Characterization of the Regulatory Process of Pyrethroid Resistance in the House Fly, *Musca domestica*

by

Ming Li

A dissertation submitted to the Graduate Faculty of
Auburn University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Auburn, Alabama
December 13, 2014

Keywords: Insecticide resistance, Regulation, Transcriptome, Cytochrome P450, Carboxylesterase, *Musca domestica*

Approved by

Nannan Liu, Chair, Professor of Entomology and Plant Pathology
Joseph W. Kloepper, Professor of Entomology and Plant Pathology
Scott R. Santos, Associate Professor of Biological Sciences
Lee Zhang, Research Fellow of Entomology and Plant Pathology

Abstract

The house fly, *Musca domestica*, is a major domestic, medical and veterinary pest that causes more than 100 human and animal intestinal diseases. The major barrier in the house fly control is their remarkable ability to develop not only resistance to the insecticide used against them, but also cross-resistance to unrelated classes of insecticides. The house fly has demonstrated to be a useful model to study and predict resistance in not only themselves but also other insect species.

The current study generated the first reference transcriptome from the adult house fly and a whole transcriptome analysis was conducted for the multiple insecticide resistant strain ALHF (wild-type) and two insecticide susceptible strains: aabys (with morphological recessive markers) and CS (wild type) to gain valuable insights into the gene interaction and complex regulation in insecticide resistance of house flies. A total of 1316 genes were identified as being co-up-regulated in ALHF in comparison to both aabys and CS. The majority of these up-regulated genes fell within the three key detailed function categories: redox detailed function category in metabolism, signal transduction and kinases/phosphatases in regulation, and proteases in intra-cellular processes. Genetic linkage analysis with house fly lines comparing different autosomal combinations from ALHF revealed that the up-regulation of gene expression occurred mainly through the co-regulation of factors among multiple autosomes, especially between autosomes 2 and 5, suggesting that signaling transduction cascades controlled by GPCRs, protein

kinase/phosphates and proteases may be involved in the regulation of P450 and carboxylesterase gene expression.

To characterize the cytochrome P450 and carboxylesterase genes that play important roles in the pyrethroid resistance of house flies, 86 P450 and 26 carboxylesterase genes were selected based on our whole transcriptome analysis of the house fly to conduct the expression profile analysis in different house fly strains with different levels of permethrin resistance and autosome combinations. Our study showed that multiple P450 and carboxylesterase genes were co-up-regulated in insecticide-resistant house flies compared to -susceptible house flies, and the expression of these genes was regulated by *cis* or *trans* regulatory factors/genes, which were mainly on autosomes 1, 2 and 5. Transgenic expression analysis of selected P450 and carboxylesterase genes in *Drosophila melanogaster* demonstrated that elevated expression of these genes confers different levels of resistance to permethrin in the transgenic *Drosophila*. Homology modeling and permethrin docking analysis further suggested potential ability of these genes to metabolize permethrin.

Taken together, the study provides a global picture of P450 and carboxylesterase gene expression, regulation, autosomal interaction, and function in insecticide resistance of house flies, indicating multiple genes are co-responsible for detoxification of insecticides, and multiple mechanisms co-work on the development of insecticide resistance in house flies.

Acknowledgments

I am highly grateful to my major professor, Dr. Nannan Liu, for her guidance, encouragement, support, timely advice and numerous opportunities that she provided over these four years. Without her patience and help, this dissertation would never be accomplished. I feel very honored and fortunate to join Dr. Liu's program and have such a great advisor in my academic life. Her attitudes towards life, perspectives of science, and ways of doing research will influence me forever. I also would like to thank my committee members, Dr. Scott R. Santos, Dr. Joseph W. Kloepper, Dr. Lee Zhang and my dissertation outside reader, Dr. Yucheng Feng, for their invaluable encouragement, advice and kind help.

I also would like to give my many thanks to my laboratory colleagues, Dr. Chuanwang Cao, Dr. Lin He, Ms. Ting Li, Mr. Li Tian, Mr. William R. Reid, Mr. Feng Liu, Ms. Youhui Gong, Ms. Xuechun Feng, Mr, Zi Ye, Ms Fang Tang, Mr Zhou Chen and Ms Caixing Xiong for their friendly help, suggestions and kindness during the period of my Ph.D study. The support and help from other faculty members, staff and students in the Department of Entomology and Plant Pathology are also greatly appreciated.

Finally, the deepest gratitude goes to my parents, wife and sister for their endless love and support. I am forever grateful for their consistent understanding and encouragement.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

RNA-Seq	Ribonucleic acid sequencing
GO	Gene Ontology
BDRC	Bloomington Drisophila Resource Center
FPKM	Fragments per kilo base of gene for every million reads mapped
ANOVA	analysis of variance
Bp	base pair
ASPCR	allele specific PCR
BC1	back-cross generation 1
cDNA	complementary DNA
DDT	dichloro-diphenyl-trichloroethane
DEF	S,S,S-tributylphosphorotrithioate
DNA	deoxyribose nucleic acid
dsRNA	double stranded RNA
EMSA	Electrophoretic mobility shift assay
GABA	gamma aminobutyric acid
GST	glutathione S-transferase
JH	juvenile hormone
kdr	knockdown resistance

LC50	lethal concentration necessary to kill 50% of a test population
LD50	lethal dosage necessary to kill 50% of a test population
MFO	mixed function oxidases
qRT-PCR	quantitative real-time polymerase chain reaction
ORF	open reading frame
mRNA	messenger RNA
GAL4:UAS	Galactose 4: Upstream Activator Sequence
NONA	Non annotated
OP	organophosphate
P450	cytochrome P450 monooxygenase
PBO	piperonyl butoxide
PCR	polymerase chain reaction
RNA	ribose nucleic acid
RR	resistance ratio
RT-PCR	reverse transcription polymerase chain reaction
SNP	single nucleotide polymorphism
FDR	False discovery rate
MFO	mixed function oxidases
PBO	piperonyl butoxide
SCOP	Structural Classification of Protein

Chapter 1: Literature Review

1.1 Insecticide resistance

Humans have achieved significant progress in controlling insects and other arthropod pests in the last 50 years due to the introduction of synthetic pesticides, including insecticides, fungicides, acaricides and nematicides. However, the decline in the effectiveness of these new chemical weapons was observed very soon, and control failures were found in many cases, due to the development of resistance by the pests to these chemicals.

Insecticide resistance was defined by the World Health Organization (WHO) as “The development of an ability in a strain of some organisms to tolerate doses of a toxicant, which would prove lethal to the majority of individuals in a normal population of the same species” (WHO, 1957). Before exposure to an insecticide, some individuals already have appropriate factors that allow them to survive, and these factors can be passed to their offspring. So, resistance is a pre-adaptive phenomenon (Liu et al. 2006).

Resistance of insecticides is a common occurrence; the first report was documented by Melander in 1914 when San Jose scale demonstrated resistance to an inorganic insecticide (Melander 1914). Between 1914 and 1946, 11 additional cases of resistance to inorganic insecticides were recorded. In 1990, more than 1000 species of pests have developed insecticide resistance (Georghiou 1990). Insecticide resistance results in increased dosage and frequency of application, environmental contamination, disruption of wildlife and increased cost (Scott 1991).

1.2 How insecticide resistance develops

Insecticide resistance develops via the process of selection by insecticide. At first, only a very small population of pests can survive after exposure to the insecticide, but after many applications of the insecticides, the proportion of resistant individuals is increased, since the resistant individuals pass the resistance genes to their offspring (Leeper et al. 1986).

The rate of insecticide resistance development depends on several factors, including how rapidly the insects reproduce, the original level of insecticide resistance of insects, the migration and host range of insects, the persistence and specificity of insects, the rate, timing and number of applications of insecticide.

1.3 Cross resistance and multiple resistances

Resistance can develop to only a single insecticide or a class of insecticides. However, it is more common for insects to develop resistance to multiple insecticides. Insects that are resistant to many insecticides cause very serious issues.

1.3.1 Cross resistance

Many resistances are conferred by a single major genetic factor that differs between resistant and susceptible insects. When a single factor confers resistance to more than one insecticide, this is cross resistance (Tabashnik et al. 1997). In the majority of cases, the insect not only has resistance to the selecting insecticide, but it also often confers cross-resistance to other chemically-related compounds. Because a group of specific compounds usually shares a common

target site and mode of action in the insect (Casida 2009), it is easy to develop resistance to the same class of chemicals. When this happens, all of compounds in this class lose their efficiency to control pests. For example, the Colorado potato beetle, *Leptinotarsa decemlineata* (say) showed cross-resistance to imidacloprid and dinotefuranin in Long Island, New York (Mota-Sanchez et al. 2006).

The key point of cross resistance is a single mechanism responsible for resistance to more than one insecticide (Brattsten et al. 1986), and cross resistance can cause resistance to insecticides that never have been used.

1.3.2 Multiple resistance

Multiple resistance is the ability of insect populations to resist more than one insecticide class. For example, Huang and Han found that field population of common cutworm, *Spodopetera litura*, (Fabricius) in China developed high resistance to pyrethroids, organophosphates and carbamates (Huang and Han 2007). The multiple resistance mechanism is acquired by at least two separate genetic modifications. In 2007, Corbel found that the *Anopheles gambiae* and *Culexquin quefasciatus* mosquitoes in West Africa showed a high frequency of resistance to permethrin and DDT due to target site insensitivity and increased metabolic enzyme activity (Corbel et al. 2007).

1.4 Insecticide resistance stability

Insecticide resistance can be either stable or unstable in the field, depending upon many factors, including the pest, the chemical, and the agricultural system that is used. A stable

resistance occurs when an insecticide is used and resistance does not decline appreciably after stopping the insecticide treatment. An unstable resistance similarly increases in response to pesticide treatments, but decreases in frequency during intervals when the pesticide is no longer used (Brévault et al. 2008).

1.5 Insecticide resistance mechanisms

The development of insecticide resistance is well documented as a major problem today. Some species of insects can tolerate many kinds of insecticide families, causing chemical control to become almost useless. So, the question is, how does insecticide resistance happen? Several mechanisms are involved in insecticide resistance development.

1.5.1 Behavioral resistance

Insects may have inherited the tendency to stay on the undersides of leaves and not venture to the upper leaf surfaces where pesticide deposition may be much greater or there may be a shift in behavior to avoid exposure to insecticides. Behavioral resistance occurs if insects are irritated or repelled by the insecticide (Georghiou 1972). This type of resistance is stimulus-dependent behavior avoidance. For stimulus-dependent behavior resistance to occur, insects must have had some type of contact with the insecticide or insecticide treatments.

Repellency is the ability of insects to detect certain repellent components of an insecticide (Fradin and Day 2002). Repellency occurs if insects are able to detect the insecticide molecules in the air before actually contacting the insecticide or the toxic substance and then avoid it (Castillochavez et al. 1988). One example of increased irritability is documented in the mosquito

Anopheles atroparvus. After 10 generations of selection by DDT, a strain of *A. atroparvus* mosquito possesses high irritability, enabling them to escape DDT residues more quickly than the normal strain (Gerold and Laarman 1964).

Another typical case of this mechanism was documented in the state of Georgia, where fly control using applications of malathion-sugar baits failed (Kilpatrick and Schoof 1958). It was observed that the flies obtained a high level of repellency to malathion (organophosphate insecticide); thus, most of the flies did not alight on the treated sugar.

Stimulus-independent behavior avoidance or resistance occurs if insects naturally avoid areas treated with insecticides. In stimulus-independent behavioral resistance, insects increase their chances of survival by moving to areas that are insecticide free. In this type of resistance, the avoidance of insects is not due to irritant or repellent effects of the insecticide (Guedes et al. 2009).

1.5.2 Physiological resistance

Insects have developed physiological modification mechanisms to help survive the presence of insecticides. Physiological resistance is any form of resistance that reduces toxicity through changes in the basic physiology of insects. In this form of resistance, the chemical is not broken down into a less toxic form instead the insect accommodates the chemical by altering one or more physiological functions. These mechanisms include decreased penetration rate, increased sequestration storage, and accelerated excretion of insecticides (Liu et al. 2006).

1.5.2.1 Reduced cuticle penetration rate

Insecticides must pass through the cuticle of the insect and reach the target site before lethal effects can occur. Reduced penetration rate of insecticides through the insect cuticle was considered as one aspect of increased insecticide resistance, since the amount of insecticide getting into the body of the insect or working on the target was reduced and became easier to digest by detoxification enzymes (Georghiou 1972). The decreased penetration rate has been found in a number of insects. A study of the house fly, *Musca domestica*, indicated that the penetration rate of 14C-labelled fipronil was slower in resistant strains than susceptible strains (Wen and Scott 1999). Other studies including the beet armyworm, *Spodoptera exigua* (Hubner) (Liu and Shen 2003) and the German cockroach, *Blattella germanica* (L) (Anspaugh et al. 1994), also found evidence of penetration resistance. The gene causing this resistance has been named *pen* (Farnham 1973). Any membrane has the potential ability to serve as a penetration barrier, however, the molecular basis for this mechanism is not clear (Liu et al. 2006; Scott 1991).

1.5.2.2 Increased sequestration storage and accelerated excretion of insecticides

Sequestration and storage as resistance mechanisms have been studied in many insect species. One example is from *C. quinquefasciatus* whose the major role of carboxylesterase A2 is to sequestrate the organphosphate insecticide (Ketterman et al. 1992). Another typical example comes from the peach potato aphid, *myzus persicae*. The resistance induced by carboxylesterase E4 in this specie is not mediated by hydrolysis, but by the storage and sequestration of a substantial proportion of a toxic dose of insecticides (Devonshire and Moores 1982). Glutathione S-transferase also offers protection against insecticides by binding to the molecules in a sequestering manner (Kostaropoulos et al. 2001).

In many insect species, a large amount of the accumulated toxic compound can be excreted or lost with the shell during molting (Zagrobelny et al. 2004). Toxic compounds also can be sequestered and subsequently used as a defensive substance against predators or pathogens (Ode 2006).

Some insects have developed selective transport and storage abilities that prevent toxins from interfering with the physiological processes of the insect (Després et al. 2007). In 2004, Kuhn reported that in studies of the molecular basis of sequestration in leaf beetles, which were feed with structurally different thioglucosides resembling natural O-gulcoside. The accumulated gulcoside in the defensive systems indicated that the larvae possess transport systems, which have evolved to adapt to the glycosides, and minor structural modifications in the aglycon can lead to drastically reduced transport rates in leaf beetles (Kuhn et al. 2004).

Wide varieties of insects have developed insecticide resistance through accelerated excretion mechanism. For example, diazinon excretion was faster in resistant thrips, *Frankliniella occidentalis* than in the susceptible strain (Zhao et al. 1994). An increased excretion rate was also observed in the carbaryl-resistant population of Western corn rootworm, *Diabrotica virgifera* (LeConte) (Scharf et al. 1999).

1.5.3 Increased activities of detoxification enzymes

Increased metabolic detoxification by enzymes is one of the most common mechanisms involved in insecticide resistance (Hemingway 2000). The metabolism of xenobiotics is a multi-step process in which lipophilic compounds are converted into more water-soluble metabolites, which can be excreted by the excretory system of insects. Three major enzyme systems play very

important roles in insecticide detoxification. They are cytochrome P450 monooxygenases (P450s), carboxylesterases (COEs) and glutathione *S*-transferases (GSTs).

1.5.3.1 Cytochrome P450 monooxygenases

Cytochrome P450s are one of the largest super gene families. They are important hemoproteins involved in monooxygenases systems, playing an essential role in the biosynthesis and metabolism of juvenile hormones and ecdysteroids, important for insect growth, development and reproduction (Feyereisen 1999; Scott and Wen 2001), as well as in the metabolism of xenobiotic substances such as drugs, pesticides and plant toxins (Feyereisen 2006).

Cytochrome P450 derives its name from an absorption peak at 450 nm when it is reduced and saturated with carbon monoxide. In the late 1980s, studies of P450 nomenclature were conducted, and a standardized P450 nomenclature system was founded which facilitated classification of the isolated cytochrome P450s (Nebert et al. 1987). Genes are designated with the abbreviation CYP, followed by a numeral for the family, a capital letter indicating the subfamily, and another numeral for the individual gene. All the members of the CYP families share >40% amino acid identity, and members of a subfamily must share >55% identity at the amino acid level. Genes are described in italics, and the gene product, mRNA or enzyme, is described in capitals (Feyereisen 1999).

P450 monooxygenase detoxification is an extremely important mechanism for insects survive an attack by insecticides. The P450 monooxygenase system includes three important components: cytochrome P450, which acts as the substrate binding protein, cytochrome P450 reductase, which transfers electrons from nicotinamide adenine dinucleotide phosphate-oxidase

(NADPH) to cytochrome P450, and cytochrome b5, which also can transfer electrons from NADH to cytochrome P450. The most common catalytic reaction of cytochromes P450 is monooxygenase reaction, which inserts one atom of oxygen into an organic substrate, and the other oxygen atom is reduced to water. The catalytic cycle is as follows: The oxidized form (P450 Fe^{III}) protein binds the substrate first, and then the P450-substrate complex receives a single electron from NADPH reductase and binds oxygen, another electron is transferred from cytochrome b5 and reduces dioxygen to a negatively charged peroxy group. The peroxy group is rapidly protonated twice, releasing one molecule of water, forming a highly reactive iron (V)-oxo species, and then the substrate in the active site reacts with the highly reactive iron (V)-oxo species, releasing a hydroxylated product. The enzyme returns to its original state with a water molecule returning to occupy the distal coordination position of the heme iron (Feyereisen 1999).

To date, more than 2000 P450 genes have been identified from various insects. The number of cytochrome P450s in sequenced genomes ranges from 37 in *Pediculus humanus* (Lee et al. 2010), to 204 in the *C. quinquefasciatus* (Yang and Liu 2011). They are assigned to four major clades: CYP2 clade, CYP3 clade, CYP4 clade and Mitochondrial CYP clade (Feyereisen 2006; Scott 1999). The role of P450 monooxygenases in insecticide resistance was first reported in the early 1960s when the study found that the resistance of house flies to carbaryl could be eliminated by the P450 inhibitor sesame (Eldefrawi et al. 1960). Since then, evidences supporting monooxygenase-mediated resistance have accumulated rapidly.

The insect P450 was first detected in the house fly by Ray in 1967 (Ray 1967), and the first insect P450 reductase was also purified and characterized from the house fly (Wilson and Hodgson 1971). In order to study the function of cytochrome p450, many attempts were made to

purify cytochrome p450s. The first insect P450 protein, CYP6A1, came from a diazinon-resistant house fly strain, Rutgers, with the expression of CYP6A1 in the Rutgers strain about 10-fold higher than in susceptible strains (Carino et al. 1994; Feyereisen et al. 1989).

The relationship and interactions of p450 with p450 reductase and cytochrome b5 have been well described (Guzov et al. 1996; Murataliev et al. 1999). It is reported that the level of P450 reductase and cytochrome b5 are increased in P450-mediated resistance (Scott and Georghiou 1986a), but genetic analysis suggests that b5 might play a role in P450-mediated resistance while P450 reductase might not (Liu and Scott 1996).

P450 monooxygenases play various functional roles in metabolism, and different expression levels have been detected in different tissues of insects. High levels of cytochrome p450 are usually detected in the midgut and fat bodies, which serve as the first defense system for absorbing foreign compounds, such as insecticides or plant chemicals (Scott 1999). The expression level of p450 also varies in different development stages; generally, cytochrome p450 cannot be detected in eggs and pupae stages, but are highly expressed in adults (Scott 1999).

The P450s showed great variation in response to inducers or repressors, which can be either endogenous compounds such as ecdysone, or xenobiotics such as insecticides. Phenobarbital (PB) is a well-studied inducer in insects. Several insect P450s coming from different species can be induced by PB, such as *CYP4D10*, *CYP6A1*, *CYP6A2*, *CYP28A1*, *CYP28A2*, and *CYP28A3* in *Drosophila*, and CYP6D1 in both the house fly and mosquito (Brun et al. 1996; Danielson et al. 1998; Dunkov et al. 1997; Jacob 1983; Lee and Scott 1989; Liu and Scott 1997). Piperonylbutoxide (PBO) is an organic compound used as pesticide synergist. It is a P450

inhibitor, which can decrease the metabolism of P450 enzymes. When house flies were treated with PBO, the resistance was reduced dramatically (Liu and Yue 2000).

Over expression of the cytochrome p450 genes in insecticide resistant insects is a common phenomenon. In 1969, research found that the level of p450 monooxygenase in insecticide-resistant strain of house flies was higher than in susceptible strains (Plapp and Casida 1969). Later research indicated that the resistance might be due to the elevation of only select cytochrome p450s and that this may not result in a significant increase in total cytochrome p450s (Fonseca-Gonzalez et al. 2009). *CYP4G8*, a P450 gene in Australian cotton bollworm, *Helicoverpa armigera* (Hubner), was found to be at a 2-fold higher level in a pyrethroid resistant strain compared to a susceptible strain (Pittendrigh et al. 1997). *CYP6F1* from the southern house mosquito, *C. quinquefasciatus* (Say), was expressed 3 times higher in a permethrin-resistant strain than in the susceptible strain (Kasai et al. 2000). *CYP6A2* was expressed 20 to 30-fold higher in a malathion-resistant strain of the fruit fly when compared to a susceptible strain (Waters et al. 1992). The study of two cytochrome P450s, *CYP6A36* and *CYP6A5v2* in house flies, revealed that they were overexpressed in a permethrin-resistant strain (Zhu et al. 2008a; Zhu and Liu 2008). *CYP6BQ9* showed more than a 200-fold higher expression in the deltamethrin-resistant *Tribolium castaneum* strain compared to the susceptible strain, and functional studies showed that *CYP6BQ9* confers deltamethrin resistance (Zhu et al. 2010). *CYP6P9a* and *CYP6P9b* in *Anopheles funestus* showed that elevated expression of either of these genes confers resistance to both type I and type II pyrethroids (Riveron et al. 2013).

Point mutations of P450 genes also have been proved to be involved in insecticide resistance. Three point mutations R335S, L336V and V476L of *CYP6A2* were responsible for

DDT resistance in the RDDT^{R} strain of *D melanogaster* (Amichot et al. 2004). Single substitution in *CYP51* of *Candida albicans* (T315A) and of *Uncinula necator* (F136Y) played a role in resistance to the fungicides fluconazole and triadimenol, respectively (Delye et al. 1998; Perea et al. 2001).

Besides detoxification, P450 enzymes can activate insecticides to more toxic substrate, for example, the activation of oxygen substituting sulphur on the double bond with phosphate (P=S to P=O) by P450 enzymes can increase the activity of organophosphorus (OP) (Sams et al. 2000).

Cytochrome P450 monooxygenase can metabolize a wide variety of substrates and induce cross resistance to unrelated compounds. Therefore, P450s-mediated detoxification is a very important resistance mechanism (Liu et al. 2006; Scott 1991; Scott and Georghiou 1986b).

1.5.3.2 Esterases and hydrolases

Hydrolases-or esterases-mediated detoxification has been found in many organisms, including carboxylesterases and phosphorotriester hydrolases. Esterases and hydrolases are involved in the metabolism of organophosphates, carbamates, pyrethroids and juvenile hormones (Kerkut and Gilbert 1985; Scott 1999). The hydrolase-mediated insecticide resistance mechanisms can be divided into two groups: quantitative changes and qualitative changes.

1.5.3.2.1 Quantitative changes

One mechanism of esterase-based insecticide resistance is the elevation of the expression of esterase. Esterase is observed to be overproduced in resistant insect strains, resulting in an increase of the amount of available esterase for sequestering penetrated insecticides in a unit of

time, thus reducing the amount of pesticide reaching the target. A study found a 64-fold amplification in resistant green peach aphid, *M. persicae* esterase gene, *E4* compared to a susceptible strain (Devonshire and Moores 1982). 250-fold amplification of B1 esterase and co-elevation of A2/B2 esterase were detected in the insecticide resistant Southern house mosquito *C. quinquefasciatus* (Vaughan and Hemingway 1995).

1.5.3.2.2 Qualitative changes

Another possible mechanism of esterase-based insecticide resistance is called non-elevated esterase mechanism, which has been documented to confer resistance to organophosphates in some species. In this mechanism, mutations of amino acids in carboxylesterase enhance the hydrolase activity of the enzyme, thus enhancing metabolism of insecticides (Russell et al. 2004). One example is the amino acid substitution from glycine to aspartic acid (Gly137 to Asp), which enhances the hydrolase activity in *ali*-esterase in the blowfly, *Lucilia cuprina* and in the housefly *M. domestica* (Campbell et al. 1998). Not all esterases have high activity in insecticide resistance strains; lower esterase activity was also founded in the OP-resistant blowfly, *Chrysomya putoria* (Wiedemann) (Townsend and Busvine 1969) and the parasitoid *Habrobracon hebetor* (Say) (Perez-Mendoza et al. 2000).

1.5.3.3 Glutathione S-transferases

Glutathine S-transferases (GSTs) are one of the major detoxification enzymes in both invertebrates and vertebrates having a broad range of substrate specificities (Jakoby 1978). With this, they play a significant role in the detoxification of a wide range of xenobiotics, including

metabolizing insecticides and catalyzing dehydrochlorination of insecticidal molecules or facilitating their conjugation with reduced glutathione (GSH) (Salinas and Wong 1999).

There are 6 different classes of GSTs: Delta, Sigma, Epsilon, Omega, Theta and Zeta, which have been identified in insects (Enayati et al. 2005). It is reported that the increased insecticide resistance of insects results from over expression of one or more GST genes (Sonoda and Tsumuki 2005). For examples, increased expression of GSTs confers resistance to DDT in *A. gambiae* (Ortelli et al. 2003; Prapanthadara et al. 1996), and the over expression of the Epsilon class of GSTs in resistant *A .gambiae* is regulated by both cis- and trans-acting factors (Ranson et al. 2000). The enhanced activity of GSTs results in resistance to permethrin in the brown plant hopper, *Nilaparvata lugens* (Vontas et al. 2002b), and the high activity of GSTs provides resistance against organophosphates in the diamond back moth *Plutella xylostella* (Chiang and Sun 1993).

1.5.4 Reduced sensitivity of target sites

Alteration of amino acids responsible for the insecticide binding site of action, results in the insecticide become less effective or even ineffective is another important mechanism that involved in insecticide resistance. The target of organophosphorus (e.g., malathion, fenitrothion) and carbamate (e.g., propoxur, sevin) insecticides is acetylcholinesterase in nerve synapses, and the target of organochlorines (e.g., DDT) and synthetic pyrethroids (e.g., permethrin, fenpropathrin) is the sodium channels of the nerve sheath.

1.5.4.1 Insensitivity of acetylcholinesterase

Acetylcholinesterase (AChE) is a serine esterase that terminates nerve impulses at cholinergic synapses by breaking down the neurotransmitter acetylcholine (Walsh et al. 2001). In insects, acetylcholinesterase is essential for life, and the inhibition by organophosphorus or carbamate insecticides is lethal. Intensive use of these insecticides over the past 60 years has led to the development of resistance in many target species that are detrimental to agriculture or serve as vectors of human and animal diseases (Georghiou 1990). The mutant forms of AChE have been characterized and showed widely differing spectra of insensitivity between different species or different compounds within a species (Byrne and Devonshire 1997; Devonshire and Moores 1984; Feyereisen 1995; Mutero et al. 1994). Point mutations are usually accompanied by a modification of the kinetic parameters of acetylcholine hydrolysis, which have been closely identified as being responsible for insecticide resistance (Walsh et al. 2001; Zhu et al. 1996).

Many potential resistance-associated point mutations may result in AChE insensitivity without significantly reducing AChE catalytic efficiency (Chen et al. 2001; Villatte et al. 2000; Vontas et al. 2002a). Insensitivity of AChE caused by multiple point mutations in structural genes have been found in different species, for examples, *D. melanogaster*, *M. domestica* , *C. pipiens*, and *A. Gambiae*, diamondback moth *Plutella xylostella*, Colorado potato beetle *L. decemlineata* and so on (Lee et al. 2007; Mutero et al. 1994; Russell et al. 2004; Zhu et al. 1996). These mutations were involved in insecticide resistance.

1.5.4.2 The GABA receptor mutation

The gamma-amniobutryric acid (GABA) receptor is a heteromultimeric gated chloride-ion channel, which is a widespread inhibitory neurotransmission channel of central nervous system

and neuromuscular junctions (Bermudez et al. 1991). GABA is released in order to wave the excitation and bind with receptors on the postsynaptic membrane, enhancing permeability to chloride ions resulting in more chloride ions with negative charge being allowed to enter the postsynaptic membrane, inducing hyperpolarization.

The GABA receptor consists of five subunits, each subunit containing a large extracellular agonist-binding N-terminal domain, and four transmembrane domains (M1–M4). The GABA receptors are the targets of numerous insecticides, including cyclodienes, trioxabicycloctanes, and pirotoxinin. These insecticides are thought to bind the pore formed by transmembrane domain M2 (Miller 1989) and inhibit the flow of chloride ions through the receptor channel complex (Le Goff et al. 2005). It was reported that an A2′S mutation in the M2 confers resistance to cyclodienes in *D. melanogaster* (ffrench-Constant et al. 1993). Two amino acid substitutions (A2′S and A2′G) have been reported to be associated with cyclodiene resistance in the aphid, cat flea, mosquito, and so on (Anthony et al. 1998; Bass et al. 2004; Du et al. 2005; Miyazaki et al. 1995; Thompson et al. 1993).

1.5.4.3 Mutations in the voltage-gated sodium channel

The voltage-gated sodium channel is a trans-membrane protein complex, which is responsible for the initiation and propagation of action potentials in almost all excitable cells. It is well known as the primary target of DDT and modern synthetic pyrethroids, which are structural derivatives of pyrethrins (Narahashi and American Chemical Society. 1979).

The structure of voltage-gated sodium channel contains α subunit and one or more auxiliary β subunits (Yu and Catterall 2003). The α subunit contains the functional domains and is believed

to be responsible for the main sodium currents when expressed in *Xenopus* oocytes without the assistance of β subunits (Catterall 2000). The α subunit is composed of four homologous domains (I to IV), each domain consisting of six transmembrane segments (S1 to S6). It has been demonstrated that S4 is the most conserved, thus making it the most likely to be involved in the voltage-sensing mechanism (Mannikko et al. 2002). The α subunit is also believed to play a role in regulating the channel gating or protein expression (Zlotkin 1999).

Multiple point mutations were found in insect sodium channels of pyrethroid-resistant insect populations (Dong 2007). In 1996, the complete coding sequence of the insect voltage-gated sodium channel in the house fly was identified (Ingles et al. 1996). Comparison of the coding sequences of a susceptible strain, knock down resistance (*kdr*) and super-*kdr* strains revealed two consistent point mutations associated with resistance; L1014F and M918T, which occur at the 1014 and 918 amino acid residues respectively. The L1014F, also called L-to-F mutation, was identified in all cases of resistance, while the M918T occurred only in the super-*kdr* phenotype.

The L1014F mutation has also been identified in other species, including malaria mosquito *A. gambiae*; German cockroach *B. germanica*; Colorado potato beetle *L. decemlineata*; the horn fly *Haematobia irritans*; the diamondback moth *Plutella xylostella*; the green peach aphid *M. persicae*, the common house mosquito *C. Pipiens* and *C. quinquefasciatus* (Guerrero et al. 1997; Lee et al. 2000; Martinez-Torres et al. 1998; Martinez-Torres et al. 1999a; Martinez-Torres et al. 1999b; Schuler et al. 1998; Soderlund and Knipple 2003; Xu et al. 2006a; Xu et al. 2006b).

L1014F and M918T are not the only two mutations on the sodium channel that are associated with knockdown resistance. There are more than 20 resistance-associated mutations on the sodium channels have been reported, and at least 10 mutations have been functionally

expressed in *Xenopus* oocytes and were confirmed to be associated with *kdr* resistance (Dong 2007; Zlotkin 1999).

1.6 Regulatory gene and gene expression regulation

The genome of an organism usually contains thousands of different genes. Some of the gene products are required by the cell under normal conditions, and other gene products are required under specific conditions (Hughes 1999). Regulation of gene expression by extracellular or intracellular signals is a fundamental mechanism of organisms to develop and adapt to the environment (Struhl 1999; Workman and Kingston 1998).

Chemicals, growth factors and neurotransmitters are all able to alter the patterns of gene expression in a cell (Nestler et al. 2001). Mechanisms that are involved in the control of gene expression include structural changes in the chromatin to make a particular gene accessible for transcription, transcription of DNA into RNA, splicing of RNA into mRNA, editing and other covalent modifications of the mRNA, translation of mRNA into protein, and, finally, post-translational modification of the protein into its mature, functional form (Fickett and Wasserman 2000; Struhl 1999; Workman and Kingston 1998).

1.6.1 Regulatory gene

A regulatory gene or regulator gene is a gene that determines the expression of other genes. Some regulatory genes can only regulate a specific gene, while others can control the expression of several genes (Dickinson 1979). A regulatory gene can be an encode protein or a RNA, which works at the RNA level, such as microRNA. The regulatory gene can act as an activator, turning a

gene or group of genes on, which increases the rate of transcription. The regulatory genes also can act as repressors, turning some genes off so that they cannot express or decrease the expression of target genes (Bellí et al. 1998). Regulatory genes are produced by all organisms, and are very important to the system of the organisms. They help check, balance and moderate genetic expression of genes so that an organism operates different systems efficiently (Desvergne et al. 2006).

Regulatory genes contain information that is used to code proteins. First, they are transcribed into RNA, which is used to build the protein by the cells. Once the protein is created, they can act in a variety of ways on the genetic material inside the cell. For example, when a mosaic virus is surrounded by antibiotics, an activator in a bacterium activates a gene for antibiotic resistance (Walsh et al. 1996). The regulatory gene also can make a protein which locks onto a section of DNA so that it cannot be transcribed (Freeman and Bassler 1999). In multicellular organisms, regulatory genes are also involved in the process of cell differentiation, development, proliferation and apoptosis, determining cell function. This allows organisms to have great diversity in cell types (Müller et al. 2001).

1.6.2 Regulation of gene expression

Regulation of gene expression is the process that influences the differential control of gene action at the level of transcription or translation by nuclear, cytoplasmic, or intercellular factors. This process includes gene activation and genetic induction. Any step of the gene expression can be modulated, from DNA-RNA initiation and transcription to the post-translational modification of a protein (Calkhoven 1996).

Gene regulation is necessary for all organisms as it increases the versatility and adaptability by allowing cells to express proteins when they are needed. Gene regulation also governs the processes of cellular differentiation and morphogenesis, leading to the creation of different cell types in multicellular organisms where the different types of cells may possess different gene expression profiles though they all possess the same genome sequence (Hogeweg 2000).

1.6.2.1 Regulation of gene expression by extracellular signals

Most genes contain response elements that respond to physiologic signals. These elements can bind with transcription factors that are activated or inhibited by specific physiologic signals. For example, activation of neurotransmitter, hormone, or neurotrophic factor receptors by physiologic signals lead to the activation of specific second messenger and protein phosphorylation pathways, and regulate expression of genes. This can be accomplished by phosphorylation of protein kinases (Marshall 1995).

There are two general mechanisms of transcriptional regulation by extracellular signal: 1 Transcription factors that are present at high levels are rapidly activated by signaling cascades to activate or repress transcription of target genes. 2 Transcription factors that are expressed at low levels can be induced by physiologic signals and can regulate expression of a series of genes (Nestler et al. 2001).

1.6.2.2 Regulation of gene expression by the structure of chromatin

In eukaryotic cells, chromosomes are extremely long molecules of DNA, which are wrapped around histone proteins to form nucleosomes, which are the major subunits of chromatin

(Workman and Kingston 1998). Chromatin can inhibit access of transcription factors to the DNA and repress gene expression. The regulation of gene expression within euchromatin requires the delivery of enzymes by DNA-bound transcription factors, which bind to the promoter of specific genes and initiate a cascade of modification events. Acetylation, methylation, phosphorylation and ubiquitination have been implicated in gene expression activation. Methylation, ubiquitination, sumoylation, deimination and proline isomerization have been thought to be involved in the repression of gene expression (Kouzarides 2007; Nestler et al. 2001).

1.6.2.3 Regulation of gene expression by transcription

Transcriptional regulation is the process in gene expression levels through the altering of transcription rates. Transcription occurs when particular activator proteins displace nucleosomes. This allows several transcription factors to bind with DNA at the core promoter region and recruit RNA polymerase. When transcription occurs, the amount of RNA produced is controlled by regulatory genes. Transcription of a gene by RNA polymerase can be regulated by at least five mechanisms:

1. Repressors regulate the expression of one or more genes by binding to the operator and blocking the attachment of RNA polymerase to the promoter, thus impeding the expression of the gene (Matthews and Nichols 1997). A good example is the lac operon of the bacteria *Escherichia coli*. The lac repressor is constitutively expressed and always bound to the operator region of the promoter, interfering with the ability of RNA polymerase to bind to the promoter and transcribe the lac operon (Bell and Lewis 2001).

2. Gene expression is regulated by specificity factors. Specificity factor is a protein that directs another protein to recognize or bind to other proteins. Altering the specificity of RNA polymerase for a given promoter or set of promoters, will make it more or less likely to bind to them. For example, sigma factor is a prokaryotic transcription initiation factor that enables specific binding of RNA polymerase to gene promoters (Gruber and Gross 2003). Mitochondrial transcription specificity factors (TFB1M and TFB2M) markedly enhance mtDNA transcription in the presence of mitochondrial RNA polymerase (Gleyzer et al. 2005).

3. General transcription factors (GTF's) or basal transcription factors are protein transcription factors which are involved in the formation of a transcription complex. When combined with RNA polymerase, they help RNA polymerase to load at the start of a protein-coding sequence and then release the polymerase to transcribe the mRNA (Deaton and Bird 2011). For example, TATA binding protein (TBP) is a GTF that binds to the TATAA box, the motif of nucleic acids that is directly upstream from the coding region in all genes, thus facilitating the binding of the transcription complex to nucleosomal DNA (Imbalzano et al. 1994).

4. Activators enhance the interaction between RNA polymerase and a particular promoter, increasing the expression of the gene. Activators increase this interaction through a connected domain that assists in the formation of the RNA polymerase holoenzyme, or they may operate through a co-activator, which binds the DNA-binding activator and contains the domain assisting holoenzyme formation. A particular activator may bind one or more specific co-activators (Brahms et al. 1985). The attraction of RNA polymerase for the promoter can be through direct interactions with subunits of the RNA polymerase or indirectly by changing the structure of the DNA. For example, the lac repressor always binds to the operator region of the promoter in the

bacteria *Escherichia coli*, interfering with the ability of RNA polymerase to bind to the promoter and transcribe the lac operon (Lin and Riggs 1975). In the presence of lactose, which acts as an activator, the repressor changes conformation and falls off the operator and RNA polymerase is able to bind to the promoter (Eschenlauer and Reznikoff 1991).

5. Gene expression is regulated by enhancers. Enhancers are short regions of DNA or sites on the DNA helix that can be bound to proteins to enhance transcription levels of genes. Enhancers can be located upstream or downstream of the gene that it regulates (Hines et al. 2004). An enhancer does not need to be located near to the transcription initiation site to affect the transcription of a gene, as some have been found to bind several hundred thousand base pairs upstream or downstream of the start site (Fraser et al. 1991). Enhancers do not act in the promoter region itself, but are bound by activator proteins.

1.6.2.4 Post-transcriptional regulation

Post-transcriptional regulation is the process that controls gene expression after mRNA is formed, between the transcription and the translation of the gene (Alberts 2008). This process is involved in modulating the capping structure (changing the five prime end of the mRNA to a three prime end by 5'-5' linkage, thus protecting the mRNA from 5' exonuclease, which degrades foreign RNA. This capping also helps in ribosomal binding), splicing (removing the introns, noncoding regions that are not transcribed into RNA, enabling the mRNA able to create proteins), adding of a Poly (A) Tail (adding Junk RNA to the 3' end which acts as a buffer to the 3' exonuclease, thus increasing the half life of mRNA), sequencing of specific nuclear export rates and sequestrating of the RNA transcript (Fabian et al. 2010; Siomi and Siomi 2010). Post-

transcriptional regulation is a major way to regulate patterns of gene expression during development.

1.6.2.5 Regulation of translation

Translational regulation refers to the control of the levels of protein synthesized from mRNA. The process can be controlled by several mechanisms, including recruitment of the small ribosomal subunit governed by an mRNA secondary structure, antisense RNA binding, or protein binding (Curtis et al. 1995; Kozak 1999; Malys and McCarthy 2011). For example, RNA binding proteins (RBPs) exist in prokaryotes and eukaryotes. The translation rate of cytochrome c mRNA was determined by the RBP TIA-1 and HuR in mammalian cells (Kawai et al. 2006).

1.6.3 Up-regulation and down-regulation

Up-regulation is a process in cells triggered by signals, either internally or externally, resulting in increased expression of one or more genes which subsequently increases the level of proteins encoded by those genes. For example, the expression of cytochrome P450 enzymes increases when xenobiotics go into the body of the cockroach (Brown et al. 2003).

Down-regulation is a process that cells decrease the quantity of a cellular component, such as RNA or protein. Troy and Shelanski found that down-regulation of copper/zinc superoxide dismutase causes apoptotic death in PC12 neuronal cells (Troy and Shelanski 1994).

1.7 Gene regulatory network

For cell survival, thousands of genes are expressed and work together, and each gene must be expressed at the proper time and in the proper amounts. The regulation and expression of some genes are highly robust, but expression of other genes is more variable from cell to cell and from individual to individual. The expression is usually induced by stresses, but individually this response can be beneficial physiologically (MacNeil and Walhout 2011).

A gene regulatory network or genetic regulatory network (GRN) is a collection of DNA segments in a cell, which interact with each other and other substances in the cell, controlling the rate of gene transcription. As regulatory genes regulate one another as well as other genes, responding to multiple inputs at the same time, the total map of their interactions formed a network (Erwin and Davidson 2009).

Gene regulatory networks play an important role in life processes, including cell differentiation, metabolism, cell cycle and signal transduction (Karlebach and Shamir 2008).

1.8 Regulation of cytochrome P450 genes

Many of P450 genes are under complex control during development, either following exposure of organisms to foreign compounds, such as insecticides and plant chemicals, or in response to endogenous signals.

1.8.1 Regulation of cytochrome P450 genes by specific promoters

A promoter is a region of DNA where the transcription initiation takes place. It usually occurs upstream from a gene coding region, providing a secure initial binding site for RNA polymerase or transcription factors, also acting as a controlling element. Mutation of the

promoter region results in the loss of cAMP induced reporter gene expression in luteal cells (Michael et al. 1997). Mutagenesis in the non-overlapping region of the promoter region of *CYP6B1v3* in *Papillo polyxenes* modulated the expression of this p450 gene (Petersen et al. 2003). Two p450 genes, *CYP28B1* and *CYP4G13V2* were isolated from the house fly, the transcription start points were mapped to 176 and 163 nucleotides upstream of the ATG translation start codon respectively. Regulatory binding sites and five conserved cis-acting elements for tissue or cell-specific transcription regulatory factors were identified in the promoter region of both P450 genes (Liu and Zhang 2002). Gfi-1 is a zinc finger protein. The *CYP6D1* promoter from the susceptible house fly strain binds with Gfi-1, which decreases the expression of *CYP6D1* (Gao and Scott 2006). The regions -1496 to -1102 bp for *CYP9A19,* and -1630 to -121- bp for *CYP9A22* were found to include several transcriptional regulatory elements, essential for basal transcriptional activity in the silkworm *Bombyx mori* (Zhao et al. 2013b). The upstream of *CYP9M10,* which included *CURE1* as a *cis*-regulatory element in *C. quinquesfasciatus,* drove a 10 fold expression compared to the susceptible *culex* mosquitoes (Wilding et al. 2012). When a 2141 bp 5' flanking region was cloned in *Spodoptera litura*, multiple putative elements for transcription factors binding sites were detected (Zhou et al. 2012). *Accord* insertion in the 5' flanking region of *CYP6G1* increased the expression of this gene and conferred nicotine resistance in .*D melanogaster* (Li et al. 2012). The increased expression of the *CYP4G61* gene was shown to be due to the mutation of *cis*-acting promoter sequences and/or *trans*-acting regulatory loci in whiteflies, *Bemisia tabaci* (Karatolos et al. 2012).

1.8.2 Regulation of cytochrome P450 genes by nuclear receptors

Nuclear receptors (NR) are a class of proteins that transfer endogenous (for example, small lipophilic hormones) or exogenous chemicals into cellular responses by triggering the transcription of NR target genes (Wang and LeCluyse 2003), thereby controlling the development, homeostasis and metabolism of the organism, including the multiple phase I and phase II metabolism enzymes and transporter systems (Honkakoski and Negishi 2000). The structural comparative analysis of NR revealed that they have several independent but interacting functional modules: N-terminal regulatory domain, two common structural domains (DNA binding domain (DBD) and ligand binding domain (LBD)) and C-terminal domain. N-terminal regulatory domain contains the activation function 1 (AF-1), which is independent of the presence of ligand. DBD is highly conserved, characterized by two C4-type zinc fingers, linking the receptor to the specific promoter regions of its target genes, such as P450 genes. LBD is a less conserved multifunctional motif, which is located in the carboxy-terminal portion of the receptor. It serves as a docking site for ligands and transcriptional activation. After the ligand binds to the LBD, the conformation of the LBD changes significantly, leading to the recruitment of co-activator proteins and co-integrators. C-terminal domain is highly variable in its sequence of various NR (Glass 1994; Horwitz et al. 1996; Mangelsdorf et al. 1995).

NR can bind with specific elements of the promoter region gene after activation by a variety of inducers, including chemicals (Wang and LeCluyse 2003), resulting in up-regulation or down-regulation of gene expression. Studies have reported that nuclear receptors, such as constitutive androstane receptor (CAR), pregnane X receptor (PXR), Hepatocyte nuclear factor 4 (HNF-4) and orphan receptor, regulate the expression of different P450 genes, which are involved in

xenobiotic detoxification in humans (Honkakoski and Negishi 2000; Maglich et al. 2002; Pascussi et al. 2000; Waxman 1999).

1.8.3 Regulation of cytochrome P450 genes by kinase

A kinase is a type of enzyme that can transfer phosphate groups from high-energy donor molecules, such as ATP, to specific substrates. Protein kinase has been shown to regulate the expression of P450 genes in catfish and humans (Bird et al. 1998; Blättler et al. 2007; Ghosh and Ray 2012; Murray et al. 2010; Rencurel et al. 2005; Yasunami et al. 2004). Salt-inducible kinase also can cooperate with the ACTH/Camp dependent protein kinase, to act on the cAMP-response element domain and repress the transcriptional activation of the *CYP11A* gene in humans (Takemori et al. 2003).Another study reported that protein kinase can not only regulate the expression of p450 genes, but also modulate the expression of GSTs and microsomal epoxide hydrolase (mEH) in rat hepatic tissue (Kim et al. 1998).

1.8.4 Regulation of cytochrome P450 genes by microRNAs

MicroRNAs are small non-coding RNA molecules, which function via base-pairing with complementary target mRNA resulting in target gene silencing via translational repression or degradation (Chen and Rajewsky 2007). They can be found in all organisms and are involved in both transcriptional and post-transcriptional regulation of gene expression (Bartel 2009). First, primary miRNA transcripts are cleaved by Rnase III, *Drosha*, in the cell nucleus into 70 to 80-nucleotide precursor miRNA hairpins and transported to the cytoplasm, where they are processed by another Rnase III, Dicer, into 19 to 25-nucleotide miRNA duplexes. One strand of duplexes is

degraded, and another strand is used as mature miRNA, which is then incorporated into the RNA-induced silencing complex to recognize the 3’ untranslated region of the target mRNA causing translational repression or mRNA cleavage. The expression of *CYP3A4* and *CYP1B1* genes, two P450 genes involved in drug-metabolizing, may be regulated by microRNAs at both the transcriptional and posttranscriptional level (Pan et al. 2009; Tsuchiya et al. 2006). MicroRNAs not only regulate the expression of cytochrome p450s, but also modulate the expression of nuclear receptors, which are in turn regulated by cellular stress and chemicals, affecting both metabolism and cellular biology (Nakajima and Yokoi 2011; Takagi et al. 2010; Takagi et al. 2008).

