TRANSCRIPTOMIC ANALYSES OF *DROSOPHILA SUZUKII RESPONSES TO PESTICIDES* AND *ANOPHELES GAMBIAE* (AG55) CELL LINE: A MODEL TO STUDY MOSQUITO PHYSIOLOGY AND PATHOGEN INTERACTIONS

by

FNU RUCHIR

(Under the Direction of MICHAEL J. ADANG)

ABSTRACT

Drosophila suzukii and *Anopheles gambiae* are of great concern to humans as they are important agricultural pests and vectors of human diseases, respectively.

D. suzukii is a devastating pest of soft and thin-skinned fruits. It causes substantial economic losses to fruit production industries, forcing farmers to apply broad-spectrum insecticides frequently which could lead to the development of insecticide resistance, thus jeopardizing their efficacy. The LC_{50} values of zeta-cypermethrin, spinosad, and malathion insecticides were determined against a Pierce and a Clarke county *D. suzukii* population. The LC_{50} values were three fold higher in the Pierce population for all treatments. Furthermore, RNA sequencing approach was used to analyze the response of Pierce and Clarke *D. suzukii* population at the transcriptome level upon insecticide treatments. A higher number of genes involved in detoxification and reduced cuticular penetration, were differentially expressed in the Pierce population. Finally, fewer nonsynonymous single nucleotide variants with deleterious effect on protein function were predicted among

detoxification, and cuticular protein encoding genes in Pierce flies. Thus a combination of increased gene expression and fewer deleterious SNVs suggest molecular mechanisms underlying the higher LC_{50} values for Pierce population.

Anopheles gambiae is an important vector of human malaria and o'nyong-nyong fever. The complexity of handling *A. gambiae* and infectious pathogens has led to the use of *A. gambiae* cell lines, including Ag55 cells, as a potential model to study vector-pathogen interactions and immune responses. The utility of cell lines can be maximized if their detailed gene expression profile, and proteome are available. Omics approaches were applied to provide a detailed gene expression profile and proteome of Ag55 cells. The transcriptome of Ag55 cells was compared to that of blood fed female adults to establish real patterns of gene expression which indicate specific functions of Ag55 cells. Gene Ontology enrichment analysis of enriched transcript suggested that Ag55 cells have phagocytic properties, a hypothesis which was confirmed by the ability of Ag55 cells to phagocytize *E. coli* bioparticles. Transcriptomic data further backed by proteomic data suggest that Ag55 cells express hemocyte like properties.

INDEX WORDS: Ag55 cells; *D. suzukii*; RNA sequencing; Insecticides; LC-MS/MS; Proteomics; Phagocytosis; Hemocyte; Detoxification genes; Insecticide target sites; Cuticular protein genes; Single nucleotide variants (SNVs)

TRANSCRIPTOMIC ANALYSES OF *DROSOPHILA SUZUKII RESPONSES TO PESTICIDES* AND *ANOPHELES GAMBIAE* (AG55) CELL LINE: A MODEL TO STUDY MOSQUITO PHYSIOLOGY AND PATHOGEN INTERACTIONS

by

FNU RUCHIR

B.S., Himachal Pradesh Agriculture University, Palampur, India, 2006

M.S., University of Sussex, Brighton, UK, 2007

M.Phil., University of Sussex, Brighton, UK, 2011

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2016

© 2016

RUCHIR

All Rights Reserved

TRANSCRIPTOMIC ANALYSES OF *DROSOPHILA SUZUKII RESPONSES TO PESTICIDES* AND *ANOPHELES GAMBIAE* (AG55) CELL LINE: A MODEL TO STUDY MOSQUITO PHYSIOLOGY AND PATHOGEN INTERACTIONS

by

FNU RUCHIR

Major Professor: Michael J. Adang
Committee: Lance Wells
Donald E. Champagne
Gang Hua

Electronic Version Approved:
Suzanne Barbour
Dean of the Graduate School
The University of Georgia
December 2016

DEDICATION

I dedicate this dissertation to my beloved daughter, my wife, my parents and family members for their love, support and encouragement. I also dedicate my thesis to my advisor Dr. Michael J Adang for his support and encouragement.

ACKNOWLEDGEMENTS

I sincerely express my immense gratitude to my supervisor, Dr. Michael Adang, for his kind and patient guidance, constant encouragement through the period of my PhD. Mike is not only an outstanding scientist but a great teacher and a wonderful person. I am indebted to him for all that I have learnt and continue to learn from him and I express my profound respect to him.

My immense gratitude to my committee members Dr. Donald Champagne, Dr. Lance Wells and Dr. Gang Hua for their valuable suggestions, guidance and encouragement throughout my PhD.

Next, I am grateful for the teaching assistantship given to me by the UGA Entomology department. I would like to thank our graduate coordinator and office staff who helped me with paper works. I would also like to thank my friends with whom I had great times during my dissertation and in India.

Finally, I want to express my deep appreciation to my parents, wife and family members in supporting me to come to the U.S. for my dissertation studies and for their love and sacrifice.

TABLE OF CONTENTS

Page

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

The dipterans have greater economic impact on humans than any other insect order because they are the important agricultural pests and also transmit pathogens which cause human diseases. Out of the many dipteran species *Anopheles gambiae* and *Drosophila suzukii* are of great concern to humans. *A. gambiae* is the major source of malaria transmission to humans. According to World Health Organisation (WHO), 214 million cases of malaria were reported world-wide in 2015 out of which 88% of those cases were reported from the African region; consequently malaria was the fourth largest cause of infant mortality in sub-Saharan Africa (WHO, 2015). On the other hand *D. suzukii* is a major invasive pest of soft and thin-skinned fruits in the US. It was first spotted in California in 2008, since then it has spread throughout the U.S. (Hauser 2011a, Walsh et al. 2011b, Burrack et al. 2012b). Economic losses caused by *D. suzukii* in the U.S. have been estimated as $718 million annually (Bolda 2011, Goodhue et al. 2011). Preventive applications of broad-spectrum chemicals such as pyrethroids, spinosyns and organophosphates have been used as a primary strategy to manage *D. suzukii* (Beers et al. 2011a, Bruck et al. 2011b, Walsh et al. 2011b). Similarly, to control *A. gambiae* the larvicides methoprene and temephos, and the adulticides synthetic pyrethroids and organophosphates have been the pesticides of choice (Adasi and Hemingway 2008, Bai et al. 2010, Marina et al. 2014, Wanjala et al. 2015).

Several cases of insecticide resistance have been reported from *A. gambiae* field populations (Adasi and Hemingway 2008, Wanjala et al. 2015). A significant level of resistance to permethrin has also been reported in a field population of *D. suzukii* (Bolda 2011).

RNA-sequencing approach to analyze for global changes at gene expression levels upon insecticide treatment in *D. suzukii* could provide valuable insight for development of strategies to delay insecticide resistance. Similarly a detailed gene expression profile of Ag55 (*A. gambiae*) cells, an alternate model to whole *A. gambiae* mosquito, could provide valuable information about *A. gambiae* and microbe interaction studies.

Mosquito cell lines have been used as models to study mosquito immune responses against microbes (Dimopoulos et al. 2002, Meister et al. 2005, Lombardo et al. 2013). Mosquito cell lines have also been used as screens to detect the effectiveness and specificity of novel compounds (insecticides) (O'Neal et al. 2013). Ag55 cells, the cell line used in my research, were established from neonate first instar larvae of *A. gambiae* (Pudney et al. 1979), and have been used for studying the effect of RNAi silencing of carbonic anhydrase on mosquito midgut pH regulation (Smith and Linser 2009a), *Plasmodium* interaction (Wilkins and Billingsley 2010) and *Lysinibacillus sphaericus* Bin toxin mode of action (Hire et al. 2015) studies.

Comparing the gene expression profile of *D. suzukii* populations upon three different insecticide treatments will help to assess the development of insecticide resistance and lead to formulation of strategies to delay the development of resistance in devastating pests such as *D. suzukii* in a timely manner. Similarly providing the detailed

gene expression profile and proteome of *A. gambiae* Ag55 cells will help researchers use and engineer cultured Ag55 cells in an efficient way by developing strategies to make it more suitable for microbe interaction and insecticide screen studies.

1.1 *Importance of next generation RNA-sequencing in mosquito research*

Next generation RNA sequencing has emerged as a very popular method for comparative gene expression profiling. RNA-sequencing has also been used for detecting single nucleotide polymorphisms (SNPs) and Indels (insertions and deletions) that may affect gene function, alternate splice forms for individual genes, and novel genes missed in genome annotation (Severson and Behura 2012).

As RNA sequencing captures only fully spliced transcripts an informative *de novo* transcriptome assembly of RNA sequences can be achieved, even in the absence of an assembled genome. Already *de novo* assembled transcriptomes have provided valuable sequence information useful for powerful molecular evolutionary analyses and quantitative gene expression profiles in the absence of an assembled genome sequence (Rinker et al. 2016). *De novo* assembly of non-blood feeding mosquito *Toxorhynchites* has provided extensive evidence for its phylogenetic relationship relative to the two fully sequenced *A. aegypti* and *C. pipiens* mosquito species (Zhou et al. 2014). Furthermore using a *de novo* assembly approach Crawford et al. assembled the transcriptome of *A. funestus* a primary vector of human malaria parasites (Zhou et al. 2014). Comparative gene expression profiles of vectors of human diseases using RNA-seq may be useful in providing important insights into the shared characteristics of common biological processes, as well as the identification of species-specific transcripts that may be used as a target for the design and development of novel control strategies (Rinker et al. 2016).

Diapause, an important mechanism for overwintering in mosquitoes, helps mosquitoes endure seasonal climatic shifts. The ability to overwinter is an attribute that directly impacts the ability of mosquitoes to transmit diseases over broad ranges (Rinker et al. 2016). Using RNA-seq transcriptomic analyses of developing oocytes researchers explored the mechanisms of photoperiodic diapause in *Aedes albopictus* and found unique preparatory signatures of photoperiodic diapause, thus providing a plausible explanation of their global invasion (Poelchau et al. 2011, Poelchau et al. 2013). Furthermore using next generation transcriptomic analysis researchers provided a plausible insight into the molecular mechanisms responsible for phenotypic divergence between the two *Culex pipiens* morphologically indistinguishable taxonomic forms (Price and Fonseca 2015).

Transcriptomic analyses have provided valuable information about host seeking behavior of mosquitoes. Using RNA-seq analysis researchers identified odorant receptor *AaegOr4* gene upregulated in F2 anthropophilic hybrid when compared to Zoophilic hybrid. The hybrids were generated by crossing anthropophilic *Aedes aegypti aegypti* with *Aedes aegypti formosus*. Further experiments suggested that Or4 odorant receptor might be linked to the anthropophily exhibited by *A. aegypti aegypti* subspecies (McBride et al. 2014). In another study, comparative antennal transcriptome profile of anthropophilic *A. gambiae* and zoophilic *Anopheles quadriannulatus* showed significant differences in the expression levels of odorant receptor genes. These enriched genes are predicted to work together for enhancing responsiveness to human-associated odors (Rinker et al. 2016).

RNA-seq analyses have also provided valuable insight into *A. gambiae* response to *Plasmodium*. Recent studies in *A. gambiae* have demonstrated that much of the known transcriptional response to *Plasmodium* infection derives from unannotated genomic regions, leading to a hypothesis that the mosquito may be employing non-coding RNAs as part of its basic response mechanism (Biryukova et al. 2014, Padron et al. 2014, Rinker et al. 2016). Comparing the gene expression profile of non-blood fed and five hours post blood-fed *A. aegypti* has provided valuable information about classes of potentially co-regulated genes, and a description of biochemical and physiological events that take place immediately after blood feeding (Bonizzoni et al. 2011).

Next generation RNA-seq has provided a powerful and efficient way to increase the phylogenetic resolution of the vast diversity of life on earth. Using sequencing data of nine *Anopheles* sp. and *Aedes aegypti* the researchers were able to retrieve the known phylogeny of these mosquitoes. Thus providing the evidence for the usefulness of RNA-seq data in resolving phylogeny (Hittinger et al. 2010, McCormack et al. 2013).

Furthermore using exome and transcriptome sequencing of *A. aegypti* Juneja et al. 2015 identified a locus that confers resistance to *Brugia malayi* and the resistance resulted in an increased immune response (Juneja et al. 2015). *Brugia malayi* is a mosquito-vectored parasite that causes lymphatic filariasis.

Transcriptomic analyses using high throughput sequencing have been used extensively to unravel possible mechanisms underlying insecticide resistance. Comparative gene expression profiles of deltamethrin-resistant and -susceptible *A. gambiae* mosquitoes from the Western Province of Kenya revealed that resistance to pyrethroids is a complex and evolving phenotype, dependent on multiple gene functions

which also include metabolic detoxification (Bonizzoni et al. 2015). Several other RNA-seq studies have revealed that insecticide resistance is more complex than previously anticipated, being associated to multiple genes rather than a single locus (Kalajdzic et al. 2012, David et al. 2014, Faucon et al. 2015, Liu 2015).

1.2 Drosophila suzukii (Spotted Wing *Drosophila*)

Drosophila suzukii belongs to the order Diptera and family Drosophilidae, and are closely related to vinegar flies. Unlike other vinegar flies *D. suzukii* causes economic loss due to the presence of a serrated ovipositor in adult females (Lee et al. 2011b).

D. suzukii adult females and males mate within 1-2 days of emergence. The females have high fecundity and can lay 300-600 eggs in their lifetime. Using the serrated ovipositor the female fly lay eggs just under the skin of the fruit. Furthermore *D. suzukii* has a short generation time and completes the whole developmental cycle in 8-10 days under optimal condition (i.e. 20˚C). It is a polyphagous insect pest and causes extensive damage to soft and thin-skinned fruits including blueberries, blackberries, raspberries, cherries, strawberries, peaches, and grapes worldwide (Dreves et al. 2009, Cini et al. 2012, Burrack et al. 2013).

1.3 Anopheles gambiae species complex

Species complexes are groups of species that are morphologically indistinguishable but differ in terms of behavior, ecology and geographic distribution (Bass et al. 2007, Lanzaro and Lee 2013). The *A. gambiae* complex has eight morphologically indistinguishable species. They are *A. gambiae* s.s., *A. arabiensis, A. bwambae, A. melas, A. merus, A. quadriannulatus, A. amharicus* and *A. comorensis* (Lanzaro and Lee 2013). The *A. gambiae* species complex was identified based on the

fact that F1 male sterility was observed sometimes between crosses of what were believed to be the same species. The chromosomal inversions were shown to differentiate the species (Gentile et al. 2001, Lanzaro and Lee 2013). *A. gambiae* s.s., *A. arabiensis* and *A. quadriannulatus* are fresh water species. In contrast *A. bwambae*, *A. melas* and *A. merus* are brackish water species (Lehmann and Diabate 2008). The *A. gambiae* complex species which prefers to feed on humans (anthropophilic) are *A. gambiae* s.s., *A. arabiensis*, *A. melas* and *A. merus* whereas *A. quadriannulatus* is zoophilic (Bass et al. 2007). The main malaria vector in the *A. gambiae* complex are *A. gambiae* s.s. and *A. arabiensis* (Bass et al. 2007), but *A. melas* and *A. merus* have also been shown to transmit malaria in coastal regions of Africa (Tsy et al. 2003, Moreno et al. 2004, Bass et al. 2007). *A. gambiae* s.s. (from now mentioned as *A. gambiae*) is undergoing speciation and this speciation event has been identified on the basis of chromosomal and molecular forms (Lehmann and Diabate 2008, Lanzaro and Lee 2013).

1.3.1 Chromosomal forms

The genome of *A. gambiae* is arranged on three chromosomes: two autosomes and one X/Y sex chromosomes. The autosomal chromosomes have submetacentric centromeres (Lanzaro and Lee 2013). Pombi et al identified 82 rare and 7 common paracentric inversions in field collected populations of *A. gambiae*. These inversions are mostly present on chromosome arm 2R (Pombi et al. 2008). Chromosomal forms of *A. gambiae* subpopulations have been distinguished based on the organization of six paracentric chromosomal inversions, five on the right arm of chromosome 2 (2Rj, b, c, d and u) and the other one on the left arm of the same chromosome (2La) (Lanzaro and Lee 2013). The name given to these chromosomal forms of *A. gambiae* are Mopti, Bamako,

Bissau, Forest and Savanna (Lanzaro and Lee 2013). The Savanna form spans throughout sub-Saharan Africa. The Mopti, Bamako and Bissau forms are present in West Africa. Bamako form prefers habitats along the Niger River and the Mopti form prefers drier habitats. In contrast the Forest form is prevalent in wetter habitats of Africa (Lanzaro and Lee 2013).

1.3.2 Molecular forms

Favia et al 2001 used a diagnostic polymerase chain reaction (PCR) assay to differentiate between the Mopti and the Savanna or Bamako chromosomal forms. They found that Mopti differed from the Savanna or Bamako forms by 10 nucleotides in a 2.3 kb fragment at the 5' end of the ribosomal DNA intergenic spacer region located on the X chromosome (Favia et al. 2001, Lanzaro and Lee 2013). The molecular forms M and S have been distinguished on the basis of SNPs present on the internal transcribed spacer (ITS) and intergenic spacer (IGS) regions of multicopy rDNA located on the X chromosome (della Torre et al. 2001, Gentile et al. 2001, Caputo et al. 2011). M and S molecular forms possess C/C and T/T genotypes, respectively (Lanzaro and Lee 2013). There are no associations between molecular forms and chromosomal forms of *A. gambiae* in parts of West Africa including western Senegal and Gambia (Lanzaro and Lee 2013). Recently the molecular form concept has overtaken the chromosomal form concept to distinguish the subpopulations of *A. gambiae*, which are to some degree reproductively isolated (Lanzaro and Lee 2013). There are two contradictory models which describe the relationship between M and S forms. These models are: I) "genomic islands of speciation" model and II) "incidental islands of divergence" model (Turner et al. 2005, Turner and Hahn 2010, Hahn et al. 2012, Lanzaro and Lee 2013). The "genomic

island of speciation" model suggests low level of divergence between M and S forms whereas "incidental islands of divergence" model treats M and S forms as distinct species (Lanzaro and Lee 2013). The S form is spread across sub-Saharan Africa and breeds mostly in alliance with rain-dependent water bodies and puddles caused by rain (Caputo et al. 2011). In contrast M form has border habitat for breeding purposes that exist throughout the year. Their breeding habitat is created by human activities like irrigation, rice cultivation and urbanization (della Torre et al. 2005, Costantini et al. 2009, Simard et al. 2009, Caputo et al. 2011). The availability of breeding habitat across seasons, has led M form to transmit malaria throughout the year (Caputo et al. 2011).

1.3.3 Implications of the differences in the members of species complex in terms of vector control

The members of *A. gambiae* species complex are morphologically identical but differ in breeding habitat, feeding and resting behavior (Mahande et al. 2007). They also differ significantly in their efficiency to transmit malaria (Bass et al. 2007).Therefore distinguishing the members of *A. gambiae* species complex is of great importance for vector control programs. Knowing accurately the species helps in the effective use of insecticides. For example *A. gambiae* s.s. an anthropophilic species if not distinguished correctly from a zoophilic *A. quadriannulatus* or non- vector species will lead to an ineffective use of insecticides and which may in turn lead to insecticide resistance development. Several cases of insecticide resistance have been reported in Africa against *A. gambiae* s.s. species (N'Guessan et al. 2003, Djouaka et al. 2008, Djogbenou et al. 2011).

1.4 Mosquito cell lines

Mosquitoes are the vectors of many filarial, protozoan, and viral pathogens and are known to transmit numerous infectious diseases to humans. Due to the rise in number of cases of mosquito borne diseases interest has been developed among the research community to explore several areas associated with pathogen-mosquito interactions. One of the main areas of the research has been the elucidation of mosquito immune response against pathogen (Walker et al. 2014). Susceptibility of mosquitoes to these pathogens is a very complex event and requires interplay of many genes and environmental conditions. The availability of well-established mosquito cell lines have become a useful tool for understanding the pathogen interaction at cellular and molecular level.

1.4.1 History of Mosquito cell lines

The first mosquito cell line was established from *Aedes aegypti* in 1965 (Grace 1966). Further in 1967 Singh established two cell lines from the larval tissues of *Aedes aegypti* and *Aedes albopictus* and were cultured in Mitsuhashi Maramorosch (MM) medium (Singh 1967). The Mitsuhashi Maramorosch (MM) medium was originally developed for culturing leafhopper cells (Mitsuhas and Maramoro 1964). Singh, further used commercially available fetal bovine serum (FBS) as a supplement instead of insect hemolymph, which was considered crucial for the growth and maintenance of mosquito cell cultures (Singh 1967, Sudeep et al. 2005). The *Aedes*, *Anopheles* and *Culex* species are the major mosquito disease vector species and several cell lines have been established from these species (Walker et al. 2014).

1.4.2 Aedes cell lines

Aedes species are competent vectors of many arboviruses including Dengue, Chikungunya, Eastern equine encephalitis, and Zika viruses (Mitchell et al. 1992, Vega-Rua et al. 2013). Recently Zika virus has grabbed lot of attention because zika virus infection during pregnancy can cause a serious birth defect called microcephaly (Mlakar et al. 2016). The first *A. albopictus* cell line, C6/36 cells (originally known as the ATC-15 cells), was established from neonate larvae (Singh 1967). Lineages of the C6/36 cell line have been extensively used to study the interaction between arboviruses and mosquito vectors. Clones derived from the ATC-15 cell line displayed heterogeneous sensitivity to virus infection (Igarashi 1978). Furthermore C6/36 (ATC-15) cells were shown to be sensitive to many arboviruses, partially due to the lack of a functional RNAi response (Brackney et al. 2010). The other cell line ATC-10 established from *A. aegypti* neonate larvae was less susceptible to arboviruses when compared to C6/36 (ATC-15) cells (Sudeep et al. 2005).

1.4.3 Anopheles cell lines

Species in the *A. gambiae* complex, and *A. stephensi* are the primary vectors of malaria in Sub-Saharan Africa and South Asia, respectively. Malaria has the highest mortality rate among all the mosquito transmitted diseases. The cell lines MOS.55, Ag55 (Marhoul and Pudney 1972, Pudney et al. 1979), Sua1B and Sua4a-3B (Muller et al. 1999) were established from neonate larvae. Sua5B cell line was derived by splitting Sua1 cell line (Rasgon et al. 2006). The three cell lines (Sua1B, Sua4a-3B, and Sua5B) are considered hemocyte-like because they have immune-responsive properties (Muller et al. 1999, Rasgon et al. 2006). Sua4a-3B is the first continuous cell line known to express

prophenoloxidases and is used as an *in vitro* model for the study of both the humoral and cellular immune defense of *An. gambiae* (Muller et al. 1999, Dong and Dimopoulos 2009). Ag55 and MOS.55 may be the same cell line (Marhoul and Pudney 1972, Pudney et al. 1979). Both cultured cell lines came from the laboratory of Mary Pudney at the London School of Tropical Medicine. In 1972 the researchers described the origin of the MOS.55 cell line (Marhoul and Pudney 1972). Further in 1979, the authors listed a number of cell lines from mosquitoes and the only *An. gambiae* line listed was called LSTM-AG-55 and the authors' state that all cell lines have been maintained in cultivation for several years and either stored in liquid nitrogen for up to 10 years (Pudney et al. 1979). The Mos.43 (Pudney and Varma 1971) and MSQ43 (Schneide.I 1969) cell lines were derived from *A. stephensi* larvae.

1.4.4 Culex cell lines

Culex species transmits Bancroftian lymphatic filariasis in tropical and sub-tropical regions, Japanese encephalitis virus (JEV) in Asia and the western Pacific, and West Nile virus (WNV) in the USA and Europe. The first cell line from *C. quinquefasciatus* was derived from ovaries (Hsu et al. 1970) and more recently a new cell line has been derived from *C. quinquefasciatus* embryos (Segura et al. 2012). Cell lines have also been established from the embryos of *C. tritaeniorhynchus* mosquitoes the primary vectors of JEV in Asia (Athawale et al. 2002, Kuwata et al. 2012). The NIID-CTR cell line established from *C. tritaeniorhynchus* is highly susceptible to both JEV and DENV (Dengue virus) infection, and provides a suitable model for virus replication in the host (Kuwata et al. 2012).

1.4.5 Applications of mosquito cell lines

In vitro cell line passage experiments can be used to deduce the host range of arbovirus in mosquitoes (Lawrie et al. 2004, Kuno 2007). In addition, cell lines can be used to study various aspects of virus-mosquito host interactions. For example C6/36 cells have been used to determine the cell entry mechanism. Acosta et al showed that DENV virus enters the C6/36 cell by low pH-dependent clathrin-mediated endocytosis (Acosta et al. 2011). Furthermore, persistently DENV infected C6/36 cells have been used to provide useful insight into the mechanisms through which arboviruses establish and maintain *in vivo* infections (Berenice Juarez-Martinez et al. 2013). Microarray analysis of DENV infected immune competent *Aedes aegypti* (Aag2) cell line provided valuable insight into the antiviral immune response of *A. aegypti* to DENV infection. The researchers showed that DENV is capable of actively suppressing the mosquito immune response in infected Aag2 cells (Sim and Dimopoulos 2010). Similarly the researches characterizing the antiviral properties of the piRNA pathway against mosquito borne Semliki Forest virus in *A. aegypti* Aag2 cells (Schnettler et al. 2013). Mosquito cell lines have been used largely to passage arboviruses for *in vivo* transmission assays that involve collection of mosquito saliva during expectoration (Walker et al. 2014).

Mosquito cell lines have played an important role in the isolation and characterization of Mosquito-only flaviviruses (MOFs) which are identified only in mosquitoes and have no known vertebrate reservoir host (Walker et al. 2014). Hobson-Peters et al isolated a new virus, tentatively named Palm Creek virus, from *Coquillettidia xanthogaster* mosquitoes in Australia using infection of cultured mosquito cells (Hobson-Peters et al. 2013). In another study researchers used C6/36 cells to isolate

Aedes flavivirus (AEFV) strain SPFLD-MO-2011-MP6 from a pool of male *A. albopictus* mosquitoes that were reared to adults from larvae collected in southwest Missouri, USA (Haddow et al. 2013).

Mosquito Mos20 and 4a-3A cell lines from *A. aegypti* and *A. gambiae*, respectively have been used for *in vitro* development of ookinetes into early oocysts (Siden-Kiamos et al. 2000, Al-Olayan et al. 2002). The *A. gambiae* cell line Ag55, a focus of my dissertation research, showed weak binding to the *P. berghei* ookinetes (Wilkins and Billingsley 2010) and the binding was not affected by exogenous carbohydrates suggesting a non-carbohydrate mediated binding (Wilkins and Billingsley 2010).

Mosquito cell lines have also been used to evaluate the potential effectiveness of novel insecticides and biopesticides. Cell-based screen platforms, established to determine the specificity of the novel insecticides (O'Neal et al. 2013), have significant advantages over whole organism screens. One of the major advantages of cell-based screens over whole organism screens is that large number of novel compounds (pesticides) can be tested for their effectiveness and specificity in very quick time and at fraction of the cost (O'Neal et al. 2013). Furthermore using *Aedes aegypti* cell line (CCL-125) researchers showed that Cry4B toxin binds to prohibitin a receptor protein found to mediate DENV entry into the same cell line (Kuadkitkan et al. 2010, Kuadkitkan et al. 2012). The researchers further showed that Cry4B pre-exposed *Aedes* cells had significantly lower DENV infected cells when compared to non-exposed *Aedes* cells (Kuadkitkan et al. 2012). Cry4B is a crystal toxin produced by gram-positive, *Bacillus thuringiensis israelensis*, bacteria. Hire et al. showed that Ag55 cells (*A. gambiae* cell

line) can be used as a model to decipher the mode of Bin toxin, produced by *Lysinibacillus sphaericus*, in mosquito larvae (Hire et al. 2015).

1.5 Diseases vectored by A. gambiae

Malaria is, currently the most important infectious parasitic disease transmitted by female *A. gambiae* mosquitoes, caused by an apicomplexan belonging to the genus *Plasmodium*. The only other disease vectored by *A.gambiae* is o'nyong'nyong fever caused by o'nyong'nyong virus. The o'nyong'nyong virus belongs to the genus alphavirus and is closely related to igbo-ora and chikungunya viruses (Posey et al. 2005). In addition to malaria and o'nyong'nyong fever *Anopheles* mosquitoes are also known to transmit *W. bancrofti* (filarial worm); the Timorese filarial (*Brugia timori*), eastern and western equine encephalitis, Venezualan equine encephalitis, and tataguine (http://www.malariasite.com/anopheles-mosquito/).

1.6 Malarial parasite life cycle in Anopheles species

The malarial parasite life cycle is highly complex and requires two hosts a mosquito vector and a vertebrate to complete its life cycle (Aly et al. 2009, Mathias et al. 2014).The asexual and sexual stages of reproduction takes place in vertebrates and mosquitoes, respectively (Sinden 2015).

The first step required in *Plasmodium* transmission is the up-take of a blood meal by female mosquito from an infected individual that contains gametocytes in peripheral circulation (Matuschewski 2006). Once in the mosquito the plasmodium gametocytes have to undergo three key transformation processes to be able to transform into infectious sporozoites. These three key steps are: I) gametogenesis, II) midgut traversal and III) salivary gland invasion (Angrisano et al. 2012).

1.6.1 Gametogenesis

In the mosquito midgut lumen the *Plasmodium* female and male gametocytes ingested during blood meal undergo maturation to form gametes after exposure to fall in temperature by 5˚C, pH change, and elevated levels of a mosquito waste product xanthurenic acid (XA) (Billker et al. 1998). XA directly or indirectly triggers a signal transduction cascade by activating guanylyl cyclase which produces cGMP that in turn is converted by phosphodiesterase to guanidine 5 monophosphate. The cGMP activates PKG (cGMP-dependent protein kinase) which is responsible for maintaining the elevated levels of cytosolic calcium (McRobert et al. 2008, Brochet et al. 2014, Sinden 2015). The elevated levels of calcium in cytoplasms of the gametocytes initiate gametogenesis and release from the red blood cells (Vlachou et al. 2006). The genes implicated in *Plasmodium* gametogenesis are male development gene-1 (MDV-1/PEG3), calcium dependent kinase 4 (CDPK4), mitogen-activated protein kinase 2 (MAPK-2) and cGMP-dependent protein kinase (PKG) (Billker et al. 2004, Rangarajan et al. 2005, Aly et al. 2009).

1.6.2 Fertilization of gametes and ookinete formation

The proteins P48/P45 and P230 on surface of male gametes and P230 on female gamete surface, belonging to 6-cysteine repeat protein family, have been shown to play an important role in fertilization (van Dijk et al. 2001, Eksi et al. 2006, Aly et al. 2009). Male and female gamete fusion leads to the formation of zygote in the mosquito midgut lumen and after that the zygote matures into a motile ookinete (Aly et al. 2009). A NIMA (never in mitosis/Aspergillus)-related protein kinase (Nek-4) has been shown to play a

crucial role in development of ookinetes from zygotes in *Plasmodium berghei* (Reininger et al. 2005).

1.6.3 Midgut epithelium recognition and midgut traversal of ookinete

The microneme, an apical ookinete organelle, possesses proteins involved in host cell recognition, attachment, motility, tissue traversal and invasion (Vlachou et al. 2006, Aly et al. 2009). For reaching to the apical surface of the mosquito midgut epithelium ookinete has to breach the peritrophic matrix and this is achieved in past by the parasitic enzyme chitinase (Aly et al. 2009). Chitinase seems to be crucial for the ookinete to traverse through the peritrophic matrix layer in *P.falciparum* and *P.gallinaceum,* but not in *P.berghei* (Vinetz et al. 1999, Vinetz et al. 2000, Dessens et al. 2001). The micronemal proteins CDPK3 (calcium-dependent protein kinase 3) and CTRP (circumsporozoite and TRAP related protein) are involved in motility and mosquito midgut invasion (Dessens et al. 1999, Siden-Kiamos et al. 2006).

For penetrating the midgut epithelium, ookinete proteins have to first interact with the proteins present on the apical end of the midgut epithelium. The proteins from ookinete implicated in mosquito midgut epithelial cell interactions are: enolase, WARP, MAOP, PPLP5, SUB2, CelTOS, SOAP, P28 and P25 (Aly et al. 2009, Vega-Rodriguez et al. 2014). The plasma membrane proteins of mosquito midgut epithelial cells typically have N-linked and O-linked glycans. Carbohydrates have been shown as key elements in the interaction of *Plasmodium falciparum* and *Plasmodium vivax* in *Anopheles tessellatus* (Ramasamy et al. 1997). Recent studies have suggested that ookinetes interact with multiple proteins present on the apical midgut plasma membrane during the invasion process (Parish et al. 2011, Vega-Rodriguez et al. 2014). Experimental evidences from

previous studies identified seven ookinete interacting proteins as AgAPN1 (Aminopeptidase N), Annexin-like (ANXB9, ANXB10B and ANXB10C), CpbAg1 (Carboxypeptidase B), SCRBQ2 (croquemort scavenger receptor homolog) and EBP (enolase-binding protein) (Parish et al. 2011, Vega-Rodriguez et al. 2014). Vega-Rodriguez et al showed that *Plasmodium* ookinete penetrates the mosquito midgut by at least two distinct mechanisms: one involves the interaction of enolase with EBP and the other is independent of the EBP-enolase interaction (Vega-Rodriguez et al. 2014). Out of the seven ookinete interacting proteins six were found to be localized in apical midgut microvilli detergent resistant membranes (DRM), which are enriched in lipid raft preparations. These results suggest that lipid rafts may be an important element for ookinete penetration of the midgut epithelium (Mathias et al. 2014). As six of the known ookinete interacting proteins were identified in DRM, Parish et al. 2011 hypothesized two models for interaction of ookinetes with lipid rafts. First model suggest that ookinete interacts with single or multiple lipid raft microdomains already at the site of penetration. In this model fusion of rafts may not be necessary for penetration of ookinete as the abundance of ookinete interacting proteins may be sufficient in each lipid raft microdomains to promote multivalent interactions. According to the second model ookinete causes multiple lipid raft microdomains to fuse to form a large platform containing ookinete interacting proteins at the site of penetration to facilitate interactions between ookinete and ookinete interacting proteins (Parish et al. 2011).

It is considered that P25 and P28, glycosylphosphatidylinositol (GPI)-anchored ookinete surface proteins, interaction with laminin of mosquito midgut basal lamina leads to the transformation of a moving ookinete into a sessile oocyst (Aly et al. 2009). The

other micronemal proteins which may be involved in ookinete to oocyst transformation are SOAP (secreted ookinete adhesive protein) (Dessens et al. 2003) and circumsporozoite and TRAP related protein (CTRP) (Mahairaki et al. 2005, Aly et al. 2009) but Nacer et al. 2008 showed that CTRP and SOAP are not required for ookinete to oocyst transformation in vitro (Nacer et al. 2008). A large number of ookinetes are lost due to mosquito immune response (Aly et al. 2009). The Pfs47 protein, a surface protein of *Plasmodium falciparum* ookinete, protects the ookinete from *Anopheles gambiae* immune responses and is essential for efficient human malaria transmission (Molina-Cruz et al. 2013).

The oocysts mature into sporozoites and this development takes place in the intercellular space between the midgut epithelium and the basal lamina (Vlachou et al. 2006). The LAPs (LCCL/lectin adhesive like proteins) are crucial for oocyst maturation and sporozoite formation (Raine et al. 2007). Wang et al showed that circumsporozoite protein (CSP) is also required for sporozoite development (Wang et al. 2005).

1.6.4 Salivary gland invasion

Sporozoites with the help of a putative enzyme ECP1 (egress cysteine protease 1) egress from the oocysts and are released into the hemocoel from there they are carried by hemolymph to salivary glands (Aly and Matuschewski 2005, Aly et al. 2009). Sporozoite ligands require specific protein receptors on the basal lamina of salivary gland to invade the salivary gland. The sporozoite proteins involved in salivary gland attachment and invasion are CSP (Barreau et al. 1995, Myung et al. 2004, Aly et al. 2009) and thrombospondin-related anonymous protein (TRAP) (Wengelnik et al. 1999, Aly et al. 2009). TRAP intraction with saglin, a salivary gland protein of *Anopheles*, has been

shown to be critical for salivary gland invasion by *Plasmodium* sporozoite (Ghosh et al. 2009). The other sporozoite proteins involved in salivary gland invasion are UOS3/TREP/S6, CRMPs (cysteine repeat modular proteins and MAEBL (Aly et al. 2009). The sporozoites present in the salivary glands of mosquitoes are motile and infectious to the vertebrate host (Aly et al. 2009).

1.7 Role of A. gambiae immune signaling pathway against microbes

The mosquito immune system depends on innate humoral and cellular reactions to fight malarial parasite infection and the parasite has to defend itself from the mosquito immune response before infecting humans (Lombardo et al. 2013). Hemocytes are the main mediators of insect cellular immune system (Pinto et al. 2009). One of the many functions of hemocytes is defense against invading pathogens (Pinto et al. 2009). In insects hemocytes have been identified and distinguished on the basis of their morphology, protein expression, and functional characteristics (Hwang et al. 2015). There are three major types of hemocytes (granulocytes, oenocytoids and prohemocytes) in *A. gambiae* which can be distinguished based on the presence or absence of certain proteins (Castillo et al. 2006). Hemocytes use phagocytosis, nodulation, and encapsulation processes as defense mechanisms against invading pathogens (Cirimotich et al. 2010). Humoral responses to microbes entail melanization and antimicrobial effector molecules. Antimicrobial effector molecules in turn are regulated by IMD, Toll, and JAK-STAT immune signaling pathways (Cirimotich et al. 2010).

In order for the immune system to respond, pathogen recognition has to occur first, where the cell relays this information through host signaling programs to stimulate effector responses (Xu and Cherry 2014). The activation of pattern recognition receptors

(PRRs) occurs in the innate immune system by sensing pathogen-associated molecular patterns (PAMPS), being able to differentiate between molecules present in the pathogen that are not usually found in the host (Kingsolver et al. 2013, Xu and Cherry 2014). Once the PAMPs are detected, signaling pathways are then activated resulting in the production of effector molecules to suppress pathogen replication (Kingsolver et al. 2013).

1.7.1 NF-κB transcription factor pathways

The Immune deficiency (IMD) and the Toll pathways are the two canonical NF-κB-dependent pattern recognition receptor (PRR) pathways involved in innate immunity in insects. Although both of these pathways have originally been categorized as antibacterial and antifungal immune signaling pathways, they have also been shown to play roles in the anti-*Plasmodium* and anti-viral immunity in mosquitoes (Kingsolver et al. 2013, Klowden 2013, Xu and Cherry 2014). According to genome sequencing analyses of *A. gambiae*, *A. aegypti, C. quinquefasciatus,* mosquito species share a conserved antimicrobial immune pathway with *Drosophila* (Kingsolver et al. 2013).

1.7.2 Immune deficiency (IMD) signaling pathway

The IMD pathway has been extensively studied in numerous insects including *Drosophila* and mosquitoes (Xu and Cherry 2014). The activation of IMD signaling is provoked by transmembrane receptors PGRP-LCs and intracellular recelptor PGRP-LE detecting diaminopimelic-containing peptidoglycan of Gram-negative bacteria and some Gram-positive *Bacilli* (Kingsolver et al. 2013). Following IMD signaling, the FAS-associated death domain protein (FADD) recruits a caspase, DREDD, leading to polyubiquitination of IMD (Kingsolver et al. 2013, Xu and Cherry 2014). TAK1 then binds to the polyubiquitin chain and activates the IKK complex, which in turn mediates

the phosphorylation of Rel2 (Relish in *Drosophila*) (Kingsolver et al. 2013, Xu and Cherry 2014). After phosphorylation followed by cleavage of Rel2 by DREDD, the N-terminal DNA-binding domain of Rel2 translocates to the nucleus and thus regulates transcription of effector genes (Kingsolver et al. 2013, Xu and Cherry 2014).

The IMD pathway has emerged as a useful pathway in terms of activity against *P. falciparum* (Garver et al. 2012). The over activation of IMD pathway by either silencing the gene encoding its negative regulator, Caspar, or over-expressing the gene encoding the REL2 transcription factor confers complete refractoriness to *P. falciparum* in laboratory reared *A. gambiae*, *A. stephensi*, and *A. albimanus* (Garver et al. 2009, Garver et al. 2012).

The IMD pathway regulates immune proteins, including APL1, TEP1, LRRD7 (APL2), FBN9, and LRIM1 and these proteins have been identified and studied with regard to their antiparasitic activity (Clayton et al. 2014).

1.7.3 Toll signaling pathway

Upon infection, the Toll pathway (Figure 3) is activated in the hemolymph by a serine protease cascade and leads to the processing of the Nerve Growth Factor-related cytokine Spätzle (Spz), the endogenous Toll ligand (Weber et al. 2003, Kingsolver et al. 2013, Klowden 2013, Xu and Cherry 2014). Following the Spätzle-Toll interaction, an intracellular signaling cascade is activated where the adaptor protein dMyD88 recruits the adaptor protein Tube through the death domain which then subsequently recruits the kinase Pelle, inducing phosphorylation and leading to the degradation of the Iκ-B-like protein Cactus and the nuclear translocation of the NF-κB-like transcription factor Rel1 which in turn up-regulates transcription of immune genes that are responsible for

microbial killing (Cirimotich et al. 2010, Clayton et al. 2014). RNAi knock-down of cactus, a negative regulator of Rel1, significantly decreases *P. berghei* burden, while knock-down of Rel1 increases infection levels in *A. gambiae* (Frolet et al. 2006).

