersburg, Russia.

[10]Institute of Cytology Russian Academy of Sciences, St. Petersburg, Russia.

[11]Department of Microbiology, University of Minnesota, Minneapolis, MN, USA.

[12]National Center for Cell Science, Pune University Campus, Ganeshkhind, Pune, India.

[13]European Molecular Biology Laboratory, European Bioinformatics Institute, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SD, UK.

[14]Department of Medical Microbiology and Immunology, University of California, Davis, CA, USA.

[15]Biological Sciences Department, California State Polytechnic University, Pomona, CA, USA.

[16]Division of Biology, Kansas State University, Manhattan, KS, USA.

[17]Department of Parasitology and Mycology, Unit of Insect Vector Genetics and Genomics, Institut Pasteur, Paris, France.

[18]CNRS Unit of Hosts, Vectors
and Pathogens (URA3012), Paris, France.

[19]Department of Computer Science and Engineering, University of Notre Dame, Notre Dame, IN, USA.

[20]Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco, CA, USA.

[21]Virginia Bioinformatics Institute, Virginia Tech, Blacksburg, VA, USA.

[22]Department of Entomology, University of Arizona, Tucson, AZ, USA.

[23]Department of Physiology, School of Medicine – CIMUS, Instituto de Investigaciones Sanitarias, University of Santiago de Compostela, Santiago de Compostela, Spain.

[24]Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, UK.

[25]Department of Biological Sciences, University of Notre Dame, Notre Dame, IN, USA.

[26]Section of Vector Biology, Laboratory of Malaria and Vector Research, National Institute of Allergy and Infectious Diseases, Rockville, MD, USA.

[27]Departments of Microbiology & Molecular Genetics and Molecular Biology & Biochemistry, University of California, Irvine, CA, USA.

Received: 27 April 2014 Accepted: 3 September 2014

Published online: 23 September 2014

References

Alonso PL, Brown G, Arevalo-Herrera M, Binka F, Chitnis C, Collins F, Doumbo OK, Greenwood B, Hall BF, Levine MM, Mendis K, Newman RD, Plowe CV, Rodríguez MH, Sinden R, Slutsker L, Tanner M: A research agenda to underpin malaria eradication. *PLoS Med* **2011**, 8:e1000406.

Anders S, Pyl PT, Huber W: HTSeq A Python framework to work with high-throughput sequencing data. *Bioinformatics* **2014**. doi:10.1093/ bioinformatics/btu638.

Anderson KV, Bokla L, Nüsslein-Volhard C: Establishment of dorsal-ventral polarity in the *Drosophila* embryo: the induction of polarity by the Toll gene product. *Cell* **1985**, 42:791–798.

Arca B, Lombardo F, Valenzuela JG, Francischetti IM, Marinotti O, Coluzzi M, Ribeiro JM: An updated catalogue of salivary gland transcripts in the adult female mosquito, *Anopheles gambiae*. *J Exp Biol* **2005**, 208:3971–3986.

Arsic D, Guerin PM: Nutrient content of diet affects the signaling activity of the insulin/target of rapamycin/p70 S6 kinase pathway in the African malaria mosquito *Anopheles gambiae*. *J Insect Physiol* **2008**, 54:1226–1235.

Baricheva EA, Berrios M, Bogachev SS, Borisevich IV, Lapik ER, Sharakhov IV, Stuurman N, Fisher PA: DNA from *Drosophila melanogaster* β-heterochromatin binds specifically to nuclear lamins in vitro and the nuclear envelope in situ. *Gene* **1996**, 171:171–176.

Baker DA, Nolan T, Fischer B, Pinder A, Crisanti A, Russell S: A comprehensive gene expression atlas of sex and tissue-specificity in the malaria vector. *Anopheles gambiae BMC Genomics* **2011**, 12:296.

Baker RH, Sakai RK: Triploids and male determination in the mosquito, *Anopheles culicifacies*. *J Hered* **1979**, 70:345–346.

Benson G: Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res* **1999**, 27:573–580.

Bhutkar A, Schaeffer SW, Russo SM, Xu M, Smith TF, Gelbart WM: Chromosomal rearrangement inferred from comparisons of 12 *Drosophila* genomes. *Genetics* **2008**, 179:1657–1680.

Bian G, Joshi D, Dong Y, Lu P, Zhou G, Pan X, Xu Y, Dimopoulos G, Xi Z: *Wolbachia* invades *Anopheles stephensi* populations and induces refractoriness to *Plasmodium* infection. *Science (New York, NY)* **2013**, 340:748–751.

Bourque G, Pevzner PA: Genome-scale evolution: reconstructing gene orders in the ancestral species. *Genome Res* **2002**, 12:26–36.

Brown AE, Bugeon L, Crisanti A, Catteruccia F: Stable and heritable gene silencing in the malaria vector *Anopheles stephensi*. *Nucleic Acids Res* **2003**, 31:e85.

Cantarel BL, Korf I, Robb SM, Parra G, Ross E, Moore B, Holt C, Sanchez Alvarado A, Yandell M: MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Res* **2008**, 18:188–196.

Capella-Gutierrez S, Silla-Martinez JM, Gabaldon T: trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* **2009**, 25:1972–1973.

Chaisson MJ, Raphael BJ, Pevzner PA: Microinversions in mammalian evolution. *Proc Natl Acad Sci U S A* **2006**, 103:19824–19829.

Chin CS, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J: Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat Methods* **2013**, 10:563–569.

Choi J, Chen J, Schreiber SL, Clardy J: Structure of the FKBP12-rapamycin complex interacting with the binding domain of human FRAP. *Science (New York, NY)* **1996**, 273:239–242.

CLC bio, a QIAGEN Company. [http://www.clcbio.com]

Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M: Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research.

Bioinformatics **2005**, 21:3674–3676.

Criscione F, Qi Y, Saunders R, Hall B, Tu Z: A unique Y gene in the Asian malaria mosquito *Anopheles stephensi* encodes a small lysine-rich protein and is transcribed at the onset of embryonic development. *Insect Mol Biol* **2013**, 22:433–441.

Dechat T, Pfleghaar K, Sengupta K, Shimi T, Shumaker DK, Solimando L, Goldman RD: Nuclear lamins: major factors in the structural organization and function of the nucleus and chromatin. *Genes Dev* **2008**, 22:832–853.

Denisov G, Walenz B, Halpern AL, Miller J, Axelrod N, Levy S, Sutton G: Consensus generation and variant detection by Celera Assembler. *Bioinformatics* **2008**, 24:1035–1040.