1.9 High-throughput sequencing studies and RNA-seq

Initial transcriptomics studies relied on hybridization-based microarray technologies or sequence-based approaches, but these methods showed a limited ability to fully catalogue and quantify the diverse RNA molecules that are expressed from genomes over wide ranges of levels (Ozsolak and Milos 2011; Wang et al. 2009). The development of high-throughput next-generation DNA sequencing (NGS) technologies revolutionized the transcriptomics analysis by allowing the analysis from RNA to cDNA sequencing on a massive scale (Ozsolak and Milos 2011), providing for a far more precise measurement of levels of transcripts and their isoforms than other previous methods.

Researchers working with small regulatory RNAs have pioneered next-generation sequencing (NGS) technologies for the analysis of RNA. RNA sequencing (RNA-seq) is also referred to as "Whole Transcriptome Shotgun Sequencing" ("WTSS") (Morin et al. 2008). A

typical RNA-seq experiment starts with mRNAs, which are converted into cDNA, subsequently used in high-throughput sequencing technologies to get RNA information of the sample. The RNA sequencing market is currently dominated by three different platforms: the FLX pyrosequencing system from 454 Life Sciences (a Roche company), the Illumina Genome Analyser (developed initially by Solexa), and the AB SOLiD system (now Life Technologies). On all three platforms, DNA fragments are sequenced in parallel, producing large numbers of relatively short sequence ‘‘reads’’ or ‘‘tags’’ (Marguerat and Bahler 2010). Studies using those methods have already altered our view of the extent and complexity of eukaryotic transcriptomes (Wang et al. 2009).

The transcriptome is the complete set of transcripts in a cell. The key aims of transcriptomics analysis are: to catalogue all transcripts of specie, including mRNAs, non-coding RNAs and small RNAs, to determine the transcriptional structure of genes, in terms of their start sites, 5′ and 3′ ends, splicing patterns and other post-transcriptional modifications, and to quantify the changing expression levels of each transcript during development and under different conditions (Wang et al. 2009).

The genomic library preparation is a key step in the RNA-seq. It determines how closely the cDNA sequence data reflect the original RNA population (Marguerat and Bahler 2010). First, the RNA (total RNA or mRNA) is converted to a library of cDNA fragments, with adaptors attached to one or both ends, which are then processed in a high-throughput manner to obtain short sequences from one end (single-end sequencing) or both ends (pair-end sequencing) after amplification. The length of reads is usually 35-100 bp, depending on the RNA-sequencing technology used (Mardis 2008). These short reads can be either aligned to a reference genome or

reference transcripts, or even assembled *de novo* without the genomic sequence to produce a genome-scale transcription map that consists of both the transcriptional structure and/or level of expression for each gene (Wang et al. 2009). The *de novo* aligner is a good strategy, because it does not rely on the genomic sequence of an organism and it can be used for non-model organism's transcriptome assembly. However, the length of the reads is not large enough to overlap between each read rendering *de novo* assembly difficult. It is not easy to reconstruct the original sequences, and the deep coverage makes the computing power to track all the possible alignments are impossible (Zerbino and Birney 2008). Another challenge comes from the relatively high error rate of RNA-Seq data, because the non-perfect matches have to be considered when mapping reads back to a genome. For example, we would encounter problems with this if interested in single nucleotide polymorphisms (SNPs) to detect allele specific expression in RNA-Seq data.

1.9.1 The wide use of High-throughput sequencing studies and RNA-seq

RNA-seq has produced highly quantifiable and reliable data, adequate sequencing depth at least as well as microarrays, and different transcripts with different UTR length allowing for quantitative analysis of transcript expression from the countable and digital data (t Hoen et al. 2008). For example, RNA-seq has been used to analyze blood-induced changes in gene expression in the mosquito (Bonizzoni et al. 2011). Beside those advances, RNA-seq is suitable for discovering new aberrant transcripts (Zhao et al. 2009), new splice junctions (Sorber et al. 2011), gene identification, polymorphism detection (Yang et al. 2011), molecular mechanisms (Brunskill et al. 2011), and metabolic interactions of organisms (Rosenthal et al. 2011).

1.9.2 Applying RNA-seq to gene regulation

From the DNA-RNA transcription step to post-translational modification of a protein, including the control of alternative splicing, RNA editing, degradation and translation, gene expression can be modulated at any step.

In molecular biology and genetics, splicing is a modification of an RNA after transcription, in which introns are removed and exons are joined together to form a mRNA. This is a necessary step for the typical eukaryotic messenger RNA before it can be used to produce a correct protein through translation. It was reported that 95% of the human multi-exon genes have been found to undergo alternative splicing, with exon skipping being the most frequent form of regulation (Pan et al. 2008). Alternative splicing has been shown to increase the diversity of the transcriptome of a single cell during embryonic development in mice (Wilhelm et al. 2008). Some research results also show that regulation of splicing is used by unicellular eukaryotes to control and diversify gene expression (Marguerat and Bahler 2010). Analysis of alternative splicing by RNA-seq has been used to quantitatively examine splicing diversity by searching for reads that span known splice junctions (Ameur et al. 2010; Gan et al. 2010; Loraine et al. 2013; Nicolae et al. 2011; Sultan et al. 2008; Wang et al. 2008; Zhao et al. 2013a).

Gene expression also can be regulated by editing of mRNA transcripts. RNA editing is a post-transcriptional molecular phenomenon that can increase proteomic diversity by modifying the sequence of completely or partially non-functional primary transcripts (Gott 2003; Picardi et al. 2010). The high-throughput RNA-Seq technology allows the detection of a comprehensive landscape of RNA editing at the genome level (Picardi et al. 2010). High-throughput sequencing

coupled with hybridization is a good strategy to enrich specific RNA populations before sequencing (Licatalosi and Darnell 2010).

Information about protein–RNA interactions are fundamental for the understanding of different layers of post-transcriptional regulatory networks (Marguerat and Bahler 2010). Researchers wish to distinguish between the primary (direct) and secondary (indirect) effects of RNA regulatory factors (Licatalosi and Darnell 2010). Protein–RNA interactions can be separated into two categories: RNA-binding proteins that are immunoprecipitated together with their intact target transcripts (RIP) (Gerber et al. 2004) or RNA-binding proteins that are crosslinked to the RNAs they interact with and treated with RNAse before immunoprecipitation (CLIP for crosslinking immunoprecipitation) (Ule et al. 2005). The second approach limits the analysis to RNA fragments protected by the binding protein. RNA-seq has been successfully applied to these approaches (Chi et al. 2009; Licatalosi et al. 2008).

High-throughput sequencing studies have provided an unprecedented ability to describe RNAs on a genome-wide scale and to suggest which *cis* elements and *trans*-acting factors are associated with RNA regulation. RNA-seq will play an important role in research in the next few years.

1.10 House fly

The house fly, *M. domestica*L. (Diptera: Muscidae), is a major domestic, medical and veterinary pest that causes irritation, spoils food and acts as a vector for many pathogenic organisms (Malik et al. 2007). The house fly has four different life stages: egg, larva, pupa and adult. Adult flies can transmit viruses to humans through their excreta and filth when they alight

on surfaces of food. They also can transmit pathogens, including bacterial agents of typhoid fever, bacillary dysentery, trachoma and protozoan amoebic dysentery (Clesceri 1987).

Insecticide resistance of the insect is one of the most serious problems in agriculture today. Resistance can lead to increased pesticide application rates, increased frequency of pesticide use and cost, and ultimately the compound may become ineffective.

Insecticides were first introduced against flies in the 1940's. Since then, many insecticides have been used, including pyrethroids, organophosphates and carbamates. However, their effectiveness decreased within a few years as fly populations developed resistance to these insecticides.

The first report of house flies insecticide resistance was in 1962 when the Rutgers strain of house flies were shown to have 98- to 125- fold resistance to diazinon (Forgash et al. 1962). By 1978, house flies were reported to be resistant to the major insecticide classes, including DDT and its analogues, cyclodienes, carbamates, organophosphates, and pyrethroids. Substantial evidence indicated that multiple resistance mechanisms or genes were involved in house fly insecticide resistance.

Resistance conferred by enhanced metabolic detoxification via cytochrome P450 monooxygenases, hydrolase, and GST, target-site (primarily sodium channel) insensitivity, and decreased penetration have been reported in house flies already (Scott 1989; Scott and Georghiou 1986a; Scott and Georghiou 1986b; Scott and Zhang 2003).

Cytochrome P450 monooxygenases-mediated detoxification is a very important mechanism of insecticide resistance in house flies (Scott 1999). Several P450 genes have been isolated from resistant house fly strains already. The *CYP6A1* is the first insect P450 gene that was cloned and

sequenced from the Rutgers strain (Feyereisen et al. 1989) with the expression of the *CYP6A1* gene in the Rutgers strain being 3-fold higher than the susceptible strain (Carino et al. 1992).

In 1982, the LPR strain of house flies was collected from a dairy. After several generations of selection with permethrin, LPR became more homozygous and had extremely high levels (>6000-fold) of resistance to pyrethroid (Scott and Georghiou 1986b).The *CYP6D1* mRNA was expressed 10-fold higher in LPR flies compared to susceptible flies (Scott 1999). Substantial evidence indicates that the *CYP6D1* is directly involved in insecticide resistance (Wheelock and Scott 1992; Zhang and Scott 1994), and the promoter of *CYP6D1v1* was able to mediate the phenobarbital induction (Lin and Scott 2011). *CYP6A1* was mapped to autosome 5 of the house fly and the factors responsible for the increased expression level were linked to autosome 2 (Feyereisen 1995). This indicated that the expression of *CYP6A1* was regulated by trans-factors. *CYP6D1* in the LPR strain of the house fly is due to an increased rate of transcription, which is regulated by factors on autosomes 1 and 2 (Liu and Scott 1996; Liu and Scott 1998). The permethrin resistance can be suppressed from 5900-fold to 32-fold by the P450 inhibitor piperonylbutoxide (PBO) (Scott and Georghiou 1986b).

The Alabama house fly strain (ALHF) was collected from a poultry farm in Alabama in 1998 after control failure with permethrin. The level of resistance to permethrin in ALHF was 1,800-fold compared to the susceptible strain. P450 monooxygenases and hydrolases were proven to be involved in resistance in ALHF (Liu and Yue 2000). *CYP28B1*, *CYP4G13v2*, *CYP4D4v2*, *CYP4G2*, *CYP6A36, CYP6A37* and *CYP6A38* were P450 genes isolated from the ALHF strain. *CYP4D4v2, CYP4G2* and *CYP6A38* were co-up-regulated by permethrin treatment in ALHF house flies (Liu and Zhang 2002; Zhu et al. 2008b). Genetic linkage analysis located *CYP4D4v2*

and *CYP6A38* on autosome 5, corresponding to the linkage of P450-mediated resistance in ALHF, whereas *CYP4G2* was located on autosome 3 (Zhu et al. 2008b). Expression of *CYP6A36* was significantly higher in ALHF compared to susceptible CS flies. Over expression of *CYP6A36* was detected in the ALHF abdomen, where the primary detoxification organs of the midgut and fat body are located. *CYP6A37* has no significant difference in expression between ALHF and the susceptible CS strain (Zhu et al. 2008a). Genetic linkage analysis located *CYP6A36* on autosome 5. Over expression of *CYP6A36* was linked to the factors on autosomes 1 and 2, corresponding to the linkage of P450-mediated resistance in ALHF (Zhu et al. 2008a).

The resistance of the NG98 house fly strain due to *kdr* was located on autosome 3 and monooxygenase-mediated resistance on autosome 1, 2, and 5, suggesting that P450 genes involved in resistance and regulatory factors controlling P450 expression are different among different populations (Scott and Kasai 2004). This evidence also suggests the importance of p450 genes in detoxification of insecticides and evolution of insecticide resistance in house flies (Zhu et al. 2008a).

The hydrolase-mediated detoxification is due to the mutation in carboxylesterase which causes the loss of its carboxylesterase activity and gain of OP hydrolase activity (Claudianos et al. 1999; Hemingway 2000). Comparisons between OP susceptible and resistant Rutgers strains indicated that resistance in the house fly was associated with a single amino acid mutation (glycine 137 to aspartic acid) in the ali-esterase E3. The same mutation has been reported in an OP-resistant strain of the sheep blowfly (Claudianos et al. 1999). Zhang et al (2007) found that carboxylesterase activities and maximal velocities to five naphthyl-substituted substrates in a

beta-cypermethrin-resistant strain were significantly higher than in a susceptible strain (Zhang et al. 2007).

In 1992, two classes of GSTs, GST1 and GST2, were identified from the OP- and carbamate-resistant house fly strain, Cornell-R (Fournier et al. 1992). The genome of *M. domestica* contains multiple intronless loci, some of which appear to have resulted from fusion between the 5' and 3' ends of different GST genes (Enayati et al. 2005), and many studies have shown GST-mediated metabolic detoxification resistance in house flies (Clark and Shamaan 1984; Clark et al. 1986; Hemingway 2000; Rauch and Nauen 2004; Wei et al. 2001).

Knockdown resistance, as exemplified by genetically defined *kdr* strains of the house fly, represents a serious threat to continued use of the pyrethroid insecticides in the field (Mullin et al. 1992). The phenomenon of *kdr* was first observed in a DDT-resistant house fly strain as an ability to resist the rapid knockdown paralysis action of DDT (Busvine 1951). The basis for *kdr* and super *kdr* comes from reduced sensitivity of sodium channels to these compounds. This reduction in sensitivity has been correlated with alterations in physicochemical properties of phospholipids and sodium channel proteins mapped to chromosome 3 of the house fly (Soderlund David et al. 2001). The L1014F point mutation in the house fly Vssc1 sodium channel confers knockdown resistance to pyrethroids (Smith et al. 1997). The M918T point mutation is associated with the super-*kdr* trait (Lee et al. 1999).

The decreased penetration rate of insecticides has been recorded in a number of resistant strains of the house fly (DeVries and Georghiou 1981; Plapp and Hoyer 1968; Scott and Georghiou 1986a; Scott and Georghiou 1986b). Penetration itself confers only a 2-to 3- fold

resistance (Plapp and Hoyer 1968). However, it is an enhancer of resistance conferred by other mechanisms.

Due to the medical and economic importance of house flies, the biochemistry and genetics of insecticide resistance have been well studied in the house fly. The size of the house-fly genome is about 184 Mb, only about 1.6-fold larger than that of *D. melanogaster*. However, the evolution of *cis*-regulatory sequences in *Drosophila* has proven difficult in some cases. Sequencing the house fly genome will provide a critical resource for the analysis of *cis*-regulatory, shed light on the immune defense systems of this important species and expand our understanding of basic house fly biology to develop new control strategies. (Scott et al. 2009).

1.11 References

Alberts, B (2008) Molecular biology of the cell, 5th edn. Garland Science, New York.

Ameur, A, Wetterbom, A, Feuk, L and Gyllensten, U (2010) Global and unbiased detection of splice junctions from RNA-seq data. Genome Biologyogy 11: R34.

Amichot, M, Tares, S, Brun-Barale, A, Arthaud, L, Bride, JM and Bergé, JB (2004) Point mutations associated with insecticide resistance in the *Drosophila* cytochrome P450 Cyp6a2 enable DDT metabolism. European Journal of Biochemistry 271: 1250-1257.

Anspaugh, DD, Rose, RL, Koehler, PG, Hodgson, E and Roe, RM (1994) Multiple Mechanisms of Pyrethroid Resistance in the German Cockroach, *Blattella germanica* (L.). Pesticide Biochemistry and Physiology 50: 138-148.

Anthony, N, Unruh, T, Ganser, D and ffrench-Constant, R (1998) Duplication of the *Rdl* GABA receptor subunit gene in an insecticide-resistant aphid, *Myzus persicae*. Molecular and

General Genetics. 260: 165-175.

Bartel, DP (2009) MicroRNAs: target recognition and regulatory functions. Cell 136: 215-233.

Bass, C, Schroeder, I, Turberg, A, Field, LM and Williamson, MS (2004) Identification of the Rdl mutation in laboratory and field strains of the cat flea, *Ctenocephalides felis* (Siphonaptera: Pulicidae). Pest Mangement Science 60: 1157-1162.

Bellí, G, Garí, E, Piedrafita, L, Aldea, M and Herrero, E (1998) An activator/repressor dual system allows tight tetracycline-regulated gene expression in budding yeast. Nucleic Acids Researchearch 26: 942-947.

Bell, CE and Lewis, M (2001) The Lac repressor: a second generation of structural and functional studies. Current Opinion in Structural Biology 11: 19-25.

Bermudez, I, Hawkins, CA, Taylor, AM and Beadle, DJ (1991) Actions of insecticides on the insect gaba receptor complex. Journal of Receptors and Signal Transduction 11: 221-232.

Bird, IM, Mason, JI and Rainey, WE (1998) Protein kinase A, protein kinase C, and Ca2+-regulated expression of 21-hydroxylase cytochrome P450 in H295R human adrenocortical cells. Journal of Clinical Endocrinology & Metabolism 83: 1592-1597.

Blättler, SM, Rencurel, F, Kaufmann, MR and Meyer, UA (2007) In the regulation of cytochrome P450 genes, phenobarbital targets LKB1 for necessary activation of AMP-activated protein kinase. Proceedings of the National Academy of Sciences 104: 1045-1050.

Bonizzoni, M, Dunn, WA, Campbell, CL, Olson, KE, Dimon, MT, Marinotti, O and James, AA (2011) RNA-seq analyses of blood-induced changes in gene expression in the mosquito vector species, *Aedes aegypti*. BMC Genomics 12: 82.

Brévault, T, Achaleke, J, Sougnabé, S and Vaissayre, M (2008) Tracking pyrethroid resistance in the polyphagous bollworm, *Helicoverpa armigera* (Lepidoptera: Noctuidae), in the shifting landscape of a cotton-growing area. Bulletin of Entomological Research 98: 565-573.

Brahms, JG, Dargouge, O, Brahms, S, Ohara, Y and Vagner, V (1985) Activation and inhibition of transcription by supercoiling. Journal of Molecular Biology 181: 455-465.

Brattsten, LB, Holyoke, CW, Leeper, JR and Raffa, KF (1986) Insecticide Resistance - Challenge to Pest-Management and Basic Research. Science 231: 1255-1260.

Brown, D, Zhang, L, Wen, ZM and Scott, JG (2003) Induction of P450 monooxygenases in the German cockroach, *Blattella germanica* L. Arch Insect Biochem Physiol 53: 119-124.

Brun, A, Cuany, A, Le Mouel, T, Berge, J and Amichot, M (1996) Inducibility of the *Drosophila melanogaster* cytochrome P450 gene, CYP6A2, by phenobarbital in insecticide susceptible or resistant strains. Insect Biochemistry and Molecular Biology 26: 697-703.

Brunskill, EW, Lai, HL, Jamison, DC, Potter, SS and Patterson, LT (2011) Microarrays and RNA-Seq identify molecular mechanisms driving the end of nephron production. BMC Developmental Biology 11: 15.

Busvine, JR (1951) Mechanism of Resistance to Insecticide in Houseflies. Nature 168: 193-195.

Byrne, FJ and Devonshire, AL (1997) Kinetics of Insensitive Acetylcholinesterases in Organophosphate-Resistant Tobacco Whitefly, *Bemisia tabaci* (Gennadius) (Homoptera: Aleyrodidae). Pesticide Biochemistry and Physiology 58: 119-124.

Calkhoven, CF and Ab, G (1996) Multiple steps in the regulation of transcription-factor level and activity. Biochemical Journal 317: 329-342.

Campbell, PM, Newcomb, RD, Russell, RJ and Oakeshott, JG (1998) Two different amino acid substitutions in the ali-esterase, E3, confer alternative types of organophosphorus insecticide resistance in the sheep blowfly, *Lucilia cuprina*. Insect Biochemistry and Molecular Biology 28: 139-150.

Carino, F, Koener, J, Plapp Jr, F and Feyereisen, R (1992) Expression of the cytochrome P450 gene CYP6A1 in the housefly, *Musca domestica*. In: Book Expression of the cytochrome P450 gene CYP6A1 in the housefly, *Musca domestica* (Editor, eds.). Vol., pp., City.

Carino, FA, Koener, JF, Plapp, FW, Jr. and Feyereisen, R (1994) Constitutive overexpression of the cytochrome P450 gene CYP6A1 in a house fly strain with metabolic resistance to insecticides. Insect Biochemistry and Molecular Biology 24: 411-418.

Casida, JE (2009) Pest Toxicology: The primary mechanisms of pesticide action. Chemical Research in Toxicology 22: 609-619.

Castillochavez, C, Levin, SA and Gould, F (1988) Physiological and behavioral adaptation to varying environments - a mathematical-model. Evolution 42: 986-994.

Catterall, WA (2000) From ionic currents to molecular mechanisms: the structure and function of voltage-gated sodium channels. Neuron 26: 13-25.

Chen, K and Rajewsky, N (2007) The evolution of gene regulation by transcription factors and microRNAs. Nature Reviews Genetics 8: 93-103.

Chen, Z, Newcomb, R, Forbes, E, McKenzie, J and Batterham, P (2001) The acetylcholinesterase gene and organophosphorus resistance in the Australian sheep blowfly, *Lucilia cuprina*. Insect Biochemistry and Molecular Biology 31: 805-816.

Chi, SW, Zang, JB, Mele, A and Darnell, RB (2009) Argonaute HITS-CLIP decodes

microRNA–mRNA interaction maps. Nature 460: 479-486.

Chiang, FM and Sun, CN (1993) Glutathione transferase isozymes of diamondback moth larvae and their role in the degradation of some organophosphorus insecticides. Pesticide Biochemistry and Physiology 45: 7-14.

Clark, AG and Shamaan, NA (1984) Evidence that DDT-dehydrochlorinase from the house fly is a glutathione S-transferase. Pesticide Biochemistry and Physiology 22: 249-261.

Clark, AG, Shamaan, NA, Sinclair, MD and Dauterman, WC (1986) Insecticide metabolism by multiple glutathione S-transferases in two strains of the house fly, *Musca domestica* (L). Pesticide Biochemistry and Physiology 25: 169-175.

Claudianos, C, Russell, RJ and Oakeshott, JG (1999) The same amino acid substitution in orthologous esterases confers organophosphate resistance on the house fly and a blowfly. Insect Biochemistry and Molecular Biology 29: 675-686.

Clesceri, LS (1987) Chemical Control of Pests and Vectors. In: Water Resources and Natural Control Processes. pp. 261-282. Springer.

Corbel, V, N'Guessan, R, Brengues, C, Chandre, F, Djogbenou, L, Martin, T, Akogbéto, M, Hougard, JM and Rowland, M (2007) Multiple insecticide resistance mechanisms in *Anopheles gambiae* and *Culex quinquefasciatus* from Benin, West Africa. Acta Tropica 101: 207-216.

Curtis, D, Lehmann, R and Zamore, PD (1995) Translational regulation in development. Cell 81: 171-178.

Danielson, P, Foster, J, McMahill, M, Smith, M and Fogleman, J (1998) Induction by alkaloids and phenobarbital of Family 4 Cytochrome P450s in *Drosophila*: evidence for involvement in host plant utilization. Molecular and General Genetics MGG 259: 54-59.

Deaton, AM and Bird, A (2011) CpG islands and the regulation of transcription. Genes Development 25: 1010-1022.

Delye, C, Bousset, L and Corio-Costet, MF (1998) PCR cloning and detection of point mutations in the eburicol 14alpha-demethylase (CYP51) gene from *Erysiphe graminis* f. sp. hordei, a "recalcitrant" fungus. Current Genetics 34: 399-403.

Després, L, David, J-P and Gallet, C (2007) The evolutionary ecology of insect resistance to plant chemicals. Trends in Ecology & Evolution 22: 298-307.

Desvergne, B, Michalik, L and Wahli, W (2006) Transcriptional regulation of metabolism. Physiological Reviews 86: 465-514.

Devonshire, AL and Moores, GD (1982) A carboxylesterase with broad substrate specificity causes organophosphorus, carbamate and pyrethroid resistance in peach-potato aphids (*Myzus persicae*). Pesticide Biochemistry and Physiology 18: 235-246.

Devonshire, AL and Moores, GD (1984) Different forms of insensitive acetylcholinesterase in insecticide-resistant house flies (*Musca domestica*). Pesticide Biochemistry and Physiology 21: 336-340.

DeVries, DH and Georghiou, GP (1981) Decreased nerve sensitivity and decreased cuticular penetration as mechanisms of resistance to pyrethroids in a (1R)-trans-permethrin-selected strain of the house fly. Pesticide Biochemistry and Physiology 15: 234-241.

Dickinson, WJ (1979) Complex cis-acting regulatory genes demonstrated in *Drosophila hybrids*. Developmental Genetics 1: 229-240.

Dong, K (2007) Insect sodium channels and insecticide resistance. Invertebrate Neuroscience 7: 17-30.

Du, W, Awolola, TS, Howell, P, Koekemoer, LL, Brooke, BD, Benedict, MQ, Coetzee, M and Zheng, L (2005) Independent mutations in the Rdl locus confer dieldrin resistance to *Anopheles gambiae* and An. arabiensis. Insect Molecular Biology 14: 179-183.

Dunkov, B, Guzov, V, Mocelin, G, Shotkoski, F, Brun, A, Amichot, M, Ffrench-Constant, R and Feyereisen, R (1997) The *Drosophila* cytochrome P450 gene Cyp6a2: structure, localization, heterologous expression, and induction by phenobarbital. DNA and Cell Biology 16: 1345-1356.

Eldefrawi, M, Miskus, R and Sutcher, V (1960) Methylenedioxyphenyl Derivatives as Synergists for Carbamate Insecticides on Susceptible, DDT-, and Parathion-resistant House Flies. Journal of Economic Entomology 53: 231-234.

Enayati, AA, Ranson, H and Hemingway, J (2005) Insect glutathione transferases and insecticide resistance. Insect Molecular Biology 14: 3-8.

Erwin, DH and Davidson, EH (2009) The evolution of hierarchical gene regulatory networks. Nature Reviews Genetics 10: 141-148.

Eschenlauer, AC and Reznikoff, WS (1991) *Escherichia coli* catabolite gene activator protein mutants defective in positive control of lac operon transcription. Journal of Bacteriology 173: 5024-5029.

Fabian, MR, Sonenberg, N and Filipowicz, W (2010) Regulation of mRNA translation and stability by microRNAs. Annual Review of Biochemistry 79: 351-379.

Farnham, AW (1973) Genetics of resistance of pyrethroid-selected houseflies, *Musca domestica* L. Pesticide Science 4: 513-520.

Feyereisen, R (1995) Molecular biology of insecticide resistance. Toxicology Letters 82–83: 83-90.

Feyereisen, R (1999) Insect P450 enzymes. Annual Review of Entomology 44: 507-533.

Feyereisen, R (2006) Evolution of insect P450. Biochemical Society Transactions 34: 1252-1255.

Feyereisen, R, Koener, JF, Farnsworth, DE and Nebert, DW (1989) Isolation and sequence of cDNA encoding a cytochrome P-450 from an insecticide-resistant strain of the house fly, *Musca domestica*. Proceedings of the National Academy of Sciences 86: 1465-1469.

ffrench-Constant, RH, Rocheleau, TA, Steichen, JC and Chalmers, AE (1993) A point mutation in a *Drosophila* GABA receptor confers insecticide resistance. Nature 363: 449-451.

Fickett, JW and Wasserman, WW (2000) Discovery and modeling of transcriptional regulatory regions. Current Opinion in Biotechnology 11: 19-24.

Fonseca-Gonzalez, I, Quinones, ML, McAllister, J and Brogdon, WG (2009) Mixed-function oxidases and esterases associated with cross-resistance between DDT and lambda-cyhalothrin in Anopheles darlingi Root 1926 populations from Colombia. Memorias do Instituto Oswaldo Cruz 104: 18-26.

Forgash, AJ, Cook, BJ and Riley, RC (1962) Mechanisms of Resistance in Diazinon-Selected Multi-Resistant *Musca domestica*1. Journal of Economic Entomology 55: 544-551.

Fournier, D, Bride, JM, Poirie, M, Berge, JB and Plapp, FW, Jr. (1992) Insect glutathione S-transferases. Biochemical characteristics of the major forms from houseflies susceptible and resistant to insecticides. Journal of Biological Chemistry 267: 1840-1845.

Fradin, MS and Day, JF (2002) Comparative Efficacy of Insect Repellents against Mosquito Bites. New England Journal of Medicine 347: 13-18.

Fraser, JD, Irving, BA, Crabtree, GR and Weiss, A (1991) Regulation of interleukin-2 gene

enhancer activity by the T cell accessory molecule CD28. Science 251: 313-316.

Freeman, JA and Bassler, BL (1999) Sequence and function of LuxU: a two-component phosphorelay protein that regulates quorum sensing in Vibrio harveyi. Journal of Bacteriology 181: 899-906.

Gan, Q, Chepelev, I, Wei, G, Tarayrah, L, Cui, K, Zhao, K and Chen, X (2010) Dynamic regulation of alternative splicing and chromatin structure in *Drosophila* gonads revealed by RNA-seq. Cell Research 20: 763-783.

Gao, J and Scott, JG (2006) Role of the transcriptional repressor mdGfi-1 in CYP6D1v1-mediated insecticide resistance in the house fly, *Musca domestica*. Insect Biochemistry and Molecular Biology 36: 387-395.

Georghiou, GP (1972) Studies on resistance to carbamate and organophosphorus insecticides in *Anopheles albimanus*. American Journal of Tropical Medicine and Hygiene 21: 797-806.

Georghiou, GP (1990) Overview of Insecticide Resistance. Acs Symposium Series 421: 18-41.

Gerber, AP, Herschlag, D and Brown, PO (2004) Extensive association of functionally and cytotopically related mRNAs with Puf family RNA-binding proteins in yeast. Plos Biology 2: E79.

Gerold, JL and Laarman, JJ (1964) Selection of Some Strains of Anopheles Atroparvus with Different Behavioural Responses to Contacts with Ddt. Nature 204: 500-501.

Ghosh, MC and Ray, AK (2012) Regulation of cytochrome P4501A by protein kinase C: the role of heat shock protein70. Journal of Cell Communication and Signaling 6: 37-44.

Glass, CK (1994) Differential recognition of target genes by nuclear receptor monomers,

dimers, and heterodimers. Endocrine reviews 15: 391-407.

Gleyzer, N, Vercauteren, K and Scarpulla, RC (2005) Control of Mitochondrial Transcription Specificity Factors (TFB1M and TFB2M) by Nuclear Respiratory Factors (NRF-1 and NRF-2) and PGC-1 Family Coactivators. Molecular and Cellular Biology 25: 1354-1366.

Gott, JM (2003) Expanding genome capacity via RNA editing. Comptes rendus biologies 326: 901-908.

Gruber, TM and Gross, CA (2003) Multiple sigma subunits and the partitioning of bacterial transcription space. Annual Review of Microbiology 57: 441-466.

Guedes, N, Guedes, R, Ferreira, G and Silva, L (2009) Flight take-off and walking behavior of insecticide-susceptible and-resistant strains of Sitophilus zeamais exposed to deltamethrin. Bulletin of Entomological Research 99: 393-400.

Guerrero, FD, Jamroz, RC, Kammlah, D and Kunz, SE (1997) Toxicological and molecular characterization of pyrethroid-resistant horn flies, Haematobia irritans: identification of kdr and super-kdr point mutations. Insect Biochemistry and Molecular Biology 27: 745-755.

Guzov, VM, Houston, HL, Murataliev, MB, Walker, FA and Feyereisen, R (1996) Molecular cloning, overexpression in *Escherichia coli*, structural and functional characterization of house fly cytochrome b5. Journal of Biological Chemistry 271: 26637-26645.

Hemingway, J (2000) The molecular basis of two contrasting metabolic mechanisms of insecticide resistance. Insect Biochemistry and Molecular Biology 30: 1009-1015.

Hines, R, Sorensen, BR, Shea, MA and Maury, W (2004) PU. 1 binding to ets motifs within the equine infectious anemia virus long terminal repeat (LTR) enhancer: regulation of LTR activity and virus replication in macrophages. Journal of virology 78: 3407-3418.

Hogeweg, P (2000) Evolving mechanisms of morphogenesis: on the interplay between differential adhesion and cell differentiation. Journal of Theoretical Biology 203: 317-333.

Honkakoski, P and Negishi, M (2000) Regulation of cytochrome P450 (CYP) genes by nuclear receptors. Biochemical Journal 347: 321-337.

Horwitz, K, Jackson, T, Bain, D, Richer, J, Takimoto, G and Tung, L (1996) Nuclear receptor coactivators and corepressors. Molecular Endocrinology 10: 1167-1177.

Huang, S and Han, Z (2007) Mechanisms for multiple resistances in field populations of common cutworm, *Spodoptera litura* (Fabricius) in China. Pesticide Biochemistry and Physiology 87: 14-22.

Hughes, AL (1999) Adaptive evolution of genes and genomes. Oxford University Press, USA.

Imbalzano, AN, Kwon, H, Green, MR and Kingston, RE (1994) Facilitated binding of TATA-binding protein to nucleosomal DNA. Nature 370 :481-485.

Ingles, PJ, Adams, PM, Knipple, DC and Soderlund, DM (1996) Characterization of voltage-sensitive sodium channel gene coding sequences from insecticide-susceptible and knockdown-resistant house fly strains. Insect Biochemistry and Molecular Biology 26: 319-326.

Jacob, ST (1983) Enzymes of nucleic acid synthesis and modification. CRC Press, Boca Raton, Fla.

Jakoby, WB (1978) The Glutathione S-Transferases: A Group of Multifunctional Detoxification Proteins. Advances in Enzymology and Related Areas of Molecular Biology, Volume 46: 383-414.

Karatolos, N, Williamson, MS, Denholm, I, Gorman, K, Ffrench-Constant, RH and Bass, C

(2012) Over-expression of a cytochrome P450 is associated with resistance to pyriproxyfen in the greenhouse whitefly *Trialeurodes vaporariorum*. PloS One 7: e31077.

Karlebach, G and Shamir, R (2008) Modelling and analysis of gene regulatory networks. Nature Reviews Molecular Cell Biology 9: 770-780.

Kasai, S, Weerashinghe, IS, Shono, T and Yamakawa, M (2000) Molecular cloning, nucleotide sequence and gene expression of a cytochrome P450 (CYP6F1) from the pyrethroid-resistant mosquito, *Culex quinquefasciatus* Say. Insect Biochemistry and Molecular Biology 30: 163-171.

Kawai, T, Lal, A, Yang, X, Galban, S, Mazan-Mamczarz, K and Gorospe, M (2006) Translational control of cytochrome c by RNA-binding proteins TIA-1 and HuR. Molecular and cellular biology 26: 3295-3307.

Kerkut, GA and Gilbert, LI (1985) Comprehensive insect physiology, biochemistry, and pharmacology, 1st edn. Pergamon Press, Oxford ; New York.

Ketterman, AJ, Jayawardena, K and Hemingway, J (1992) Purification and characterization of a carboxylesterase involved in insecticide resistance from the mosquito *Culex quinquefasciatus*. Biochemical Journal 287: 355-360.

Kilpatrick, JW and Schoof, HF (1958) A Field Strain of Malathion-Resistant House Flies. Journal of Economic Entomology 51: 18-19.

Kim, SG, Cho, JY, Chung, YS, Ahn, E-T, Lee, K-Y and Han, Y-B (1998) Suppression of Xenobiotic-Metabolizing Enzyme Expression in Rats by Acriflavine, a Protein Kinase C Inhibitor Effects on Epoxide Hydrolase, GlutathioneS-Transferases, and Cytochromes P450. Drug Metabolism and Disposition 26: 66-72.

Kostaropoulos, I, Papadopoulos, AI, Metaxakis, A, Boukouvala, E and Papadopoulou-Mourkidou, E (2001) Glutathione S–transferase in the defence against pyrethroids in insects. Insect Biochemistry and Molecular Biology 31: 313-319.

Kouzarides, T (2007) Chromatin modifications and their function. Cell 128: 693-705.

Kozak, M (1999) Initiation of translation in prokaryotes and eukaryotes. Gene 234: 187-208.

Kuhn, J, Pettersson, EM, Feld, BK, Burse, A, Termonia, A, Pasteels, JM and Boland, W (2004) Selective transport systems mediate sequestration of plant glucosides in leaf beetles: A molecular basis for adaptation and evolution. Proceedings of the National Academy of Sciences of the United States of America 101: 13808-13813.

Le Goff, G, Hamon, A, Berge, JB and Amichot, M (2005) Resistance to fipronil in *Drosophila* simulans: influence of two point mutations in the RDL GABA receptor subunit. J Neurochem 92: 1295-1305.

Lee, D-W, Choi, JY, Kim, WT, Je, YH, Song, JT, Chung, BK, Boo, KS and Koh, YH (2007) Mutations of acetylcholinesterase1 contribute to prothiofos-resistance in *Plutella xylostella* (L.). Biochemical and Biophysical Research Communications 353: 591-597.

Lee, S and Scott, J (1989) Microsomal cytochrome P450 monooxygenases in the house fly (*Musca domestica* L.): biochemical changes associated with pyrethroid resistance and phenobarbital induction. Pesticide Biochemistry and Physiology 35: 1-10.

Lee, SH, Kang, JS, Min, JS, Yoon, KS, Strycharz, JP, Johnson, R, Mittapalli, O, Margam, VM, Sun, W and Li, HM (2010) Decreased detoxification genes and genome size make the human body louse an efficient model to study xenobiotic metabolism. Insect Molecular Biology 19: 599-615.

Lee, SH, Smith, TJ, Knipple, DC and Soderlund, DM (1999) Mutations in the house fly Vssc1 sodium channel gene associated with super-kdr resistance abolish the pyrethroid sensitivity of Vssc1/tipE sodium channels expressed in *Xenopus oocytes*. Insect Biochemistry and Molecular Biology 29: 185-194.

Lee, SH, Yoon, K-S, Williamson, MS, Goodson, SJ, Takano-Lee, M, Edman, JD, Devonshire, AL and Marshall Clark, J (2000) Molecular Analysis of kdr-like Resistance in Permethrin-Resistant Strains of Head Lice, *Pediculus capitis*. Pesticide Biochemistry and Physiology 66: 130-143.

Leeper, JR, Roush, RT and Reynolds, HT (1986) Preventing or managing resistance in arthropods. Pesticide resistance: strategies and tactics for management: 335-346.

Li, X, Bai, S and Cass, BN (2012) Accord insertion in the 5′ flanking region of CYP6G1 confers nicotine resistance in *Drosophila melanogaster*. Gene 502: 1-8.

Licatalosi, DD and Darnell, RB (2010) RNA processing and its regulation: global insights into biological networks. Nature Reviews Genetics 11: 75-87.

Licatalosi, DD, Mele, A, Fak, JJ, Ule, J, Kayikci, M, Chi, SW, Clark, TA, Schweitzer, AC, Blume, JE and Wang, X (2008) HITS-CLIP yields genome-wide insights into brain alternative RNA processing. Nature 456: 464-469.

Lin, GG-H and Scott, JG (2011) Investigations of the constitutive overexpression of CYP6D1 in the permethrin resistant LPR strain of house fly (*Musca domestica*). Pesticide Biochemistry and Physiology 100: 130-134.

Lin, S and Riggs, A (1975) The general affinity of lac repressor for *E. coli* DNA: implications for gene regulation in procaryotes and eucaryotes. Cell 4: 107-111.

Liu, N and Scott, JG (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*. Biochemical Genetics 34: 133-148.

Liu, N and Scott, JG (1997) Phenobarbital induction of CYP6D1 is due to a trans acting factor on autosome 2 in house flies, *Musca domestica*. Insect Molecular Biology 6: 77-81.

Liu, N and Scott, JG (1998) Increased transcription of CYP6D1 causes cytochrome P450-mediated insecticide resistance in house fly. Insect Biochemistry and Molecular Biology 28: 531-535.

Liu, N and Yue, X (2000) Insecticide resistance and cross-resistance in the house fly (Diptera: Muscidae). Journal of Economic Entomology 93: 1269-1275.

Liu, N and Zhang, L (2002) Identification of two new cytochrome P450 genes and their 5′-flanking regions from the housefly, *Musca domestica*. Insect Biochemistry and Molecular Biology 32: 755-764.

Liu, N, Zhu, F, Xu, Q, Pridgeon, JW and Gao, X (2006) Behavioral change, physiological modification, and metabolic detoxification: mechanisms of insecticide resistance. Acta Entomologica Sinica 48: 671-679.

Liu, Y and Shen, J (2003) Cuticular penetration mechanism of resistance to lambda--cyhalothrin in *Spodoptera exigua* (Hubner). Acta Entomologica Sinica 46: 288-291.

Loraine, AE, McCormick, S, Estrada, A, Patel, K and Qin, P (2013) RNA-Seq of Arabidopsis Pollen Uncovers Novel Transcription and Alternative Splicing. Plant physiology 162: 1092-1109.

Müller, H, Bracken, AP, Vernell, R, Moroni, MC, Christians, F, Grassilli, E, Prosperini, E,

Vigo, E, Oliner, JD and Helin, K (2001) E2Fs regulate the expression of genes involved in differentiation, development, proliferation, and apoptosis. Genes & development 15: 267-285.

MacNeil, LT and Walhout, AJ (2011) Gene regulatory networks and the role of robustness and stochasticity in the control of gene expression. Genome Researchearch 21: 645-657.

Maglich, JM, Stoltz, CM, Goodwin, B, Hawkins-Brown, D, Moore, JT and Kliewer, SA (2002) Nuclear pregnane x receptor and constitutive androstane receptor regulate overlapping but distinct sets of genes involved in xenobiotic detoxification. Molecular Pharmacology 62: 638-646.

Malik, A, Singh, N and Satya, S (2007) House fly (*Musca domestica*): a review of control strategies for a challenging pest. Journal of Environmental Science B 42: 453-469.

Malys, N and McCarthy, JE (2011) Translation initiation: variations in the mechanism can be anticipated. Cellular and Molecular Life Sciences 68: 991-1003.

Mangelsdorf, DJ, Thummel, C, Beato, M, Herrlich, P, Schutz, G, Umesono, K, Blumberg, B, Kastner, P, Mark, M and Chambon, P (1995) The nuclear receptor superfamily: the second decade. Cell 83: 835-839.

Mannikko, R, Elinder, F and Larsson, HP (2002) Voltage-sensing mechanism is conserved among ion channels gated by opposite voltages. Nature 419: 837-841.

Mardis, ER (2008) The impact of next-generation sequencing technology on genetics. Trends in Genetics 24: 133-141.

Marguerat, S and Bahler, J (2010) RNA-seq: from technology to biology. Cellular and Molecular Life Sciences 67: 569-579.

Marshall, C (1995) Specificity of receptor tyrosine kinase signaling: transient versus

sustained extracellular signal-regulated kinase activation. Cell 80: 179-185.

Martinez-Torres, D, Chandre, F, Williamson, MS, Darriet, F, Berge, JB, Devonshire, AL, Guillet, P, Pasteur, N and Pauron, D (1998) Molecular characterization of pyrethroid knockdown resistance (kdr) in the major malaria vector *Anopheles gambiae* s.s. Insect Molecular Biology 7: 179-184.

Martinez-Torres, D, Chevillon, C, Brun-Barale, A, Bergé, JB, Pasteur, N and Pauron, D (1999a) Voltage-dependent Na+ channels in pyrethroid-resistant Culex pipiens L mosquitoes. Pesticide Science 55: 1012-1020.

Martinez-Torres, D, Foster, SP, Field, LM, Devonshire, AL and Williamson, MS (1999b) A sodium channel point mutation is associated with resistance to DDT and pyrethroid insecticides in the peach-potato aphid, *Myzus persicae* (Sulzer) (Hemiptera: Aphididae). Insect Molecular Biology 8: 339-346.

Matthews, KS and Nichols, JC (1997) Lactose Repressor Protein: Functional Properties and Structure. In: Progress in Nucleic Acid Research and Molecular Biology (Kivie, M, ed. Vol. Volume 58, pp. 127-164. Academic Press.

Melander, AL (1914) Can insects become resistant to sprays. Journal of Economic Entomology 7: 167-173.

Michael, MD, Michael, LF and Simpson, ER (1997) A CRE-like sequence that binds CREB and contributes to cAMP-dependent regulation of the proximal promoter of the human aromatase P450 (CYP19) gene. Molecular and Cellular Endocrinology 134: 147-156.

Miller, C (1989) Genetic manipulation of ion channels: a new approach to structure and mechanism. Neuron 2: 1195-1205.

Miyazaki, M, Matsumura, F and Beeman, RW (1995) DNA sequence and site of mutation of the GABA receptor of cyclodiene-resistant red flour beetle, *Tribolium castaneum*. Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology 111: 399-406.

Morin, R, Bainbridge, M, Fejes, A, Hirst, M, Krzywinski, M, Pugh, T, McDonald, H, Varhol, R, Jones, S and Marra, M (2008) Profiling the HeLa S3 transcriptome using randomly primed cDNA and massively parallel short-read sequencing. Biotechniques 45: 81-94.

Mota-Sanchez, D, Hollingworth, RM, Grafius, EJ and Moyer, DD (2006) Resistance and cross-resistance to neonicotinoid insecticides and spinosad in the Colorado potato beetle, *Leptinotarsa decemlineata* (Say) (Coleoptera: Chrysomelidae). Pest Management Science 62: 30-37.

Mullin, CA, Scott, JG, American Chemical Society. Division of Agrochemicals. and American Chemical Society. Meeting (1992) Molecular mechanisms of insecticide resistance : diversity among insects. American Chemical Society, Washington, DC. xiii, 322 p. pp.

Murataliev, MB, Arino, A, Guzov, VM and Feyereisen, R (1999) Kinetic mechanism of cytochrome P450 reductase from the house fly (*Musca domestica*). Insect Biochemistry and Molecular Biology 29: 233-242.

Murray, M, Cui, PH and Zhou, F (2010) Roles of mitogen-activated protein kinases in the regulation of CYP genes. Current Drug Metabolism 11: 850-857.

Mutero, A, Pralavorio, M, Bride, JM and Fournier, D (1994) Resistance-associated point mutations in insecticide-insensitive acetylcholinesterase. Proceedings of the National Academy of Sciences 91: 5922-5926.

Nakajima, M and Yokoi, T (2011) MicroRNAs from biology to future pharmacotherapy:

Regulation of cytochrome P450s and nuclear receptors. Pharmacology & Therapeutics 131: 330-337.

Narahashi, T and American Chemical Society. (1979) Neurotoxicology of insecticides and pheromones. Plenum Press, New York. viii, 308 p. pp.

Nebert, DW, Adesnik, M, Coon, MJ, Estabrook, RW, Gonzalez, FJ, Guengerich, FP, Gunsalus, IC, Johnson, EF, Kemper, B, Levin, W and et al. (1987) The P450 gene superfamily: recommended nomenclature. DNA 6: 1-11.

Nestler, EEJ, Hyman, SE and Malenka, RC (2001) Molecular neuropharmacology. McGraw-Hill, Medical Pub. Division.

Nicolae, M, Mangul, S, Mandoiu, II and Zelikovsky, A (2011) Estimation of alternative splicing isoform frequencies from RNA-Seq data. Algorithms for Molecular Biology 6: 9.