1.7.4 JAK-STAT pathway

Although the JAK/STAT pathway was previously characterized for its role in development and hemocyte proliferation, it has been found to also contribute in antiviral, antiplasmodial and antibacterial responses in mosquitoes, by regulating the production of downstream effector molecules including antimicrobial peptides (AMPs) (Paradkar et al. 2012, Kingsolver et al. 2013, Xu and Cherry 2014). Initial recognition of infection by PRRs triggers the activation of the pathway through the binding of secreted ligands (Paradkar et al. 2012, Kingsolver et al. 2013, Xu and Cherry 2014).

The JAK/STAT pathway is initiated by the binding of unpaired (UPD) ligand to their receptor Domeless (Dome) (Kingsolver et al. 2013). After dimerization of the receptor, the associated Janus kinases Hopscotch (hop) transphosphorylate each other, recruit, and phosphorylate STAT, thus resulting in translocation of STAT to the nucleus and transcriptional activation of immune effector genes (Clayton et al. 2014). *A. gambiae* possesses two STAT genes (*STAT1*/*AgSTAT-B* and *STAT2*/*AgSTAT-A*).

The JAK/STAT pathway can be regulated by various methods such as dephosphorylation, nuclear export, or negative feedback from the inhibitor of hop, the suppressor of cytokine signaling 36E, as well as the negative feedback from the inhibitor of STAT, the *Drosophila* protein inhibitor of activated STAT (dPIAS) (Kingsolver et al. 2013).

Gupta et al. 2009 showed that silencing of *STAT-A* increases mature oocyst development in *P. berghei* and *P. falciparum*-infected *A. gambiae* mosquitoes (Gupta et al. 2009). The authors further showed that knock-down of SOCS reduces parasite infection levels.

1.8 Insecticide resistance in dipteran species

The dipterans have developed resistance to many insecticides in fields, as well as in laboratory upon selection. Insecticides, the key weapons for the control of insect pests, mostly target important receptors or enzymes in the insect nervous system whose poisoning result in paralysis and death (Ffrench-Constant et al. 2004a, Hemingway et al. 2004, Li et al. 2007b). The most important challenge to agriculture production and control of insect vectors of pathogens is the development of insecticide resistance, a widespread well-documented genetic phenomenon in insect pests. Understanding insecticide resistance at molecular level would aid in elucidating the mechanism of resistance as well as devising new methods for pest management. Major resistance mechanisms are categorised into two groups. 1) Metabolic, which involves detoxification of insecticides mainly involving enzymes cytochrome P450s, carboxylesterases and glutathione S-transferases. The over expression of these detoxifying enzymes cause the sequesteration or break down of the insecticides. The metabolic detoxifying enzymes carboxylesterase, GSTs and Cytochrome P450s are ubiquitously expressed in insects but the over expression of these enzymes especially in midgut, Malpighian tubules and fat body have been shown to be responsible for insecticide resistance. For example increased expression of CYP6G1 in midgut, fat body and malpighian tubules were responsible for DDT resistance in *D.melanogaster* (Hemingway et al. 2004, Chung et al. 2007, Dow 2009, Bass and Field 2011). 2)

Alterations in the target site of insecticides which encompasses mutations in the voltage-gated sodium channel, acetylcholinesterase and GABA receptor genes. The other two mechanisms implicated in insecticide resistance are: reduced cuticular penetration and insecticide avoidance (Mallet 1989, Hemingway 2000, Ffrench-Constant et al. 2004a, Hemingway et al. 2004, ffrench-Constant et al. 2006, Li et al. 2007b, Bass and Field 2011, Perry et al. 2011).

1.8.1 Metabolic resistance mechanism

The main enzyme families implicated in metabolic resistance include carboxylesterases, Glutathione S-Transferases (GSTs) and cytochrome P450s. Several studies (Hemingway et al. 2004, Bass and Field 2011) have demonstrated that amplifications, upregulations and coding mutations of the structural genes, encoding these enzymes have contributed to insecticide resistance. To demonstrate this several molecular techniques were used to compare sequences, copy numbers and expression levels of the genes encoding metabolic enzymes between susceptible and resistant strains. The substrate binding preferences of the metabolic enzymes were also compared between susceptible and resistant strains to determine the metabolic mediated resistance (Li et al. 2007b).

1.8.1.1 Carboxylesterases

Carboxylesterases belong to the hydrolase super family. The classical classification which is still in use divides the carboxylesterases into three types A, B and C on the basis of their interaction with organophosphates. Carboxylesterases belonging to type A hydrolyse organophosphates, carboxylesterases belonging to type B are irreversibly inhibited by these molecules, type C carboxylesterases neither hydrolyse nor are irreversibly inhibited by organophosphates (Oakeshott et al. 2005). In *Drosophila* and

mosquitoes carboxylesterases are grouped into α and β based on their preferential hydrolysis of the isomeric artificial substrate α- and β-naphthyl acetate (Oakeshott et al. 2005).

In *Drosophila* genome database there are 35 genes encoding carboxylesterase enzymes. Among them ten are the genes *(α-Est1-10)* of the *α-Esterase* cluster and are positioned at the right arm of the chromosome 3 (Tweedie et al. 2009, Birner-Gruenberger et al. 2012). The *α-Esterase* genes have diverse function. The *α-Est1, α-Est2* and *α-Est8* genes are expressed in *Drosophila* heads, and transcriptional up-regulation of *α-Est1* and *α-Est8 genes* in adult male's heads have been linked to aggressive behaviour (Dierick and Greenspan 2006, Birner-Gruenberger et al. 2012). The α-EST2 and α-EST7 proteins belong to fat body lipid droplet proteome, thus suggesting their role in lipid metabolism (Birner-Gruenberger et al. 2012). The orthologs of *α-EST7* gene have also been shown to be linked to insecticide resistance in some higher dipterans (Birner-Gruenberger et al. 2012).

Esterase 6, a protein which is encoded by the *β-esterase* cluster gene, has been shown to be expressed in the male ejaculatory duct and from there it is transferred as a component of the seminal fluid to the female during mating. This esterase 6 protein moves to the hemolymph of the female and stimulates egg laying (Birner-Gruenberger et al. 2012).

Juvenile hormone esterase is a carboxylesterase which is involved in the hydrolysis of the ester linkage of the juvenile hormone (Campbell et al. 2001). Juvenile hormone esterase plays an important role in metamorphosis by clearing the juvenile hormone from the hemolymph, thus preparing the last instar larva to undergo metamorphosis (Oakeshott et al. 2005).

The increased production of carboxylesterases due to gene amplification, upregulation and coding sequence mutation have been related to organophosphates, carbamates and pyrethroids resistance in variety of insects. The increased production of carboxylesterases cause sequestration or breakdown of insecticides for example increased amount of carboxylesterases in the midgut of mosquitoes (*Culex pipiens, Culex quinquefasciatus, Culex tarsalis* and *Culex tritaeniorhynchus*) (Newcomb et al. 1997, Cui et al. 2007, Li et al. 2007b, Bass and Field 2011).

1.8.1.1.1 Resistance via carboxylesterase gene amplification

The most common cause for increased carboxylesterase production is gene amplification. Resistance to organophosphates in *Culex pipiens* has been shown to be due to the co-amplification of two esterase genes *Est-3* encoding alpha esterase and *Est-2* encoding beta esterase. These genes were located at two different loci on chromosome II (Bass and Field 2011). The genes are amplified singly or are coamplified as allelic pairs at these loci in resistant mosquitoes. Resistant *Culex* has at least eight distinct amplicons (Hemingway 2000, Bass and Field 2011). Co-amplification of *estα2-estβ2* genes has been reported to be the most common reason for insecticide resistance in *Culex* population (Bass and Field 2011).

Hemingway and her colleagues showed that *estα2-estβ2* coamplicon of *Culex quinquefasciatus* resistant strain PelRR also possesses a full length gene encoding aldehyde oxidase (AO), immediately upstream (5') of *estα2*. Resistant strain has greater AO activity. AO in mammals is believed to have detoxification activity against several environmental pollutants thus AO might play an important role in metabolic detoxification of insecticides

in PelRR strain. The role of AO in insecticide resistance has not yet been established (Hemingway et al. 2000, Bass and Field 2011).

1.8.1.1.2 Resistance via carboxylesterase gene mutations

Mutations in carboxylesterase sequence have also been related to organophosphate resistance in *Culex pipiens* (mosquito) and *Lucilia cuprina* (blow fly) (Newcomb et al. 1997, Cui et al. 2007, Bass and Field 2011). In susceptible strain PelSS of *Culex quinquefasciatus, estα2* and *estβ2* genes were found in a head to head orientation. These two genes were separated by an intergenic spacer of 1.7 kb which possesses the core promoter sequences of both the genes. However in the resistant strain PelRR, due to the insertion of 3 indels (insertion and/or deletion), intergenic spacer increased to 2.7 kb. The insertion of indels occurred in resistant strain during the amplification event. These insertions possess a number of regulatory elements for example two *zeste* binding sites located within the first insertion of the resistant intergenic spacer. Intergenic spacer region of susceptible and resistance strain has been shown to act as a promoter for driving the expression of luciferase reporter gene (Hawkes and Hemingway 2002, Bass and Field 2011).

Cui and his colleagues mutated carboxylesterase B1 of *C. pipiens in vitro* at two positions namely 110 and 224 (Cui et al. 2007). Position 224 is highly conserved across all insect carboxylesterases. The mutation at 224th position from tryptophan to leucine converted carboxylesterase B1 to an organophosphate hydrolase and increased its malathion carboxylesterase activity. Thus they hypothesized that functional changes in carboxylesterases due to mutation may cause natural population of mosquitoes to acquire organophosphate resistance (Cui et al. 2007).

Carboxylesterase E3 enzymes encoded by *LcαE7* gene of susceptible and resistance strains differ from each other at 5 amino acid positions. They are Gly137Asp, Ala 267 Val, Met 283 Leu, Thr 335 His, Ile 35 Phe. In vitro studies have shown that Gly137Asp substitution is alone responsible for the conversion of carboxylesterase activity to an organophosphorous hydrolase activity and confers dianizon (organophosphate) resistance on *Lucilia cuprina* (blow fly). Substitution of Gly137Asp has been shown to be related to Organophosphate resistance across 15 isogenic strains of blow fly. Knowledge of the structure of related enzyme (acetylcholinesterase) suggests that $Gly^{137}$is present in the active site (Newcomb et al. 1997).

1.8.1.1.3 Resistance via carboxylesterase genes upregulation

Over production of esterase A1, due to upregulation, has been reported as a mechanism of resistance in *Culex pipiens* (Raymond et al. 1998). In *Culex quinquefasciatus* resistant strain, involving coamplification of *Estβ2 and Estα2* gene, upregulation due to expansion of their common promoter region was also responsible for the over production of *Estβ2 and Estα2* transcripts leading to resistance (Hawkes and Hemingway 2002, Li et al. 2007b).

1.8.1.2 Glutathione S-Transferases (GSTs)

The glutathione-s-transferases are a large group of multifunctional enzyme involved in the metabolism of wide range of xenobiotics including insecticides (Enayati et al. 2005).

GSTs are mainly cytosolic enzyme. They are also present in mitochondria and microsome. Microsomal GSTs are membrane associated proteins involved in eicosanoid and glutathione metabolism (Sheehan et al. 2001).

The GST catalyses the conjugation of reduced GSH and xenobiotics. It is a nucleophilic addition reaction. Conjugation makes the product more water soluble and therefore it can be easily excreted (Enayati et al. 2005).

Insect cytosolic GSTs are grouped into six classes: delta (δ), epsilon (ε), sigma (σ), theta (θ), omega (ω) and zeta (ζ). They are grouped into the above mentioned six classes on the basis of their amino acid sequence identity and phylogenetic relationship. Chromosomal location and immunological properties if known are also taken into consideration for classification (Chen et al. 2003, Ranson and Hemingway 2005) .

In insects, GST genes are clustered. In *Anopheles gambiae*, 28 genes encoding GST have been identified. Out of which 20 genes belong to delta and epsilon classes of GST. Genes belonging to delta class of GST are clustered on chromosome arm 2L and genes belonging to epsilon are clustered on chromosome arm 3R. In *Drosophila* there are 37 putative GST genes (Ding et al. 2003).

The δ, ε classes of GSTs play an important role in detoxification, thus δ, ε classes of GSTs are involved in insecticide resistance in insects. Zeta class GSTs have highly conserved structure suggesting an essential housekeeping role for the members of this family. In *A.gambiae* there is only one Zeta class GST i.e. GSTZ1-1 and this enzyme catalyses an important step in the tyrosine degradation pathway (Ranson and Hemingway 2005). GSTs protect insects against reactive oxygen species (ROS), including hydrogen peroxide, superoxide anions, and hydroxyl radicals that are generated during aerobic respiration. To terminate the sensory response in mosquitoes, odorant molecule must be degraded and GSTs play an important role in degradation process (Ranson and Hemingway 2005).

GSTs are metabolic detoxifying enzymes involved in insecticide resistance. GSTs have been shown to be involved in organophosphate, organochlorines and pyrethroids resistance in insects. GST mediated insecticide resistance has been shown to be due increased production of GSTs caused by amplification or upregulation of GST genes (Li et al. 2007b, Bass and Field 2011).

GSTs detoxify organophosphates via two distinct pathways a) *O*-dealkylation and b) *O*-dearylation. In *O*-dealkylation glutathione is conjugated with alkyl group of the insecticide whereas in *O*-dearylation glutathione reacts with the leaving group. GSTs do not metabolise pyrethroid molecule directly instead they detoxify lipid peroxide products formed by pyrethroids. Sequesteration of pyrethroids by GSTs has also been reported as a mechanism of resistance. Resistance to DDT in *Aedes aegypti* and *Anopheles gambiae* was due to the increased rate of dehydrochlorination of DDT by GST (Hemingway et al. 2004, Bass and Field 2011).

In most of the cases increased production of GSTs were suggested to be due to the upregulation of GST encoding genes but in few cases amplification of the structural GST genes were suggested for increased production of GST and in turn insecticide resistance (Bass and Field 2011).

1.8.1.2.1 Resistance via GST gene upregulation

In a DDT resistant strain of *Aedes aegypti , m*utation in the trans acting repressor element of Delta class GST gene was the suggested mechanism for overexpression of delta class GST(Hemingway et al. 2004). Mutations in the promoter elements of the Epsilon GST cluster was also the suggested mechanism of DDT resistance in *A.gambiae*

(Hemingway et al. 2004). Overexpression of *DmGSTD1* was the proposed mechanism of DDT resistance in the PSU-R strain of *Drosophila melanogaster (Tang and Tu 1994)*.

1.8.1.2.2 Resistance via GST gene amplification

In organophosphate resistant strain, Cornell –R, of *Musca domestica* four GST genes where overexpressed when compared with susceptible strain, but out of the 4 overexpressed genes *MdGSTD3* gene was the only gene linked to organophosphate resistance. The resistant strain possessed 12 *MdGSTD3*-like loci but the susceptible strain possessed 3 *MdGSTD3*-like loci suggesting the role of gene amplification in organophosphate resistance in Cornell-R strain of *M.domestica*((Syvanen et al. 1994, Li et al. 2007b)

1.8.1.3 Cytochrome P450s

The cytochrome P450 superfamily is one of the diverse groups of enzymes. Cytochrome P450s belong to the superfamily of proteins containing a heme cofactor and, therefore, are hemoproteins. Metabolic intermediates such as lipids and steroidal hormones, xenobiotics are the various substrates of cytochrome P450. The monooxygenase reaction is the most common reaction catalyzed by cytochromes P450 (Feyereisen 2005).

CYP is an abbreviation used, not only to designate the enzymes themselves but also the genes encoding CYP enzymes, which is usually followed by an Arabic numeral indicating the gene family. In addition, subfamily is indicated with the help of a capital letter followed by another numeral for the individual gene. It has been suggested by the current nomenclature guidelines that >40% amino acid identity is shared by the

members of new CYP families while members of subfamilies usually share >55% amino acid identity (Feyereisen 2005).

More than 660 CYP genes have been identified in insects and they are grouped into families and subfamilies. Mostly genes in the *CYP4, CYP6, CYP9* and *CYP12* groups have been implicated in insecticide resistance. Complete genome sequence of *Drosophila melanogaster* showed that 90 cytochrome P450 (CYP) genes are present, out of which 85 are potentially functional CYP genes and 5 are pseudo genes. Similarly in *Anopheles gambiae* 111 CYP genes are identified out of which 5 are thought to be pseudo genes. Complete genome sequence also showed that CYP genes were present in clusters. In *A. gambiae*, out of the 111 genes only 22 are present as singletons, rest of them are present in clusters. The largest cluster is of 14 genes which belong to the CYP6 family, and this cluster is present on the right arm of Chromosome 3. Sequential gene duplication is the reason behind the formation of gene clusters. Gene duplication is one of the main reasons, which is responsible for CYP gene diversity (Feyereisen 2005, Li et al. 2007b).

Cytochrome P450 enzymes are also involved in hydroxylation, N-, O- and S-dealkylation sulphoxidation, epoxidation, deamination, desulphuration, dehalogenation, peroxidation, and N-oxide reduction reactions (Bernhardt 2006). Insect cytochrome P450 metabolises both endogenous and exogenous substrates. The endogenous substrates include Juvenile hormone, ecdysteroids and pheromones. The exogenous substrate includes allelochemicals, insecticides and promutagens (Scott and Wen 2001).

Cytochrome P450s have been shown to be involved in conversion of ecdysone to 20 hydroxyecdysone and the conversion takes place in many peripheral tissues such as fat body, midgut and Malpighian tubules (Feyereisen 2005). Petryk and his colleagues showed

that CYP314A1, which is encoded by *Shade* gene, catalyses the NADPH- dependent hydroxylation of ecdysone to 20 hydroxyecdysone in *Drosophila* S2 cells (Petryk et al. 2003).

Insect P450s are present in microsomal membrane as well as mitochondria in all the cells of insect body (Feyereisen 2005). Mostly microsomal cytochrome P450s have been shown to be involved in xenobiotics (which include insecticides) metabolism for example *Drosophila*, microsomal cytochrome P450, CYP6A2 when expressed in vitro were able to efficiently metabolise diazonin and cyclodienes. Similarly CYP6D1 has been shown to catalyse the hydroxylation of pyrethroids (permethrin, deltamethrin and cypermethrin) (Feyereisen 2005, Li et al. 2007b). In some cases mitochondrial cytochrome P450s have also been shown to be involved in xenobiotics metabolism for example Guzov and his colleagues showed that insect CYP12 enzymes, which are mitochondrial enzymes, are capable of metabolising xenobiotics (Guzov et al. 1998, Feyereisen 2005). Insect midgut has been shown to be a rich source of P450 activity (Feyereisen 2005).

Cytochrome P450s are another important group of metabolic enzymes involved in insecticide resistance. Cytochrome P450s have been implicated in pyrethroid, organophosphate and organocholrine resistance in insects (Li et al. 2007b). Until recently overproduction of cytochrome P450s via upregulation have been implicated in cytochrome P450 mediated resistance however *CYP* gene amplification has now been shown to be related to insecticide resistance (Li et al. 2007b, Bass and Field 2011).

1.8.1.3.1 Resistance via upregulation of cytochrome P450 genes

Field based resistance to DDT were linked to over expression of *CYP6G1 gene* in *Drosophila melanogaster*. The overexpression of this gene was due to alterations in *cis*-

acting regulatory loci (Daborn et al. 2002b, Bass and Field 2011). Pyrethroid resistance in *Culex quinquefasciatus*, JPal-per strain, was related to overexpression of *CYP9M10 gene*. The overexpression of *CYP9M10* gene was caused by a *cis* acting mutation (Itokawa et al. 2010, Komagata et al. 2010).

Over expression of many *CYP* genes, in resistance strains, do not mean that all over expressed genes are related to resistance. One reason for this could be that if there is a common transcription factor which controls the expression of these genes and a mutation in the transcription factor could cause concomitant over expression of these genes. One of the prime example was pyrethroid resistance in *Musca domestica* LPR strain, in which over expression of *CYP6D1* gene was the major mechanism responsible for pyrethroid resistance but there were two other genes *CYP6A1* and *CYP6D3* which were also over expressed. The researchers showed that these three over expressed genes were regulated by a common trans- acting factor found on chromosome 2 (Carino et al. 1994, Liu and Scott 1996, Komagata et al. 2010).

1.8.1.3.2 Resistance via amplification of cytochrome P450 genes

The overexpression of *CYP9M10* gene in *Culex quinquefasciatus*, JPal-per strain, was mediated by gene duplication and *cis* acting mutation (Itokawa et al. 2010, Bass and Field 2011).

The overexpression of *CYP6P9* and *CYP6P4* genes were implicated in pyrethroid resistance in FUMOZ-R strain of *Anopheles funestus* and the overexpressions were due to the tandem duplications of these two genes (Wondji et al. 2009).

1.8.2 Resistance due to target site mutation

Another important mechanism involved in insecticide resistance is mutations in the target sites which render insensitivity to the target sites. The γ-aminobutyric acid (GABA), the voltage-gated sodium channel and acetylcholinesterase are the major targets of conventional insecticides and mutations in these three major target sites have been implicated in resistance to the conventional insecticides (ffrench-Constant et al. 1998b). To determine mutations in three major targets of insecticides, researchers cloned the genes of these targets (proteins) and to check the functional significance of a range of mutations in these genes, the genes were expressed in a range of heterologous expression systems for example *Xenopus* oocyte. Expressed mutant proteins were then directly tested for their insensitivity to a variety of different insecticides or insecticide inhibitors (ffrench-Constant et al. 1998b).

1.8.2.1 Acetylcholinesterase (AChE)

Acetylcholinesterase is a serine esterase that hydrolyses the neurotransmitter acetylcholine. It is GPI anchored to the surface of the postsynaptic membrane. It is found mainly at neuromuscular junctions and cholinergic brain synapses (Oakeshott et al. 2005, Voet and Voet 2010).

Structure of acetylcholinesterase is similar to that of carboxylesterase i.e . it also belongs to α/β hydrolase fold family. It has catalytic triad residues that are highly conserved in this family. A nucleophile (serine, cysteine or aspartic acid) is located after β5 strand, an acidic residue is almost always located after β7 strand and an absolutely conserved histidine residue is located after the last β strand. The active site of the acetylcholinesterase, present around the catalytic triad, has a conserved GXSXG pentapeptide termed the

nucleophilic elbow and 3 subsites namely leaving group pocket (P1 subsite), the acyl pocket (P2 subsite) and oxyanion hole (Nardini and Dijkstra 1999).

Acetylcholinesterase is the target site for organophosphates and carbamates. The insect acetylcholinesterase is encoded by *ace* gene. Mutations in the *ace* gene have been implicated in insecticide resistance. *Drosophila* and other higher Diptera appear to have one *ace* gene and all other insects studied have 2 *ace* genes (Oakeshott et al. 2005).

The *ace1* gene encodes acetylcholinesterase -1, which is responsible for the hydrolysis of neural acetylcholine. Mutations in the *ace1* gene have been linked to Organophosphate and carbamate resistance in insects. The *ace2* gene encodes acetylcholinesterase-2, and the function of acetylcholinesterase-2 is unknown in the insects which have two acetylcholinesterase enzymes. *Drosophila* acetylcholinesterase-2 is involved in various cell-cell communication (ffrench-Constant et al. 1998b, Oakeshott et al. 2005).

1.8.2.1.1 Mutations in AChE and insecticide resistance

Acetylcholinesterase insensitivity towards the target site of organophosphate and carbamate have been reported as a mechanism of resistance to these insecticides in *C. p. pipiens, C. p. quinquefasciatus, C. tritaeniorhynchus, A. nigerimus, A. atroparvus , A. sacharovi* and *A.gambiae* (Hemingway et al. 2004).

Mutations at only few particular positions in *ace1*gene have been related to acetylcholinesterase1 insensitivity. Mostly substitutions of the amino acid residues lining the active site cause the acetylcholinesterase insensitivity (Alout et al. 2007).

Alout and his colleagues showed that insecticide resistance in ACE-R strain of *C.pipiens* was associated with a single amino acid substitution, Phe290Val, in

acetylcholinesterase-1. In mosquitoes only three acetylcholinesterase-1 amino acid substitutions identified so far: G119S, F290V and F331W have been implicated in resistance and all these positions are lined near the active site of the enzyme (Alout et al. 2007, Alout et al. 2009).

In flies there is only one *ace gene* (*ace2*). Mutations at several positions in *ace2* gene led to the moderate acetylcholinesterase insensitivity (Alout et al. 2007).

1.8.2.2 GABA receptor

GABA receptors are ionotropic receptors for the γ-aminobutyric acid (GABA). They belong to the dicysteine loop (Cys-loop) super family. They are present in the nervous system of vertebrates and invertebrates and are mediators of rapid neurotransmission (Hosie et al. 1997, Buckingham et al. 2005).

GABA receptors are transmembrane proteins made up of five subunits. These subunits arrange in circle to form a channel. Each subunit possesses a long N-terminal extracellular domain. The long N-terminal domain possesses the cystein loop and 4 transmembrane regions (M1-M4). Many of the amino acid residues of the M2 transmembrane region line the integral chloride channel. GABA binding site is also present on the N-terminal extracellular domain (Hosie et al. 1997, Buckingham et al. 2005).

In *Drosophila* 3 GABA receptor subunits have been cloned and they are named as RDL, GRD and LCCH3 because they are encoded by three different genes: *Rdl (resistance to dieldrin), Grd (GABA and glycine like receptor of Drosophila) and Lcch3 (ligand-gated chloride channel homologue),* respectively (Hosie et al. 1997). The RDL subunit of GABA receptor, encode by *Rdl* gene, is the target site for cyclodienes and fipronils. Mutations in

RDL subunit has been related to cyclodiene and fipronil resistance in insects (ffrench-Constant et al. 1998b).

1.8.2.2.1 GABA receptor mutation and insecticide resistance

GABA receptors are the target sites of cyclodiene insecticides such as dieldrin. Insect resistance to cyclodienes are very common and the cause of this insensitivity is very well studied in *Drosophila melanogaster* (Hemingway et al. 2004).

Single amino acid substitution from alanine to serine at 302^{nd} position of RDL subunit has been documented in almost all the dieldrin resistant insect species. In rare cases, like in resistant strain of *D.simulans*, alanine to glycine substitution at 302^{nd} position of RDL subunit was observed (ffrench-Constant et al. 1998b, Hemingway et al. 2004).

Only the substitution of A302S/G in RDL subunit has been documented as the mechanism responsible for cyclodiene resistance in insects. The reason for the insensitivity is that A302S/G substitution stabilises the channel in an open confirmation, considerably decreasing the rate of desensitization. The predicted structure of RDL shows that A^{302} is present on M2 transmembrane and lines the chloride ion channel, suggesting its criticalness (Hosie et al. 1997, ffrench-Constant et al. 1998b).

1.8.2.3 Voltage-gated sodium channel

Voltage-gated sodium channels are transmembrane multimeric proteins that play an important role in the initiation and propagation of action potentials in neurons and other electrically excitable cells like myocytes and endocrine cells (Yu and Catterall 2003).

Voltage-gated sodium channels consist of a large α subunit and one or more auxillary β subunits. The α subunit consists of four domains (I-IV). These four domains are similar to each other and each of the domains possesses six α-helical transmembrane

segments (S1-S6). The S4 segment is the most conserved segment and contains voltage sensor and has positively charged amino acid residue at every third position (Zlotkin 1999, Yu and Catterall 2003).

Drosophila brain possesses two sodium channel genes. One gene (DSCI) is situated on the right arm of the second chromosome and is near to the seizure locus, whereas the second gene *para* is present on the paralysis (*para*) locus of the X chromosome (Zlotkin 1999). The mutations in *para* gene which encodes a voltage-gated sodium channel (PARA) have been implicated in insecticide resistance (ffrench-Constant et al. 1998b).

1.8.2.3.1 Mutations in voltage-gated sodium channel and insecticide resistance

Voltage-gated sodium channels are the primary target site of DDT and synthetic pyrethroids. oxadiazines, a new class of pyrazoline-like insecticides, also target sodium channels. Pyrethroids change the gating kinetics of voltage sensitive sodium channels by slowing both the activation and inactivation of the channels. Pyrethroids keep the channel open at the resting potential and cause the depolarisation of the neuronal membrane which initiates the repetitive discharges in motor and sensory axons. The repetitive discharges of motor and sensory axons lead to paralysis and death of insects (Hemingway et al. 2004, Dong 2007).

One of the major mechanisms of pyrethroid resistance is reduced target site insensitivity of sodium channels and it is referred to as knockdown resistance *(kdr)*. Single or multiple amino acid substitutions in the sodium channel, caused by the mutations in sodium channel gene *para*, reduces the binding affinity between the insecticide and sodium channel, thus resulting in knockdown resistance (*kdr*). Knockdown resistance (*kdr*) was first found in *Musca domestica* (Hemingway et al. 2004, Tan et al. 2012).

Comparisons of partial and complete sequences from 15 susceptible and kdr and kdr-like resistant housefly strains showed two substitutions (L1014F and/or M918T) were related with knockdown resistance. In many pyrethroid resistant insect species L1014F substitution has been reported (Martinez-Torres et al. 1999, Tan et al. 2012). In the African malarial vector *Anopheles gambiae* two different *kdr* substitutions, L1014F and L1014S were reported. Single nucleotide polymorphisms in the 6th segment of domain II of the para-type sodium channels were the cause for these two substitutions (Tan et al. 2012).

1.8.3 Other emerging molecular mechanism implicated in insecticide resistance

In recent times researchers have documented the role of transposable elements in insecticide resistance. Transposable elements cause resistance by inserting itself either upstream and/or downstream of the genes or within the genes, encoding metabolic detoxifying enzymes thereby causing overexpression of the metabolic detoxifying enzymes (Li et al. 2007b).

1.8.3.1 Role of transposable elements in insecticide resistance

Transposable elements are DNA sequences that can move and replicate within the genome. They are grouped into two classes based on their transposition strategies. Class 1 transposable elements use RNA as their intermediates for transposition and are called retrotransposons and class 2 transposable elements use DNA as their intermediates for transposition and are called DNA transposons (Feschotte 2008).

Many researchers have shown that transposable elements play an important role in genome evolution. In humans it has been shown that transposable elements contain substantial number of regulatory elements, and form the part of the coding sequences of some genes.

Transposable elements possess regulatory elements such as enhancers or insulators and as transposable elements get inserted into the coding regions of the genes they affect the genes expression or function. Transposable elements insertions cause numerous types of mutations (Chung et al. 2007).

Over expression of *CYP6G1* gene was shown to be associated with DDT resistance in *D.melanogaster*. Insertion of *Accord* (a long terminal repeat retrotransposon) upstream of *CYP6G1* gene was shown to be responsible for the over expression of this gene. The *Accord possesses* a regulatory sequence which increases the expression of *CYP6G1*gene in the tissues, midgut, malpighian tubule and fat body, responsible for detoxification. The *Accord* was mapped to the DDT resistant locus (DDT-R) which confers resistance to DDT and many newer classes of insecticides like neonicotinoids (Daborn et al. 2002b, ffrench-Constant et al. 2006).

Aminetzach and co-workers showed that insertion of *DOC1420* (a LINE like retrotransposon) at the 5' end of second exon of *CG10618* gene (which they named *CHKov1*) helps this gene to generate protein with a putative choline kinase function. Expressions of *CHKov1* paralogs were higher in adults and larvae as compared to pupae and embryos. From this the researchers assumed that *CHKov1* might be involved in digestion or detoxification. The researchers therefore hypothesized that *CHKov1* might be involved in choline metabolism and changes in *CHKov1* gene might affect choline metabolism and function of acetylcholinesterase, the target of organophosphates. Hence they suggested that *DOC1420* insertion into *CHKov1* might cause resistance to organophosphates (Aminetzach et al. 2005a, ffrench-Constant et al. 2006).

In *Musca domestica* mariner like elements (type of transposable element) were found in 5' and 3', 5' flanking sequences of pyrethroid resistant *CYP6D3* and *CYP6D1* paralogs, respectively (Kasai and Scott 2001b, Li et al. 2007b).

1.8.3.2 Cuticular penetration and insecticide resistance

Decreased rate of cuticular penetration has also been reported as the mechanism of resistance to insecticides. Lower nerve sensitivity and reduced cuticular penetration was shown to be the potential mechanism of resistance to pyrethroids in a (1R)-trans-permethrin selected strain of *Musca domestica.* In *Musca domestica* decreased cuticular penetration of insecticide has been reported to be controlled by *pen* gene present on chromosome 3(Devries and Georghiou 1981).

Pan and his colleagues showed that *Laccase2* gene was significantly over expressed in fenvalrate resistant strain of *Culex pipiens pallens. Laccase2* gene encodes Laccase enzyme which has p-diphenol oxidase activity. Laccases have been shown to be involved in cuticular sclerotization in mosquitoes. Thus the researchers suggested that overexpression of *Laccase2* gene reduced the rate of cuticular penetration of fenvalrate in *Culex pipien pallens* by increasing the sclerotization of cuticle (Pan et al. 2009).

1.9 Rationale and goals of the research

The dipterans are of great concern to humans. They not only vector human diseases but also damage important cash crops. Out of several dipterans *D. suzukii* and *A. gambiae* are of great concer. *D. suzukii* damages soft and thin skinned fruits there by causing significant economic loss to the fruit growing farmers. Similarly, *A. gambiae* transmits malaria to humans and malaria is the fourth largest cause of infant mortality in sub- Saharan Africa (WHO, 2015). Mosquito cell lines have helped in the understanding

of mosquito immune response and interaction to microbes (Dimopoulos et al. 2002, Meister et al. 2005, Lombardo et al. 2013).

The overall goal of my research is to use next generation omics methods to provide insight into the global changes in the gene expression profile of *D. suzukii* populations upon three different classes of insecticide treatments and also provide insight into the gene expression profile and proteome of *A. gambiae* Ag55 cells. The transcriptomic and proteomic analyses was performed to enhance the utility of Ag55 cells for pathogen interaction, genetic and immune response studies as these analyses provides insight into the characteristics of the cells. My research will help researchers and farmers to assess the development of insecticide resistance and lead to formulation of strategies to delay the development of resistance in devastating pests such as *D. suzukii* in a timely manner. Furthermore, providing gene expression profile and proteome of Ag55 cells will help researchers use and engineer Ag55 cell line in an efficient way by developing strategies to make it more suitable for microbe interaction and insecticide screen studies.

Chapter 2 of my thesis presents results of the gene expression profile of *D. suzukii* Pierce and Clarke populations upon three different classes of insecticide treatments. Pairwise comparison was performed between two sets of samples: (1) Clarke population that was treated or untreated (control); and (2) Pierce population that was treated or untreated (control). Further single nucleotide variants (SNVs), and insertions and deletions (indels) in *D. suzukii* Pierce and Clarke populations were predicted to reveal genetic changes that may have occurred while the Pierce and Clarke populations of D. suzukii were subjected to different selective pressures (Pierce population was continuously under insecticide selection pressure). By correlating our analysis on

genomic sequence differences between the Pierce and Clarke *D. suzukii* populations, and their observed differential response to multiple classes of insecticides, our results have shed light on potential mechanisms of differences in LC_{50} values in *D. suzukii* Pierce and Clarke populations.

Chapter 3 presents results of gene expression profile and proteome of Ag55 cells, and pairwise comparison of Ag55 cells with 5 days old *A. gambiae* female adults 3 hrs post blood feeding. By comparing the gene expression profile of Ag55 cells with 5 days old *A. gambiae* female adults 3 hrs post blood feeding I found that phagocytosis biological processes gene ontology term was significantly enriched in Ag55 cells. Phagocytic characteristics of Ag55 cells were further, validated using confocal imaging. Our transcriptomic data backed by proteomics data and confocal microscopy have shed light on the characteristics and expression profile of Ag55 cells.

1.10 References

Acosta, E. G., V. Castilla, and E. B. Damonte. 2011. Infectious dengue-1 virus entry into mosquito C6/36 cells. Virus Res 160: 173-179.

Adasi, K., and J. Hemingway. 2008. Susceptibility to three pyrethroids and detection of knockdown resistance mutation in Ghanaian *Anopheles gambiae* sensu stricto. J Vector Ecol 33: 255-262.

Al-Olayan, E. M., A. L. Beetsma, G. A. Butcher, R. E. Sinden, and H. Hurd. 2002. Complete development of mosquito phases of the malaria parasite in vitro. Science 295: 677-679.

Alout, H., A. Berthomieu, A. Hadjivassilis, and M. Weill. 2007. A new amino-acid substitution in acetylcholinesterase 1 confers insecticide resistance to *Culex pipiens* mosquitoes from Cyprus. Insect Biochem Mol Biol 37: 41-47.

Alout, H., P. Labbe, A. Berthomieu, N. Pasteur, and M. Weill. 2009. Multiple duplications of the rare ace-1 mutation F290V in *Culex pipiens* natural populations. Insect Biochem Mol Biol 39: 884-891.

Aly, A. S. I., and K. Matuschewski. 2005. A malarial cysteine protease is necessary for *Plasmodium* sporozoite egress from oocysts. J Exp Med 202: 225-230.

Aly, A. S. I., A. M. Vaughan, and S. H. I. Kappe. 2009. Malaria Parasite Development in the Mosquito and Infection of the Mammalian Host. Annu Rev Microbiol 63: 195-221.

Aminetzach, Y. T., J. M. Macpherson, and D. A. Petrov. 2005. Pesticide resistance via transposition-mediated adaptive gene truncation in *Drosophila*. Science 309: 764-767.

Angrisano, F., Y.-H. Tan, A. Sturm, G. I. McFadden, and J. Baum. 2012. Malaria parasite colonisation of the mosquito midgut - Placing the *Plasmodium* ookinete centre stage. Int J Parasitol 42: 519-527.

Athawale, S. S., A. B. Sudeep, P. V. Barde, R. Jadi, U. Pant, A. C. Mishra, and D. T. Mourya. 2002. A new cell line from the embryonic tissues of *Culex tritaeniorhynchus* and its susceptibility to certain flaviviruses. Acta Virol 46: 237-240.

Bai, H., D. B. Gelman, and S. R. Palli. 2010. Mode of action of methoprene in affecting female reproduction in the African malaria mosquito, *Anopheles gambiae*. Pest Manag Sci 66: 936-943.

Barreau, C., M. Touray, P. F. Pimenta, L. H. Miller, and K. D. Vernick. 1995. *Plasmodium gallinaceum*: Sporozoite invasion of *Aedes aegypti* salivary glands is inhibited by anti-gland antibodies and by lectins. Exp Parasitol 81: 332-343.

Bass, C., and L. M. Field. 2011. Gene amplification and insecticide resistance. Pest Manag Sci 67: 886-890.

Bass, C., M. S. Williamson, C. S. Wilding, M. J. Donnelly, and L. M. Field. 2007. Identification of the main malaria vectors in the *Anopheles gambiae* species complex using a TaqMan real-time PCR assay. Malar J 6.

Beers, E. H., R. A. Van Steenwyk, P. W. Shearer, W. W. Coates, and J. A. Grant. 2011. Developing *Drosophila suzukii* management programs for sweet cherry in the western United States. Pest Manag Sci 67: 1386-1395.

Berenice Juarez-Martinez, A., T. Olivia Vega-Almeida, M. Salas-Benito, M. Garcia-Espitia, M. De Nova-Ocampo, R. Maria del Angel, and J. Santiago Salas-Benito. 2013. Detection and sequencing of defective viral genomes in C6/36 cells persistently infected with dengue virus 2. Arch Virol 158: 583-599.

Bernhardt, R. 2006. Cytochromes P450 as versatile biocatalysts. J Biotechnol 124: 128-145.

Billker, O., S. Dechamps, R. Tewari, G. Wenig, B. Franke-Fayard, and V. Brinkmann. 2004. Calcium and a calcium-dependent protein kinase regulate gamete formation and mosquito transmission in a malaria parasite. Cell 117: 503-514.

Billker, O., V. Lindo, M. Panico, A. E. Etienne, T. Paxton, A. Dell, M. Rogers, R. E. Sinden, and H. R. Morris. 1998. Identification of xanthurenic acid as the putative inducer of malaria development in the mosquito. Nature 392: 289-292.

Birner-Gruenberger, R., I. Bickmeyer, J. Lange, P. Hehlert, A. Hermetter, M. Kollroser, G. N. Rechberger, and R. P. Kuhnlein. 2012. Functional fat body proteomics and gene targeting reveal in vivo functions of *Drosophila melanogaster* alpha-Esterase-7. Insect Biochem Mol Biol 42: 220-229.

Biryukova, I., T. Ye, and E. Levashina. 2014. Transcriptome-wide analysis of microRNA expression in the malaria mosquito *Anopheles gambiae*. BMC Genomics 15.

Bolda, M. 2011. Suspected tolerance to pyganic (pyrethrin) found in spotted wing Drosophila. Posted on August 23.

Bonizzoni, M., W. A. Dunn, C. L. Campbell, K. E. Olson, M. T. Dimon, O. Marinotti, and A. A. James. 2011. RNA-seq analyses of blood-induced changes

in gene expression in the mosquito vector species, *Aedes aegypti*. BMC Genomics 12.