Dong Y, Das S, Cirimotich C, Souza-Neto JA, McLean KJ, Dimopoulos G: Engineered *Anopheles* immunity to *Plasmodium* infection. *PLoS Pathog* **2011**, 7:e1002458.

Edgar RC: MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **2004**, 32:1792–1797.

Eilbeck K, Lewis SE, Mungall CJ, Yandell M, Stein L, Durbin R, Ashburner M: The Sequence Ontology: a tool for the unification of genome annotations. *Genome Biol* **2005**, 6:R44.

English AC, Richards S, Han Y, Wang M, Vee V, Qu J, Qin X, Muzny DM, Reid JG, Worley KC, Gibbs RA: Mind the gap: upgrading genomes with Pacific Biosciences RS long-read sequencing technology. *PLoS One* **2012**, 7:e47768.

Ensembl Metazoa. [http://metazoa.ensembl.org]

EnsEMBL Hive a system for creating and running pipelines on a distributed compute resource. [https://github.com/Ensembl/ensembl-hive]

The Ensembl Variation Perl API and SQL schema. [https://github.com/ Ensembl/ensembl-variation/]

Falcon S, Gentleman R: Using GOstats to test gene lists for GO term association. *Bioinformatics* **2007**, 23:257–258.

Feachem RGA, Phillips AA, Hwang J, Cotter C, Wielgosz B, Greenwood BM, Sabot O, Rodriguez MH, Abeyasinghe RR, Ghebreyesus TA, Snow RW: Shrinking the malaria map: Progress and prospects. *Lancet* **2010**, 376:1566–1578.

Frisch M, Frech K, Klingenhoff A, Cartharius K, Liebich I, Werner T: In silico prediction of scaffold/matrix attachment regions in large genomic sequences. *Genome Res* **2002**, 12:349–354.

Faulde MK, Rueda LM, Khaireh BA: First record of the Asian malaria vector *Anopheles stephensi* and its possible role in the resurgence of malaria in Djibouti, Horn of Africa. *Acta Trop* **2014**, 139C:39–43.

Gakhar SK, Sharma R, Sharma A: Population genetic structure of malaria vector *Anopheles stephensi* Liston (Diptera: Culicidae). *Indian J Exp Biol* **2013**, 51:273–279.

Garver LS, Dong Y, Dimopoulos G: Caspar controls resistance to *Plasmodium falciparum* in diverse anopheline species. *PLoS Pathog* **2009**, 5:e1000335

Gibson G, Warren B, Russell IJ: Humming in tune: sex and species recognition by mosquitoes on the wing. *J Assoc Res Otolaryngol* **2010**, 11:527–540.

Göpfert MC, Robert D: Active auditory mechanics in mosquitoes. *Proc Roy Soc Lond B Biol Sci* **2001**, 268:333–339.

Grewal SS: Insulin/TOR signaling in growth and homeostasis: A view from the fly world. *Int J Biochem Cell Biol* **2009**, 41:1006–1010.

Griffiths-Jones S, Bateman A, Marshall M, Khanna A, Eddy SR: Rfam: an RNA family database. *Nucleic Acids Res* **2003**, 31:439–441.

Guy L, Kultima JR, Andersson SG: genoPlotR: comparative gene and genome visualization in R. *Bioinformatics* **2010**, 26:2334–2335.

Hall AB, Qi Y, Timoshevskiy V, Sharakhova MV, Sharakhov IV, Tu Z: Six novel Y chromosome genes in *Anopheles* mosquitoes discovered by independently sequencing males and females. *BMC Genomics* **2013**, 14:273.

Hoffmann AA, Sgrò CM, Weeks AR: Chromosomal inversion polymorphisms and adaptation.

Trends Ecol Evol **2004**, 19:482–488.

Holt RA, Subramanian GM, Halpern A, Sutton GG, Charlab R, Nusskern DR, Wincker P, Clark AG, Ribeiro JMC, Wides R, Salzberg SL, Loftus B, Yandell M, Majoros WH, Rusch DB, Lai Z, Kraft CL, Abril JF, Anthouard V, Arensburger P, Atkinson PW, Baden H, de Berardinis V, Baldwin D, Benes V, Biedler J, Blass C, Bolanos R, Boscus D, Barnstead M, *et al*: The genome sequence of the malaria mosquito *Anopheles gambiae*. *Science (New York, NY)* **2002**, 298:129–149.

Horton AA, Wang B, Camp L, Price MS, Arshi A, Nagy M, Nadler SA, Faeder JR, Luckhart S: The mitogen-activated protein kinome from *Anopheles gambiae*: identification, phylogeny and functional characterization of the ERK, JNK and p38 MAP kinases. *BMC Genomics* **2011**, 12:574.

Hunter S, Jones P, Mitchell A, Apweiler R, Attwood TK, Bateman A, Bernard T, Binns D, Bork P, Burge S, de Castro E, Coggill P, Corbett M, Das U, Daugherty L, Duquenne L, Finn RD, Fraser M, Gough J, Haft D, Hulo N, Kahn D, Kelly E, Letunic I, Lonsdale D, Lopez R, Madera M, Maslen J, McAnulla C, McDowall J, *et al*: InterPro in 2011: new developments in the family and domain prediction database. *Nucleic Acids Res* **2012**, 40:D306–D312.

Kasprzyk A: BioMart: driving a paradigm change in biological data management. *Database (Oxford)* 2011, **2011**:bar049.

Korf I: Gene finding in novel genomes. *BMC Bioinformatics* **2004**, 5:59.

Kumar S, Blaxter ML: Comparing de novo assemblers for 454 transcriptome data. *BMC Genomics* **2010**, 11:571.

Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu C, Salzberg SL: Versatile and open software for comparing large genomes. *Genome Biol* **2004**, 5:R12.

Lawniczak MK, Emrich SJ, Holloway AK, Regier AP, Olson M, White B, Redmond S, Fulton L, Appelbaum E, Godfrey J, Farmer C, Chinwalla A, Yang SP, Minx P, Nelson J, Kyung K, Walenz BP, Garcia-Hernandez E, Aguiar M, Viswanathan LD, Rogers YH, Strausberg RL, Saski CA, Lawson D, Collins FH, Kafatos FC,

Christophides GK, Clifton SW, Kirkness EF, Besansky NJ: Widespread divergence between incipient *Anopheles gambiae* species revealed by whole genome sequences. *Science* **2010**, 330:512–514.