Ode, PJ (2006) Plant chemistry and natural enemy fitness: effects on herbivore and natural enemy interactions. Annual Review of Entomology. 51: 163-185.

Ortelli, F, Rossiter, LC, Vontas, J, Ranson, H and Hemingway, J (2003) Heterologous expression of four glutathione transferase genes genetically linked to a major insecticide-resistance locus from the malaria vector *Anopheles gambiae*. Biochemical Journal 373: 957-963.

Ozsolak, F and Milos, PM (2011) RNA sequencing: advances, challenges and opportunities. Nature Reviews Genetics 12: 87-98.

Pan, Q, Shai, O, Lee, LJ, Frey, BJ and Blencowe, BJ (2008) Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. Nature Genetics 40: 1413-1415.

Pan, Y-Z, Gao, W and Yu, A-M (2009) MicroRNAs regulate CYP3A4 expression via direct

and indirect targeting. Drug Metabolism and Disposition 37: 2112-2117.

Pascussi, J-M, Gerbal-Chaloin, S, Fabre, J-M, Maurel, P and Vilarem, M-J (2000) Dexamethasone enhances constitutive androstane receptor expression in human hepatocytes: consequences on cytochrome P450 gene regulation. Molecular Pharmacology 58: 1441-1450.

Perea, S, López-Ribot, JL, Kirkpatrick, WR, McAtee, RK, Santillán, RA, Martínez, M, Calabrese, D, Sanglard, D and Patterson, TF (2001) Prevalence of molecular mechanisms of resistance to azole antifungal agents in Candida albicans strains displaying high-level fluconazole resistance isolated from human immunodeficiency virus-infected patients. Antimicrobial Agents and Chemotherapy 45: 2676-2684.

Perez-Mendoza, J, Fabrick, JA, Zhu, KY and Baker, JE (2000) Alterations in esterases are associated with malathion resistance in *Habrobracon hebetor* (Hymenoptera : Braconidae). Journal of Economic Entomology 93: 31-37.

Petersen, RA, Niamsup, H, Berenbaum, MR and Schuler, MA (2003) Transcriptional response elements in the promoter of CYP6B1, an insect P450 gene regulated by plant chemicals. Biochimica et Biophysica Acta (BBA) - General Subjects 1619: 269-282.

Picardi, E, Horner, DS, Chiara, M, Schiavon, R, Valle, G and Pesole, G (2010) Large-scale detection and analysis of RNA editing in grape mtDNA by RNA deep-sequencing. Nucleic Acids Research 38: 4755-4767.

Pittendrigh, B, Aronstein, K, Zinkovsky, E, Andreev, O, Campbell, B, Daly, J, Trowell, S and Ffrench-Constant, RH (1997) Cytochrome P450 genes from *Helicoverpa armigera*: expression in a pyrethroid-susceptible and -resistant strain. Insect Biochemistry and Molecular Biology 27: 507-512.

Plapp, FW, Jr. and Casida, JE (1969) Genetic control of house fly NADPH-dependent oxidases: relation to insecticide chemical metabolism and resistance. Journal of Economic Entomology 62: 1174-1179.

Plapp, FW, Jr. and Hoyer, RF (1968) Insecticide resistance in the house fly: decreased rate of absorption as the mechanism of action of a gene that acts as an intensifier of resistance. Journal of Economic Entomology 61: 1298-1303.

Prapanthadara, L-A, Koottathep, S, Promtet, N, Hemingway, J and Ketterman, AJ (1996) Purification and characterization of a major glutathione S-transferase from the mosquito Anopheles dirus (Species B). Insect Biochemistry and Molecular Biology 26: 277-285.

Ranson, H, Jensen, B, Wang, X, Prapanthadara, L, Hemingway, J and Collins, FH (2000) Genetic mapping of two loci affecting DDT resistance in the malaria vector *Anopheles gambiae*. Insect Molecular Biology 9: 499-507.

Rauch, N and Nauen, R (2004) Characterization and molecular cloning of a glutathione S-transferase from the whitefly *Bemisia tabaci* (Hemiptera: Aleyrodidae). Insect Biochemistry and Molecular Biology 34: 321-329.

Ray, JW (1967) The epoxidation of aldrin by housefly microsomes and its inhibition by carbon monoxide. Biochemical Pharmacology 16: 99-107.

Rencurel, F, Stenhouse, A, Hawley, SA, Friedberg, T, Hardie, DG, Sutherland, C and Wolf, CR (2005) AMP-activated protein kinase mediates phenobarbital induction of CYP2B gene expression in hepatocytes and a newly derived human hepatoma cell line. Journal of Biological Chemistry 280: 4367-4373.

Riveron, JM, Irving, H, Ndula, M, Barnes, KG, Ibrahim, SS, Paine, MJI and Wondji, CS

(2013) Directionally selected cytochrome P450 alleles are driving the spread of pyrethroid resistance in the major malaria vector *Anopheles funestus*. Proceedings of the National Academy of Sciences 110: 252-257.

Rosenthal, AZ, Matson, EG, Eldar, A and Leadbetter, JR (2011) RNA-seq reveals cooperative metabolic interactions between two termite-gut spirochete species in co-culture. ISME JOURNAL 5: 1133-1142.

Russell, RJ, Claudianos, C, Campbell, PM, Horne, I, Sutherland, TD and Oakeshott, JG (2004) Two major classes of target site insensitivity mutations confer resistance to organophosphate and carbamate insecticides. Pesticide Biochemistry and Physiology 79: 84-93.

Salinas, AE and Wong, MG (1999) Glutathione S-transferases--a review. Current Medicinal Chemistry 6: 279-309.

Sams, C, Mason, HJ and Rawbone, R (2000) Evidence for the activation of organophosphate pesticides by cytochromes P450 3A4 and 2D6 in human liver microsomes. Toxicology letters 116: 217-221.

Scharf, ME, Meinke, LJ, Siegfried, BD, Wright, RJ and Chandler, LD (1999) Carbaryl susceptibility, diagnostic concentration determination, and synergism for US populations of western corn rootworm (Coleoptera: Chrysomelidae). Journal of Economic Entomology 92: 33-39.

Schuler, TH, Martinez-Torres, D, Thompson, AJ, Denholm, I, Devonshire, AL, Duce, IR and Williamson, MS (1998) Toxicological, Electrophysiological, and Molecular Characterisation of Knockdown Resistance to Pyrethroid Insecticides in the Diamondback Moth, *Plutella xylostella* (L.). Pesticide Biochemistry and Physiology 59: 169-182.

Scott, JG (1989) Cross-resistance to the biological insecticide abamectin in pyrethroid-resistant house flies. Pesticide Biochemistry and Physiology 34: 27-31.

Scott, JG (1991) Investigating mechanisms of insecticide resistance: methods, strategies, and pitfalls. In: Pesticide resistance in arthropods. pp. 39-57. Springer.

Scott, JG (1999) Cytochromes P450 and insecticide resistance. Insect Biochemistry and Molecular Biology 29: 757-777.

Scott, JG and Georghiou, GP (1986a) The biochemical genetics of permethrin resistance in the Learn-PyR strain of house fly. Biochemical Genetics 24: 25-37.

Scott, JG and Georghiou, GP (1986b) Mechanisms responsible for high levels of permethrin resistance in the house fly. Pesticide Science 17: 195-206.

Scott, JG and Kasai, S (2004) Evolutionary plasticity of monooxygenase-mediated resistance. Pesticide Biochemistry and Physiology 78: 171-178.

Scott, JG, Liu, N, Kristensen, M and Clark, AG (2009) A case for sequencing the genome of *Musca domestica* (Diptera: Muscidae). Journal of Medical Entomology 46: 175-182.

Scott, JG and Wen, Z (2001) Cytochromes P450 of insects: the tip of the iceberg. Pest Mangement Science 57: 958-967.

Scott, JG and Zhang, L (2003) The house fly aliesterase gene (MdαE7) is not associated with insecticide resistance or P450 expression in three strains of house fly. Insect Biochemistry and Molecular Biology 33: 139-144.

Siomi, H and Siomi, MC (2010) Posttranscriptional Regulation of MicroRNA Biogenesis in Animals. Molecular Cell 38: 323-332.

Smith, TJ, Hyeock Lee, S, Ingles, PJ, Knipple, DC and M Soderlund, D (1997) The L1014F

point mutation in the house fly Vssc1 sodium channel confers knockdown resistance to pyrethroids. Insect Biochemistry and Molecular Biology 27: 807-812.

Soderlund David, M, Ingles Patricia, J, Lee Si, H, Smith Timothy, J and Knipple Douglas, C (2001) The Molecular Mechanism of Knockdown Resistance. In: Agrochemical Resistance. Vol. 808, pp. 77-89.

Soderlund, DM and Knipple, DC (2003) The molecular biology of knockdown resistance to pyrethroid insecticides. Insect Biochemistry and Molecular Biology 33: 563-577.

Sonoda, S and Tsumuki, H (2005) Studies on glutathione S-transferase gene involved in chlorfluazuron resistance of the diamondback moth, *Plutella xylostella* L. (Lepidoptera: Yponomeutidae). Pesticide Biochemistry and Physiology 82: 94-101.

Sorber, K, Dimon, MT and DeRisi, JL (2011) RNA-Seq analysis of splicing in Plasmodium falciparum uncovers new splice junctions, alternative splicing and splicing of antisense transcripts. Nucleic Acids Research 39: 3820-3835.

Struhl, K (1999) Fundamentally different logic of gene regulation in eukaryotes and prokaryotes. Cell 98: 1-4.

Sultan, M, Schulz, MH, Richard, H, Magen, A, Klingenhoff, A, Scherf, M, Seifert, M, Borodina, T, Soldatov, A and Parkhomchuk, D (2008) A global view of gene activity and alternative splicing by deep sequencing of the human transcriptome. Science 321: 956-960.

t Hoen, PA, Ariyurek, Y, Thygesen, HH, Vreugdenhil, E, Vossen, RH, de Menezes, RX, Boer, JM, van Ommen, GJ and den Dunnen, JT (2008) Deep sequencing-based expression analysis shows major advances in robustness, resolution and inter-lab portability over five microarray platforms. Nucleic Acids Research 36: e141-e141.

Tabashnik, BE, Liu, YB, Finson, N, Masson, L and Heckel, DG (1997) One gene in diamondback moth confers resistance to four Bacillus thuringiensis toxins. Proceedings of the National Academy of Sciences of the United States of America 94: 1640-1644.

Takagi, S, Nakajima, M, Kida, K, Yamaura, Y, Fukami, T and Yokoi, T (2010) MicroRNAs regulate human hepatocyte nuclear factor 4α, modulating the expression of metabolic enzymes and cell cycle. Journal of Biological Chemistry 285: 4415-4422.

Takagi, S, Nakajima, M, Mohri, T and Yokoi, T (2008) Post-transcriptional regulation of human pregnane X receptor by micro-RNA affects the expression of cytochrome P450 3A4. Journal of Biological Chemistry 283: 9674-9680.

Takemori, H, Doi, J, Horike, N, Katoh, Y, Min, L, Lin, X-z, Wang, Z-n, Muraoka, M and Okamoto, M (2003) Salt-inducible kinase-mediated regulation of steroidogenesis at the early stage of ACTH-stimulation. The Journal of Steroid Biochemistry and Molecular Biology 85: 397-400.

Thompson, M, Steichen, JC and Ffrench-Constant, RH (1993) Conservation of cyclodiene insecticide resistance-associated mutations in insects. Insect Molecular Biology 2: 149-154.

Townsend, MG and Busvine, JR (1969) The mechanism of malathion-resistance in the blowfly, *Chrysomya putoria*. Entomologia Experimentalis et Applicata 12: 243-267.

Troy, CM and Shelanski, ML (1994) Down-regulation of copper/zinc superoxide dismutase causes apoptotic death in PC12 neuronal cells. Proceedings of the National Academy of Sciences 91: 6384-6387.

Tsuchiya, Y, Nakajima, M, Takagi, S, Taniya, T and Yokoi, T (2006) MicroRNA regulates the expression of human cytochrome P450 1B1. Cancer research 66: 9090-9098.

Ule, J, Jensen, K, Mele, A and Darnell, RB (2005) CLIP: A method for identifying protein–RNA interaction sites in living cells. Methods 37: 376-386.

Vaughan, A and Hemingway, J (1995) Mosquito carboxylesterase Est alpha 2(1) (A2). Cloning and sequence of the full-length cDNA for a major insecticide resistance gene worldwide in the mosquito *Culex quinquefasciatus*. Journal of Biological Chemistry 270: 17044-17049.

Villatte, F, Ziliani, P, Marcel, V, Menozzi, P and Fournier, D (2000) A High Number of Mutations in Insect Acetylcholinesterase May Provide Insecticide Resistance. Pesticide Biochemistry and Physiology 67: 95-102.

Vontas, J, Hejazi, M, Hawkes, N, Cosmidis, N, Loukas, M and Hemingway, J (2002a) Resistance ssociated point mutations of organophosphate insensitive acetylcholinesterase, in the olive fruit fly *Bactrocera oleae*. Insect Molecular Biology 11: 329-336.

Vontas, JG, Small, GJ, Nikou, DC, Ranson, H and Hemingway, J (2002b) Purification, molecular cloning and heterologous expression of a glutathione S-transferase involved in insecticide resistance from the rice brown planthopper, *Nilaparvata lugens*. Biochemical Journal 362: 329-337.

Walsh, CT, Fisher, SL, Park, IS, Prahalad, M and Wu, Z (1996) Bacterial resistance to vancomycin: five genes and one missing hydrogen bond tell the story. Chemistry & Biology 3: 21-328.

Walsh, SB, Dolden, TA, Moores, GD, Kristensen, M, Lewis, T, Devonshire, AL and Williamson, MS (2001) Identification and characterization of mutations in housefly (*Musca domestica*) acetylcholinesterase involved in insecticide resistance. Biochemical Journal 359: 175-181.

Wang, ET, Sandberg, R, Luo, S, Khrebtukova, I, Zhang, L, Mayr, C, Kingsmore, SF, Schroth, GP and Burge, CB (2008) Alternative isoform regulation in human tissue transcriptomes. Nature 456: 470-476.

Wang, H and LeCluyse, EL (2003) Role of orphan nuclear receptors in the regulation of drug-metabolising enzymes. Clinical pharmacokinetics 42: 1331-1357.

Wang, Z, Gerstein, M and Snyder, M (2009) RNA-Seq: a revolutionary tool for transcriptomics. Nature Reviews Genetics 10: 57-63.

Waters, LC, Zelhof, AC, Shaw, BJ and Ch'ang, LY (1992) Possible involvement of the long terminal repeat of transposable element 17.6 in regulating expression of an insecticide resistance-associated P450 gene in *Drosophila*. Proceedings of the National Academy of Sciences USA 89: 4855-4859.

Waxman, DJ (1999) P450 gene induction by structurally diverse xenochemicals: central role of nuclear receptors CAR, PXR, and PPAR. Archives of Biochemistry and Biophysics 369: 11-23.

Wei, SH, Clark, AG and Syvanen, M (2001) Identification and cloning of a key insecticide-metabolizing glutathione S-transferase (MdGST-6A) from a hyper insecticide-resistant strain of the housefly *Musca domestica*. Insect Biochemistry and Molecular Biology 31: 1145-1153.

Wen, Z and Scott, JG (1999) Genetic and biochemical mechanisms limiting fipronil toxicity in the LPR strain of house fly, *Musca domestica*. Pesticide Science 55: 988-992.

Wheelock, GD and Scott, JG (1992) The role of cytochrome P450lpr in deltamethrin metabolism by pyrethroid-resistant and susceptible strains of house flies. Pesticide Biochemistry and Physiology 43: 67-77.

Wilding, CS, Smith, I, Lynd, A, Yawson, AE, Weetman, D, Paine, MJI and Donnelly, MJ (2012) A cis-regulatory sequence driving metabolic insecticide resistance in mosquitoes: Functional characterisation and signatures of selection. Insect Biochemistry and Molecular Biology 42: 699-707.

Wilhelm, BT, Marguerat, S, Watt, S, Schubert, F, Wood, V, Goodhead, I, Penkett, CJ, Rogers, J and Bahler, J (2008) Dynamic repertoire of a eukaryotic transcriptome surveyed at single-nucleotide resolution. Nature 453: 1239-1243.

Wilson, TG and Hodgson, E (1971) Microsomal NADPH-cytochrome c reductase from the housefly, *Musca domestica*: Properties of the purified enzyme. Insect Biochemistry 1: 171-180.

Workman, JL and Kingston, RE (1998) Alteration of nucleosome structure as a mechanism of transcriptional regulation. Annual Review of Biochemistry 67: 545-579.

Xu, Q, Wang, H, Zhang, L and Liu, N (2006a) Kdr allelic variation in pyrethroid resistant mosquitoes, *Culex quinquefasciatus* (S.). Biochemical and Biophysical Research Communications 345: 774-780.

Xu, Q, Wang, H, Zhang, L and Liu, N (2006b) Sodium channel gene expression associated with pyrethroid resistant house flies and German cockroaches. Gene 379: 62-67.

Yang, SS, Tu, ZJ, Cheung, F, Xu, WW, Lamb, JF, Jung, HJ, Vance, CP and Gronwald, JW (2011) Using RNA-Seq for gene identification, polymorphism detection and transcript profiling in two alfalfa genotypes with divergent cell wall composition in stems. BMC Genomics 12: 199.

Yang, T and Liu, N (2011) Genome analysis of cytochrome P450s and their expression profiles in insecticide resistant mosquitoes, *Culex quinquefasciatus*. PloS One 6: e29418.

Yasunami, Y, Hara, H, Iwamura, T, Kataoka, T and Adachi, T (2004) C-jun N-terminal

kinase modulates 1,25-dihydroxyvitamin D3-induced cytochrome P450 3A4 gene expression. Drug Metabolism and Disposition 32: 685-688.

Yurank, HY and Catterall, WA (2003) Overview of the voltage-gated sodium channel family. Genome Biology 4: 207-212.

Zagrobelny, M, Bak, S, Rasmussen, AV, Jørgensen, B, Naumann, CM and Lindberg Møller, B (2004) Cyanogenic glucosides and plant–insect interactions. Phytochemistry 65: 293-306.

Zerbino, DR and Birney, E (2008) Velvet: algorithms for de novo short read assembly using de Bruijn graphs. Genome Research 18: 821-829.

Zhang, L, Gao, X and Liang, P (2007) Beta-cypermethrin resistance associated with high carboxylesterase activities in a strain of house fly, *Musca domestica* (Diptera: Muscidae). Pesticide Biochemistry and Physiology 89: 65-72.

Zhang, M and Scott, JG (1994) Cytochrome b5 involvement in cytochrome P450 monooxygenase activities in house fly microsomes. Arch Insect Biochem Physiol 27: 205-216.

Zhao, C, Waalwijk, C, de Wit, PJ, Tang, D and van der Lee, T (2013a) RNA-Seq analysis reveals new gene models and alternative splicing in the fungal pathogen Fusarium graminearum. BMC Genomics 14: 21.

Zhao, G, Liu, W and Knowles, CO (1994) Mechanisms associated with diazinon resistance in western flower thrips. Pesticide Biochemistry and Physiology 49: 13-23.

Zhao, Q, Caballero, OL, Levy, S, Stevenson, BJ, Iseli, C, de Souza, SJ, Galante, PA, Busam, D, Leversha, MA, Chadalavada, K, Rogers, YH, Venter, JC, Simpson, AJ and Strausberg, RL (2009) Transcriptome-guided characterization of genomic rearrangements in a breast cancer cell line. Proceedings of the National Academy of Sciences 106: 1886-1891.

Zhao, S-S, Zhao, G-D, Di, T-Y, Ding, H, Wan, X-L, Li, B, Chen, Y-H, Xu, Y-X, Shen, W-D and Wei, Z-G (2013b) Analyzing the promoters of two CYP9A genes in the silkworm Bombyx mori by dual-luciferase reporter assay. Molecular Biology Reports 40: 1701-1710.

Zhou, J, Zhang, G and Zhou, Q (2012) Molecular characterization of cytochrome P450 CYP6B47 cDNAs and 5′-flanking sequence from *Spodoptera litura* (Lepidoptera: Noctuidae): Its response to lead stress. Journal of Insect Physiology 58: 726-736.

Zhu, F, Feng, JN, Zhang, L and Liu, N (2008a) Characterization of two novel cytochrome P450 genes in insecticide-resistant house-flies. Insect Molecular Biology 17: 27-37.

Zhu, F, Li, T, Zhang, L and Liu, N (2008b) Co-up-regulation of three P450 genes in response to permethrin exposure in permethrin resistant house flies, *Musca domestica*. BMC Physiol 8: 18.

Zhu, F and Liu, N (2008) Differential expression of CYP6A5 and CYP6A5v2 in pyrethroid-resistant house flies, *Musca domestica*. Arch Insect Biochem Physiol 67: 107-119.

Zhu, F, Parthasarathy, R, Bai, H, Woithe, K, Kaussmann, M, Nauen, R, Harrison, DA and Palli, SR (2010) A brain-specific cytochrome P450 responsible for the majority of deltamethrin resistance in the QTC279 strain of *Tribolium castaneum*. Proceedings of the National Academy of Sciences 107: 8557-8562.

Zhu, KY, Lee, SH and Clark, JM (1996) A Point Mutation of Acetylcholinesterase Associated with Azinphosmethyl Resistance and Reduced Fitness in Colorado Potato Beetle. Pesticide Biochemistry and Physiology 55: 100-108.

Zlotkin, E (1999) The insect voltage-gated sodium channel as target of insecticides. Annual Review of Entomology 44: 429-455.

Chapter 2: Research Goal and Specific Objectives

2.1 The goal of the research and objectives

To fill the gaps in our knowledge about the importance of P450 genes in insecticide resistance and to improve understanding of P450 gene regulation pathways, the long term goal of my research is to characterize the P450 gene expression profiles and elucidate the P450 gene regulation pathways in insecticide-resistant house flies. To achieve this long term goal, the following objectives will be addressed in this research: 1) The whole transcriptomal linkage analysis of gene co-regulation in insecticide-resistant house flies. 2) Characterization of the expression profiles of P450 genes in different house fly strains. 3) Genetic mapping and autosomal linkage analysis of the up-regulated P450 genes in insecticide resistant house flies; 4) Characterization of the key P450 genes function in insecticide-resistant house flies. 5) Homology modeling and permethrin docking study of P450s

2.1.1 Objective 1: The whole transcriptomal linkage analysis of gene co-regulation in insecticide resistant house flies

In my study, three house fly strains, ALHF, aabys and CS, will be used. ALHF is a multi-insecticide resistant strain (Liu and Yue 2000) collected from a poultry farm in Alabama in 1998. This strain was further selected with permethrin for six generations after collection to reach a high level of resistance and has been subsequently maintained under biannual selection with

permethrin (Liu and Yue 2001; Tian et al. 2011). CS is a wild-type insecticide-susceptible strain kept in laboratory breeding for more than five decades. aabys is an insecticide-susceptible strain with recessive morphological markers of ali-curve (*ac*), aristapedia (*ar*), brown body (*bwb*), yellow eyes (*ye*), and snipped wings (*snp*) on autosomes 1, 2, 3, 4, and 5, respectively. In addition, five house fly lines, A2345, A1345, A1245, A1235, and A1234 generated from genetic crosses of aabys and ALHF. These five lines contain different autosomal combinations from ALHF and will be used in the research to characterize the genetic linkages of insecticide resistance.

The RNA extraction from each strain or line was performed three times with different fly samples on different days to provide biological replications for the RNA-Seq experiments (ALHF) or, later, as the replications of qRT-PCR experiments. Illumina HiSeq 2000 RNA-Seq was conducted by the Hudson Alpha Institute of Biotechnology (Huntsville, AL). The first reference transcriptome of the house fly was assembled *de novo* using Trinity. The reference transcriptome was then used as the common reference for the estimation of the gene expression values for each of the strains. Several genes with general function categories of metabolism, such as P450s, and regulation, such as G-protein coupled receptors (GPCRs), adenylate and guanylatecyclases, protein kinases (PKs), proteases, and DNA binding domain genes were identified from the reference transcriptome analysis of house flies.

2.1.2 Objective 2: Characterization of the expression profiles of P450 genes in different house fly strains

P450s associated insecticide resistance is due to increase the rate of detoxification of

insecticides by constitutively transcriptional up-regulation of P450 genes, resulting in increased levels of P450 proteins and P450 activities. *CYP4D4V2*, *CYP6A1*, *CYP6A5*, *CYP6A5V2*, *CYP6A24*, *CYP6D1*, *CYP6D3*, *CYP6A36* and *CYP12A1* have been found to be induced by insecticides or over-expressed in different insecticide-resistant house fly strains. The expression profiles of all P450 genes identified from objective 1 will be conducted by qRT-PCR. I hypothesize that several of P450 genes are over-expressed in the insecticide-resistant house flies when compared to susceptible house flies, and these P450 genes may play critical roles in the development of insecticide resistance in insecticide-resistant house flies.

2.1.3 Objective 3: Genetic mapping and autosomal linkage analysis of the up-regulated P450 genes in insecticide-resistant house flies

My third objective will focus on the autosome location and autosomal linkage of up-regulated P450 genes detected from objective 2. ALHF, aabys, A2345, A1345, A1245, A1235, and A1234 house fly strains and lines will be used to conduct the genetic mapping and autosomal linkage of up-regulated P450 genes. Allele specific PCR will be conducted using the cDNA from the 7 house fly strains/lines. Two rounds of PCR will be conducted. For the first PCR reaction, we will use the regular primer pairs to generate cDNA fragments, respectively. The second PCR will employ 0.5 µl of the first round PCR reaction solution and an allele specific primer pair, designed by placing a specific nucleotide polymorphism at the 3’ end of the primer to permit preferential amplification of the allele from ALHF based its gene specific sequence. Then, qRT-PCR will be used to detect the autosomal interaction of P450 genes. I hypothesize that several up-regulated P450 genes will be mapped on different autosomes in the ALHF strain, especially

autosome 5, the importance of factors of which has been extensively reported in insecticide resistance. We also expect that the expressions of the majority of up-regulated P450 genes are controlled by multiple factors/autosomes.

2.1.4 Objective 4: Characterization of the key P450 genes function in insecticide-resistant house flies

Once we have mapped the up-regulated P450 genes on the different autosomes, we will characterize the function of the up-regulated genes in insecticide-resistant house flies. We will select key P450 genes that identified from objective 1, 2 and 3 to investigate their function by transgenic technology. Construction of transgenic *Drosophila melanogaster* lines will be applied to validate the functions of the P450 genes. Briefly, selected P450 genes will be constructed into a pUASTattB vector and transformed into the germline of *D.melanogaster.* Permethrin toxicity bioassays will be then conducted on a 2-3 day post eclosion female *Drosophila* to examine the toxicity of permethrin to transgenic flies. Bioassay will be independently replicated three times. We predict that the expression of any one of the transformed key P450 genes in *Drosophila* will increase the resistance level of *Drosophila* to permethrin.

2.1.5 Objective 5: Homology modeling and permethrin docking study of P450s

We will select key p450 genes from objective 4 to conduct homology modeling and permehtrin docking study. Structural modeling will be performed by the I-TASSER server with the combined methods of threading, *ab initio* modeling (Roy et al. 2010; Zhang 2008). Five models will be predicted by the I-TASSER for each P450. We will submit the top scoring model

to the FG-MD server for fragment guided molecular dynamics structure refinement (Zhang et al. 2011). The Model quality will be controlled by Ramachandran plots generated with Procheck (http://services.mbi.ucla.edu/SAVES/) (Laskowski et al. 1993) and ProSA-web (https://prosa.services.came.sbg.ac.at/prosa.php) (Sippl 1993; Wiederstein and Sippl 2007). Channels of P450s will be calculated by CAVER 3.0 (http://caver.cz/index.php) (Chovancova et al. 2012; Medek et al. 2007), the passage of a sphere of maximal radius greater than 1.2 Å are considered tabulated, and named according to nomenclature of Cojocaru et al. (Cojocaru et al. 2007). The volume of the substrate binding cavity will be characterized by VOIDOO with a 1.4 Å probe (Kleywegt et al. 2001), Proteins and ligands will be prepared for docking with Autodock Tools v1.5.6 (http://mgltools.scripps.edu/downloads). Molecular docking will be performed by Autodock 4.2. (Morris et al. 2009). Ligand permethrin structures will be retrieved from the ZINC database (Irwin et al. 2012). For all dockings, a search space with a grid box of 60 x 60 x 60 Å, centered at the heme iron will be set corresponding to substrate recognition sites (SRSs) following those of the CYP2 family proposed by Gotoh (Gotoh 1992). The figures will be produced by Pymol (http://www.pymol.org/) (DeLano 2002).

2.2 Significance of research

Insect cytochrome P450s are critical for detoxification of xenobiotics and play a fundamental role in biosynthesis and degradation of endogenous compounds. Elucidation of the regulation pathway and P450s that involved in insecticide resistance will provide us with a better understanding of the post-regulation processes involved in insecticide resistance development in the house fly and will be important for developing new control strategies for many insect pests,

including house flies.

2.3 Reference

Chovancova, E, Pavelka, A, Benes, P, Strnad, O, Brezovsky, J, Kozlikova, B, Gora, A, Sustr, V, Klvana, M and Medek, P (2012) CAVER 3.0: a tool for the analysis of transport pathways in dynamic protein structures. PLoS Computational Biology **8**: e1002708.

Cojocaru, V, Winn, PJ and Wade, RC (2007) The ins and outs of cytochrome P450s. Biochimica et Biophysica Acta (BBA)-General Subjects 1770: 390-401.

DeLano, WL (2002) The PyMOL molecular graphics system. CA. USA.

Gotoh, O (1992) Substrate recognition sites in cytochrome P450 family 2 (CYP2) proteins inferred from comparative analyses of amino acid and coding nucleotide sequences. Journal of Biological Chemistry 267: 83-90.

Kleywegt, G, Zou, J-Y, Kjeldgaard, M and Jones, T (2001) Around O. International Tables for Crystallography Volume F: Crystallography of Biological Macromolecules: 353-356.

Laskowski, RA, MacArthur, MW, Moss, DS and Thornton, JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. Journal of Applied Crystallography 26: 283-291.

Liu, N and Yue, X (2000) Insecticide resistance and cross-resistance in the house fly (Diptera: Muscidae). Journal of Economic Entomology 93: 1269-1275.

Liu, N and Yue, X (2001) Genetics of pyrethroid resistance in a strain (ALHF) of house flies (Diptera: Muscidae). Pesticide Biochemistry and Physiology 70: 151-158.

Medek, P, Benes, P and Sochor, J (2007) Computation of Tunnels in Protein Molecules

using Delaunay Triangulation. Journal of WSCG 15: 107-114.

Morris, GM, Huey, R, Lindstrom, W, Sanner, MF, Belew, RK, Goodsell, DS and Olson, AJ (2009) AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. Journal of Computational Chemistry 30: 2785-2791.

Roy, A, Kucukural, A and Zhang, Y (2010) I-TASSER: a unified platform for automated protein structure and function prediction. Nature protocols 5: 725-738.

Scott, HG and Littig, KS (1962) Flies of public health importance and their control. US Department of Health, Education, and Welfare, Public Health Service, Communicable Disease Center.

Scott, JG (1999) Cytochromes P450 and insecticide resistance. Insect Biochemistry and Molecular Biology 29: 757-777.

Tian, L, Cao, C, He, L, Li, M, Zhang, L, Liu, H and Liu, N (2011) Autosomal interactions and mechanisms of pyrethroid resistance in house flies, *Musca domestica*. International Journal of Biological Sciences 7: 902-911.

Wiederstein, M and Sippl, MJ (2007) ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Researchearch 35: W407-W410.

Zhang, J, Liang, Y and Zhang, Y (2011) Atomic-level protein structure refinement using fragment-guided molecular dynamics conformation sampling. Structure 19: 1784-1795.

Zhang, Y (2008) I-TASSER server for protein 3D structure prediction. BMC Bioinformatics 9: 40.

Zhu, F, Feng, JN, Zhang, L and Liu, N (2008a) Characterization of two novel cytochrome

P450 genes in insecticide-resistant house-flies. Insect Molecular Biology 17: 27-37.

Zhu, F, Li, T, Zhang, L and Liu, N (2008b) Co-up-regulation of three P450 genes in response to permethrin exposure in permethrin resistant house flies, *Musca domestica*. BMC Physiol 8: 18.

Zhu, F and Liu, N (2008) Differential expression of CYP6A5 and CYP6A5v2 in pyrethroid-resistant house flies, *Musca domestica*. Arch Insect Biochem Physiol 67: 107-119.

Chapter 3: A Whole Transcriptomal Linkage Analysis of Gene Co-Regulation in Insecticide Resistant House Flies, *Musca domestica*

3.1 Abstract

The first reference transcriptome from the adult house fly was generated and a whole transcriptome analysis was conducted for the multiple insecticide resistant strain ALHF (wild-type) and two insecticide susceptible strains: aabys (with morphological recessive markers) and CS (wild type) to gain valuable insights into the gene interaction and complex regulation involved in the development of insecticide resistance in house flies, *Musca domestica*. Over 56 million reads were used to assemble the adult female *M. domestica* transcriptome reference and 14488 contigs were generated from the *de novo* transcriptome assembly. A total of 6159 (43%) of the contigs contained coding regions, among which 1316 genes were identified as being co-up-regulated in ALHF in comparison to both aabys and CS. The majority of these up-regulated genes fell within the SCOP categories of metabolism, general, intra-cellular processes, and regulation, and covered three key SCOP detailed function categories: redox detailed function category in metabolism, signal transduction and kinases/phosphatases in regulation, and proteases in intra-cellular processes. The redox detailed function group contained gene superfamilies for detoxification, including multiple cytochrome P450s, glutathione S-transferases, and esterases. The signal transduction and kinases/phosphatases detailed function groups contained gene

families including 7 transmembrane receptors (rhodopsin-like GPCR family), adenylate and guanylate cyclases, protein kinases and phosphatases. The proteases detailed function group contained genes with digestive, catalytic, and proteinase activities. Genetic linkage analysis with house fly lines comparing different autosomal combinations from ALHF revealed that the up-regulation of gene expression in the three key SCOP detailed function categories occurred mainly through the co-regulation of factors among multiple autosomes, especially between autosomes 2 and 5, suggesting that signaling transduction cascades controlled by GPCRs, protein kinase/phosphates and proteases may be involved in the regulation of resistance P450 gene regulation.

3.2 Introduction

Insecticides have a major impact on agriculture, economy, and public health due to their outstanding contribution towards controlling agriculturally, medically, and economically important insect pests worldwide. Nevertheless, the development of resistance to insecticides in diverse insect pests is becoming a global problem in the insect pest control battle (Hemingway et al. 2002). Resistance is thought to be a pre-adaptive phenomenon, in that prior to insecticide exposure rare individuals already exist who carry an altered genome that results in one or more possible mechanisms (factors) allowing survival from the selection pressure of insecticides (Brattsten et al. 1986; Sawicki and Denholm 1984), and overall, the rate of development of resistance in field populations of insects depends upon the levels of genetic variability in a population (Liu and Scott 1995; Liu and Yue 2001). Efforts to characterize the genetic variation involved in insecticide resistance have thus focused on building a better fundamental

understanding of the development of resistance and studying resistance mechanisms, both of which are vital for practical applications such as the design of novel strategies to prevent or minimize the spread and evolution of resistance development and the control of insect pests (Roush et al. 1990).

There is considerable evidence to suggest that the interaction of multiple insecticide resistance mechanisms or genes is responsible for the development of insecticide resistance (Georghiou 1971; Liu et al. 2005; Liu et al. 2007; Liu and Scott 1996; Liu and Yue 2000; Plapp Jr et al. 1987; Pridgeon and Liu 2003; Raymond et al. 1989; Vontas et al. 2005; Xu et al. 2005; Xu et al. 2006b). While altering target site sensitivity to insecticides has been shown to reduce insects' response to insecticides, transcriptional up-regulation of the detoxification machinery, increasing metabolism of insecticides into less harmful substances and facilitating insecticide excretion are all known to play a role in allowing insects to defend themselves against insecticides (Xu et al. 2005). This detoxification machinery in insects has been mainly attributed to three enzyme systems, namely cytochrome P450s, esterases, and glutathione S-transferases (GSTs), the up-regulation of which underlies the development of insecticide resistance in many insect species. It has been suggested that new patterns of gene expression may arise via a variety of mechanisms involving changes to upstream regulators (change in trans) and mutations of the noncoding regulatory DNA sequences (e.g., enhancers) of a gene (change in cis) (Rebeiz et al. 2011). Indeed, many studies on the development of insecticide resistance in insects have demonstrated different patterns of gene expression between resistant and susceptible insect populations. Many studies have also found that the different patterns of gene expression in metabolic detoxification of insecticide-resistant insects are regulated by trans and/or cis factors (Liu and Scott 1995; Liu and Scott 1996; Liu and Yue 2001;

Plapp Jr et al. 1987). Taken together, these studies suggest that not only is insecticide resistance conferred via multiple gene up-regulation, but it is mediated through the interaction of regulatory factors and resistance genes. However, no regulatory factors in insecticide resistance have yet been identified, and there has been no examination of the regulatory interaction of resistance genes. Recent advances in DNA sequencing technology have provided an opportunity for genome/whole transcriptome-wide gene discovery in organisms, including those genes suspected of involvement in insecticide resistance and the factors that may be involved in the regulation of resistance genes and mechanisms.

The house fly, *Musca domestica*, is a major domestic, medical and veterinary pest that causes more than 100 human and animal intestinal diseases, including bacterial infections such as salmonellosis, anthrax ophthalmia, shigellosis, typhoid fever, tuberculosis, cholera and infantile diarrhea; protozoan infections such as amebic dysentery; helminthic infections such as pinworms, roundworms, hookworms and tapeworms; and both viral and rickettsial infections (Greenberg 1965; Keiding 1986; Scott and Littig 1962; Scott et al. 2009). Current approaches to control house flies rely primarily on source reduction and the application of insecticides, generally pyrethroids, organophosphates, neonicotinoids, as well as chitin synthesis inhibiting/disrupting larvicides. However, house flies have shown a remarkable ability to develop not only resistance to the insecticide used against them but also cross-resistance to unrelated classes of insecticides (Greenberg 1965; Liu and Yue 2000; Pap and Tóth 1995; Wen and Scott 1997). Because of this ability and the relatively rapid rate at which they develop resistance and cross-resistance to insecticides, their well described linkage map for five autosomes and two sex chromosomes (X and Y) (Hiroyoshi 1964; Milani et al. 1967; Nickel and Wagoner 1974; Tsukamoto et al. 1961), and

their relatively well studied biochemistry and genetics of insecticide resistance, the house fly has proven to be a useful model for understanding and predicting resistance in other insect species.

The house fly strain ALHF has demonstrated the ability to develop resistance and/or cross-resistance to not only pyrethroids and organophosphates (OPs), but also relatively new insecticides such as fipronil and imidacloprid (Liu and Yue 2000; Tian et al. 2011). Genetic studies have linked pyrethroid resistance to autosomes 1, 2, 3 and 5 (Liu and Yue 2001). The major mechanisms governing pyrethroid resistance in this strain include increased detoxification mediated by P450 monooxygenases and decreased sensitivity of voltage-gated sodium channels (kdr) (Xu et al. 2006b; Zhu et al. 2008; Zhu and Liu 2008). Previous genetic studies of ALHF have linked pyrethroid resistance primarily to autosomes 2, 3 and 5, with a minor role played by factor(s) on autosome 1 (Liu and Yue 2000; Tian et al. 2011). Furthermore, multiple P450 genes, CYP6A5, CYP6A5v2, CYP6A36, CYP6A37, CYP4D4v2 and CYP6A38, that are known to be overexpressed in ALHF have been located on autosome 5 and the regulation of these P450 genes have been linked to autosomes 1 and 2 (Zhu et al. 2008; Zhu and Liu 2008). However, the precise nature of the interaction between the regulatory factors and resistance genes such as P450s is unclear. In an effort to better understand the genetic variation relation to resistance and gain valuable insights into the gene interaction and regulation involved in the development of permethrin resistance in the house fly, the current study generated the first adult transcriptome of the house fly *M. domestica* using Illumina RNA-Seq. Whole transcriptome comparative analyses were conducted for the resistant ALHF strain, susceptible CS and aabys strains, which enabled us to investigate the complete transcriptome of *M. domestica* and identify the genes that are most likely to be involved in pyrethroid resistance and their autosomal interactions.

3.3 Materials and methods

3.3.1 House fly strains and lines

Three house fly strains were used in this study. ALHF, a multi-insecticide resistant strain (Liu and Yue 2001) collected from a poultry farm in Alabama in 1998. This strain was further selected with permethrin in the laboratory for six generations after collection, reaching to a high level of resistance (Liu and Yue 2000; Tian et al. 2011). This strain has been maintained under biannual selection with permethrin. CS is a wild type insecticide-susceptible strain kept in laboratory breeding for more than five decades. aabys is an insecticide-susceptible strain with recessive morphological markers ali-curve (ac), aristapedia (ar), brown body (bwb), yellow eyes (ye), and snipped wings (snp) on autosomes 1, 2, 3, 4, and 5, respectively. Both CS and aabys were obtained from Dr. J. G. Scott (Cornell University).

A cross of ALHF female and aabys male was performed with each of ~400 flies. The F1 males (~400 flies) were then backcrossed to aabys female (Figure 3.1). Five back-cross (BC1) lines with the following genotypes were isolated: ac/ac, +/ar, +/bwb, +/ye, +/snp (A2345); +/ac, ar/ar, +/bwb, +/ye, +/snp (A1345); +/ac, +/ar, bwb/bwb, +/ye, +/snp (A1245); +/ac, +/ar, +/bwb, ye/ye, +/snp (A1235); and +/ac, +/ar, +/bwb, +/ye, snp/snp (A1234). Homozygous lines (+/+, +/+, bwb/bwb, +/+, +/+ (A1245); +/+, + /+, +/+, ye/ye, +/+ (A1235); +/+, +/+, +/+, +/+, snp/snp (A1234); +/+, ar /ar, +/+, +/+,+/+ (A1345); and ac/ac, +/+, +/+, +/+, +/+ (A2345)) were accomplished by sorting for appropriate phenotypic markers and selecting with permethrin at a corresponding dose that caused ~70% mortality for each of lines for three generations. One hundred single-pair crossing (n=100) of each of lines for the desired phenotype and genotype

were then set up (Liu and Yue 2000; Tian et al. 2011). The name of each line indicates which of its autosomes bear wild-type markers from ALHF. For instance, the A2345 strain has wild-type markers on autosomes 2, 3, 4, 5 from ALHF and the mutant marker on autosome 1 from aabys. A1235 strain (flies with a recessive mutant marker on autosome 4 from aabys) showed no significant differences in resistance levels compared with ALHF based on the overlapping 95% confidence intervals for the two strains (Liu and Yue 2000; Tian et al. 2011). A2345, A1345, A1245, or A1234 house fly lines with recessive morphological markers on autosomes 1, 2, 3 or 5, respectively, from aabys had significantly decreased levels of permethrin resistance compared with ALHF, implying that factors on autosomes 1, 2, 3 and 5 play important roles in pyrethroid resistance in ALHF (Tian et al. 2011).

3.3.2 RNA extraction

A total of 20 3-day old adult female house flies from each of three house fly strains (ALHF, aabys and CS) and five house fly lines (A2345, A1345, A1245, A1235, and A1234) were flash frozen on dry ice and immediately processed for RNA extraction. Total RNA was extracted using the hot acid phenol extraction method as outlined by Chomczynski and Sacchi (1987) (Chomczynski and Sacchi 1987). The RNA extraction from each strain or line was performed three times with different fly samples on different days to provide biological replications for the RNA-Seq experiments (ALHF) or, later, as the replications of qRT-PCR experiments for the validation of the up-regulated genes. A total of 30 μg of RNA was subsequently treated with DNase I using the DNA-Free kit from Ambion (Austin, TX) to remove any remaining DNA and then extracted over two successive acid phenol: chloroform (1:1) steps, followed by a final

chloroform extraction to remove any residual phenol. The RNA was then precipitated over ethanol at -80 °C overnight, pelleted, dried, and suspended in sterile distilled water, after which a 1µg aliquot of RNA was visually inspected for quality and for DNA contamination on a 1% agarose gel. Total RNA was sent for RNA-Seq analysis to Hudson Alpha Institute of Biotechnology (HAIB), in Huntsville, Alabama.

3.3.3 RNA library preparation and RNA-Seq

RNA quality for each sample of house fly strains and lines was assessed using a Qubit fluorimeter and an Agilent 2100 Bioanalyzer at HAIB. Libraries were then prepared using Illumina Tru-Seq RNA Sample Prep Kits for mRNA-Seq using a 3' poly A selection method. Samples for the mRNA-Seq were run using the PE-50 module (HAIB) on an Illumina HiSeq 2000 instrument to generate 50 nucleotide paired end libraries. Base calling and barcode parsing were also conducted at HAIB. Data were processed to remove any reads not passing the Chastity filter and then further trimmed for the adapter using Trimmomatic (Lohse et al. 2012). Two biological replications of RNA-Seq sequencing were conducted on independent samples of the resistant ALHF strain to validate the gene expression values. All sequence traces have been submitted to the National Center for Biotechnology Information (NCBI) Short Read Archive (SRA) as accessions [NCBI:SRR521286], [NCBI:SRR521288], [NCBI:SRR521289], and [NCBI:SRR521290], and are part of Bioproject #170716 with additional information in the NCBI Gene Expression Omnibus [NCBI:GSE39327].

3.3.4 Contig generation, gene annotation, gene expression determination, and data analysis

The bioinformatic analysis of the *M. domestica* transcriptomic data generated in this study was performed as illustrated in the flowchart in Figure 3.2. To generate the *M. domestica* reference transcriptome, the raw data for the two ALHF samples were pooled and then assembled *de novo* using Trinity version r2012-05-18 (Grabherr et al. 2011). The standard program settings were modified to increase the Java memory to 20 GB. The contigs obtained from the Trinity build were then compressed, to reduce redundancy, using CAP3 (Huang and Madan 1999) at a 95% similarity level. Compressed sequences that were <500 nt in length were discarded and the remaining contigs were further annotated to predict the gene coding region within each transcript using Augustus (Stanke et al. 2006). Within the Augustus program (Augustus version 2.5.5), the species model "FLY" based on *Drosophila melanogaster*'s genome was selected as the reference species due to the relatively close phylogenetic relationship of *D. melanogaster* and *M. domestica*. All predicted coding regions that were ≥300 nt in length were retained as the *M. domestica* ALHF strain reference transcriptome and were used for further gene expression comparisons with all of the *M. domestica* strains in our study. Within the ALHF reference transcriptome, 90% of the contigs with coding region were predicted to consist of full-length ORFs, and 10% expressed sequence tags (EST) that were missing either the 5' or the 3' ends of the predicted sequence. The functional annotations of the sequences within the ALHF reference transcriptome were then predicted using HMMScan in HMMER (v 3.0) at an e-value cut off of 10^{-20} against the Pfam-A (v26.0) hidden Markov model (HMM) database from the Wellcome Trust Sanger Institute, which is a manually-curated database of known protein domains that can be used to predict the function of an unknown protein by homology (Johnson et al. 2010; Punta et al. 2012). In addition, the ALHF reference sequences were annotated using blastx (Altschul et al.

1997) against the *D. melanogaster* proteome (v. r5.46) (McQuilton et al. 2012) at an e-value cut off of 10^{-20}, and enzyme functions were annotated by the Kyoto Encyclopedia of Genes and Genomes (KEGG) automatic annotation server (KAAS) at an e-value cut off of 10^{-5} (http://www.genome.jp/tools/kaas/; (Moriya et al. 2007)). All predicted sequences have been submitted to the TSA of NCBI as accession numbers [NCBI:KA644422] through [NCBI:KA650580].