Bonizzoni, M., E. Ochomo, W. A. Dunn, M. Britton, Y. Afrane, G. Zhou, J. Hartsel, M.-C. Lee, J. Xu, A. Githeko, J. Fass, and G. Yan. 2015. RNA-seq analyses of changes in the *Anopheles gambiae* transcriptome associated with resistance to pyrethroids in Kenya: identification of candidate-resistance genes and candidate-resistance SNPs. Parasit Vectors 8.

Brackney, D. E., J. C. Scott, F. Sagawa, J. E. Woodward, N. A. Miller, F. D. Schilkey, J. Mudge, J. Wilusz, K. E. Olson, C. D. Blair, and G. D. Ebel. 2010. C6/36 *Aedes albopictus* cells have a dysfunctional antiviral RNA interference response. PLoS Negl Trop Dis 4: e856.

Brochet, M., M. O. Collins, T. K. Smith, E. Thompson, S. Sebastian, K. Volkmann, F. Schwach, L. Chappell, A. R. Gomes, M. Berriman, J. C. Rayner, D. A. Baker, J. Choudhary, and O. Billker. 2014. Phosphoinositide Metabolism Links cGMP- Dependent Protein Kinase G to Essential Ca (2+) Signals at Key Decision Points in the Life Cycle of Malaria Parasites. Plos Biol 12.

Bruck, D. J., M. Bolda, L. Tanigoshi, J. Klick, J. Kleiber, J. DeFrancesco, B. Gerdeman, and H. Spitler. 2011. Laboratory and field comparisons of insecticides to reduce infestation of *Drosophila suzukii* in berry crops. Pest Manag Sci 67: 1375-1385.

Buckingham, S. D., P. C. Biggin, B. M. Sattelle, L. A. Brown, and D. B. Sattelle. 2005. Insect GABA receptors: splicing, editing, and targeting by antiparasitics and insecticides. Mol Pharmacol 68: 942-951.

Burrack, H. J., G. E. Fernandez, T. Spivey, and D. A. Kraus. 2013. Variation in selection and utilization of host crops in the field and laboratory by *Drosophila suzukii* Matsumara (Diptera: Drosophilidae), an invasive frugivore. Pest Manag Sci 69: 1173-1180.

Burrack, H. J., J. P. Smith, D. G. Pfeiffer, G. Koeher, and J. Laforest. 2012. Using volunteer-based networks to track *Drosophila suzukii* (Diptera: Drosophilidae) an invasive pest of fruit crops. J Integr Pest Manag 3: B1-B5.

Campbell, P. M., R. L. Harcourt, E. J. Crone, C. Claudianos, B. D. Hammock, R. J. Russell, and J. G. Oakeshott. 2001. Identification of a juvenile hormone esterase gene by matching its peptide mass fingerprint with a sequence from the Drosophila genome project. Insect Biochem Mol Biol 31: 513-520.

Caputo, B., F. Santolamazza, J. L. Vicente, D. C. Nwakanma, M. Jawara, K. Palsson, T. Jaenson, B. J. White, E. Mancini, V. Petrarca, D. J. Conway, N. J.

Besansky, J. Pinto, and A. della Torre. 2011. The "Far-West" of *Anopheles gambiae* Molecular Forms. Plos One 6.

Carino, F. A., J. F. Koener, F. W. Plapp, Jr., and R. Feyereisen. 1994. Constitutive overexpression of the cytochrome P450 gene CYP6A1 in a house fly strain with metabolic resistance to insecticides. Insect Biochem Mol Biol 24: 411-418.

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

Chen, L., P. R. Hall, X. E. Zhou, H. Ranson, J. Hemingway, and E. J. Meehan. 2003. Structure of an insect delta-class glutathione S-transferase from a DDT-resistant strain of the malaria vector Anopheles gambiae. Acta crystallogr. Sect D, Biol crystallogr 59: 2211-2217.

Chung, H., M. R. Bogwitz, C. McCart, A. Andrianopoulos, R. H. Ffrench-Constant, P. Batterham, and P. J. Daborn. 2007. Cis-regulatory elements in the Accord retrotransposon result in tissue-specific expression of the *Drosophila melanogaster* insecticide resistance gene Cyp6g1. Genetics 175: 1071-1077.

Cini, A., C. Ioriatti, and G. Anfora. 2012. A review of the invasion of *Drosophila suzukii* in Europe and a draft research agenda for integrated pest management. Bull Insectol 65: 149-160.

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

Clayton, A. M., Y. Dong, and G. Dimopoulos. 2014. The Anopheles Innate Immune System in the Defense against Malaria Infection. J Innate Immun 6: 169-181.

Costantini, C., D. Ayala, W. M. Guelbeogo, M. Pombi, C. Y. Some, I. H. N. Bassole, K. Ose, J.-M. Fotsing, N. F. Sagnon, D. Fontenille, N. J. Besansky, and F. Simard. 2009. Living at the edge: biogeographic patterns of habitat segregation conform to speciation by niche expansion in *Anopheles gambiae*. BMC Ecol 9: 16-Article No.: 16.

Cui, F., H. Qu, J. Cong, X. L. Liu, and C. L. Qiao. 2007. Do mosquitoes acquire organophosphate resistance by functional changes in carboxylesterases? Faseb J 21: 3584-3591.

Daborn, P. J., J. L. Yen, M. R. Bogwitz, G. Le Goff, E. Feil, S. Jeffers, N. Tijet, T. Perry, D. Heckel, P. Batterham, R. Feyereisen, T. G. Wilson, and R. H. ffrench-Constant. 2002. A single p450 allele associated with insecticide resistance in *Drosophila*. Science 297: 2253-2256.

David, J.-P., F. Faucon, A. Chandor-Proust, R. Poupardin, M. A. Riaz, A. Bonin, V. Navratil, and S. Reynaud. 2014. Comparative analysis of response to selection with three insecticides in the dengue mosquito *Aedes aegypti* using mRNA sequencing. BMC Genomics 15.

della Torre, A., Z. J. Tu, and V. Petrarca. 2005. On the distribution and genetic differentiation of *Anopheles gambiae* s.s. molecular forms. Insect Biochem Mol Biol 35: 755-769.

della Torre, A., C. Fanello, M. Akogbeto, J. Dossou-yovo, G. Favia, V. Petrarca, and M. Coluzzi. 2001. Molecular evidence of incipient speciation within *Anopheles gambiae* s.s. in West Africa. Insect Mol Biol 10: 9-18.

Dessens, J. T., A. L. Beetsma, G. Dimopoulos, K. Wengelnik, A. Crisanti, F. C. Kafatos, and R. E. Sinden. 1999. CTRP is essential for mosquito infection by malaria ookinetes. Embo J 18: 6221-6227.

Dessens, J. T., J. Mendoza, C. Claudianos, J. M. Vinetz, E. Khater, S. Hassard, G. R. Ranawaka, and R. E. Sinden. 2001. Knockout of the rodent malaria parasite chitinase PbCHT1 reduces infectivity to mosquitoes. Infect Immun 69: 4041-4047.

Dessens, J. T., I. Siden-Kiamos, J. Mendoza, V. Mahairaki, E. Khater, D. Vlachou, X. J. Xu, F. C. Kafatos, C. Louis, G. Dimopoulos, and R. E. Sinden. 2003. SOAP, a novel malaria ookinete protein involved in mosquito midgut invasion and oocyst development. Mol Microbiol 49: 319-329.

Devries, D. H., and G. P. Georghiou. 1981. Decreased Nerve Sensitivity and Decreased Cuticular Penetration as Mechanisms of Resistance to Pyrethroids in a (1r)-Trans-Permethrin-Selected Strain of the Housefly. Pestic Biochem Phys 15: 234-241.

Dierick, H. A., and R. J. Greenspan. 2006. Molecular analysis of flies selected for aggressive behavior. Nat Genet 38: 1023-1031.

Dimopoulos, G., G. K. Christophides, S. Meister, J. Schultz, K. P. White, C. Barillas-Mury, and F. C. Kafatos. 2002. Genome expression analysis of *Anopheles gambiae*: Responses to injury, bacterial challenge, and malaria infection. Proc. Natl. Acad. Sci. U. S. A. 99: 8814-8819.

Ding, Y., F. Ortelli, L. C. Rossiter, J. Hemingway, and H. Ranson. 2003. The *Anopheles gambiae* glutathione transferase supergene family: annotation, phylogeny and expression profiles. BMC Genomics 4: 35.

Djogbenou, L., N. Pasteur, M. Akogbeto, M. Weill, and F. Chandre. 2011. Insecticide resistance in the *Anopheles gambiae* complex in Benin: a nationwide survey. Med Vet Entomol 25: 256-267.

Djouaka, R. F., A. A. Bakare, O. N. Coulibaly, M. C. Akogbeto, H. Ranson, J. Hemingway, and C. Strode. 2008. Expression of the cytochrome P450s, CYP6P3 and CYP6M2 are significantly elevated in multiple pyrethroid resistant populations of *Anopheles gambiae* s.s. from Southern Benin and Nigeria. BMC Genomics 9.

Dong, K. 2007. Insect sodium channels and insecticide resistance. Invert Neurosci : IN 7: 17-30.

Dong, Y., and G. Dimopoulos. 2009. Anopheles Fibrinogen-related Proteins Provide Expanded Pattern Recognition Capacity against Bacteria and Malaria Parasites. J Biol Chem 284: 9835-9844.

Dow, J. A. T. 2009. Insights into the Malpighian tubule from functional genomics. J Exp Biol 212: 435-445.

Dreves, A. J., V. M. Walton, and G. C. Fisher. 2009. A new pest attacking healthy ripening fruit in Oregon: spotted wing Drosophila: *Drosophila suzukii* (Matsumura). Corvallis, Or.: Extension Service, Oregon State University.

Eksi, S., B. Czesny, G.-J. van Gemert, R. W. Sauerwein, W. Eling, and K. C. Williamson. 2006. Malaria transmission-blocking antigen, Pfs230, mediates human red blood cell binding to exflagellating male parasites and oocyst production. Mol Microbiol 61: 991-998.

Enayati, A. A., H. Ranson, and J. Hemingway. 2005. Insect glutathione transferases and insecticide resistance. Insect Mol Biol 14: 3-8.

Faucon, F., I. Dusfour, T. Gaude, V. Navratil, F. Boyer, F. Chandre, P. Sirisopa, K. Thanispong, W. Juntarajumnong, R. Poupardin, T. Chareonviriyaphap, R. Girod, V. Corbel, S. Reynaud, and J.-P. David. 2015. Identifying genomic changes associated with insecticide resistance in the dengue mosquito *Aedes aegypti* by deep targeted sequencing. Genome Res 25: 1347-1359.

Favia, G., A. Lanfrancotti, L. Spanos, I. Siden-Kiamos, and C. Louis. 2001. Molecular characterization of ribosomal DNA polymorphisms discriminating among chromosomal forms of *Anopheles gambiae* s.s. Insect Mol Biol 10: 19-23.

Feschotte, C. 2008. Transposable elements and the evolution of regulatory networks. Nat Rev Genet 9: 397-405.

Feyereisen, R. 2005. 4.1 - Insect Cytochrome P450, pp. 1-77. In I. G. Editors-in-Chief: Lawrence, I. Kostas and S. G. Sarjeet (eds.), Comprehensive Molecular Insect Science. Elsevier, Amsterdam.

ffrench-Constant, R., P. Daborn, and R. Feyereisen. 2006. Resistance and the jumping gene. BioEssays : news and reviews in molecular, cellular and developmental biology 28: 6-8.

Ffrench-Constant, R. H., P. J. Daborn, and G. Le Goff. 2004. The genetics and genomics of insecticide resistance. Trends Genet 20: 163-170.

ffrench-Constant, R. H., B. Pittendrigh, A. Vaughan, and N. Anthony. 1998. Why are there so few resistance-associated mutations in insecticide target genes? Philos. Trans R Soc Lond B Biol Sci 353: 1685-1693.

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

Garver, L. S., Y. Dong, and G. Dimopoulos. 2009. Caspar Controls Resistance to *Plasmodium falciparum* in Diverse *Anopheline* Species. Plos Pathog 5.

Garver, L. S., A. C. Bahia, S. Das, J. A. Souza-Neto, J. Shiao, Y. Dong, and G. Dimopoulos. 2012. *Anopheles* Imd Pathway Factors and Effectors in Infection Intensity-Dependent Anti-Plasmodium Action. Plos Pathog 8.

Gentile, G., M. Slotman, V. Ketmaier, J. R. Powell, and A. Caccone. 2001. Attempts to molecularly distinguish cryptic taxa in Anopheles gambiae s.s. Insect Mol Biol 10: 25-32.

Ghosh, A. K., M. Devenport, D. Jethwaney, D. E. Kalume, A. Pandey, V. E. Anderson, A. A. Sultan, N. Kumar, and M. Jacobs-Lorena. 2009. Malaria Parasite Invasion of the Mosquito Salivary Gland Requires Interaction between the Plasmodium TRAP and the Anopheles Saglin Proteins. Plos Pathog 5.

Goodhue, R. E., M. Bolda, D. Farnsworth, J. C. Williams, and F. G. Zalom. 2011. Spotted wing drosophila infestation of California strawberries and raspberries: economic analysis of potential revenue losses and control costs. Pest Manag Sci 67: 1396-1402.

Grace, T. D. C. 1966. Establishment of a line of mosquito (*Aedes aegypti* L.) cells grown in vitro. Nature 211: 366–367.

Gupta, L., A. Molina-Cruz, S. Kumar, J. Rodrigues, R. Dixit, R. E. Zamora, and C. Barillas-Mury. 2009. The STAT Pathway Mediates Late-Phase Immunity against Plasmodium in the Mosquito *Anopheles gambiae*. Cell Host Microbe 5: 498-507.

Guzov, V. M., G. C. Unnithan, A. A. Chernogolov, and R. Feyereisen. 1998. CYP12A1, a mitochondrial cytochrome P450 from the house fly. Arch Biochem Biophys 359: 231-240.

Haddow, A. D., H. Guzman, V. L. Popov, T. G. Wood, S. G. Widen, A. D. Haddow, R. B. Tesh, and S. C. Weaver. 2013. First isolation of *Aedes* flavivirus in the Western Hemisphere and evidence of vertical transmission in the mosquito *Aedes* (Stegomyia) *albopictus* (Diptera: Culicidae). Virology 440: 134-139.

Hahn, M. W., B. J. White, C. D. Muir, and N. J. Besansky. 2012. No evidence for biased co-transmission of speciation islands in *Anopheles gambiae*. Philos. Trans. R. Soc. B. Biol. Sci. 367: 374-384.

Hauser, M. 2011. A historic account of the invasion of *Drosophila suzukii* (Matsumura) (Diptera: Drosophilidae) in the continental United States, with remarks on their identification. Pest Manag. Sci. 67: 1352-1357.

Hawkes, N. J., and J. Hemingway. 2002. Analysis of the promoters for the beta-esterase genes associated with insecticide resistance in the mosquito *Culex quinquefasciatus*. BBA-Gene Struct. Expr. 1574: 51-62.

Hemingway, J. 2000. The molecular basis of two contrasting metabolic mechanisms of insecticide resistance. Insect Biochem Mol Biol 30: 1009-1015.

Hemingway, J., N. J. Hawkes, L. McCarroll, and H. Ranson. 2004. The molecular basis of insecticide resistance in mosquitoes. Insect Biochem Mol Biol 34: 653-665.

Hemingway, J., M. Coleman, M. Paton, L. McCarroll, A. Vaughan, and D. DeSilva. 2000. Aldehyde oxidase is coamplified with the World's most common Culex mosquito insecticide resistance-associated esterases. Insect Mol. Biol. 9: 93-99.

Hire, R. S., G. Huab, Q. Zhang, R. Mishra, and M. J. Adang. 2015. *Anopheles gambiae* Ag55 cell line as a model for *Lysinibacillus sphaericus* Bin toxin action. J Invertebr Pathol 132: 105-110.

Hittinger, C. T., M. Johnston, J. T. Tossberg, and A. Rokas. 2010. Leveraging skewed transcript abundance by RNA-Seq to increase the genomic depth of the tree of life. Proc. Natl. Acad. Sci. U. S. A. 107: 1476-1481.

Hobson-Peters, J., A. W. Y. Yam, J. W. F. Lu, Y. X. Setoh, F. J. May, N. Kurucz, S. Walsh, N. A. Prow, S. S. Davis, R. Weir, L. Melville, N. Hunt, R. I. Webb, B. J. Blitvich, P. Whelan, and R. A. Hall. 2013. A New Insect-Specific Flavivirus from Northern Australia Suppresses Replication of West Nile Virus and Murray Valley Encephalitis Virus in Co-infected Mosquito Cells. Plos One 8.

Hosie, A. M., K. Aronstein, D. B. Sattelle, and R. H. ffrench-Constant. 1997. Molecular biology of insect neuronal GABA receptors. Trends NeuroSci. 20: 578-583.

Hsu, S. H., W. H. Mao, and J. H. Cross. 1970. Establishment of a line of cells derived from ovarian tissue of *culex-quinquefasciatus* say. J Med. Entomol. 7: 703-707

Hwang, S., K. Bang, J. Lee, and S. Cho. 2015. Circulating hemocytes from larvae of the Japanese rhinoceros beetle *Allomyrina dichotoma* (Linnaeus) (Coleoptera: Scarabaeidae) and the ceelular immune response to microorganisms. Plos One 10.

Igarashi, A. 1978. Isolation of a singhs *Aedes-albopictus* cell clone sensitive to dengue and chikungunya viruses. J. Gen. Virol. 40: 531-544.

Itokawa, K., O. Komagata, S. Kasai, Y. Okamura, M. Masada, and T. Tomita. 2010. Genomic structures of Cyp9m10 in pyrethroid resistant and susceptible strains of *Culex quinquefasciatus*. Insect Biochem. Mol. Biol. 40: 631-640.

Juneja, P., C. V. Ariani, Y. S. Ho, J. Akorli, W. J. Palmer, A. Pain, and F. M. Jiggins. 2015. Exome and transcriptome sequencing of *Aedes aegypti* identifies a locus that confers resistance to *Brugia malayi* and alters the immune response. PLoS Pathog. 11.

Kalajdzic, P., S. Oehler, M. Reczko, N. Pavlidi, J. Vontas, A. G. Hatzigeorgiou, and C. Savakis. 2012. Use of mutagenesis, genetic mapping and next generation transcriptomics to investigate insecticide resistance mechanisms. PLoS One 7: e40296.

Kasai, S., and J. G. Scott. 2001. Cytochrome P450s CYP6D3 and CYP6D1 are part of a P450 gene cluster on autosome 1 in the house fly. Insect Mol. Biol. 10: 191-196.

Kingsolver, M. B., Z. Huang, and R. W. Hardy. 2013. Insect antiviral innate immunity: pathways, effectors, and connections. J. Mol. Biol. 425: 4921-4936.

Klowden, M. J. 2013. Physiological systems in insects / Marc J. Klowden, Amsterdam ; San Diego, CA : Academic Press, 2013.Third edition.

Komagata, O., S. Kasai, and T. Tomita. 2010. Overexpression of cytochrome P450 genes in pyrethroid-resistant *Culex quinquefasciatus*. Insect Biochem. Mol. Biol. 40: 146-152.

Kuadkitkan, A., D. R. Smith, and C. Berry. 2012. Investigation of the Cry4B-prohibitin interaction in *Aedes aegypti* cells. Curr. Microbiol. 65: 446-454.

Kuadkitkan, A., N. Wikan, C. Fongsaran, and D. R. Smith. 2010. Identification and characterization of prohibitin as a receptor protein mediating DENV-2 entry into insect cells. Virology 406: 149-161.

Kuno, G. 2007. Host range specificity of flaviviruses: Correlation with in vitro replication. J. Med. Entomol. 44: 93-101.

Kuwata, R., K. Hoshino, H. Isawa, Y. Tsuda, S. Tajima, T. Sasaki, T. Takasaki, M. Kobayashi, and K. Sawabe. 2012. Establishment and characterization of a cell line from the mosquito *Culex tritaeniorhynchus* (Diptera: Culicidae). In vitro cell Dev. Biol. Anim. 48: 369-376.

Lanzaro, G. C., and Y. Lee. 2013. Speciation in *Anopheles gambiae*—The distribution of genetic polymorphism and patterns of reproductive isolation among natural populations. *Anopheles Mosquitoes*: New Insights into Malaria Vector, ed Manguin S (InTech, Rijeka, Croatia).

Lawrie, C. H., N. Y. Uzcategui, M. Armesto, L. Bell-Sakyi, and E. A. Gould. 2004. Susceptibility of mosquito and tick cell lines to infection with various flaviviruses. Med. Vet. Entomol. 18: 268-274.

Lee, J. C., D. J. Bruck, H. Curry, D. Edwards, D. R. Haviland, R. A. Van Steenwyk, and B. M. Yorgey. 2011. The susceptibility of small fruits and cherries to the spotted-wing drosophila, *Drosophila suzukii*. Pest Manag. Sci. 67: 1358-1367.

Lehmann, T., and A. Diabate. 2008. The molecular forms of *Anopheles gambiae*: A phenotypic perspective. Infect. Genet. Evol. 8: 737-746.

Li, X., M. A. Schuler, and M. R. Berenbaum. 2007. Molecular mechanisms of metabolic resistance to synthetic and natural xenobiotics. Annu. Rev. Entomol. 52: 231-253.

Liu, N. 2015. Insecticide resistance in mosquitoes: impact, mechanisms, and research directions. Annu. Rev. Entomol. 60: 537-559.

Liu, N., and J. G. Scott. 1996. Genetic analysis of factors controlling high-level expression of cytochrome P450, CYP6D1, cytochrome b5, P450 reductase, and monooxygenase activities in LPR house flies, *Musca domestica*. Biochem. Genet. 34: 133-148.

Lombardo, F., Y. Ghani, F. C. Kafatos, and G. K. Christophides. 2013. Comprehensive genetic dissection of the hemocyte immune response in the malaria mosquito *Anopheles gambiae*. PLoS Pathog. 9.

Mahairaki, V., T. Voyatzi, I. Siden-Kiamos, and C. Louis. 2005. The *Anopheles gambiae* gamma1 laminin directly binds the *Plasmodium berghei* circurnsporozoite- and TRAP-related protein (CTRP). Mol. Biochem. Parasit**.** 140: 119-121.

Mahande, A., F. Mosha, J. Mahande, and E. Kweka. 2007. Feeding and resting behaviour of malaria vector, *Anopheles arabiensis* with reference to zooprophylaxis. Malar. J. 6.

Mallet, J. 1989. The evolution of insecticide resistance: Have the insects won? Trends Ecol. Evol. 4: 336-340.

Marhoul, Z., and M. Pudney. 1972. Mosquito cell line (mos 55) from *Anopheles-gambiae* larvae. Trans. R. Soc. Trop. Med. Hyg. 66: 183-184

Marina, C. F., J. G. Bond, J. Munoz, J. Valle, R. Novelo-Gutierrez, and T. Williams. 2014. Efficacy and non-target impact of spinosad, Bti and temephos larvicides for control of *Anopheles* spp. in an endemic malaria region of southern Mexico. Parasit Vectors 7: 55

Martinez-Torres, D., C. Chevillon, A. Brun-Barale, J. B. Berge, N. Pasteur, and D. Pauron. 1999. Voltage-dependent Na+ channels in pyrethroid-resistant Culex *pipiens* L mosquitoes. Pestic. Sci. 55: 1012-1020.

Mathias, D. K., J. G. Jardim, L. A. Parish, J. S. Armistead, H. V. Trinh, C. Kumpitak, J. Sattabongkot, and R. R. Dinglasan. 2014. Differential roles of an *Anopheline* midgut GPI-anchored protein in mediating *Plasmodium falciparum* and *Plasmodium vivax* ookinete invasion. **Infect. Genet. Evol.** 28: 635-647.

Matuschewski, K. 2006. Getting infectious: formation and maturation of *Plasmodium* sporozoites in the *Anopheles* vector. Cell. Microbiol. 8: 1547-1556.

McBride, C. S., F. Baier, A. B. Omondi, S. A. Spitzer, J. Lutomiah, R. Sang, R. Ignell, and L. B. Vosshall. 2014. Evolution of mosquito preference for humans linked to an odorant receptor. Nature 515: 222-U151.

McCormack, J. E., S. M. Hird, A. J. Zellmer, B. C. Carstens, and R. T. Brumfield. 2013. Applications of next-generation sequencing to phylogeography and phylogenetics. Mol. Phylogenet. Evol. 66: 526-538.

McRobert, L., C. J. Taylor, W. Deng, Q. L. Fivelman, R. M. Cummings, S. D. Polley, O. Billker, and D. A. Baker. 2008. Gametogenesis in malaria parasites is mediated by the cGMP-dependent protein kinase. PLoS Biol. 6: 1243-1252.

Meister, S., S. M. Kanzok, X. L. Zheng, C. Luna, T. R. Li, N. T. Hoa, J. R. Clayton, K. P. White, F. C. Kafatos, G. K. Christophides, and L. B. Zheng. 2005. Immune signaling pathways regulating bacterial and malaria parasite infection of the mosquito *Anopheles gambiae*. Proc. Natl. Acad. Sci. U. S. A.102: 11420-11425.

Mitchell, C. J., M. L. Niebylski, G. C. Smith, N. Karabatsos, D. Martin, J. P. Mutebi, G. B. Craig, and M. J. Mahler. 1992. Isolation of Eastern equine encephalitis-virus from *Aedes-albopictus* in Florida. Science 257: 526-527.

Mitsuhas, J., and K. Maramoro. 1964. Leafhopper tissue culture - embryonic nymphal and imaginal tissues from aseptic insects. Contributions from Boyce Thompson Institute 22: 435-460.

Mlakar, J., M. Korva, N. Tul, M. Popovic, M. Poljsak-Prijatelj, J. Mraz, M. Kolenc, K. R. Rus, T. V. Vipotnik, V. F. Vodusek, A. Vizjak, J. Pizem, M. Petrovec, and T. A. Zupanc. 2016. Zika Virus Associated with Microcephaly. New Engl. J. Med**.** 374: 951-958.

Molina-Cruz, A., L. S. Garver, A. Alabaster, L. Bangiolo, A. Haile, J. Winikor, C. Ortega, B. C. L. van Schaijk, R. W. Sauerwein, E. Taylor-Salmon, and C. Barillas-Mury. 2013. The human malaria parasite Pfs47 *gene* mediates evasion of the mosquito immune system. Science 340: 984-987.

Moreno, M., J. Cano, S. Nzambo, L. Bobuakasi, J. N. Buatiche, M. Ondo, F. Micha, and A. Benito. 2004. Malaria panel assay versus PCR: detection of naturally infected *Anopheles* melas in a coastal village of Equatorial Guinea. Malar. J. 3- 20

Muller, H. M., G. Dimopoulos, C. Blass, and F. C. Kafatos. 1999. A hemocyte-like cell line established from the malaria vector *Anopheles gambiae* expresses six prophenoloxidase genes. J Biol. Chem. 274: 11727-11735.

Myung, J. M., P. Marshall, and P. Sinnis. 2004. The *Plasmodium* circumsporozoite protein is involved in mosquito salivary gland invasion by sporozoites. Mol. Biochem. Parasit. 133: 53-59.

N'Guessan, R., F. Darriet, P. Guillet, P. Carnevale, M. Traore-Lamizana, V. Corbel, A. A. Koffi, and F. Chandre. 2003. Resistance to carbosulfan in *Anopheles gambiae* from Ivory Coast, based on reduced sensitivity of acetylcholinesterase. Med. Vet. Entomol. 17: 19-25.

Nacer, A., A. Underhill, and H. Hurd. 2008. The microneme proteins CTRP and SOAP are not essential for *Plasmodium berghei* ookinete to oocyst transformation in vitro in a cell free system. **Malar. J.** 7: 82.

Nardini, M., and B. W. Dijkstra. 1999. Alpha/beta hydrolase fold enzymes: the family keeps growing. Curr. Opin. Struct. Biol. 9: 732-737.

Newcomb, R. D., P. M. Campbell, D. L. Ollis, E. Cheah, R. J. Russell, and J. G. Oakeshott. 1997. A single amino acid substitution converts a carboxylesterase to an organophosphorus hydrolase and confers insecticide resistance on a blowfly. Proc. Natl. Acad. Sci. U.S.A 94: 7464-7468.

O'Neal, M. A., B. A. Posner, C. J. Coates, and J. M. Abrams. 2013. A cell-based screening platform identifies novel mosquitocidal toxins. J. Biomol. Screen 18: 688-694.

Oakeshott, J. G., C. Claudianos, P. M. Campbell, R. D. Newcomb, and R. J. Russell. 2005. 5.10 - Biochemical genetics and genomics of insect esterases, pp. 309-381. In I. G. Editors-in-Chief: Lawrence, I. Kostas and S. G. Sarjeet (eds.), Comprehensive Molecular Insect Science. Elsevier, Amsterdam.

Padron, A., A. Molina-Cruz, M. Quinones, J. M. C. Ribeiro, U. Ramphul, J. Rodrigues, K. Shen, A. Haile, J. L. Ramirez, and C. Barillas-Mury. 2014. In depth annotation of the *Anopheles gambiae* mosquito midgut transcriptome. BMC Genomics 15.

Pan, C., Y. Zhou, and J. Mo. 2009. The clone of laccase gene and its potential function in cuticular penetration resistance of *Culex pipiens* pallens to fenvalerate. Pestic. Biochem. Phys. 93: 105-111.

Paradkar, P. N., L. Trinidad, R. Voysey, J. B. Duchemin, and P. J. Walker. 2012. Secreted vago restricts West Nile virus infection in *Culex* mosquito cells by activating the Jak-STAT pathway. Proc. Natl. Acad. Sci. U.S.A 109: 18915-18920.

Parish, L. A., D. R. Colquhoun, C. U. Mohien, A. E. Lyashkov, D. R. Graham, and R. R. Dinglasan. 2011. Ookinete-Interacting Proteins on the microvillar surface are partitioned into detergent resistant membranes of *Anopheles gambiae* midguts. **J. Proteome Res.** 10: 5150-5162.

Perry, T., P. Batterham, and P. J. Daborn. 2011. The biology of insecticidal activity and resistance. Insect Biochem Mol Biol 41: 411-422.

Petryk, A., J. T. Warren, G. Marques, M. P. Jarcho, L. I. Gilbert, J. Kahler, J. P. Parvy, Y. Li, C. Dauphin-Villemant, and M. B. O'Connor. 2003. Shade is the *Drosophila* P450 enzyme that mediates the hydroxylation of ecdysone to the steroid insect molting hormone 20-hydroxyecdysone. Proc. Natl. Acad. Sci. U. S. A. 100: 13773-13778.

Pinto, S. B., F. Lombardo, A. C. Koutsos, R. M. Waterhouse, K. McKay, C. An, C. Ramakrishnan, F. C. Kafatos, and K. Michel. 2009. Discovery of *Plasmodium* modulators by genome-wide analysis of circulating hemocytes in *Anopheles gambia*e. Proc. Natl. Acad. Sci. U. S. A. 106: 21270-21275.

Poelchau, M. F., J. A. Reynolds, D. L. Denlinger, C. G. Elsik, and P. A. Armbruster. 2011. A de novo transcriptome of the Asian tiger mosquito, *Aedes albopictus*, to identify candidate transcripts for diapause preparation. BMC Genomics 12.

Poelchau, M. F., J. A. Reynolds, C. G. Elsik, D. L. Denlinger, and P. A. Armbruster. 2013. RNA-Seq reveals early distinctions and late convergence of gene expression between diapause and quiescence in the Asian tiger mosquito, *Aedes albopictus*. J Exp Biol 216: 4082-4090.

Pombi, M., B. Caputo, F. Simard, M. A. Di Deco, M. Coluzzi, A. della Torre, C. Costantini, N. J. Besansky, and V. Petrarca. 2008. Chromosomal plasticity and evolutionary potential in the malaria vector *Anopheles gambiae* sensu stricto: insights from three decades of rare paracentric inversions. BMC Evol. Biol. 8: 309-309.

Posey, D. L., T. O'Rourke, J. T. Roehrig, R. S. Lanciotti, M. Weinberg, and S. Maloney. 2005. Short report: O'Nyong-Nyong fever in West Africa. **BMC Evol. Biol.** 73: 32-32.

Price, D. C., and D. M. Fonseca. 2015. Genetic divergence between populations of feral and domestic forms of a mosquito disease vector assessed by transcriptomics. Peerj 3.

Pudney, M., and M. G. R. Varma. 1971. *Anopheles-stephensi* var mysorensis - establishment of a larval cell line (mos 43). Exp. Parasitol. 29: 7-12

Pudney, M., M. G. R. Varma, and C. J. Leake. 1979. Establishment of cell lines from larvae of culicine (*Aedes* species) and anopheline mosquitoes. Tca. Manual 5: 997-1002.

Raine, J. D., A. Ecker, J. Mendoza, R. Tewari, R. R. Stanway, and R. E. Sinden. 2007. Female inheritance of malarial lap genes is essential for mosquito transmission. Plos Pathog 3.

Ramasamy, M. S., R. Kulasekera, I. C. Wanniarachchi, K. A. Srikrishnaraj, and R. Ramasamy. 1997. Interactions of human malaria parasites, *Plasmodium vivax* and P-falciparum, with the midgut of *Anopheles* mosquitoes. Med. Vet. Entomol. 11: 290-296.

Rangarajan, R., A. K. Bei, D. Jethwaney, P. Maldonado, D. Dorin, A. A. Sultan, and C. Doerig. 2005. A mitogen-activated protein kinase regulates male

gametogenesis and transmission of the malaria parasite *Plasmodium berghei.* Embo Reports 6: 464-469.

Ranson, H., and J. Hemingway. 2005. Mosquito glutathione transferases. Methods Enzymol. 401: 226-241.

Rasgon, J. L., X. Ren, and M. Petridis. 2006. Can *Anopheles gambiae* be infected with *Wolbachia pipientis*? Insights from an in vitro system. Appl Environ Microbiol 72: 7718-7722.

Raymond, M., C. Chevillon, T. Guillemaud, T. Lenormand, and N. Pasteur. 1998. An overview of the evolution of overproduced esterases in the mosquito *Culex pipiens*. Philos Trans R Soc Lond B Biol Sci 353: 1707-1711.

Reininger, L., O. Billker, R. Tewari, A. Mukhopadhyay, C. Fennell, D. Dorin-Semblat, C. Doerig, D. Goldring, L. Harmse, L. Ranford-Cartwright, and J. Packer. 2005. A NIMA-related protein kinase is essential for completion of the sexual cycle of malaria parasites. J. Biol. Chem. 280: 31957-31964.

Rinker, D. C., R. J. Pitts, and L. J. Zwiebel. 2016. Disease vectors in the era of next generation sequencing. Genome Biol 17.

Schneide.I. 1969. Establishment of 3 diploid cell lines of *Anopheles stephensi* (Diptera - Culicidae). J. Cell Biol. 42: 603-606.

Schnettler, E., C. L. Donald, S. Human, M. Watson, R. W. C. Siu, M. McFarlane, J. K. Fazakerley, A. Kohl, and R. Fragkoudis. 2013. Knockdown of piRNA pathway proteins results in enhanced Semliki Forest virus production in mosquito cells. **J. Gen. Virol.** 94: 1680-1689.

Scott, J. G., and Z. Wen. 2001. Cytochromes P450 of insects: the tip of the iceberg. **Pest Manag. Sci.** 57: 958-967.

Segura, N. A., E. Santamaria, O. L. Cabrera, and F. Bello. 2012. Establishment and characterisation of a new cell line derived from *Culex quinquefasciatus* (Diptera: Culicidae). **Mem. Inst. Oswaldo Cruz.** 107: 89-95.

Severson, D. W., and S. K. Behura. 2012. Mosquito genomics: Progress and challenges. Annu Rev Entomol. 57: 143-166.

Sheehan, D., G. Meade, V. M. Foley, and C. A. Dowd. 2001. Structure, function and evolution of glutathione transferases: implications for classification of non-mammalian members of an ancient enzyme superfamily. **Biochem. J.** 360: 1-16.

Siden-Kiamos, I., A. Ecker, S. Nyback, C. Louis, R. E. Sinden, and O. Billker. 2006. *Plasmodium berghei* calcium-dependent protein kinase 3 is required for ookinete gliding motility and mosquito midgut invasion. **Mol. Microbiol.** 60: 1355-1363.

Siden-Kiamos, I., D. Vlachou, G. Margos, A. Beetsma, A. P. Waters, R. E. Sinden, and C. Louis. 2000. Distinct roles for Pbs21 and Pbs25 in the in vitro ookinete to oocyst transformation of *Plasmodium berghei*. **J. Cell Sci.** 113: 3419-3426.

Sim, S., and G. Dimopoulos. 2010. Dengue Virus Inhibits Immune Responses in *Aedes aegypt*i Cells. PLoS One 5.

Simard, F., D. Ayala, G. C. Kamdem, M. Pombi, J. Etouna, K. Ose, J.-M. Fotsing, D. Fontentille, N. J. Besansky, and C. Costantini. 2009. Ecological niche partitioning between *Anopheles gambiae* molecular forms in Cameroon: the ecological side of speciation. BMC Ecol. 9: 24.

Sinden, R. E. 2015. The cell biology of malaria infection of mosquito: advances and opportunities. Cell Microbiol. 17: 451-466.

Singh, K. R. P. 1967. Cell cultures derived from larvae of *Aedes albopictus* (skuse) and *Aedes aegypti* (l). Curr. Sci. 36: 506-508.

Smith, K. E., and P. J. Linser. 2009. Silencing of carbonic anhydrase in an *Anopheles gambiae* larval cell line, Ag55. J. RNAi Gene Silenc**.** 5: 345-350.

Sudeep, A. B., D. T. Mourya, and A. C. Mishra. 2005. Insect cell culture in research: Indian scenario. Indian J. Med. Res.121: 725-738.

Syvanen, M., Z. H. Zhou, and J. Y. Wang. 1994. Glutathione transferase gene family from the housefly *Musca domestica*. MGG 245: 25-31.

Tan, W. L., Z. M. Wang, C. X. Li, H. L. Chu, Y. Xu, Y. D. Dong, Z. C. Wang, D. Y. Chen, H. Liu, D. P. Liu, N. Liu, J. Sun, and T. Zhao. 2012. First report on co-occurrence knockdown resistance mutations and susceptibility to beta-cypermethrin in *Anopheles sinensis* from Jiangsu Province, China. PLoS One 7: e29242.

Tang, A. H., and C. P. Tu. 1994. Biochemical characterization of Drosophila glutathione S-transferases D1 and D21. J. Biol. Chem. 269: 27876-27884.

Tsy, J., J. B. Duchemin, L. Marrama, P. Rabarison, G. Le Goff, V. Rajaonarivelo, and V. Robert. 2003. Distribution of the species of the *Anopheles gambiae* complex and first evidence of *Anopheles merus* as a malaria vector in Madagascar. Malar J. 2.

Turner, T. L., and M. W. Hahn. 2010. Genomic islands of speciation or genomic islands and speciation? Mol Ecol. 19: 848-850.

Turner, T. L., M. W. Hahn, and S. V. Nuzhdin. 2005. Genomic islands of speciation in *Anopheles gambiae*. **PLoS Biol**. 3: 1572-1578.

Tweedie, S., M. Ashburner, K. Falls, P. Leyland, P. McQuilton, S. Marygold, G. Millburn, D. Osumi-Sutherland, A. Schroeder, R. Seal, and H. Zhang. 2009. FlyBase: enhancing *Drosophila* Gene Ontology annotations. Nucleic Acids Res. 37: D555-559.

van Dijk, M. R., C. J. Janse, J. Thompson, A. P. Waters, J. A. M. Braks, H. J. Dodemont, H. G. Stunnenberg, G. J. van Gemert, R. W. Sauerwein, and W. Eling. 2001. A central role for P48/45 in malaria parasite male gamete fertility. Cell 104: 153-164.

Vega-Rodriguez, J., A. K. Ghosh, S. M. Kanzok, R. R. Dinglasan, S. Wang, N. J. Bongio, D. E. Kalume, K. Miura, C. A. Long, A. Pandey, and M. Jacobs-Lorena. 2014. Multiple pathways for *Plasmodium* ookinete invasion of the mosquito midgut. Proc. Natl. Acad. Sci. U. S. A 111: E492-E500.

Vega-Rua, A., K. Zouache, V. Caro, L. Diancourt, P. Delaunay, M. Grandadam, and A.-B. Failloux. 2013. High efficiency of temperate *Aedes albopictus* to transmit Chikungunya and Dengue viruses in the Southeast of France. Plos One 8.

Vinetz, J. M., S. K. Dave, C. A. Specht, K. A. Brameld, B. Xu, R. Hayward, and D. A. Fidock. 1999. The chitinase PfCHT1 from the human malaria parasite Plasmodium falciparum lacks proenzyme and chitin-binding domains and displays unique substrate preferences. Proc. Natl. Acad. Sci. U. S. A 96: 14061-14066.

Vinetz, J. M., J. G. Valenzuela, C. A. Specht, L. Aravind, R. C. Langer, J. M. C. Ribeiro, and D. C. Kaslow. 2000. Chitinases of the avian malaria parasite *Plasmodium gallinaceum*, a class of enzymes necessary for parasite invasion of the mosquito midgut. J. Biol. Chem. 275: 10331-10341.