Love MI, Huber W, Anders S: Moderated estimation of fold change and dispersion for RNA-Seq data with DESeq2. **2014**.

Lowe TM, Eddy SR: tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res* **1997**, 25:955–964.

Isaacs AT, Jasinskiene N, Tretiakov M, Thiery I, Zettor A, Bourgouin C, James AA: PNAS Plus: Transgenic *Anopheles stephensi* coexpressing single-chain antibodies resist *Plasmodium falciparum* development. *Proc Natl Acad Sci U S A* **2012**, 109:E1922–E1930.

Laplante M, Sabatini DM: mTOR signaling in growth control and disease. *Cell* **2012**, 149:274–293.

Luckhart S, Giulivi C, Drexler AL, Antonova-Koch Y, Sakaguchi D, Napoli E, Wong S, Price MS, Eigenheer R, Phinney BS, Pakpour N, Pietri JE, Cheung K, Georgis M, Riehle M: Sustained activation of Akt elicits mitochondrial dysfunction to block *Plasmodium falciparum* infection in the mosquito host. *PLoS Pathog* **2013**, 9:e1003180.

Mahmood F, Sakai RK: Inversion polymorphisms in natural populations of *Anopheles stephensi*. *Can J Genet Cytol* **1984**, 26:538–546.

Marinotti O, Cerqueira GC, de Almeida LG, Ferro MI, Loreto EL, Zaha A, Teixeira SM, Wespiser AR, Almeida ESA, Schlindwein AD, Pacheco AC, Silva AL, Graveley BR, Walenz BP, Lima Bde A, Ribeiro CA, Nunes-Silva CG, de Carvalho CR, Soares CM, de Menezes CB, Matiolli C, Caffrey D, Araújo DA, de Oliveira DM, Golenbock D, Grisard EC, Fantinatti-Garboggini F, de Carvalho FM, Barcellos FG, Prosdocimi F, *et al*: The genome of *Anopheles darlingi*, the main neotropical malaria vector. *Nucleic Acids Res* **2013**, 41:7387–7400.

McLaren W, Pritchard B, Rios D, Chen Y, Flicek P, Cunningham F: Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor.

Bioinformatics **2010**, 26:2069–2070.

Mitri C, Thiery I, Bourgouin C, Paul REL: Density-dependent impact of the human malaria parasite *Plasmodium falciparum* gametocyte sex ratio on mosquito infection rates. *Proc Roy Soc Lond B Biol Sci* **2009**, 276:3721–3726.

Murray CJL, Rosenfeld LC, Lim SS, Andrews KG, Foreman KJ, Haring D, Fullman N, Naghavi M, Lozano R, Lopez AD: Global malaria mortality between 1980 and 2010: A systematic analysis. *Lancet* **2012**, 2012:413–431.

Nawrocki EP, Kolbe DL, Eddy SR: Infernal 1.0: inference of RNA alignments. *Bioinformatics* **2009**, 25:1335–1337.

Neafsey DE, Christophides GK, Collins FH, Emrich SJ, Fontaine MC, Gelbart W, Hahn MW, Howell PI, Kafatos FC, Lawson D, Muskavitch MA, Waterhouse RM, Williams LJ, Besansky NJ: The evolution of the *Anopheles* 16 genomes project. *G3 (Bethesda)* **2013**, 3:1191–1194.

Nene V, Wortman JR, Lawson D, Haas B, Kodira C, Tu ZJ, Loftus B, Xi Z, Megy K, Grabherr M, Ren Q, Zdobnov EM, Lobo NF, Campbell KS, Brown SE, Bonaldo MF, Zhu J, Sinkins SP, Hogenkamp DG, Amedeo P, Arensburger P, Atkinson PW, Bidwell S, Biedler J, Birney E, Bruggner RV, Costas J, Coy MR, Crabtree J, Crawford M, *et al*: Genome sequence of *Aedes aegypti*, a major arbovirus vector. *Science* **2007**, 316:1718–1723.

Nolan T, Bower TM, Brown AE, Crisanti A, Catteruccia F: piggyBac-mediated germline transformation of the malaria mosquito *Anopheles stephensi* using the red fluorescent protein dsRED as a selectable marker. *J Biol Chem* **2002**, 277:8759–8762.

O'Brochta DA, Alford RT, Pilitt KL, Aluvihare CU, Harrell RA: *piggyBac* transposon remobilization and enhancer detection in *Anopheles* mosquitoes. *Proc Natl Acad Sci U S A* **2011**, 108:16339–16344.

Pakpour N, Corby-Harris V, Green GP, Smithers HM, Cheung KW, Riehle MA, Luckhart S: Ingested human insulin inhibits the mosquito NF-κB-dependent immune response to *Plasmodium falciparum*. *Infect Immun* **2012**, 80:2141–2149.

Parra G, Bradnam K, Korf I: CEGMA: a pipeline to accurately annotate core genes in eukaryotic genomes. *Bioinformatics* **2007**, 23:1061–1067.

Peng Q, Pevzner PA, Tesler G: The fragile breakage versus random breakage models of chromosome evolution. *PLoS Comput Biol* **2006**, 2:e14.

Pop M, Kosack DS, Salzberg SL: Hierarchical scaffolding with Bambus. *Genome Res* **2004**, 14:149–159.

Price I, Ermentrout B, Zamora R, Wang B, Azhar N, Mi Q, Constantine G, Faeder JR, Luckhart S, Vodovotz Y: In vivo, in vitro, and in silico studies suggest a conserved immune module that regulates malaria parasite transmission from mammals to mosquitoes. *J Theor Biol* **2013**, 334:173–186.

Quevillon E, Silventoinen V, Pillai S, Harte N, Mulder N, Apweiler R, Lopez R: InterProScan: protein domains identifier. *Nucleic Acids Res* **2005**, 33:W116–W120.

Rafinejad J, Vatandoost H, Nikpoor F, Abai MR, Shaeghi M, Duchen S, Rafi F: Effect of washing on the bioefficacy of insecticide-treated nets (ITNs) and long-lasting insecticidal nets (LLINs) against main malaria vector *Anopheles stephensi* by three bioassay methods. *J Vector Borne Dis* **2008**, 45:143–150.

Rai KS, Black WC IV: Mosquito genomes: structure, organization, and evolution. *Adv Genet* **1999**, 41:1–33.