The *de novo* ALHF reference transcriptome was then used as the common reference for the estimation of the gene expression values for each of the strains. The paired end reads within each strain were mapped as paired end mate pairs using RSEM (Li and Dewey 2011) to estimate the fragments per kilo base of reference gene length per million reads mapped (FPKM). The differential gene expression was then determined using EdgeR from Bioconductor at the α=0.05 (0.05%) false discovery rate (FDR) (Robinson et al. 2010; Robinson and Smyth 2007; Robinson and Smyth 2008). To examine sequence coverage and to verify that the sequences of the ALHF reference transcriptome were present within each of the strains tested, the reads from each strain were mapped against the ALHF transcriptome reference using TopHat, with the no-novel-juncs option, to estimate the percentage of gene coverage within each strain (Trapnell et al. 2012) and the resulting alignment files were converted to nucleotide sequences using Samtools pileup (v0.1.13) (Li et al. 2009) and Seqret in EMBOSS (Rice et al. 2000). The percentage of gene coverage for each ALHF predicted gene for each house fly strain tested was then determined using faCount (http://hgwdev.cse.ucsc.edu/~kent/src/unzipped/utils/faCount/).

3.3.5 Real-time quantitative RT-PCR validation of RNA-Seq data

A total of 70 genes that were differentially expressed among the different house fly strains/lines were chosen for the validation study using real-time quantitative PCR, with primers designed according to the RNA-Seq sequencing data (Table S3.1). Total RNA was extracted from samples of 20 3-day old post-eclosion female *M. domestica* as previously described. The total RNA (0.5 μg/per sample) from each house fly sample was reverse-transcribed using SuperScript II reverse transcriptase (Stratagene) in a total volume of 20 μL. The quantity of cDNAs was measured using a spectrophotometer prior to qRT-PCR, which was performed with the SYBR Green master mix Kit and ABI 7500 Real Time PCR system (Applied Biosystems). Each qRT-PCR reaction (15 μL final volumes) contained 1× SYBR Green master mix, 1 μL of cDNA, and a gene specific primer pair at a final concentration of 0.3–0.5 μM. A 'no-template' negative control and all samples were performed in triplicate. Relative expression levels for specific genes were calculated by the $2^{-\Delta\Delta Ct}$ method using SDS RQ software (Livak and Schmittgen 2001). The β-actin gene, an endogenous control, was used for internal normalization in the qRT-PCR assays (Aerts et al. 2004). Preliminary qRT-PCR experiments with the primer pair for the β-actin gene (Table S3.1) designed according to the sequences of the β-actin gene had revealed that the β-actin gene expression remained constant in the house fly strains, so the β-actin gene was used. Each experiment was performed three times with different preparations of RNA samples. The statistical significance of the gene expressions was calculated using a Welch's t-test for all pairwise sample comparisons against the ALHF strain at a value of α=0.05 (RDevelopment 2012).

3.3.6 Genetic linkage analysis of up-regulated genes

BC_1 lines of *ac/ac, +/ar, +/bwb, +/ye, +/snp* (A2345); *+/ac, ar/ar, +/bwb, +/ye, +/snp* (A1345); *+/ac, +/ar, bwb/bwb, +/ye, +/snp* (A1245); *+/ac, +/ar, +/bwb, ye/ye, +/snp* (A1235); and *+/ac, +/ar, +/bwb, +/ye, snp/snp* (A1234) were used to determine genetic linkage of up-regulated genes. Briefly, allele specific PCR was conducted using the cDNA from five BC_1 house fly lines (Liu and Yue 2001; Zhu et al. 2008; Zhu and Liu 2008) for genetic mapping of the genes (Liu and Scott 1996). Two rounds of PCR were conducted. For the first PCR reaction, the allele-independent primer pairs (Table S3.1) were designed for generating P450 (ALHF_04445.g2939 (*CYP6A36*), (Zhu et al. 2008)) and ALHF_04553.g3033, carboxylesterase (ALHF_03407.g2111), adenylate cyclase (ALHF_01050.g580), protein kinase (ALHF_10712.g5974), G-protein coupled receptor (ALHF_06811.g4468), and peptidase (ALHF_07511.g4836 and ALHF_05334.g3663) cDNA fragments, respectively. The first PCR solution with cDNA template and an allele-independent primer pair was heated to 95°C for 3 min, followed by 35 cycles of 95°C for 30 s, 60°C for 30 s, and 72°C for 1 min, then 72°C for 10 min. The PCR product from this reaction was then used as the template to determine autosomal linkage. The second PCR was employed with 0.5 µL of the first round PCR reaction solution and the allele specific primer pair (Table S3.1). The second PCR reaction was heated to 95°C for 3 min, followed by 35 cycles of 95°C for 30 s, 58°C for 30 s, and 72°C for 30 s, then 72°C for 10 min. One of each allele specific primer pair was designed based on the specific sequence of the genes from ALHF by placing a specific nucleotide polymorphism at the 3’ end of the primer to permit preferential amplification of the allele from ALHF. Each experiment was repeated three times with different mRNAs to ensure that the same autosomal linkage could be consistently repeated. The PCR products were sequenced at least once for each gene to confirm the consistency of the tested gene fragments.

3.4 Results

3.4.1 Illumina sequencing, transcriptome assembling and annotation of ALHF house flies

The number of paired end reads for each of the house fly strains ranged from 25-37 million, with an average of 7% of the reads discarded due to low quality (Table 3.1). The two RNA-Seq biological replicates from the ALHF strain ([NCBI:SRR521288] and [NCBI:SRR521289]) had sequence depths of 25 and 28 million reads, respectively. After the sequence cleaning steps, the two RNA-Seq sequences from ALHF were pooled, resulting in 53 million reads that could then be used to assemble the ALHF transcriptome reference. After the Trinity *de novo* transcriptome assembling (Grabherr et al. 2011) and CAP3 processing steps (Huang and Madan 1999), 14488 contigs were generated from the adult female ALHF *M. domestica de novo* transcriptome assembly (Table 3.2). The majority of the contigs ranged from 500 to 1500 nt in length (Figure 3.3). The N50, the central tendency of the contig length, was 1039 nt, indicating that half of the total number of nucleotides used for the entire transcriptome assembly were contained within contigs with ≥1039 nt in length (Miller et al. 2010). Within the 14488 contigs, a total of 6159 (43%) of them contained coding regions with >500 nt length, 5469 of which had complete putative open reading frames (ORFs). The nucleotide sequence information for house flies has been submitted to the NCBI Transcriptome Shotgun Assembly (TSA) (http://www.ncbi.nlm.nih.gov/genbank/tsa/). The complete annotation spreadsheet for the *M. domestica* predicted gene set is provided in Table S3.2, including the predicted gene name, nucleotide length, the TSA accession number from NCBI, the *D. melanogaster* blastx homology, the Superfamily general and detailed function annotations, the Pfam-A HMM homology, the

KEGG homology, and the putative GO terms based on the *M. domestica* predicted gene homology to the *D. melanogaster* proteome. The predicted gene set had an N50 of 2043 nt and the majority of the genes ranged in length from 1000 to 3000 nt (Figure 3.3). Among the 6159 annotated sequences, 5975 had significant hits to three different databases. A total of 5549 sequences had significant hits for the Pfam-A HMM library (v26.0), representing 2147 gene families (Johnson et al. 2010; Punta et al. 2012), while 5730 sequences had significant hits for *D. melanogaster* (v. r5.46) (McQuilton et al. 2012) (Figure 3.4). Since ~93% of the sequences had significant (e-value < 10^{-20}) matches to *D. melanogaster*, we used the bioinformatic information available for the *D. melanogaster* Structural Classification of Proteins (SCOP) functional annotation (http://supfam.cs.bris.ac.uk/SUPERFAMILY/) as a reference for the *M. domestica* transcriptome. To provide a general overview of the gene discovery in the adult *M. domestica* transcriptome, the predicted genes in *M. domestica* were thus classified according to their sequence homology to the functional categories of *D. melanogaster* in the SCOP database (Table 3.3) (http://supfam.cs.bris.ac.uk/SUPERFAMILY/). The SCOP functional category annotation sorts the genes from *D. melanogaster* into eight general function categories, which are then divided into detailed functional categories. A total of 1963 genes, which was approximately one-third of the 6159 genes, were placed into the non-annotated category of the SCOP general function category, and represented the largest SCOP general function category. The second most abundant general function category was the metabolism category, encompassing 17% of all predicted genes, followed by general function (15%), regulation (15%), intracellular processes (13%), information (4%), extracellular processes (2%), and “other” (1%). The Kyoto Encyclopedia of Genes and Genomes (KEGG) was also used for gene annotation in order to

identify those genes with putative enzymatic function. Taken together, the use of multiple databases for the functional annotation of the predicted genes in the *M. domestica* adult transcriptome allowed us to categorize the *M. domestica* genes into higher (general and detailed function) and lower (family and gene function) levels of annotation.

3.4.2 House fly transcriptome reference and gene expression profiles

The ALHF transcriptome was used as the reference for the comparison of gene expression between the resistant ALHF and the susceptible aabys and CS strains of *M. domestica*. To verify that the predicted genes within the ALHF transcriptome also provided good coverage for the other *M. domestica* strains tested, we independently mapped the raw Illumina reads from each of the *M. domestica* strains to the ALHF reference transcriptome using Tophat and then determined the percentage of gene coverage for each of the genes within each *M. domestica* strain. The results showed that the median nucleotide coverage for the 6159 genes within the ALHF transcriptome was >99% for all of the *M. domestica* strains tested (Figure 3.5), demonstrating that the transcriptome from ALHF was indeed a suitable reference for the determination of gene expression levels for all the strains.

The program RSEM (Li and Dewey 2011) was therefore used to estimate the gene expression values (FPKM) for all the *M. domestica* strains using the ALHF transcriptome as the reference. The gene expression values from the susceptible aabys and CS strains were then compared with the gene expression values of the ALHF strain to determine differential gene expression using EdgeR (Robinson and Smyth 2007) with a 0.05 false discovery rate (α=0.05, (Robinson and Smyth 2007; Robinson and Smyth 2008)). In addition to testing the two

susceptible strains of *M. domestica* for differential gene expression, we further tested the gene expression values of the ALHF strain against an additional biologically-independent sample of the ALHF strain to ensure that the gene expression values were reproducible. When the gene expression values of the two ALHF samples were compared, the results showed a strong 1:1 correlation (r^2=0.95); the correlation coefficients for the aabys and CS strains were 0.62 and 0.49, respectively (Figure 3.6). In addition, <10% (606) of the genes tested as differentially expressed between the two ALHF replicates, while the aabys and CS strains had 3428 and 4792 genes that were differentially expressed, respectively (Figure 3.6). Since gene over-expression has been linked to insecticide resistance (Cao et al. 2008; Djouaka et al. 2008; Liu and Scott 1995; Ortelli et al. 2003; Vontas et al. 2002), the genes identified as differentially up-regulated in the ALHF strain when compared to both the insecticide-susceptible aabys and CS strains represent the genes putatively involved in pyrethroid resistance.

Overall, a total of 1316 genes were identified as being co-up-regulated in ALHF in comparison to both aabys and CS (Table S3.3). While one-third of these genes (452 genes) were distributed within the SCOP general function category of “no annotation” (Table S3.3), the majority (777 genes) fell within the SCOP categories of metabolism, general, intra-cellular processes, and regulation, containing 275, 178, 174, and 150 genes, respectively (Table S3.3). The metabolism SCOP general function category had the greatest number of detailed function groups (13 groups), among which the redox detailed function group contained the second largest number of genes and detoxification enzymes such as cytochrome P450s, glutathione-S-transferases (GSTs) and esterases (Table S3.3). Interestingly, within the regulation category the detailed function groups with the greatest number of genes were involved with signal

transduction, kinases/phosphatases, and DNA binding, while in the intra-cellular processes category the proteases were the most abundant detailed function group. A total of 1440 genes were identified as down-regulated in the ALHF strain when compared to the susceptible aabys and CS strains (Table S3.3), among which one-third (458) of the genes had no SCOP annotation. The rest of the down-regulated genes were distributed within the SCOP general function category of regulation (254 genes), general (242 genes), metabolism (194 genes), intra-cellular processes (152 genes), information (105 genes), other (22 genes), and extra-cellular processes (13 genes).

3.4.3 Validation of the expression of up-regulated genes in house fly strains/lines

A total of 70 genes were selected from the predominant groups of up-regulated genes identified by RNA-Seq, including multiple cytochrome P450s, GSTs, and esterases in metabolism; kinases/phosphatases, 7 transmembrane receptors (rhodopsin-like G-protein coupled receptor (GPCR) family), adenylate and guanylate cyclases in regulation; and serpins and carboxypeptidases in intracellular processes for further validation by qRT-PCR. We examined the expression of these 70 genes in resistant ALHF, susceptible aabys and five house fly homozygous lines A2345, A1345, A1245, A1235, and A1234, the lines represent the ALHF strain where autosomes 1, 2, 3, 4, and 5, respectively, have been replaced by the autosome from aabys. Overall, the biological replication of qRT-PCR results showed that the expression of the majority of genes (81%) was consistent with the RNA-Seq data, being highly expressed in resistant ALHF compared with the susceptible aabys (Table 3.4). We also examined the expression levels of the up-regulated genes in ALHF for the five house fly lines to determine the effects of factors from the different autosomes of ALHF on the up-regulation of the genes. Clear changes in the gene

expression levels were identified when each autosome in ALHF was replaced by the corresponding aabys autosome, i.e., lines A2345, A1345, A1245, A1235, and A1234 (Figure 3.7, Table 3.4). In general, no significant change in the level of expression was observed for most of the selected genes, when autosome 4 of ALHF (i.e., line A1235, Figure 3.7, Table 3.4) was replaced with that from aabys except for four protease genes. Previous research by Tian et al., (Tian et al. 2011) identified that when autosome 4 in ALHF was replaced by the one from aabys, there was no significant decrease in resistance. Liu and Scott (1995) also demonstrated that replacement of autosome 4 in the resistant LPR house flies with the one from aabys, the resistance level was not changed (Liu and Scott 1995). Liu and Yue (2001) reported the similar results in house flies (Liu and Yue 2001). Taken together, these results strongly revealed that factors/genes on autosome 4 do not have a major role in the up-regulation of genes in ALHF, although further investigation of the up-regulated protease genes in resistance is needed. The majority of the selected up-regulated genes exhibited no change in expression when autosome 1 or 3 in ALHF was replaced by the corresponding autosome from aabys (i.e., lines A2345 or A1245, respectively). However, significant changes in the gene expression for most of the selected up-regulated genes (>90%) were observed when autosome 2 or 5 in ALHF was replaced by the corresponding autosome from aabys (i.e., lines A1345 or A1234 respectively). These results suggest the importance of factors on autosome 2 and/or 5 for the expression of up-regulated genes in ALHF and/or that several of the up-regulated loci reside on the replaced chromosomes.

3.4.4 Autosome co-regulation in up-regulation gene expression in resistant house flies

We next examined the autosomal linkage of factors from different autosomes on the 70 up-regulated genes that have been validated by qRT-PCR to determine the effects of the co-regulation on the expression of the up-regulated genes among five house fly lines of A2345, A1345, A1245, A1235 and A1234. Analyzing the gene expression changes resulting from autosome replacement in ALHF enabled us to evaluate the role of genes or factors on each autosome plays in gene overexpression in ALHF. We conducted Venn diagram analyses on the autosome interaction for the expression of genes in each of the SCOP general function categories of metabolism, regulation, and intracellular processes (Figure 3.8). The results revealed that apart from the 11 genes up-regulated solely by factor(s) on a single autosome (four in autosome 2, six in autosome 5 and one in autosome 3), the expression of the rest of the up-regulated genes were all linked to factors on more than one autosome (Figure 3.8). This result suggests that factors on different autosomes are capable of co-regulation of some genes. This was most commonly observed for autosomes 2 and 5. Almost one-third of the tested genes (n=21 genes) were up-regulated by co-regulation of factors on autosome 2 and 5 only, including cytochrome P450s, GSTs, and esterases in metabolism; kinases/phosphatases, 7 transmembrane receptors (rhodopsin-like GPCR family), adenylate and guanylate cyclases in regulation; and serpin and carboxypeptidases in intracellular processes (Figure 3.8, Table 3.4). Nine genes were co-up-regulated by factors on autosomes 1, 2, and 5, with the functions of these genes being linked to metabolism and regulation categories, suggesting that factors on autosome 1, besides those on autosomes 2 and 5, were also involved in the regulation of some of the gene expression in metabolism and regulation. Six genes were up-regulated by the interaction of factors on autosomes 2, 3, and 5 and these genes were mainly located in the regulation and protease

categories and none of the metabolic genes were involved in the interactions by factors among autosomes 2, 3, and 5, suggesting that, besides the factors on autosomes 2 and 5, factors on autosome 3 might have a role in the functions of regulation and proteolysis. A few genes were co-regulated by factors on autosomes 1 and 2 (including one P450 gene, one carboxylesterase gene, and one GPCR gene), or autosomes 1, 2, 3, and 5 (including one P450 gene and two protein kinase genes) (Table 3.4). No gene interactions between 1 and 3; 1 and 5; 2 and 3; 3 and 5; 1, 2 and 3; 1, 3 and 5; 2. 3, and 4 or 2, 3, 4 and 5 were observed. None of the genes were found to be up-regulated solely by factors on autosome 1 or 4 alone.

To better understand the *cis/trans* regulation of the up-regulated genes in resistant house flies, autosomal location analyses were conducted for eight up-regulated genes scattered among all three important functional categories. An allele specific PCR (AS-PCR) determination was performed to examine the autosomal location of the genes with five house fly lines. The ALHF allele specific primer pair was designed based on the specific sequence of the genes from ALHF by placing a specific nucleotide polymorphism at the 3' end of each primer to permit preferential amplification of specific alleles from ALHF. Our results showed that the ALHF allele-specific primer sets for P450 genes of ALHF_04445.g2939 (*CYP6A36*) and ALHF_04553.g3033 and protein kinase gene ALHF_10712.g5974 amplified specific DNA fragments only in flies having the autosome 5 wild-type marker from ALHF (Figure 3.9), which demonstrated that these three genes were located on autosome 5. Whereas, carboxylesterase gene ALHF_03407.g2111, adenylate cyclase gene ALHF_01050.g580, G-protein coupled receptor gene ALHF_06811.g4468, and peptidase genes ALHF_07511.g4836 and ALHF_05334.g3663 were located on autosome 2. These results were consistent with our autosomal linkage map (Figure

3.7, Table 3.4).

3.5 Discussion

The central hypothesis guiding this research is that normal biological and physiological pathways and gene expression signatures are varied in resistant insects through changes in multiple gene expression, thus enabling resistant house flies to adapt to environmental or insecticidal stress, and that these changes are controlled by a regulatory network and perhaps by signaling transduction. This hypothesis is grounded in evidence from the considerable body of research that has been done in this field. Results from previous studies by ourselves (Liu et al. 2011; Liu et al. 2007; Xu et al. 2005; Xu et al. 2006a; Yang and Liu 2011), and others (David et al. 2010; David et al. 2005; Gregory et al. 2011; Kalajdzic et al. 2012; Karatolos et al. 2012; Liu and Scott 1996; Marcombe et al. 2009; Martinez-Torres et al. 1999; MÜLler et al. 2008; Plapp Jr et al. 1987; Pridgeon and Liu 2003; Strode et al. 2006; Strode et al. 2008; Tao et al. 2012; Vontas et al. 2005) all indicate that the interaction of multiple genes and complex mechanisms are responsible for the development of insecticide resistance in insects. Indeed, many studies have demonstrated different patterns of gene expression between resistant and susceptible insect populations and the up-regulation of P450 and GST genes in resistant insects. Many studies have also found that the overexpression of resistant metabolic genes is regulated by *trans* and/or *cis* factors in insecticide-resistant insects (Brown et al. 2005; Carino et al. 1994; Gao and Scott 2006; Grant and Hammock 1992; Liu and Scott 1997; Liu and Scott 1998; Maitra et al. 2000b; Maitra et al. 1996; Misra et al. 2011; Wilding et al. 2012). The up-regulation of a GST gene (GST-2) in the mosquito *Aedes aegypti* is controlled by a *trans*-acting factor (Grant and Hammock 1992),

while the up-regulation of two P450 genes, *CYP6A1* and *CYP6D1*, in the house fly *M. domestica* are known to be *trans*-regulated by one or more factors on autosome 2 (Carino et al. 1994; Gao and Scott 2006; Liu and Scott 1997; Liu and Scott 1996). The up-regulation of *CYP6A2* and *CYP6A8* in the fruit fly *D. melanogaster* is transcriptionally regulated by *trans*-regulatory factors (Maitra et al. 2000a; Maitra et al. 1996). The up-regulation of *CYP6G1* and *CYP6D1* is controlled by *cis/trans* regulatory factors (Gao and Scott 2006; Liu and Scott 1997; Liu and Scott 1998).

Taken together, these findings suggest that not only is insecticide resistance conferred via multi-resistance mechanisms or genes, but it is mediated through the interaction of regulatory genes and resistance genes such as P450s, esterases and GSTs. However, a global understanding of the complex processes resulting from gene interaction and regulation remains elusive. None of the regulatory factors responsible for insecticide resistance have yet been identified, and no regulation pathways have been examined. Nevertheless, these gaps will soon be filled following the availability of whole transcriptome analyses, which have begun to provide new ways of assessing how insects respond to the environment and insecticides (Schuler 2012).

To define the key genes and their *trans/cis*- or co-regulation involved in insecticide resistance, and thus gain fresh insights into the overall picture of how molecular mechanisms in resistant house flies function, we began by assembling and annotating the adult house fly transcriptome, providing the first reference transcriptome for adult house flies. Using the house fly transcriptome as a reference, our RNA-Seq of the resistant ALHF strain revealed a set of 1316 genes that were up-regulated relative to the susceptible aabys and CS strains, and a total of 1440 genes that were down-regulated. These results may not only reveal equally dynamic changes in abundance for both the increases and decreases in the total gene expression for different

categories in resistant house flies, but also indicate an important feature of resistance gene regulation by both activators (the up-regulated genes) and perhaps, repressors (those down regulated genes). Several hypotheses have been proposed for the harmonizing of up- and down-regulation, e.g., homeostatic responses for protecting the cell from the harmful effects of oxidizing species from metabolic enzymes (Morgan 2001; White and Coon 1980); homeostatic responses to provocative processes (Morgan 1989); and/or an essential for the tissue to utilize its transcriptional machinery and energy for the synthesis of other components involved in the inflammatory response (Morgan 1989). Whether the down-regulated genes identified in the resistant house flies by our study reflects a regulation feature or homeostatic response of mosquitoes to insecticides needs to be further studied.

Deciphering the up-regulated genes among the SCOP general categories into detailed functions uncovered three key SCOP detailed function categories, namely the redox detailed function category in metabolism, signal transduction and kinases/phosphatases in regulation, and proteases in intra-cellular processes. The redox detailed function group contained a number of superfamilies that have been linked to detoxification, including multiple cytochrome P450s, glutathione S-transferases, and esterases. The signal transduction and kinases/phosphatases detailed function groups were found to contain several gene families with signal transduction and regulation functions, including 7 transmembrane receptors (rhodopsin-like GPCR family), adenylate and guanylate cyclases, protein kinases and phosphatases. The proteases detailed function group contained genes with digestive, catalytic, and proteinase activities. Since co-regulation provides valuable insights into altered categories/pathways, thereby aiding functional interpretation (Blalock et al. 2004), this finding suggests that these co-up-regulated functional

groups of genes may share co-regulation features. Among the three key detailed function categories, the roles of the detoxification superfamilies of P450s, GSTs and esterase in insecticide resistance have been extensively studied and up-regulation of their expression has been demonstrated to be associated with enhanced metabolic detoxification of insecticides, resulting in the development of insecticide resistance in insects (Carino et al. 1994; Feyereisen 2006; Hemingway et al. 2002; Hemingway and Karunaratne 1998; Itokawa et al. 2010; Liu et al. 2011; Liu and Scott 1998; Oakeshott et al. 2005; Ortelli et al. 2003; Pavek and Dvorak 2008; Ranson and Hemingway 2005; Scott 1999; Small and Hemingway 2000; Vontas et al. 2005; Yang and Liu 2011; Zhu et al. 2008; Zhu and Liu 2008; Zhu et al. 2010). In contrast to the well-known role played by the detoxification system in insecticide resistance, the functions of genes in two other key detailed function categories, the signaling transduction system and proteases/serine proteases, such as GPCRs, protein kinase/phosphatases and proteases, in insecticide resistance are less well understood, although a few studies have reported the up-regulation of protease genes in insecticide resistant insects (Ahmed et al. 1998; Pedra et al. 2004; Reid et al. 2012; Vontas et al. 2005; Wilkins et al. 1999; Wu et al. 2004). Nevertheless, the genes in these two key categories are well known as key intracellular signaling regulators and share common functions in the signaling pathway, playing an important role in transmitting information from extracellular polypeptide signals to target gene promoters in the nucleus and in the regulation of gene expression, activation/termination intracellular signaling transduction, and regulating numerous diverse cellular and biological/physiological processes (Betke et al. 2012; Cottrell et al. 2012; de la Nuez Veulens and Rodríguez 2009; Goupil et al. 2012; Krishna and Narang 2008; Lagerström and Schiöth 2008; Lawan et al. 2012; Lemberg 2011; Marrs et al. 2010; Oldham and Hamm

2008; Pedra et al. 2004; Ramsay et al. 2008; Singh and Aballay 2012; Spehr and Munger 2009; Trejo 2003; Verrier et al. 2011; Yang et al. 2013).

To test the co-regulation of the up-regulated genes in these three key categories in resistant house flies, a novel approach was applied in the study, in which the gene expression profile in the house fly genetic lines was characterized in terms of different autosome combinations from the resistant ALHF strain, thus illustrating the co-regulation of autosomes in the expression of individual genes. This research approach not only provides a catalog of genes and information about their potential functions in insecticide resistance (Donnell and Strand 2006), but also serves as a stepping stone towards filling important gaps in our knowledge of transcriptional interaction and the regulation networks that are involved in insecticide resistance. Our gene co-regulation analysis revealed that the up-regulated gene expression in resistant ALHF house flies occurred primarily as a result of the co-regulation of factors between autosomes 2 and 5, although a few genes had their expression regulated by factors among autosomes 1, 2, and 5, or among autosomes 2, 3, and 5. These findings strongly suggest that multiple factor/autosome co-regulation, especially those related to autosomes 2 and 5, are key determinants for individual gene expression in resistant house flies. Among the up-regulated genes, cytochrome P450s, GSTs, and esterases in metabolism; kinases/phosphatases, 7 transmembrane receptors (rhodopsin-like GPCR family), adenylate and guanylate cyclases in regulation; and serpin and carboxypeptidases in intracellular processes as major groups of genes were up-regulated by the interactions of factors on autosomes 2 and 5 (Table 3.4). Our genetic mapping study further located two P450 genes and a protein kinase gene on autosome 5, and mapped a carboxylesterase gene, an adenylate cyclase gene, a G-protein coupled receptor gene and two peptidase genes on

autosome 2. With the exception of one P450 gene, whose up-regulation was controlled by *cis* factor(s) on the same autosome on which the gene was located, all the genes tested in the genetic mapping study showed their expression being controlled by *cis* and *trans* factors. i.e., factors not only on the autosomes on which the genes were located but also other autosomes as well. Taken together, our findings suggested that that not only is insecticide resistance conferred via multi-resistance mechanisms or up-regulated genes, but it is mediated through the *trans* and/or *cis* co-regulations of resistance genes. Whether the signaling transduction cascades controlled by GPCRs, protein kinase/phosphatases and proteases are indeed involved in the regulation of resistance P450 genes and of resistance development remains an urgent topic for investigation.

3.6 Acknowledgements

The authors are grateful to the Hudson Alpha Institute of Biotechnology for their expertise in conducting the RNA sequencing work and for all of their help and support with this study. The project described was supported by Award AAES Hatch/Multistate Grants ALA08-045 and ALA015-1-10026 to N.L.

3.7 Reference

Aerts, JL, Gonzales, MI and Topalian, SL (2004) Selection of appropriate control genes to assess expression of tumor antigens using real-time RT-PCR. Biotechniques 36: 84-97.

Ahmed, S, Wilkins, RM and Mantle, D (1998) Comparison of proteolytic enzyme activities in adults of insecticide resistant and susceptible strains of the housefly *M. domestica*. Insect Biochemistry and Molecular Biology 28: 629-639.

Altschul, SF, Madden, TL, Schäffer, AA, Zhang, J, Zhang, Z, Miller, W and Lipman, DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Researchearch 25: 3389-3402.

Betke, KM, Wells, CA and Hamm, HE (2012) GPCR mediated regulation of synaptic transmission. Progress in Neurobiology 96: 304-321.

Blalock, EM, Geddes, JW, Chen, KC, Porter, NM, Markesbery, WR and Landfield, PW (2004) Incipient Alzheimer's disease: microarray correlation analyses reveal major transcriptional and tumor suppressor responses. Proceedings of the National Academy of Sciences 101: 2173-2178.

Brattsten, LB, Holyoke, CW, Leeper, JR and Raffa, KF (1986) Insecticide resistance: challenge to pest management and basic research. Science (New York, N.Y.) 231: 1255-1260.

Brown, RP, McDonnell, CM, Berenbaum, MR and Schuler, MA (2005) Regulation of an insect cytochrome P450 monooxygenase gene (CYP6B1) by aryl hydrocarbon and xanthotoxin response cascades. Gene 358: 39-52.

Cao, C-W, Zhang, J, Gao, X-W, Liang, P and Guo, H-L (2008) Overexpression of carboxylesterase gene associated with organophosphorous insecticide resistance in cotton aphids, *Aphis gossypii* (Glover). Pesticide Biochemistry and Physiology 90: 175-180.

Carino, FA, Koener, JF, Plapp, FW, Jr. and Feyereisen, R (1994) Constitutive overexpression of the cytochrome P450 gene CYP6A1 in a house fly strain with metabolic resistance to insecticides. Insect Biochemistry and Molecular Biology 24: 411-418.

Chomczynski, P and Sacchi, N (1987) Single-step RNA isolation from cultured cells or tissues. Anal Biochem 162: 156-259.

Cottrell, GA, Lambert, JJ and Peters, JA (2012) Modulation of GABAA receptor activity by alphaxalone. British journal of pharmacology 90: 491-500.

David, J-P, Coissac, E, Melodelima, C, Poupardin, R, Riaz, MA, Chandor-Proust, A and Reynaud, S (2010) Transcriptome response to pollutants and insecticides in the dengue vector *Aedes aegypti* using next-generation sequencing technology. BMC Genomics 11: 216.

David, J-P, Strode, C, Vontas, J, Nikou, D, Vaughan, A, Pignatelli, PM, Louis, C, Hemingway, J and Ranson, H (2005) The *Anopheles gambiae* detoxification chip: a highly specific microarray to study metabolic-based insecticide resistance in malaria vectors. Proceedings of the National Academy of Sciences of the United states of America 102: 4080-4084.

de la Nuez Veulens, A and Rodríguez, R (2009) Protein-coupled receptors as targets for drug design. Biotecnol Apl 26: 24-33.

Djouaka, RF, Bakare, AA, Coulibaly, ON, Akogbeto, MC, Ranson, H, Hemingway, J and Strode, C (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: 538.

Donnell, DM and Strand, MR (2006) Caste-based differences in gene expression in the polyembryonic wasp *Copidosoma floridanum*. Insect Biochemistry and Molecular Biology 36: 141-153.

Feyereisen, R (2006) Evolution of insect P450. Biochemical Society Transactions 34: 1252-1255.

Gao, J and Scott, J (2006) Role of the transcriptional repressor mdGfi-1 in CYP6D1v1-

mediated insecticide resistance in the house fly, *Musca domestica*. Insect Biochemistry and Molecular Biology 36: 387-395.

Georghiou, GP (1971) Isolation, characterization and re-synthesis of insecticide resistance factors in the housefly, *Musca domestica*. In: Book Isolation, characterization and re-synthesis of insecticide resistance factors in the housefly, *Musca domestica* (Editor, eds.). Vol. 2, pp. 77. City.

Goupil, E, A Laporte, S and E Hebert, T (2012) Functional Selectivity in GPCR Signaling: Understanding the Full Spectrum of Receptor Conformations. Mini Reviews in Medicinal Chemistry 12: 817-830.

Grabherr, MG, Haas, BJ, Yassour, M, Levin, JZ, Thompson, DA, Amit, I, Adiconis, X, Fan, L, Raychowdhury, R and Zeng, Q (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nature biotechnology 29: 644-652.

Grant, D and Hammock, B (1992) Genetic and molecular evidence for a trans-acting regulatory locus controlling glutathione S-transferase-2 expression in *Aedes aegypti*. Molecular and General Genetics MGG 234: 169-176.

Greenberg, B (1965) Flies and Disease. Scientific American 213: 92-99.

Gregory, R, Darby, AC, Irving, H, Coulibaly, MB, Hughes, M, Koekemoer, LL, Coetzee, M, Ranson, H, Hemingway, J and Hall, N (2011) A de novo expression profiling of *Anopheles funestus*, malaria vector in Africa, using 454 pyrosequencing. PloS One 6: e17418.

Hemingway, J, Field, L and Vontas, J (2002) An Overview of Insecticide Resistance. Science 298: 96-97.

Hemingway, J and Karunaratne, SH (1998) Mosquito carboxylesterases: a review of the molecular biology and biochemistry of a major insecticide resistance mechanism. Medical and

veterinary entomology 12: 1-12.

Hiroyoshi, T (1964) Sex-limited inheritance and abnormal sex ratio in strains of the houseflye. Genetics 50: 373-385.

Huang, X and Madan, A (1999) CAP3: A DNA sequence assembly program. Genome Researchearch 9: 868-877.

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

Johnson, LS, Eddy, SR and Portugaly, E (2010) Hidden Markov model speed heuristic and iterative HMM search procedure. BMC Bioinformatics 11: 431.

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

Karatolos, N, Williamson, MS, Denholm, I, Gorman, K, Ffrench-Constant, RH and Bass, C (2012) Over-expression of a cytochrome P450 is associated with resistance to pyriproxyfen in the greenhouse whitefly Trialeurodes vaporariorum. PloS One 7: e31077.

Keiding, J (1986) House-fly; biology and control. OMS.

Krishna, M and Narang, H (2008) The complexity of mitogen-activated protein kinases (MAPKs) made simple. Cellular and molecular life sciences 65: 3525-3544.

Lagerström, MC and Schiöth, HB (2008) Structural diversity of G protein-coupled receptors and significance for drug discovery. Nature reviews Drug discovery 7: 339-357.

Lawan, A, Shi, H, Gatzke, F and Bennett, AM (2012) Diversity and specificity of the

mitogen-activated protein kinase phosphatase-1 functions. Cellular and Molecular Life Sciences: 1-15.

Lemberg, MK (2011) Intramembrane proteolysis in regulated protein trafficking. Traffic 12: 1109-1118.

Li, B and Dewey, CN (2011) RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics 12: 323.

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

Liu, H, Xu, Q, Zhang, L and Liu, N (2005) Chlorpyrifos Resistance in Mosquito *Culex quinquefasciatus*. Journal of Medical Entomology 42: 815-820.

Liu, N, Li, T, Reid, WR, Yang, T and Zhang, L (2011) Multiple Cytochrome P450 genes: their constitutive overexpression and permethrin induction in insecticide resistant mosquitoes, *Culex quinquefasciatus*. PloS One 6: e23403.

Liu, N, Liu, H, Zhu, F and Zhang, L (2007) Differential expression of genes in pyrethroid resistant and susceptible mosquitoes, *Culex quinquefasciatus* (S.). Gene 394: 61-68.

Liu, N and Scott, J (1997) Phenobarbital induction of CYP6D1 is due to atransacting factor on autosome 2 in house flies, *Musca domestica*. Insect Molecular Biology 6: 77-81.

Liu, N and Scott, JG (1995) Genetics of resistance to pyrethroid insecticides in the house fly, *Musca domestica*. Pesticide Biochemistry and Physiology 52: 116-124.

Liu, N and Scott, JG (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*. Biochemical Genetics 34: 133-148.

Liu, N and Scott, JG (1998) Increased transcription of CYP6D1 causes cytochrome P450-mediated insecticide resistance in house fly. Insect Biochemistry and Molecular Biology 28: 531-535.

Liu, N and Yue, X (2000) Insecticide resistance and cross-resistance in the house fly (Diptera: Muscidae). Journal of Economic Entomology 93: 1269-1275.

Liu, N and Yue, X (2001) Genetics of pyrethroid resistance in a strain (ALHF) of house flies (Diptera: Muscidae). Pesticide Biochemistry and Physiology 70: 151-158.

Livak, KJ and Schmittgen, TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25: 402-408.

Lohse, M, Bolger, AM, Nagel, A, Fernie, AR, Lunn, JE, Stitt, M and Usadel, B (2012) RobiNA: a user-friendly, integrated software solution for RNA-Seq-based transcriptomics. Nucleic Acids Researchearch 40: W622-W627.

Maitra, S, Dombrowski, S, Basu, M, Raustol, O, Waters, L and Ganguly, R (2000a) Factors on the third chromosome affect the level of Cyp6a2 and Cyp6a8 expression in *Drosophila melanogaster*. Gene 248: 147-156.

Maitra, S, Dombrowski, SM, Waters, LC and Ganguly, R (1996) Three second chromosome-linked clustered Cyp6 genes show differential constitutive and barbital-induced expression in DDT-resistant and susceptible strains of *Drosophila melanogaster*. Gene 180: 165-171.

Marcombe, S, Poupardin, R, Darriet, F, Reynaud, S, Bonnet, J, Strode, C, Brengues, C, Yébakima, A, Ranson, H and Corbel, V (2009) Exploring the molecular basis of insecticide resistance in the dengue vector *Aedes aegypti*: a case study in Martinique Island (French West

Indies). BMC Genomics 10: 494.

Marrs, WR, Blankman, JL, Horne, EA, Thomazeau, A, Lin, YH, Coy, J, Bodor, AL, Muccioli, GG, Hu, SS-J and Woodruff, G (2010) The serine hydrolase ABHD6 controls the accumulation and efficacy of 2-AG at cannabinoid receptors. Nature neuroscience 13: 951-957.

Martinez-Torres, D, Chevillon, C, Brun-Barale, A, Bergé, JB, Pasteur, N and Pauron, D (1999) Voltage-dependent Na+ channels in pyrethroid-resistant Culex pipiens L mosquitoes. Pesticide Science 55: 1012-1020.

McQuilton, P, Pierre, SES and Thurmond, J (2012) FlyBase 101–the basics of navigating FlyBase. Nucleic Acids Researchearch 40: D706-D714.

Milani, R, Rubini, P and Franco, M (1967) Sex determination in the housefly. Genetica Agraria 21: 385-411.

Miller, JR, Koren, S and Sutton, G (2010) Assembly algorithms for next-generation sequencing data. Genomics 95: 315-327.

Misra, JR, Horner, MA, Lam, G and Thummel, CS (2011) Transcriptional regulation of xenobiotic detoxification in *Drosophila*. Genes & Development 25: 1796-1806.

Morgan, E (1989) Suppression of constitutive cytochrome P-450 gene expression in livers of rats undergoing an acute phase response to endotoxin. Molecular Pharmacology 36: 699-707.

Morgan, ET (2001) Regulation of cytochrome p450 by inflammatory mediators: why and how? Drug Metabolism and Disposition 29: 207-212.

Moriya, Y, Itoh, M, Okuda, S, Yoshizawa, AC and Kanehisa, M (2007) KAAS: an automatic genome annotation and pathway reconstruction server. Nucleic Acids Researchearch 35: W182-W185.

MÜLler, PIE, ChouaÏBou, M, Pignatelli, P, Etang, J, Walker, ED, Donnelly, MJ, Simard, F and Ranson, H (2008) Pyrethroid tolerance is associated with elevated expression of antioxidants and agricultural practice in *Anopheles arabiensis* sampled from an area of cotton fields in Northern Cameroon. Molecular Ecology 17: 1145-1155.

Nickel, C and Wagoner, D (1974) Mutants on Linkage Groups 3 and 4 of the House Fly. Annals of the Entomological Society of America 67: 775-776.

Oakeshott, J, Claudianos, C, Newcomb, R and Russell, R (2005) Biochemical genetics and genomics of insect esterases. 309-381.

Oldham, WM and Hamm, HE (2008) Heterotrimeric G protein activation by G-protein-coupled receptors. Nature Reviews Molecular Cell Biology 9: 60-71.

Ortelli, F, Rossiter, LC, Vontas, J, Ranson, H and Hemingway, J (2003) Heterologous expression of four glutathione transferase genes genetically linked to a major insecticide-resistance locus from the malaria vector *Anopheles gambiae*. Biochemical Journal 373: 957-63.

Pap, L and Tóth, A (1995) Development and characteristics of resistance in the susceptible who/srs house fly (*Musca domestica*) strain subjected to selection with beta-cypermethrin. Pesticide Science 45: 335-349.

Pavek, P and Dvorak, Z (2008) Xenobiotic-induced transcriptional regulation of xenobiotic metabolizing enzymes of the cytochrome P450 superfamily in human extrahepatic tissues. Current Drug Metabolism 9: 129-143.

Pedra, JHF, McIntyre, LM, Scharf, ME and Pittendrigh, BR (2004) Genome-wide transcription profile of field- and laboratory-selected dichlorodiphenyltrichloroethane (DDT)-resistant *Drosophila*. Proceedings of the National Academy of Sciences of the United States of

America 101: 7034-7039.

Plapp Jr, F, McWhorter, G and Vance, W (1987) Monitoring for pyrethroid resistance in the tobacco budworm in Texas: 1986. In: Book Monitoring for pyrethroid resistance in the tobacco budworm in Texas: 1986 (Editor, eds.). Vol., pp., City.

Pridgeon, JW and Liu, N (2003) Overexpression of the cytochrome c oxidase subunit I gene associated with a pyrethroid resistant strain of German cockroaches, *Blattella germanica* (L.). Insect Biochemistry and Molecular Biology 33: 1043-1048.

Punta, M, Coggill, PC, Eberhardt, RY, Mistry, J, Tate, J, Boursnell, C, Pang, N, Forslund, K, Ceric, G and Clements, J (2012) The Pfam protein families database. Nucleic Acids Researchearch 40: D290-D301.

Ramsay, AJ, Dong, Y, Hunt, ML, Linn, M, Samaratunga, H, Clements, JA and Hooper, JD (2008) Kallikrein-related peptidase 4 (KLK4) initiates intracellular signaling via protease-activated receptors (PARs). Journal of Biological Chemistry 283: 12293-12304.

Ranson, H and Hemingway, J (2005) Mosquito glutathione transferases. Methods in enzymology 401: 226-241.

Raymond, M, Beyssat-Arnaouty, V, Sivasubramanian, N, Mouches, C, Georghiou, GP and Pasteur, N (1989) Amplification of various esterase B's responsible for organophosphate resistance in *Culex* mosquitoes. Biochemical Genetics 27: 417-423.

RDevelopment, C (2012) TEAM (2008): R: A language and environment for statistical computing. Vienna, Austria. Internet: http://www. R-project. org.

Rebeiz, M, Jikomes, N, Kassner, VA and Carroll, SB (2011) Evolutionary origin of a novel gene expression pattern through co-option of the latent activities of existing regulatory sequences.

Proceedings of the National Academy of Sciences 108: 10036-10043.

Reid, WR, Zhang, L, Liu, F and Liu, N (2012) The Transcriptome Profile of the Mosquito *Culex quinquefasciatus* following Permethrin Selection. PloS One 7: e47163.

Rice, P, Longden, I and Bleasby, A (2000) EMBOSS: the European molecular biology open software suite. Trends in genetics 16: 276-277.

Robinson, MD, McCarthy, DJ and Smyth, GK (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26: 139-140.

Robinson, MD and Smyth, GK (2007) Moderated statistical tests for assessing differences in tag abundance. Bioinformatics 23: 2881-2887.

Robinson, MD and Smyth, GK (2008) Small-sample estimation of negative binomial dispersion, with applications to SAGE data. Biostatistics 9: 321-332.

Roush, RT, Hoy, CW, Ferro, DN and Tingey, WM (1990) Insecticide Resistance in the Colorado Potato Beetle (Coleoptera: Chrysomelidae): Influence of Crop Rotation and Insecticide Use. Journal of Economic Entomology 83: 315-319.

Sawicki, R and Denholm, I (1984) Adaptation of insects to insecticides. Origins and Development of Adaptation: 1-152.

Schuler, MA (2012) Insect P450s: mounted for battle in their war against toxins. Molecular Ecology 21: 4157-4159.

Scott, HG and Littig, KS (1962) Flies of public health importance and their control. US Department of Health, Education, and Welfare, Public Health Service, Communicable Disease Center.

Scott, JG (1999) Cytochromes P450 and insecticide resistance. Insect Biochemistry and

Molecular Biology 29: 757-777.

Scott, JG, Liu, N, Kristensen, M and Clark, AG (2009) A case for sequencing the genome of *Musca domestica* (Diptera: Muscidae). Journal of Medical Entomology 46: 175-182.

Singh, V and Aballay, A (2012) Endoplasmic reticulum stress pathway required for immune homeostasis is neurally controlled by arrestin-1. Journal of Biological Chemistry 287: 33191-33197.

Small, GJ and Hemingway, J (2000) Molecular characterization of the amplified carboxylesterase gene associated with organophosphorus insecticide resistance in the brown planthopper, *Nilaparvata lugens*. Insect Molecular Biology 9: 647-653.

Spehr, M and Munger, SD (2009) Olfactory receptors: G protein-coupled receptors and beyond. Journal of neurochemistry 109: 1570-1583.

Stanke, M, Schöffmann, O, Morgenstern, B and Waack, S (2006) Gene prediction in eukaryotes with a generalized hidden Markov model that uses hints from external sources. BMC Bioinformatics 7: 62.

Strode, C, Steen, K, Ortelli, F and Ranson, H (2006) Differential expression of the detoxification genes in the different life stages of the malaria vector *Anopheles gambiae*. Insect Molecular Biology 15: 523-530.

Strode, C, Wondji, CS, David, J-P, Hawkes, NJ, Lumjuan, N, Nelson, DR, Drane, DR, Karunaratne, SHPP, Hemingway, J, Black Iv, WC and Ranson, H (2008) Genomic analysis of detoxification genes in the mosquito *Aedes aegypti*. Insect Biochemistry and Molecular Biology 38: 113-123.

Tao, XY, Xue, XY, Huang, YP, Chen, XY and Mao, YB (2012) Gossypol-enhanced P450

gene pool contributes to cotton bollworm tolerance to a pyrethroid insecticide. Molecular Ecology 21: 4371-4385.

Tian, L, Cao, C, He, L, Li, M, Zhang, L, Liu, H and Liu, N (2011) Autosomal interactions and mechanisms of pyrethroid resistance in house flies, *Musca domestica*. International Journal of Biological Sciences 7: 902-911.

Trapnell, C, Roberts, A, Goff, L, Pertea, G, Kim, D, Kelley, DR, Pimentel, H, Salzberg, SL, Rinn, JL and Pachter, L (2012) Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nature Protocols 7: 562-578.

Trejo, J (2003) Protease-activated receptors: new concepts in regulation of G protein-coupled receptor signaling and trafficking. Journal of Pharmacology and Experimental Therapeutics 307: 437-442.

Tsukamoto, M, Baba, Y and Hiraga, S (1961) Mutations and linkage groups in Japanese strains of the housefly. Japanese Journal of Genetics 36: 168-174.