Vlachou, D., T. Schlegelmilch, E. Runn, A. Mendes, and F. C. Kafatos. 2006. The developmental migration of *Plasmodium* in mosquitoes. Curr. Opin. Genet. Dev. 16: 384-391.

Voet, D., and J. G. Voet. 2010. Biochemistry, John Wiley & Sons.

Walker, T., C. L. Jeffries, K. L. Mansfield, and N. Johnson. 2014. Mosquito cell lines: history, isolation, availability and application to assess the threat of arboviral transmission in the United Kingdom. **Parasit Vectors** 7.

Walsh, D. B., M. P. Bolda, R. E. Goodhue, A. J. Dreves, J. Lee, D. J. Bruck, V. M. Walton, S. D. O'Neal, and F. G. Zalom. 2011. *Drosophila suzukii* (Diptera: Drosophilidae): invasive pest of ripening soft fruit expanding its geographic range and damage potential. J. Integrated Pest Manag. 2: G1-G7.

Wang, Q., H. Fujioka, and V. Nussenzweig. 2005. Mutational analysis of the GPI-anchor addition sequence from the circumsporozoite protein of *Plasmodium*. Cell Microbiol. 7: 1616-1626.

Wanjala, C. L., J. P. Mbugi, E. Ototo, M. Gesuge, Y. A. Afrane, H. E. Atieli, G. Zhou, A. K. Githeko, and G. Yan. 2015. Pyrethroid and DDT resistance and organophosphate susceptibility among *Anopheles* spp. Mosquitoes, Western Kenya. Emerg. Infect. Diseases 21: 2178-2181.

Weber, A. N. R., S. Tauszig-Delamasure, J. A. Hoffmann, E. Lelièvre, H. Gascan, K. P. Ray, M. A. Morse, J.-L. Imler, and N. J. Gay. 2003. Binding of the *Drosophila* cytokine Spätzle to Toll is direct and establishes signaling. Nat Immunol 4: 794-800.

Wengelnik, K., R. Spaccapelo, S. Naitza, K. J. H. Robson, C. J. Janse, F. Bistoni, A. P. Waters, and A. Crisanti. 1999. The A-domain and the thrombospondin-related motif of *Plasmodium falciparum* TRAP are implicated in the invasion process of mosquito salivary glands. EMBO J. 18: 5195-5204.

Wilkins, S., and P. F. Billingsley. 2010. Mosquito cell line glycoproteins: an unsuitable model system for the *Plasmodium* ookinete-mosquito midgut interaction? Parasit Vectors 3.

Wondji, C. S., H. Irving, J. Morgan, N. F. Lobo, F. H. Collins, R. H. Hunt, M. Coetzee, J. Hemingway, and H. Ranson. 2009. Two duplicated P450 genes are associated with pyrethroid resistance in *Anopheles funestus*, a major malaria vector. Genome Res. 19: 452-459.

Xu, J., and S. Cherry. 2014. Viruses and antiviral immunity in *Drosophila*. Dev. Comp. Immunol. 42: 67-84.

Yu, F. H., and W. A. Catterall. 2003. Overview of the voltage-gated sodium channel family. Genome Biol. 4: 207.

Zhou, X., D. C. Rinker, R. J. Pitts, A. Rokas, and L. J. Zwiebel. 2014. Divergent and conserved elements comprise the chemoreceptive repertoire of the nonblood-feeding mosquito toxorhynchites amboinensis. **Genome Biol. Evol.** 6: 2883-2896.

Zhao, Y., W. Zhang, Y. Kho, and Y. Zhao. 2004. Proteomic analysis of integral plasma membrane proteins. Anal. Chem. 76: 1817-1823.

Zlotkin, E. 1999. The insect voltage-gated sodium channel as target of insecticides. Annu. Rev. Entomol. 44: 429-455.

CHAPTER 2

HIGH THROUGHPUT SEQUENCING REVEALS *DROSOPHILA SUZUKII* RESPONSES TO INSECTICIDES

Ruchir., Joana C. Chiu., Gang Hua., Nilesh R. Tawari., Michael J. Adang, and Ashfaq A. Sial

To be submitted to *Insect Molecular Biology*

Abstract

Global climate change and acquired resistance to insecticides are threats to world food security. *Drosophila suzukii,* a devastating invasive pest in many parts of the world, causes substantial economic losses to fruit production industries, forcing farmers to apply broad-spectrum insecticides frequently. This could lead to the development of insecticide resistance. We determined the LC_{50} (median lethal concentration) values of zeta-cypermethrin, spinosad, and malathion insecticides against *D. suzukii* colonies established from Clarke and Pierce county Georgia, U.S. The LC_{50} values were three-fold higher in the Pierce county population for all insecticide treatments. We then used RNA sequencing to analyze the responses of Pierce and Clarke population flies surviving an LC_{50} dose of the three insecticides. We identified a high number of differentially expressed genes (DEGs) that are likely involved in detoxification and reduced cuticular penetration, especially in the Pierce population, with a high degree of overlap in DEGs between the three insecticide treatments. Finally, we predicted fewer nonsynonymous single nucleotide variants with deleterious effect on protein function among detoxification, insecticide target, and cuticular protein encoding genes in Pierce flies. Thus a combination of increased gene expression and fewer deleterious SNVs suggest molecular mechanisms underlying the higher LC_{50} values for Pierce population.

Keywords: *Drosophila suzukii*, insecticide resistance, cross-resistance, high throughput sequencing, differentially expressed genes, single nucleotide variants, insertions and deletions

Abbreviations: **LCR**: Lethal Concentration Ratio, **LC_{50}**: Lethal concentration at which 50% test organisms are dead, ***Cyp***: Cytochrome P450, ***Est***: Esterases, ***GST***: Glutathione S-transferase, ***UGT***: UDP glucuronosyltransferase, ***GABA***: Gamma-aminobutyric acid, ***nAchR***: nicotinic acetylcholine receptors, ***Mdr***: Multidrug resistance, ***Cpr***: Cuticular protein, ***SCGs***: Sodium channel encoding genes, ***Cplc***: Cuticular protein of low complexity, **DDT**: dichloro-diphenyl-trichloroethane, **GO:** gene ontology, **DAVID**: **D**atabase for **A**nnotation, **V**isualization and **I**ntegrated **D**iscovery, **STRING**: Search Tool for the Retrieval of Interacting Genes/Proteins, **FDR:** False Discovery Rate.

2.1. Introduction

Global fruit production is being threatened by *Drosophila suzukii* Matsumura (Diptera: Drosophilidae), a recently invasive insect pest commonly known as spotted wing Drosophila (SWD) (Lee et al. 2011a). *D. suzukii* is indigenous to Eastern and South Eastern Asia and over the past five years it has dramatically expanded its range globally to include Europe, North America, and South America (Walsh et al. 2011a, Cini et al. 2012, Deprá et al. 2014). *D. suzukii* was first detected in California in 2008 (Hauser 2011b, Walsh et al. 2011a). Since then it has spread throughout the U.S. (Burrack et al. 2012a) causing significant losses in crop yield and quality. *D. suzukii* is a polyphagous insect pest, which causes significant damage to soft and thin-skinned fruits including blueberries, blackberries, raspberries, cherries, strawberries, peaches, and grapes worldwide.

Preventive applications of insecticides have been used as a primary strategy to manage *D. suzukii* (Beers et al. 2011b, Bruck et al. 2011a, Walsh et al. 2011a, Haviland and Beers 2012). The most effective insecticides used against *D. suzukii* are relatively broad-spectrum chemicals such as pyrethroids, spinosyns and organophosphates. These chemicals are neurotoxins and their primary targets are sodium channels, nicotinic acetylcholine receptors (nAChR), and acetylcholinesterase, respectively (ffrench-Constant et al. 1998a, Baxter et al. 2010a). A zero damage tolerance policy among growers for *D. suzukii* has led to the frequent application of broad-spectrum insecticides, which could lead to resistance development in *D. suzukii*. This is of particular concern in *D. suzukii* because it has a short generation time (Kopp and True 2002) comparable to the closely related *D. melanogaster*. A species which , has been shown to develop insecticide resistance at a much faster rate than anticipated by prior population models (Karasov et al. 2010).

Alarmingly, a significant level of resistance to permethrin in a field population of *D. suzukii* has been reported (Bolda 2011).

Four types of insecticide resistance mechanisms have been documented in insects, which include metabolic resistance, target-site/receptor mutations, reduced cuticular penetration, and avoidance behavior (Li et al. 2007a, Rivero et al. 2010, Silva et al. 2012). Of those, metabolic resistance and resistance due to target-site/receptor mutations have been the most extensively studied at both genetic and molecular levels (Hemingway and Ranson 2000). Metabolic resistance involves overproduction of detoxification enzymes that degrade or sequester insecticides (Hemingway et al. 1998, Vontas et al. 2001, ffrench-Constant et al. 2004b, Li et al. 2007a), which may be achieved by gene amplification (Karunaratne et al. 1998, Puinean et al. 2010), mutations in coding sequences (Claudianos et al. 1999, Zhu et al. 1999, Aminetzach et al. 2005b, Magwire et al. 2012), or mutations and insertions/deletions (Indels) in cis-acting promoter sequences (Daborn et al. 2002a, Seifert and Scott 2002, Guio et al. 2014) and/or trans-acting regulatory loci (Kasai and Scott 2001a, Maitra et al. 2002, Feyereisen et al. 2005, Mateo et al. 2014).

The effective management of insecticide resistance depends ultimately on a thorough understanding of resistance mechanisms at the genetic and molecular levels. Since the invasion of *D. suzukii* in the U.S. and the heavy use of insecticides to control it in fruit crops is relatively recent phenomenon, it is therefore timely to take a proactive approach to access the risk of resistance development in *D. suzukii*. Recently, transcriptomic analyses using high throughput sequencing and microarray have been used extensively to unravel possible mechanisms underlying insecticide resistance (Pedra et al. 2004, Puinean et al. 2010, Kalajdzic et al. 2012, Silva et al. 2012, David et al. 2014). Such

studies have revealed that insecticide resistance is more complex than previously anticipated, often being associated to multiple genes rather than a single locus.

In the current study, we used RNA sequencing-based transcriptomic analyses to determine the responses of two populations of *D. suzukii* that survived LC_{50}-level treatments of three broad-spectrum insecticides commonly used against *D. suzukii*. We further predicted single nucleotide variants (SNVs) and insertions and deletions (indels) in these *D. suzukii* populations to reveal possible reasons for differences in their respective LC_{50} values. By correlating our analysis on genomic sequence differences between the Pierce and Clarke county *D. suzukii* population, and their observed differential response to multiple classes of insecticides, our results have shed light on potential mechanisms underlying lower LC_{50} values of Clarke county population against three different classes of insecticides.

2.2. Materials and Methods

2.2.1. D. suzukii rearing and field collection

The Clarke population was established in July 2013 by collecting *D. suzukii* from unsprayed blueberries from Clarke County (Georgia, U.S., approximately 250 flies were collected to start the colony) and at the time of bioassays flies had been in the colony for approximately 12 generations. The flies were reared in plastic fly bottles using a cornmeal and molasses-based artificial diet [9]. The colony was maintained in upright growth chambers (model I-41 LLVLC8, Percival Scientific, Perry, IA) operating at 50% RH, 25° C ± 2° C with a 14:10 (L: D) photoperiod. A stock of minimum 2000 flies was maintained for each generation of the Clarke population. The *D. suzukii* Pierce population used in this study was collected in June 2014 from a commercial blueberry orchard in Pierce County

(Georgia, U.S.). Approximately 200 flies were collected and maintained in the laboratory for one generation using the same methods as for the Clarke colony. In order to collect the Pierce population, blueberries infested with *D. suzukii* larvae were collected into 52 oz ventilated plastic containers and placed into upright growth chambers. Adult flies were aspirated and transferred into plastic fly bottles with artificial diet and reared for one generation to obtain pupae. Pupae were then separated into individual 1.5 oz Solo® clear polystyrene cups (Solo® Cup Co., Urbana, IL) until adults emerged. Newly emerged adults were separated by sex and 5-7 d old females were used for the bioassays.

2.2.2. Insecticide bioassays

Standardized glass vial bioassay protocols were utilized to assess the susceptibility of adult *D. suzukii* females to zeta-cypermethrin, spinosad, and malathion. As these three insecticides have contact activity, a residual bioassay protocol was developed. Insecticide was diluted in acetone for zeta-cypermethrin and malathion, or deionized water for spinosad to prepare 20-100 mL of stock solution. A series of dilutions at desired concentrations were prepared for each of the selected insecticides. The bioassay chambers, 225 mL glass jars (cat. no. 02-911-460, Fisher Scientific, Pittsburgh, PA) and their lids were labeled with appropriate dilution using labeling tape. The glass jars used for spinosad bioassays contained approximately one-centimeter deep layer of fly diet at the bottom. One milliliter of appropriate insecticide dilution was added to each of the pre-labeled glass jars. The lids were put back on the corresponding glass jars. The glass jars were then swirled and inverted to insure that all surfaces inside the bottle were coated with insecticide residue. The glass jars were then opened and placed inside a fume hood to air dry for 30 mins to 1 hr depending on the selected insecticide. While glass jars were drying in the fume hood, 5-

7 d old *D. suzukii* females were aspirated from the vials used to maintain *D. suzukii* colonies in the laboratory into 22 mL glass vials. Once glass jars and lids were completely dry, a set of 10 *D. suzukii* females was transferred to each of the pre-treated glass jars. Glass jars treated with acetone or deionized water served as the controls. Fly mortality was evaluated after 2-6 hours of exposure depending on nature of the insecticides tested. The exposure time was 2 hrs for zeta-cypermethrin and malathion, and 6 hrs for spinosad. All bioassays were replicated 6-10 times. The LC_{50} values were generated using a range of concentrations of 0.01-3 ppm, 0.3-100 ppm and 0.3-100 ppm for zeta-cypermethrin, spinosad and malathion, respectively. Median lethal concentration (LC_{50}) values were estimated using probit option of the POLO software (Software 1987) and lethal concentration ratios (LCR) at LC_{50} values and their corresponding 95% confidence limits (CL) were calculated using a lethal concentration ratio significance test (Robertson et al.). The Clarke colony served as the reference susceptible population for comparison purposes and was assigned a ratio of 1.0. The LC_{50} values of the Pierce population were considered significantly different from those of the Clarke population if the 95% CL of their corresponding LCR did not include the value of 1.0 ($\alpha = 0.05$).

2.2.3. Sample collection, RNA extraction, and high throughput sequencing

Once the LC_{50} values were established for zeta-cypermethrin, spinosad, and malathion, the 5-7 days old *D. suzukii* Pierce and Clarke populations female flies were treated with their respective LC_{50} concentration values with the same exposure time as for bioassays. For controls, flies were placed in jars coated with either acetone for zeta-cypermethrin and malathion or water for spinosad. A total of 10 survivors per replicate were collected from both insecticide treated (i.e. survivors of LC_{50} dose) and control groups

of Clarke and Pierce populations, and immediately stored at -80 °C. All bioassays for extraction of RNA from flies were replicated three times.

Total RNA was extracted from survivors of treated and control groups using TRIzol (Ambion) as follows. Ten survivors from each group of Pierce and Clarke populations were pooled separately and homogenized in 200 µl of TRIzol reagent using a cordless motor-driven pellet pestle (Grainger) and processed for total RNA according to the manufacturer's instructions (Ambion). The purity and concentrations of RNA samples were determined using NanoDrop spectrophotometer (N-1000) and samples were submitted to the Georgia Genomic Facility (GGF), University of Georgia. GGF performed RNA integrity determination, Poly (A) enrichment of mRNA, cDNA synthesis, library preparations, and sequencing using 75 bp paired-end Illumina NextSeq 500 platform.

2.2.4. Transcriptome assembly, differential expression and statistical analysis

Due to high sequencing base quality at both 3' and 5' ends according to the FastQC output (Mortazavi et al. 2008) and the availability of a reference genome, preprocessing steps were not performed on raw reads. RNA-seq raw datasets were analyzed using the protocol developed by Trapnell et al (Trapnell et al. 2012a). The 24 samples (paired-end reads, including 3 biological replicates) were independently mapped onto the *D. suzukii* genome (SpottedWingFlybase v.1) (Chiu et al. 2013) downloaded from (http://spottedwingflybase.oregonstate.edu/) by using Tophat v2.0.13 (Trapnell et al. 2009, Kim et al. 2013b) which uses Bowtie2 (Langmead et al. 2009b) as an aligner. After the alignment, Cufflinks v2.2.1 (Trapnell et al. 2010b) was used to estimate the expression values of the transcripts in FPKM (Fragments Per Kilobase per Million mapped reads) with the Cuffdiff 2 default geometric normalization. Differentially

expressed genes (FDR < 0.05 after Benjamini-Hochberg correction for multiple-testing) were identified for insecticide-treated or untreated control for either (1) Clarke population or (2) Pierce population. CummeRbund tools (Trapnell et al. 2012a), a package in R v2.15.3, were used to generate scatter, volcano, and squared coefficient of variation plots and to calculate Pearson's correlation coefficient for biological replicates. All sequences are deposited in the NCBI database (GEO series accession number: GSE73595) and SRA accession number (SRP064328).

2.2.5. Comparative analyses across insecticide treatments as well as between Pierce and Clarke populations

In-house written Perl scripts, Awk and Linux commands were used to pull out significantly differentially expressed genes, across three insecticide treatments in Pierce and Clarke population flies, from Cuffdiff outputs. Awk command was used to extract significantly differentially expressed genes from the Cuffdiff outputs. Only genes with FDR<0.05, FPKM values ≥ 0.1 were used for differential expression analyses. An in-house written Perl script was used to generate a list of significantly differentially expressed detoxification, receptor, and defense response genes. Another Perl script was written to identify significantly up and down-regulated *D. suzukii* genes with *D. melanogaster* orthologs. Venn Diagram Plotter (Littlefield and Monroe 2008) was used to generate diagrams illustrating overlapping differential expression of genes between treatments.

2.2.6. GO enrichment analyses

GO Biological processes (BP) and Molecular Function (MF) terms enrichment analyses were performed using DAVID v6.7 algorithm (Huang et al. 2009b, a) with the Benjamini-Hochberg correction for multiple testing and fold change. The Flybase gene ID

for the *D. melanogaster* orthologs to significantly up-regulated and down-regulated *D.suzukii* genes were used as input for GO enrichment analysis. A maximum of top 15 GO terms were selected on the basis of number of sequences with corrected P value (Benjamini) < 0.05 and fold enrichment >= 1.5 to be presented in our results.

2.2.7. SNV and Indel analyses

Bam files (output of Tophat v2.0.13) from individual replicates were first sorted and then combined using SAMtools (Li et al. 2009) mpileup for Pierce and Clarke untreated populations separately. VarScan v2.3.8 (Koboldt et al. 2012), which uses SAMtools mpileup data as input, was used to call for SNVs and indels. Selected VarScan v2.3.8 (Koboldt et al. 2012) options included: minimum read depth at a position to make a call (10), minimum base quality at a position to count a read (25), p-value threshold (0.01), minimum supporting reads at a position to call variants (2) and for minimum variant frequency the threshold default option of 0.2.. We further used SnpEff v4.1i (build 2015-08-14) (Cingolani et al. 2012), in which our *D. suzukii* database was built, to annotate and predict the effects of SNVs and indels.

To predict the effect of nonsynonymous SNVs on protein function, SIFT4G (Vaser et al. 2016) was used. Before using SIFT4G, *D. suzukii* database was built to be used in SIFT4G. The nonsynonymous SNV is predicted to be deleterious on protein function when the SIFT score is ≤0.05. SIFT median measures the diversity of the sequences used for prediction. A warning with low confidence occurs when the sift median is greater than 3.5 because this indicates that the prediction was based on closely related sequences. The low confidence in SIFT score means that the protein alignment does not have enough sequence

diversity because the position artificially appears to be conserved, an amino acid substitution may incorrectly be predicted to be damaging.

An in house python script was written to pull out the SNVs and indels on genes of interest from SnpEff and SIFT4G vcf outputs.

2.3. Results

2.3.1. Insecticide bioassays indicate that Pierce population is less susceptible to three classes of insecticides than Clarke population

The LC_{50} values of zeta-cypermethrin, spinosad, and malathion for *D. suzukii* females from the Clarke colony were 0.49, 2.78 and 10.25 ppm and those for the Pierce population were 1.50, 7.60 and 27.34 ppm, respectively (Table 1). Based on LCR significance test, the Pierce population was significantly less susceptible to zeta-cypermethrin, spinosad, and malathion than the Clarke colony with LCRs of 3.07, 2.73 and 2.66, respectively.

2.3.2. Examination of transcriptomic changes upon insecticide treatments by RNA sequencing

To reveal possible trends that could explain the significant differences in the LC_{50} values of *D. suzukii* Clarke and Pierce populations treated with zeta-cypermethrin, spinosad, or malathion, we used high throughput RNA sequencing to examine transcriptomic changes by performing pairwise comparison between: (1) Clarke population that was treated with an insecticide versus untreated (control); and (2) Pierce population that was treated versus untreated (control). The total paired-end reads obtained per sample ranged from 28 million to 51 million, of which 17 million to 27 million reads had aligned pairs (Additional file 1: Table S1). Pearson's correlation coefficients (r) were calculated

among the three biological replicates of each condition using CummeRbund to evaluate the quality of our identification of differentially expressed genes. Estimates of gene expression abundance expressed as FPKM were highly correlated with r values ranging from 0.89 to 0.99 (Figure 1).

The expression scatter plots show strong positive correlation with Pearson's correlation coefficient (r) values ranging from 0.88 to 0.99, between control and treated samples in Pierce and Clarke population across 3 treatments (Figure 1). The r values between treated and control samples are lower in Pierce population as compared to Clarke population suggesting higher variance in Pierce population. Similar results were obtained when squared coefficient of variance (SCV) were calculated (Additional file 2: Figure S1), suggesting more genetic variability in the Pierce population.

2.3.3. Comparison of differentially expressed genes upon treatments with three classes of insecticide classes in Clarke and Pierce populations

A total of 2411, 2330, and 1310 genes were significantly up-regulated in flies from the *D. suzukii* Pierce population surviving the respective zeta-cypermethrin, spinosad, or malathion treatments (Figure 2A). Of the up-regulated genes, 1223 were common between the 3 insecticide treatments, while 737, 56, and 21 genes overlapped exclusively between zeta-cypermethrin and spinosad, zeta-cypermethrin and malathion, and spinosad and malathion treatments, respectively (Figure 2A). In contrast, only 168, 186, and 86 genes were significantly up-regulated in zeta-cypermethrin, spinosad, and malathion-treated Clarke population, respectively (Figure 2C), with only 29 genes common between the 3 treatments. A total of 20, 28 and 10 genes were common exclusively between zeta-cypermethrin and spinosad, zeta-cypermethrin and malathion, and spinosad and malathion

treatments, respectively (Figure 2C). Similar trends were observed in significantly down-regulated genes of Pierce and Clarke populations (Figure 2B and D). A greater number of genes were down-regulated in Pierce than Clarke population and the number of common genes down-regulated across the three insecticide treatments was also higher in Pierce population flies.

Comparisons of the total number of up-regulated genes in the Pierce and Clarke populations after zeta-cypermethrin, spinosad and malathion treatments illustrates a more substantial transcriptomic response in Pierce population flies to the insecticides (Figure 3A-C). While the Clarke population response was lower, the majority of up-regulated genes in the Clarke population showed were also up-regulated in the Pierce population (Figure 3A, C, and E, yellow region). As with up-regulated genes, the total number of down-regulated genes was greater in Pierce as compared to Clarke population flies (Figure 3B, D and E). In contrast to up-regulated genes, there were relatively few down-regulated genes common between Pierce and Clarke populations after insecticide treatments (Figure 3B, D and F, yellow region).

2.3.4. GO analysis identified enriched functional gene classes in differentially expressed genes upon insecticide treatments

We performed GO enrichment analysis on significantly up-regulated and down-regulated genes in zeta-cypermethrin, spinosad, and malathion-treated Pierce and Clarke populations to identify significantly enriched biological processes (BP) and molecular functions (MF) terms. Using DAVID (Huang et al. 2009b, a), we selected the top 15 significantly enriched BP and MF terms on the basis of number of genes included in the term with Benjamini corrected FDR <0.05 and fold change ≥ 1.5.

2.3.4.1. GO enrichment analysis of up-regulated genes. The top enriched BP and MF terms for significantly up-regulated genes, across 3 treatments in Pierce flies, were oxidation-reduction process (BP) and transporter activity (MF) (Additional file 3: Tables S2 and S3). Those terms, oxidation-reduction process (BP) and transporter activity (MF), were not enriched in Clarke flies insecticide treatments (Additional file 3: Tables S2 and S3).

Significantly enriched GO terms such as oxidoreductase activity and receptor activity, in Pierce population were of particular interest (Additional file 3: Tables S2 and S3). Oxidoreductase activity (MF term) includes cytochrome P450 monooxygenases (CYP), which are involved in phase I detoxification process of insecticides (Mitchell et al. 2014) and have been implicated in insecticide resistance across a wide variety of insects (Rivero et al. 2010). Similarly receptor activity, a MF term, includes nAcR and GABA receptors, which act as target sites for insecticides (ffrench-Constant et al. 1998a, Thompson et al. 2000, Baxter et al. 2010a).

2.3.4.2. GO enrichment analysis of down-regulated genes. Similar BP and MF terms were enriched in significantly down-regulated genes across 3 treatments of Pierce flies (Additional file 3: Tables S4 and S5). The significantly enriched GO terms for insecticide-treated Pierce population include genes involved in cell cycle, chromosome organization, macro-molecule, nitrogen compound metabolic process, binding, hydrolase, and helicase activity (Additional file 3: Tables S4 and S5). In comparison to Pierce population flies, very few BP and MF terms were enriched for significantly down-regulated genes of zeta-cypermethrin, spinosad, and malathion-treated Clarke population. Hydrolase activity was the common enriched MF term across the 3 pesticide treatments in Clarke

population flies, whileoxidoreductase activity was enriched after malathion and spinosad treatments.

2.3.5. Differentially expressed genes upon insecticide treatments are involved in metabolic detoxification

We observed higher number of significantly up-regulated detoxification genes (*Est*, *Gst*, and *Ugt*) in zeta-cypermethrin, spinosad and malathion treated Pierce population with fold changes ≥2 (Table 2). Many of these significantly up-regulated detoxification genes were common across three different classes of insecticide treatments (Table 2, Additional file 4: Tables S6-S17) in Pierce flies. The *Cyp* genes, which encode for enzymes involved in phase I detoxification, were more responsive to three different classes of insecticide treatments in Pierce population relative to other detoxification gene classes (*Est*, *Gst*, and *Ugt*) (Table 2).

The significantly up-regulated (with fold change ≥2) *Cyp* genes in zeta-cypermethrin, spinosad, and malathion treated Pierce population belong to the *Cyp* families 4, 6, 9, 12, 28, 49, 301, 304, 308, and 310 (Table 2, Additional file 4: Tables S9-S11). The up-regulation of members of *Cyp* families 4, 6, 9 and 12 have been implicated in insecticide resistance (Li et al. 2007a).

In contrast only 2 detoxification genes (*Cyp4e3* and *Cyp6w1*) were significantly up-regulated with a fold change ≥2 in Clarke population insecticide treatments (Table 2). The *Cyp4e3* and *Cyp6w1* genes were significantly up-regulated with a fold change ≥2 in spinosad and zeta-cypermethrin treated Clarke population, respectively (Table 2).

We found few detoxification genes, which were significantly up-regulated in insecticide treated Pierce population to be down-regulated in Clarke population. The

Cyp4s3, and *Cyp4p2* genes which were significantly up-regulated across 3 treatments in Pierce population were found to be down-regulated in the spinosad and malathion-treated Clarke population, respectively (Additional file 4: Tables S9-S11, Additional file 5: Tables S21-S23). Similarly, *Est-6* (spinosad and malathion) and *alpha-Est8* (spinosad) were significantly down-regulated in Clarke population treatments were found to be up-regulated in Pierce population treatments (Additional file 4: Tables S4-S6, Additional file 5: Tables S18-S20).

2.3.6. Identification of differentially expressed insecticide target genes upon insecticide treatments

We further analyzed the expression levels of insecticide target genes in Pierce and Clarke insecticides treated samples relative to their respective control samples as G-protein-coupled receptor (GPCR), a potential target for novel insecticides and the target for Amitraz pesticide (Corley et al. 2013), have been shown to be significantly differentially expressed in permethrin resistant *C. quinquefasciatus* strain relative to its susceptible strain (Li et al. 2014). In Pierce population insecticide treatments, we observed many insecticide target (GABA receptor, nicotinic acetylcholine receptor, and sodium channel**)** encoding genes significantly up-regulated with a fold change ≥2 (Table 3, Additional file 6: Tables S30-S38). In contrast, no insecticide target genes were significantly up-regulated with a fold change ≥2 in Clarke population insecticide treatments (Table 3, Additional file 7: Tables S42-S50).

2.3.7. ATP-binding cassette transporters are up-regulated upon insecticide treatments in D. suzukii

We also analyzed the expression levels of *Mdr* genes in Pierce and Clarke population insecticide treated samples relative to their respective control samples. *Mdr* genes encodes for P-glycoproteins (P-gp) (also called ATP-binding cassette (ABC) transporters) which have recently gained attention for their contributions to insecticide and Bt Cry resistance in insects (Heckel 2012, Merzendorfer 2014). P-gps have been shown to protect *Culex pipiens* larvae from cypermethrin, endosulfan, and ivermectin toxicity (Buss et al. 2002). We found *Mdr65* and *Mdr50* genes significantly up-regulated in zeta-cypermethrin, and spinosad treated Pierce population with a fold change ≥2 (Table 3, Additional file 6: Tables S39-S41). No *Mdr* genes were significantly up-regulated with a fold change ≥2 in Clarke population treatments (Table 3, Additional file 6: Tables S39-S41). In malathion treated Clarke population we found *Mdr65* and *Mdr49* genes significantly down-regulated (Additional file 7: Tables S39-41),

2.3.8. Altered expression of cuticular protein genes is a potential mechanism to reduce penetration of pesticides

We analyzed the *Cpr* genes response to insecticide treatments in Pierce, and Clarke populations as significantly higher expression levels of cuticular protein genes have been observed in pyrethroid resistant bed bugs (*Cimex lectularius L.*) (Koganemaru et al. 2013). A total of 13, 12, and 13 *Cpr* genes were significantly up-regulated in zeta-cypermetrin, spinosad, and malathion treatments, respectively, with a fold change ≥ 2 in Pierce population (Table 4). The *Cpr92F*, *Cpr100A*, *Cpr47Ec*, *Cpr62Bc*, *Cpr49Ah*, *Cpr76Bd*, and *Cpr49Ae* genes were significantly up-regulated with a fold change ≥ 5 across 3 treatments

in Pierce population (Additional file 8: Tables S54-S56). In contrast no *Cpr* genes were significantly up-regulated with fold change ≥ 2 in zeta-cypermethrin, spinosad, and malathion treated Clarke population (Table 4). The *Cpr78Cc* and *Cpr49Ae* genes were significantly down-regulated in zeta-cypermethrin-treated Pierce and spinosad-treated Clarke populations, respectively (Additional file 8: Tables S54-S56, Additional file 9: Tables S57-S59).

2.3.9. Comparing the basal levels of gene expression in Pierce and Clarke D. suzukii without insecticide treatments

In total, 2860 and 3302 genes were significantly up- and down-regulated respectively in untreated Pierce population samples relative to Clarke population samples (Additional file 10: Table S66, Additional file 11: Table S67). The expression levels of genes encoding proteins involved in detoxification process and target sites of insecticides were significantly down-regulated in *D. suzukii* Pierce population when compared to Clarke population in control samples with the exceptions of *Cyp6a22*, *Cyp6v1*, *GstO2*, *GstD5* and *rpk*, which were significantly up-regulated in Pierce population control samples (Additional file 12: Table S68). The genes encoding cuticular proteins were significantly down-regulated in Pierce population (Additional file 13: Table S69).

2.3.10. Higher number of predicted nonsynonymous SNVs and indels observed in genes encoding detoxification enzymes, insecticide targets and cuticular proteins in Clarke population

In addition to assaying gene expression level, we performed single nucleotide variant (SNV) and indel analysis as a means to predict the functionality of the DEGs. We predicted a total of 19,432 and 18,315 nonsynonymous SNVs in all expressed genes in

Pierce and Clarke population, respectively (Figure 4). We then focused our analysis on genes encoding detoxification enzymes, insecticide targets and cuticular proteins, specifically identifying SNVs and indels in genes that are significantly differentially expressed in Pierce and Clarke populations upon treatment with at least one insecticide. We examined this group of genes because mutations or overexpression among these genes have been implicated in insecticide resistance [17]. In comparison to Clarke population, fewer nonsynonymous SNVs were predicted in *CYPs*, *ESTs*, *GSTs*, *UGTs*, *MDRs*, *GABAs*, *nAcRs*, sodium channel and cuticular protein encoding genes in Pierce population (Figure 5). Furthermore we predicted the effects of nonsynonymous SNVs on the protein function in the above mentioned group of genes using SIFT4G(Vaser et al. 2016). Nonsynonymous SNVs in *Cyp12d1-d*, *Cyp9b2*, *Cyp6d5* (DS10_00010146), *Cyp4d1*, *Est-6*, *Ugt36Bb*, *Ugt86Dd*, *Mdr49*, *Rdl* (DS10_00003821), *para*, *rpk* (DS10_00010216) and *Cpr49Ac* genes in Clarke population were predicted to have deleterious effect on protein function (Table 5). In contrast, only 6 genes (*Cyp9b2*, *Cyp4p1*, *Cyp4d1*, *Ugt86Dd*, *Ugt* and *rpk* (DS10_00010216)) were predicted to have nonsynonymous SNVs leading to deleterious effect in the Pierce population (Table 5). Among them, *Cyp9b2*, *Cyp4d1*, *Ugt86Dd* and *rpk* genes have common nonsynonymous SNVs with deleterious effect, in Pierce and Clarke population. Furthermore *Cyp9b2* has additional predicted deleterious SNVs in Clarke population (Table 5). *Mdr49* gene has 2 nonsynonymous SNVs with predicted deleterious effect in the Clarke population but none were predicted in the Pierce population (Table 5, Additional file 14: Table S71, Additional file 15: Table S74). Coincidently, these 2 nonsynonymous SNVs predicted to have deleterious effect on protein function are present

in the highly conserved Walker B domain of MDR49 protein (based on NCBI conserved domain database search).

Indels (insertions and deletions) were predicted for the above-mentioned group of genes in untranslated (UTR) and coding sequence (CDS) regions, which could lead to deleterious effects on expression pattern or protein function. *Cyp4s3* gene has 2 insertions and 2 deletions in 3'-UTR region of Clarke population. In contrast only 1 insertion was predicted in 3'-UTR region of *Cyp4s3* gene Pierce population (Table 6). *Cpr92F* gene has a predicted frameshift deletion in the CDS region in Clarke population but not in Pierce population (Table 6).

2.4. Discussion

We identified differentially expressed genes (DEGs) upon insecticide treatments in Clarke and Pierce county D. suzukii populations after treatments with zeta-cypermethrin, spinosad, or malathion insecticides. Gene Ontology (GO) enrichment analysis identified gene classes that are overrepresented among the DEGs. We then focused our analysis on detoxification and cuticle related genes due to their known functions in conferring insecticide resistance (Li et al. 2007a, Rivero et al. 2010, Silva et al. 2012) and found that many of these genes were significantly up-regulated upon insecticide treatment. Based on enriched categories of up-regulated detoxification genes, Cytochrome P450 monooxygenases are the likely candidates for primary detoxification process in zeta-cypermethrin, spinosad, and malathion-treated Pierce population. However other families of detoxification enzymes, which were significantly up-regulated, such as ESTs, UGTs and GSTs may have a secondary role in detoxification of these three classes of insecticides. The relative significance of cytochrome P450 genes could be experimentally tested by

inhibiting the enzymes with pipronyl butoxide (PBO) and comparing LC_{50} values of insecticides.

In *D. suzukii* Pierce population treatments 7 *Cpr* genes were overexpressed with a fold change ≥5 and none of the *Cpr* genes were overexpressed with fold change ≥2 in Clarke population treatments. Cuticular protein encoding genes have previously been shown to be overexpressed in pyrethroid resistant *Anopheles* species (Awolola et al. 2009, Gregory et al. 2011, Bonizzoni et al. 2012, Nkya et al. 2014). This makes *Cpr genes* interesting candidates to track for insecticide resistance development in *D. suzukii*.

Perhaps one of the more interesting findings is that many of the DEGs were regulated in the same manner upon treatment with multiple insecticide classes, suggesting these genes are involved in mechanisms that are commonly induced to counteract and cope with the negative effects of different classes of insecticides. In the event that genetic changes result in altered expression of these DEGs, e.g. up-regulation of a gene that is normally induced upon insecticide treatment, this may lead to the development of cross-resistance against more than one class of insecticides. Our results therefore indicate a high risk for cross-resistance development in the field.

Moreover, a higher number of significantly DEGs were identified in Pierce population when compared to Clarke population. This correlates with the higher LC_{50} values observed in Pierce county flies and suggest that Pierce population has stronger capacity to detoxify insecticides and cope with insecticide treatments by other mechanisms. Specifically, we observed a substantially higher number of up-regulated detoxification genes in zeta-cypermethrin, spinosad, and malathion-treated Pierce population as

compared to insecticide-treated Clarke population, suggesting that Pierce population has a much higher potential to detoxify these three classes of insecticides upon treatment.

It is important to point out that we identified genes that are commonly differentially expressed (DE) in Pierce and Clarke populations, although we observed a higher number of DEGs in Pierce population upon insecticide treatments. This suggests that the DEGs that are exclusively identified in Pierce population may be responsible for the lower susceptibility of Pierce population towards zeta-cypermethrin, spinosad and malathion. In fact, results from insecticide bioassays showed higher LC_{50} values in field population, suggesting they are indeed less susceptible.

Interestingly and rather surprisingly, the baseline expression level of many genes, including those known to confer insecticide resistance, are higher in Clarke population samples when compared to the Pierce population untreated samples. These results are quite puzzling; since the LC_{50} value in Clarke population is lower as compared to the Pierce population, we initially expected the Pierce population to have a higher baseline expression of genes, especially those that are involved in detoxification and conferring insecticide resistance. Our results suggest that the induction level of these genes could be more important than the baseline gene expression level when considering the action of these protein products to counteract the effects of insecticides.

Our analysis on the occurrence of deleterious mutations, specifically nonsynonymous SNVs and indels, may also provide insight into our observation that untreated Clarke population have higher baseline expression level of many genes, including those involved in insecticide response, as compared to Pierce population. Although we identified a slightly higher number of predicted nonsynonymous SNVs in

Pierce population when compared to Clarke population when we considered all expressed genes. In contrast we observed a higher number of nonsynonymous SNVs in Clarke population samples when only focusing on genes encoding detoxification enzymes, insecticide targets, and cuticular proteins. Furthermore, as many insecticide detoxification and response genes in the Clarke population carry nonsynonymous SNVs that are predicted to have deleterious effects to protein function, it is perhaps not surprising that many of these genes were not differentially expressed upon insecticide treatment. This suggests that mechanisms exist to prevent the induction of genes producing non-functional protein products.

The results from our combined gene expression and SNV analysis in Clarke and Pierce populations suggest that this approach may be effective in identifying potential candidates involved in insecticide response for insect population. In our analysis, these would be the insecticide response genes that lack nonsynonymous mutations with deleterious effects and are significantly differentially expressed in the Pierce population.

Finally, an important implication of our findings for future studies is that genes that are characterized for resistance mechanism in laboratory-developed resistant strains may not always be related to resistance mechanism in field-evolved resistant strains.

2.5. Concluding remarks

Our transcriptomic analysis not only identified the genes that are inducible with specific insecticide treatments and possible resistance mechanisms, but also serve as a baseline for comparing similar data from field-evolved resistant *D. suzukii* populations in the future. We anticipate that high throughput sequencing can be used as an effective tool

to assess the risk of insecticide resistance development and lead to strategies to delay the development of resistance in devastating pests such as *D. suzukii* in a timely manner.

Availability of supporting data

The data discussed in the submission have been submitted to NCBI's Gene Expression Omnibus (GEO) and are accessible through GEO series accession number GSE73595 (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE73595). BioProject Id: PRJNA297377 and SRA Id: SRP064328.

SnpEff and SIFT4G output data available from following link (http://spottedwingflybase.org/downloads).

Note: Field and Laboratory population in the NCBI GEO has been re-named as Pierce and Clarke, respectively in the manuscript.

2.6. References

Aminetzach, Y. T., J. M. Macpherson, and D. A. Petrov. 2005. Pesticide resistance via transposition-mediated adaptive gene truncation in Drosophila. Science 309: 764-767.

Awolola, T. S., O. A. Oduola, C. Strode, L. L. Koekemoer, B. Brooke, and H. Ranson. 2009. Evidence of multiple pyrethroid resistance mechanisms in the malaria vector Anopheles gambiae sensu stricto from Nigeria. Trans R Soc of Trop Med Hyg 103: 1139-1145.