Ranz JM, Maurin D, Chan YS, Von Grotthuss M, Hillier LW, Roote J, Ashburner M, Bergman CM: Principles of genome evolution in the *Drosophila melanogaster* species group. *PLoS Biol* **2007**, 5:1366–1381.

Ranz JM, Casals F, Ruiz A: How malleable is the eukaryotic genome? Extreme rate of chromosomal rearrangement in the genus *Drosophila*. *Genome Res* **2001**, 11:230–239.

Ribeiro JMC, Mans BJ, Arcà B: An insight into the sialome of blood-feeding Nematocera. *Insect Biochem Mol Biol* **2010**, 40:767–784.

Schaeffer SW, Bhutkar A, McAllister BF, Matsuda M, Matzkin LM, O'Grady PM, Rohde C, Valente VLS, Aguadé M, Anderson WW, Edwards K, Garcia AC, Goodman J,

Hartigan J, Kataoka E, Lapoint RT, Lozovsky ER, Machado CA, Noor MA, Papaceit M, Reed LK, Richards S, Rieger TT, Russo SM, Sato H, Segarra C, Smith DR, Smith TF, Strelets V, Tobari YN, *et al*: Polytene chromosomal maps of 11 *Drosophila* species: the order of genomic scaffolds inferred from genetic and physical maps. *Genetics* **2008**, 179:1601–1655.

Sharakhova MV, Xia A, Leman SC, Sharakhov IV: Arm-specific dynamics of chromosome evolution in malaria mosquitoes. *BMC Evol Biol* **2011**, 11:91.

Sharakhova MV, George P, Brusentsova IV, Leman SC, Bailey JA, Smith CD, Sharakhov IV: Genome mapping and characterization of the *Anopheles gambiae* heterochromatin. *BMC Genomics* **2010**, 11:459.

Sharakhova MV, Xia A, Tu Z, Shouche YS, Unger MF, Sharakhov IV: A physical map for an Asian malaria mosquito, Anopheles stephensi. *Am J Trop Med Hyg* **2010**, 83:1023–1027.

Sharakhov IV, Serazin AC, Grushko OG, Dana A, Lobo N, Hillenmeyer ME, Westerman R, Romero-Severson J, Costantini C, Sagnon N, Collins FH, Besansky NJ: Inversions and gene order shuffling in *Anopheles gambiae* and *A. funestus*. *Science* **2002**, 298:182–185.

Sharma VP: Current scenario of malaria in India. *Parassitologia* **1999**, 41:349–353.

Si Y, Liu P, Li P, Brutnell TP: Model-based clustering for RNA-seq data. *Bioinformatics* **2014**, 30:197–205.

Smidler AL, Terenzi O, Soichot J, Levashina EA, Marois E: Targeted mutagenesis in the malaria mosquito using TALE nucleases. *PLoS One* **2013**, 8:e74511.

Stamatakis A: RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **2014**, 30:1312–1313.

Stanke M, Steinkamp R, Waack S, Morgenstern B: AUGUSTUS: a web server for gene finding in eukaryotes. *Nucleic Acids Res* **2004**, 32:W309–W312.

Tempel S: Using and understanding RepeatMasker. *Methods Mol Biol* **2012**, 859:29–51.

Tesler G: GRIMM: genome rearrangements web server. *Bioinformatics* **2002**, 18:492–493.

Timoshevskiy VA, Kinney NA, de Bruyn BS, Mao C, Tu Z, Severson DW, Sharakhov IV, Sharakhova MV: Genomic composition and evolution of *Aedes aegypti* chromosomes revealed by the analysis of physically mapped supercontigs. *BMC Biol* **2014**, 12:27.

Trapnell C, Pachter L, Salzberg SL: TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* **2009**, 25:1105–1111.

Valenzuela JG, Francischetti IMB, Pham VM, Garfield MK, Ribeiro JMC: Exploring the salivary gland transcriptome and proteome of the *Anopheles stephensi* mosquito. *Insect Biochem Mol Biol* **2003**, 33:717–732.

VectorBase, *Anopheles stephensi* Indian strain. [https://www.vectorbase. org/Anopheles_stephensiI/Info/Index]

VectorBase, Gene AGAP012184 Expression Report. [http://funcgen. vectorbase.org/expression-browser/gene/AGAP012184]

Vectorbase. [https://www.vectorbase.org/]

Waterhouse RM, Tegenfeldt F, Li J, Zdobnov EM, Kriventseva EV: OrthoDB: a hierarchical catalog of animal, fungal and bacterial orthologs. *Nucleic Acids Res* **2013**, 41:D358–D365.

White MT, Conteh L, Cibulskis R, Ghani AC: Costs and cost-effectiveness of malaria control interventions–a systematic review. *Malar J* **2011**, 10:337.

Wickham H: ggplot2. *Wiley Interdiscipl Rev Comput Stat* **2011**, 3:180–185.

Wu TD, Watanabe CK: GMAP: a genomic mapping and alignment program for mRNA and EST sequences. *Bioinformatics* **2005**, 21:1859–1875.

Wurm Y, Wang J, Riba-Grognuz O, Corona M, Nygaard S, Hunt BG, Ingram KK, Falquet L, Nipitwattanaphon M, Gotzek D, Dijkstra MB, Oettler J, Comtesse F, Shih CJ, Wu WJ, Yang CC, Thomas J, Beaudoing E, Pradervand S, Flegel V, Cook ED, Fabbretti R, Stockinger H, Long L, Farmerie WG, Oakey J, Boomsma JJ, Pamilo P, Yi SV, Heinze J, *et al*: The genome of the fire ant *Solenopsis invicta*. *Proc*

Natl Acad Sci U S A **2011**, 108:5679–5684.

Xi Z, Ramirez JL, Dimopoulos G: The *Aedes aegypti* toll pathway controls dengue virus infection. *PLoS Pathog* **2008**, 4:e1000098.

Xia A, Sharakhova MV, Leman SC, Tu Z, Bailey JA, Smith CD, Sharakhov IV: Genome landscape and evolutionary plasticity of chromosomes in malaria mosquitoes. *PLoS One* **2010**, 5:e10592.

Zhou D, Zhang D, Ding G, Shi L, Hou Q, Ye Y, Xu Y, Zhou H, Xiong C, Li S, Yu J, Hong S, Yu X, Zou P, Chen C, Chang X, Wang W, Lv Y, Sun Y, Ma L, Shen B, Zhu C: Genome sequence of *Anopheles sinensis* provides insight into genetics basis of mosquito competence for malaria parasites. *BMC Genomics* **2014**, 15:42.