Verrier, F, An, S, Ferrie, AM, Sun, H, Kyoung, M, Deng, H, Fang, Y and Benkovic, SJ (2011) GPCRs regulate the assembly of a multienzyme complex for purine biosynthesis. Nature Chemical Biology 7: 909-915.

Vontas, J, Blass, C, Koutsos, AC, David, JP, Kafatos, FC, Louis, C, Hemingway, J, Christophides, GK and Ranson, H (2005) Gene expression in insecticide resistant and susceptible *Anopheles gambiae* strains constitutively or after insecticide exposure. Insect Molecular Biology 14: 509-521.

Vontas, JG, Small, GJ, Nikou, DC, Ranson, H and Hemingway, J (2002) Purification, molecular cloning and heterologous expression of a glutathione S-transferase involved in

insecticide resistance from the rice brown planthopper, *Nilaparvata lugens*. Biochemical Journal 362: 329-337.

Wen, Z and Scott, JG (1997) Cross-Resistance to Imidacloprid in Strains of German Cockroach (*Blattella germanica*) and House Fly (*Musca domestica*). Pesticide Science 49: 367-371.

White, RE and Coon, MJ (1980) Oxygen activation by cytochrome P-4501. Annual review of biochemistry 49: 315-356.

Wilding, CS, Smith, I, Lynd, A, Yawson, AE, Weetman, D, Paine, MJI and Donnelly, MJ (2012) A cis-regulatory sequence driving metabolic insecticide resistance in mosquitoes: Functional characterisation and signatures of selection. Insect Biochemistry and Molecular Biology 42: 699-707.

Wilkins, RM, Ahmed, S and Mantle, D (1999) Comparative effect of fenitrothion treatment on intracellular protease activities in insecticide-resistant and susceptible strains of *Musca domestica* L. Comparative Biochemistry and Physiology Part C: Pharmacology, Toxicology and Endocrinology 124: 337-343.

Wu, H-W, Tian, H-S, Wu, G-L, Langdon, G, Kurtis, J, Shen, B, Ma, L, Li, X-L, Gu, Y, Hu, X-B and Zhu, C-L (2004) Culex pipiens pallens: identification of genes differentially expressed in deltamethrin-resistant and -susceptible strains. Pesticide Biochemistry and Physiology 79: 75-83.

Xu, Q, Liu, H, Zhang, L and Liu, N (2005) Resistance in the mosquito, *Culex quinquefasciatus*, and possible mechanisms for resistance. Pest Management Science 61: 1096-1102.

Xu, Q, Wang, H, Zhang, L and Liu, N (2006a) Kdr allelic variation in pyrethroid resistant mosquitoes, *Culex quinquefasciatus* (S.). Biochemical and Biophysical Research Communications 345: 774-780.

Xu, Q, Wang, H, Zhang, L and Liu, N (2006b) Sodium channel gene expression associated with pyrethroid resistant house flies and German cockroaches. Gene 379: 62-67.

Yang, S-H, Sharrocks, AD and Whitmarsh, AJ (2013) MAP kinase signalling cascades and transcriptional regulation. Gene 513: 1-13.

Yang, T and Liu, N (2011) Genome analysis of cytochrome P450s and their expression profiles in insecticide resistant mosquitoes, *Culex quinquefasciatus*. PloS One 6: e29418.

Zhu, F, Feng, JN, Zhang, L and Liu, N (2008) Characterization of two novel cytochrome P450 genes in insecticide-resistant house-flies. Insect Molecular Biology 17: 27-37.

Zhu, F and Liu, N (2008) Differential expression of CYP6A5 and CYP6A5v2 in pyrethroid-resistant house flies, *Musca domestica*. Arch Insect Biochem Physiol 67: 107-119.

Zhu, F, Parthasarathy, R, Bai, H, Woithe, K, Kaussmann, M, Nauen, R, Harrison, DA and Palli, SR (2010) A brain-specific cytochrome P450 responsible for the majority of deltamethrin resistance in the QTC279 strain of Tribolium castaneum. Proceedings of the National Academy of Sciences 107: 8557-8562.

Table 3.1 Total Illumina reads for each of the *Musca domestica* strains/lines tested.

Sample	SRA* accession	Total PE** reads passing Chastity filter	Discarded‡	Reads used
aabys	SRR521286	40284931	2794419	37490512
CS	SRR521290	34589399	2561540	32027859
ALHF	SRR521289	30329576	2208815	28120761
ALHF- replicate	SRR521288	26151304	1406309	24744995

*Sequence Read Archive, National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/sra/)

**51nt Paired-end reads with an average insert size of 200 nt.

‡Discarded reads were removed using Trimmomatic based on a sliding window quality cut off of (4:15) and a minimum length of 36nt after adapter removal. In addition all mate pairs were excluded from analysis if one of the mate pairs was rejected by Trimmomatic.

Table 3.2 Results of the homology testing for the transcriptome assembly for the adult female *Musca domestica* ALHF strain.

Assembly results	**Total number**	**Unique annotation**
Contigs generated		
Non annotated contigs[†]	8229	-
Predicted ORFs	5469	-
5' Partial ESTs	519	-
3' Partial ESTs	171	
Contigs annotated		
No annotation[*]	184	-
D. melanogaster[‡]	5730	4265
Pfam family[**]	5549	2148
KEGG[††]	2795	1967

[†] "Non annotated contig" indicates that no ORF was identified in the contig

[*] "No annotation" indicates the presence of an ORF, a start, and a stop codon, but no homology to the Pfam family HMM search, *Drosophila melanogaster*, or KEGG

[**] Pfam HMM A (v26.0) Nov 2011. http://www.pfam.sanger.ac.uk/

[‡] version r5.46 20 Feb, 2009. http://www.flybase.org/

[††]Release 62.0 April 1, 2012. http://www.genome.jp/kegg/

Table 3.3 Higher level SCOP annotation for the predicted genes from the adult *Musca domestica* transcriptome based on sequence homology to *Drosophila melanogaster*[§].

General function	Detailed function	Superfamilies	Predicted genes
Metabolism	Amino acids metabolism /transport	6	17
	Carbohydrate metabolism /transport	11	94
	Coenzyme metabolism /transport	13	45
	Electron transfer	7	28
	Energy	23	45
	Lipid metabolism /transport	4	18
	Nitrogen metabolism /transport	1	2
	Nucleotide metabolism /transport	14	74
	Other enzymes	51	344
	Photosynthesis	1	2
	Polysaccharide metabolism /transport	2	19
	Redox	28	180
	Secondary metabolism	6	68
	Transferases	14	143
Regulation	DNA-binding	32	218
	Kinases/phosphatases	7	245
	Other regulatory function	5	12
	Receptor activity	2	5
	RNA binding, metabolism /transport	12	113
	Signal transduction	30	301
Information	Chromatin structure	4	4
	DNA replication/repair	15	141
	RNA processing	6	25
	Transcription	7	17
	Translation	44	77
Extra-cellular processes	Blood clotting	1	13
	Cell adhesion	18	101
	Immune response	4	22
	Toxins/defense	2	5
Intra-cellular processes	Cell cycle, Apoptosis	10	26
	Cell motility	11	38
	Ion metabolism /transport	13	173
	Phospholipid metabolism /transport	4	36
	Proteases	21	304
	Protein modification	15	105
	Transport	22	118
General	General	12	348
	Ion binding	1	6
	Ligand binding	1	4
	Protein interaction	19	178
	Small molecule binding	11	403
Other	Unknown function	26	75

	Viral proteins	1	4
NONA	not annotated	1	1963
	TOTAL	537	6159

[§]SCOP Superfamily annotation for *D. melanogaster* (v. 66_539)

Table 3.4 Gene expression values and the predicted autosomal interactions for the selected genes linked to pyrethroid resistance in *Musca domestica* as assayed by qPCR.

SCOP† Functional annotation		Predicted gene function	Accession number	Relative gene expression ± SE					
General Metabolism	Detailed Other	Pfam annotation§	Gene	ALHF	A2345	A1345	A1245	A1235	A1234
		Carboxylesterase	ALHF_03407.g2111	3.1±0.2	2.4±0.1	1.8±0.12*	3.0±0.1	2.9±0.1	1.1±0.11*
			ALHF_05628.g3847	6.6±1.3	2.1±0.3*	1.0±0.1*	4.0±0.4	3.2±0.1	4.4±0.8
			ALHF_00771.g422	26±5.5	14±1.2	0.24±0.07*	18±1.5	26±2	3.1±0.3*
	Redox	Cytochrome P450	ALHF_04553.g3033	2.7±0.09	2.5±0.1	2.7±0.2	2.5±0.08	3.0±0.2	0.8±0.03*
			ALHF_05265.g3608	510±28	300±5.0	15±2.6*	530±16	500±18	300±4.3
			ALHF_03088.g1882	310±9.3	310±10	190±9.0	150±4.4	310±18	120±6.4*
			ALHF_02791.g1651	12±0.2	7.4±0.7	7.9±0.04	8.8±0.2	13±0.9	2.4±0.09*
			ALHF_07553.g4857	6.2±0.6	2.6±0.4*	3.1±0.6*	9.5±2.3	6.2±0.5	2.6±0.3*
			ALHF_04445.g2939	3.0±0.1	1.6±0.2*	1.8±0.4*	4.0±0.4	3.0±0.04	1.0±0.2*
			ALHF_04444.g2938	3.6±0.06	1.9±0.1	2.1±0.09	5.6±2.1	3.6±0.1	1.0±0.2*
			ALHF_03006.g1816	2.3±0.09	1.9±0.3	1.1±0.08*	3.8±0.4	2.1±0.06	0.9±0.1*
			ALHF_01822.g1025	4.3±1.0	1.3±0.2*	1.9±0.4*	5.6±1.2	4.2±0.2	0.6±0.1*
			ALHF_04730.g3176	2.1±0.02	1.2±0.2	2.6±0.3	7.1±2.7	1.8±0.09	0. 7±0.09*
			ALHF_03063.g1860††	-	-	-	-	-	-
			ALHF_05136.g3505	2.4±0.3	0.9±0.1*	0.4 ±0.2*	3.2±0.4	2.5±0.2	1.5±0.1
			ALHF_07623.g4891	1.9±0.2	1.8±0.2	0.4±0.2*	4.6±1.2	2.2±0.1	1.0±0.1 *
			ALHF_08221.g5182	2.6±0.2	1.6±0.04	1.8±0.3	0.6 ±0.04	2.9±0.04	0.6±0.05*
			ALHF_04665.g3125	2.9±0.08	1.0±0.05*	1.1±0.06*	0.8 ±0.03*	2.8±0.09	1.0±0.03*
			ALHF_01339.g731	2.0±0.08	2.3±0.07	0.2±0.03*	1.1±0.03	2.1±0.07	0.8±0.04*
			ALHF_04736.g3182††	-	-	-	-	-	-
			ALHF_03849.g2446††	-	-	-	-	-	-
		Glutathione-S-transferase	ALHF_04900.g3328	2.8±0.4	2.1±0.2	0.9±0.08*	2.3±0.2	2.8±0.2	2.2±0.2
			ALHF_04476.g2964	2.4±0.2	1.7±0.2	0.6 ±0.07*	1.9±0.2	1.8±0.08	0.7±0.02*
			ALHF_03731.g2351	2.4±0.3	1.8±0.3	0.4±0.03*	1.2±0.08	1.5±0.07	0.6±0.08*
			ALHF_04477.g2965	1.5±0.02	1.1±0.11	0.6±0.07*	1.7±0.1	1.3±0.06	0.8±0.1*
			ALHF_03145.g1917	1.9±0.2	0.7±0.06*	0.5±0.01*	1.6±0.2	1.8±0.06	0.9±0.02*
Regulation	Kinase / phosphatase	Protein kinase domain	ALHF_02546.g1487	2.9±0.09	3.8±0.75	0.4±0.04*	2.2±0.3	2.6±0.2	0.5±0.08*
			ALHF_00685.g381	3.6±0.4	3.2±0.7	0.8 ±0.2*	1.1±0.09*	3.3±0.3	0.5±0.04*
			ALHF_03462.g2147	3.7±0.7	4.1±0.4	0.5±0.01*	1.5±0.2*	3.5±0.2	0.2±0.2*
			ALHF_02885.g1722	1.8±0.2	1.8±0.2	0.6 ±0.2*	1.5±0.8	1.6±0.1	0.8 ±0.1*
			ALHF_00823.g452	2.0±0.2	1.8±0.2	0.8±0.2*	1.2±0.09	1.8±0.2	0.7±0.1 *
			ALHF_04500.g2986	1.7±0.09	1.1±0.04	0.7±0.2*	1.5±0.1	1.7±0.04	0.9±0.06*
			ALHF_04095.g2646	1.7±0.1	1.1±0.06	0.9 ±0.2	1.0±0.08	1.5±0.03	1.0±0.2
			ALHF_01595.g882	2.3±0.2	0.8±0.06*	0.8±0.1*	1.4±0.2	2.5±0.2	1.0±0.1*
			ALHF_01832.g1033	2.3±0.2	2.8±0.6	0.4±0.09*	1.0±0.06*	1.9±0.1	0.4±0.06*
			ALHF_08078.g5122	2.6±0.7	1.0±0.04	0.5±0.03*	1.7±0.1	2.4±0.2	0.6±0.06*
			ALHF_11277.g6269	2.3±0.1	0.7±0.05*	0.2±0.02*	1.8±0.1	2.1±0.2	0.8±0.2*
			ALHF_11442.g6384††	-	-	-	-	-	-
			ALHF_00727.g395††	-	-	-	-	-	-
		Protein tyrosine kinase	ALHF_11829.g6650	1.7±0.08	1.3±0.09	1.5±0.17	2.2±0.4	1.9±0.1	1.3±0.1
			ALHF_11144.g6194	4.0±0.8	3.2±0.2	1.4±0.2*	2.5±0.7	5.1±0.2	0.8±0.04*
			ALHF_09312.g5609	1.9±0.1	1.3±0.1	0.7±0.1*	1.7±0.2	1.7±0.06	0.7±0.1*

			ALHF_10712.g5974	1.8±0.1	0.7±0.04*	0.6±0.2*	1.7±0.3	1.8±0.08	0.7±0.1*
			ALHF_07173.g4665	1.5±0.1	1.3±0.1	0.8±0.05*	1.2±0.2	1.6±0.07	0.5±0.1*
			ALHF_03649.g2289	2.1±0.1	0.9±0.07*	0.4±0.2*	1.5±0.2	2.1±0.2	1.1±0.06*
			ALHF_05773.g3933††	-	-	-	-	-	-
			ALHF_11245.g6252	1.7±0.07	0.6±0.03*	0.2±0.06*	0.5±0.2*	1.9±0.09	0.9±0.09*
		Protein-tyrosine phosphatase	ALHF_11768.g6612	39±4.9	1.6±0.07*	0.6±0.05*	0.4±0.1*	36±3.8	0.5±0.03*
			ALHF_03863.g2457	1.5±0.06	0.9±0.09	0.5±0.04*	1.2±0.2	1.7±0.06	0.9±0.08*
	Signal transduction	GPCR (rhodopsin family)	ALHF_01760.g986	1.6±0.09	1.0±0.08	0.9±0.1*	1.1±0.03	1.7±0.1	1.0±0.1
			ALHF_02400.g1393	1.7±0.07	0.9±0.07	0.7±0.05*	1.2±0.2	1.6±0.1	0.9±0.1*
			ALHF_06811.g4468	1.7±0.1	0.9±0.03*	0.6±0.07*	1.2±0.2	1.7±0.1	0.8±0.08*
			ALHF_07519.g4838	1.7±0.07	1.1±0.09	0.2±0.06*	0.9±0.2	1.6±0.06	0.9±0.04
			ALHF_02706.g1581	1.4±0.1	0.52±0.1*	0.37±0.1*	1.2±0.3	1.3±0.04	1.2±0.11
			ALHF_04422.g2918††	-	-	-	-	-	-
		Adenylate and Guanylate cyclase catalytic domain	ALHF_01050.g580	2.6±0.3	1.3±0.1	0.8±0.09*	1.1±0.2*	2.2±0.21	1.3±0.3*
			ALHF_07748.g4948	2.1±0.2	1.2±0.1	0.6±0.05*	1.3±0.2	1.8±0.05	1.0±0.2*
		Serpentine type 7TM GPCR chemoreceptor Srw	ALHF_01902.g1074	-	-	-	-	-	-
Intra-cellular processes	Proteases	Serpins	ALHF_07374.g4763	4.5±0.8	1.8±0.2	0.6±0.2*	1.3±0.1*	2.3±0.2	0.5±0.08*
			ALHF_01182.g646	1.7±0.1	0.8±0.07	0.3±0.07*	1.1±0.08	1.8±0.1	0.7±0.08*
		Carboxypeptidases	ALHF_04057.g2616	2.2±0.3	1.5±0.1	0.5±0.1*	1.3±0.2	1.8±0.07	0.9±0.1*
			ALHF_05871.g3981††	-	-	-	-	-	-
		Subtilase	ALHF_00530.g295††	-	-	-	-	-	-
		Aspartyl protease	ALHF_06529.g4317	2.5±0.1	2.3±0.2	1.4±0.2*	1.3±0.1*	2.1±0.1	1.4±0.01*
		Peptidases	ALHF_00761.g417	780±50	530±100	630±60	23±2*	650±110	530±20
			ALHF_03218.g1970	29±2.1	30±3	35±5	2.6±0.4*	17±3*	8.2±1*
			ALHF_02207.g1267	145±10	150±20	120±3	5.0±0.3*	4.6±2*	4.9±0.6*
			ALHF_07511.g4836	2.1±0.1	1.5±0.2	0.52±0.04*	3.2±0.2	0.52±0.07*	0.41±0.08*
			ALHF_01861.g1049††	-	-	-	-	-	-
			ALHF_05334.g3663	1.4±0.04	1.0±0.1	0.66±0.06*	1.5±0.1	0.75±0.07*	0.40±0.1*

†Structural Classification of Proteins (SCOP) http://supfam.cs.bris.ac.uk/SUPERFAMILY/

‡Autosomal interactions are named after the combination of the autosomes donated by the aabys strain in each case where a *M. domestica* autosomal combination line had a level of gene expression significantly lower than the ALHF strain at the α=0.05 level of significance.

§Gene predicted function as evidenced by the Pfam HMM-A homology (http://pfam.sanger.ac.uk/)

*Indicates that the gene expression value within a given *M. domestica* autosomal line was significantly lower than the expression in the parental ALHF strain at the α=0.05 level of significance

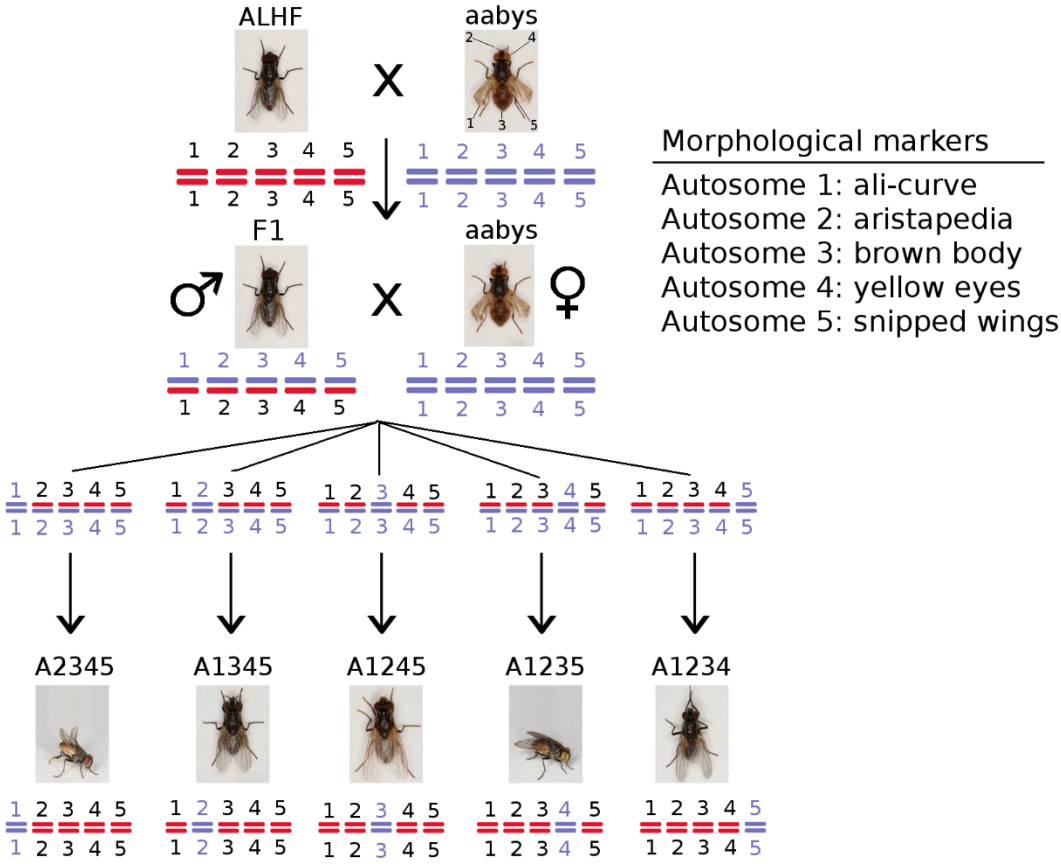


Figure 3.1 Schematic for the generation of the *M. domestica* combination strains used in our study. Strain ALHF is a highly insecticideresistant strain while the aabys strain is an insecticide-susceptible strain that possesses five recessive morphological markers, with each morphological marker being uniquely present on one autosome. The images along the bottom row show the recessive morphological markers unique to each of the combination strains.

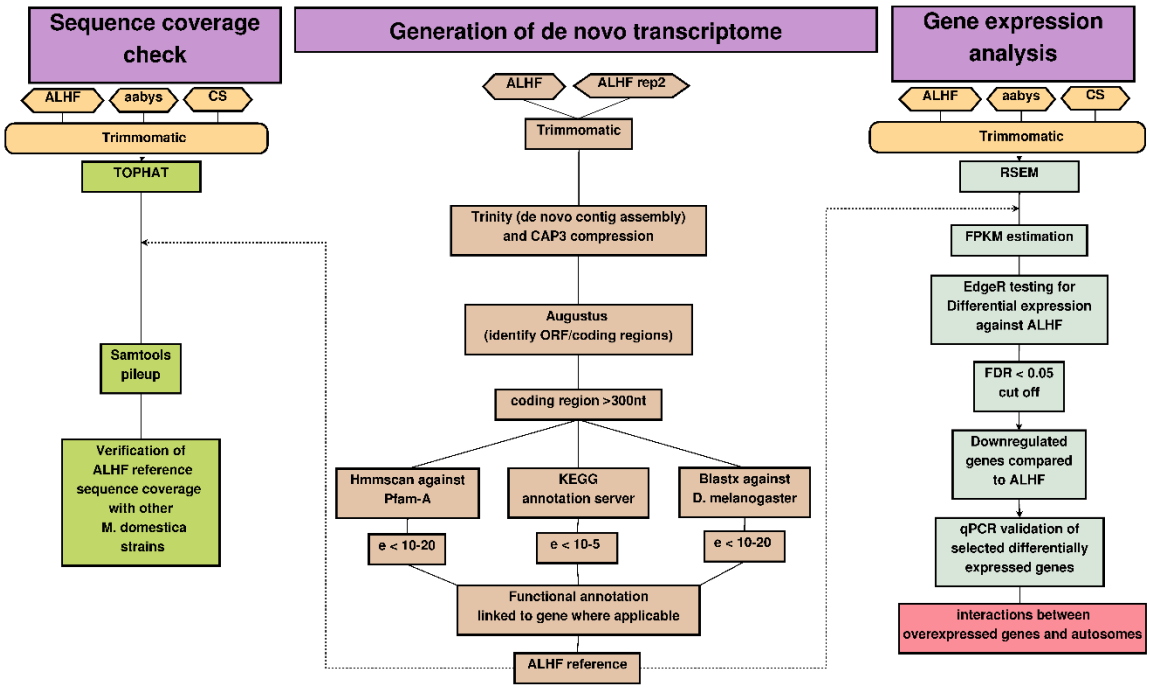


Figure 3.2. Data analysis pipeline for the generation of the *M. domestica* predicted gene set and differential gene expression testing. Hexagons represent the raw data used in this study, while terms within boxes represent either the programs, or the filtering steps used in the data analysis. The direction of the arrows indicates the flow of data processing.

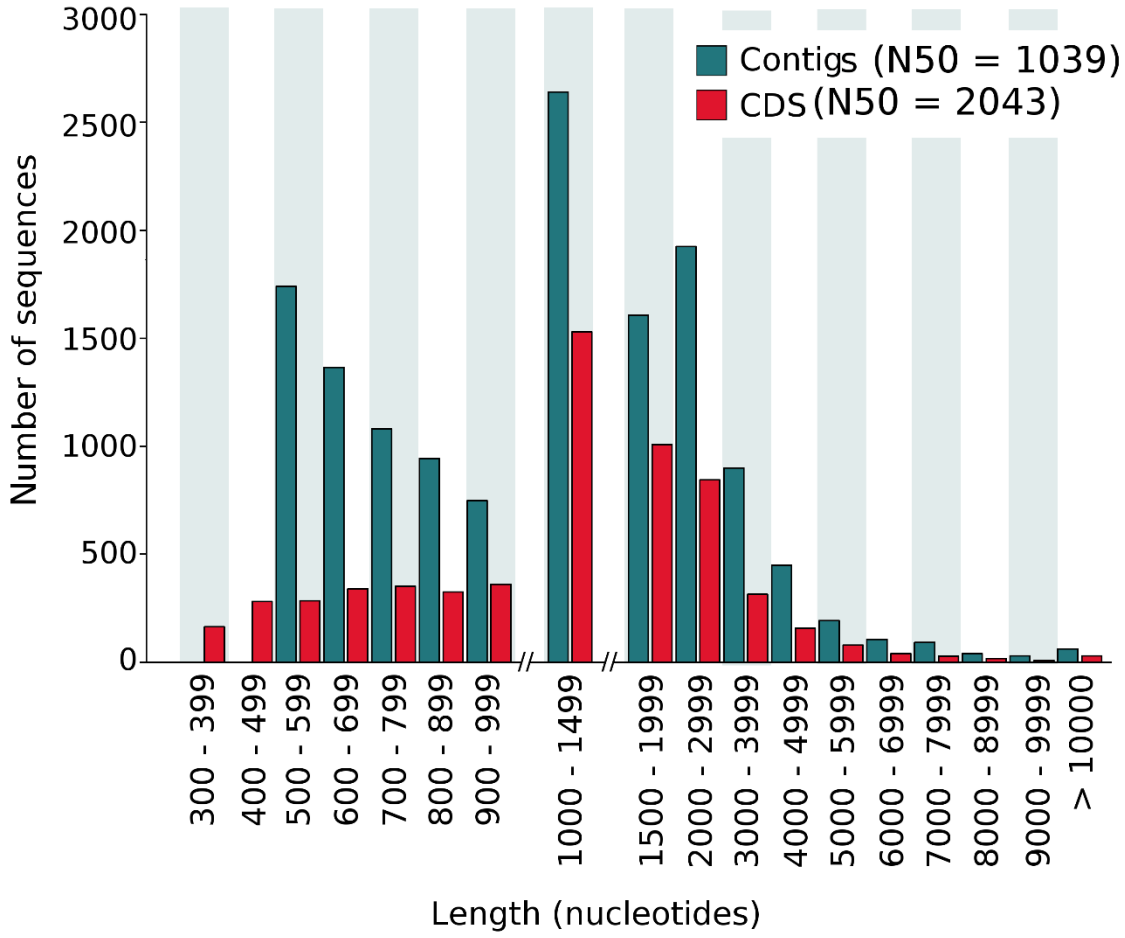


Figure 3.3 Nucleotide length distributions for the *M. domestica* ALHF strain raw assembled contigs and predicted coding regions (CDS). Coding region lengths were predicted using Augustus (version 2.5.5) under the "fly" model and include both complete (5469 sequences) open reading frames (ORFs) and partial ORFs (690 sequences). A partial ORF means any sequence that is predicted to be missing either the start, or the stop codon, but not both.

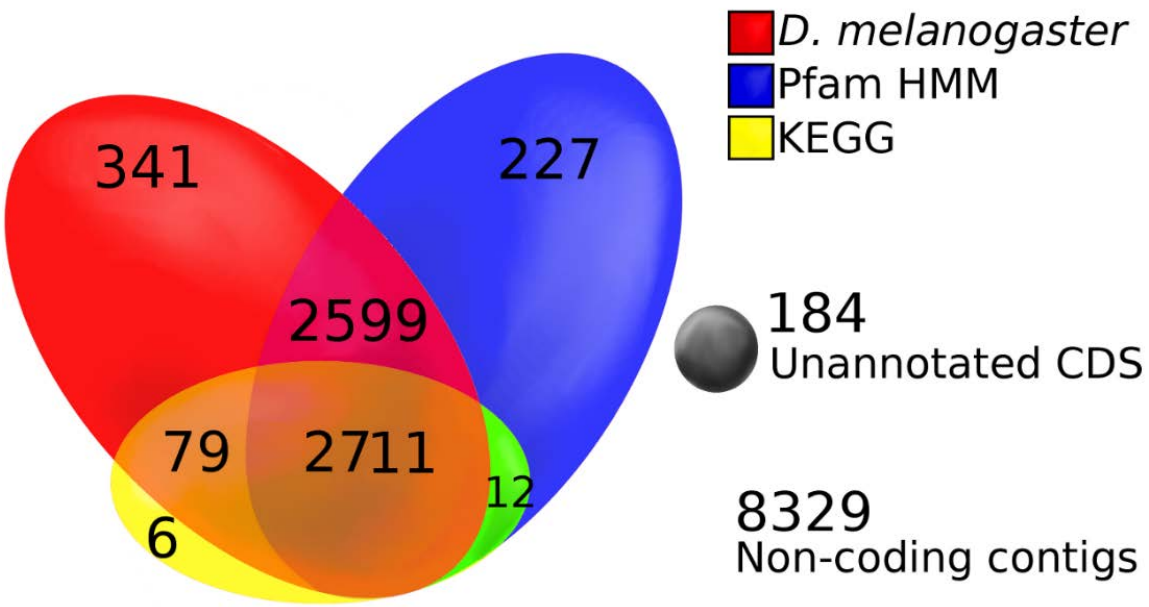


Figure 3.4 Venn diagram for the annotation obtained for the *M. domestica* ALHF strain predicted gene set. Overlapping ellipses represent predicted genes from the Pfam-A (v26.0), the *D. melanogaster* proteome (v. r5.46), and the Kyoto Encyclopedia of Genes and Genomes (KEGG) automatic annotation server (KAAS) that could be annotated in two or more of the databases used to predict gene function. An e-value threshold for homology detection was fixed at 10^{-20} for all databases. The circle excluded from the overlapping ellipses represent sequences which contained coding region, but had no homology to any of the three databases used for gene prediction.

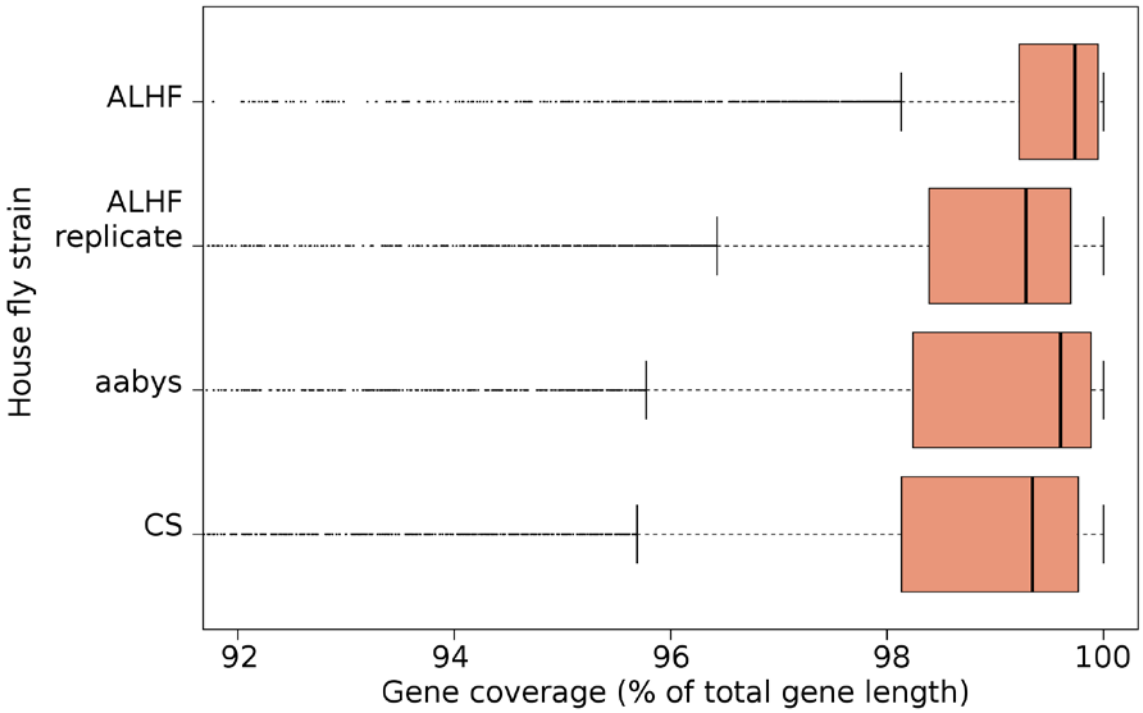


Figure 3.5 Box and whisker plots representing the interquartile ranges (IQR) for the nucleotide coverage for each of the ALHF strain predicted genes for each of the *M. domestica* strains tested. The dependent axis has been broken to make the IQR and median values discernible. The solid line within each of the boxes represents the median value for the gene coverage for each of the house fly strains.

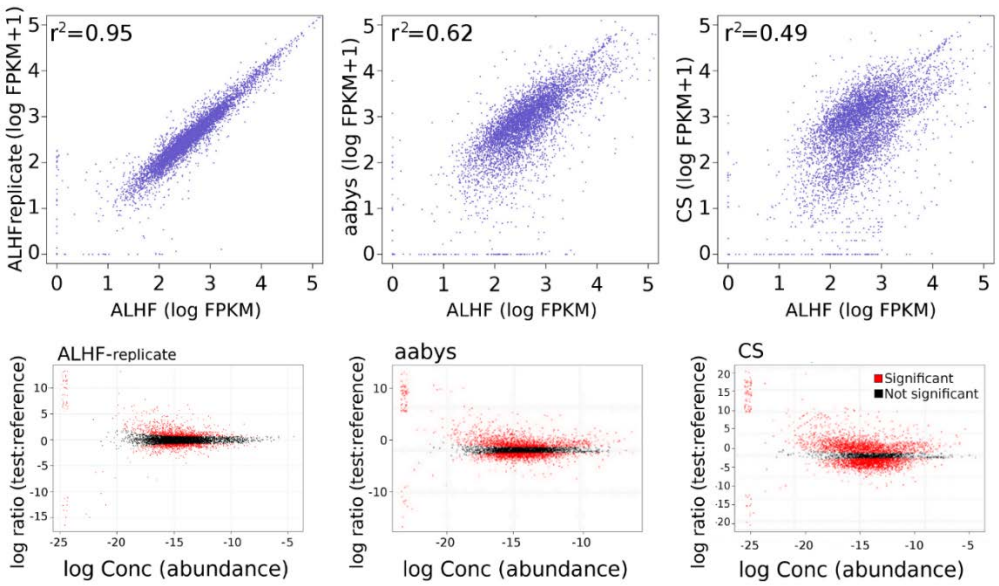


Figure 3.6 Correlation of the gene expression levels (FPKM) for all of the *M. domestica* strains tested versus the ALHF pyrethroid-resistant strain (upper panels). Scatterplots for the testing of the differential gene expression compared to the ALHF strain (lower panels). In the upper panels, the points closest to the 1:1 line represent genes that had the same gene expression value for the ALHF strain and the tested *M. domestica* strain. In the lower panels, each point represents a gene, with red points below the central axis indicating the genes that were down-regulated in the tested *M. domestica* strain compared to the ALHF strain, thus the red points below the horizontal axis on the lower panels represent the genes that were up-regulated in the pyrethroid-resistant ALHF strain and putatively linked to insecticide resistance.

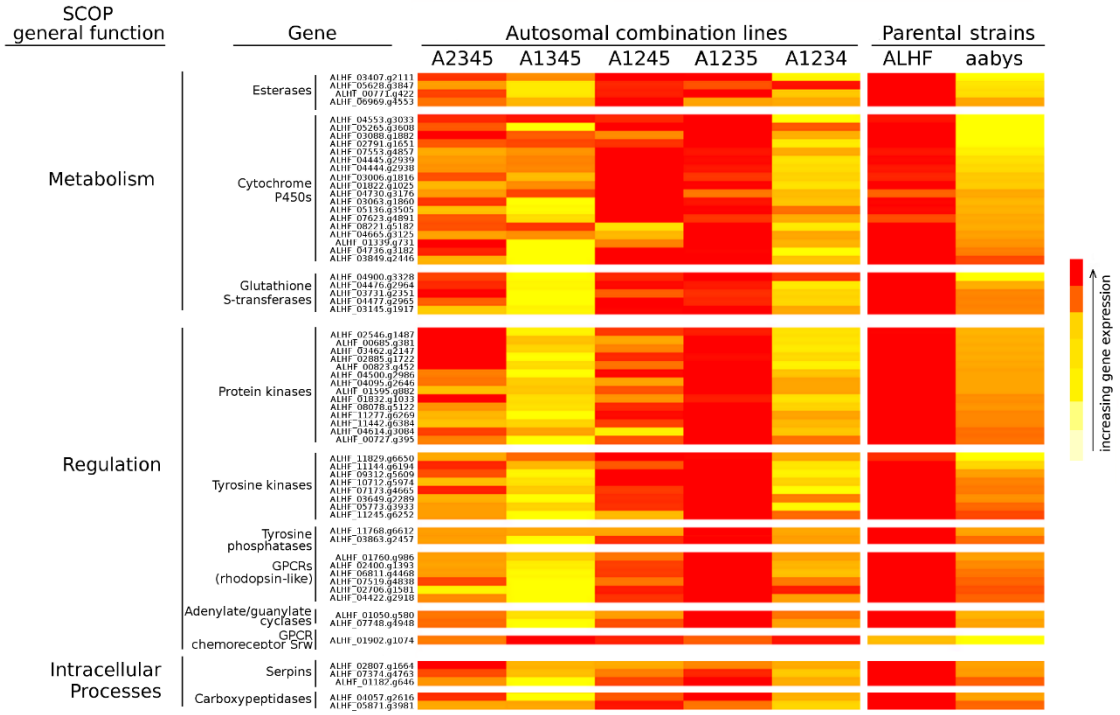


Figure 3.7 Heat map of the gene expression values (within gene) for each of the genes tested by qPCR to validate the gene expression levels within the different *M. domestica* lines and the parental ALHF and aabys strains. Colors scaled from yellow to red indicate low to higher gene expression, respectively.

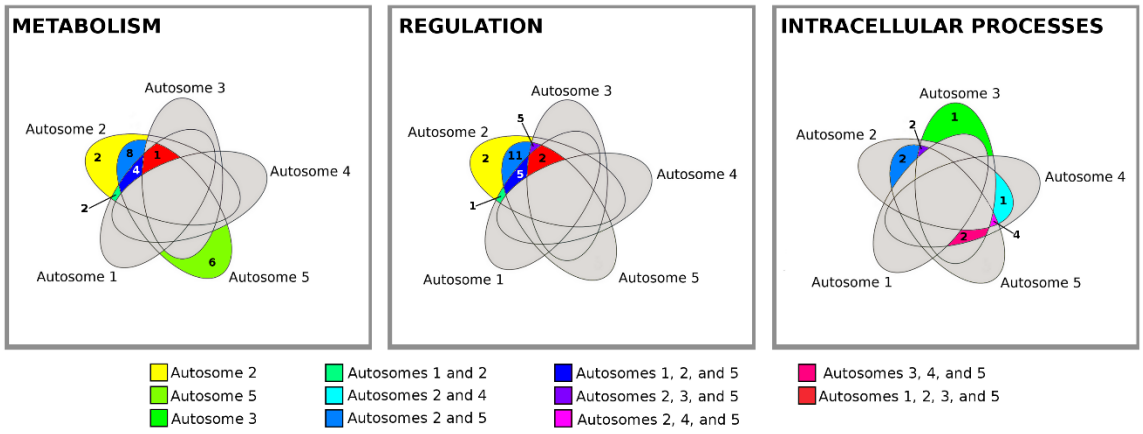


Figure 3.8 Interaction of the autosomal-linked genes associated with pyrethroid resistance in the ALHF strain of *M. domestica.* The overlapping areas between the ellipses indicate the autosomal interaction that existed for the genes that were up-regulated in the ALHF strain for two, or more, of the autosomal combination lines. One gene, ALHF_02807.g1664, having an autosome 2,3,4,5 interaction is not shown in this figure.

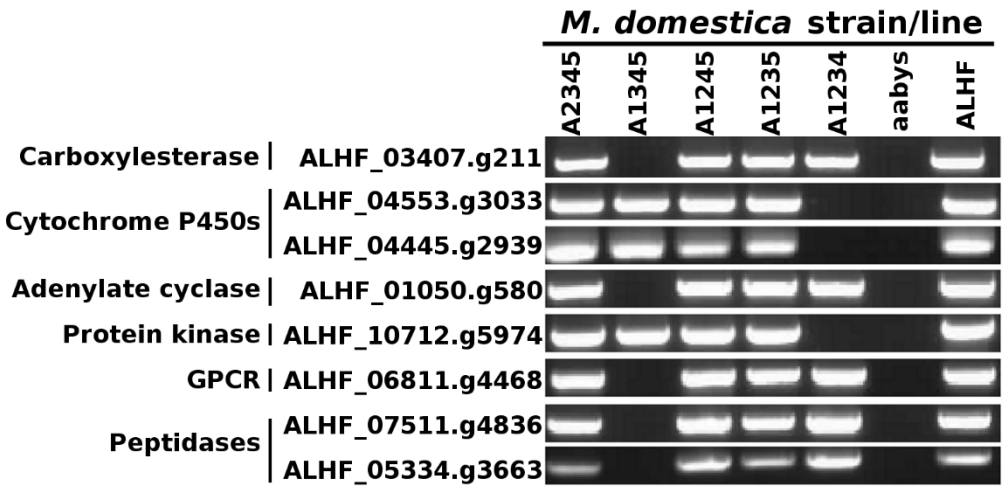


Figure 3.9 Allele-specific RT-PCR autosomal mapping of the *M. domestica* genes used for the transgenic *D. melanogaster* functional assays. The presence of a band within *M. domestica* stain/line indicates that the gene is present. The absence of a band within both the aabys and an autosomal combination line indicates that the gene is located on the autosome donated by aabys (eg/ the absence of a band in the A1234 autosomal combination line indicates that the gene was present on autosome 5).

Chapter 4: Roles of P450 Genes in Permethrin Resistance of the House Fly, *Musca domestica*

4.1 Abstract

To characterize the P450 genes that play important roles in the pyrethroid resistance of house flies, 86 cytochrome P450 genes were selected based on our whole transcriptome analysis of the house fly, to conduct the expression profile analysis in different house fly strains with different levels of permethrin resistance and autosome combinations. Our study showed that 4 P450 genes, *CYP6A36, CYP6A52, CYP6D,* and *CYP4S24*, were co-up-regulated in insecticide-resistant house flies compared to -susceptible house flies, and the expression of these genes was regulated by *cis* or *trans* regulatory factors/genes, which were mainly on autosomes 1, 2 and 5. Transgenic expression analysis of these four P450 genes in *Drosophila melanogaster* demonstrated that elevated expression of each of these 4 genes confers different levels of resistance to permethrin in the transgenic *Drosophila*, with relative lower level in *CYP4S24* transgenic

flies. Homology modeling and permethrin docking analysis further suggested potential abilities of all 4 P450 genes to metabolize permethrin, generating multiple metabolites except *CYP4S24*, which may have low degree metabolism of permethrin, consistent with lower level of resistance to permethrin in the *CYP4S24* transgenic flies. Taken together, the study provides a global picture of P450 gene expression, regulation, autosomal interaction, and function in insecticide resistance of house flies.

4.2 Introduction

The development of resistance to insecticides in insect pests is becoming a global challenge in the insect pest control confrontation. Efforts to characterize the molecular mechanisms involved in insecticide resistance have thus been focused to build better understanding of resistance development, which is vital for practical applications such as the design of novel strategies to prevent or minimize the spread and evolution of resistance development and the control of insect pests (Roush et al., 1990). While the interaction of multiple mechanisms or genes in response to the development of insecticide resistance has been extensively recognized, transcriptional up-regulation of the detoxification machinery, increasing metabolism of insecticides into less harmful

substances and facilitating insecticide excretion are known to play an important role in allowing insects to defend themselves against insecticides (Xu et al., 2005). Cytochrome P450s are attributed as one of the important components in the detoxification machinery of insects, the up-regulation of which underlies the development of insecticide resistance in many insect species. It has been suggested that new patterns of gene expression may arise via a variety of mechanisms involving changes to upstream regulators (change in *trans*) and mutations of the noncoding regulatory DNA sequences (e.g., enhancers) of a gene (change in *cis*) (Rebeiz et al., 2011). Indeed, different patterns of gene expression between resistant and susceptible insect populations and the up-regulation of P450 has also been found to be controlled by unknown *trans*- or *cis*-regulatory factors have demonstrated (Grant and Hammock, 1992; Carino et al., 1994; Liu and Scott, 1996; Maitra 1996, 2000), suggesting that not only is insecticide resistance conferred via multiple P450 gene up-regulation, but it is mediated through the interaction of regulatory factors and resistance P450 genes. However, it is not yet clear how many P450 genes precisely are involved in insecticide resistance in a single insect, no regulatory factors in insecticide resistance have yet been identified, and no resistance gene interaction has been examined. Recent advances in genome/whole transcriptome sequencing technology

have provided opportunities for discovery of gene/mechanism/regulatory factor involvement in insecticide resistance (Li et al. 2013).

The house fly, *Musca domestica*, is a major domestic, medical and veterinary pest that causes more than 100 human and animal intestinal diseases (Scott and Lettig, 1962; Keiding, 1986; Greenberg, 1965, Scott et al. 2000). However, the major barrier in the house fly control, as other insect species, is their remarkable ability to develop not only resistance to the insecticide used against them but also cross-resistance to unrelated classes of insecticides (Greenberg, 1965; Liu and Yue, 2000; Pap and Tóth, 1995; Wen and Scott, 1997). Because of their ability in rapid development of resistance and cross-resistance to insecticides, their well described linkage map for five autosomes and two sex chromosomes (X and Y) (Hiroyoshi, 1960; Tsukamoto et al. 1961, Nickel and Wagoner, 1974; Hiroyoshi, 1977, Milani et al. 1967), and their relatively well studied biochemistry and genetics of insecticide resistance, the house fly has demonstrated to be a useful model to study and predict resistance in not only themselves but also other insect species. With the availability of the first adult transcriptome database of the house fly *M. domestica* (Li et al. 2013), the current study focused on characterizing key P450 genes that were involved in permethrin resistance through expression profiles analysis of a total

of 86 cytochrome P450 genes in house fly strains with different levels of resistance to pyrethroid/permethrin; examined interaction of the resistant key P450 genes and factors on different autosomes through house fly lines with different combination of autosomes from a resistant house fly strain, ALHF; exploring the function of key P450 genes in insecticide resistance using *Drosophila* transgenic techniques; and analyzing P450 modeling and permethrin docking to investigate the roles of P450s that are involved in pyrethroid resistance in house flies.