Baxter, S. W., M. Chen, A. Dawson, J.-Z. Zhao, H. Vogel, A. M. Shelton, D. G. Heckel, and C. D. Jiggins. 2010. Mis-Spliced Transcripts of Nicotinic Acetylcholine Receptor alpha 6 Are Associated with Field Evolved Spinosad Resistance in Plutella xylostella (L.). Plos Genet 6.

Beers, E. H., R. A. Van Steenwyk, P. W. Shearer, W. W. Coates, and J. A. Grant. 2011. Developing Drosophila suzukii management programs for sweet cherry in the western United States. Pest management science 67: 1386-1395.

Bolda, M. 2011. Suspected tolerance to pyganic (pyrethrin) found in spotted wing drosophila. Posted on August 23.

Bonizzoni, M., Y. Afrane, W. A. Dunn, F. K. Atieli, G. Zhou, D. Zhong, J. Li, A. Githeko, and G. Yan. 2012. Comparative Transcriptome Analyses of Deltamethrin-Resistant and - Susceptible Anopheles gambiae Mosquitoes from Kenya by RNA-Seq. Plos One 7: e44607.

Bruck, D. J., M. Bolda, L. Tanigoshi, J. Klick, J. Kleiber, J. DeFrancesco, B. Gerdeman, and H. Spitler. 2011. Laboratory and field comparisons of insecticides to reduce infestation of Drosophila suzukii in berry crops. Pest Manag Sci 67: 1375-1385.

Burrack, H. J., J. P. Smith, D. G. Pfeiffer, G. Koeher, and J. Laforest. 2012. Using volunteer-based networks to track Drosophila suzukii (Diptera: Drosophilidae) an invasive pest of fruit crops. J Integ Pest Mngmt 3: B1-B5.

Buss, D. S., A. R. McCaffery, and A. Callaghan. 2002. Evidence for p-glycoprotein modification of insecticide toxicity in mosquitoes of the Culex pipiens complex. Med Vet Entomol 16: 218-222.

Chiu, J. C., X. Jiang, L. Zhao, C. A. Hamm, J. M. Cridland, P. Saelao, K. A. Hamby, E. K. Lee, R. S. Kwok, G. Zhang, F. G. Zalom, V. M. Walton, and D. J. Begun. 2013. Genome of Drosophila suzukii, the Spotted Wing Drosophila. G3 3: 2257-2271.

Cingolani, P., A. Platts, L. L. Wang, M. Coon, T. Nguyen, L. Wang, S. J. Land, X. Lu, and D. M. Ruden. 2012. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly 6: 80-92.

Cini, A., C. Ioriatti, and G. Anfora. 2012. A review of the invasion of Drosophila suzukii in Europe and a draft research agenda for integrated pest management. Bull Insectol 65: 149-160.

Claudianos, C., R. J. Russell, and J. G. Oakeshott. 1999. The same amino acid substitution in orthologous esterases confers organophosphate resistance on the house fly and a blowfly. Insect Biochem Mol Biol 29: 675-686.

Corley, S. W., N. N. Jonsson, E. K. Piper, C. Cutullé, M. J. Stear, and J. M. Seddon. 2013. Mutation in the RmβAOR gene is associated with amitraz resistance in the cattle tick Rhipicephalus microplus. Proc Natl Acad Sci USA 110: 16772-16777.

Daborn, P. J., J. L. Yen, M. R. Bogwitz, G. Le Goff, E. Feil, S. Jeffers, N. Tijet, T. Perry, D. Heckel, P. Batterham, R. Feyereisen, T. G. Wilson, and R. H. ffrench-Constant. 2002. A single P450 allele associated with insecticide resistance in Drosophila. Science 297: 2253-2256.

David, J.-P., F. Faucon, A. Chandor-Proust, R. Poupardin, M. A. Riaz, A. Bonin, V. Navratil, and S. Reynaud. 2014. Comparative analysis of response to selection with three insecticides in the dengue mosquito Aedes aegypti using mRNA sequencing. BMC Genomics 15.

Deprá, M., J. L. Poppe, H. J. Schmitz, D. C. De Toni, and V. L. Valente. 2014. The first records of the invasive pest Drosophila suzukii in the South American continent. J Pest Sci 87: 379-383.

Feyereisen, R., L. I. Gilbert, K. Iatrou, and S. S. Gill. 2005. Insect cytochrome P450. Comprehensive molecular insect science. Volume four. Biochemistry and molecular biology.: 1-77.

ffrench-Constant, R. H., P. J. Daborn, and G. Le Goff. 2004. The genetics and genomics of insecticide resistance. Trends Genet 20: 163-170.

ffrench-Constant, R. H., B. Pittendrigh, A. Vaughan, and N. Anthony. 1998. Why are there so few resistance-associated mutations in insecticide target genes? Philos Trans R Soc Lond B Biol Sci 353: 1685-1693.

Gregory, R., A. C. Darby, H. Irving, M. B. Coulibaly, M. Hughes, L. L. Koekemoer, M. Coetzee, H. Ranson, J. Hemingway, N. Hall, and C. S. Wondji. 2011. A De Novo Expression Profiling of Anopheles funestus, Malaria Vector in Africa, Using 454 Pyrosequencing. Plos One 6: e17418.

Guio, L., M. G. Barron, and J. Gonzalez. 2014. The transposable element Bari-Jheh mediates oxidative stress response in Drosophila. Mol Ecol 23: 2020-2030.

Hauser, M. 2011. A historic account of the invasion of Drosophila suzukii (Matsumura) (Diptera: Drosophilidae) in the continental United States, with remarks on their identification. Pest Manag Sci 67: 1352-1357.

Haviland, D. R., and E. H. Beers. 2012. Chemical control programs for Drosophila suzukii that comply with international limitations on pesticide residues for exported sweet cherries. J Integ Pest Mngmt 3: F1-F6.

Heckel, D. G. 2012. Learning the ABC's of Bt: ABC transporters and insect resistance to *Bacillus thuringiensis* provide clues to a crucial step in toxin mode of action. Pestic Biochem Physiol 104: 103-110.

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

Hemingway, J., N. Hawkes, L. A. Prapanthadara, K. G. I. Jayawardenal, and H. Ranson. 1998. The role of gene splicing, gene amplification and regulation in mosquito insecticide resistance. Trans R Soc Lond B Biol Sci 353: 1695-1699.

Huang, D. W., B. T. Sherman, and R. A. Lempicki. 2009a. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 4: 44-57.

Huang, D. W., B. T. Sherman, and R. A. Lempicki. 2009b. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res 37: 1-13.

Kalajdzic, P., S. Oehler, M. Reczko, N. Pavlidi, J. Vontas, A. G. Hatzigeorgiou, and C. Savakis. 2012. Use of Mutagenesis, Genetic Mapping and Next Generation Transcriptomics to Investigate Insecticide Resistance Mechanisms. Plos One 7: e40296.

Karasov, T., P. W. Messer, and D. A. Petrov. 2010. Evidence that Adaptation in Drosophila Is Not Limited by Mutation at Single Sites. Plos Genet 6: e1000924.

Karunaratne, S., A. Vaughan, M. G. Paton, and J. Hemingway. 1998. Amplification of a serine esterase gene is involved in insecticide resistance in Sri Lankan Culex tritaeniorhynchus. Insect Mol Biol 7: 307-315.

Kasai, S., and J. G. Scott. 2001. A house fly gene homologous to the zinc finger proto-oncogene Gfi-1. Biochem Biophys Res Commun 283: 644-647.

Kim, D., G. Pertea, C. Trapnell, H. Pimentel, R. Kelley, and S. L. Salzberg. 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol 14.

Koboldt, D. C., Q. Zhang, D. E. Larson, D. Shen, M. D. McLellan, L. Lin, C. A. Miller, E. R. Mardis, L. Ding, and R. K. Wilson. 2012. VarScan 2: Somatic mutation and copy number alteration discovery in cancer by exome sequencing. Genome Res 22: 568-576.

Koganemaru, R., D. M. Miller, and Z. N. Adelman. 2013. Robust cuticular penetration resistance in the common bed bug (Cimex lectularius L.) correlates with increased steady-state transcript levels of CPR-type cuticle protein genes. Pestic Biochem Physiol 106: 190-197.

Kopp, A., and J. R. True. 2002. Evolution of male sexual characters in the Oriental Drosophila melanogaster species group. Evol Dev 4: 278-291.

Langmead, B., C. Trapnell, M. Pop, and S. L. Salzberg. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 10.

Lee, J. C., D. J. Bruck, H. Curry, D. Edwards, D. R. Haviland, R. A. Van Steenwyk, and B. M. Yorgey. 2011. The susceptibility of small fruits and cherries to the spotted-wing drosophila, Drosophila suzukii. Pest Manag Sci 67: 1358-1367.

Li, H., B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, and R. Durbin. 2009. The sequence alignment/map format and SAMtools. Bioinformatics 25: 2078-2079.

Li, T., L. Liu, L. Zhang, and N. Liu. 2014. Role of G-protein-coupled receptor-related genes in insecticide resistance of the mosquito, Culex quinquefasciatus. Sci. Rep. 4: 6474.

Li, X., M. A. Schuler, and M. R. Berenbaum. 2007. Molecular mechanisms of metabolic resistance to synthetic and natural xenobiotics. Annu Rev Entomol 52: 231-253.

Littlefield, K., and M. Monroe. 2008. Venn Diagram Plotter. Pacific Northwest National Laboratory, Richland, WA.

Magwire, M. M., D. K. Fabian, H. Schweyen, C. Cao, B. Longdon, F. Bayer, and F. M. Jiggins. 2012. Genome-Wide Association Studies Reveal a Simple Genetic Basis of Resistance to Naturally Coevolving Viruses in Drosophila melanogaster. Plos Genet 8.

Maitra, S., C. Price, and R. Ganguly. 2002. Cyp6a8 of Drosophila melanogaster: gene structure, and sequence and functional analysis of the upstream DNA. Insect Biochem Mol Biol 32: 859-870.

Mateo, L., A. Ullastres, and J. Gonzalec. 2014. A Transposable Element Insertion Confers Xenobiotic Resistance in Drosophila. Plos Genet 10.

Merzendorfer, H. 2014. ABC transporters and their role in protecting insects from pesticides and their metabolites. Adv Insect Physiol 46: 1-72.

Mitchell, C. L., M. C. Saul, L. Lei, H. Wei, and T. Werner. 2014. The Mechanisms Underlying alpha-Amanitin Resistance in Drosophila melanogaster: A Microarray Analysis. Plos One 9: e93489.

Mortazavi, A., B. A. Williams, K. McCue, L. Schaeffer, and B. Wold. 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods 5: 621-628.

Nkya, T. E., R. Poupardin, F. Laporte, I. Akhouayri, F. Mosha, S. Magesa, W. Kisinza, and J.-P. David. 2014. Impact of agriculture on the selection of insecticide resistance in the malaria vector Anopheles gambiae: a multigenerational study in controlled conditions. Parasit Vectors 7: 480.

Pedra, J. H. F., L. M. McIntyre, M. E. Scharf, and B. R. Pittendrigh. 2004. Genom-wide transcription profile of field- and laboratory-selected dichlorodiphenyltrichloroethane (DDT)-resistant Drosophila. Proc Nat Acad Sci USA 101: 7034-7039.

Puinean, A. M., S. P. Foster, L. Oliphant, I. Denholm, L. M. Field, N. S. Millar, M. S. Williamson, and C. Bass. 2010. Amplification of a Cytochrome P450 Gene Is Associated with Resistance to Neonicotinoid Insecticides in the Aphid Myzus persicae. Plos Genet 6: e1000999.

Rivero, A., J. Vezilier, M. Weill, A. F. Read, and S. Gandon. 2010. Insecticide Control of Vector-Borne Diseases: When Is Insecticide Resistance a Problem? Plos Pathog 6: e1001000.

Robertson, J. L., N. Savin, H. K. Preisler, and R. M. Russell. Binary quantal response with one explanatory variable. Bioassays with Arthropods, 2nd edition, CRC press, Boca Raton, FL; 2007. pp. 21-34

Seifert, J., and J. G. Scott. 2002. The CYP6D1v1 allele is associated with pyrethroid resistance in the house fly, Musca domestica. Pestic Biochem Physiol 72: 40-44.

Silva, A. X., G. Jander, H. Samaniego, J. S. Ramsey, and C. C. Figueroa. 2012. Insecticide Resistance Mechanisms in the Green Peach Aphid Myzus persicae (Hemiptera: Aphididae) I: A Transcriptomic Survey. Plos One 7: e36366.

Software, L. 1987. POLO-PC: a user's guide to probit or logit analysis. LeOra Software Berkeley, CA.

Thompson, G. D., R. Dutton, and T. C. Sparks. 2000. Spinosad - a case study: an example from a natural products discovery programme. Pest Manag Sci 56: 696-702.

Trapnell, C., L. Pachter, and S. L. Salzberg. 2009. TopHat: discovering splice junctions with RNA-Seq. Bioinformatics 25: 1105-1111.

Trapnell, C., B. A. Williams, G. Pertea, A. Mortazavi, G. Kwan, M. J. van Baren, S. L. Salzberg, B. J. Wold, and L. Pachter. 2010. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol 28: 511-U174.

Trapnell, C., A. Roberts, L. Goff, G. Pertea, D. Kim, D. R. Kelley, H. Pimentel, S. L. Salzberg, J. L. Rinn, and L. Pachter. 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc 7: 562-578.

Vaser, R., S. Adusumalli, S. N. Leng, M. Sikic, and P. C. Ng. 2016. SIFT missense predictions for genomes. Nat protoc 11: 1-9.

Vontas, J. G., G. J. Small, and J. Hemingway. 2001. Glutathione S-transferases as antioxidant defence agents confer pyrethroid resistance in Nilaparvata lugens. The Biochem J 357: 65-72.

Walsh, D. B., M. P. Bolda, R. E. Goodhue, A. J. Dreves, J. Lee, D. J. Bruck, V. M. Walton, S. D. O'Neal, and F. G. Zalom. 2011. Drosophila suzukii (Diptera: Drosophilidae): invasive pest of ripening soft fruit expanding its geographic range and damage potential. J Integ Pest Mngmt 2: G1-G7.

Zhu, Y. C., A. K. Dowdy, and J. E. Baker. 1999. Detection of single-base substitution in an esterase gene and its linkage to malathion resistance in the parasitoid Anisopteromalus calandrae (Hymenoptera : Pteromalidae). Pestic Sci 55: 398-404.

Figure legends

Figure 2.1. Scatter plots of log_{10} (FPKM). Scatter plots showing pairwise comparisons between control and insecticide-treated *D. suzukii* Pierce and Clarke populations**.** Results for Pierce population are shown in (A-C) and comparisons for Clarke population are shown in (D-F)**.** The FPKM values for all transcripts were plotted for control and treated samples by averaging across the biological replicates and normalization. Pairwise comparisons were generated include **(**A and D) control vs. zeta-cypermethrin, (B and E) control vs. spinosad, and (C and F) control vs. malathion. Each dot on the scatterplot represents a gene. CummeRbund was used for statistical analyses and generation of plots.

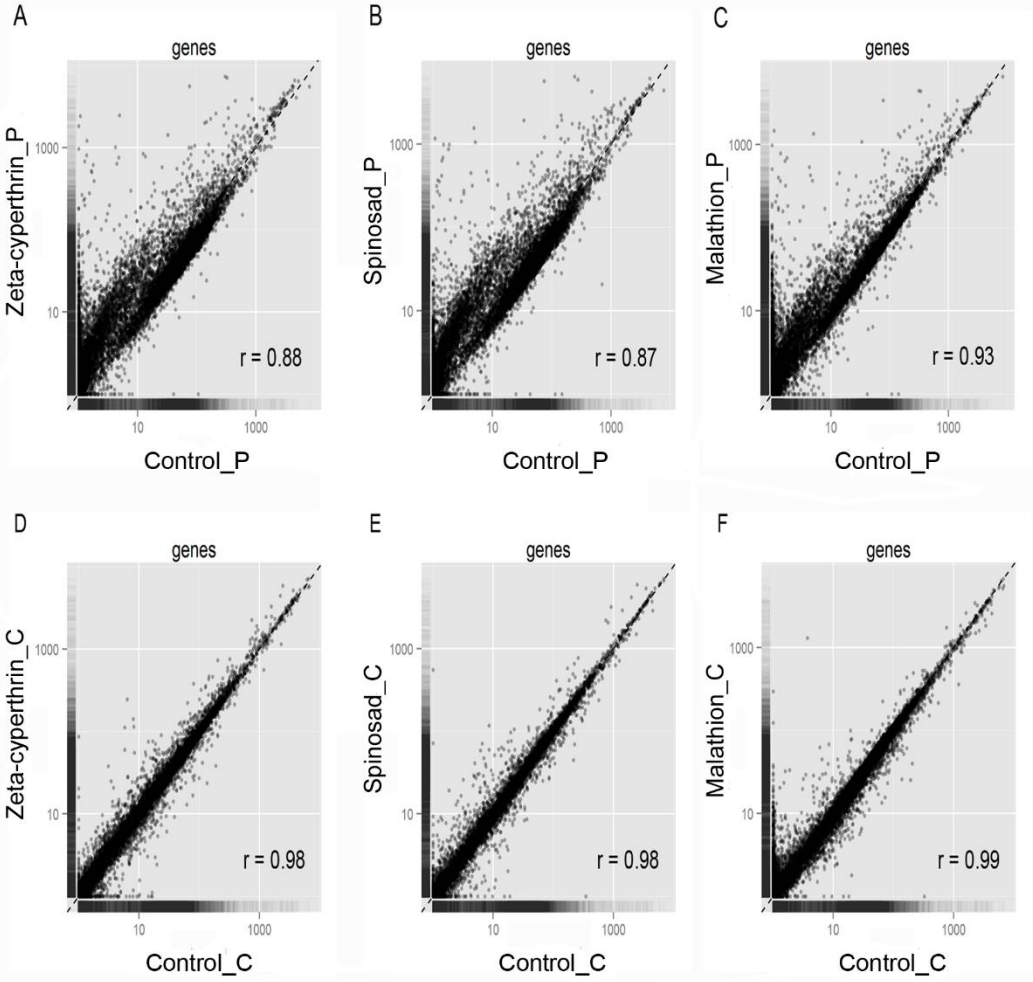

A
genes
Zeta-cyperthrin_P
Control_P
r = 0.88
B
genes
Spinosad_P
Control_P
r = 0.87
C
genes
Malathion_P
Control_P
r = 0.93
D
genes
Zeta-cyperthrin_C
Control_C
r = 0.98
E
genes
Spinosad_C
Control_C
r = 0.98
F
genes
Malathion_C
Control_C
r = 0.99
10
1000

Figure 2.2. Venn diagrams depicting relationship of differentially expressed genes between three distinct insecticide treatments. Comparisons between the up-regulated genes in zeta-cypermethrin, spinosad, and malathion-treated *D. suzukii* Pierce (A) and Clarke (C) populations. Similarly comparisons between the down-regulated genes in zeta-cypermethrin, spinosad, and malathion-treated *D. suzukii* Pierce (B) and Clarke (D) populations. Color representation for up- and down-regulated genes in *D. suzukii* Pierce and Clarke population are as follow: Red, Green, and Blue: genes that are exclusively differentially regulated in zeta-cypermethrin, spinosad, and malathion-treated flies, respectively; White: Common to three treatments; Yellow: Exclusive to zeta-cypermethrin and spinosad; Cyan: Exclusive to spinosad and malathion; Magenta: Exclusive to zeta-cypermethrin and malathion. Venn diagram plotter was used to generate proportional Venn diagrams.

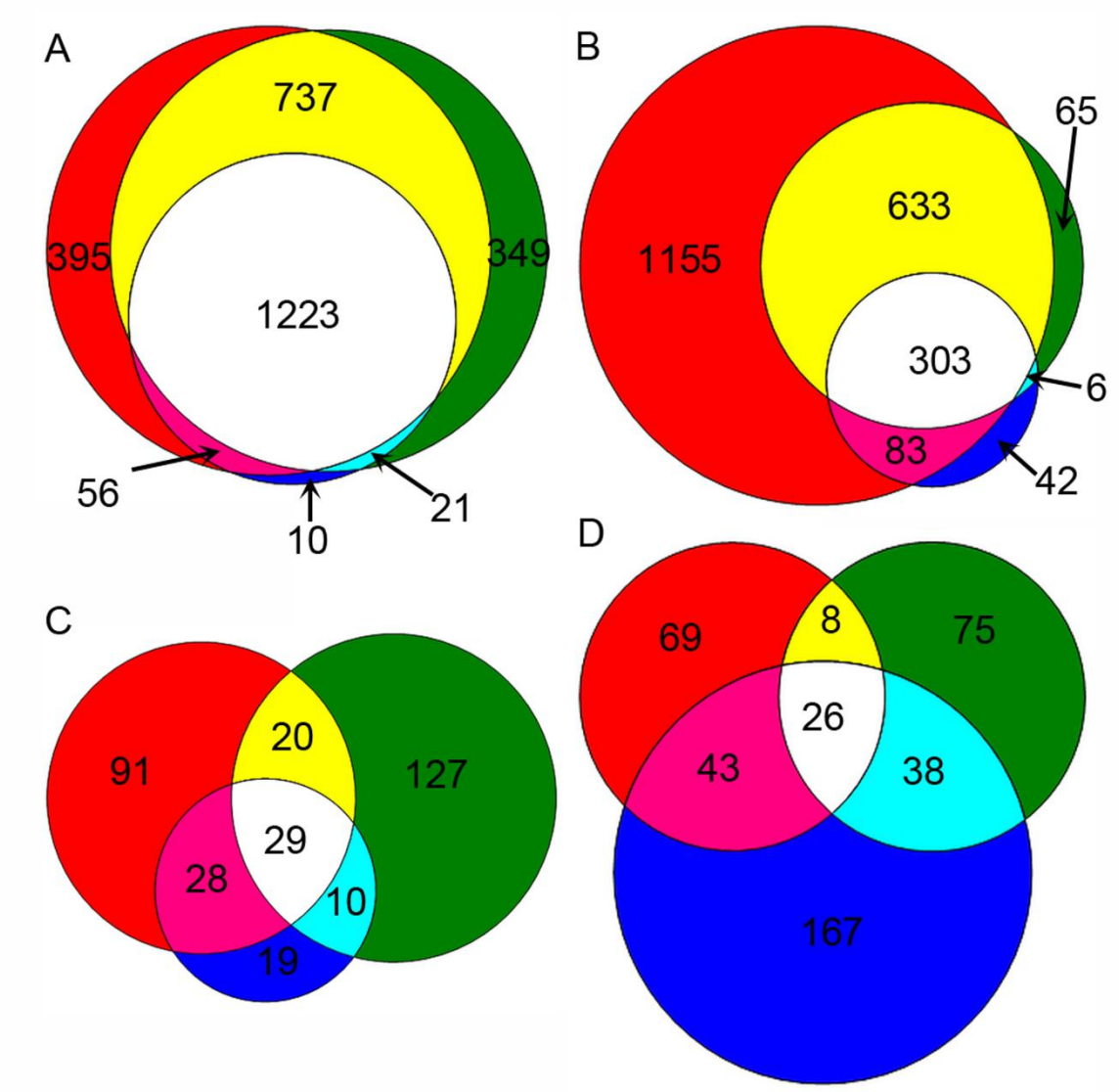

A
737
395
349
1223
56
10
21
B
65
633
1155
303
6
83
42
C
20
91
127
29
28
10
19
D
69
8
75
26
43
38
167

Figure 2.3. Venn diagrams depicting relationship of differentially expressed genes between *D. suzukii* Pierce and Clarke population. (A, C, and E): Comparisons of up-regulated genes between *D. suzukii* Pierce and Clarke population in zeta-cypermethrin (A), spinosad (C), and malathion (E) treatments. (B, D, and F): Comparisons of down-regulated genes between *D. suzukii* Pierce and Clarke population in zeta-cypermethrin (B), spinosad (D), and malathion (F) treatments. Color representation for up- and down-regulated genes comparisons between Pierce and Clarke population are as follows: Red: Exclusive to Pierce population; Green: Exclusive to Clarke population; Yellow: common to Pierce and Clarke population.

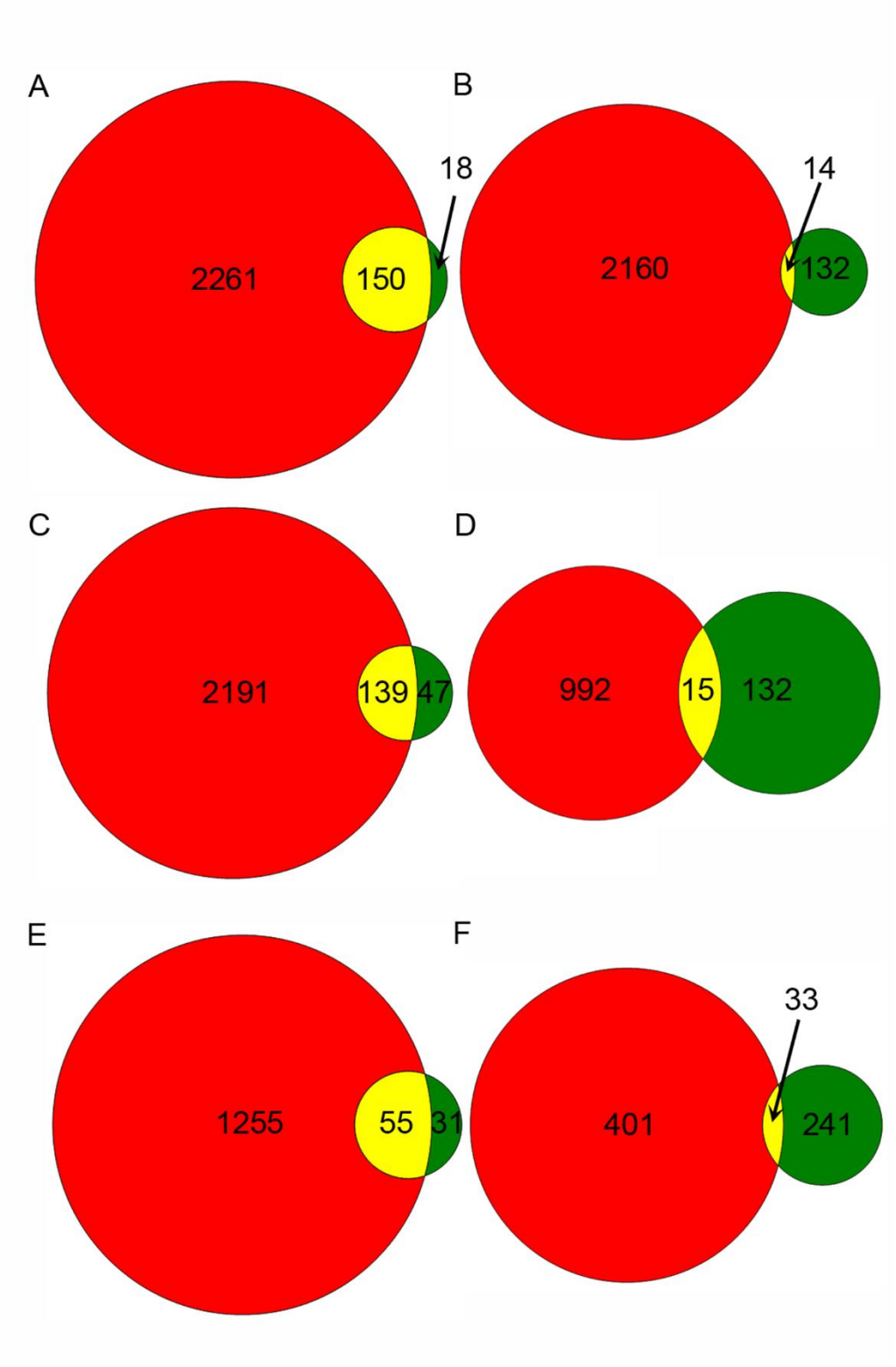

A
2261
150
18
B
2160
132
14
C
2191
139
47
D
992
15
132
E
1255
55
31
F
401
241
33

Figure 2.4. Predicted Nonsynonymous SNVs in Pierce and Clarke population. Total number of nonsynonymous SNVs predicted in all expressed genes in Pierce and Clarke population. Figure was generated using R v3.2.3.

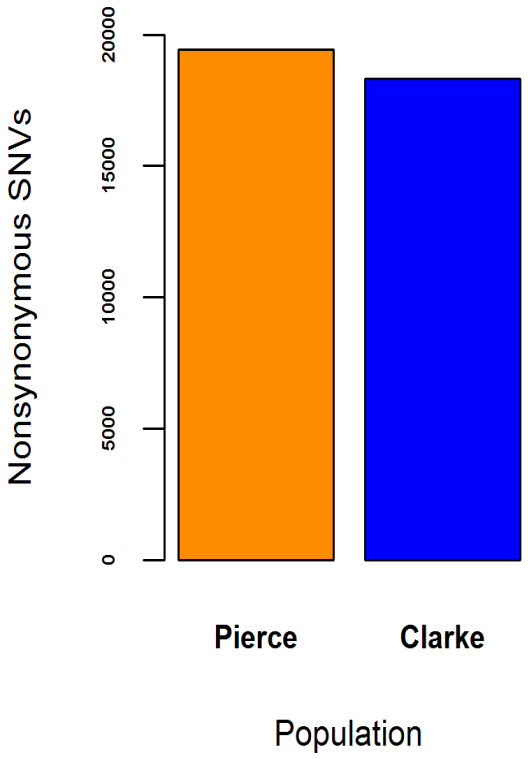

Nonsynonymous SNVs
0
5000
10000
15000
20000
Pierce
Clarke
Population

Figure 2.5. Number of predicted Nonsynonymous SNVs in detoxification, insecticide target and cuticular protein encoding genes in Pierce and Clarke population.

Significantly differentially expressed genes in the above mentioned gene groups in Pierce and Clarke population were selected for analyses. Figure was generated using R v3.2.3.

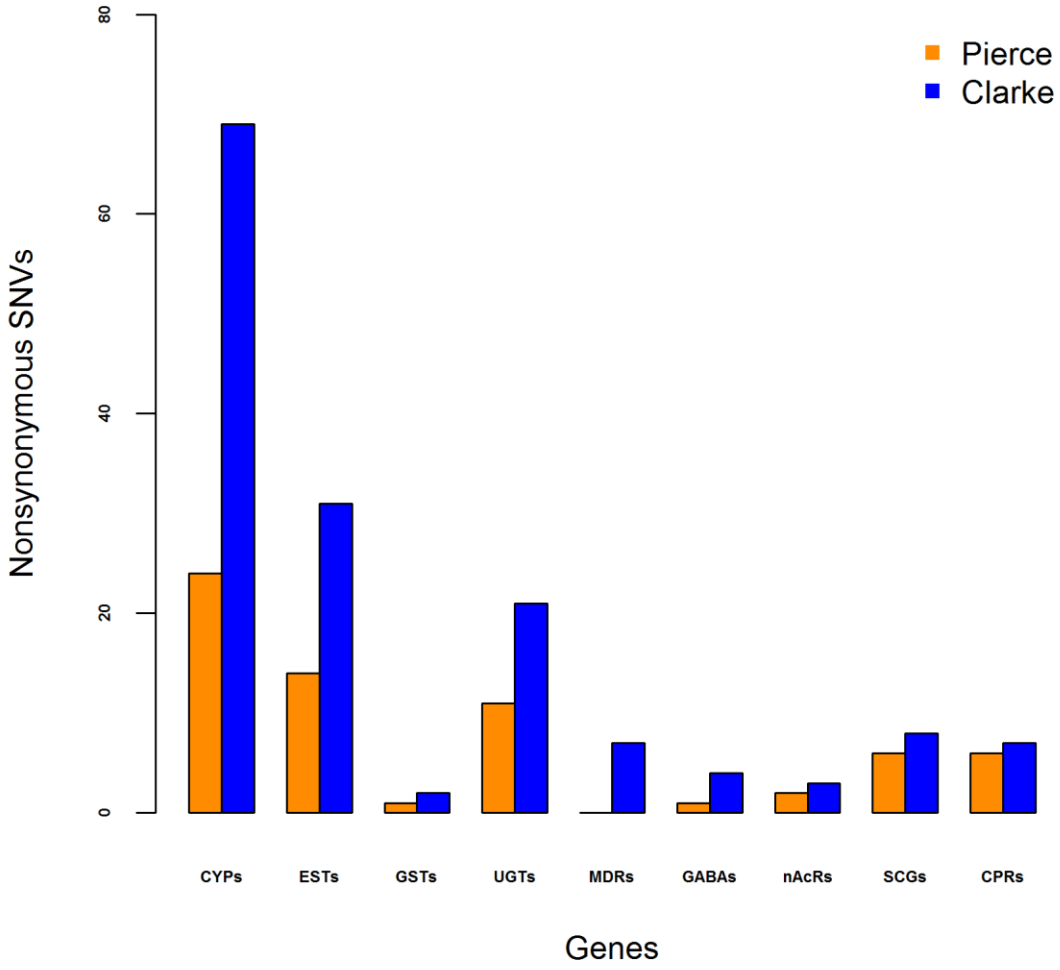

Pierce
Clarke
Nonsynonymous SNVs
0
20
40
60
80
CYPs
ESTs
GSTs
UGTs
MDRs
GABAs
nAcRs
SCGs
CPRs
Genes

Table 1. The median lethal concentration (LC_{50}) values of zeta-cypermethrin, spinosad, and malathion against Pierce and Clarke county *D. suzukii* population

Populations	N[a]	Slope (±SE)	χ^2	LC_{50} (ppm) (95% FL)[b]	LCR[c] at LC_{50} (95% CL)[d]
Zeta-cypermethrin					
Pierce	100	2.27 (±0.52)	0.97	1.50 (0.87-2.66)	3.07 (1.58-5.96)*
Clarke	100	3.18 (±0.85)	0.45	0.49 (0.30-0.80)	
Spinosad					
Pierce	100	1.66 (±0.33)	1.10	7.60 (4.072-14.32)	2.73 (1.29-5.81)*
Clarke	100	2.69 (±0.67)	0.44	2.78 (1.64-4.67)	
Malathion					
Pierce	70	2.37 (±0.56)	0.45	27.34 (15.84-47.41)	2.66 (1.41-5.05)*
Clarke	60	3.40 (±0.83)	3.34	10.25 (6.49-15.94)	

[a]N = number of adults assayed
[b]95% fiducial limits estimated using POLO (LeOra Software 1987)
[c]LCR, lethal concentration ratio = LC_{50} (Pierce population)/LC_{50} (Clarke population)
[d]95% confidence limits estimated using lethal concentration ratio significance test.
* LC_{50} of Pierce collected population significantly different from that of the Clarke population at $\alpha = 0.05$.
χ^2 = Chi-squared values indicate that the bioassay data fit the probit model and the distribution of responses is binomial.

Table 2.2. Significantly up-regulated detoxification genes with corrected P-value (Benjamini) < 0.05 and fold change ≥ 2

Significantly up-regulated detoxification genes						
Detox genes	**Zeta-cypermethrin**		**Spinosad**		**Malathion**	
	Pierce	**Clarke**	**Pierce**	**Clarke**	**Pierce**	**Clarke**
Cyp	*Cyp4e3*	*Cyp6w1*	*Cyp4e3*	*Cyp4e3*	*Cyp4e3*	None
	Cyp6a14		*Cyp49a1*		*Cyp304a1*	
	Cyp308a1		*Cyp4d14*		*Cyp28c1*	
	Cyp310a1		*Cyp4g15*		*Cyp6a14*	
	Cyp304a1		*Cyp9b2*			
	Cyp4p2		*Cyp28c1*			
	Cyp6a20		*Cyp318a1*			
	Cyp6a21		*Cyp301a1*			
	Cyp28c1		*Cyp6a14*			
	Cyp6w1		*Cyp4d8*			
	Cyp6d5		*Cyp308a1*			
	Cyp12d1		*Cyp6d5*			
	Cyp4d14		*Cyp313b1*			
	Cyp4g15		*Cyp304a1*			
	Cyp49a1		*Cyp6d5*			
	Cyp4s3					
	Cyp9b2					
	Cyp6d5					
	Cyp301a1					
Est	*DS10_00002107*	None	*DS10_00002107*	None	None	None
Gst	*GstD10*	None	*GstD10*	None	*GstD10*	None
	GstZ2		*GstZ2*		*GstZ2*	
	GstE14		*GstE14*			
			GstE10			
Ugt	*Ugt86Dd*	None	*Ugt36Ba*	None	None	None
	Ugt58Fa		*Ugt58Fa*			
	Ugt36Bb		*Ugt37b1*			
			Ugt35b			
			Ugt58Fa			
			Ugt36Bb			

Table 2.3. Significantly up-regulated insecticide targeted receptor and multi-drug resistant genes (*Mdr*) with corrected P-value (Benjamini) < 0.05 and fold change ≥ 2

Significantly up-regulated receptor encoding and *Mdr* genes						
Genes	**Zeta-cypermethrin**		**Spinosad**		**Malathion**	
	Pierce	**Clarke**	**Pierce**	**Clarke**	**Pierce**	**Clarke**
GABA	*GABA-B-R3*	None	*GABA-B-R3*	None	*GABA-B-R3*	None
	GABA-B-R1		*GABA-B-R1*		*GABA-B-R1*	
	Rdl		*Rdl*		*Rdl*	
	DS10_00005233		*DS10_00005233*		*DS10_00005876*	
	DS10_00005876		*DS10_00005876*		*GABA-B-R2*	
	GABA-B-R2		*GABA-B-R2*			
Mdr	*Mdr50*	None	*Mdr50*	None	None	None
	Mdr65		*Mdr65*			
nAcR	*nAcRalpha-34E*	None	*nAcRalpha-34E*	None	*nAcRbeta-64B*	None
	nAcRalpha-96Aa		*nAcRbeta-64B*		*nAcRbeta-96A*	
	nAcRbeta-64B		*nAcRbeta-96A*		*nAcRalpha-96Ab*	
	nAcRbeta-96A		*nAcRalpha-96Ab*			
	nAcRalpha-96Ab					
Sodium channel	*DS10_00000507*	None	*DS10_00000507*	None	*Teh3*	None
	para		*para*			
	Teh3		*Teh3*			
			ppk19			

Table 2.4. Significantly up-regulated cuticular protein encoding genes with corrected P-value (Benjamini) < 0.05 and fold change ≥ 2.

Significantly up-regulated cuticular protein encoding genes						
Defense genes	**Zeta-cypermethrin**		**Spinosad**		**Malathion**	
	Pierce	**Clarke**	**Pierce**	**Clarke**	**Pierce**	**Clarke**
Cpr	*Cpr92F*	None	*Cpr92F*	None	*Cpr92F*	None
	Cpr100A		*Cpr100A*		*Cpr100A*	
	Cpr47Ec		*Cpr47Ec*		*Cpr47Ec*	
	Cpr62Bc		*Cpr62Bc*		*Cpr62Bc*	
	Cpr49Ah		*Cpr49Ah*		*Cpr49Ah*	
	Cpr76Bd		*Cpr76Bd*		*Cpr76Bd*	
	Cpr49Ae		*Cpr49Ae*		*Cpr49Ae*	
	Cpr47Ee		*Cpr47Ee*		*Cpr47Ee*	
	Cpr50Cb		*Cpr50Cb*		*Cpr50Cb*	
	Cpr62Bb		*Cpr62Bb*		*Cpr62Bb*	
	Cpr49Ac		*Cpr49Ac*		*Cpr49Ac*	
	Cpr67B		*Cpr73D*		*Cpr73D*	
	Cpr73D					

Table 2.5: Predicted effect of Nonsynonymous (missense) SNPs on protein function in Clarke and Pierce populations.