Figures

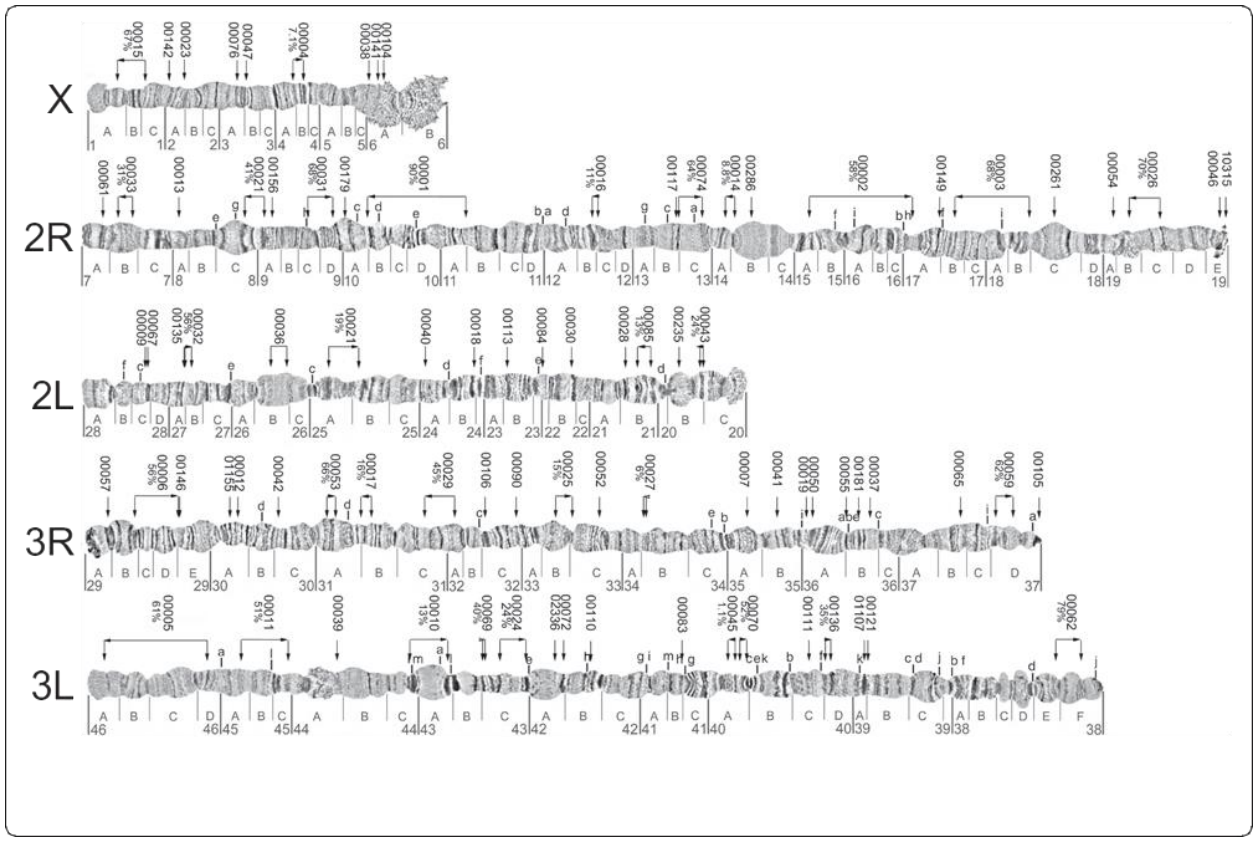

Figure C-1 Physical map

A physical map of the *An. stephensi* genome was created from FISH on polytene chromosomes comprising 227 probes and 86 scaffolds. These 86 scaffolds comprise 137.14 Mb or 62% of the *An. stephensi* genome. Orientation was assigned to 32 of the 86 scaffolds. The physical map includes 28 of the 30 largest scaffolds.