4.3 Materials and Methods

4.3.1 House fly strains and lines

Three house fly strains were used in this study. ALHF is a wild-type insecticide-resistant strain collected from a poultry farm in Alabama in 1998. This strain was further selected with permethrin for six generations after collection to reach a high level of resistance, and has been maintained under biannual selection with permethrin (Liu and Yue 2000; Tian et al. 2011). aabys is an insecticide-susceptible strain with recessive morphological markers ali-curve (*ac*), aristapedia (*ar*), brown body (*bwb*), yellow eyes (*ye*), and snipped wings (*snp*) on autosomes 1, 2, 3, 4, and 5, respectively. CS is a wild-

type insecticide-susceptible strain, which has been kept in laboratory breeding for more than five decades. Both aabys and CS strains were originally obtained from Dr. J. G. Scott (Cornell University).

A cross of ALHF female and aabys male was performed with each of ~400 flies, and the F1 males (~400 flies) were then backcrossed to aabys female. Five back-cross (BC_1) lines with the following genotypes were isolated: *ac/ac*, +/*ar*, +/*bwb*, +/*ye*, +/*snp*; +/*ac*, *ar/ar*, +/*bwb*, +/*ye*, +/*snp*; +/*ac*, +/*ar*, *bwb/bwb*, +/*ye*, +/*snp*; +/*ac*, +/*ar*, +/*bwb*, *ye/ye*, +/*snp*; and +/*ac*, +/*ar*, +/*bwb*, +/*ye*, *snp/snp* (Li et al. 2013). Homozygous lines *ac*/*ac*, +/+, +/+, +/+, +/+ (A2345); +/+, *ar* /*ar*, +/+, +/+,+/+ (A1345); (+/+, +/+, *bwb*/*bwb*, +/+, +/+ (A1245); +/+, + /+, +/+, *ye*/*ye*, +/+ (A1235) and +/+, +/+, +/+, +/+, *snp*/*snp* (A1234) were accomplished by sorting for appropriate phenotypic markers and selecting with permethrin at a corresponding dose that caused ~70% mortality for each of lines for three generations. One hundred single-pair crossing of each of lines for the desired phenotype and genotype were then set up (Liu and Yue 2000; Tian et al. 2011). The name of each line indicates which of its autosomes bear wild-type markers from ALHF. For instance, the A2345 strain has wild-type markers on autosomes 2, 3, 4 and 5 from ALHF and the mutant marker on autosome 1 from aabys.

4.3.2 RNA extraction, cDNA preparation and gene expression detection

A total of 20 3-day old adult female house flies from each of three house fly strains (ALHF, aabys and CS) and five house fly lines (A2345, A1345, A1245, A1235, A1234) were flash frozen on dry ice and immediately processed for RNA extraction. Total RNA (0.5 μg/sample) from each house fly sample was reverse-transcribed using SuperScript II reverse transcriptase (Stratagene) in a total volume of 20 μl. The quantity of cDNAs was measured using a spectrophotometer prior to qRT-PCR, which was performed with the SYBR Green master mix Kit and ABI 7500 Real Time PCR system (AppliedBiosystems). Each qRT-PCR reaction (15 μl final volumes) contained 1× SYBR Green master mix, 1 μl of cDNA, and a gene specific primer pair at a final concentration of 0.3-0.5 μM (Table S4.1). A 'no-template' negative control and all samples were performed in triplicate. Relative expression levels of specific genes were calculated by the $2^{-\Delta\Delta Ct}$ method using SDS RQ software (Livak and Schmittgen 2001). The β-actin gene, an endogenous control, was used to normalize expression of target genes (Zhu et al. 2008a). Each experiment was repeated three times with different preparations of RNA samples. The statistical significance of the gene expressions was calculated using Student's t-test for all

pairwise sample comparisons and a one-way analysis of variance (ANOVA) for multiple sample comparisons (SPSS v19.0), a value of $P < 0.05$ was considered statistically significant.

4.3.3 Autosomal mapping of P450 genes in *M. domestica*

Five house fly BC_1 lines were used to determine genetic linkage of up-regulated P450 genes. Allele specific PCR was conducted using the cDNA from 5 BC_1 lines (Liu et al. 1995). The ALHF allele specific primer pair was designed based on the specific sequence of the genes from ALHF by placing a specific nucleotide polymorphism at the 3' end of each primer to permit preferential amplification of specific alleles from ALHF. Two rounds of PCR were conducted. For the first PCR reaction, the allele-independent primer pairs (Table S4.1) were used for generating P450 cDNA fragments, respectively. The first PCR solution with cDNA template and a primer pair were heated to 95°C for 3 min, followed by 35 cycles of 95°C for 30 s, 60°C for 30 s, and 72°C for 1 min, then 72°C for 10 min. The second PCR was employed with 0.5 µl of the first round PCR reaction solution and the allele specific primer pair (Table S4.1). The second PCR reaction was heated to 95°C for 3 min, followed by 35 cycles of 95°C for 30 s, 62°C for

30 s, and 72°C for 30 s, then 72°C for 10 min. One of each allele specific primer pair was designed based on the specific sequence of the genes from ALHF by placing a specific nucleotide polymorphism at the 3' end of the primer to permit preferential amplification of the allele from ALHF. Each experiment was repeated three times with different mRNAs, and the PCR products were sequenced at least once for each gene.

4.3.4 Permethrin induction experiment

Topical applications were performed by dropping 0.5-µl of permethrin solution (10 ug/fly, dissolved in acetone) onto the thoracic notum of 2-day old female flies (Zhu et al. 2008b). The surviving flies were collected for RNA extraction after 6, 12, 24, 48 h exposure to permethrin. Control house flies were exposed to acetone only and collected at the same time points as their permethrin treatment counterparts. All tests were replicated at least three times.

4.3.5 Transgenic expression of candidate P450 genes in *Drosophila melanogaster* and toxicity of permethrin to the transgenic lines

The full length of four up-regulated P450 genes were amplified from cDNA of ALHF house fly using Platinum Taq DNA polymerase High Fidelity (Invitrogen) with specific primer pairs (Table S4.1) based on the 5' and 3' end sequences of the genes. PCR products were purified using a QIAquick Gel Extraction Kit (Qiagen). The purified PCR products were ligated into pCR 2.1 vector using the Original TA Cloning kit (Invitrogen) as described by the manufacturer. The full lengths of P450 genes were cloned in One Shot TOPO 10F' cells using the One Shot TOP10F' Chemically Competent *E. coli* kit (Invitrogen). Cloning and sequence analyses were repeated at least three times and three TA clones from each replication were verified by sequencing. The clones were then sub-cloned into a pUASTattB vector (a gift from Dr. Johannes Bischof, University of Zurich). The plasmid of each pUASTattB-up-regulated P450 gene was transformed into the germ line of the M{vas-int.Dm}ZH-2A, M{3xP3-RFP.attP'}ZH-58A strain of *D. melanogaster* (Bloomington stock #24484), resulting in site specific integration on chromosome 2R (Rainbow Transgenic Flies Inc.). Flies were then reciprocally-crossed against a W^{1118} strain to obtain a transgenic line with the orange eye phenotype, then balanced against the *D. melanogaster* balancer strain w[1118]/Dp(1;Y)y[+]; sna[Sco]/CyO, P{ry[+t7.2]=sevRas1.V12}FK1 (Bloomington stock #6312) to generate a homozygous

line containing the Cytochrome P450 transgene. The insertion of the up-regulated P450 genes in the transgenic fly lines were further confirmed using RT-PCR. The homozygous lines were crossed with the GAL4-expressing *D. melanogaster* strain P{Act5C-GAL4}17bFO1 (Bloomington stock #3954) which expresses GAL4 under control of the Act5C promoter, resulting in ubiquitous non-tissue-specific expression. The F1 generation of these crosses expressed GAL4 and contained a single copy of the Cytochrome P450 transgene which was under control of the UAS enhancer. The expression of the transgenes in transgenic *Drosophila* flies was confirmed using qRT-RCR. The ribosomal protein L11 (RPL11) of *D. melanogaster* was used as an endogenous control to normalize expression of target genes.

Permethrin toxicity bioassays were then conducted on 3-day posteclosion female *D. melanogaster* of F1 UAS-GAL4 crosses to examine the toxicity of permethrin to transgenic flies. Briefly, a serial concentrations of permethrin solution in acetone, ranging from 3 ng/µL to 150 ng/µL that gave >0 and <100% mortality to the tested flies were prepared, two hundred microliter of each permethrin solution were evenly coated on the inside of individual 20 mL glass scintillation vials. Twenty female flies were transferred to each of the prepared vials, which were plugged with cotton balls soaked with 15%

sucrose. The vials for the control groups were coated with acetone alone and plugged with identical 15% sucrose-soaked cotton balls. The mortality was scored after 24 h exposure to permethrin. Flies that did not move were scored as dead. Each bioassay was independently replicated three times using only flies that exhibited the correct morphological marker eyes. The *D. melanogaster* line (Bollmington stock #24484), which containing the empty pUAST vector donated insert, but no transgene from *M. domestica* were used as the control reference line. All tests were run at 25±2^{o}. Bioassay data were pooled and probit analysis was conducted. Significant difference in the resistance levels of the *D. melanogaster* lines were determined based on non-overlap of 95% confidence intervals (CI). All *D. melanogaster* were reared on Jazz-Mix *D. melanogaster* food (Fisher Scientific, Kansas City, MO) at 25 ± 2 ^{o}C under a photoperiod of 12:12 (L:D) h.

4.3.6 Homology modeling and permethrin docking

Structural modeling was performed by the I-TASSER server with the combined methods (Roy et al. 2010; Zhang 2008). Five models were predicted by the I-TASSER for each P450. The top scoring model was submitted to the FG-MD server for fragment

guided molecular dynamics structure refinement (Zhang et al. 2011). Model quality was controlled by Ramachandran plots generated with Procheck (http://services.mbi.ucla.edu/SAVES/) (Laskowski et al. 1993) and ProSA-web (https://prosa.services.came.sbg.ac.at/prosa.php) (Sippl 1993; Wiederstein and Sippl 2007). Channels of P450s were calculated by CAVER 3.0 (http://caver.cz/index.php) (Chovancova et al. 2012; Medek et al. 2007), the passage of a sphere of maximal radius greater than 1.2 Å were considered tabulated, and named according to nomenclature of Cojocaru et al. (Cojocaru et al. 2007). The volume of the substrate binding cavity was characterized by VOIDOO with a 1.4 Å probe (Kleywegt et al. 2001). Proteins and ligands were prepared for docking with Autodock Tools v1.5.6 (http://mgltools.scripps.edu/downloads). Molecular docking was performed by Autodock 4.2. (Morris et al. 2009). Ligand permethrin structures were retrieved from the ZINC database (Irwin et al. 2012). For all dockings, a search space with a grid box of 60 x 60 x 60 Å, centered at the heme iron was set corresponding to substrate recognition sites (SRSs) following those of the CYP2 family proposed by Gothoh (Gotoh 1992). The figures were produced by Pymol (http://www.pymol.org/) (DeLano 2002).

4.4 Results

4.4.1 Cytochrome P450 genes in *M. domestica*

To characterize the P450 genes that play important roles in the pyrethroid resistance of house flies, 86 cytochrome P450 genes, whose expression can be detected in ALHF house flies based on our whole transcriptome analysis of ALHF *M. domestica* (Li et al. 2013), were selected to conduct the current study. These genes were distributed into four major clans of CYP2, CYP3, CYP4, and mitochondrial (Figure 4.1 and S4.2). Of the 86 P450 genes, the majority of them assemble in clans 3 and 4: 41 P450 genes were found in clan 3, with 29 in family CYP6, 6 in family CYP9 family, 3 in family CYP28, and 1 each in families CYP310 and CYP317. Thirty-one P450 genes were in clan 4, with 26 in family 4, 3 in family 313, and 1 each in families 311 and 318. Seven P450 genes were found in clan 2, in the families of 304, 305, 306 and 18. The remaining 13 P450 genes were found in the mitochondrial clan with 5 families of 12, 301, 302, 314 and 315. These detectable P450 genes in ALHF house flies showed a clear expansion in families 4, 6 and 9 compared to other families. P450 genes in these three families have been implicated in environmental response, xenobiotics metabolism and resistance in insect. This expansion

may provide a clue to study the ecological and/or physiological environments adaption strategy of ALHF house flies.

4.4.2 Expression profile of P450 genes in the insecticide-resistant (ALHF) and -susceptible (aabys and CS) *M. domestica* strains

The relative expression profile of 86 P450 genes was examined in ALHF, aabys and CS house fly strains. The expression level of 12 P450 genes was found significantly ($P < 0.05$) up-regulated in the ALHF house flies compared to aabys and CS house flies. The up-regulated levels were from 1.9-fold to 274.4-fold (Figure 4.2, Table S4.3). These genes were distributed into two clans (clan 3 and 4) with 5 genes in family 4, 6 in family 6 and 1 in family 9. Meanwhile, 14 P450 genes were found significantly down-regulated in the ALHF strain compared to both aabys and CS strain. The down-regulated levels were from 1.43-fold to 16.7-fold. For the remaining 59 P450 genes, no significant difference was observed between the transcriptional level in the ALHF strain and that in the aabys and CS strains.

4.4.3 Autosome co-regulation in up-regulation P450 gene expression in insecticide-resistant house flies

We next examined the autosomal linkage of factors from different autosomes to determine the effects of the co-regulation on the expression of the up-regulated P450 genes among the five house fly homozygous lines of A2345, A1345, A1245, A1235 and A1234. Analyzing the gene expression changes resulting from autosome replacement in ALHF house flies enabled us to evaluate the role that genes or factors on each autosome play in P450 gene overexpression in ALHF house flies. The results showed that no significant change in the level of expression was observed for these genes when autosome 4 of ALHF was replaced with that from aabys (i.e., line A1235) (Table 4.2). Suggesting none of the up-regulated P450 genes in ALHF house flies was regulated by factors on autosome 4. In addition, apart from the *CYP6A52* and *CYP9F7*, whose expression level were regulated solely by factor(s) on one autosome only, the expression of the rest of the up-regulated P450 genes were all linked to factors on multiple autosomes (Table 4.2). This result suggests that factors on different autosomes are capable of co-regulation of the P450 genes. This was most commonly observed for autosomes 1, 2 and 5, with eight genes regulated by factors on autosome 1, ten genes

were up-regulated by factors on autosome 2 and seven genes were regulated by factors on autosome 5. Furthermore, *CYP4G99* was regulated by the interaction of factors on autosomes 1 and 3; *CYP4G13* and *CYP4S24* were regulated by the interaction of factors on autosomes 1, 2 and 3, suggesting that factors on autosome 3, besides those on autosomes 1, 2 and 5, were also involved in the expression regulation of up-regulated P450 genes in ALHF house flies (Table 4.2).

To better understand the *cis/trans* regulation of the up-regulated P450 genes in the ALHF house flies, autosomal location analyses were conducted for these up-regulated P450 genes. Our results showed the ALHF allele-specific primer sets for *CYP6D3* and *CYP6D10* amplified specific DNA fragments only in flies with the autosome 1 wild-type marker from ALHF (Figure 4.3), demonstrating that these two P450 genes were on autosome 1. Whereas, *CYP9F7* was located on autosome 2, *CYP4G13*, *CYP4G99* and *CYP4S24* were located on autosome 3. *CYP4E10, CYP4E11, CYP6A36, CYP6A40, CYP6A52* and *CYP6A58* were on autosome 5. Taken together, these results strongly suggested that the expression of the up-regulated P450 genes in ALHF house flies was regulated by *cis* or *trans* regulatory factors/genes.

4.4.4 Response of P450 genes to permethrin challenge in ALHF house flies

Our results showed the expression of 16 P450 genes was significantly induced by permethrin at varying levels and time ranges compared to the control house flies (Figure 4.4, Table S4.4), with 3 in family 4, 9 in family 6, 1 in family 9 and 2 in family 12. Meanwhile, the expression of 23 P450 genes was significantly decreased at least one of the investigation times after treatment with permethrin (Table S4.4), suggesting potential active regulation of P450 gene expression and contribution of these genes to the permethrin resistance in ALHF house flies. The expression level of the rest of P450 genes showed no significant difference between permethrin treated and non-treated house flies.

4.4.5 Functional study of four up-regulated P450 genes in the transgenic *D. melanogaster*

To further characterize the function of the up-regulated P450 genes in ALHF house flies, four P450 genes *CYP4S24, CYP6A36, CYP6A52* and *CYP6D10,* whose expression was not only constitutively up-regulated in ALHF strain compared to aabys and CS strains, but also can be induced by permethrin, were selected to conduct transgenic study by using the GAL4-UAS enhancer trap system of *D. melanogaster*. The P450 gene,

CYP4S24, was located on autosome 3 and expression was regulated by factors on autosomes 1, 2 and 3. *CYP6A36* located on autosome 5 and regulated by factors on autosomes 1, 2 and 5. *CYP6A52* was regulated solely by factors on autosome 5, where the gene itself was located. *CYP6D10* was mapped on autosome 1 and regulated by factors on autosomes 1 and 2 (Figure 4.3, Table 4.2).

We first determined the presence of the house fly P450 genes in transgenic lines of *D. melanogaster* by RT-PCR, the result showed that all the transgenic lines obtained target transgenes (Figure S4.1). To test whether the GAL4-UAS expressing system of *D. melanogaster* can increase the expression of target house fly P450 genes, qRT-PCR was employed to detect the expression level of three lines: Control (Bloomington stock #24484 *D. melanogaster* line containing the empty pUAST vector), GAL4 (ubiquitous Act5C driver line) and Control + GAL4 (the F1 progeny from the cross between the control females and the GAL4 males), which were not transformed with the house fly P450 gene recombinant plasmid, and two transgene lines, P450 (P450 homozygous transgene line) and P450+GAL4 (the F1 progeny from the cross between the P450 homozygous transgene line females and the GAL4 males). Our results showed that the expression of four P450 genes was not detected in the non-transgenic *D. melanogaster*

lines (Control, GAL4 and Control+GAL4), but was detected in transgenic lines (P450 and P450+GAL4). The expression of these P450 genes was enhanced by the GAL4-UAS expressing system of *D. melanogaster* (P450+GAL4 > P450) (Figure 4.5). In addition, there was no significant difference in P450 gene expression observed among different transgenic *Drosophila* P450 homozygous lines or P450+GAL4 lines.

We next characterized the sensitivities of the non-transgenic and transgenic *D. melanogaster* lines to permethrin. The bioassay results showed that there was no significant difference of permethrin toxicity among three non-transgenic lines based on the overlapping 95% confidence intervals (Figure 4.6). However, after the house fly P450 genes were expressed in *D. melanogaster* respectively, the permethrin resistance level was significantly increased compared to non-transgenic lines (Figure 4.5 and 4.6). These results indicated that these P450 genes of house flies are capable of conferring permethrin tolerance in *D. melanogaster*, suggesting that these up-regulated P450 genes play important roles in permethrin resistance of *M. domestica*. In addition, CYP6A36 conferred the highest permethrin tolerance ability in the transgenic flies among the four P450s, followed by CYP6A52, CYP6D10 and CYP4S24.

4.4.6 Homology modeling and permethrin docking

Homology modeling and permethrin docking studies were conducted to investigate the potential permethrin metabolic detoxification differences of selected P450s. The deduced amino acid sequence of four P450 genes from ALHF and aabys house flies were aligned respectively. Sequences contained several conserved P450 characteristics, such as hydrophobic N-terminal domain that acts as a transmembrane anchor, heme-binding region (FXXGXRXCXG) near the C-terminal and six predicted substrate recognition sites (SRSs) (Figure S4.2, S4.3, S4.4 and S4.5). In total, zero, four, thirteen and nine amino acid variations, generated by non-synonymous single nucleotide polymorphisms (SNPs) were detected in *CYP4S24*, *CYP6A36*, *CYP6A52* and *CYP6D10* respectively. The tertiary structure of P450s showed that basic P450 folds were detected in all P450 models, such as helices A to L, commencing from the N terminus, helices F', G' and F/G loop putatively making contact with the membrane region, and cysteine-pocket attaching heme (Figure 4.7).

The active site of P450 is buried deep within the enzyme structure. It is connected to the surrounding environment by a network of channels, which serve as access/egress paths and may determine the substrate specificity of P450s (Cojocaru et al. 2007;

Otyepka et al. 2012). The most probable candidate substrate access/egress channels (solvent and family 2) were investigated in our study. Seven channels (2a, 2ac, 2b, 2c, 2d, 2f and S) were detected in CYP4S24, and eight channels (2a, 2ac, 2b, 2c, 2e 2d, 2f and S) were discovered in CYP6A36, CYP6A52 and CYP6D10 (Figure 4.8, Table S4.5). In addition, the geometry analysis revealed that channel 2c has the largest bottleneck radii among all channels in CYP6A36, CYP6A52 and CYP6D10. Channel 2a has the largest bottleneck radii among all channels in CYP4S24. Furthermore, the analysis showed that CYP6A36, CYP6A52 and CYP6D10 have large openings to the heme prosthetic group and large volumes of active cavity, while the CYP4S24 has a restrained narrow opening to the heme prosthetic group and a smaller active cavity, attributed to the protrusion of two amino acid (Glutamate 313 and Threonine 317) of I helix structure (Figure 4.8, table S6).

Our permethrin docking study showed that there were three predicted permethrin binding conformations in CYP6A36, CYP6A52 and CYP6D10 (Figure 4.9 A, B, C), corresponding with 3 predicted metabolic sites of permethrin (Figure S4.6). All permethrin binding models presented favorable binding affinity and putative permethrin hydroxylation sites directly over the heme iron within 6.0 Å (Feenstra et al. 2007;

Vasanthanathan et al. 2008; Zhu et al. 2013) (Figure 4.9, Table S4.6). These metabolic sites of pyrethroid have been experimentally verified or predicted in several studies (Boonsuepsakul et al. 2008; Lertkiatmongkol et al. 2011; Stevenson et al. 2011; Zhu et al. 2013). These results indicated that CYP6A36, CYP6A52 and CYP6D10 have the potential ability to degrade permethrin and generate multiple metabolites. The short distance from the substrate metabolic sites to P450 heme iron corresponds to the high electrophilic character of this position (Karunker et al. 2009) and high catalytic reactivity (Zimmer et al. 2013). The shortest distance from the metabolic sites to heme iron and the highest permehtrin binding affinity were observed in the CYP6A36 binding models (Table S4.6), implying stronger nucleophilic and electrophilic attacks (Coon et al. 1998), and permethrin metabolism ability of CYP6A36. These results were consistent with our bioassay results that CYP6A36 provided the highest permethrin tolerance ability to transgenic flies among the four P450s. Our study also revealed a single permethrin binding model in CYP4S24 (Figure 4.9 D); however, the catalytic site of permethrin was far away from the heme iron (7.41 Å), indicating CYP4S24 may not metabolize permethrin.

Sequence analysis of *CYP6A36* revealed that there were four amino acid variations between ALHF and aabys house flies (Figure 4.7, Figure S4.3). The geometry analysis and permethrin docking study indicated that the backbone of protein structure, the active site volume, permethrin binding energy and the distance from the permethrin putative hydroxylation sites to the heme iron of CYP6A36 in ALHF house flies were same as that in the aabys house flies (Table S4.6). These observations strongly suggested that the involvement of CYP6A36 and CYP4S24 (no amino acid mutation) in permethrin resistance of ALHF house flies is mainly from the increased expression of these genes compared to aabys house flies. On the contrary, although the protein backbones of CYP6A52 and CYP6D10 in aabys house flies were very similar to that in ALHF house flies (root-mean-square deviation was 0.285 Å and 0.373 Å, respectively), larger active site cavities, higher binding affinity and shorter distance from the permethrin metabolic sites to P450 heme iron of these two enzyme were observed in the ALHF house flies compared to that in aabys house flies (Table S4.6). These results indicated that the amino acid differences of CYP6A52 and CYP6D10 between ALHF and aabys house flies influenced the interactions between permethrin and P450s, suggesting the involvement of the two P450 enzymes in permethrin resistance of ALHF house flies was not only from

the increased expression of the genes, but also from the more efficient form of the enzymes compared to aabys house flies.

Residues that contacted to permethrin in the active sites of P450s were examined. ILE122, ASN216, ALA311, THR315 and ILE377 of CYP6A36; ARG105, ILE214, GLU215, ALA309 and ILE490 of CYP6A52; ARG99, LYS211 and VAL319 of CYP6D10 were conserved in permethrin binding models of each P450 enzyme, respectively (Table S4.5), suggesting the importance of these residues for permethrin binding.

4.5 Discussion

Previous studies showed that permethrin resistance in house flies could be largely suppressed by PBO, indicating P450s are the primary enzymes involved in detoxifying permethrin and conferring permethrin resistance in house flies (Carino et al. 1992; Carino et al. 1994; Feyereisen 2006; Liu and Yue 2000; Liu and Yue 2001; Zhu et al. 2008a; Zhu et al. 2008b; Zhu and Liu 2008). An expansion of these P450 gene members in families 4, 6 and 9, which groups were most commonly involved in xenobiotics metabolism and resistance in insects (Feyereisen 2005; Li et al. 2007; Pavlidi et al. 2012; Yang et al.

2008). This expansion implied a clue to study the ecological and/or physiological environments adaption strategy of ALHF house flies, such as the house fly may have benefited from this expansion by obtaining ability to metabolize insecticides.

Up-regulation of P450 gene expression led to increased levels of total P450s and P450 activities that can significantly affect the disposition of xenobiotics or endogenous compounds in the tissues of organisms and alter their toxicological effects (Feyereisen 2006; Pavek and Dvorak 2008). Many studies have suggested that both constitutively increased expression and induction of P450s in insects are responsible for the insecticide resistance (Gong et al. 2013; Liu et al. 2011; Pavek and Dvorak 2008; Scharf et al. 2001; Yang and Liu 2011; Zhu et al. 2008b). We therefore expected that the P450 genes, which were involved in permethrin resistance would be highly expressed or could be induced by permethrin in ALHF house flies. The approaches to this study, which compared P450 gene expression profile in different house fly strains and permethrin induction, revealed that these up-regulated or inducible P450 genes were also mainly distributed in families 4, 6 and 9, suggesting the importance of these P450 genes in permethrin resistance in ALHF house flies. Interestingly, four P450 genes, *CYP4S24*, *CYP6A36*, *CYP6A52* and *CYP6D10*, whose expression was not only constitutively overexpressed, but also could be

induced by permethrin in the ALHF house flies, were observed in our study. The concurrence of the constitutive overexpression and induction of these four genes in the ALHF strain strongly suggested that these genes play important roles in permethrin resistance in the ALHF house flies.

P450 gene down-regulation has been reported in several insect species, such as *C. quinquefasciatus, Anopheles gambiae, Spodoptera littoralis* and *D. melanogaster* (Davies et al. 2006; Marinotti et al. 2005; Yang and Liu 2011). In the current study, we also found several P450 genes were down-regulated in ALHF house flies compared to both aabys and CS house flies, or the expression was repressed by permethrin. These results may not only reveal equally dynamic changes in abundance for both the increased and decreased P450 gene expression in resistant house flies, but also indicate an important feature of the gene regulation system of the house fly in response to environmental changes (Li et al. 2013). Several hypotheses have been proposed for the harmonizing of up- and down-regulation, e.g., homeostatic responses for protecting the cell from the harmful effects of oxidizing species, nitric oxide, or arachidonic acid metabolites from catalytic or metabolic enzymes (Morgan 2001); homeostatic responses to provocative processes; and/or an essential need for the tissue to utilize its transcriptional machinery and energy

for the synthesis of other components involved in the inflammatory response (Li et al. 2013; Morgan 1989; Reid et al. 2012). Whether the down-regulated P450 genes identified in the ALHF house flies reflects a regulation feature or homeostatic response of house flies to insecticides needs further study.

Examination of the autosomal linkage of overexpressed P450 genes is an important step in establishing a causal link between P450 genes and their roles in insecticide resistance (Carino et al. 1992; Liu and Scott 1996; Maitra et al. 2000). The current genetic mapping of the up-regulated P450 genes in the ALHF house flies showed that these genes were on autosome 1, 2, 3 and 5. Gene co-regulation analysis revealed that the up-regulated P450 gene expression in ALHF house flies also occurred among autosomes 1, 2, 3 and 5. Interestingly, these results were consistent with our previous synergism studies that factors/genes on autosome 1, 2, 3 and 5 play very important roles in insecticide resistance of the ALHF house fly (Li et al. 2013; Liu and Scott 1995; Liu and Yue 2001; Tian et al. 2011), indicating the importance of these P450 genes in pyrethroid resistance of ALHF house flies. The expression of the up-regulated P450 genes were linked to regulatory factors/genes on multiple autosomes except for the *CYP6A52* and *CYP9F7*, whose expression was regulated solely by factor(s) on one autosome that they

were located on also. These results strongly indicated a common feature of up-regulated P450 genes were that they were regulated by *trans*- and/or *cis*-acting factors in ALHF house flies. Characterizing co-regulated P450 genes will represent a good starting point for characterizing the transcriptional regulatory network and pathways in house flies.

The model insect, *D. melanogaster*, is a useful tool for functionally studying the role of metabolic enzymes in conferring metabolism-based insecticide resistance of insects (Daborn et al. 2012). The function of several P450 genes has been studied through the transgenic expression of P450 genes in *D. melanogaster*, such as *CYP6BQ9* from *Tribolium castaneum* (Zhu et al. 2010), *CYP6P9a* and *CYP6P9b* from *Anopheles funestus* (Riveron et al. 2013b), and *CYP6CM1* from *Bemisia tabaci* (Daborn et al. 2012). The important roles played by *CYP4S24*, *CYP6A36*, *CYP6A52* and *CYP6D10* in permethrin resistance of ALHF house flies were well supported by the transgenic expression of these genes in *D. melanogaster* demonstrating these genes are able to confer permethrin resistance in house flies. Interestingly, although no significant difference of P450 gene expression was detected among different transgenic *Drosophila* lines (P450+GAL4), different permethrin resistance levels were observed, indicating the contributions of these P450s to permethrin resistance in ALHF house flies were different.

SRSs are the most variable regions in P450s, with SRS1, SRS4, SRS5 and SRS6 involved in the catalytic site formation while SRS2 and SRS3 are related to the substrate access channel formation (Chiu et al. 2008; Fishelovitch et al. 2009; Schuler and Berenbaum 2013). Differences in the substrate channels and active sites were responsible for the differences in binding to pyrethroids (Lertkiatmongkol et al. 2011). Several modeling and docking studies, such as CYP6AA3 of *A. minumus*, CYP6M2 of *A. gambiae* have predicted that pyrethroid insecticides metabolism of these enzymes can produce multiple metabolites due to different catalyze sites, and these predictions have been proven in the corresponding *in vitro* metabolism studies (Boonsuepsakul et al. 2008; Lertkiatmongkol et al. 2011; Stevenson et al. 2011). In our study, multiple permethrin binding models were also detected in CYP6A36, CYP6A52 and CYP6D10 due to the huge active site cavity, strongly indicating multiple permethrin metabolites can be produced by these enzymes. In addition, the binding models showed a shorter distance from the metabolic sites of permethrin to the heme iron and higher permethrin binding affinity in CYP6A36 compared to CYP6A52 and CYP6D10, implying stronger nucleophilic and electrophilic attacks and permethrin metabolism ability of CYP6A36, perhaps explaining why CYP6A36 provided the highest permethrin tolerance ability in

transgenic flies among the four P450s. The increased gene expression of *CYP6Z2* was shown to elevate the pyrethroid resistance in *An. gambiae* (McLaughlin et al. 2008; Nikou et al. 2003), and modeling study showed that CYP6Z2 can bind with permethrin but lacks the ability to metabolize permethrin because of structural constraints within the active site (McLaughlin et al. 2008). The current transgenic *D. melanogaster* study showed that increased expression of CYP4S24 was capable of conferring permethrin tolerance in *D. melanogaster*. However, compared to CYP6A36, CYP6A52 and CYP6D10, CYP4S24 has a narrow opening to heme iron, which prevented permethrin access to the iron, resulting in the catalytic site of permethrin being far away from the heme-oxygen (more than 6.0 Å) , this may cause the absence of the metabolic detoxification activity toward permethrin, suggesting that metabolic activity differences among these four P450 enzymes in house flies come from active site differences allowing different binding models to form. Beside detoxification, P450s also have other functions, such as hormone and pigment biosynthesis (Scott 2008). Interestingly, P450s in the CYP4G family, such as CYP4G2 in the house fly have been reported to produce hydrocarbons from aldehydes (Qiu et al. 2012), which are involved in cuticle formation. Decreased insecticide penetration due to the increased cuticle thickness has been reported

to be involved in insecticide resistance (Koganemaru et al. 2013; McKenzie and Batterham 1994; Strycharz et al. 2013). To determine whether the contribution of CYP4S24 to permethrin resistance in house flies was related to cuticle formation or storage/sequestration, similar to carboxyl esterase E4 of *Myzus persicae* (Devonshire and Moores 1982), more studies are needed.

The active site of P450 is buried within the protein structure. To arrive at the active site, substrates will undergo chemical modification and channel conduction (Kirchmair et al. 2012). In our study, the most probable substrate access/egress channels (solvent and family 2 channels) of CYP4S24, CYP6A36, CYP6A52 and CYP6D10 were investigated. Interestingly, channel 2e was not detected in CYP4S24, and this enzyme showed the lowest contribution to permethrin resistance among the four P450s based on transgenic and permethrin docking studies. To determine whether the absence of channel 2e is responsible for the low activity of CYP4S24 on permethrin, more studies are needed. The channel 2f always pointed toward the lipid interior, which may serve as common access channels for lipophilic substrates. The solvent channel always pointed into the water-membrane interface and served as an entrance/release for hydrophilic substrates. However, the position of substrate access channels and residues of channel openings were

contributed to substrate preferences (Berka et al. 2013; Kirchmair et al. 2012), the bottleneck diameter also fluctuates due to the flexibility of P450s (Hendrychová et al. 2011; Hendrychova et al. 2012; Skopalík et al. 2008), and different P450s may use different channels to conduct substrates because of their structural differences (Cojocaru et al. 2007; Otyepka et al. 2012). Without further deep gating mechanisms and molecular dynamics simulation studies, it is still unknown which channels are involved in permethrin access and corresponding metabolites egress from the active site of these P450s.

Enzymatic activities of P450s are affected by residue mutations. Mutations in catalytic site and substrate access channels were linked to the range of substrate metabolized, the rate of substrate entry/exit and metabolism directly (Hiratsuka 2011; Sansen et al. 2007; Wen et al. 2005; Zhao and Halpert 2007). The mutations in the proximal surface of helix C and the C-D loop region related to the interactions between cytochrome P450 and cytochrome b_5 or cytochrome P450 reductase (Ahuja et al. 2013; Bridges et al. 1998; Im and Waskell 2011; Skopalík et al. 2008) can affect electron conduction and kinetic properties due to alterations in interaction between P450 and their partners (Kenaan et al. 2011; Schenkman and Jansson 1999; Schuler and Berenbaum

2013; Zhang et al. 2005). Many studies have also shown that mutations outside of the active site changed the active site cavity or binding affinity of P450s, influencing the metabolism activity of P450s (Kumar et al. 2005; Kumar et al. 2006; Tiong et al. 2014; Wilderman et al. 2012). Sequence analysis of CYP6A36 revealed that there were four amino acid variations in ALHF house flies compared with aabys house flies. The geometry analysis and permethrin docking study indicated that the involvement of CYP6A36 and CYP4S24 in permethrin resistance of ALHF house flies compared to aabys house flies is mainly from the increased expression of these genes in ALHF house flies. However, the contribution of CYP6A52 and CYP6D10 to the permethrin resistance of ALHF house flies is not only from the increased expression of gene but also from the more efficient form of enzymes. Many studies also showed that mutations located outside of the active site of P450 can enhance P450 expression or stability (Kumar et al. 2007; Talakad et al. 2010). To determine whether the mutations that were identified in our study relate to P450 expression or stability, further studies are needed.

Taken together, our findings suggest that P450 constitutive over-expression or induction is responsible for permethrin resistance in house flies, and these P450 genes are regulated by *trans*- and/or *cis*-acting factors. Functional characterization of house fly

P450s in the transgenic *D. melanogaster*, modeling and permethrin docking analyses have provided important information on the pivotal roles of CYP4S24, CYP6A36, CYP6A52 and CYP6D10 in permethrin resistance of house flies, and opened the avenue for P450 *in vitro* study. In addition, our results confirm and support the co-overexpression of multiple P450 genes as likely to be key factors enhancing permethrin resistance in house flies. Furthermore, our study indicates that besides detoxification, other roles of P450 may be involved in permethrin resistance in house flies. Future studies of identification and characterization of the promoter regions and regulatory factors that are involved in the P450 gene expression regulation will shed new light on the molecular basis of insecticide resistance in house flies or even in other insect species.

4.6 Reference

Adams, MD, Celniker, SE, Holt, RA, Evans, CA, Gocayne, JD, Amanatides, PG, Scherer, SE, Li, PW, Hoskins, RA and Galle, RF (2000) The genome sequence of *Drosophila melanogaster*. Science 287: 2185-2195.

Ahmad, A, Nagaraja, T and Zurek, L (2007) Transmission of *Escherichia coli* O157: H7 to cattle by house flies. Preventive Veterinary Medicine 80: 74-81.

Ahuja, S, Jahr, N, Im, S-C, Vivekanandan, S, Popovych, N, Le Clair, SV, Huang, R, Soong, R, Xu, J and Yamamoto, K (2013) A model of the membrane-bound cytochrome b5-cytochrome P450 complex from NMR and mutagenesis data. Journal of Biological Chemistry 288: 22080-22095.

Anspaugh, DD, Rose, RL, Koehler, PG, Hodgson, E and Roe, RM (1994) Multiple Mechanisms of Pyrethroid Resistance in the German Cockroach, *Blattella germanica* (L.). Pesticide Biochemistry and Physiology 50: 138-148.

Berka, K, Paloncýová, Mt, Anzenbacher, P and Otyepka, M (2013) Behavior of Human Cytochromes P450 on Lipid Membranes. The Journal of Physical Chemistry B 117: 11556-11564.

Boonsuepsakul, S, Luepromchai, E and Rongnoparut, P (2008) Characterization of Anopheles minimus CYP6AA3 expressed in a recombinant baculovirus system. Archives of Insect Biochemistry and Physiology 69: 13-21.

Brévault, T, Achaleke, J, Sougnabé, S and Vaissayre, M (2008) Tracking pyrethroid resistance in the polyphagous bollworm, *Helicoverpa armigera* (Lepidoptera: Noctuidae), in the shifting landscape of a cotton-growing area. Bulletin of Entomological Research 98: 565-573.

Bridges, A, Gruenke, L, Chang, Y-T, Vakser, IA, Loew, G and Waskell, L (1998) Identification of the binding site on cytochrome P450 2B4 for cytochrome b5 and cytochrome P450 reductase. Journal of Biological Chemistry 273: 17036-17049.

Carino, F, Koener, J, Plapp Jr, F and Feyereisen, R (1992) Expression of the cytochrome P450 gene CYP6A1 in the housefly, *Musca domestica*. In: Book Expression of the cytochrome P450 gene CYP6A1 in the housefly, *Musca domestica* (Editor, ed.eds.). Vol., pp., City.

Carino, FA, Koener, JF, Plapp, FW, Jr. and Feyereisen, R (1994) Constitutive overexpression of the cytochrome P450 gene CYP6A1 in a house fly strain with metabolic resistance to insecticides. Insect Biochemistry and Molecular Biology 24: 411-418.

Carvalho, RA, Omoto, C, Field, LM, Williamson, MS and Bass, C (2013) Investigating the Molecular Mechanisms of Organophosphate and Pyrethroid Resistance in the Fall Armyworm Spodoptera frugiperda. PloS One 8: e62268.

Chiu, T-L, Wen, Z, Rupasinghe, SG and Schuler, MA (2008) Comparative molecular modeling of *Anopheles gambiae* CYP6Z1, a mosquito P450 capable of metabolizing DDT. Proceedings of the National Academy of Sciences 105: 8855-8860.

Chovancova, E, Pavelka, A, Benes, P, Strnad, O, Brezovsky, J, Kozlikova, B, Gora, A, Sustr, V, Klvana, M and Medek, P (2012) CAVER 3.0: a tool for the analysis of transport pathways in dynamic protein structures. PLoS Computational Biology 8: e1002708.

Cojocaru, V, Winn, PJ and Wade, RC (2007) The ins and outs of cytochrome P450s. Biochimica et Biophysica Acta (BBA)-General Subjects 1770: 390-401.

Coon, MJ, Vaz, AD, McGinnity, DF and Peng, H-M (1998) Multiple Activated Oxygen Species in P450 Catalysis Contributions to Specificity in Drug Metabolism. Drug Metabolism and Disposition 26: 1190-1193.

Daborn, PJ, Lumb, C, Harrop, TW, Blasetti, A, Pasricha, S, Morin, S, Mitchell, SN, Donnelly, M, Müller, P and Batterham, P (2012) Using *Drosophila melanogaster* to validate metabolism-based insecticide resistance from insect pests. Insect Biochemistry and Molecular Biology 42: 918-924.

Davies, L, Williams, D, Aguiar-Santana, I, Pedersen, J, Turner, P and Rees, H (2006) Expression and down-regulation of cytochrome P450 genes of the CYP4 family by ecdysteroid agonists in Spodoptera littoralis and *Drosophila melanogaster*. Insect Biochemistry and Molecular Biology 36: 801-807.

DeLano, WL (2002) The PyMOL molecular graphics system. CA. USA

Devonshire, AL and Moores, GD (1982) A carboxylesterase with broad substrate specificity causes organophosphorus, carbamate and pyrethroid resistance in peach-potato aphids (*Myzus persicae*). Pesticide Biochemistry and Physiology 18: 235-246.

Estabrook, RW (2003) A passion for P450s (remembrances of the early history of research on cytochrome P450). Drug Metabolism and Disposition 31: 1461-1473.

Feenstra, KA, Starikov, EB, Urlacher, VB, Commandeur, JN and Vermeulen, NP (2007) Combining substrate dynamics, binding statistics, and energy barriers to rationalize regioselective hydroxylation of octane and lauric acid by CYP102A1 and mutants. Protein Science 16: 420-431.

Feyereisen, R (2005) Insect cytochrome P450. Protein Science 4: 1-77.

Feyereisen, R (2006) Evolution of insect P450. Biochemical Society Transactions 34: 1252-1255.

Fishelovitch, D, Shaik, S, Wolfson, HJ and Nussinov, R (2009) Theoretical characterization of substrate access/exit channels in the human cytochrome P450 3A4 enzyme: involvement of phenylalanine residues in the gating mechanism. The Journal of Physical Chemistry B 113: 13018-13025.

Gong, Y, Li, T, Zhang, L, Gao, X and Liu, N (2013) Permethrin induction of multiple cytochrome P450 genes in insecticide resistant mosquitoes, *Culex quinquefasciatus*. International Journal of Biological Sciences 9: 863.

Gotoh, O (1992) Substrate recognition sites in cytochrome P450 family 2 (CYP2) proteins inferred from comparative analyses of amino acid and coding nucleotide sequences. Journal of Biological Chemistry 267: 83-90.

Gouet, P, Robert, X and Courcelle, E (2003) ESPript/ENDscript: extracting and rendering sequence and 3D information from atomic structures of proteins. Nucleic Acids Researchearch 31: 3320-3323.

Grant, D and Hammock, B (1992) Genetic and molecular evidence for a trans-acting regulatory locus controlling glutathione S-transferase-2 expression in *Aedes aegypti*. Molecular and General Genetics MGG 234: 169-176.

Greenberg, B (1965) Flies and Disease. Scientific American 213: 92-99.

Hendrychová, T, Anzenbacherová, E, Hudeček, J, Skopalík, J, Lange, R, Hildebrandt, P, Otyepka, M and Anzenbacher, P (2011) Flexibility of human cytochrome P450 enzymes: Molecular dynamics and spectroscopy reveal important function-related variations. Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics 1814: 58-68.

Hendrychova, T, Berka, K, Navratilova, V, Anzenbacher, P and Otyepka, M (2012) Dynamics and hydration of the active sites of mammalian cytochromes P450 probed by molecular dynamics simulations. Current Drug Metabolism 13: 177-189.

Hiratsuka, M (2011) In vitro assessment of the allelic variants of cytochrome P450. Drug Metabolism and Pharmacokinetics 27: 68-84.

Hiroyosh, T (1960) Some New Mutants and Linkage Groups of the House Fly1. Journal of Economic Entomology 53: 985-990.

Hiroyosh, T (1977) Some new mutants and revised linkage maps of the housefly, *Musca domestica* L. Japanese Journal of Genetics 52: 275-288.

Im, S and Waskell, L (2011) The interaction of microsomal cytochrome P450 2B4 with its redox partners, cytochrome P450 reductase and cytochrome b (5). Archives of Biochemistry and Biophysics 507: 144-153.

Irwin, JJ, Sterling, T, Mysinger, MM, Bolstad, ES and Coleman, RG (2012) ZINC: a free tool to discover chemistry for biology. Journal of Chemical Information and Modeling 52: 1757-1768.

Karunker, I, Morou, E, Nikou, D, Nauen, R, Sertchook, R, Stevenson, B, Paine, M, Morin, S and Vontas, J (2009) Structural model and functional characterization of the

Bemisia tabaci CYP6CM1vQ, a cytochrome P450 associated with high levels of imidacloprid resistance. Insect Biochemistry and Molecular Biology 39: 697-706.

Kaufman, PE, Nunez, SC, Mann, RS, Geden, CJ and Scharf, ME (2010) Nicotinoid and pyrethroid insecticide resistance in houseflies (Diptera: Muscidae) collected from Florida dairies. Pest Mangement Science 66: 290-294.

Keiding, J (1986) House-fly; biology and control. OMS.

Kenaan, C, Zhang, H, Shea, EV and Hollenberg, PF (2011) Uncovering the Role of Hydrophobic Residues in Cytochrome P450-Cytochrome P450 Reductase Interactions. Biochemistry 50: 3957-3967.

Kirchmair, J, Williamson, MJ, Tyzack, JD, Tan, L, Bond, PJ, Bender, A and Glen, RC (2012) Computational prediction of metabolism: sites, products, SAR, P450 enzyme dynamics, and mechanisms. Journal of Chemical Information and Modeling 52: 617-648.

Kleywegt, G, Zou, J-Y, Kjeldgaard, M and Jones, T (2001) Around O. International Tables for Crystallography Volume F: Crystallography of Biological Macromolecules: 353-356.

Klingenberg, M (1958) Pigments of rat liver microsomes. Archives of Biochemistry and Biophysics 75: 376-386.

Koganemaru, R, Miller, DM and Adelman, ZN (2013) Robust cuticular penetration resistance in the common bed bug (*Cimex lectularius*) correlates with increased steady-state transcript levels of CPR-type cuticle protein genes. Pesticide Biochemistry and Physiology 106: 190-197.

Kumar, S, Chen, C, Waxman, D and Halpert, J (2005) Directed evolution of mammalian cytochrome P450 2B1: mutations outside of the active site enhance the metabolism of several substrates, including the anticancer prodrugs cyclophosphamide and ifosfamide. The Journal of Biological Chemistry 280: 19569-19575.

Kumar, S, Sun, L, Liu, H, Muralidhara, B and Halpert, JR (2006) Engineering mammalian cytochrome P450 2B1 by directed evolution for enhanced catalytic tolerance to temperature and dimethyl sulfoxide. Protein Engineering Design and Selection 19: 547-554.

Kumar, S, Zhao, Y, Sun, L, Negi, SS, Halpert, JR and Muralidhara, BK (2007) Rational engineering of human cytochrome P450 2B6 for enhanced expression and stability: importance of a Leu264→ Phe substitution. Molecular Pharmacology 72: 1191-1199.