Population	Gene_Id/name	SNPs[1]	Ref AA/Alt AA[2]	AA position[3]	SIFT score[4]	SIFT median	SIFT prediction
Clarke	DS10_00002643/Cyp12d1	scaffold2: 6068291_T/A	I/F	981	0.007	4.32	Deleterious (low conf)[5]
	DS10_00005250/Cyp4d1	Scaffold5:1948347_C/G	A/G	15	0.040	4.32	Deleterious (low conf)[5]
	DS10_00005633/Cyp9b2	scaffold6: 2068271_G/T	F/L	450	0.027	2.68	Deleterious
		scaffold6: 2068759_T/G	I/L	358	0.045	2.68	Deleterious
	DS10_00010146/ Cyp6d5	scaffold99: 37973_C/A	A/S	159	0.043	4.32	Deleterious (low conf)[5]
	DS10_00004134/ Est-6	scaffold3: 6208049_C/A	D/Y	389	0.005	3.39	Deleterious
	DS10_00008658/Ugt86Dd	Scaffold123: 382551_A/C	Q/P	840	0.015	4.32	Deleterious (low conf)[5]
	DS10_00012705/Ugt36Bb	Scaffold1724: 69775_C/T	T/I	72	0.009	2.74	Deleterious
		Scaffold1724: 70315_G/T	V/F	232	0.011	2.74	Deleterious
	DS10_00005769/Mdr49	scaffold6: 3398679_G/A	V/I	1225	0.027	3.56	Deleterious (low conf)[5]
		scaffold6: 3407477_G/A	V/I	2522	0.007	4.32	Deleterious (low conf)[5]
	DS10_00003821/Rdl	scaffold3: 1855481_T/A	D/V	309	0.008	2.41	Deleterious

	DS10_00008225/para	scaffold23: 585370_A/G	K/R	426	0.026	2.55	Deleterious
	DS10_00010216/rpk	scaffold305: 184502_C/G	S/R	379	0.033	2.66	Deleterious
	DS10_00002609/ Cpr49Ac	scaffold2: 5702262_G/A	E/K	310	0.037	4.32	Deleterious (low conf)[5]
Pierce	DS10_00005633/ Cyp9b2	scaffold6: 2068759_T/G	I/L	358	0.045	2.68	Deleterious
	DS10_00003335/ Cyp4p1	scaffold2: 11817769_C/T	A/V	489	0.013	4.32	Deleterious (low conf)[5]
	DS10_00005250/Cyp4d1	scaffold5:1948347_C/G	A/G	15	0.040	4.32	Deleterious
	DS10_00008658/Ugt86Dd	scaffold123: 382551_A/C	Q/P	840	0.015	4.32	Deleterious (low conf)[5]
	DS10_00008811/Ugt	Scaffold97: 25366_C/G	T/R	112	0.001	2.65	Deleterious
	DS10_00010216/rpk	scaffold305: 184502_C/G	S/R	379	0.033	2.66	Deleterious

SNPs[1]: a description of single nucleotide polymorphism Scaffold2 and 6068291 represent chromosome and position at which substitution occurred, respectively and T/A represents (T replaced by A). Ref AA/Alt AA[2]: I/F (Reference amino acid/ Alternate amino acid, i.e I substituted by F). AA position[3]: Amino acid position where substitution occurred. SIFT score[4]: Amino acid substitution predicted deleterious when score is ≤0.05. (low conf)[5]: A warning low confidence occurs when SIFT median score is greater than 3.5 because this indicates that the prediction is based on closely related sequences. SIFT4G was used to predict the effect of nonsynonymous SNPs on protein function.

Table 2.6. Predicted insertions and deletions in UTR and CDS regions of detoxification, insecticide receptor and cuticular proteins encoding genes in Clarke and Pierce populations.

Population	Gene_Id/name	Indels[1]	Region[2]	Var type[3]	AA_indel[4]
Clarke	DS10_00006248/ Cyp4s3	scaffold11: 772157_C/CA	UTR_3	Insertion	N/A
		scaffold11: 772668_A/ATT	UTR_3	Insertion	N/A
		scaffold11: 772718_AT/A	UTR_3	Deletion	N/A
		Scaffold11: 773151_GC/G	UTR_3	Deletion	N/A
	DS10_00012547/ GSTZ2	scaffold459: 69561_ GT/G	UTR_3	Deletion	N/A
	DS10_00002615/ Cpr49Ae	scaffold2: 5731703_CT/C	UTR_3	Deletion	N/A
	DS10_00003845/ Cpr67B	scaffold3: 2134602_C/CA	UTR_3	Insertion	N/A
	DS10_00004445/ Cpr62Bb	scaffold3: 9885475_AT/A	UTR_5	Deletion	N/A
	DS10_00009142/ Cpr100A	scaffold290: 138696_ T/TA	UTR_3	Insertion	N/A
	DS10_00012013/ Cpr92F	scaffold282: 132664_ TA/T	CDS	Frameshift deletion	p.Asp382fs[5]
Pierce	DS10_00006248/ Cyp4s3	scaffold11: 772157_C/CA	UTR_3	Insertion	N/A
	DS10_00002615/ Cpr49Ae	scaffold2: 5731982_ AT/A	UTR_3	Deletion	N/A
	DS10_00003845/ Cpr67B	scaffold3: 2134602_C/CA	UTR_3	Deletion	N/A

Indels[1]: a description of Insertions and deletions, Scaffold11 and 772157 represent chromosome and position after which insertion occurred, respectively and C/CA represents an A is inserted after C. Region[2]: The region where insertion or deletion occurred, UTR_3: 3'prime untranslated region, UTR5: 5' untranslated region and CDS: coding sequence. Var type[3]: Variant type. AA_indel[4]: amino acid position where insertion or deletion occurred. p.Asp382fs[5]: p.Asp382fs (HGVS nomenclature) denotes frameshifting change deleting Asp382 amino acid.

CHAPTER 3

TRANSCRIPTOMIC AND PROTEOMIC ANALYSES OF *ANOPHELES GAMBIAE* (AG55) CELLS: A POTENTIAL MODEL TO STUDY MOSQUITO-PATHOGEN INTERACTION

Ruchir., Robert L Bridger., Ujwal Bagal., Gang Hua., Lance Wells., Donald E. Champagne., and Michael J. Adang

To be submitted to *BMC Genomics*

Abstract

Anopheles gambiae is the predominant vector of malaria, the fourth largest cause of infant mortality, in sub- Saharan Africa. Furthermore, *A. gambiae* is also the primary vector of o'nyong-nyong virus. The complexity of handling *A. gambiae* and infectious pathogens has led to the use of *A. gambiae* cell lines, including Ag55 cells, as a potential model to study vector-pathogen interactions and immune responses. The utility of cell lines can be maximized if their detailed gene expression profile, and proteome are available.We provide detailed gene expression and proteome profiles of Ag55 cells. We further compared the gene expression profiles of Ag55 cells and blood fed female adults. Gene Ontology enrichment analysis of our transcriptomic data suggested that Ag55 cells have phagocytic properties, a hypothesis which we confirmed using confocal imaging. Transcriptomic data further backed by proteomic data suggest that Ag55 cells express hemocyte like properties, and are immune competent.

As Ag55 cells are immune competent and express hemocyte like properties they can be used as a model to study vector-pathogen immune response. Furthermore, the availability of transcriptomic and proteomic data of Ag55 cells will help researchers use and engineer the Ag55 cell line in an efficient way, for example by developing strategies to make it more suitable for studies of interactions with *Plasmodium* and other microbes.

3.1. Introduction

Anopheles gambiae is the predominant vector of malaria in Africa which is caused by an apicomplexan belonging to the genus *Plasmodium*. In 2015, 214 million cases of malaria were reported around the world out of which 88% of those cases were reported from Africa and was the fourth largest cause of infant mortality in sub- Saharan Africa (WHO 2015). *Anopheles gambiae* also vectors O'nyong-nyong arbovirus (Keene et al. 2004) and other viruses that pose emerging threats to human health.

The vectorial capacity of mosquitoes to transmit infectious pathogens depends on many pathogen-host interactions such as pathogen entry and development in the host, each of which is countered by the innate immune response of the host (Dimopoulos et al. 1997, Waldock et al. 2012). Environmental factors provide challenges for studying mosquito-infectious microbe interactions at the whole animal level. The handling of mosquitoes and infectious pathogens requires a range of skills and facilities including the ability to raise large numbers of mosquitoes, facilities for handling both the insect vector and the pathogen and expertise in vector and pathogen biology.

Mosquito cell lines have been used to investigate mosquito pathogen interactions (Sim and Dimopoulos 2010, Wilkins and Billingsley 2010, Berenice Juarez-Martinez et al. 2013), as models for deducing the host range of arboviruses (Lawrie et al. 2004, Kuno 2007), for isolating and characterizing mosquito-specific flaviviruses (Haddow et al. 2013, Hobson-Peters et al. 2013), for the potential *in vitro* development of *Plasmodium* ookinetes (Siden-Kiamos et al. 2000, Al-Olayan et al. 2002, Wilkins and Billingsley 2010), and for screening of insecticides (O'Neal et al. 2013). High throughput RNAi screens have made cell lines useful models for reverse genetic studies (Boutros et al.

2004a, Echeverri and Perrimon 2006, Cherry 2008). Furthermore, experiments with *Drosophila melanogaster* cell lines have been shown to correlate well with experiments using whole flies for genetic and developmental studies (Fehon et al. 1990, Cherbas et al. 2003, Hu et al. 2003, Cherbas and Gong 2014). *A. gambiae* cells lines have been used as models to study mosquito immune responses (Dimopoulos et al. 1997, Muller et al. 1999, Dimopoulos et al. 2002, Meister et al. 2005, Lombardo et al. 2013). Further, these cells have been shown to express immune factors upon microbial challenge, and perform complex immune tasks such as phagocytosis of beads and bacteria (Levashina et al. 2001, Lombardo et al. 2013). The discovery of *A. gambiae* densonucleovirus (AgDNV) is important as this virus replicates in adult mosquitoes and cultured *An. gambiae* MOS55 cells with minimal physiological effects on the host. It is likely that AgDNV can be engineered as a transducing virus (Ren et al. 2008).

Several cell lines have been derived from *A. gambiae* as it is the primary vector of malaria in sub-Saharan Africa. The Ag55 (Pudney et al. 1979), Sua1B and Sua4a-3B (Muller et al. 1999) cell lines have been derived from neonate first instar larvae. Further, a Sua5B cell line was derived by splitting Sua1 cell line (Rasgon et al. 2006). The Sua1B, Sua4a-3B, and Sua5B cell lines are considered to have hemocyte-like properties (Muller et al. 1999, Rasgon et al. 2006).

The Ag55 cells were tested as a model to study *Plasmodium* ookinete interaction (Wilkins and Billingsley 2010), and *Lysinibacillus sphaericus* Bin toxin mode of action (Hire et al. 2015). Ag55 cells are amenable to RNA inhibition-based knock-down of targeted mosquito genes (Konet et al. 2007, Smith and Linser 2009b). Towards the goal of enhancing the utility of Ag55 cells for pathogen interaction, genetic and immune

response studies, we used transcriptomic and proteomic approaches to provide gene expression and proteome profiles of Ag55 cells. Ag55 cells express genes encoding proteins involved in ookinete interaction, immune signaling pathways, and glycan biosynthesis. Considering that adult female mosquitoes encounter the human pathogens they transmit after blood feeding we compared the gene expression profiles of Ag55 cells with 5 day old female adult mosquitoes 3 hours post blood feeding (BF_Adult). We found phagocytosis biological process related transcripts with higher levels in Ag55 cells as compared to BF_Adult. This finding from transcriptomic data led us to establish that Ag55 cells have functional phagocytic properties. Furthemore, both transcriptomic and proteomic data identified hemocyte markers in Ag55 cells suggesting that Ag55 cells have hemocyte like characteristics.

3.2. Materials and Methods

3.2.1. Ag55 cell growth and maintenance

Ag55 cell line was originally derived from *Anopheles gambiae* neonate first instar larvae (Pudney et al. 1979) and gifted by Dr. Paul Linser, University of Florida, USA to Dr. Michael J Adang, University of Georgia. Ag55 cells were grown in Leibovitz's L-15 cell culture media (from Sigma) with additional 10% (v/v) fetal bovine serum (FBS, Atlanta Biologicals) and 1% (v/v) penicillin-streptomycin solution (10,000 U/ml and 10 mg/ml, respectively) (Sigma) in 25 cm^2 flasks (corning) at 28°C. The cells were grown in the flask until 80% confluency was reached after that 0.5×10^6 cells were transferred to new flask with fresh media for next (new) generation culture.

3.2.2. RNA extraction, sequencing and data analysis

RNA extraction, sequencing and data analysis was performed according to Hire et al. (Hire et al. 2015) In brief, RNA was extracted from Ag55 cells using TRIzol (Ambion). Approximately, 1×10^7 confluent Ag55 cells (1 flask of 25 cm2) were homogenized in TRIzol reagent (200 μl) and further processed for RNA extraction following the manufacturer's instructions (Ambion). Extraction of RNA from Ag55 cells was replicated three times. The purity and concentrations of RNA samples were determined using a NanoDrop spectrophotometer (N-1000) before dispatching samples for quality control and sequencing to the Georgia Genome Facility (GGF), University of Georgia.

At the GGF, pre-sequencing steps such as RNA integrity determination, Poly (A) enrichment of mRNA followed by cDNA synthesis and library preparation were performed. Sequencing was carried out using 100 bp paired-end Illumina HiSeq 2000 v3 platform.

Four replicate data sets of 5 day old whole female adult mosquitoes 3h post blood feeding (BF_Adult) were downloaded from NCBI GEO (Series GSE55453, SRA id : SRP039058). The raw sequenced data consisted of 7 sample sets (3 biological replicates of Ag55 cells and 4 biological replicates of BF_Adult). For RNA-seq analysis, Trapnell et al. (Trapnell et al. 2012b) protocol was followed. Based on the FastQC result, the pre-processing step of trimming the 3` and 5` end sequences was omitted. *A. gambiae* genome (*Anopheles-gambiae*-PEST_CHROMOSOMES_AgamP4.fa.gz) was downloaded from VectorBase and used as a reference for mapping the transcripts. Each dataset was independently mapped to the reference using Tophat v2.0.13 (Trapnell et al.

2009, Kim et al. 2013a), which uses Bowtie2 (Langmead et al. 2009a) as an aligner. Post alignment, Cufflinks v2.2.1 (Trapnell et al. 2010a) was used to produce a single file of assembled transfrags (transcripts.gtf) for each replicate (Trapnell et al. 2012b) and cufflinks v2.2.1 (cuffmerge) was used to merge all transcripts.gtf files from Ag55 cells and BF_Adult. Further, cufflinks v2.2.1 was used to estimate the expression values of the transcripts in FPKM (Fragments Per Kilobase per Million mapped reads) with the Cuffdiff 2 default geometric normalization.

3.2.3. Comparative analysis between Ag55 cells and BF_Adult

In-house written Perl scripts, Awk and Linux commands were used to perform comparative analyses between Ag55 cells and BF_Adult. Awk command was used to extract significantly differentially expressed genes from the Cuffdiff (gene_exp.diff and isoform_exp.diff) output files. Genes or isoforms are considered to be significantly differentially expressed in cuffdiff output files when q_value (FDR) <0.05. An in-house written Perl script was used to pull out information about genes of interest from cuffdiff output files. The expression levels (in FPKM) of genes of interest, which were grouped with other genes, were extracted from isoforms.fpkm_tracking or isoform_exp.diff files.

3.2.4. Gene Ontology (GO) enrichment analysis

GO Biological processes (BP) term enrichment analysis was performed using the DAVID v6.7 algorithm (Huang et al. 2009c, d) with the Benjamini-Hochberg correction for multiple testing and fold change. Only the transcripts which were significantly up-regulated (FDR<0.05) in Ag55 cells relative to BF_Adult with FPKM values ≥ 0.1 were selected for enrichment analysis.

3.2.5. Confocal imaging for determining the morphology of live Ag55 cells

Ag55 cells (1×10^6) were seeded in chambered coverslips (ibidi, GmbH) and after three hours the Ag55 cells were incubated with Hochest stain (ThermoFisher) and Nile red for 10 min. After 10 min incubation, live Ag55 cells were observed with a LSM 710 confocal microscope (Carl Zeiss).

3.2.6. Ag55 cell phagocytosis assay

Ag55 cells (0.5×10^6) were seeded in 4 chambered glass slides. After adhesion of the cells to the chambered glass slide, the cells were incubated with 10 nM Lysotracker for 2 hr at 28 ˚C. The cells were washed three times with cell media and incubated with FITC-labeled *E. coli* bioparticles (ThermoFisher scientific) at a ratio of 1:50 (cell/FITC-labeled *E. coli* bioparticles) for 35 min. After incubation, cells were washed three times with 1× PBS (phosphate buffer saline (137 mM NaCl, 2.7 mM KCl, 4.3 mM Na_2HPO_4, and1.47 mM KH_2PO_4)). Subsequently the cells were fixed with 4% paraformaldehyde (Electron Microscope grade,) for 20 min, washed times with 1× PBS, and incubated with 8.1 µM Hoechst stain (nuclear stain) for 10 min in the dark. The cells were then washed 3 times with 1× PBS and observed under a LSM 710 confocal microscope (Carl Zeiss).

3.2.7. Sample preparation for LC-MS/MS analyses

To maximize the identification of proteins three different sample preparation methods: cell surface biotinylation (CSB), total membrane preparation (TM), and whole cell (WC) digestion were used. All the experimental procedures were performed using two biological replicates.

3.2.7.1. Cell surface biotinylation

Cell surface biotinylation was performed according to Zhao et al. 2004 with slight modifications. The flasks containing approximately 5×10^7cells were tapped to detach. The cells were collected in an Eppendorf tube and centrifuged at 350 × g for 10 min. The cells were washed three times with ice cold 1X- PBS and cell surface biotinylation was performed using EZ-Link Sulfo-NHS-LC-Biotin (ThermoFisher) according to the instructions provided by the manufacturer (https://tools.thermofisher.com/content/sfs/manuals/MAN0011583_EZ_Sulfo_NHS_LC_Biotin_UG.pdf). After surface biotinylation, cells were collected by centrifugation at 350 × g for 10 min and re-suspended in ice-cold hypotonic buffer (10mM HEPES, pH 7.5, 1.5 mM $MgCl_2$, 10 mM KCl, and 1× protease inhibitor cocktail (complete Mini, EDTA-free, Roche). The cells were broken (30 passes) using a Teflon homogenizer (Wheaton) and kept on ice for 30 mins. Unbroken cells and nuclei were pelleted from the cell homogenate by centrifugation at 1000 × g for 10 min at 4°C. This step was repeated twice to generate postnuclear supernatant (PNS). Further, the KCl concentration of PNS was adjusted to 150 mM. A 300 µl aliquot of suspended streptavidin magnetic beads (BcMag, Bioclone Inc) prewashed 4 times with ice-cold hypotonic buffer) was added to the PNS. The resulting magnetic beads and PNS mixture was rotated at 4°C for 1 hr. The magnetic beads were collected using a magnetic stand (Promega), washed 3 times each by 1ml ice cold high salt wash (1 M KCl), and 1ml ice cold high pH wash (100 mM Na_2CO_3, pH 11.5) and the final wash was given using 1ml ice cold hypotonic buffer. The magnetic streptavidin beads bearing the enriched fraction was re-suspended in 100 µl ProteaseMax/Urea mix (0.2% ProteaseMax (Promega) and 8 M Urea) containing 100

mM DTT and 1× protease inhibitor cocktail, and incubated for 1hr at room temperature. After incubation magnetic stand (Promega) was used to separate magnetic beads and supernatant. Proteins were precipitated from supernatant by cold acetone according to ThermoFisher scientific protocol (TR0049).

3.2.7.2. Total membrane preparation

For total membrane preparation approximately 5×10^7cells were collected by centrifugation at 350 × g for 10 min and re-suspended in ice-cold hypotonic buffer (10 mM HEPES, pH 7.5, 1.5 mM $MgCl_2$, 10 mM KCl, and 1× protease inhibitor cocktail. The cells were broken by teflon homogenizer (30 passes, homogenizer from Wheaton) and kept on ice for 30 min. Unbroken cells and nuclei were pelleted from the cell homogenate by centrifugation at 1000g for 10 min at 4°C. This step was repeated twice to generate PNS. The PNS was further subjected to ultra-centrifugation at 100,000 × g for 90 min to obtain total membrane. Total membrane preparation was re-suspended in 100 µl ProteaseMax surfactant/Urea mix (0.2% ProteaseMax (Promega) and 8 M Urea) and 1× protease inhibitor cocktail and incubated for 1hr at room temperature. Proteins were precipitated from total membrane suspension using cold acetone.

3.2.7.3. Whole Cell digestion

For whole cell digestion approximately 1×10^7cells were collected by centrifugation at 350 × g for 10 min and re-suspended in 500 µl of ProteaseMax/Urea mix (0.2% ProteaseMax (Promega) and 8 M Urea) and 1× protease inhibitor cocktail, and further incubated for 1 hr at room temperature. After incubation cells were subjected to centrifugation at 16,000 × g for 10 min at room temperature. The supernatant was collected, and proteins were precipitated from supernatant using cold acetone.

3.2.8. Protein in-solution digestion

Precipitated proteins from CSB, TM, and WC methods were re-suspended in 15 µl of Urea followed by addition of 20 µl 0.2% ProteaseMax acid-labile surfactant. Tubes were vortexed gently in intervals until the pellet got dissolved (~ 30 min). Concentrations of the total proteins from three methods were analyzed using Bradford protein assay (Bradford 1976). Approximately 30 µg proteins from each method was used for in-solution digestion. Proteins were reduced with DTT (dithiothreitol) and carbamidomethylated with iodoacetamide (IAA) according to the ProteaseMax technical bulletin (ProteaseMax surfactant, TB373, Promega). Tryptic digestion was perfomed over-night using mass-spec grade Trypsin (promega) at (1:50, Trypsin: Protein) ratio. Tryptic digestion was stopped by adding 90 % formic acid to obtain a final in-digestion concentration of 5% formic acid. Peptides were cleaned using Pierce C18 reverse phase columns.

3.2.9. LC-MS/MS analysis

Peptides were dried down and re-suspended with 39uL of buffer A (0.1% formic acid) and 1uL of buffer B (80% acetonitrile and 0.1% formic acid). The samples were spun through a 0.2 µm filter (Nanosep, Pall Corp) before being loaded into an autosampler tube and racked into an Ultimate 3000 LC System (Thermo Scientific – Dionex).

LC-MS/MS analysis was performed on an Orbitrap Fusion Tribrid Mass Spectrometer (Thermo Scientific) utilizing a nanospray ionization source and the Acclaim PepMap RSLC analytical column (Thermo Scientific - Dionex). For each sample, 5 µl was injected and separated via a 180-minute gradient of increasing buffer B

at a flow rate of 300 nL per minute. An instrument method was used to collect a full mass spectrum every three seconds and continuously fragment the most intense precursor ions with 38% collision-induced dissociation (CID) and record the resulting MS/MS spectra. Dynamic exclusion was utilized to exclude precursors ions from selection process for 30 seconds following a second selection within a 10 second window.

Database searches were performed on all raw data files using the SEAQUEST (Proteome Discoverer 1.4, Thermo Fisher Scientific) search algorithm with the following parameters: 10-ppm tolerance for monoisotopic precursor mass and 0.3 Da tolerance for fragment masses; Trypsin (fully specific – C terminus) specified as enzyme; a maximum of 2 missed cleavage sites and a maximum of four modifications per peptide. All searches were performed against Uniprot culicidae database (updated April 14, 2014) allowing for oxidation (+15.995 Da) at methionine residues and carbamidomethylation (+57.021 Da) on cysteine residues. Search results were filtered at 5% peptide-level FDR and 1% protein-level FDR. Proteins identified with at least 2 unique peptides were selected for further down-stream analyses. Phobius (Kall et al. 2004, Kaell et al. 2007) analyses were performed on the identified proteins to predict transmembrane domains and signal peptide.

3.3. Results

3.3.1. Comparing the transcriptome profile of Ag55cells with blood-fed adult mosquitoes

The transcriptome of Ag55 cells was defined by DNA sequencing of cDNA prepared from RNA isolated from three biological replicates of Ag55 cell replicates using the Illumina HighSeq2000 platform. The total paired-end reads obtained per sample from

the Illumina HighSeq2000 platform ranged from 16.29 million to 23.11 million, of which 11.76 million to 16.99 million reads had aligned pairs.

To reveal possible cell-type specific functions, the Ag55 gene expression profile was compared with a similar transcriptomic analysis of 5day old female mosquito adult whole bodies harvested 3 hrs after blood feeding (BF_Adult). The BF_Adult data sets, which were derived from 4 biological replicates, were downloaded from NCBI GEO (Series GSE55453, SRA id : SRP039058) (Vannini et al. 2014). Mapping of Ag55 cell transcripts to an *A. gambiae* reference genome detected the expression of 10, 320 genes. The expression values of transcripts from Ag55 cell BF_adults were calculated as Fragments Per Kilobase per Million mapped (FPKM) reads. About twenty-five percent (2, 307 transcripts) of the expressed genes in Ag55 cells had higher levels in the cultured cells relative to BF_Adult mosquitoes using a Benjamini corrected FDR <0.05. Gene ontology (GO) biological processes (BP) enrichment analysis was performed using DAVID functional annotation clustering [1, 2], on the significantly enriched transcripts of Ag55 cells relative to BF_Adult mosquitoes. Figure 1 shows enriched BP GO terms and the number of genes expressed for each term. Vesicle mediated transport, protein catabolic processes, membrane organization and phagocytosis were enriched GO BP terms that suggested specialized cellular functions.

3.3.2. Ag55 cells have a morphology and gene expression profile indicative of hemocyte-like cells

At a visual level confocal imaging of non-confluent live Ag55 cells showed leaf-shaped cells with pseudopodia-like structures (Figure 2). Insect hemocytes use pseudopodia to contact and surround the foreign particles during initial stages of

phagocytosis (Kwon et al. 2014). Furthermore, *A. gambiae* hemocytes have a profile of enriched gene transcripts (Baton et al. 2009) associated with A. gambiae hemocytes. Baton et al. (Baton et al. 2009) identified 279 genes having significantly higher expression levels in *A. gambiae* adult mosquito hemocytes relative to whole adults. Of the 279 gene transcripts that were designated by ENSANGT and Ensembl identifications in (Baton et al. 2009), we were able to update 215 genes to their current AGAP designations. Of the 215 *A. gambiae* hemocyte-enriched genes, 172 were expressed in Ag55 cells (Additional file 1: TableS1-S2). We also analyzed the expression of other reported hemocyte marker genes (Muller et al. 1999, Pinto et al. 2009, Lombardo et al. 2013, Evans et al. 2014) (Table 1, Additional file 1: Table S3-S4). We found hemocyte markers: *nimrod* family genes (*eater, draper*) (Estevez-Lao and Hillyer 2014), *HPX4/IRID3* (ortholog of peroxidasin) (Nelson et al. 1994, Garver et al. 2008), *PSMD3*/*DoxA2* and *PDGF-VEGF receptor* gene ortholog expressed at transcript level (Table 1) in Ag55 cells. Furthermore, we also identified the above-mentioned hemocyte markers at protein level (Table 2) from proteomic (LC-MS/MS) analyses of Ag55 cells.

The peroxidasin and nimrod family genes have been implicated in phagocytosis [38-40]. In addition to peroxidasin, nimrod family genes we also analyzed the expression levels of AGAP006769, FBN9, AGAP009459, AGAP003879, AGAP000095 and AGAP008500 genes as knock-down of these genes have been shown to significantly increase or decrease phagocytosis of *E. coli* bioparticles in A. gambiae Sua5.1 cultured cells [23]. All 6 genes were expressed in Ag55 cells (Figure 3, Additional file 1: Table S5).

The enrichment of the phagocytosis BP term and presence of hemocyte markers at both transcript and protein levels prompted us to test Ag55 cells for phagocytic properties. Ag55 cells were exposed to FITC-labeled *E. coli* bioparticles and then the locations of fluorescent particles imaged by confocal microscopy. Figure 4 shows the clustering of *E. coli* bioparticles in Ag55 cells (Figure 4, Additional file 2: Figures S1-S2). Before Ag55 cells were exposed to FITC-*E. coli* they were treated with Lysotracker, an acidophilic dye that localizes to late endosomes and lysosomes (Via et al. 1998). The FITC-labeled *E. coli* bioparticles internalized and co-localized with Lysotracker deep red in Ag55 cells.

3.3.3. Ag55 cells express genes involved in immune response pathways

Toll, JAK-STAT, and the IMD pathways are the three major immune pathways in insects. These three pathways have been implicated in the defense against at least one species of *Plasmodium* (Meister et al. 2005, Frolet et al. 2006, Garver et al. 2009, Gupta et al. 2009, Jaramillo-Gutierrez et al. 2010). Similarly, IMD and JAK-STAT have also been implicated in defense against o'nyong-nyong virus (Carissimo et al. 2015). We analyzed the expression levels of the genes and their isoforms encoding proteins involved in IMD, JAK-STAT, and the Toll pathways. We found *PGRPLC*, *IMD*, *FADD*, *CASPL1*, *MAP3K7IP1*, *IKK1*, *Faf1* and *REL2* genes encoding proteins involved in IMD pathway, expressed in Ag55 cells (Table 3, Additional file 3: Tables S6-S7).

Similarly, genes such as *Dome, JAK*, *STAT*, *CBP*, *PIAS*, *SOCS*, and *c-Myc* involved in the JAK-STAT pathway were expressed in Ag55 cells (Table 3, Additional file 3: Tables S8-S9).

Activation of the Toll pathway has been shown to initiate a strong immune response that greatly reduces *Plasmodium berghei* infection in *Anopheles gambiae* (Frolet et al. 2006, Garver et al. 2009). The genes involved in the Toll signaling pathway were found to be expressed in Ag55 cells (Table 3, Additional file 3: Tables S10-S11). The *TOLL11, TOLL*5*A* genes were expressed in Ag55 cells with FPKM values ≥1. Furthermore, *TOLL10* gene was expressed with an FPKM value ≥0.1. Other *TOLL* genes such as *TOLL1A, TOLL6, TOLL7, TOLL8, TOLL9* were expressed with FPKM values <0.1. *TOLL5B* and *TOLL1B* genes were not detected in Ag55 cells (Additional file 3: Tables S10-S11).

Furthermore, we also examined the expression levels of genes encoding immune-related proteins including *FBNs*, C-type lectins *(CTLs)*, beta-glucan binding proteins (*GNBPs*) , scavenger receptors (*SCRBQs*), and leucine-rich repeat containing proteins (*LRIM/APLs*) as proteins belonging to these groups have been shown to play important roles during *Plasmodium* infection (Parish et al. 2013). Most of the *FBNs, CTLs, GNBPs, SCRBQs* and *LRIM/APLs* genes were expressed with FPKM values ≥0.1 (Additional file 3: Tables S10-S13) in Ag55 cells.

3.3.4. Ag55 cells express known ookinete interacting proteins

Wilkins and Billingsley (Wilkins and Billingsley 2010) have shown that *Plasmodium berghei* binds weakly to the Ag55 cells and suggested that the binding is non-carbohydrate mediated. The findings of Wilkins and Billingsley led us to analyze the expression levels of the known ookinete interacting protein genes in Ag55 cells. These genes encode proteins that have experimentally been shown to facilitate *Plasmodium* ookinete invasion of the midgut of *A. gambiae* adult female (Kotsyfakis et

al. 2005, Dinglasan et al. 2007a, Dinglasan et al. 2007b, Lavazec et al. 2007, Parish et al. 2013, Mathias et al. 2014, Vega-Rodriguez et al. 2014, Zhang et al. 2015). Out of these nine (ookinete interacting protein) genes, eight genes were expressed in Ag55 cells (Figure 5, Additional file 4: Tables S14-S15). *FREP1* gene transcripts were not detected in Ag55 cells (Figure 6, Additional file 4: Tables S14-S15). Furthermore, Baton et al. (Baton et al. 2009) microarray data showed that the genes encoding ookinete interacting proteins were also expressed in *A. gambiae* adult hemocytes with the exception of *ANXB10C*, and *AgSGU*.

3.3.5. Genes encoding enzymes involved in glycosylation and sulfation process expressed in Ag55 cells

Several studies have shown the critical role of carbohydrates in *Plasmodium* ookinete attachment and invading process and the ookinete-oocyst transition stages within the mosquito vector (Zieler et al. 1999, Zieler et al. 2000, Wilkins and Billingsley 2001, Dinglasan et al. 2005, Dinglasan et al. 2007a). To evaluate the suitability of Ag55 cells for studies of *Plasmodium* interaction, expression level of genes and their isoforms associated with glycosylation processes were examined. We also compared the expression levels of genes encoding enzymes involved in carbohydrate biosynthesis as Wilkins and Billingsley 2010 suggested that the Ag55 cells have different glycosylation pattern than that of the adult mosquito midgut. We found genes encoding enzymes involved in N-linked glycosylation processes expressed in Ag55 cells (Table 4, Additional file 5: Tables S16-S22). The expression level of genes involved in N-linked glycan precursor biosynthesis, in Ag55 cells were comparable to BF_Adult with the exceptions of AGAP003928 (*ALG8*), AGAP011324 (*ALG11*), and AGAP007168 (*ALG3*)

which had significantly higher expression levels in BF_Adult) and AGAP003551 (*ALG1*) which was enriched in Ag55 cells (Additional file 5: Tables S16-S17). Most of the genes involved in N-glycan precursor trimming and core structure formation in *A. gambiae* had significantly higher expression levels in Ag55 cells with no significant higher expressions observed in BF_Adult (Additional file 5: Tables S18-S19). In contrast most of the genes involved in N-glycan core structure elongation had significantly higher expression levels in BF_Adult but none of these genes had significantly higher expressions in Ag55 cells (Additional file 5: Tables S20-S21). The genes involved in N-linked glycan biosynthesis (precursor biosynthesis, trimming and core structure formation, and core structure elongation) were expressed in Ag55 cells and BF_Adult with the exception of AGAP006903 (*ST6GAL*) gene, which was not detected in Ag55 cells (Additional file 5: Tables S20-S21).

We found two isoforms of AGAP007080 (hexoseaminidase) gene, homolog of *fdl*, expressed in Ag55 cells and BF_Adult. The expression level of one of the isoforms (TCONS_00002552) was significantly higher in Ag55 cells (FPKM values of 45.12) than in BF_Adult (FPKM=6.68). On the other hand, isoform (TCONS_00006016) had significantly higher expression level in BF_Adult (FPKM values of 6.53) relative to Ag55 cells (FPKM value 2.45) (Additional file 5: Tables S20). The overall expression of AGAP007080 gene was higher in Ag55 cells (Additional file 5: Tables S22).

O-glycans have also been shown to mediate *Plasmodium* interaction with the midgut (Billingsley 1994, Dinglasan et al. 2005). Most of the genes in *A. gambiae* known to participate in mucin type O-glycan biosynthesis were expressed with FPKM values ≥0.1 (Table 4, Additional file 5: Table S23) in Ag55 cells as well as BF_Adult. There

were more number of genes (involved in O-glycan biosynthesis) which had significantly higher expression levels in Ag55 cells relative to BF_Adult (Additional file 5: Tables S23-24).

Chondroitin sulfate glycosaminoglycans, present on the midgut surface of *A. gambiae*, are putative ligands for *Plasmodium falciparum* ookinete binding (Dinglasan et al. 2007a, Mathias et al. 2013). We analyzed the expression level of genes encoding enzymes involved in chondroitin sulfate biosynthesis in Ag55 cells and BF_Adult. We found most of the chondroitin sulfate biosynthesis genes having significantly higher expression levels in BF_Adult (Additional file 5: Tables S25-S26). All the known genes associated with chondroitin sulfate biosynthesis in *A. gambiae* were expressed with a FPKM values ≥0.1 in Ag55 cells, with the exception of AGAP001050 (*CHPF*) gene which was expressed with FPKM value (0.026) (Table 4, Additional file 5: Tables S25-S26).

3.3.6. Proteomic analysis identified hemocyte markers, ookinete interacting proteins in Ag55 cells

We performed LC-MS/MS (proteomic) analysis to provide a detailed proteome of Ag55 cells and identify characteristic properties of Ag55 cells at the protein level. We used three different methods (whole cell (WC), total membrane preparation (TM), and cell surface biotinylation (CSB)) for Ag55 cells sample preparations. These three different methods were used to identify as many proteins as possible including membrane proteins from LC-MS/MS runs. Identification of membrane proteins are important as they carry out many critical functions, including cell-cell interactions, initiation of signaling cascades, vesicle trafficking, protein translocation, and transport of ions and solutes

(Elschenbroich et al. 2010, Weekes et al. 2010). Furthermore, ookinete interacting proteins and large repertoire of proteins involved in identifying and phagocytosing foreign particles and microbes are membrane proteins (Parish et al. 2011, Evans et al. 2014).

We identified 1260, 1087 and 696 proteins from WC, TM and CSB methods, respectively (Table 3). Out of the three methods, TM method identified the highest percentage of integral membrane proteins (16.65%) (Table 5). In total (from all three methods), we identified 1682 unique proteins (Additional file 6: Table S27).

In our proteomics data we further analyzed for integral membrane proteins as many hemocyte markers and proteins involved in phagocytosis are integral to membrane (i.e. they have transmembrane domain) (Estevez-Lao and Hillyer 2014, Evans et al. 2014). There were 225 unique proteins predicted to have transmembrane domains (Additional file 7: Table S28). We identified several integral membrane proteins that are known hemocyte markers and involved in phagocytosis (Estevez-Lao and Hillyer 2014, Evans et al. 2014) in our Ag55 cell proteomics data (Table 2). In addition, we also identified peroxidasin and misshapen (not predicted to have transmembrane domain) in our proteomics data (Table 2). The peroxidasin and misshapen proteins are also known hemocyte markers (Marcu et al. 2011, Evans et al. 2014). We further analyzed whether or not the hemocyte enriched transcripts (Baton et al. 2009) are expressed at protein level in Ag55 cells. We identified 32 proteins encoded by their respective hemocyte enriched transcripts (Baton et al. 2009) (Additional file 8: Table S29).

Our proteomics data further revealed the presence of ookinete interacting proteins: Annexin 10B (AGAP003721), Annexin 10C (AGAP003722), and Class B

Scavenger Receptor (CD36 domain, AGAP010133) in Ag55 cells (Additional file 6: Table S27).

3.3.7. Comparing the gene expression values with proteins identified

We further compared the 1682 unique proteins identified from the LC-MS/MS analysis to their respective genes expression values in FPKM from RNA-seq data. A majority of the proteins identified has genes expression values ≥1 but we also identified a few proteins with gene expression values < 0.1 (Additional file 9: Figure S3, Additional file 10: Figure S4).

3.4. Discussion

Our comparative transcriptomic analysis showed an enrichment of phagocytosis biological process GO term, which led us to hypothesize that Ag55 cells have phagocytic properties. This hypothesis was confirmed using confocal imaging which showed that Ag55 cells were phagocytosing FITC labeled *E. coli* bioparticles. We also identified most of the known hemocyte enriched genes (Baton et al. 2009) at transcript and some at protein levels using RNA-seq and LC-MS/MS analyses, respectively. Confocal images showed the presence of pseudopodia like structure on Ag55 cells. Pseudopodia have been shown to be present on the plasma membrane of insect hemocytes (Kwon et al. 2014). Hemocytes use phagocytosis, nodulation and encapsulation as defense response against invading microbes (Lavine and Strand 2002, Browne et al. 2013). Our proteomics (LC-MS/MS) data further identified hemocyte marker proteins (Table 2). These proteins are the known markers of hemocytes in *D. melanogaster* (Nelson et al. 1994, Garver et al. 2008, Bretscher et al. 2015) and *A. gambiae* (Castillo et al. 2006).

In *D. melanogaster nimrod* family genes have been divided into 3 sub-classes *nimrodA, B,and C*. The *nimrodA* genes encode for 1 NIM domain, several EGF domains and transmembrane domain. In contrast *nimrodB* genes encode for several NIM repeats and no transmembrane domain. Genes in *nimrodC* subclass encode multiple NIM repeats and a transmembrane domain (Estevez-Lao and Hillyer 2014). The nimrod family genes (*draper,* and *eater orthologs*), identified at transcript and protein level in Ag55 cells, have been shown to be expressed in hemocyte and are involved in mosquito antibacterial immune response (Estevez-Lao and Hillyer 2014). In *D. melanogaster draper*, *nimC1* and, *eater* participate in bacterial binding which in turn leads to phagocytosis by hemocytes (Estevez-Lao and Hillyer 2014). Estevez-Lao and Hillyer have characterized AGAP009762 as ortholog of nimrodB2 based on phylogenetic analysis and their bioinformatics analysis of full length sequence of AGAP009762 failing to predict any transmembrane domain. But, when we used the full length sequence of AGAP009762 (uniprot Id Q7PM27) and predicted for transmembrane domain using Phobius, TMHMM and SMART we found the presence of transmembrane domain (from amino acid positions 1197-1217) which is in contrast to Estevez-Lao and Hillyer result. Further, HomoloGene (NCBI) characterized AGAP009762 to be the ortholog of *eater*. We also identified several EGF and EGF-like domains suggesting that AGAP009762 belongs to nimrodC subclass to which eater belongs.

Similarly peroxidasin has been shown to be expressed in maturing plasmatocyte (Marcu et al. 2011, Evans et al. 2014) and involved in extracellular matrix consolidation, phagocytosis and defense (Nelson et al. 1994). Another protein Class B Scavenger Receptor having CD36 domain (hemocyte marker) recognizes dying cells (Franc et al.

1996). Further, PDGF-VEGF receptor have been show to control hemocyte survival in *D. melanogaster* (Brückner et al. 2004). The presence of hemocyte markers at transcript and protein level, pseudopodia like structure, and phagocytosis property suggest that Ag55 cells have hemocyte like properties.

We also found intact machinery of the IMD, JAK-STAT and Toll signaling pathways in Ag55 cells. IMD signaling pathway has emerged as the most effective pathway for controlling the malarial parasite *P. falciparum* in *A. gambiae* (Garver et al. 2012). Similarly *Anopheles* species use JAK-STAT and Toll signaling pathways as defense mechanisms against *Plasmodium* development (Meister et al. 2005, Frolet et al. 2006, Garver et al. 2009, Gupta et al. 2009, Jaramillo-Gutierrez et al. 2010). Most of the known *TOLL genes* of *A. gambiae* were expressed in Ag55 cells of which the *TOLL11* gene was significantly up-regulated in Ag55 cells. RNAi mediated knock-down studies have revealed that *TOLL 11* mediates significant protection against *P. falciparum* oocyst infection (Redmond et al. 2015). *TOLL5* mediates anti-bacterial and anti-fungal functions in *A. gambiae* (Christophides et al. 2002) and *A. aegypti* (Shin et al. 2006). Our results suggest that Ag55 cells can be used as a useful model to study immune related response against microbes.