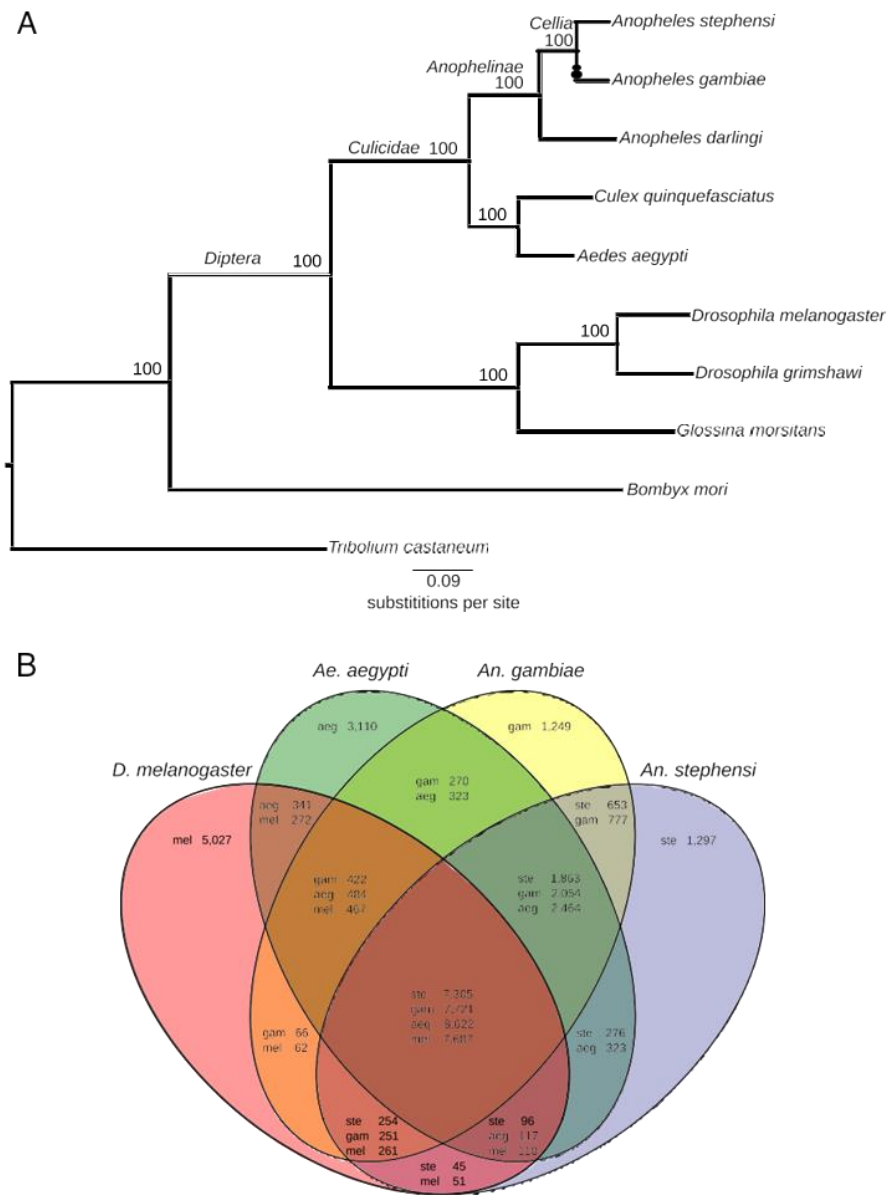

Figure C-2 Molecular species phylogeny and orthology

(A) The maximum likelihood molecular species phylogeny estimated from universal single-copy orthologs supports the recognized species relationships with *An. stephensi* and *An. gambiae* in subgenus Cellia within the genus *Anopheles*. (B) Comparative analysis of orthologs from *An. stephensi*, *An. gambiae*, *Ae. aegypti*, and *D. melanogaster*. Orthologous genes were retrieved from OrthoDB. A total of 7,305 genes were shared among all four species, 1,297 genes were

specific to *An. stephensi*, 653 genes were *Anopheles*-specific, and 1,863 genes were mosquito-specific.

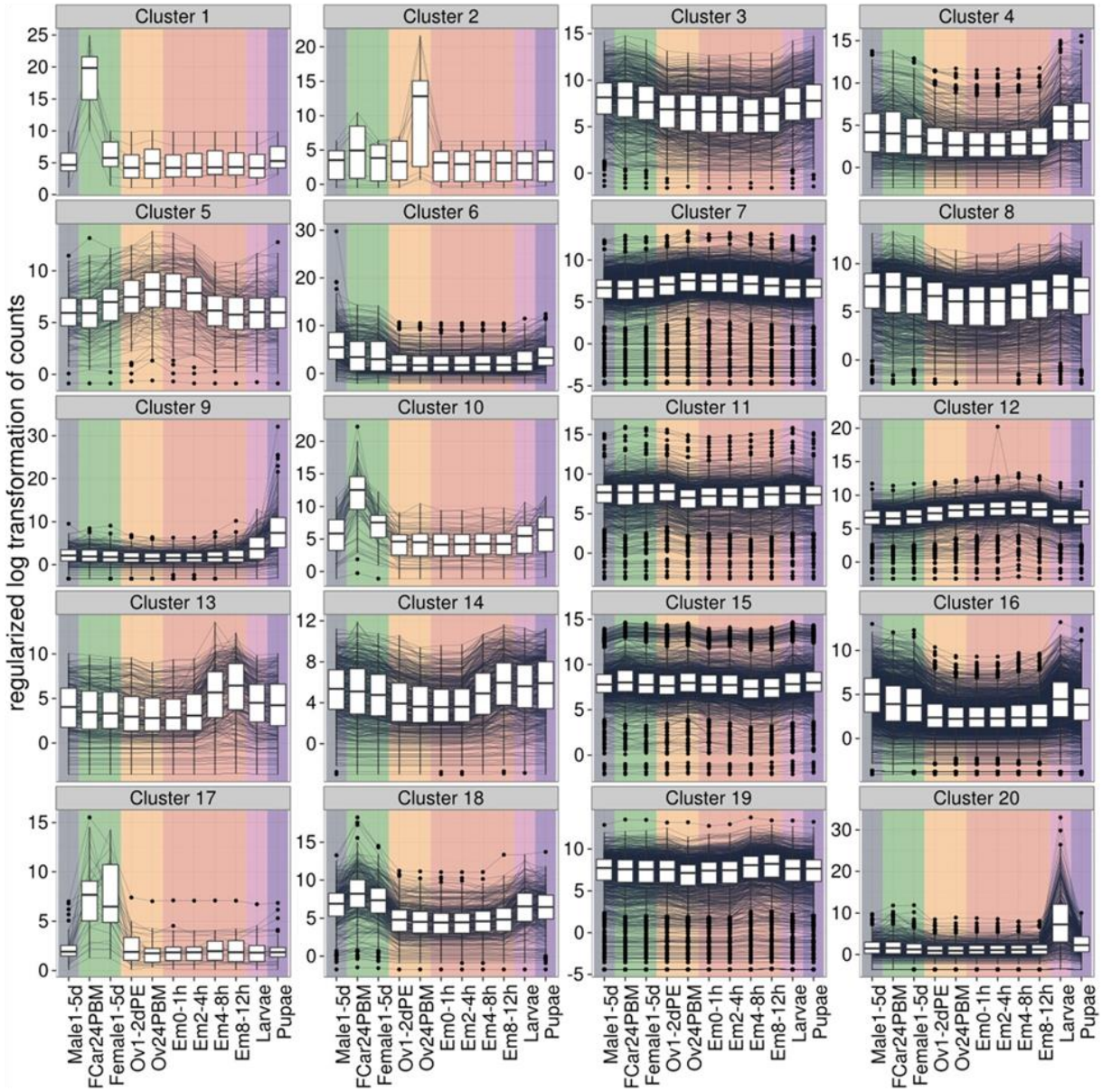

Figure C-3 Gene clustering according to expression profile

Twenty groups of genes were clustered by expression profile. The expression profiles used for grouping were generated using 11 RNA-seq samples spanning developmental time points including: 0 to 1, 2 to 4, 4 to 8, and 8 to 12 h embryos, larva, pupa, adult males, adult females, non-blood-fed ovaries, blood-fed ovaries, and 24 h post-blood-fed female carcass without ovaries. Male stage are colored blue, female stages are colored

green, ovary samples are colored yellow, embryo samples are colored red, larva samples are colored pink, and pupa samples are colored purple. Many of these clusters correspond to either a specific developmental stage or specific sex.