Laskowski, RA, MacArthur, MW, Moss, DS and Thornton, JM (1993)

PROCHECK: a program to check the stereochemical quality of protein structures. Journal of Applied Crystallography 26: 283-291.

Lertkiatmongkol, P, Jenwitheesuk, E and Rongnoparut, P (2011) Homology modeling of mosquito cytochrome P450 enzymes involved in pyrethroid metabolism: insights into differences in substrate selectivity. BMC Research Notes 4: 321.

Li, M, Reid, WR, Zhang, L, Scott, JG, Gao, X, Kristensen, M and Liu, N (2013) A whole transcriptomal linkage analysis of gene co-regulation in insecticide resistant house flies, *Musca domestica*. BMC Genomics 14: 803.

Li, X, Schuler, MA and Berenbaum, MR (2007) Molecular mechanisms of metabolic resistance to synthetic and natural xenobiotics. Annual Review of Entomology 52: 231-253.

Liu, N, Li, T, Reid, WR, Yang, T and Zhang, L (2011) Multiple Cytochrome P450 genes: their constitutive overexpression and permethrin induction in insecticide resistant mosquitoes, *Culex quinquefasciatus*. PloS One 6: e23403.

Liu, N and Scott, JG (1995) Genetics of resistance to pyrethroid insecticides in the house fly, *Musca domestica*. Pesticide Biochemistry and Physiology 52: 116-124.

Liu, N and Scott, JG (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*. Biochemical Genetics 34: 133-148.

Liu, N, Tomita, T and Scott, JG (1995) Allele-specific PCR reveals that CYP6D1 is on chromosome 1 in the house fly, *Musca domestica*. Cellular and Molecular Life Sciences 51: 164-167.

Liu, N and Yue, X (2000) Insecticide resistance and cross-resistance in the house fly (Diptera: Muscidae). Journal of Economic Entomology 93: 1269-1275.

Liu, N and Yue, X (2001) Genetics of pyrethroid resistance in a strain (ALHF) of house flies (Diptera: Muscidae). Pesticide Biochemistry and Physiology 70: 151-158.

Livak, KJ and Schmittgen, TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25: 402-408.

Maitra, S, Dombrowski, SM, Basu, M, Raustol, O, Waters, LC and Ganguly, R (2000) Factors on the third chromosome affect the level of Cyp6a2 and Cyp6a8 expression in *Drosophila melanogaster*. Gene 248: 147-156.

Maitra, S, Dombrowski, SM, Waters, LC and Ganguly, R (1996) Three second

chromosome-linked clustered Cyp6 genes show differential constitutive and barbital-induced expression in DDT-resistant and susceptible strains of *Drosophila melanogaster*. Gene 180: 165-171.

Marinotti, O, Nguyen, QK, Calvo, E, James, AA and Ribeiro, J (2005) Microarray analysis of genes showing variable expression following a blood meal in *Anopheles gambiae*. Insect Molecular Biology 14: 365-373.

McKenzie, JA and Batterham, P (1994) The genetic, molecular and phenotypic consequences of selection for insecticide resistance. Trends in Ecology & Evolution 9: 166-169.

McLaughlin, L, Niazi, U, Bibby, J, David, JP, Vontas, J, Hemingway, J, Ranson, H, Sutcliffe, M and Paine, M (2008) Characterization of inhibitors and substrates of *Anopheles gambiae* CYP6Z2. Insect Molecular Biology 17: 125-135.

Medek, P, Benes, P and Sochor, J (2007) Computation of Tunnels in Protein Molecules using Delaunay Triangulation. Journal of WSCG 15: 107-114.

Milani, R, Rubini, P and Franco, M (1967) Sex determination in the housefly. Genetica Agraria 21: 385-411.

Morgan, E (1989) Suppression of constitutive cytochrome P-450 gene expression in

livers of rats undergoing an acute phase response to endotoxin. Molecular Pharmacology 36: 699-707.

Morgan, ET (2001) Regulation of cytochrome p450 by inflammatory mediators: why and how? Drug Metabolism and Disposition 29: 207-212.

Morris, GM, Huey, R, Lindstrom, W, Sanner, MF, Belew, RK, Goodsell, DS and Olson, AJ (2009) AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. Journal of Computational Chemistry 30: 2785-2791.

Nene, V, Wortman, JR, Lawson, D, Haas, B, Kodira, C, Tu, ZJ, Loftus, B, Xi, Z, Megy, K and Grabherr, M (2007) Genome sequence of *Aedes aegypti*, a major arbovirus vector. Science 316: 1718-1723.

Nickel, C and Wagoner, D (1974) Mutants on linkage groups 3 and 4 of the house fly. Annals of the Entomological Society of America 67: 775-776.

Nikou, D, Ranson, H and Hemingway, J (2003) An adult-specific CYP6 P450 gene is overexpressed in a pyrethroid-resistant strain of the malaria vector, *Anopheles gambiae*. Gene 318: 91-102.

Otyepka, M, Berka, K and Anzenbacher, P (2012) Is there a relationship between the substrate preferences and structural flexibility of cytochromes P450? Current Drug

Metabolism 13: 130-142.

Pap, L and Tóth, A (1995) Development and characteristics of resistance in the susceptible who/srs house fly (*Musca domestica*) strain subjected to selection with beta-cypermethrin. Pesticide Science 45: 335-349.

Pavek, P and Dvorak, Z (2008) Xenobiotic-induced transcriptional regulation of xenobiotic metabolizing enzymes of the cytochrome P450 superfamily in human extrahepatic tissues. Current Drug Metabolism 9: 129-143.

Pavlidi, N, Monastirioti, M, Daborn, P, Livadaras, I, Van Leeuwen, T and Vontas, J (2012) Transgenic expression of the *Aedes aegypti* CYP9J28 confers pyrethroid resistance in *Drosophila melanogaster*. Pesticide Biochemistry and Physiology 104: 132-135.

Qiu, Y, Tittiger, C, Wicker-Thomas, C, Le Goff, G, Young, S, Wajnberg, E, Fricaux, T, Taquet, N, Blomquist, GJ and Feyereisen, R (2012) An insect-specific P450 oxidative decarbonylase for cuticular hydrocarbon biosynthesis. Proceedings of the National Academy of Sciences 109: 14858-14863.

Rebeiz, M, Jikomes, N, Kassner, VA and Carroll, SB (2011) Evolutionary origin of a novel gene expression pattern through co-option of the latent activities of existing

regulatory sequences. Proceedings of the National Academy of Sciences 108: 10036-10043.

Reid, WR, Zhang, L, Liu, F and Liu, N (2012) The Transcriptome Profile of the Mosquito *Culex quinquefasciatus* following Permethrin Selection. PloS One 7: e47163.

Riveron, JM, Irving, H, Ndula, M, Barnes, KG, Ibrahim, SS, Paine, MJ and Wondji, CS (2013) Directionally selected cytochrome P450 alleles are driving the spread of pyrethroid resistance in the major malaria vector *Anopheles funestus*. Proceedings of the National Academy of Sciences 110: 252-257.

Roush, RT, Hoy, CW, Ferro, DN and Tingey, WM (1990) Insecticide Resistance in the Colorado Potato Beetle (Coleoptera: Chrysomelidae): Influence of Crop Rotation and Insecticide Use. Journal of Economic Entomology 83: 315-319.

Roy, A, Kucukural, A and Zhang, Y (2010) I-TASSER: a unified platform for automated protein structure and function prediction. Nature protocols 5: 725-738.

Sansen, S, Hsu, M-H, Stout, CD and Johnson, EF (2007) Structural insight into the altered substrate specificity of human cytochrome P450 2A6 mutants. Archives of Biochemistry and Biophysics 464: 197-206.

Sasaki, T, Kobayashi, M and Agui, N (2000) Epidemiological potential of excretion

and regurgitation by *Musca domestica* (Diptera: Muscidae) in the dissemination of *Escherichia coli* O157: H7 to food. Journal of Medical Entomology 37: 945-949.

Scharf, M, Parimi, S, Meinke, LJ, Chandler, L and Siegfried, BD (2001) Expression and induction of three family 4 cytochrome P450 (CYP4)* genes identified from insecticide-resistant and susceptible western corn rootworms, Diabrotica virgifera virgifera. Insect Molecular Biology 10: 139-146.

Schenkman, J and Jansson, I (1999) Interactions between cytochrome P450 and cytochrome b5. Drug Metabolism Reviews 31: 351-364.

Schuler, MA and Berenbaum, MR (2013) Structure and Function of Cytochrome P450S in Insect Adaptation to Natural and Synthetic Toxins: Insights Gained from Molecular Modeling. Journal of Chemical Ecology 39: 1232-1245.

Scott, JG (1999) Cytochromes P450 and insecticide resistance. Insect Biochemistry and Molecular Biology 29: 757-777.

Scott, JG (2008) Insect cytochrome P450s: thinking beyond detoxification. Recent advances in insect physiology, Toxicology and Molecular Biology 1: 117-124.

Scott, HG and Littig, KS (1962) Flies of public health importance and their control. US Department of Health, Education, and Welfare, Public Health Service, Communicable

Disease Center.

Scott, JG, Alefantis, TG, Kaufman, PE and Rutz, DA (2000) Insecticide resistance in house flies from caged-layer poultry facilities. Pest Management Science 56: 147-153.

Sippl, MJ (1993) Recognition of errors in three-dimensional structures of proteins. Proteins: Structure, Function, and Bioinformatics 17: 355-362.

Skopalík, J, Anzenbacher, P and Otyepka, M (2008) Flexibility of human cytochromes P450: molecular dynamics reveals differences between CYPs 3A4, 2C9, and 2A6, which correlate with their substrate preferences. The Journal of Physical Chemistry B 112: 8165-8173.

Stevenson, BJ, Bibby, J, Pignatelli, P, Muangnoicharoen, S, O'Neill, PM, Lian, L-Y, Müller, P, Nikou, D, Steven, A, Hemingway, J, Sutcliffe, MJ and Paine, MJI (2011) Cytochrome P450 6M2 from the malaria vector *Anopheles gambiae* metabolizes pyrethroids: Sequential metabolism of deltamethrin revealed. Insect Biochemistry and Molecular Biology 41: 492-502.

Strycharz, JP, Lao, A, Li, H, Qiuc, X, Lee, SH, Sun, W, Yoon, KS, Doherty, JJ, Pittendrigh, BR and Clark, JM (2013) Resistance in the highly DDT-resistant 91-R strain of *Drosophila melanogaster* involves decreased penetration, increased metabolism, and

direct excretion. Pesticide Biochemistry and Physiology 107: 207-217.

Talakad, JC, Wilderman, PR, Davydov, DR, Kumar, S and Halpert, JR (2010) Rational engineering of cytochromes P450 2B6 and 2B11 for enhanced stability: Insights into structural importance of residue 334. Archives of Biochemistry and Biophysics 494: 151-158.

Tian, L, Cao, C, He, L, Li, M, Zhang, L, Liu, H and Liu, N (2011) Autosomal interactions and mechanisms of pyrethroid resistance in house flies, *Musca domestica*. International Journal of Biological Sciences 7: 902-911.

Tiong, KH, Yunus, NAM, Yiap, BC, Tan, EL, Ismail, R and Ong, CE (2014) Inhibitory Potency of 8-Methoxypsoralen on Cytochrome P450 2A6 (CYP2A6) Allelic Variants CYP2A6* 15, CYP2A6* 16, CYP2A6* 21 and CYP2A6* 22: Differential Susceptibility Due to Different Sequence Locations of the Mutations. PloS One 9: e86230.

Tsukamoto, M, Baba, Y and Hiraga, S (1961) Mutations and linkage groups in Japanese strains of the housefly. Japanese Journal of Genetics 36: 168-174.

Vasanthanathan, P, Hritz, J, Taboureau, O, Olsen, L, Steen Jørgensen, F, Vermeulen, NP and Oostenbrink, C (2008) Virtual screening and prediction of site of metabolism for

cytochrome P450 1A2 ligands. Journal of Chemical Information and Modeling 49: 43-52.

Wang, B, Shahzad, M, Zhang, Z, Sun, H, Han, P, Li, F and Han, Z (2014) Genome-wide analysis reveals the expansion of Cytochrome P450 genes associated with xenobiotic metabolism in rice striped stem borer, *Chilo suppressalis*. Biochemical and Biophysical Research Communications 443: 756-760.

Wen, Z, Baudry, J, Berenbaum, MR and Schuler, MA (2005) Ile115Leu mutation in the SRS1 region of an insect cytochrome P450 (CYP6B1) compromises substrate turnover via changes in a predicted product release channel. Protein Engineering Design and Selection 18: 191-199.

Wen, Z and Scott, JG (1997) Cross-Resistance to Imidacloprid in Strains of German Cockroach (*Blattella germanica*) and House Fly (*Musca domestica*). Pesticide Science 49: 367-371.

Wiederstein, M and Sippl, MJ (2007) ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Researchearch 35: W407-W410.

Wilderman, PR, Gay, SC, Jang, H-H, Zhang, Q, Stout, CD and Halpert, JR (2012) Investigation by site-directed mutagenesis of the role of cytochrome P450 2B4 non-

active-site residues in protein-ligand interactions based on crystal structures of the ligand-bound enzyme. FEBS Journal 279: 1607-1620.

Witzig, C, Parry, M, Morgan, J, Irving, H, Steven, A, Cuamba, N, Kerah-Hinzoumbé, C, Ranson, H and Wondji, C (2013) Genetic mapping identifies a major locus spanning P450 clusters associated with pyrethroid resistance in kdr-free *Anopheles arabiensis* from Chad. Heredity. 110: 389-397.

Wondji, CS, Irving, H, Morgan, J, Lobo, NF, Collins, FH, Hunt, RH, Coetzee, M, Hemingway, J and Ranson, H (2009) Two duplicated P450 genes are associated with pyrethroid resistance in *Anopheles funestus*, a major malaria vector. Genome Researchearch 19: 452-459.

Xu, Q, Liu, H, Zhang, L and Liu, N (2005) Resistance in the mosquito, *Culex quinquefasciatus*, and possible mechanisms for resistance. Pest Management Science 61: 1096-1102.

Yang, T and Liu, N (2011) Genome analysis of cytochrome P450s and their expression profiles in insecticide resistant mosquitoes, *Culex quinquefasciatus*. PloS One 6: e29418.

Yang, Y, Yue, L, Chen, S and Wu, Y (2008) Functional expression of *Helicoverpa*

armigera CYP9A12 and CYP9A14 in Saccharomyces cerevisiae. Pesticide Biochemistry and Physiology 92: 101-105.

Zhang, H, Myshkin, E and Waskell, L (2005) Role of cytochrome b (5) in catalysis by cytochrome P4502B4. Biochemical and Biophysical Research Communications 338: 499-506.

Zhang, J, Liang, Y and Zhang, Y (2011) Atomic-level protein structure refinement using fragment-guided molecular dynamics conformation sampling. Structure 19: 1784-1795.

Zhang, Y (2008) I-TASSER server for protein 3D structure prediction. BMC Bioinformatics 9: 40.

Zhao, Y and Halpert, JR (2007) Structure-function analysis of cytochromes P450 2B. Biochimica et Biophysica Acta (BBA)-General Subjects 1770: 402-412.

Zhu, F, Feng, JN, Zhang, L and Liu, N (2008a) Characterization of two novel cytochrome P450 genes in insecticide-resistant house-flies. Insect Molecular Biology 17: 27-37.

Zhu, F, Li, T, Zhang, L and Liu, N (2008b) Co-up-regulation of three P450 genes in response to permethrin exposure in permethrin resistant house flies, *Musca domestica*.

BMC Physiol 8: 18.

Zhu, F and Liu, N (2008) Differential expression of CYP6A5 and CYP6A5v2 in pyrethroid-resistant house flies, *Musca domestica*. Arch Insect Biochem Physiol 67: 107-119.

Zhu, F, Moural, TW, Shah, K and Palli, SR (2013) Integrated analysis of cytochrome P450 gene superfamily in the red flour beetle, *Tribolium castaneum*. BMC Genomics 14: 174.

Zhu, F, Parthasarathy, R, Bai, H, Woithe, K, Kaussmann, M, Nauen, R, Harrison, DA and Palli, SR (2010) A brain-specific cytochrome P450 responsible for the majority of deltamethrin resistance in the QTC279 strain of *Tribolium castaneum*. Proceedings of the National Academy of Sciences 107: 8557-8562.

Zimmer, C, Bass, C, Williamson, M, Kaussmann, M, Wölfel, K, Gutbrod, O and Nauen, R (2013) Molecular and functional characterization of CYP6BQ23, a cytochrome P450 conferring resistance to pyrethroids in European populations of pollen beetle, Meligethes aeneus. Insect Biochemistry and Molecular Biology 45:18-29.

Table 4.1 Relative P450 gene expression values and the predicted autosomal interactions of the up-regulated P450 genes in ALHF *Musca domestica*

	Relative gene expression[§]						Autosome	Autosomal
Gene	ALHF	A2345	A1345	A1245	A1235	A1234	location	interaction
CYP4E10	274.42±22.40	125.67±10.39*	144.62±12.62*	257.81±23.61	268.61±25.8	66.6±6.74*	5	125
CYP4E11	5.86±0.55	6.31±0.48	2.61±0.17*	5.99±0.71	5.80±0.36	2.71±0.26*	5	25
CYP4G13	1.93±0.13	1.04±0.10*	1.22±0.09*	1.04±0.04*	2.01±0.12	1.91±0.08	3	123
CYP4G99	4.52±0.31	2.12±0.22*	4.28±0.41	1.38±0.13*	4.43±0.45	4.43±0.23	3	13
CYP4S24	2.94±0.08	1.03±0.05*	1.14±0.06*	0.83±0.03*	2.83±0.49	3.04±0.53	3	123
CYP6A36	7.03±0.64	2.60±0.21*	2.80±0.41*	6.63±0.62	7.01±1.04	1.72±0.19*	5	125
CYP6A40	2.95±0.22	2.90±0.32	1.16±0.10*	2.99±0.41	2.68±0.36	1.17±0.05*	5	25
CYP6A52	2.74±0.15	2.50±0.22	2.70±0.17	2.54±0.08	3.01±0.24	0.81±0.03*	5	5
CYP6A58	4.32±1.02	1.32±0.23*	1.86±0.36*	5.63±2.17	4.19±0.22	0.64±0.15*	5	125
CYP6D3	2.40±0.13	1.38±0.07*	1.34±0.06*	2.25±0.13	2.40±0.37	2.34±0.18	1	12
CYP6D10	6.24±0.63	2.55±0.37*	3.06±0.55*	5.99±0.25	6.19±0.51	2.62±0.32*	1	125
CYP9F7	2.01±0.19	2.13±0.18	1.54±0.11	1.90±0.17	2.08±0.10	2.23±0.29	2	2

[§] The relative levels of gene expression were shown as a ratio in comparison with that in aabys flies, the data are shown as the mean ± SEM.

* Gene expression value within a given *M. domestica* autosomal line was significantly lower than the expression in the parental ALHF strain at the $P < 0.05$ level of significance.

Table S4.1 List and sequences of the primers used

Gene	Forward primer (5'-3')	Reverse primer (5'-3')
Primers used for quantitative real time PCR		
CYP18A1	TTAAGACCACCCTGCTGTGGATCA	GTGGCCAATGGTACAATGCTGGAA
CYP304A2	CAATGTTGTCGGCTCAGGTCGTTT	AAGGTACCAGCGTCTCAATTCGCA
CYP3073A1	AAACGACAGCGGAGGGTTTGAGAT	TAACCACACCATTGCTCAGTCGGA
CYP3073A2	GGTGCATGTATTTAGTGCGGCCTT	AGCTCGAACTCTCCGTCATCGTTT
CYP3073B1	ACAATGTTGGGTCACAATGCCAGG	TTCAATGGTCTCCCGACATTCCCT
CYP305A1	TCACAAACCACCAGCACCACAATG	TAGGCCTTGTTACCCAAACTGGGA
CYP306A1	TGCCATTCTATTCGCGATCGGTCA	GCACTGAAGGCCACAAAGCTGAAT
CYP28B1	TACACCACCGAAGTTGTCAGCGAT	TCATCACCTTCTTTAGGCTGGGCA
CYP28B2	GAAACCCGGAGACCCAAGAGAAAT	AGTTTAGAGTTCACAGCCACGGGT
CYP28G6	TACTGGGTCTTTCGGCTGATGCTT	AAACTCCTCCACGGGCTTTGGTAT
CYP310B2	GATAGCATCGACGGCAGTATTT	CTTTGTTTCACTGTTCGCCATTAT
CYP317A3	TATGCAGCAAAGGGTACGGGAAGA	AGAGTGCGACGCATTACATAGGGT
CYP6A1	GTCTACGCGAGGAGGTTAATG	GCGGAGTGTTTCATTCAATACC
CYP6A24	TTCGAATACCCAAGGGAACACCCA	AAATCGGGCGCCAATACAGTTTCG
CYP6A25	TTCCCGGTCATCCGAAATATG	CGTCGGGATTGGGATAATACTG
CYP6A36	TTTATCCTTTGGTGATGGTCCC	CGCCACAATGATGCTCTTCTTA
CYP6A37	ATGAGTTGGCCCAGAATCAGGAGA	ATTCAGAATGGGCAGGACGGTGTA
CYP6A4	AAGGGCATAGATTTGTCGCATGGC	CCTGGTGCCTAGCCAATTCATAGA
CYP6A40	TGAGGGAGAGCAAGCAAATC	TGTTGTGGAGGAGGTCTCATA
CYP6A52	ATGTGATTGGTCGCTGTGCCTTTG	TACTTCTCCACATCGGGCATGGTT
CYP6A7	GCATTGTCCGCGAAACTGTGGAAT	AACACAAAGGCCTGGCCAGTTATC
CYP6A54	GAAGGTATTCGGTTCGGAGAAA	TGGCTCCAAAGTAGCGTAAAT
CYP6A63P	AGGCCATGATGGAGATGGCCTATT	ATCCTCCATGCAGACACGCTGTAA
CYP6A57	GCTGCCCAGGTGTTTGTCTTCTTT	AAAGTCTCTTTGACCTCCTGGCGT
CYP6A58	AATCGCCAAGCTGTCGAAGACTAC	TCCGGATTGAAGACATTGGGTTGG
CYP6A59	CCTATGAGTCGCTGAAGGATATG	GGGACCTGATAGTCTTTGACAC
CYP6A5	CAAATGGCCGCTCAGACCTTTGTT	AGGCCCTCATAGGTTATTTCGCCA
CYP6C2	TCTGTGGCGTAAGGTGAGAACCAA	TGCCAATGACATCGGTGGTGAAAC
CYP6D10	ACTGCTCGCAAATATCCTGGCCTA	TCCGGCTTATAATCCATGGGCTGA
CYP6D11	CAGCGGCTACCATAGCATTTA	CCCTCAACTTATGTTCCGTCAA
CYP6D3	AACTGCCTCAAGTACCTCATTTA	TACGTCCATCCGGCTTTATTC

CYP6D8	TGGTCGGCGGTAAAGACAAAGGAT	CGGCATTGCGTGCCAATTCATAGA
CYP6EK2	CCTTTCCGAATTTGGCCCGGAAAT	GGCCGCTATTTGTTCCAACGTCAA
CYP6FS2	AACCTTTAGCAACTCTCTGCGGGA	ATTGTGTGTTGCGATAACGCACGG
CYP6FT3	GACGATGACAGGGATATGGATAAG	GGCCAATTCAAAGAGAGCATTT
CYP6G4	TGACTGCTGGCTTTGAGACATCCT	ATTTCTCGACGCAAACGTTCCTGC
CYP6GU1	TGCTCTCTATGAATTGGCCCGGAA	AGGACGGCATATTTGCGTAGGGTT
CYP6GW1	GTTCCAGTTGACCATGGATATGA	CTCATATAGGGCCTGACACAAAG
CYP6V3	CGCAAGTACCCGATTGTGCCATTT	AAAGAGGTCTGGTTCGGGCCAATA
CYP9F10	GTTCTACGCAAATGGCCGGTGAAT	AAACGTTCCGGATCAAAGGCACTG
CYP9F11	AAATGGTCGTGTCGGAGGTGTTGA	TTCCGGATCGAAGGCATCAGGATT
CYP9F12	CGATGAGGCCATGTCGTATT	CAAATGCCGTCGAAGCAATAA
CYP9F7	CGATCAACGACAACCAATGCTGCT	TGGCGACAAAGTGTTGCGCATATC
CYP9F8v1	ATGCGTTCCATGTTCCAGCTGATG	GGCCAAATGCAGTCGAGGCAATAA
CYP9F9	AAACAAGGCAAGGGAGAGGATGGA	GACTTGCAGACCAAATGCAGTGCT
CYP311A1	CCAATACCTCGGATCTCAACAA	AATGTTACGCCAGGTGGATAG
CYP313D1	GGCTTTCGATGTGACCTATGCCAA	TTGATGACACATTGCACACCAGCC
CYP313D2	TGCACAGACGCAAGGATATATGGG	TTAGCAACTTTGCCAAGGCCACCT
CYP318B1	TGCATCGTAGTGCAGATGTTTGGG	GCCAAGTACATGCTATAGCGGCTT
CYP438A4	CGGTCGATAAGGCAACGATAG	GGCATACTGATACAAAGCGAATG
CYP4AA1	AACACTGGGATAGAGTCGCAACGA	AGGCCATGATGGAGATGGCCTATT
CYP4AC6v1	GAGGAGTATCGCAAGAAGAAGG	CATCCAGCAGCGTATCCAA
CYP4AD1	GCCCTCTTCACCACAGATTTA	AATGCCACAGAGAAGTCGATAG
CYP4AE3	TTGCCTCTATGCACTATCGCGTCA	CCAACAGCCGGAATTGGTGGAAAT
CYP4C74	TGAGGAGGTTGACACGTTTATG	TCGACAACACGTTCCTGATATT
CYP4D3	TTCGGCCGTAACGTTCTGCTTCTA	GCAGCGTCTCCTTGATGCACAAAT
CYP4D36	GTCGAACCATTTACGCTGGCCAAA	ATTGTGTAGACGCCGGAACAATGC
CYP4D4	GCGAATGGCTTTCCTTGATGTGCT	TCGTGGTGTCATGTCCTTCGAACA
CYP4D54	CAGTTGCCGGGAATAACAACACCA	AGATGCGAACACTGCTCCCACATA
CYP4D55	GCCTTGTTGGATGTGCTATTG	TAAGGGCCTGCATCAAAGAG
CYP4D56	TGTTTCCAGTTGTGTGTCGCATGG	TTGGCGAGGTGAGTTTGAAGAGGA
CYP4D58v1	ATTTGCCACCGCCGAAATGAAGAG	ACCCAATTGGACATCAGTCTCGGA
CYP4D9	CAAGCGTTGCCGACATAATAG	TCCCTCATCATCTTGCGATAC
CYP4E10	GCACCACCCTGCTACTATTT	CTCGAAGGGATTGTAGGTCATT
CYP4E11	CAAGATCATAACACCGGCATTTC	GCCTCCCTCAATTTGTCCATA
CYP4E7	ATCAACATGCGTGCCTTCAATCCC	ATTCTTTGCCTCCTGCTTGGCTTC

CYP4G13	ATTGCCCGTAAAGCGGAAGAGGAT	TAGATGTCGGCACGACGATGAACA
CYP4G2	ATGACCCGTAAAGTGGTGACGGAA	TCATCCAAATCATCGCGCAAACCC
CYP4G99	TGAGACATCGTAGCCTCTTTCTG	CGTTGCCATTGGAGTTTAGAC
CYP4P8	AATGCGCGAAGCAGTCGAAGAAAC	AAAGCATACGGATGACGTCCCACA
CYP4P10	ACGACATACATCGTAATCCCAGGC	AAATGCAAACGGGTGCCTCTGTTC
CYP4S23	TGCCATTCTATTCGCGATCGGTCA	GCACTGAAGGCCACAAAGCTGAAT
CYP4S24	GGTCCAACATTTCGCATTTG	GCTTCACCACTACTCGTCAGCA
CYP12A1	AAGTGGAGTGCCAGAGAAAC	GTATCCACACCAGCCAATATCA
CYP12A3	CCCACAGCAAAGCGTATGTTCCAA	GCCGAACAATGGCCTCTTCAACAA
CYP12A14	AGGGCCTGCATTAAGGAATCGCTA	AGCTCTCAACCATCGTTCGGGTAA
CYP12A16	GCCAATGGCAACATCGAACCAAGT	AAGAGCAAGCCGGTAAAGGCAGAT
CYP12A17	TAAGGGCCTGCATCAAAGAG	GATAACCGCTGAGGACAACAT
CYP12A13	CCCTGCGCATGTATCCATTGACAT	ACTTCCTGATTCGGTTGGCCTCAA
CYP12A2	AGTGGCTATCGTGTTCCCAAAGGT	GGCATGTGGACATTCAGCGGATTT
CYP12G2	ATGGGCGGTGTAGAAATCACCTCA	ATGGAGAAATCTTCGTGCCGTCCA
CYP12G4	TGTTTGGCAATATGCGAGCCTTGG	TTTCGGGCTGTGATATCTTCGGCA
CYP301A1	ACGGATCGCCTAAAGGTTCAAGGA	CAAACCGCCATGGAAATGGTGTCA
CYP302A1	TGCTTGTCTGAAGGAGGTATTC	TCCTTCGGCACCAAATATCC
CYP314A1	ACCGAACAGCCGGAGAAGATCAAA	TTTCCTGTAGGTCTGCGTGGGAAA
CYP315A1	TTTATACTGCCGGTCGTGATCCCA	ATAGGGCCAATTTGCGGCCAATAC
Actin	ATGAGGCTCAGAGCAAACGTGGTA	AGTCATCTTCTCGCGATTGGCCTT
Dm RPL11	CGATCCCTCCATCGGTATCT	AACCACTTCATGGCATCCTC
Primers used for autosome mapping		
CYP4E10	F1: CGATTTGAAGAACCAGAAGC	R: ACTCATTGTTGTGGCTCTCA
	F2: AAGCTGATTAAGGCGGAAC	
CYP4E11	F1: CCAGGCCTGTAACGGCAATCC	R: ACTCATTGTTGTGGCTCTCA
	F2: GGCCGATAACTATGCCACCGTT	
CYP4G13	F1: GTCTTCAAGGATTGTGGTGAAAC	R: TATTGTCACGCTCACGTAGCTTG
	F2: TTTGTGCCCACATTTGTCAAG	
CYP4G99	F1: TTGGACTTGCTCTTGGAGA	R: TGGTAGCCACTGTGATGGT
	F2: TGGTGCCACCATTACGGACACT	
CYP4S24	F1: AGAGAATCTGCCACATCTTGC	R: TGCCGAACTTGTCGTATCG
	F2: CGAAGATGCCAAGATACAAAAT	
CYP6A36	F1:TTTAAGGGGTCTCTCGACGAGT	R: AATGGAAACATGGCCTTCATA
	F2:TTATGAGAAATATAGAAATTCCGCC	

CYP6A52	F1: ATGTGATTGGTCGCTGTGCCTTTG	R: CATACCCTTGCGTATGACATAG
	F2: CCCAAATCGGAATTCCGC	
CYP6A40	F1: CATTCTCACCAGGGATTTCA	R: AAAGGCACAGCAGCCAAT
	F2: GCTGACACCCACCTTCAGTTCA	
CYP6A58	F1: CACCTTTACCTCGGGCAAA	R: CGCCATCCTCACTCTTCATC
	F2: AACCGAAACCATGCCCGATGT	
CYP6D3	F1: GACGGCTACAAATCTCAAGG	R: GTTTCATCGGCAAATCGTTC
	F2: TGAATTGGCCATGAATCCAGAT	
CYP6D10	F1: CTACCCGTTTGGAATCGTG	R: TCCTTGCGTGGCTGTATT
	F2: AAAGCTGCCCTGGTAAAAATG	
CYP9F10	F1: GCTTTTCTTATTGCCGGCATAG	R: GTCACTGCCTTCGGTAAATAC
	F2: ATGCCAAGCCTTCGTTCTTTTA	

Primers used for functional study

CYP4S24	CCGGGATCCCAAAATGGATTTACTAACAATCAACAG	CTAGCTAGCATTGCCACGCGATTTTATG
CYP6A36	CCGAGATCTCAAAATGGTTTTTCTAACGCTT	CTAGTCTAGACAACTTCTCCACCTTCAA
CYP6A52	CCGGAATTCCAAAATGATCGCTTTGACAATAC	CTAGCTCGAGAATCCTCTCCACACGCAAATAAAT
CYP6D10	CCGGAATTCCAAAATGTTTTTATATTTGGCTATATTCG	CTAGGCTAGCACATCGTTTTATGAGTGTAATTTTC

Table S4.2 List of selected P450 genes in the ALHF *M. domestica*

Name[†]	**Clan**	**SC_number**[‡]	**XM_number**[‡]	**XP_number**[‡]
CYP18A1	2	NW_004765049	XM_005183375	XP_005183432
CYP304A2	2	NW_004765002	XM_005183057	XP_005183114
CYP305A1	2	NW_004764745	XM_005180589	XP_005180646
CYP306A1	2	NW_004765049	XM_005183378	XP_005183435
CYP28B1	3	NW_004764738	XM_005180394	XP_005180451
CYP28B2	3	NW_004764738	XM_005180398	XP_005180455
CYP28G6	3	NW_004765174	XM_005184255	XP_005184312
CYP310B2	3	NW_004765160	XM_005184125	XP_005184182
CYP317A3	3	NW_004765183	XM_005184339	XP_005184396
CYP438A4	3	NW_004765049	XM_005183376	XP_005183433
CYP6A1	3	NW_004765183	XM_005184331	XP_005184388
CYP6A24	3	NW_004768817	XM_005190469	XP_005190526
CYP6A25	3	NW_004768817	XM_005190472	XP_005190529
CYP6A36	3	NW_004765183	XM_005184332	XP_005184389
CYP6A37	3	NW_004765183	XM_005184336	XP_005184393
CYP6A4	3	NW_004765183	XM_005184338	XP_005184395
CYP6A40	3	NW_004765183	XM_005184343	XP_005184400
CYP6A5	3	NW_004765183	XM_005184346	XP_005184405
CYP6A52	3	NW_004765183	XM_005184334	XP_005184391
CYP6A54	3	NW_004760864	XM_005175565	XP_005175622
CYP6A57	3	NW_004768817	XM_005190468	XP_005190525
CYP6A58	3	NW_004765183	XM_005184341	XP_005184398
CYP6A59	3	NW_004765183	XM_005184344	XP_005184401
CYP6A63P	3	NW_004760864	XM_005175566	XP_005175623
CYP6A7	3	NW_004765183	XM_005184333	XP_005184390
CYP6C2	3	NW_004765183	XM_005184350	XP_005184407
CYP6D10	3	NW_004765160	XM_005184128	XP_005184185
CYP6D11	3	NW_004765015	XM_005183145	XP_005183202
CYP6D3	3	NW_004765160	XM_005184123	XP_005184180
CYP6D8	3	NW_004765436	XM_005185673	XP_005185730
CYP6EK2	3	NW_004765349	XM_005185208	XP_005185265
CYP6FS2	3	NW_004764906	XM_005182161	XP_005182218

CYP6FT3	3	NW_004768760	XM_005190452	XP_005190509
CYP6G4	3	NW_004766190	XM_005188667	XP_005188724
CYP6GU1	3	NW_004765183	XM_005184337	XP_005184394
CYP6GW1	3	NW_004768817	XM_005190471	XP_005190528
CYP6V3	3	NW_004764478	XM_005176346	XP_005176403
CYP9F10	3	NW_004764700	XM_005180062	XP_005180119
CYP9F11	3	NW_004764700	XM_005180054	XP_005180111
CYP9F12	3	NW_004764700	XM_005180063	XP_005180120
CYP9F7	3	NW_004764700	XM_005180052	XP_005180109
CYP9F8v1	3	NW_004764700	XM_005180050	XP_005180107
CYP9F9	3	NW_004764700	XM_005180053	XP_005180110
CYP3073A1	4	NW_004765739	XM_005187010	XP_005187067
CYP3073A2	4	NW_004765739	XM_005187012	XP_005187069
CYP3073B1	4	NW_004765739	XM_005187011	XP_005187068
CYP311A1	4	NW_004764740	XM_005180423	XP_005180480
CYP313D1	4	NW_004767316	XM_005189825	XP_005189882
CYP313D2	4	NW_004766063	XM_005188279	XP_005188336
CYP318B1	4	NW_004766506	XM_005189277	XP_005189334
CYP4AA1	4	NW_004764464	XM_005175977	XP_005176034
CYP4AC6v1	4	NW_004765632	XM_005186465	XP_005186522
CYP4AD1	4	NW_004765578	XM_005186278	XP_005186335
CYP4AE3	4	NW_004764514	XM_005177255	XP_005177311
CYP4C74	4	NW_004765515	XM_005185973	XP_005186030
CYP4D3	4	NW_004764514	XM_005177258	XP_005177315
CYP4D36	4	NW_004764744	XM_005180553	XP_005180610
CYP4D4	4	NW_004765144	XM_005183986	XP_005184043
CYP4D54	4	NW_004764514	XM_005177250	XP_005177307
CYP4D55	4	NW_004764514	XM_005177252	XP_005177309
CYP4D56	4	NW_004764514	XM_005177251	XP_005177308
CYP4D58v1	4	NW_004765144	XM_005183988	XP_005184045
CYP4D9	4	NW_004764514	XM_005177345	XP_005177402
CYP4E10	4	NW_004765578	XM_005186272	XP_005186329
CYP4E11	4	NW_004765578	XM_005186268	XP_005186325
CYP4E7	4	NW_004765578	XM_005186267	XP_005186324
CYP4G13	4	NW_004764475	XM_005176292	XP_005176349

CYP4G2	4	NW_004764475	XM_005176294	XP_005176351
CYP4G99	4	NW_004764542	XM_005177736	XP_005177793
CYP4P10	4	NW_004764771	XM_005180896	XP_005180953
CYP4P8	4	NW_004764771	XM_005180909	XP_005180966
CYP4S23	4	NW_004764524	XM_005177495	XP_005177552
CYP4S24	4	NW_004764524	XM_005177488	XP_005177545
CYP12A1	Mito	NW_004764697	XM_005180004	XP_005180061
CYP12A13	Mito	NW_004764697	XM_005180007	XP_005180064
CYP12A14	Mito	NW_004764697	XM_005179996	XP_005180053
CYP12A16	Mito	NW_004769267	XM_005190663	XP_005190720
CYP12A17	Mito	NW_004764512	XM_005177016	XP_005177073
CYP12A2	Mito	NW_004764697	XM_005179998	XP_005180055
CYP12A3	Mito	NW_004764697	XM_005179997	XP_005180054
CYP12G2	Mito	NW_004764745	XM_005180644	XP_005180701
CYP12G4	Mito	NW_004765031	XM_005183241	XP_005183298
CYP301A1	Mito	NW_004764517	XM_005177409	XP_005177466
CYP302A1	Mito	NW_004764628	XM_005179205	XP_005179262
CYP314A1	Mito	NW_004764603	XM_005178726	XP_005178783
CYP315A1	Mito	NW_004765301	XM_005184970	XP_005185027

[†] Nomenclature provided by the cytochrome P450 nomenclature committee, David R. Nelson.

[‡] RefSeq accession number, National Center for Biotechnology Information, Bethesda, MD. SC_number: supercontig number; XM_number: mRNA sequence number; XP_number: amino acid sequence number.

Table S4.3 Relative expression profile of 86 P450 genes in three strains of *Musca domestica*

	Relative gene expression ± SEM*		
Gene	aabys	CS	ALHF
CYP4E10	1.00	1.56±0.08	274.42±22.40
CYP4E11	1.00	1.22±0.12	5.86±0.55
CYP4G13	1.00	0.99±0.11	1.93±0.13
CYP4G99	1.00	0.54±0.07	4.52±0.31
CYP4S24	1.00	1.04±0.11	2.94±0.08
CYP6A36	1.00	1.11±0.03	7.03±0.64
CYP6A40	1.00	1.08±0.11	2.95±0.22
CYP6A52	1.00	1.01±0.11	2.74±0.15
CYP6A58	1.00	0.84±0.18	4.32±1.02
CYP6D10	1.00	1.77±0.07	6.24±0.63
CYP6D3	1.00	0.26±0.05	2.40±0.13
CYP9F10	1.00	1.23±0.09	2.01±0.19
CYP12A3	1.00	0.74±0.01	0.35±0.15
CYP12G4	1.00	1.21±0.12	0.35±0.05
CYP28G6	1.00	1.18±0.17	0.21±0.04
CYP3073A1	1.00	1.03±0.11	0.12±0.03
CYP4AD1	1.00	0.94±0.08	0.07±0.01
CYP4D4	1.00	2.25±0.19	0.66±0.08
CYP6A24	1.00	1.08±0.11	0.69±0.07
CYP6A5	1.00	1.44±0.08	0.67±0.07
CYP6A57	1.00	1.89±0.17	0.66±0.03
CYP6A63P	1.00	0.22±0.05	0.08±0.03
CYP6D8	1.00	0.74±0.09	0.18±0.04
CYP6EK2	1.00	1.39±0.13	0.41±0.11
CYP6FT3	1.00	0.83±0.09	0.06±0.01
CYP6GW1	1.00	0.84±0.07	0.08±0.01
CYP12A1	1.00	1.28±0.12	1.02±0.19
CYP12A13	1.00	0.04±0.04	0.54±0.06
CYP12A14	1.00	0.36±0.06	0.74±0.08

CYP12A16	1.00	1.96±0.13	1.85±0.17
CYP12A17	1.00	3.13±0.11	2.81±0.24
CYP12A2	1.00	1.67±0.08	1.53±0.14
CYP12G2	1.00	2.99±0.25	2.16±0.19
CYP18A1	1.00	2.75±0.23	2.64±0.22
CYP28B1	1.00	0.95±0.11	1.04±0.12
CYP28B2	1.00	1.68±0.08	1.55±0.14
CYP301A1	1.00	0.78±0.08	0.69±0.07
CYP302A1	1.00	0.20±0.05	0.68±0.07
CYP304A2	1.00	0.15±0.03	0.74±0.16
CYP305A1	1.00	4.27±0.34	2.32±0.18
CYP306A1	1.00	0.95±0.10	1.09±0.11
CYP3073A2	1.00	0.34±0.06	0.37±0.05
CYP3073B1	1.00	0.74±0.13	0.88±0.09
CYP310B2	1.00	1.25±0.31	1.51±0.14
CYP311A1	1.00	0.96±0.11	0.86±0.09
CYP313D1	1.00	530.5±46.61	512.12±27.76
CYP313D2	1.00	0.93±0.12	0.79±0.08
CYP314A1	1.00	1.65±0.15	1.26±0.12
CYP315A1	1.00	1.88±0.13	2.07±0.18
CYP317A3	1.00	0.98±0.13	1.14±0.11
CYP318B1	1.00	0.91±0.13	0.99±0.11
CYP438A4	1.00	0.89±0.11	0.88±0.09
CYP4AA1	1.00	1.13±0.11	1.30±0.06
CYP4AC6v1	1.00	27.45±3.72	23.59±1.94
CYP4AE3	1.00	0.44±0.06	0.62±0.08
CYP4C74	1.00	5.16±0.41	1.60±0.15
CYP4D3	1.00	0.89±0.13	1.06±0.13
CYP4D36	1.00	1.11±0.08	1.03±0.12
CYP4D54	1.00	0.41±0.06	0.59±0.06
CYP4D55	1.00	0.53±0.07	1.04±0.16
CYP4D56	1.00	0.79±0.09	0.51±0.06
CYP4D9	1.00	1.02±0.11	1.08±0.13
CYP4D58V1	1.00	5.87±0.39	4.33±0.93
CYP4E7	1.00	1.11±0.11	1.05±0.12

CYP4G2	1.00	1.08±0.05	1.20±0.23
CYP4P10	1.00	2.98±0.26	2.93±0.11
CYP4P8	1.00	1.02±0.07	0.81±0.08
CYP4S23	1.00	11.69±1.31	6.61±0.55
CYP6A1	1.00	2.94±0.42	2.91±0.25
CYP6A25	1.00	3.67±0.32	1.05±0.10
CYP6A37	1.00	1.04±0.11	0.88±0.06
CYP6A4	1.00	2.02±0.18	1.18±0.05
CYP6A54	1.00	3.55±0.29	2.57±0.23
CYP6A59	1.00	1.17±0.12	0.86±0.09
CYP6A7	1.00	52.34±3.8	10.41±0.86
CYP6C2	1.00	1.51±0.14	0.82±0.08
CYP6D11	1.00	0.43±0.06	0.77±0.08
CYP6FS2	1.00	0.97±0.11	1.24±0.12
CYP6G4	1.00	3.22±0.26	2.41±0.31
CYP6GU1	1.00	1.79±0.16	1.91±0.21
CYP6V3	1.00	0.96±0.10	0.92±0.09
CYP9F11	1.00	1.24±0.05	0.99±0.09
CYP9F12	1.00	2.17±0.19	1.02±0.10
CYP9F7	1.00	1.23±0.09	1.31±0.12
CYP9F8v1	1.00	1.24±0.12	1.18±0.09
CYP9F9	1.00	1.06±0.11	0.81±0.08

* The relative levels of gene expression were shown as a ratio in comparison with that in aabys flies, the data are shown as the mean ± SEM.