Wilkins and Billingsley 2010 have shown that *P. berghei* binds weakly to the Ag55 cells, and exogenous carbohydrates do not affect this binding, which suggests a non-carbohydrate mediated binding. We found all the known ookinete interacting protein encoding genes expressed in Ag55 cells with the exception of *FREP1*. Our proteomic data also identified the presence of three ookinete interacting proteins (annexin 10B, annexin 10C and Class B Scavenger Receptor (CD36 domain)) further validating the

expression of ookinete interacting proteins in Ag55 cells. These proteins likely account for the weak binding of *P. berghei* ookinetes to these cells. Furthermore, Baton et al. (Baton et al. 2009) have identified these ookinete interacting proteins to be expressed in *A. gambiae* adult hemocytes. For future studies, it would be interesting to analyze the roles of ookinete interacting proteins in malarial parasite- hemocyte interactions.

Several studies have highlighted the importance of carbohydrates in Plasmodium ookinete interaction with the adult female midgut and invasion, process (Zieler et al. 1999, Wilkins and Billingsley 2001, Dinglasan et al. 2005, Dinglasan et al. 2007a). Ookinete interaction with Ag55 cells is not carbohydrate dependent, which led Wilkins and Billingsley 2010 to suggest that these cells differ from the mosquito midgut in glycosylation, though the specific differences were not identified. Accordingly, we analyzed the expression levels of genes encoding enzymes involved in N-linked, mucin type O-linked and chondroitin sulfate glycosylation process.

We found differences in expression pattern (i.e. levels) of genes involved in N-glycan biosynthesis in Ag55 cells when compared to blood fed female adults. We found most of the genes, encoding enzymes, involved in N-glycan precursor trimming and core structure formation having higher expression levels in Ag55 cells relative to BF_Adult. In contrast most of the genes, encoding enzymes, involved in N-glycan core structure elongation have higher expression levels in blood fed female adults (BF_Adult). These trends suggest that Ag55 cells may be not efficient in synthesizing branched N-glycan structure. Furthermore, we found higher expression level of hexoseaminidase (AGAP007080), an ortholog of *fdl* gene. The *fdl* gene in *Drosophila melanogaster* encodes for an N-acetylglucosaminidase enzyme involved in the removal of terminal N-

acetylglucoseamine (GlcNAc) which leads to the formation of paucimannosidic N-glycan structure (Geisler et al. 2008). The lack of terminal GlcNAc residue was suggested to be the cause for poor binding of ookinete with Ag55 cells (Wilkins and Billingsley 2010). Thus our transcriptomic data further supports the hypothesis made by Wilkins and Billingsley (Wilkins and Billingsley 2010) that the differences in the glycosylation pattern and the apparent lack of abundant N-acetylglucoseamine (GlcNAc) and N-acetylgalactoseamine (GalNAc) in Ag55 cell line compared to that of the mosquito midgut epithelial cell could be the reason for poor binding of the *P. berghei* ookinete.

Chondroitin sulfate (CS) present on the apical surfaces of mosquito midgut are require for *P. falciparum* and *P. berghei* invasion (Dinglasan et al. 2007a, Mathias et al. 2013). We observed weaker expression levels of genes involved in CS biosynthesis pathway when compared to blood fed female adults. Further suggesting the weak interaction of *P. berghei* ookinete with Ag55 cells.

3.5. Conclusion

We have provided the gene expression profile and proteome of Ag55 cells. Enrichment analysis of our transcriptomic data suggested that Ag55 cells have phagocytic properties, a hypothesis which we confirmed using confocal imaging. Further proteomic data showed the presence of hemocyte markers. Altogether, Ag55 cells express hemocyte like properties, although the ultimate tissue origin of this cell line is still uncertain. The availability of transcriptomic and proteomic data of Ag55 cells will help researchers use and engineer Ag55 cell line in an efficient way, for example by developing strategies to make it suitable for *Plasmodium* and other microbes interaction studies.

Availability of supporting data

The data discussed in the submission have been submitted to NCBI's Gene Expression Omnibus (GEO) and are accessible through GEO series accession number GSE85643 (http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE85643).

3.6. References

Al-Olayan, E. M., A. L. Beetsma, G. A. Butcher, R. E. Sinden, and H. Hurd. 2002. Complete development of mosquito phases of the malaria parasite in vitro. Science 295: 677-679.

Baton, L. A., A. Robertson, E. Warr, M. R. Strand, and G. Dimopoulos. 2009. Genome-wide transcriptomic profiling of Anopheles gambiae hemocytes reveals pathogen-specific signatures upon bacterial challenge and Plasmodium berghei infection. BMC Genomics 10.

Berenice Juarez-Martinez, A., T. Olivia Vega-Almeida, M. Salas-Benito, M. Garcia-Espitia, M. De Nova-Ocampo, R. Maria del Angel, and J. Santiago Salas-Benito. 2013. Detection and sequencing of defective viral genomes in C6/36 cells persistently infected with dengue virus 2. Arch Virol 158: 583-599.

Billingsley, P. F. 1994.. Approaches to vector control - new and trusted .2. molecular targets in the insect midgut. Trans R Soc Trop Med Hyg 88(2):136-140.

Boutros, M., A. A. Kiger, S. Armknecht, K. Kerr, M. Hild, B. Koch, S. A. Haas, R. Paro, N. Perrimon, and C. Heidelberg Fly Array. 2004. Genome-wide RNAi analysis of growth and viability in Drosophila cells. Science 303: 832-835.

Bradford, M. M..1976. Rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 72: 248–254.

Bretscher, A. J., V. Honti, O. Binggeli, O. Burri, M. Poidevin, E. Kurucz, J. Zsamboki, I. Ando, and B. Lemaitre. 2015. The Nimrod transmembrane receptor Eater is required for hemocyte attachment to the sessile compartment in Drosophila melanogaster. Biol Open 4: 355-363.

Browne, N., M. Heelan, and K. Kavanagh. 2013. An analysis of the structural and functional similarities of insect hemocytes and mammalian phagocytes. Virulence 4: 597-603.

Brückner, K., L. Kockel, P. Duchek, C. M. Luque, P. Rørth, and N. Perrimon. 2004. The PDGF/VEGF receptor controls blood cell survival in Drosophila. Dev Cell 7: 73-84.

Carissimo, G., E. Pondeville, M. McFarlane, I. Dietrich, C. Mitri, E. Bischoff, C. Antoniewski, C. Bourgouin, A.-B. Failloux, A. Kohl, and K. D. Vernick. 2015. Antiviral immunity of Anopheles gambiae is highly compartmentalized, with distinct roles for RNA interference and gut microbiota. Proc Natl Acad Sci US . 112: E176-E185.

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

Cherbas, L., and L. Gong. 2014. Cell lines. Methods 68: 74-81.

Cherbas, L., X. Hu, I. Zhimulev, E. Belyaeva, and P. Cherbas. 2003. EcR isoforms in Drosophila: testing tissue-specific requirements by targeted blockade and rescue. Development 130: 271-284.

Cherry, S. 2008. Genomic RNAi screening in Drosophila S2 cells: what have we learned about host-pathogen interactions? Curr Opin Microbiol 11: 262-270.

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

Dimopoulos, G., A. Richman, H. M. Muller, and F. C. Kafatos. 1997. Molecular immune responses of the mosquito Anopheles gambiae to bacteria and malaria parasites. Proc Natl Acad Sci USA 94: 11508-11513.

Dimopoulos, G., G. K. Christophides, S. Meister, J. Schultz, K. P. White, C. Barillas-Mury, and F. C. Kafatos. 2002. Genome expression analysis of Anopheles gambiae: Responses to injury, bacterial challenge, and malaria infection. Proc Natl Acad Sci USA 99: 8814-8819.

Dinglasan, R. R., J. G. Valenzuela, and A. F. Azad. 2005. Sugar epitopes as potential universal disease transmission blocking targets. Insect Biochem Mol Biol 35: 1-10.

Dinglasan, R. R., A. Alaganan, A. K. Ghosh, A. Saito, T. H. van Kuppevelt, and M. Jacobs-Lorena. 2007a. Plasmodium faiciparum ookinetes require mosquito midgut chondroitin sulfate proteoglycans for cell invasion. Proc Natl Acad Sci U SA 104: 15882-15887.

Dinglasan, R. R., D. E. Kalume, S. M. Kanzok, A. K. Ghosh, O. Muratova, A. Pandey, and M. Jacobs-Lorena. 2007b. Disruption of Plasmodium falciparum development by antibodies against a conserved mosquito midgut antigen. Proc Natl Acad Sci USA 104: 13461-13466.

Echeverri, C. J., and N. Perrimon. 2006. High-throughput RNAi screening in cultured cells: a user's guide. Nat Rev Genet 7: 373-384.

Elschenbroich, S., Y. Kim, J. A. Medin, and T. Kislinger. 2010. Isolation of cell surface proteins for mass spectrometry-based proteomics. Expert Rev Proteomics 7: 141-154.

Estevez-Lao, T. Y., and J. F. Hillyer. 2014. Involvement of the Anopheles gambiae Nimrod gene family in mosquito immune responses. Insect Biochem Mol Biol 44: 12-22.

Evans, C. J., T. Liu, and U. Banerjee. 2014. Drosophila hematopoiesis: Markers and methods for molecular genetic analysis. Methods 68: 242-251.

Fehon, R. G., P. J. Kooh, I. Rebay, C. L. Regan, T. Xu, M. A. T. Muskavitch, and S. Artavanistsakonas. 1990. Molecular-interactions between the protein products of the neurogenic loci notch and delta, 2 egf-homologous genes in Drosophila. Cell 61: 523-534.

Franc, N. C., J.-L. Dimarcq, M. Lagueux, J. Hoffmann, and R. A. B. Ezekowitz. 1996. Croquemort, a novel Drosophila hemocyte/macrophage receptor that recognizes apoptotic cells. Immunity 4: 431-443.

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

Garver, L. S., Z. Xi, and G. Dimopoulos. 2008. Immunoglobulin superfamily members play an important role in the mosquito immune system. Dev Comp Immunol 32: 519-531.

Garver, L. S., Y. Dong, and G. Dimopoulos. 2009. Caspar Controls Resistance to Plasmodium falciparum in Diverse Anopheline Species. Plos Pathog. 5: e1000335.

Garver, L. S., A. C. Bahia, S. Das, J. A. Souza-Neto, J. Shiao, Y. Dong, and G. Dimopoulos. 2012. Anopheles Imd Pathway Factors and Effectors in Infection Intensity-Dependent Anti-Plasmodium Action. Plos Pathog 8 : e1002737.

Geisler, C., J. J. Aumiller, and D. L. Jarvis. 2008. A fused lobes gene encodes the processing beta-N-acetylglucosaminidase in Sf9 cells. J. Biol. Chem. 283: 11330-11339.

Gupta, L., A. Molina-Cruz, S. Kumar, J. Rodrigues, R. Dixit, R. E. Zamora, and C. Barillas-Mury. 2009. The STAT Pathway Mediates Late-Phase Immunity against Plasmodium in the Mosquito Anopheles gambiae. Cell Host Microbe 5: 498-507.

Haddow, A. D., H. Guzman, V. L. Popov, T. G. Wood, S. G. Widen, A. D. Haddow, R. B. Tesh, and S. C. Weaver. 2013. First isolation of Aedes flavivirus in the Western Hemisphere and evidence of vertical transmission in the mosquito Aedes (Stegomyia) albopictus (Diptera: Culicidae). Virology 440: 134-139.

Hire, R. S., G. Huab, Q. Zhang, R. Mishra, and M. J. Adang. 2015. Anopheles gambiae Ag55 cell line as a model for Lysinibacillus sphaericus Bin toxin action. J Invertebr Pathol132: 105-110.

Hobson-Peters, J., A. W. Y. Yam, J. W. F. Lu, Y. X. Setoh, F. J. May, N. Kurucz, S. Walsh, N. A. Prow, S. S. Davis, R. Weir, L. Melville, N. Hunt, R. I. Webb, B. J. Blitvich, P. Whelan, and R. A. Hall. 2013. A New Insect-Specific Flavivirus from Northern Australia Suppresses Replication of West Nile Virus and Murray Valley Encephalitis Virus in Co-infected Mosquito Cells. PLoS One 8.

Hu, X., L. Cherbas, and P. Cherbas. 2003. Transcription activation by the ecdysone receptor (EcR/USP): Identification of activation functions. Mol Endocrinol 17: 716-731.

Huang, D. W., B. T. Sherman, and R. A. Lempicki. 2009a. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res. 37: 1-13.

Huang, D. W., B. T. Sherman, and R. A. Lempicki. 2009b. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 4: 44-57.

Jaramillo-Gutierrez, G., A. Molina-Cruz, S. Kumar, and C. Barillas-Mury. 2010. The Anopheles gambiae Oxidation Resistance 1 (OXR1) Gene Regulates Expression of Enzymes That Detoxify Reactive Oxygen Species. PLoS One 5.

Kaell, L., A. Krogh, and E. L. L. Sonnhammer. 2007. Advantages of combined transmembrane topology and signal peptide prediction - the Phobius web server. Nucleic Acids Res. 35: W429-W432.

Kall, L., A. Krogh, and E. L. L. Sonnhammer. 2004. A combined transmembrane topology and signal peptide prediction method. J Mol Biol 338: 1027-1036.

Kim, D., G. Pertea, C. Trapnell, H. Pimentel, R. Kelley, and S. L. Salzberg. 2013. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol 14.

Konet, D. S., J. Anderson, J. Piper, R. Akkina, E. Suchman, and J. Carlson. 2007. Short-hairpin RNA expressed from polymerase III promoters mediates RNA interference in mosquito cells. Insect Mol Biol 16: 199-206.

Kotsyfakis, M., L. Ehret-Sabatier, I. Siden-Kiamos, J. Mendoza, R. E. Sinden, and C. Louis. 2005. Plasmodium berghei ookinetes bind to Anopheles gambiae and Drosophila melanogaster annexins. Mol Microbiol 57: 171-179.

Kuno, G. 2007. Host range specificity of flaviviruses: Correlation with in vitro replication. J Med Entomol 44: 93-101.

Kwon, H., K. Bang, and S. Cho. 2014. Characterization of the Hemocytes in Larvae of Protaetia brevitarsis seulensis: Involvement of Granulocyte-Mediated Phagocytosis. PLoS One 9.

Langmead, B., C. Trapnell, M. Pop, and S. L. Salzberg. 2009. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 10.

Lavazec, C., C. Boudin, R. Lacroix, S. Bonnet, A. Diop, S. Thiberge, B. Boisson, R. Tahar, and C. Bourgouin. 2007. Carboxypeptidases B of Anopheles gambiae as targets for a Plasmodium falciparum transmission-blocking vaccine. Infect Immun 75: 1635-1642.

Lavine, M. D., and M. R. Strand. 2002. Insect hemocytes and their role in immunity. Insect Biochem Mol Biol 32: 1295-1309.

Lawrie, C. H., N. Y. Uzcategui, M. Armesto, L. Bell-Sakyi, and E. A. Gould. 2004. Susceptibility of mosquito and tick cell lines to infection with various flaviviruses. Med Vet Entomol 18: 268-274.

Levashina, E. A., L. F. Moita, S. Blandin, G. Vriend, M. Lagueux, and F. C. Kafatos. 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.

Lombardo, F., Y. Ghani, F. C. Kafatos, and G. K. Christophides. 2013. Comprehensive Genetic Dissection of the Hemocyte Immune Response in the Malaria Mosquito Anopheles gambiae. Plos Pathog 9.

Marcu, O., M. P. Lera, M. E. Sanchez, E. Levic, L. A. Higgins, A. Shmygelska, T. F. Fahlen, H. Nichol, and S. Bhattacharya. 2011. Innate Immune Responses of Drosophila melanogaster Are Altered by Spaceflight. PLoS One 6.

Mathias, D. K., J. G. Jardim, L. A. Parish, J. S. Armistead, H. V. Trinh, C. Kumpitak, J. Sattabongkot, and R. R. Dinglasan. 2014. Differential roles of an Anopheline midgut GPI-anchored protein in mediating Plasmodium falciparum and Plasmodium vivax ookinete invasion. Infect Genet Evol 28: 635-647.

Mathias, D. K., R. Pastrana-Mena, E. Ranucci, D. Tao, P. Ferruti, C. Ortega, G. O. Staples, J. Zaia, E. Takashima, T. Tsuboi, N. A. Borg, L. Verotta, and R. R. Dinglasan. 2013. A Small Molecule Glycosaminoglycan Mimetic Blocks Plasmodium Invasion of the Mosquito Midgut. Plos Pathog 9.

Meister, S., S. M. Kanzok, X. L. Zheng, C. Luna, T. R. Li, N. T. Hoa, J. R. Clayton, K. P. White, F. C. Kafatos, G. K. Christophides, and L. B. Zheng. 2005. Immune signaling pathways regulating bacterial and malaria parasite infection of the mosquito Anopheles gambiae. Proc Natl Acad Sci USA 102: 11420-11425.

Muller, H. M., G. Dimopoulos, C. Blass, and F. C. Kafatos. 1999. A hemocyte-like cell line established from the malaria vector Anopheles gambiae expresses six prophenoloxidase genes. J Biol Chem 274: 11727-11735.

Nelson, R. E., L. I. Fessler, Y. Takagi, B. Blumberg, D. R. Keene, P. F. Olson, C. G. Parker, and J. H. Fessler. 1994. Peroxidasin - a novel enzyme-matrix protein of Drosophila development. EMBO J 13: 3438-3447.

O'Neal, M. A., B. A. Posner, C. J. Coates, and J. M. Abrams. 2013. A Cell-Based Screening Platform Identifies Novel Mosquitocidal Toxins. J Biomol Screen 18: 688-694.

Organization, W. H. 2015. World malaria report 2015. 2015. WHO: Geneva.

Parish, L. A., D. R. Colquhoun, C. U. Mohien, A. E. Lyashkov, D. R. Graham, and R. R. Dinglasan. 2011. Ookinete-Interacting Proteins on the Microvillar Surface are Partitioned into Detergent Resistant Membranes of Anopheles gambiae Midguts. J Proteome Res10: 5150-5162.

Parish, L. A., L. S. Garver, D. R. Colquhoun, C. U. Mohien, E. Weissbrod, and R. R. Dinglasan. 2013. Dissecting Mosquito-Parasite Interactions through Molecular Biology and Biochemistry: Genomic, Proteomic and Glycomic Analyses. Malaria Parasites: Comparative Genomics , Evolution and Molecular Biology: 221-248.

Pinto, S. B., F. Lombardo, A. C. Koutsos, R. M. Waterhouse, K. McKay, C. An, C. Ramakrishnan, F. C. Kafatos, and K. Michel. 2009. Discovery of Plasmodium modulators by genome-wide analysis of circulating hemocytes in Anopheles gambiae. Proc Natl Acad Sci USA 106: 21270-21275.

Pudney, M., M. G. R. Varma, and C. J. Leake. 1979. Establishment of cell lines from larvae of culicine (Aedes species) and anopheline mosquitoes. Tca (Tissue Culture Association) Manual 5: 997-1002.

Rasgon, J. L., X. Ren, and M. Petridis. 2006. Can Anopheles gambiae be infected with Wolbachia pipientis? Insights from an in vitro system. Appl Environ Microbiol 72: 7718-7722.

Redmond, S. N., K. Eiglmeier, C. Mitri, K. Markianos, W. M. Guelbeogo, A. Gneme, A. T. Isaacs, B. Coulibaly, E. Brito-Fravallo, G. Maslen, D. Mead, O. Niare, S. F. Traore, N. F. Sagnon, D. Kwiatkowski, M. M. Riehle, and K. D. Vernick. 2015. Association mapping by pooled sequencing identifies TOLL 11 as a protective factor against Plasmodium falciparum in Anopheles gambiae. BMC Genomics 16.

Ren X., Hoiczyk E., Rasgon J.L. 2008. Viral Paratransgenesis in the Malaria vector *Anopheles gambiae*. PLoS Pathog 4(8): e1000135.

Schnettler, E., C. L. Donald, S. Human, M. Watson, R. W. C. Siu, M. McFarlane, J. K. Fazakerley, A. Kohl, and R. Fragkoudis. 2013. Knockdown of piRNA pathway proteins results in enhanced Semliki Forest virus production in mosquito cells. J Gen Virol 94: 1680-1689.

Shin, S. W., G. Bian, and A. S. Raikhel. 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.

Siden-Kiamos, I., D. Vlachou, G. Margos, A. Beetsma, A. P. Waters, R. E. Sinden, and C. Louis. 2000. Distinct roles for Pbs21 and Pbs25 in the in vitro ookinete to oocyst transformation of Plasmodium berghei. J. Cell Sci.113: 3419-3426.

Sim, S., and G. Dimopoulos. 2010. Dengue Virus Inhibits Immune Responses in Aedes aegypti Cells. PLoS One 5.

Smith, K. E., and P. J. Linser. 2009. Silencing of carbonic anhydrase in an Anopheles gambiae larval cell line, Ag55. J RNAi Gene Silencing 5: 345.

Trapnell, C., L. Pachter, and S. L. Salzberg. 2009. TopHat: discovering splice junctions with RNA-Seq. Bioinformatics 25: 1105-1111.

Trapnell, C., B. A. Williams, G. Pertea, A. Mortazavi, G. Kwan, M. J. van Baren, S. L. Salzberg, B. J. Wold, and L. Pachter. 2010. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol 28: 511-U174.

Trapnell, C., A. Roberts, L. Goff, G. Pertea, D. Kim, D. R. Kelley, H. Pimentel, S. L. Salzberg, J. L. Rinn, and L. Pachter. 2012. Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc 7: 562-578.

Vannini, L., W. A. Dunn, T. W. Reed, and J. H. Willis. 2014. Changes in transcript abundance for cuticular proteins and other genes three hours after a blood meal in Anopheles gambiae. Insect Biochem Mol Biol 44: 33-43.

Vega-Rodriguez, J., A. K. Ghosh, S. M. Kanzok, R. R. Dinglasan, S. Wang, N. J. Bongio, D. E. Kalume, K. Miura, C. A. Long, A. Pandey, and M. Jacobs-Lorena. 2014. Multiple pathways for Plasmodium ookinete invasion of the mosquito midgut. Proc Natl Acad Sci USA 111: E492-E500.

Via, L., R. Fratti, M. McFalone, E. Pagan-Ramos, D. Deretic, and V. Deretic. 1998. Effects of cytokines on mycobacterial phagosome maturation. J Cell Sci 111: 897-905.

Waldock, J., K. E. Olson, and G. K. Christophides. 2012. Anopheles gambiae Antiviral Immune Response to Systemic O'nyong-nyong Infection. PLoS Negl Trop Dis 6.

Walker, T., C. L. Jeffries, K. L. Mansfield, and N. Johnson. 2014. Mosquito cell lines: history, isolation, availability and application to assess the threat of arboviral transmission in the United Kingdom. Parasit Vectors 7.

Weekes, M. P., R. Antrobus, J. R. Lill, L. M. Duncan, S. Hor, and P. J. Lehner. 2010. Comparative analysis of techniques to purify plasma membrane proteins. J Biomol Tech.: JBT 21: 108-115.

Wilkins, S., and P. F. Billingsley. 2001. Partial characterization of oligosaccharides expressed on midgut microvillar glycoproteins of the mosquito, Anopheles stephensi Liston. Insect Biochem Mol Biol 31: 937-948.

Wilkins, S., and P. F. Billingsley. 2010. Mosquito cell line glycoproteins: an unsuitable model system for the Plasmodium ookinete-mosquito midgut interaction? Parasit Vectors 3.

Zhang, G., G. Niu, C. M. Franca, Y. Dong, X. Wang, N. S. Butler, G. Dimopoulos, and J. Li. 2015. Anopheles Midgut FREP1 Mediates Plasmodium Invasion. J Biol Chem 290: 16490-16501.

Zhao, Y., W. Zhang, Y. Kho, and Y. Zhao. 2004. Proteomic analysis of integral plasma membrane proteins. Anal Chem 76: 1817-1823.

Zieler, H., J. P. Nawrocki, and M. Shahabuddin. 1999. Plasmodium gallinaceum ookinetes adhere specifically to the midgut epithelium of Aedes aegypti by interaction with a carbohydrate ligand. J Exp Biol 202: 485-495.

Zieler, H., C. F. Garon, E. R. Fischer, and M. Shahabuddin. 2000. A tubular network associated with the brush-border surface of the Aedes aegypti midgut: implications for pathogen transmission by mosquitoes. J Exp Biol 203: 1599-1611.

Figure legends

Figure 3.1. Enriched Biological process terms in Ag55 cells relative to blood fed female adult (BF_Adult). Only the terms with Benjamini corrected FDR <0.05 were selected. The number on the right hand side of the bar denotes the number of genes in each GO BP term

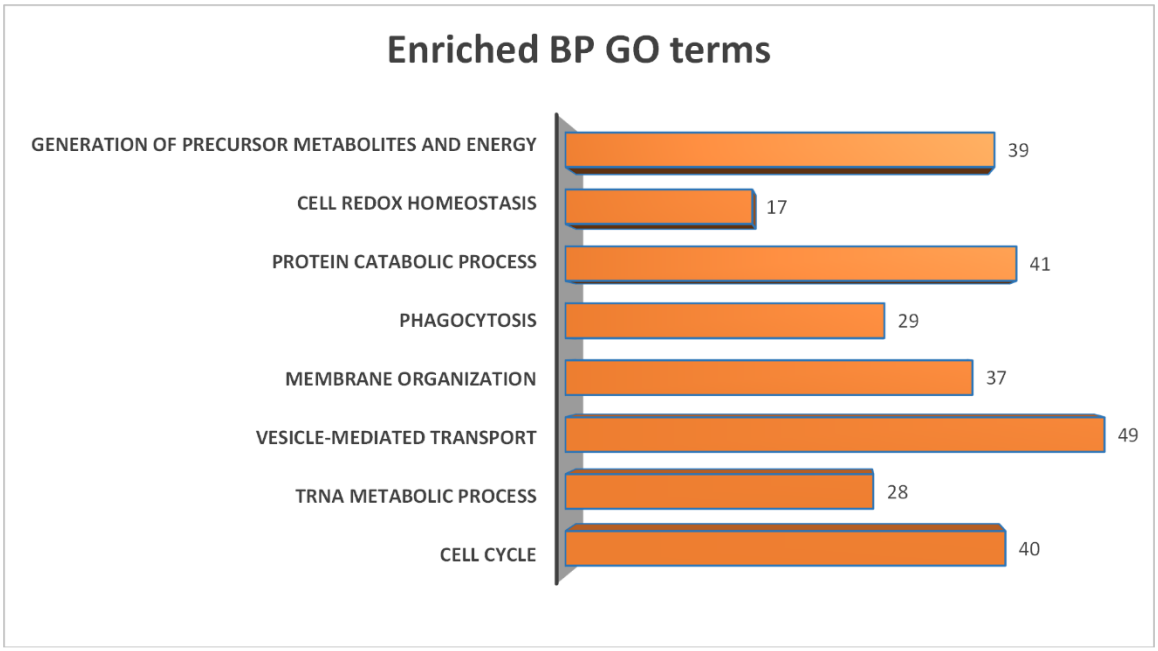

Enriched BP GO terms
GENERATION OF PRECURSOR METABOLITES AND ENERGY
39
CELL REDOX HOMEOSTASIS
17
PROTEIN CATABOLIC PROCESS
41
PHAGOCYTOSIS
29
MEMBRANE ORGANIZATION
37
VESICLE-MEDIATED TRANSPORT
49
TRNA METABOLIC PROCESS
28
CELL CYCLE
40

Figure 3.2. Confocal image of live Ag55 cells showing their morphology. 20 X (A) 63 X (B) magnification. Nile red: used to stain the lipids and Hoechst stain: used to stain the nucleus.

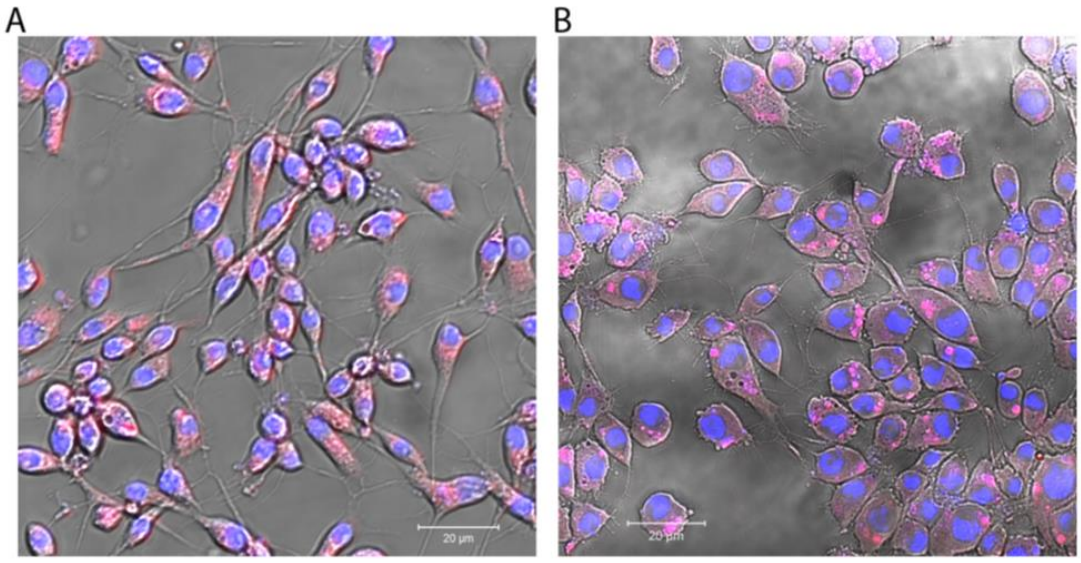

A
B
20 µm
20 µm

Figure 3.3. Heatmap showing the expression level of gene transcripts involved in phagocytosis in Ag55 cells

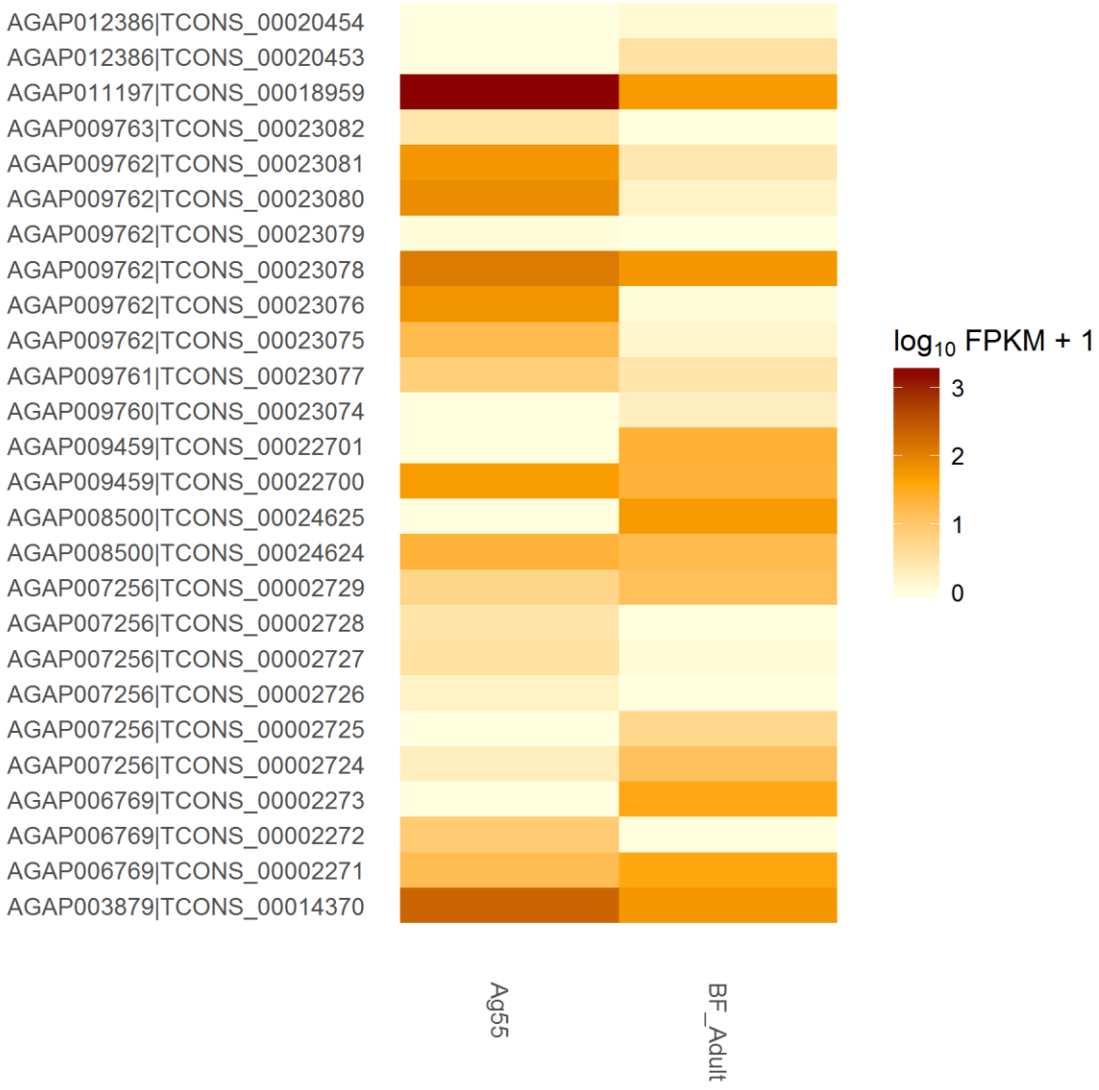

AGAP012386|TCONS_00020454
AGAP012386|TCONS_00020453
AGAP011197|TCONS_00018959
AGAP009763|TCONS_00023082
AGAP009762|TCONS_00023081
AGAP009762|TCONS_00023080
AGAP009762|TCONS_00023079
AGAP009762|TCONS_00023078
AGAP009762|TCONS_00023076
AGAP009762|TCONS_00023075
AGAP009761|TCONS_00023077
AGAP009760|TCONS_00023074
AGAP009459|TCONS_00022701
AGAP009459|TCONS_00022700
AGAP008500|TCONS_00024625
AGAP008500|TCONS_00024624
AGAP007256|TCONS_00002729
AGAP007256|TCONS_00002728
AGAP007256|TCONS_00002727
AGAP007256|TCONS_00002726
AGAP007256|TCONS_00002725
AGAP007256|TCONS_00002724
AGAP006769|TCONS_00002273
AGAP006769|TCONS_00002272
AGAP006769|TCONS_00002271
AGAP003879|TCONS_00014370
Ag55
BF_Adult
log10 FPKM + 1
3
2
1
0

Figure 3.4. Confocal image showing Ag55 cells phagocytosing FITC-labeled E. coli bioparticles. A: Hoechst stain, B: Bright field, C: FITC-labeled *E. coli* bioparticles, D: Lysotracker deep red E: Merged image. 20 X magnification. Fixed cell imaging.

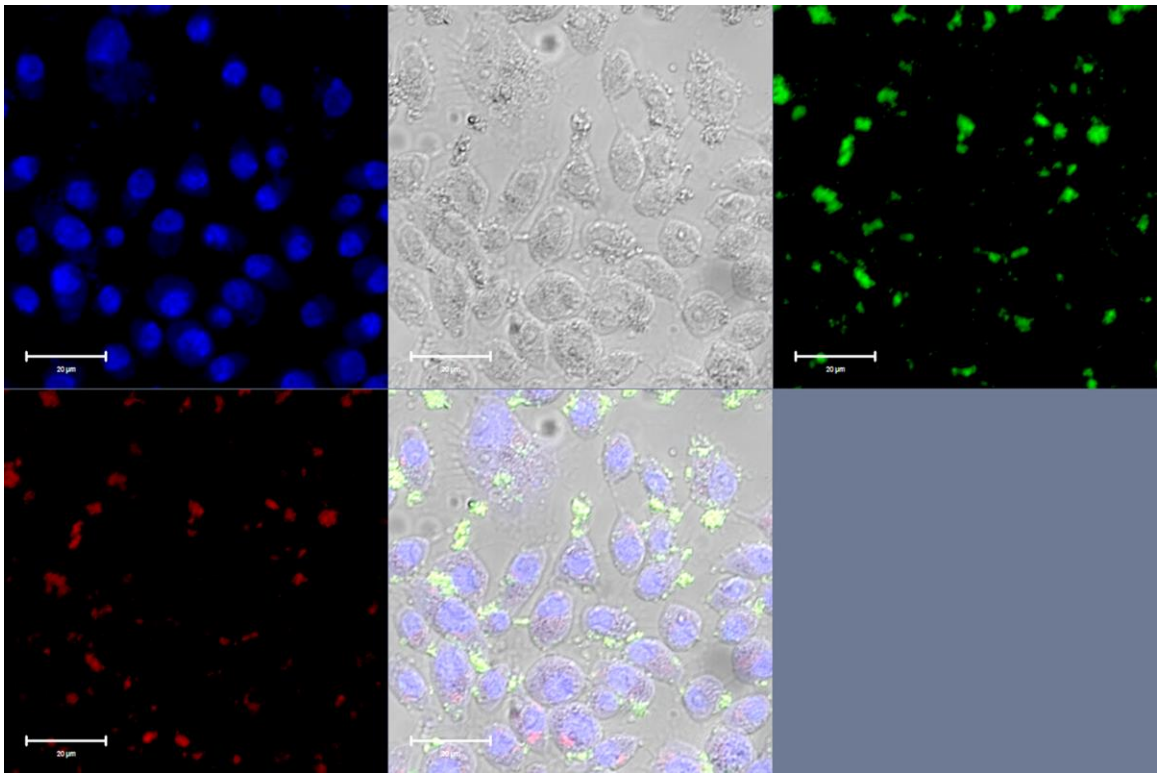

A
B
C
D
E

Figure 3.6. Expression Barplot of genes involved in encoding ookinete interacting proteins in Ag55 cells.

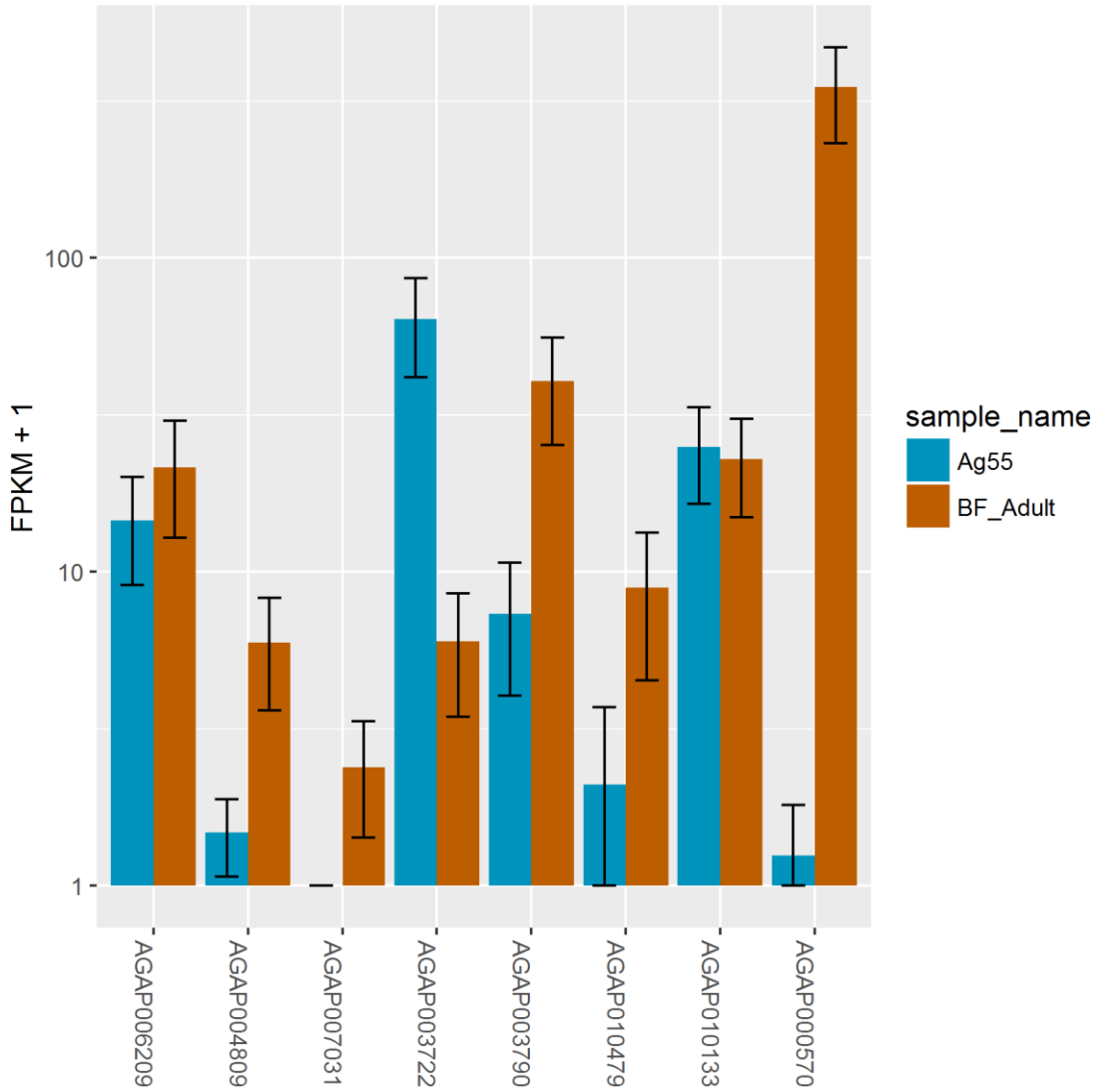

FPKM + 1
100
10
1
AGAP006209
AGAP004809
AGAP007031
AGAP003722
AGAP003790
AGAP010479
AGAP010133
AGAP000570
sample_name
Ag55
BF_Adult

Table 3.1. Expression level of hemocyte marker genes (FPKM) in Ag55 cells.