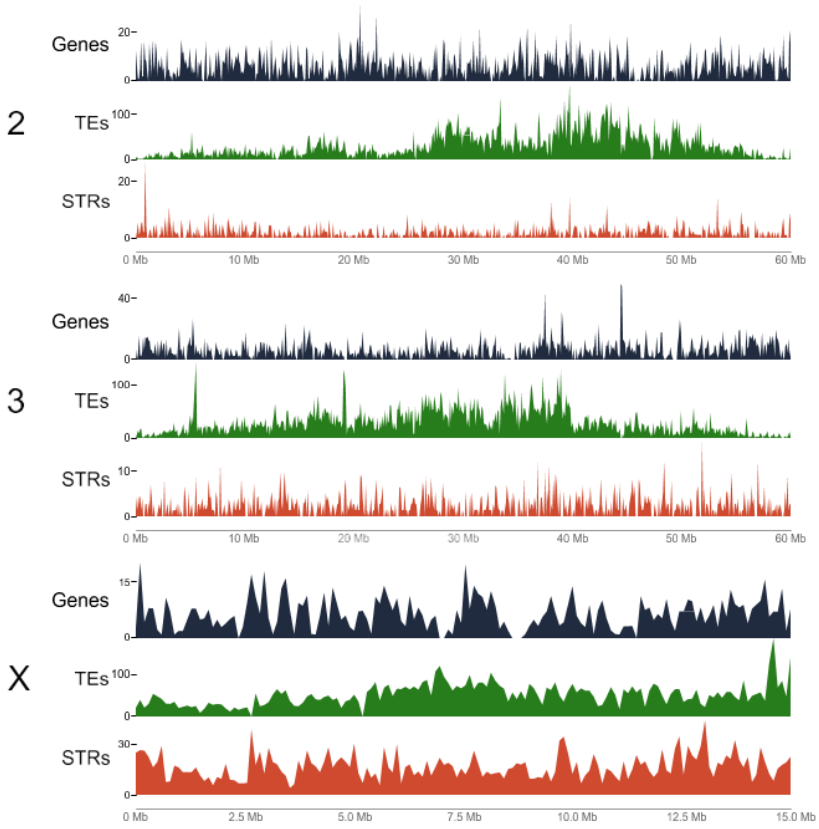

Figure C-4 Genome landscape

Density of genes (black vertical lines), transposable elements (TEs; green vertical lines), and short tandem repeats (STRs; red vertical lines) in 100 kb windows of mapped scaffolds. Based on the physical map, scaffolds were ordered and oriented respective to their position in the chromosomes and then 100 kb non-overlapping windows were generated for each scaffold (X-axis). The density of genes and TEs (Y-axis) was determined using coverageBed. Satellite sequences were identified using TandemRepeatFinder. The short tandem repeats track is a combination of the number of microsatellites, minisatellites, and satellites per 100 kb window.

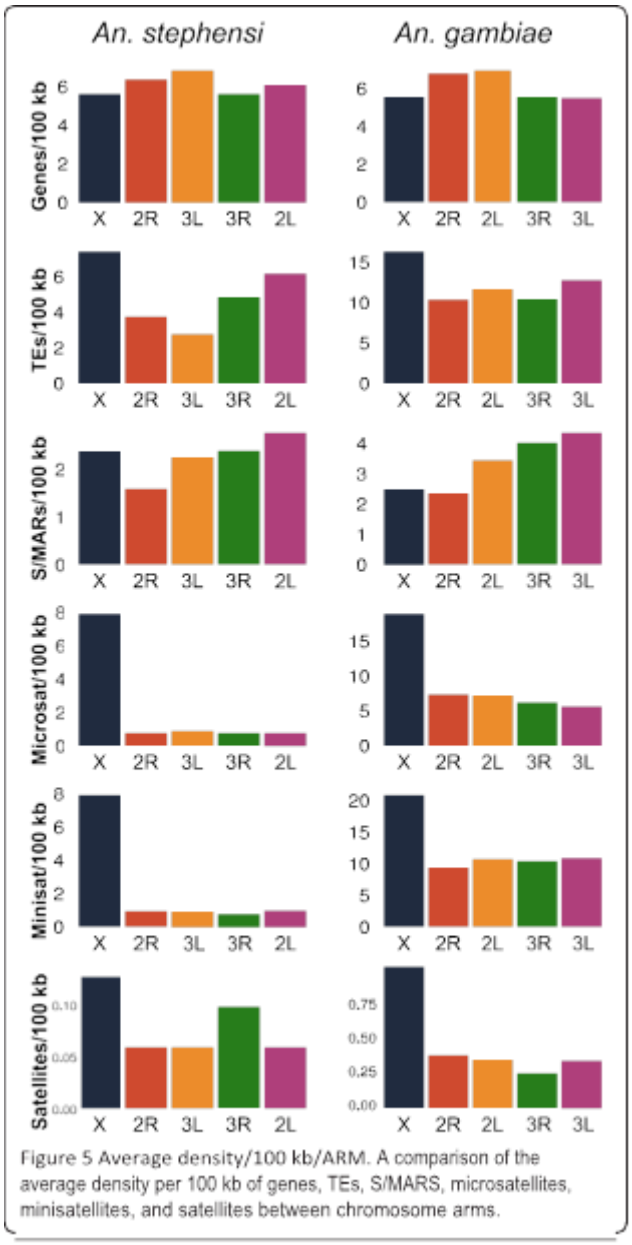

Figure 5 Average density/100 kb/ARM. A comparison of the average density per 100 kb of genes, TEs, S/MARS, microsatellites, minisatellites, and satellites between chromosome arms.

Figure C-5 Average density/100 kb/ARM

A comparison of the average density per 100 kb of genes, TEs, S/MARS, microsatellites, minisatellites, and satellites between chromosome arms.

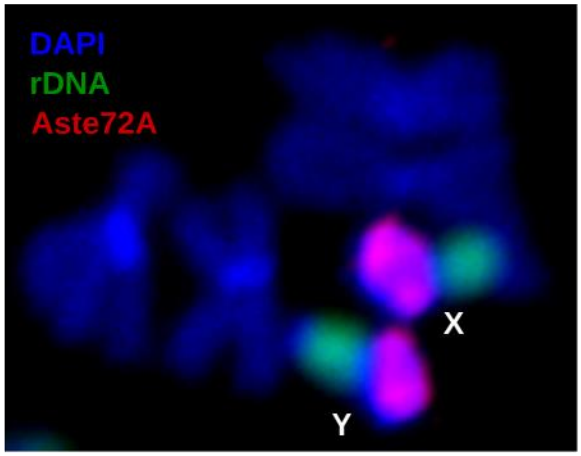

Figure C-6 FISH with Aste72A, rDNA, and DAPI on mitotic chromosomes

The pattern of hybridization for satellite DNA Aste72A on mitotic sex chromosomes of *An. stephensi*. Aste72A hybridizes to pericentric heterochromatin in both X and Y chromosomes while ribosomal DNA locus maps next to the heterochromatin band in sex chromosomes.

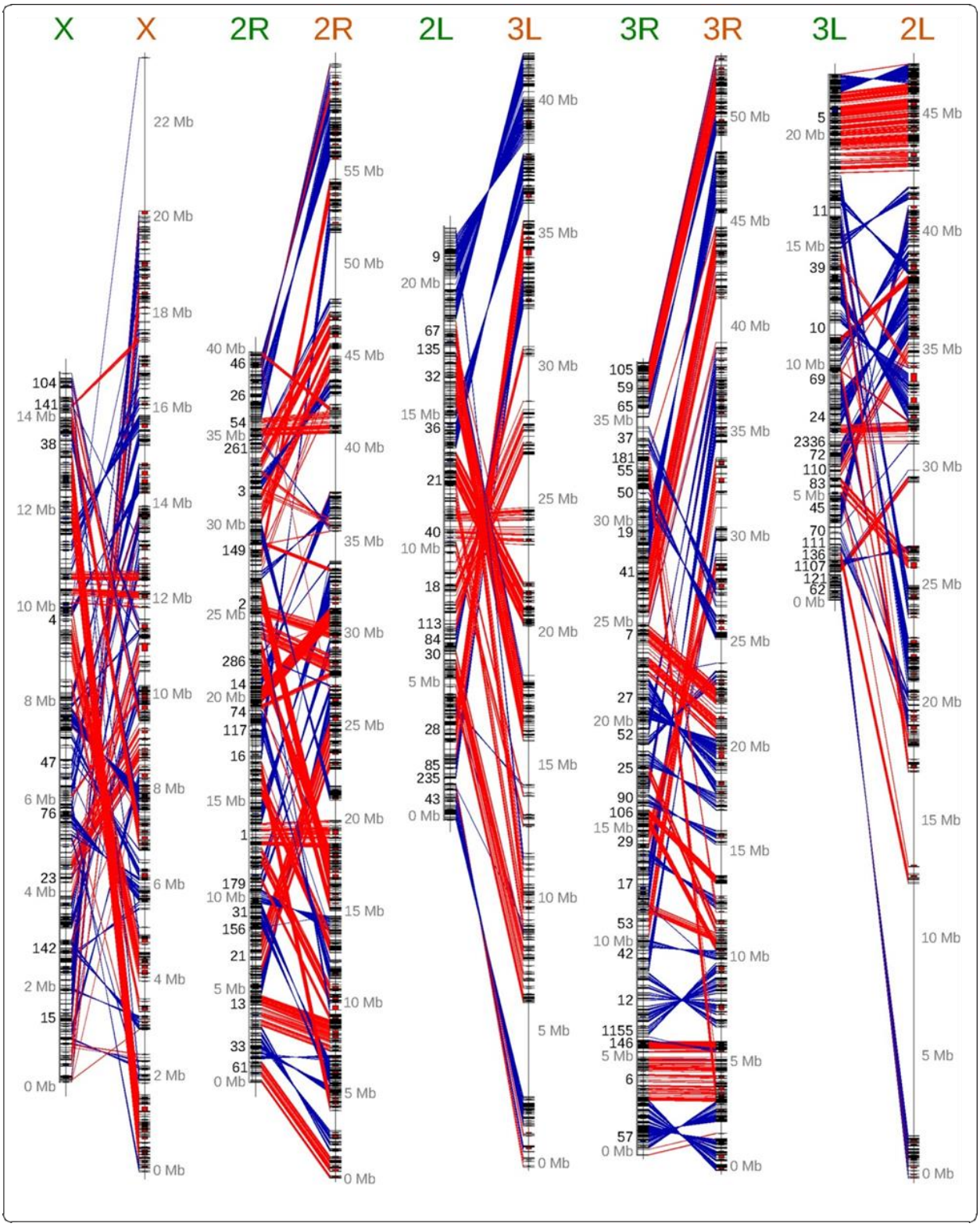

Figure C-7 Synteny

Synteny between *An. stephensi* and *An. gambiae* based on 6,448 single-copy orthologs. Orthologs with the same orientation in *An. stephensi* and *An. gambiae* are connected with red lines and orthologs with the opposite orientation are connected with blue lines. Orthologous genes from *An. stephensi* and *An. gambiae* were retrieved from OrthoDB. The physical map was used to identify the relative locations of genes on the *An. stephensi* chromosomes. The relationship of the position between the *An. stephensi* and *An. gambiae* orthologs were plotted with GenoPlotR. 66 syntenic blocks were identified on the X chromosome. A total of 104 and 64 syntenic blocks were identified on 2R and 2L (3L in *An. stephensi*). A total of 104 and 42 syntenic blocks were identified on 3R and 3L (2L in *An. stephensi*). Therefore, the X chromosome has undergone the most rearrangements per megabase.

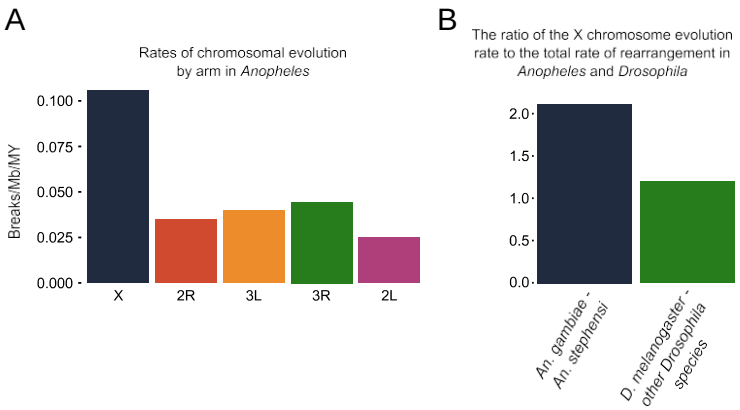

Figure C-8 Chromosome evolution in *Anopheles* and *Drosophila*

(A) Higher rates of rearrangement on the X chromosome compared to autosomes between *An. stephensi* and *An. gambiae*. Arm designations for the figure are according to *An. stephensi.* (B) The ratio of the X chromosome evolution rate to the total rate of rearrangement is higher in *Anopheles* than in *Drosophila*.