FPKM range				
≥100	**<100 & ≥1**	**<1 & ≥0.1**	**<0.1 & ≥0.01**	**<0.01**
misshapen (AGAP006340)	***HPX4/IRID3*** (AGAP007237)	***Lozenge homolog*** (AGAP002506)	***PPO3*** (AGAP004975)	
PGRPLC (AGAP005203)	***draper ortholog*** (AGAP007256)	***PPO4*** (AGAP004981)	***PPO8*** (AGAP004976)	
LYSC1 (AGAP007347)	(AGAP007314)	***CLIPB14*** (AGAP010833)	***PPO6*** (AGAP004977)	
***GATA* (*pannier homolog*)** (AGAP002235)	***SRPN10*** (AGAP005246)	AGAP007783	***PPO2*** (AGAP006258)	
collagen, type IV, alpha (AGAP009201)	AGAP001243	AGAP009231	***PPO1*** (AGAP002825)	
collagen, type IV, alpha (AGAP009200)	***Prophenoloxidase activating factor*** (AGAP001964)	AGAP009549	***GATA* (*pannier homolog*)** (AGAP004228)	
eater (AGAP009762)	***APOD*** (AGAP002593)	AGAP000806	***Gcm ortholog*** (AGAP007782)	
SCRASP1*/*Sp22D (AGAP005625)	***Prophenoloxidase activating factor*** (AGAP010730)		***SCRBQ3*** (AGAP008179)	
SCRC1/D-SR-C1 (AGAP011974)	***Spondin-1*** (AGAP011765)		***PPO5*** (AGAP012616)	
	PSMD3/DoxA2 (AGAP009082)			
	CLIPB15 (AGAP009844)			
	SCRBQ2 (AGAP010133)			
	PDGF-VEGF receptor ortholog (AGAP008813)			
	SRPN6 (AGAP009212)			
	histone arginine demethylase (AGAP000158)			
	Nimrod family (AGAP009763)			
	Dscam ortholog (AGAP010884)			

Table 3.2. List of Hemocyte marker proteins identified in Ag55 cells proteomic analysis

Uniprot_ID	Sequence_name	Length_(AA)	Total_Score	Uniq_Peptides	Coverage	PSMs	TMD	Signal_P
O61470	Probable 26S proteasome non-ATPase regulatory subunit 3 (Dox-A2)	496	23.41	5	12.7	13	0	No
A7UTA1	PGRPLC AGAP005203-PC	464	9.44	2	9.91	2	1	No
A7UU73	misshapen/NIK-related kinase AGAP006340-PB	1188	15.92	3	6.14	6	0	No
Q5TWT0	Multiple epidermal growth factor-like domains 10 (Draper ortholog) AGAP007256-PA	1171	10.83	3	3.5	4	1	Yes
Q7PM27	eater AGAP009762-PA	1300	11.06	2	2.69	6	1	Yes
Q7PMH9	Class B Scavenger Receptor (CD36 domain) AGAP010133-PA (SCRBQ2)	492	18.79	4	13.41	15	2	No
Q7PZL9	Class C Scavenger Receptor AGAP011974-PA (SCRC1) (D-SR-C1 ortholog)	582	14.06	3	11	4	1	Yes
Q7Q840	FMS-like tyrosine kinase 1 AGAP008813-PA (PDGF-VEGF receptor ortholog)	1247	9.31	2	3.53	2	1	No
Q7QGT8	Dscam ortholog AGAP010884-PA	1951	18.92	4	4.31	8	1	No
Q7QJ29	Heme peroxidase 4 AGAP007237-PA (HPX4) (Peroxidasin ortholog)	1514	29.5	6	7.53	13	0	No

PSMs: Peptide spectral matches, Total score: Xcorr score, TMD: Transmembrane domain, Signal_P: Signal peptide. TMD and Signal_P predicted using Phobius.

Table 3.3. Expression (FPKM) range of genes encoding enzymes involved in immune signaling pathways in Ag55 cells.

Immune signaling pathways	FPKM range				
	≥100	**<100 & ≥1**	**<1 & ≥0.1**	**<0.1 & ≥0.01**	**<0.01**
IMD	*PGRPLC* (AGAP005203)	*IMD* (AGAP004959)			
		Faf1 (AGAP006473)			
		REL2 (AGAP006747)			
		FADD (AGAP007173)			
		MAP3K7IP1 (AGAP002953)			
		CASPL1 (AGAP011693)			
		IKK1 (AGAP009166)			
JAK-STAT		*SOCS* (AGAP004844)			
		PIAS (AGAP005031)			
		SOCS (AGAP011042)			
		STAM (AGAP008494)			
		JAK (AGAP008354)			
		DOME (AGAP010083)			
		STAT2 (AGAP000099)			
		SOCS (AGAP000880)			
		CBP/P300 (AGAP000029)			
TOLL	*SPZ4* (AGAP007866)	*PGRPS3* (AGAP006342)	*PGRPS2* (AGAP006343)	*TOLL9* (AGAP006974)	*TOLL* (AGAP012385)
	PGRPS1 (AGAP000536)	*SPZ5* (AGAP007177)	*TOLL10* (AGAP011187)	*TOLL7* (AGAP012326)	*TOLL6* (AGAP012387)
		MYD (AGAP005252)	*SPZ3* (AGAP008360)	*TOLL1A* (AGAP001004)	
		TUBE (AGAP003062)			
		PELLE (AGAP002966)			
		TOLL11 (AGAP011186)			
		CACT (AGAP007938)			

		REL1 (AGAP009515)			
		SPZ1 (AGAP000346)			
		TOLL5A (AGAP000999)			

Table 3.4. Expression (FPKM) range of genes encoding enzymes involved in glycan biosynthesis in Ag55 cells.

Glycan Biosynthesis	**FPKM range**				
	≥100	**<100 & ≥1**	**<1 & ≥0.1**	**<0.1 & ≥0.01**	**<0.01**
N-glycan	*MAN1A2* (AGAP000558)	*ALG2* (AGAP001232)	*ALG13* (AGAP003697)		
		ALG10 (AGAP002420)	*ALG14* (AGAP003461)		
		ALG1 (AGAP003551)			
		ALG9 (AGAP003601)			
		ALG13 (AGAP003699)			
		ALG8 (AGAP003928)			
		ALG11 (AGAP011324)			
		ALG6 (AGAP008946)			
		DPAGT1 (AGAP008131)			
		ALG12 (AGAP000102)			
		MAN1B1 (AGAP003884)			
		MAN2A2 (AGAP004020)			
		GANAB (AGAP000862)			
		MGAT1 (AGAP005347)			
		FUCT6 (AGAP001888)			
		MGAT2 (AGAP004397)			
		MGAT3 (AGAP008132)			
		MGAT4 (AGAP012440)			

		FDL (AGAP007080)			
O-glycan		*GALNT* (AGAP006925)	*GALNT* (AGAP006881)		
		GALNT (AGAP004429)			
		GALNT (AGAP011984)			
		GALNT (AGAP012414)			
		GALNT (AGAP008229)			
		GALNT (AGAP008613)			
		C1GALT1/C1GALT1C1 (AGAP009979)			
		GALNT (AGAP010078)			
		GALNT (AGAP000656)			
Chondroitin sulfate		*XYLT1,2* (AGAP005811)	*CSGALNACT1,2* (AGAP007720)	*CHPF* (AGAP001050)	
		BGAT3 (AGAP002801)	*CHSY1,3* (AGAP001010)		
		B3GALT6 (AGAP009175)			

Table 3.5. Proteins identified using three different methods

	Methods		
Proteins identified	**Cell surface biotinylation**	**Total membrane preparation**	**Whole cell**
Total	696	1087	1260
Integral Membrane Proteins (IMP)	87	181	111
%IMP	12.5	16.651	8.809

CHAPTER 4

GENERAL DISCUSSION AND CONCLUSION

Insects belonging to the order Diptera have greater economic impact on humans than any other insect Order due to their impact as economically important agricultural pests and as vectors of major human diseases. The malarial mosquito Anopheles gambiae and the spotted wing Drosophila, *D. suzukii* were the subjects of my dissertation research. *A. gambiae* is the major vector of malaria in Africa and *D. suzukii* has become a major pest of soft and thin skinned fruits in the US. The applications of broad-spectrum chemicals such as pyrethroids, spinosyns and organophosphates have been used as a primary strategy to manage *D. suzukii* and *A. gambiae* (Darriet et al. 2005, Bai et al. 2010, Beers et al. 2011a, Bruck et al. 2011b, Walsh et al. 2011b, Marina et al. 2014). Due to the frequent use of these broad spectrum chemicals, several cases of insecticide resistance have been reported in field population of *A. gambiae* (Adasi and Hemingway 2008, Wanjala et al. 2015). Similarly, in *D. suzukii* field population a significant level of permethrin resistance has been reported (Bolda 2011). RNA-sequencing has been used extensively to elucidate possible mechanisms underlying insecticide resistance (David et al. 2014, Bonizzoni et al. 2015, Faucon et al. 2015).

The vectorial capacity of mosquitoes to transmit infectious pathogens depends on many pathogen-host interactions such as pathogen entry and development in the host, and these interactions are countered by the innate immune response of the host (Dimopoulos et al. 1997, Waldock et al. 2012, Parish et al. 2013, Walker et al. 2014). Environmental

factors provide challenges for studying mosquito-infectious microbe interactions at the whole animal level. The handling of mosquitoes and infectious pathogens requires repertoire of skills and facilities including the ability to raise large numbers of mosquitoes, facilities for handling both the insect vector and the pathogen and expertise in vector and pathogen biology. To overcome the above mentioned problems, mosquito cell lines can be developed and used as models to investigate mosquito pathogen interactions. Several mosquito cell lines have been used for host-microbe interactions (reviewed in Walker et al. 2014).

A. gambiae cell lines can be useful in understanding the vector-pathogen interactions as high-throughput RNAi screens have been highly successful on cell lines (Boutros et al. 2004b, Echeverri and Perrimon 2006). Availability of gene expression profile of cell line enhances the effectiveness of high-throughput RNAi screens and RNA-sequencing is the most efficient and economical method for providing detailed gene expression profile. Furthermore, using RNA-sequencing data researchers have provided valuable insight into the host seeking behavior of *A. gambiae* (McBride et al. 2014) and also their interaction with *Plasmodium* (Biryukova et al. 2014).

In chapter two I used RNA-sequencing method to elucidate the response of LC_{50} doses of three different insecticides (Zeta-cypermethrin, Spinosad and Malathion) on Pierce and Clarke population of *D. suzukii*. Pierce population was collected from a commercial blueberry orchard located in Pierce County, Georgia, U.S.A. and a Clarke population previously collected from unmanaged blueberry bushes located in Clarke County, Georgia, U.S.A. The rationale for this study was to look for the responses of commonly used three different insecticide classes against Pierce and Clarke population as

the Pierce population has previously been exposed to the insecticides regularly but the Clarke population was not exposed to the insecticide when collected from the field. The LC_{50} values were significantly higher in the Pierce population for all treatments relative to Clarke population, indicating that it is less susceptible to these insecticides. Pairwise comparison on RNA-sequencing data was performed between insecticide treated versus untreated (control) in Clarke as well as Pierce population. Higher number of differentially expressed genes (DEGs) were identified in *D. suzukii* Pierce population upon three different insecticide treatments. Furthermore, a high number of significantly DEGs were involved in detoxification and reduced cuticular penetration, especially in Pierce population, with a high degree of overlap in DEGs when *D. suzukii* were treated with these three different insecticide classes. The higher number of significantly DEGs involved in detoxification and reduced cuticular penetration in Pierce population correlates with the higher LC_{50} values observed in Pierce county flies suggesting that Pierce population has stronger capacity to detoxify insecticides and cope with insecticide treatments. It is important to point out that there were some genes that were commonly differentially expressed (DE) in Pierce and Clarke populations, although a higher number of DEGs were identified in Pierce population upon insecticide treatments. This suggests that the DEGs that were exclusively identified in Pierce population might be responsible for the lower susceptibility of Pierce population towards three different insecticides. Furthermore, comparing the basal expression level of genes in Pierce and Clarke untreated samples identified that most of the genes involved in detoxification and cuticular penetration, including those known to confer insecticide resistance, have significantly higher baseline expression level in Clarke population samples relative to the

Pierce population untreated samples. These results were quite puzzling since the LC_{50} value in Clarke population was lower as compared to the Pierce population, initially it was expected from the Clarke population to have a lower baseline expression of genes, especially those that are involved in detoxification and conferring insecticide resistance. These results led to the assumption that the induction level of these genes could be more important to baseline gene expression level when considering the action of these protein products to counteract the effects of insecticides. Furthermore, SNVs and indels analysis was performed on Pierce and Clarke population (control samples) in order to provide valuable insight into the finding that untreated Clarke population has higher baseline expression level of many genes, including those involved in insecticide response, as compared to the Pierce population. The SNVs and indels have been associated with insecticide resistance in insects (Li et al. 2007c, Baxter et al. 2010b, David et al. 2014). Analyzing the SNVs and indels associated with Pierce and Clarke population will improve our understanding of roles played by these SNVs and indels towards the susceptibility (i.e less or more susceptible) against different insecticides. Slightly higher numbers of predicted nonsynonymous SNVs were identified in the Pierce population when compared to the Clarke population when all expressed genes were taken into consideration. In contrast a higher number of nonsynonymous SNVs were identified in Clarke population samples when the genes encoding detoxification enzymes, insecticide targets, and cuticular proteins were taken into consideration. Presence of higher number of nonsynonymous mutations in genes encoding detoxification enzymes might have led to the expression of non-functional detoxification enzymes upon insecticide treatment, which may be the reason for why the Clarke population is much more susceptible to

insecticides as compared to Pierce population in my bioassays. Furthermore, there were higher number of nonsynonymous SNVs in detoxification genes that were predicted to have deleterious effects on protein function in Clarke population and it is perhaps not surprising that many of these genes were not differentially expressed upon insecticide treatment. The results suggest that combined approach of gene expression and SNV analysis may be effective in identifying potential candidate genes involved in insecticide response in insect population.

In chapter 3, I further used omics approaches to provide detailed gene expression profile, and proteome of Ag55 (*A. gambiae*) cell line. Furthermore, gene expression profile of Ag55 cells were compared with blood fed female adults to establish whether or not Ag55 cell line is a good model to study mosquito-pathogen interaction and immune response.

Phagocytosis biological processes GO term was significantly enriched in Ag55 cells relative to blood feed female adults leading to the hypothesis, that Ag55 cells have phagocytic properties, which was further validated using confocal imaging. Baton et al. identified 4047 genes expressed in hemocytes, using *A. gambiae* genome-wide microarrays and also found 279 genes having significantly higher expression in hemocytes relative to whole adult female mosquitoes (Baton et al. 2009). Out of 279, 211 genes were expressed in Ag55 cells. Proteomics data also identified 35 proteins encoded by enriched genes. Furthermore, hemocyte markers genes such as *nimrodB2*, *draper*, *peroxidasin*, *SCRBQ2*, *PGDF-VEGF receptor* homologs, and *Dox-A2* were expressed in Ag55 cells. These expressed genes were also identified at protein level using LC-MS/MS analysis. Apart from the expressed hemocyte markers Ag55 cells also exhibited

phagocytosis property. Hemocyte uses phagocytosis as defense mechanism against invading microbes (Lavine and Strand 2002, Browne et al. 2013). Furthermore, Ag55 cells express ookinete interacting proteins which are expressed in the midgut surface. These results suggest that Ag55 cells predominantly express hemocyte like properties but also have some properties of midgut tissue.

In conclusion, results from my dissertation research highlight the usefulness of omics approaches. The results from RNA-sequencing analysis not only identified the genes that are inducible with specific insecticide treatments and possible resistance mechanisms, but also have provided a baseline for comparing similar data from laboratory-selected and field-evolved resistant *D. suzukii* populations in the future. Furthermore, RNA-sequencing can be used as an effective tool to assess the development of insecticide resistance and formation of strategies to delay the development of resistance in devastating insect pests in a timely manner. Similarly making available the detailed gene expression profile and proteome of Ag55 cells will help researchers use and engineer Ag55 cell line in a useful way, for example by developing strategies to make it suitable for *Plasmodium* and other microbes interaction studies.

4.1. References

Adasi, K., and J. Hemingway. 2008. Susceptibility to three pyrethroids and detection of knockdown resistance mutation in Ghanaian Anopheles gambiae sensu stricto. J Vector Ecol 33: 255-262.

Bai, H., D. B. Gelman, and S. R. Palli. 2010. Mode of action of methoprene in affecting female reproduction in the African malaria mosquito, Anopheles gambiae. Pest Manag Sci 66: 936-943.

Baton, L. A., A. Robertson, E. Warr, M. R. Strand, and G. Dimopoulos. 2009. Genome-wide transcriptomic profiling of Anopheles gambiae hemocytes reveals pathogen-specific signatures upon bacterial challenge and Plasmodium berghei infection. BMC Genomics 10.

Baxter, S. W., M. Chen, A. Dawson, J.-Z. Zhao, H. Vogel, A. M. Shelton, D. G. Heckel, and C. D. Jiggins. 2010. Mis-Spliced Transcripts of Nicotinic Acetylcholine Receptor alpha 6 Are Associated with Field Evolved Spinosad Resistance in Plutella xylostella (L.). Plos Genet 6.

Beers, E. H., R. A. Van Steenwyk, P. W. Shearer, W. W. Coates, and J. A. Grant. 2011. Developing Drosophila suzukii management programs for sweet cherry in the western United States. Pest Manag Sci 67: 1386-1395.

Biryukova, I., T. Ye, and E. Levashina. 2014. Transcriptome-wide analysis of microRNA expression in the malaria mosquito Anopheles gambiae. BMC Genomics 15.

Bolda, M. 2011. Suspected tolerance to pyganic (pyrethrin) found in spotted wing drosophila. Posted on August 23.

Bonizzoni, M., E. Ochomo, W. A. Dunn, M. Britton, Y. Afrane, G. Zhou, J. Hartsel, M.-C. Lee, J. Xu, A. Githeko, J. Fass, and G. Yan. 2015. RNA-seq analyses of changes in the Anopheles gambiae transcriptome associated with resistance to pyrethroids in Kenya: identification of candidate-resistance genes and candidate-resistance SNPs. Parasit Vectors 8.

Boutros, M., A. A. Kiger, S. Armknecht, K. Kerr, M. Hild, B. Koch, S. A. Haas, R. Paro, N. Perrimon, and C. Heidelberg Fly Array. 2004. Genome-wide RNAi analysis of growth and viability in Drosophila cells. Science 303: 832-835.

Browne, N., M. Heelan, and K. Kavanagh. 2013. An analysis of the structural and functional similarities of insect hemocytes and mammalian phagocytes. Virulence 4: 597-603.

Bruck, D. J., M. Bolda, L. Tanigoshi, J. Klick, J. Kleiber, J. DeFrancesco, B. Gerdeman, and H. Spitler. 2011. Laboratory and field comparisons of insecticides to reduce infestation of Drosophila suzukii in berry crops. Pest Manag Sci 67: 1375-1385.

Darriet, F., S. Duchon, and J. M. Hougard. 2005. Spinosad: A new larvicide against insecticide-resistant mosquito larvae. J Am Mosq Control Assoc 21: 495-496.

David, J.-P., F. Faucon, A. Chandor-Proust, R. Poupardin, M. A. Riaz, A. Bonin, V. Navratil, and S. Reynaud. 2014. Comparative analysis of response to selection with three insecticides in the dengue mosquito Aedes aegypti using mRNA sequencing. BMC Genomics 15.

Dimopoulos, G., A. Richman, H. M. Muller, and F. C. Kafatos. 1997. Molecular immune responses of the mosquito Anopheles gambiae to bacteria and malaria parasites. Proc Nat Acad Sci USA 94: 11508-11513.

Echeverri, C. J., and N. Perrimon. 2006. High-throughput RNAi screening in cultured cells: a user's guide. Nat Rev Genet 7: 373-384.

Faucon, F., I. Dusfour, T. Gaude, V. Navratil, F. Boyer, F. Chandre, P. Sirisopa, K. Thanispong, W. Juntarajumnong, R. Poupardin, T. Chareonviriyaphap, R. Girod, V. Corbel, S. Reynaud, and J.-P. David. 2015. Identifying genomic changes associated with insecticide resistance in the dengue mosquito Aedes aegypti by deep targeted sequencing. Genome Res 25: 1347-1359.

Lavine, M. D., and M. R. Strand. 2002. Insect hemocytes and their role in immunity. Insect Biochem Mol Biol 32: 1295-1309.

Li, X., M. A. Schuler, and M. R. Berenbaum. 2007. Molecular mechanisms of metabolic resistance to synthetic and natural xenobiotics. Annu Rev Entomol 52: 231-253.

Marina, C. F., J. G. Bond, J. Munoz, J. Valle, R. Novelo-Gutierrez, and T. Williams. 2014. Efficacy and non-target impact of spinosad, Bti and temephos larvicides for control of Anopheles spp. in an endemic malaria region of southern Mexico. Parasit Vectors 7.

McBride, C. S., F. Baier, A. B. Omondi, S. A. Spitzer, J. Lutomiah, R. Sang, R. Ignell, and L. B. Vosshall. 2014. Evolution of mosquito preference for humans linked to an odorant receptor. Nature 515: 222-U151.

Parish, L. A., L. S. Garver, D. R. Colquhoun, C. U. Mohien, E. Weissbrod, and R. R. Dinglasan. 2013. Dissecting Mosquito-Parasite Interactions through Molecular Biology and Biochemistry: Genomic, Proteomic and Glycomic Analyses. Malaria Parasites: Comparative Genomics , Evolution and Molecular Biology: 221-248.

Waldock, J., K. E. Olson, and G. K. Christophides. 2012. Anopheles gambiae Antiviral Immune Response to Systemic O'nyong-nyong Infection. Plos Negl Trop Dis 6.

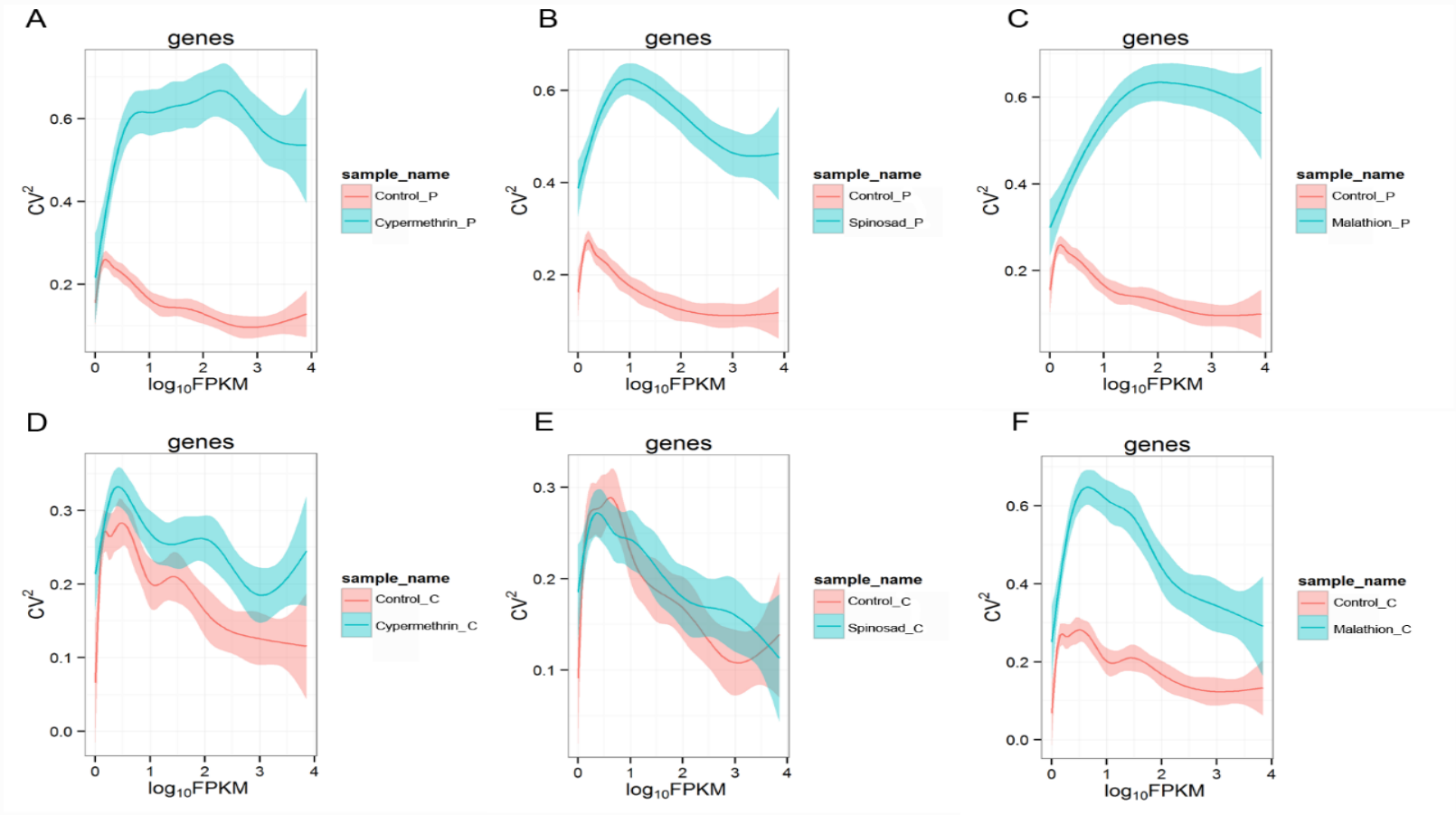
Walker, T., C. L. Jeffries, K. L. Mansfield, and N. Johnson. 2014. Mosquito cell lines: history, isolation, availability and application to assess the threat of arboviral transmission in the United Kingdom. Parasit Vectors 7.

Walsh, D. B., M. P. Bolda, R. E. Goodhue, A. J. Dreves, J. Lee, D. J. Bruck, V. M. Walton, S. D. O'Neal, and F. G. Zalom. 2011. Drosophila suzukii (Diptera: Drosophilidae): invasive pest of ripening soft fruit expanding its geographic range and damage potential. J Integr Pest Manag 2: G1-G7.

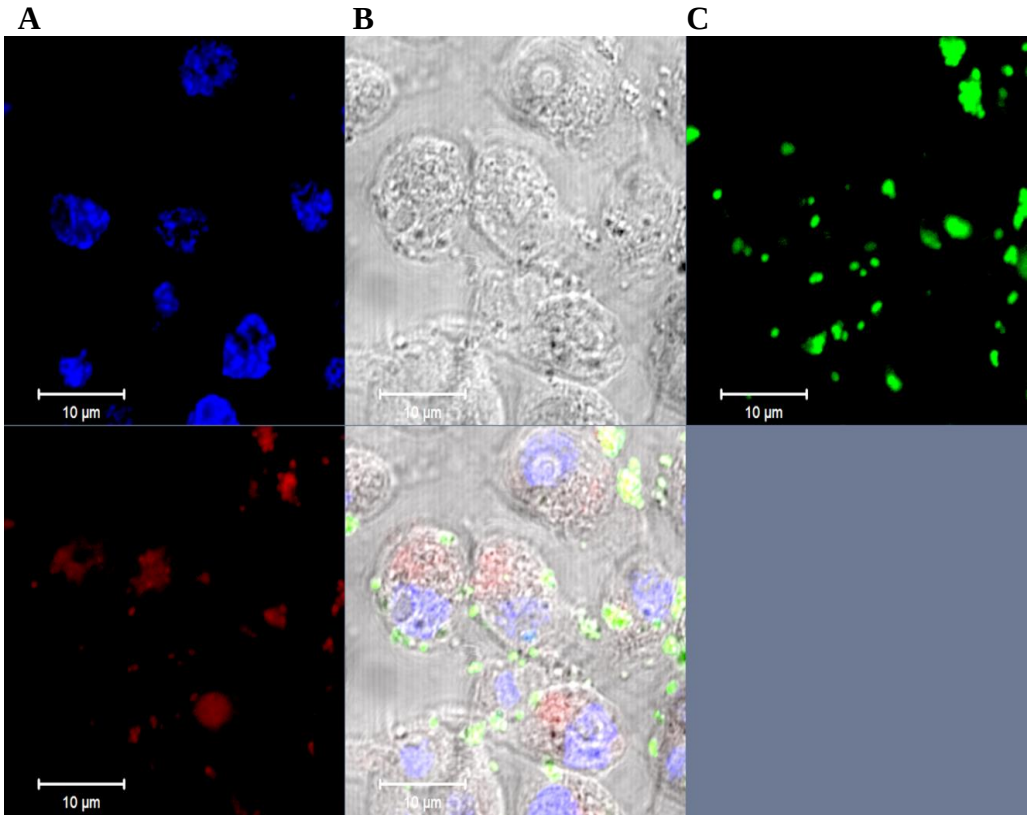
Wanjala, C. L., J. P. Mbugi, E. Ototo, M. Gesuge, Y. A. Afrane, H. E. Atieli, G. Zhou, A. K. Githeko, and G. Yan. 2015. Pyrethroid and DDT Resistance and Organophosphate Susceptibility among Anopheles spp. Mosquitoes, Western Kenya. Emerg Infect Diseases 21: 2178-2181.

Appendices

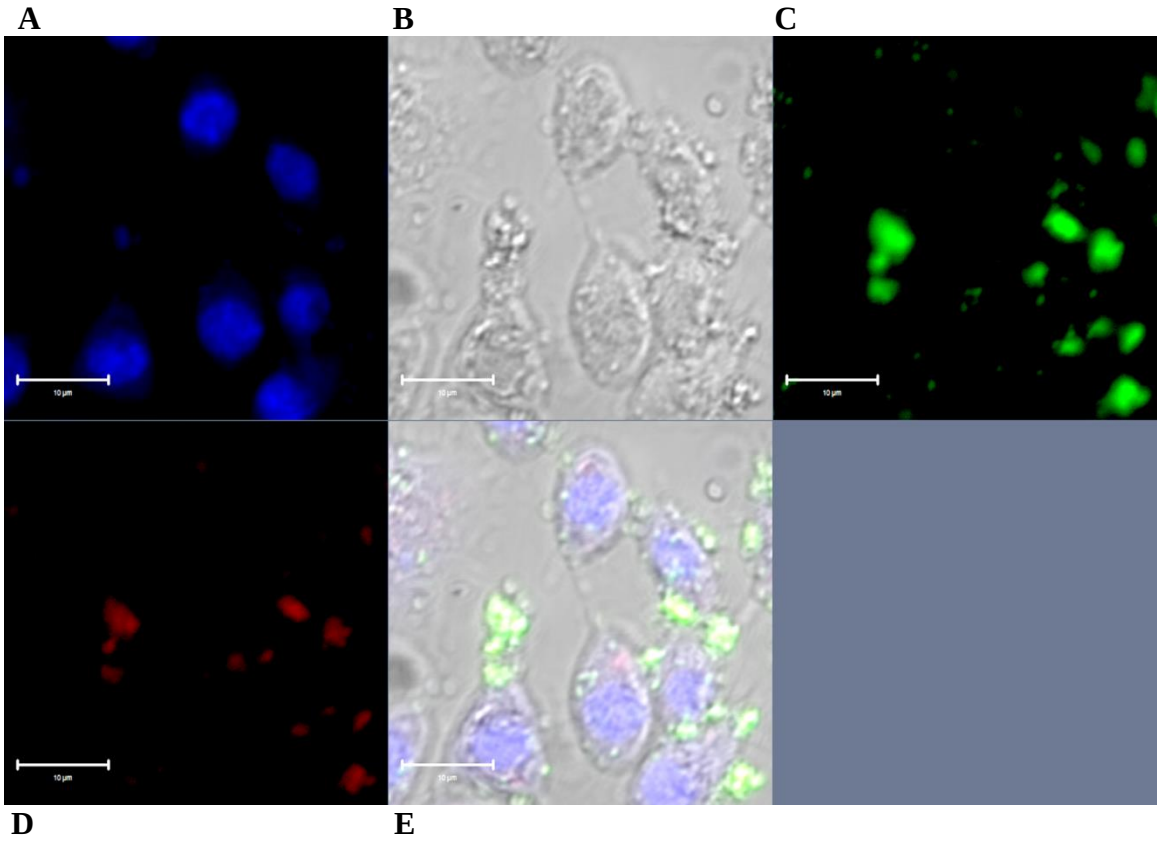
Chapter_2_Additional_file_2. Figure S1. Squared coefficient of Variation plot to assess cross replicate variability between control and treated SWD field and laboratory population. Field population: cypermethrin (A), spinosad (B) and malathion (C) treatment. Laboratory population: cypermethrin (D) spinosad (E) and malathion (F). CummeRbund was used for statistical analyses and generating plots.


A
genes
CV2
log10FPKM
sample_name
Control_P
Cypermethrin_P
B
genes
CV2
log10FPKM
sample_name
Control_P
Spinosad_P
C
genes
CV2
log10FPKM
sample_name
Control_P
Malathion_P
D
genes
CV2
log10FPKM
sample_name
Control_C
Cypermethrin_C
E
genes
CV2
log10FPKM
sample_name
Control_C
Spinosad_C
F
genes
CV2
log10FPKM
sample_name
Control_C
Malathion_C

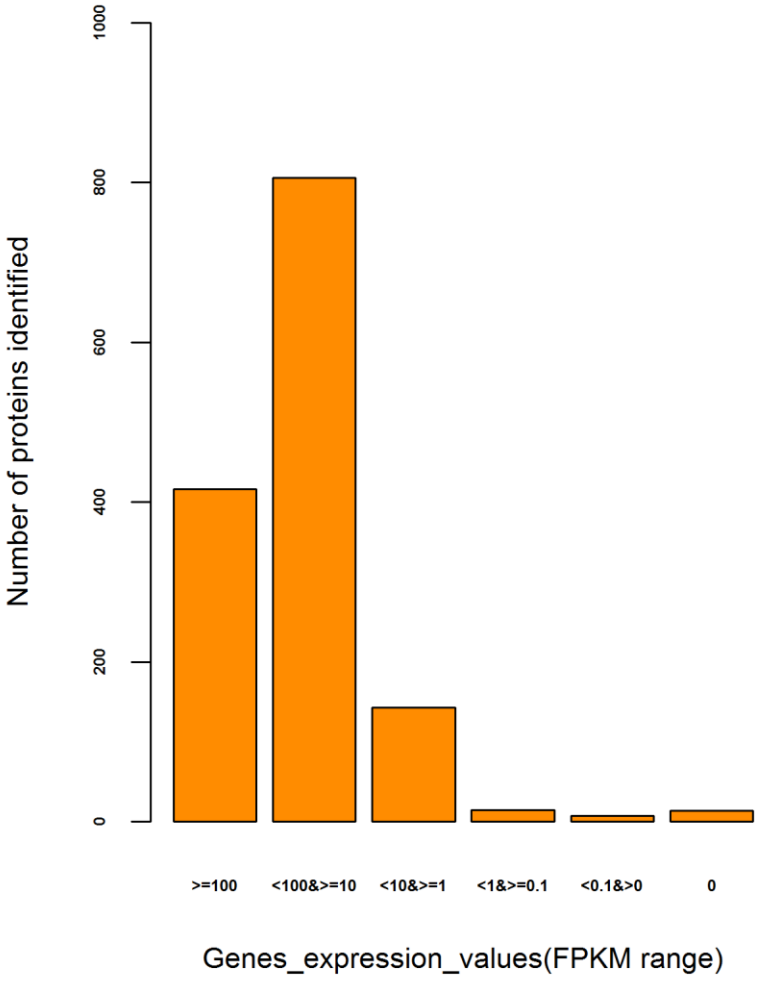
Chapter_3_Additional file_2. Figures S1. Confocal image showing Ag55 cells phagocytosing FITC-labeled E. coli bioparticles. A: Hoechst stain, B: Bright field, C: FITC-labeled E.coli bioparticles, D: Lysotracker deep red, E: Merged image. 63 X magnification. Fixed cell imaging.


A
B
C
10 µm
10 µm
10 µm
10 µm
10 µm
D
E

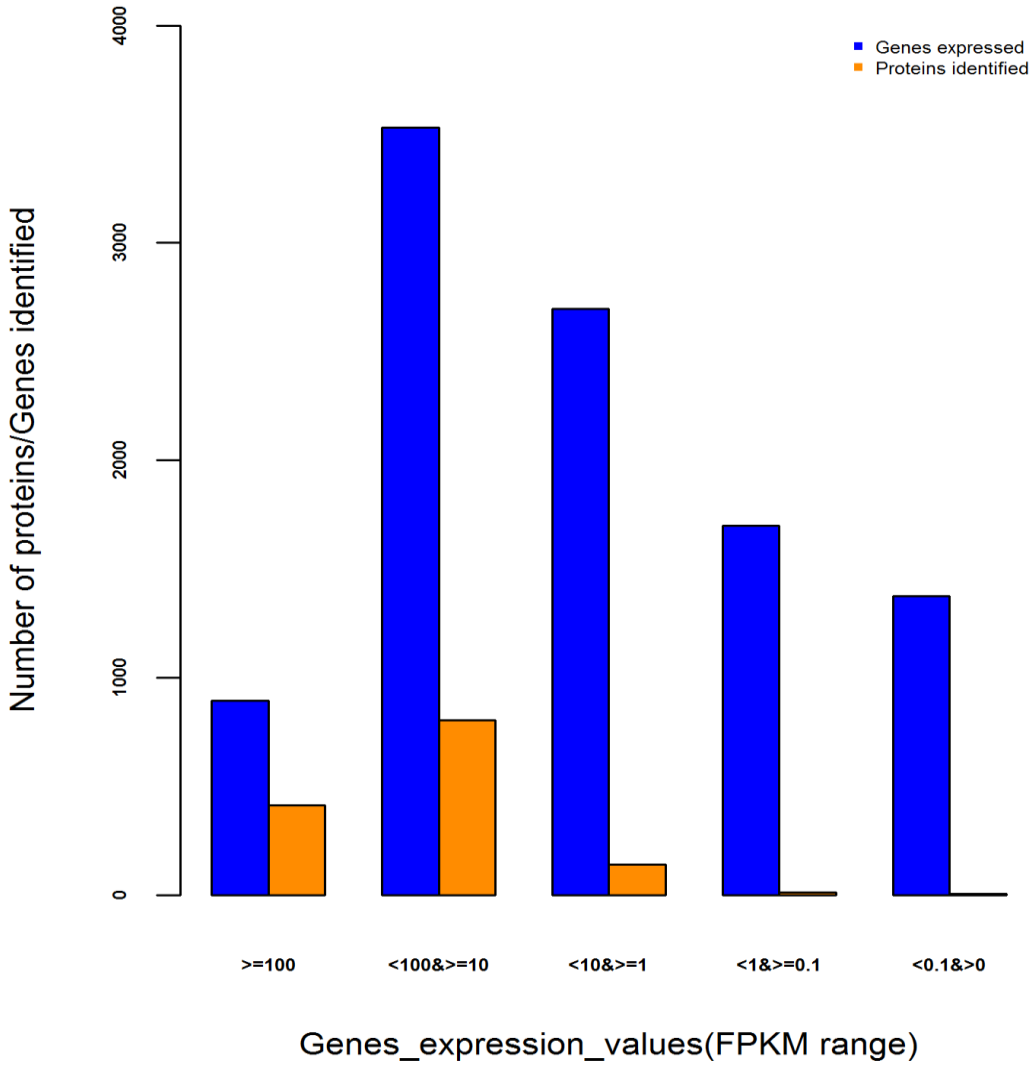
Chapter_3_Additional file_2. Figures S2. Confocal image showing Ag55 cells phagocytosing FITC-labeled E. coli bioparticles. A: Hoechst stain, B: Bright field, C: FITC-labeled E.coli bioparticles, D: Lysotracker deep red, E: Merged image. 63 X magnification. Fixed cell imaging.


A
B
C
D
E

Chapter_3_Additional file_9. Figures S3. Comparing the gene expression (in FPKM) to the proteins identified from the LC-MS/MS proteomic data


Number of proteins identified
0
200
400
600
800
1000
>=100
<100&>=10
<10&>=1
<1&>=0.1
<0.1&>0
0
Genes_expression_values(FPKM range)

Chapter_3_Additional file_10. Figures S4. Comparing the gene expression (in FPKM) to the proteins identified from the LC-MS/MS proteomic data


Genes expressed
Proteins identified
Number of proteins/Genes identified
0
1000
2000
3000
4000
>=100
<100&>=10
<10&>=1
<1&>=0.1
<0.1&>0
Genes_expression_values(FPKM range)