Table S4.4 The relative expression of P450 genes in ALHF *Musca domestica* following permethrin treatment

	Relative gene expression ± SEM*				
Gene	0h	6h	12h	24h	48h
CYP4D4	1.00	1.23±0.09	1.71±0.23	2.65±0.34	1.63±0.16
CYP4G2	1.00	1.31±0.17	1.65±0.21	2.52±0.12	0.98±0.22
CYP4S23	1.00	2.04±0.36	1.82±0.28	0.95±0.15	1.01±0.25
CYP4S24	1.00	1.09±0.12	2.25±0.32	2.53±0.32	0.77±0.13
CYP6A24	1.00	1.41±0.16	1.96±0.29	3.19±0.17	0.94±0.23
CYP6A36	1.00	1.04±0.11	2.05±0.30	2.12±0.27	1.34±0.18
CYP6A52	1.00	2.46±0.29	3.01±0.40	2.22±0.28	2.07±0.33
CYP6A54	1.00	4.46±0.55	3.71±0.47	3.27±0.43	0.81±0.44
CYP6A58	1.00	2.03±0.22	2.43±0.24	0.99±0.15	0.65±0.12
CYP6D8	1.00	4.03±0.53	3.25±0.42	7.88±0.54	0.98±0.15
CYP6D10	1.00	1.36±0.25	5.22±0.92	2.82±0.35	1.14±0.16
CYP6V3	1.00	1.59±0.26	1.96±0.31	2.80±0.35	1.05±0.16
CYP6EK2	1.00	4.83±0.39	8.73±0.97	6.98±0.81	2.02±0.48
CYP9F9	1.00	2.94±0.36	6.54±0.75	2.65±0.33	1.83±0.23
CYP12A13	1.00	2.10±0.25	1.95±0.29	0.85±0.14	1.21±0.13
CYP12G4	1.00	3.12±0.52	3.50±0.44	1.17±0.09	1.24±0.29
CYP12A14	1.00	0.84±0.09	0.68±0.17	0.56±0.17	0.49±0.11
CYP12A17	1.00	0.75±0.08	0.22±0.12	0.43±0.09	0.21±0.08
CYP12A2	1.00	0.55±0.16	0.46±0.15	0.52±0.10	0.38±0.11
CYP28G6	1.00	0.55±0.08	0.49±0.15	0.81±0.13	0.84±0.14
CYP302A1	1.00	0.58±0.08	0.43±0.14	0.43±0.09	0.51±0.11
CYP304A2	1.00	0.57±0.18	0.95±0.19	0.69±0.12	1.29±0.18
CYP306A1	1.00	0.68±0.08	0.45±0.14	0.91±0.14	0.75±0.13
CYP3073A2	1.00	0.48±0.06	0.46±0.15	0.23±0.06	0.56±0.11
CYP3073B1	1.00	0.59±0.02	0.66±0.16	0.84±0.13	1.34±0.18
CYP311A1	1.00	0.58±0.12	0.56±0.15	0.91±0.14	1.05±0.16
CYP4AA1	1.00	0.84±0.19	0.51±0.15	0.99±0.22	0.66±0.12
CYP4AC6v1	1.00	0.16±0.05	0.14±0.04	0.59±0.11	0.55±0.11
CYP4D3	1.00	0.64±0.09	0.28±0.13	0.52±0.13	0.93±0.15

CYP4D36	1.00	0.41±0.06	0.21±0.12	0.41±0.09	1.26±0.17
CYP4D55	1.00	0.22±0.05	0.15±0.05	0.35±0.08	0.48±0.11
CYP4D58v1	1.00	1.13±0.13	0.75±0.17	0.58±0.11	0.91±0.14
CYP6A1	1.00	0.49±0.22	0.38±0.14	0.27±0.07	0.33±0.09
CYP6A25	1.00	0.24±0.05	0.19±0.12	0.13±0.06	0.21±0.08
CYP6A7	1.00	0.75±0.08	0.63±0.16	0.69±0.12	0.27±0.09
CYP6D11	1.00	1.03±0.23	1.38±0.21	0.98±0.15	0.52±0.11
CYP6GU1	1.00	0.96±0.20	1.02±0.24	0.63±0.11	1.19±0.17
CYP9F8v1	1.00	0.58±0.07	0.34±0.13	0.35±0.08	0.85±0.14
CYP9F12	1.00	0.65±0.07	0.52±0.15	0.67±0.12	0.51±0.11
CYP12A1	1.00	1.26±0.24	1.36±0.23	1.17±0.08	1.30±0.18
CYP12A16	1.00	0.95±0.21	0.98±0.10	0.95±0.25	1.03±0.15
CYP12A3	1.00	1.18±0.13	1.11±0.22	1.32±0.19	1.01±0.15
CYP12G2	1.00	1.25±0.16	1.09±0.22	0.75±0.29	0.85±0.14
CYP18A1	1.00	0.99±0.11	0.81±0.18	0.80±0.19	0.92±0.19
CYP28B1	1.00	1.07±0.14	1.08±0.26	0.88±0.14	1.11±0.16
CYP28B2	1.00	1.01±0.13	0.91±0.24	0.79±0.19	1.05±0.16
CYP301A1	1.00	1.04±0.29	1.21±0.22	1.14±0.17	1.01±0.15
CYP305A1	1.00	0.99±0.25	1.29±0.23	1.01±0.15	1.03±0.15
CYP3073A1	1.00	1.12±0.14	1.22±0.22	1.26±0.18	1.21±0.17
CYP310B2	1.00	1.01±0.13	0.84±0.18	1.04±0.15	1.12±0.21
CYP313D1	1.00	0.98±0.06	1.27±0.22	1.33±0.19	1.31±0.18
CYP313D2	1.00	1.06±0.33	1.18±0.21	0.97±0.15	0.73±0.43
CYP314A1	1.00	0.81±0.09	1.09±0.21	1.15±0.17	0.84±0.14
CYP315A1	1.00	0.89±0.21	0.89±0.28	0.88±0.25	1.16±0.19
CYP317A3	1.00	0.81±0.09	1.15±0.21	1.17±0.17	0.77±0.13
CYP318B1	1.00	1.09±0.18	0.86±0.18	1.29±0.18	0.96±0.15
CYP438A4	1.00	0.86±0.13	1.06±0.26	1.04±0.16	1.21±0.17
CYP4AD1	1.00	1.17±0.31	1.01±0.24	1.28±0.18	1.27±0.18
CYP4AE3	1.00	0.83±0.11	0.94±0.19	1.16±0.17	1.24±0.17
CYP4C74	1.00	0.95±0.12	1.09±0.02	1.01±0.15	1.27±0.18
CYP4D54	1.00	0.98±0.07	1.07±0.23	0.78±0.13	0.86±0.19
CYP4D56	1.00	1.04±0.12	0.93±0.24	0.82±0.13	1.21±0.17
CYP4D9	1.00	1.16±0.15	1.16±0.21	1.34±0.22	0.84±0.14
CYP4E10	1.00	1.05±0.23	0.92±0.19	0.91±0.19	0.73±0.23

CYP4E11	1.00	1.37±0.36	1.26±0.22	1.26±0.18	1.34±0.18
CYP4E7	1.00	0.87±0.06	0.92±0.19	1.16±0.17	0.91±0.14
CYP4G13	1.00	1.35±0.27	0.87±0.18	0.94±0.24	1.25±0.17
CYP4G99	1.00	0.95±0.23	0.94±0.19	0.87±0.14	0.89±0.09
CYP4P10	1.00	1.14±0.13	1.18±0.17	1.31±0.19	1.21±0.17
CYP4P8	1.00	1.32±0.15	1.23±0.18	1.33±0.29	1.12±0.16
CYP6A37	1.00	1.29±0.16	1.23±0.22	0.86±0.14	1.17±0.19
CYP6A4	1.00	1.09±0.22	1.01±0.20	1.23±0.18	0.86±0.14
CYP6A40	1.00	1.03±0.33	1.04±0.25	1.22±0.35	1.04±0.32
CYP6A5	1.00	1.06±0.21	1.35±0.15	1.15±0.17	1.14±0.16
CYP6A57	1.00	1.18±0.14	1.08±0.26	1.11±0.16	0.99±0.15
CYP6A59	1.00	1.21±0.09	0.79±0.18	1.39±0.19	0.84±0.24
CYP6A63P	1.00	1.22±0.31	1.07±0.21	0.95±0.20	1.16±0.17
CYP6C2	1.00	1.38±0.24	1.24±0.22	0.83±0.13	1.17±0.07
CYP6D3	1.00	0.87±0.11	0.95±0.12	0.92±0.14	0.75±0.13
CYP6FS2	1.00	1.12±0.24	1.11±0.21	1.10±0.16	0.81±0.13
CYP6FT3	1.00	1.12±0.22	1.21±0.22	1.13±0.16	1.29±0.18
CYP6G4	1.00	1.14±0.15	0.96±0.15	1.26±0.23	1.03±0.21
CYP6GW1	1.00	1.12±0.13	1.08±0.21	1.13±0.17	1.22±0.19
CYP9F10	1.00	1.15±0.13	0.87±0.24	1.25±0.18	0.88±0.14
CYP9F11	1.00	1.04±0.17	1.47±0.35	1.18±0.17	0.83±0.14
CYP9F7	1.00	1.07±0.12	1.18±0.21	1.12±0.16	0.75±0.17

* The relative levels of gene expression following permethrin treatment are shown as the mean ± SEM.

Table S4.5 Bottleneck values of channels detected by Cavity 3.0

	Bottleneck radii values of channels detected by Cavity 3.0 (Å)							
Protein	2a	2ac	2b	2c	2d	2e	2f	Solvent
CYP4S24	2.14	1.64	1.61	1.66	1.21	-	1.42	1.23
CYP6A36	1.55	1.22	1.37	2.04	1.32	1.47	1.39	1.26
CYP6A52	1.31	1.35	1.36	1.85	1.36	1.63	1.27	1.24
CYP6D10	1.33	1.59	1.36	1.93	1.89	1.77	1.31	1.27

- Indicated channel was not detected by Cavity 3.0

Table S4.6 Docking results of selected P450 homology models

Protein	Active cavity volume (Å^3) [§]	Putative metabolic site[#]	Estimate free energy (kcal/mol)	Distance from heme iron (Å) *	Predicted permethrin contact residues[$]
CYP6A36-aabys	574	gem	-10.98	2.67	[1](ILE122), [2](ASN216), [4](LEU310, ALA311, GLU314, THR315), [5](THR376, ILE377, PRO378), [6](ILE492)
		C5-PB	-9.91	2.64	[1](ARG106, ILE122), [2](ASN216), [4](ALA311, THR315), [5](THR376, ILE377, THR379), [6](ILE492)
		C4′-PB	-10.58	2.85	[1](ASP114, ILE122), [2](ASN216, GLU217), [3](GLU242), [4](ALA311, THR315), [5](ILE377)
CYP6A36-ALHF	574	gem	-10.98	2.67	[1](ILE122), [2](ASN216), [4](LEU310, ALA311, GLU314, THR315), [5](THR376, ILE377, PRO378), [6](ILE492)
		C5-PB	-9.91	2.64	[1](ARG106, ILE122), [2](ASN216), [4](ALA311, THR315), [5](THR376, ILE377, THR379), [6](ILE492)
		C4′-PB	-10.58	2.85	[1](ASP114, ILE122), [2](ASN216, GLU217), [3](GLU242), [4](ALA311, THR315), [5](ILE377)
CYP6A52-aabys	535	gem	-9.71	3.21	[1](ARG105, PHE121), [2](ILE214, GLU215), [4](LEU308, ALA309), [6](ILE490)
		C5-PB	-9.15	3.43	[1](ARG105, PHE121), [2](ILE214, GLU215), [4](ALA309, THR313), [6](ILE490, ILE491)
		C4′-PB	-8.67	3.19	[1](ARG105, PHE108, PHE121), [2](ILE214, GLU215), [4](ALA309, THR313, ILE374, ALA375), [6](ILE490)
CYP6A52-ALHF	482	gem	-10.16	2.94	[1](ARG105), [2](MET213, ILE214, GLU215), [4](LEU308, ALA309), [6](ILE490)
		C5-PB	-9.51	3.25	[1](ARG105, PHE121), [2](ILE214, GLU215), [4](LEU305, LEU308, ALA309, THR313), [6](ILE490)

		C4′-PB	-8.93	3.01	[1](ARG105, PHE108, PHE121), [2](ILE214, GLU215), [4](LEU305, LEU308, ALA309, THR313), [6](ILE490)
CYP6D10-aabys	593	gem	-9.73	3.09	[1](ARG99, SER111), [2](LYS211), [4](ALA319, THR323), [5](LEU389, PRO390, VAL391)
		C5-PB	-9.78	3.23	[1](ARG99, SER111, ALA112), [2](LYS211, THR212), [4](LEU315, ALA319, THR323), [5](LEU389, VAL391)
		C4′-PB	-9.45	3.16	[1](ARG99, GLU105, LYS107, PHE115, SER116), [2](LYS211), [4](LEU315, ILE318, ALA319), [5](LEU389)
CYP6D10-ALHF	537	gem	-9.71	2.97	[1](ARG99, TYR102, PHE115), [2](LYS211), [4](ALA319, THR323), [5](LEU389, PRO390, VAL391)
		C5-PB	-10.11	2.91	[1](ARG99, GLU105, LYS107, PHE115), [2](LYS211, THR212), [4](LEU315, ILE318, ALA319, THR323), [5](LEU389)
		C4′-PB	-9.75	2.98	[1](ARG99, TYR102, GLU105, LYS107, PHE115, SER116), [2](LYS211), [4](LEU315, ILE318, ALA319), [5](LEU389)
CYP4S24	279	C5-PB	-9.86	7.41	[1](HIS101, ASN105, TYR106, LEU109), [5](SER370, VAL371, PRO372, THR373, ALA375), [6](LEU479, VAL480)

[§] Active site cavity was calculated using the VOIDOO program with conventional probe radius of 1.4 Å. Å, angstrom. 1 Å=10^{-10} m

[#] Predicted metabolic sites are indicated as following: Gem, germinal-dimethyl group; C5-PB, carbon 5 in alcohol moiety; C4′-PB, carbon 4′ in terminal aromatic ring

[*] The distance between the heme iron and putative metabolic sites.

[$] Predicted contact residues of permethrin are grouped based on SRSs of P450s. Superscript presents order of SRS

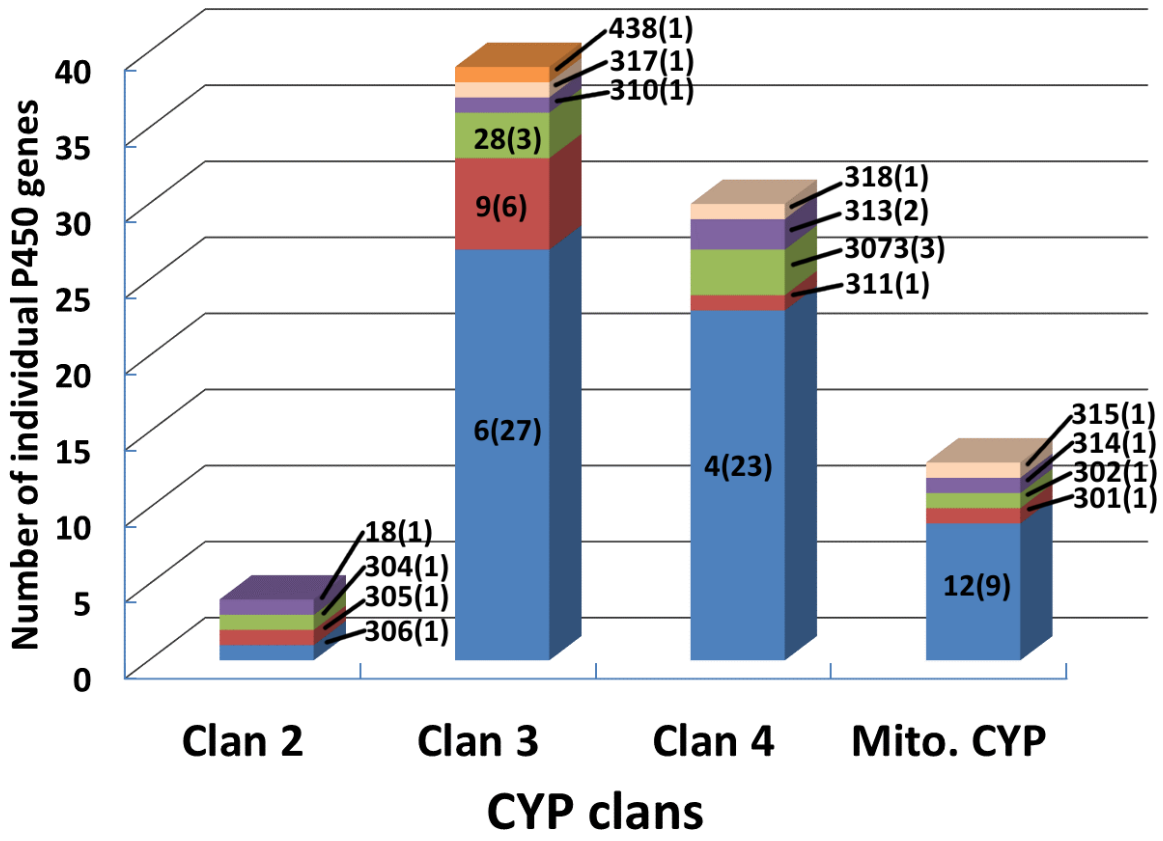


Figure 4.1 Number, family and clan distribution of cytochrome P450 genes in house flies, *Musca domestica*. The number shown along each column represents the P450 family, the number in parenthesis is the number of individual genes in the corresponding family. The P450 gene sequence information was generated from the RNA-seq ([NCBI:SRR521288], [NCBI:SRR521289]) and genome (PRJNA210139, PRJNA176013) analysis of *Musca domestica*.

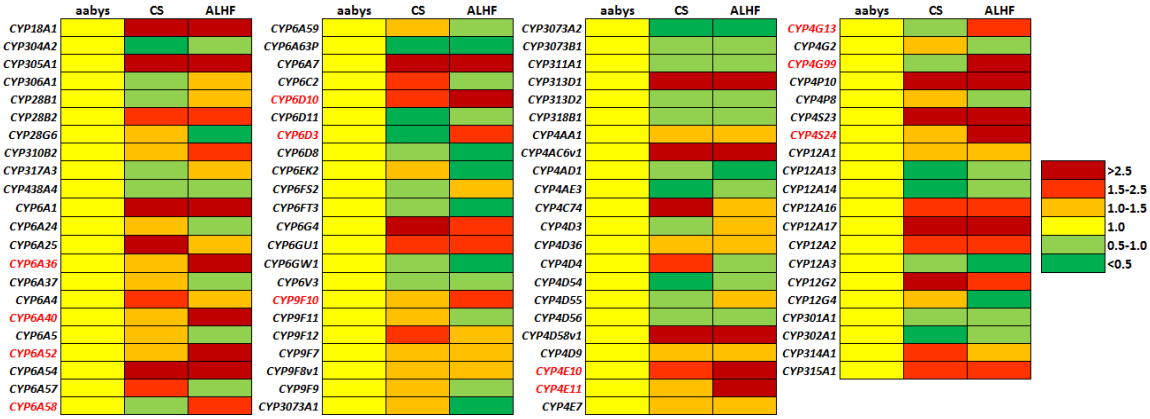


Figure 4.2 mRNA levels of 86 P450 genes in aabys, CS and ALHF *Musca domestica* strains. mRNA levels were shown as the mean fold relative to their levels in aabys strain. Colors scaled from green to red indicate low to higher mRNA level. P450 genes were highlighted with red indicated significant up-regulated expression of P450 gene in ALHF strain compared to aabys and CS strains ($P < 0.05$).

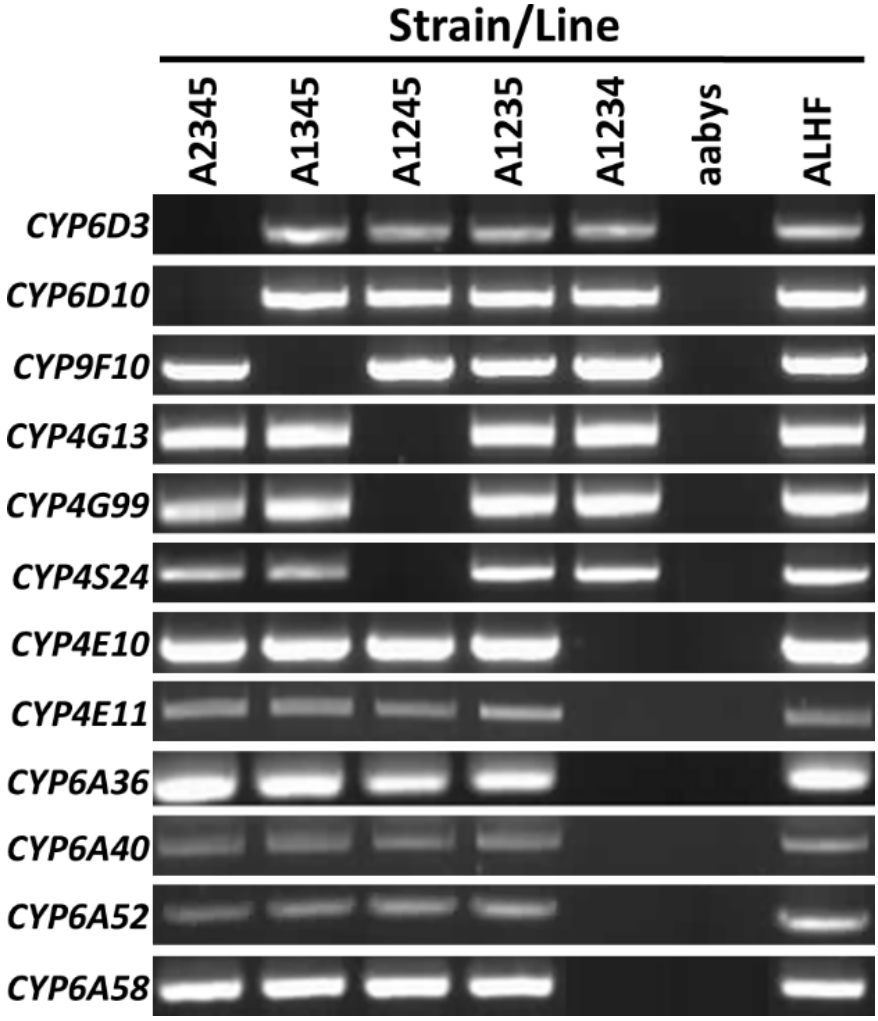


Figure 4.3 Allele-specific RT-PCR autosomal mapping of the *Musca domestica* P450 genes. PCR fragments were generated using the allele-specific primer set according to the sequence of each gene from ALHF. The absence of a PCR product in a house fly line indicated that the gene was located on the corresponding autosome from aabys (i.e. the absence of a band in the A1234 line indicates that the gene was present on autosome 5).

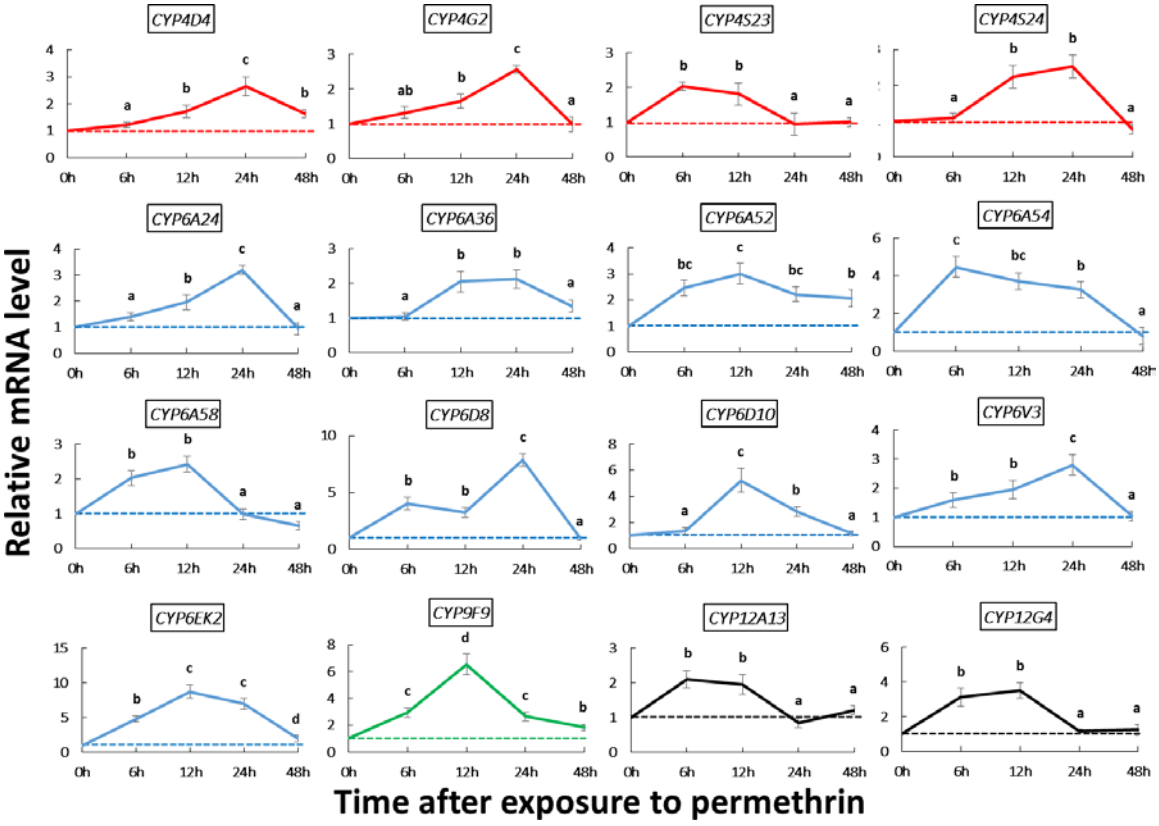


Figure 4.4 Induction of P450 gene expression in ALHF house flies following permethrin treatment. The relative expression of P450 genes was analyzed by qRT-PCR as described in the methods. Y axis represents the ratio of the gene expression in each treatment compared with that in the acetone treated control house flies. The results are shown as the mean ± SEM (n = 3). There is no significant difference in the expression level with the same alphabetic letter ($P < 0.05$).

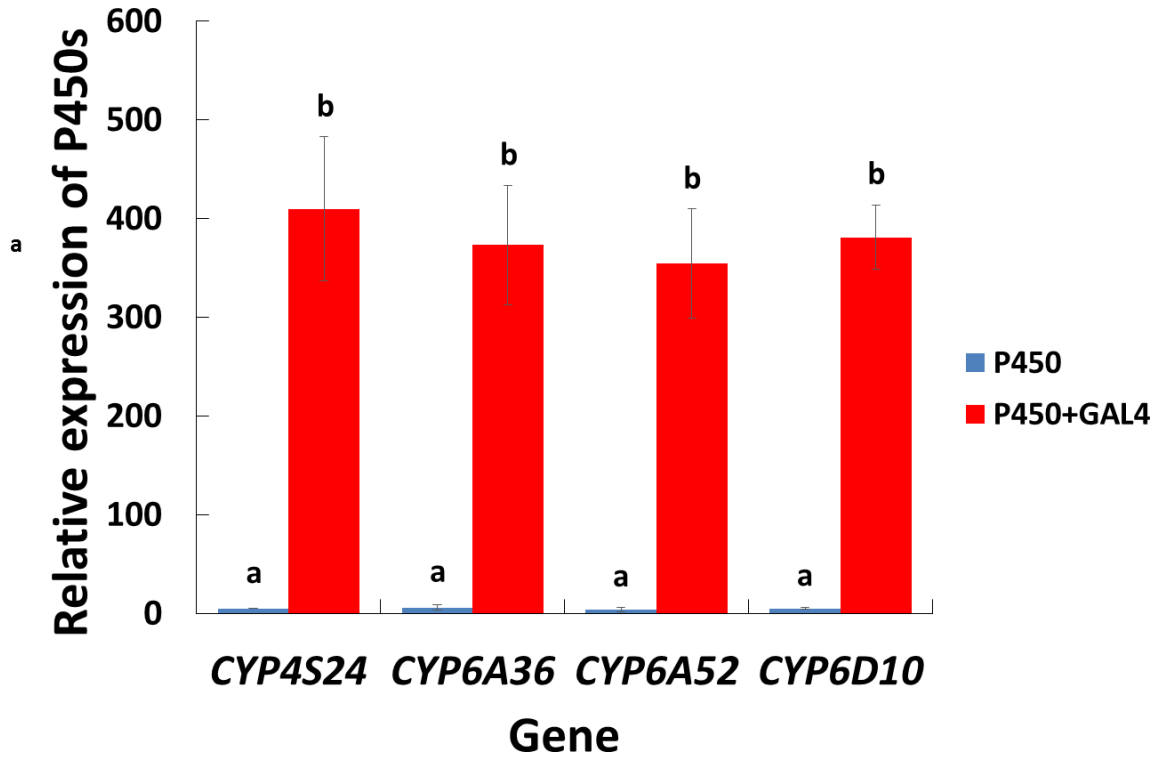


Figure 4.5 Transgenic expression of *CYP4S24, CYP6A36, CYP6A52* and *CYP6D10* in *Drosophila melanogaster*. The relative expression of the four transgenes were quantified by qRT-PCR. "P450" represents the homozygous transgenic *Drosophila* line with house fly P450 genes, "P450+GAL" represents the F1 generation of homozygous transgenic *Drosophila* line crossed with GAL4 driver line. Data shown are mean ± SEM (n = 3). There was no significant difference in the expression level with the same alphabetic letter ($P < 0.05$).

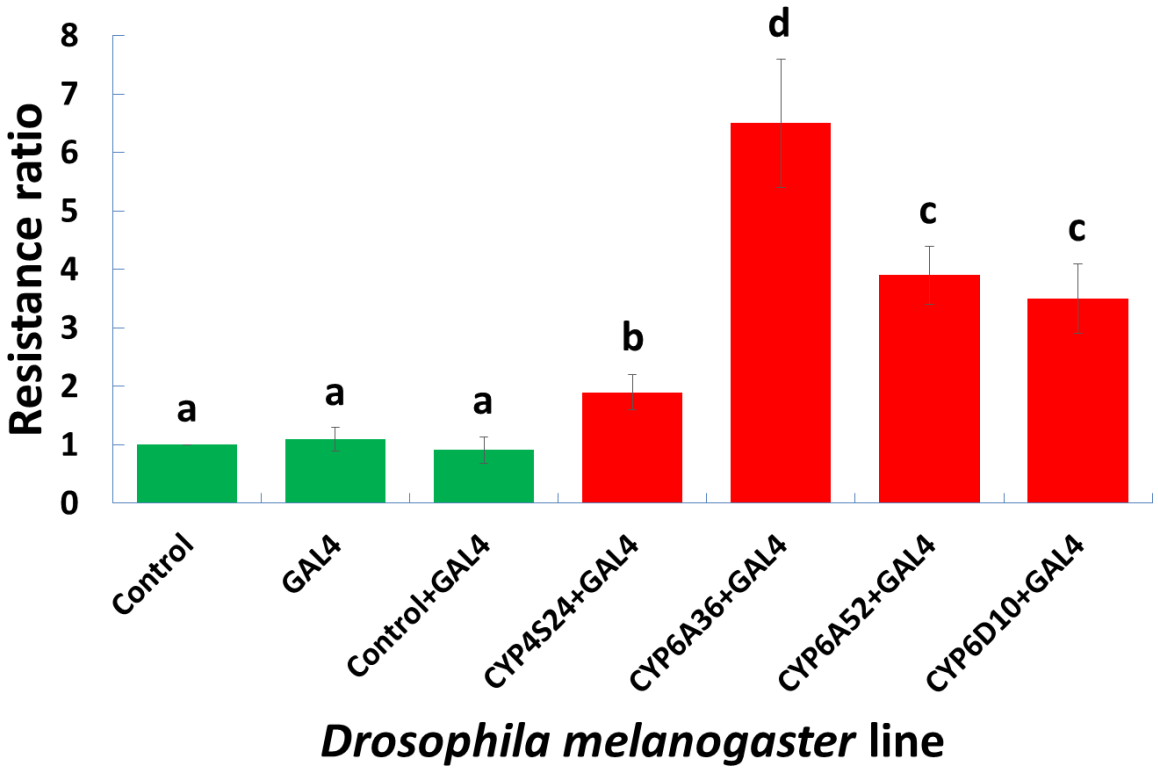


Figure 4.6 Toxicity of permethrin to non-transgenic and transgenic *Drosophila melanogaster* lines. Resistance ratios = LC_{50} of *Drosophila melanogaster* lines/LC_{50} of Control line. There was no significant difference in the expression level with the same alphabetic letter ($P < 0.05$).

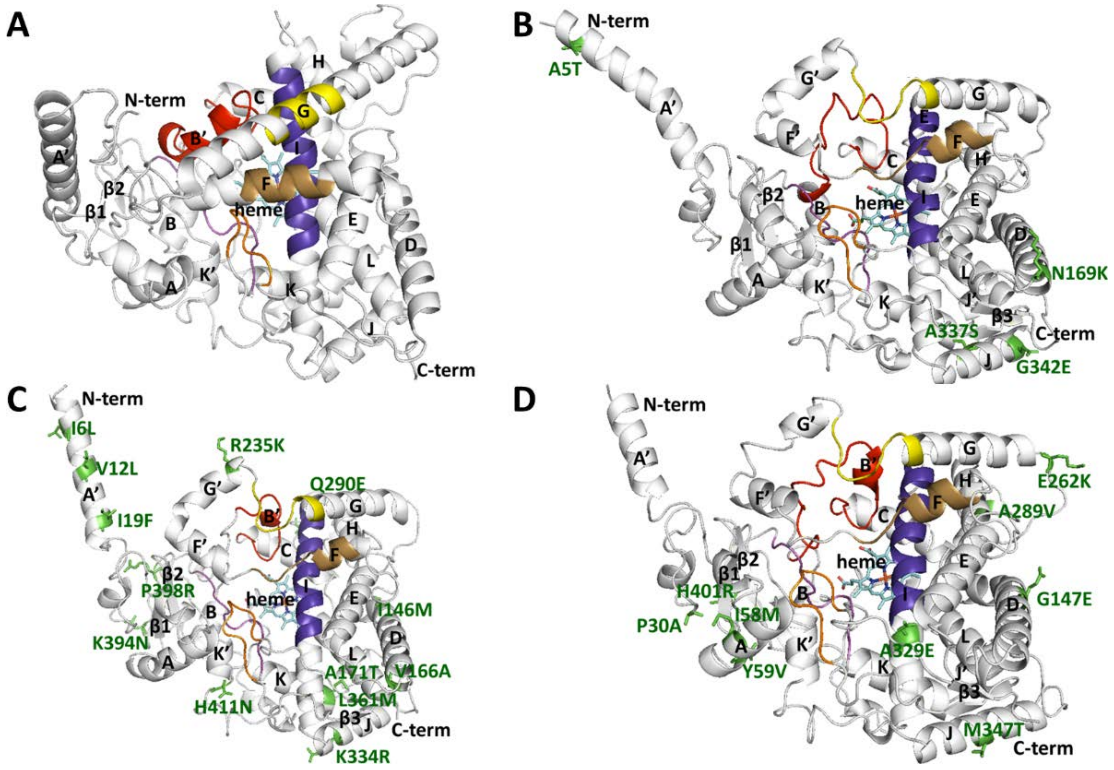


Figure 4.7 Topology of P450s. Secondary structure of helices and sheets are labeled. The mutated residues are shown in green sticks and listed with aabys residue first and ALHF residue second. Six putative SRSs were colored and predicted, according to Gotoh's predicted models (Gotoh 1992). SRS 1 to 6 is represented by red, wheat, yellow, blue, purple and orange color, respectively. The heme group is represented by cyan sticks. (A) CYP4S24, (B) CYP6A36, (C) CYP6A52, (D) CYP6D10.

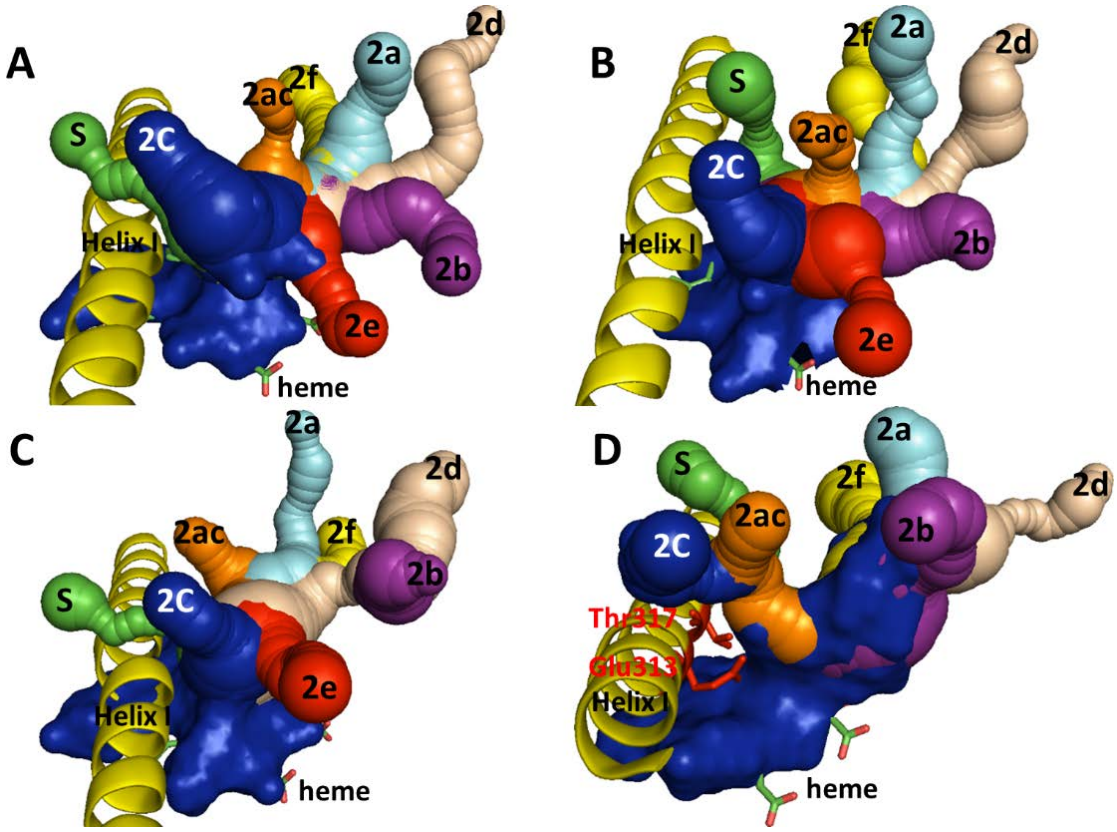


Figure 4.8 Solvent channel, family 2 channels and active site of P450s. Channels are shown as connected spheres with different color: cyan (2a), orange (2ac), purple (2b), blue (2c), wheat (2d), red (2e) and yellow (2f). Helix I and heme of P450s are labeled. (A) CYP6A36, (B) CYP6A52, (C) CYP6D10, (D) CYP4S24. Glutamate 313 and Threonine 317 are labeled.

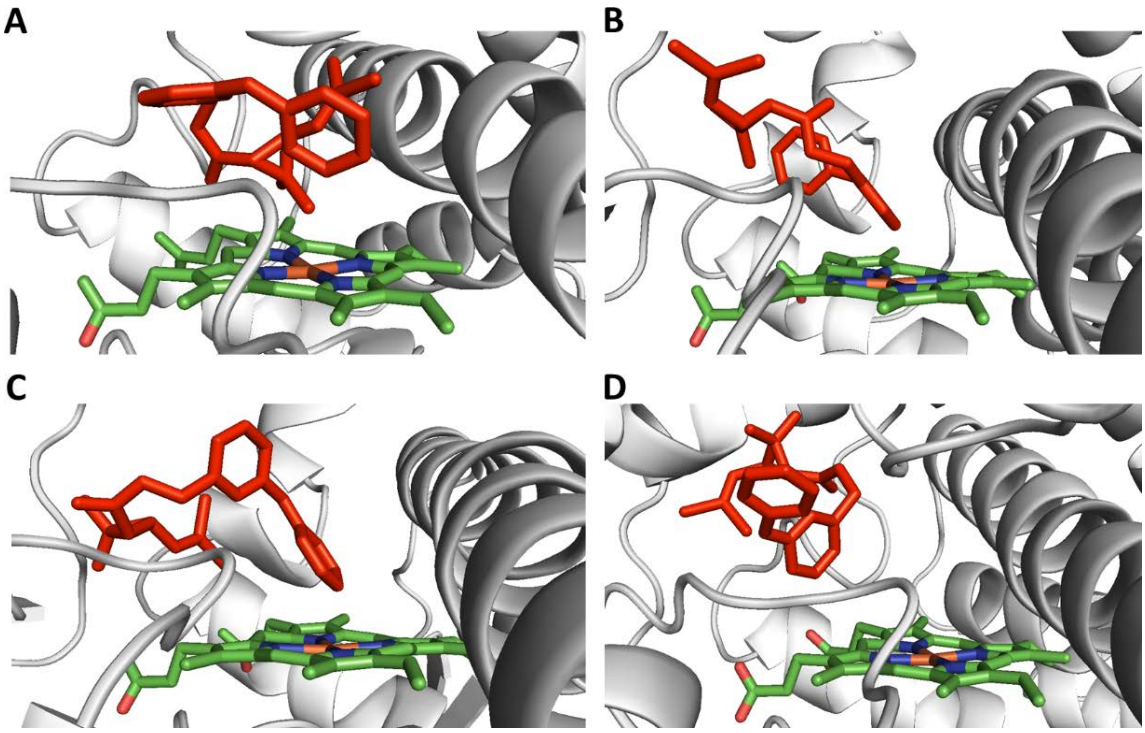


Figure 4.9 Permethrin binding models in active site of P450s. The heme group is represented by green sticks. The permethrin is represented by red sticks. CYP6A36, CYP6A52 and CYP6D10 exhibited 3 permethrin binding models close to heme iron: (A) germinal-dimethyl group, (B) 5-phenoxybenzyl carbon, (C) 4′-phenoxybenzyl carbon. Single binding mode, (D) 5-phenoxybenzyl carbon was obtained in CYP4S24.

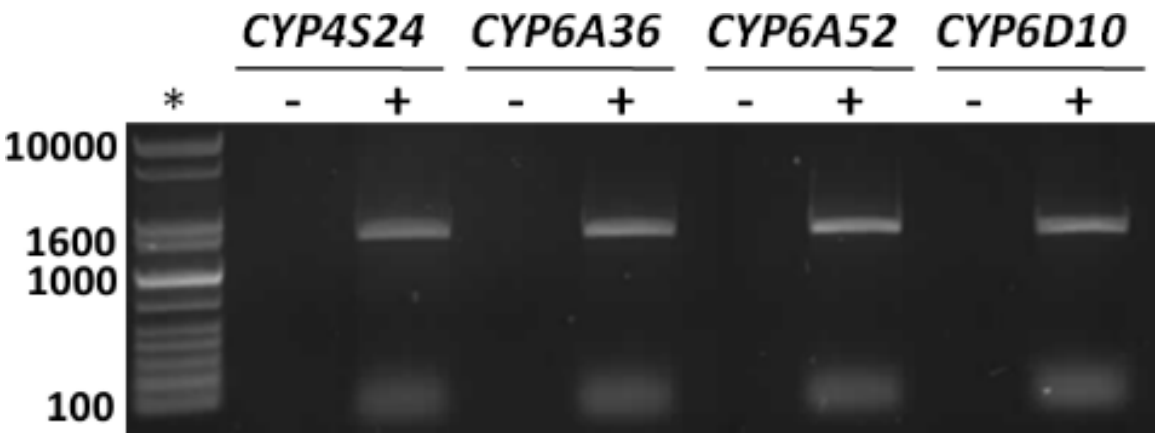
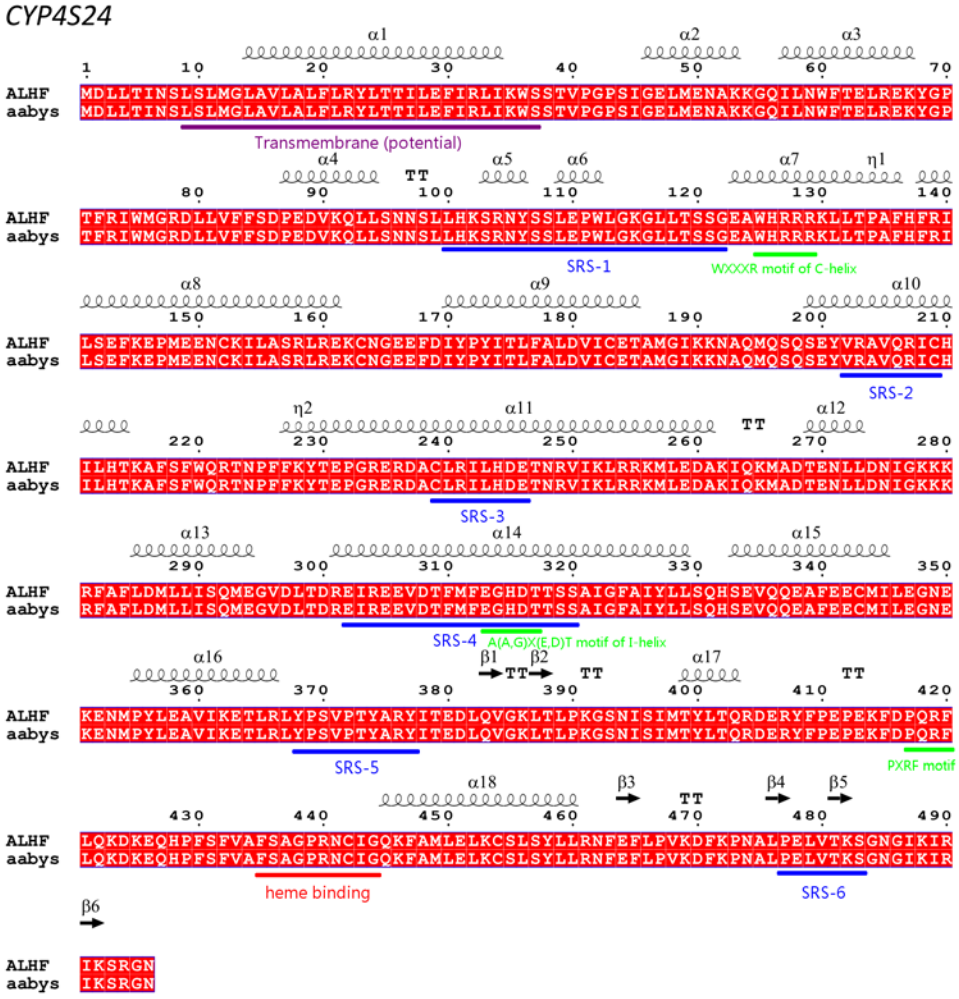
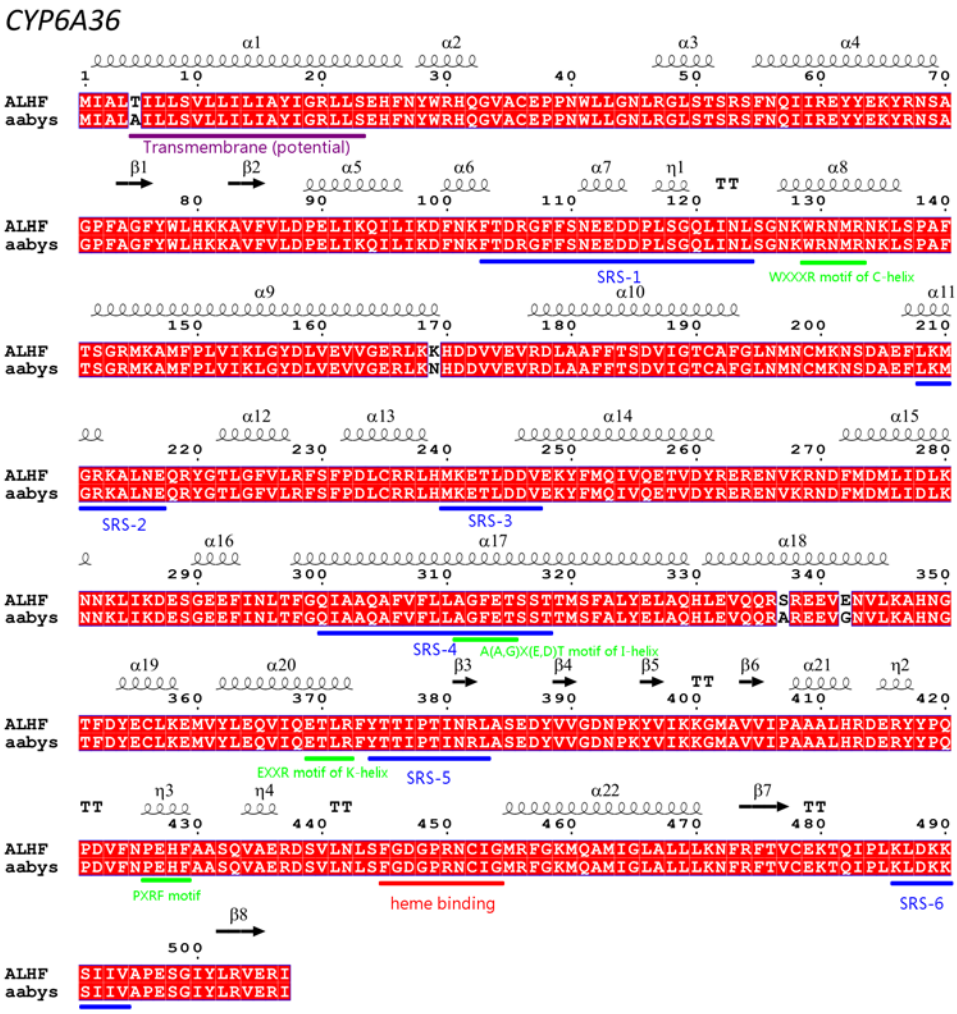
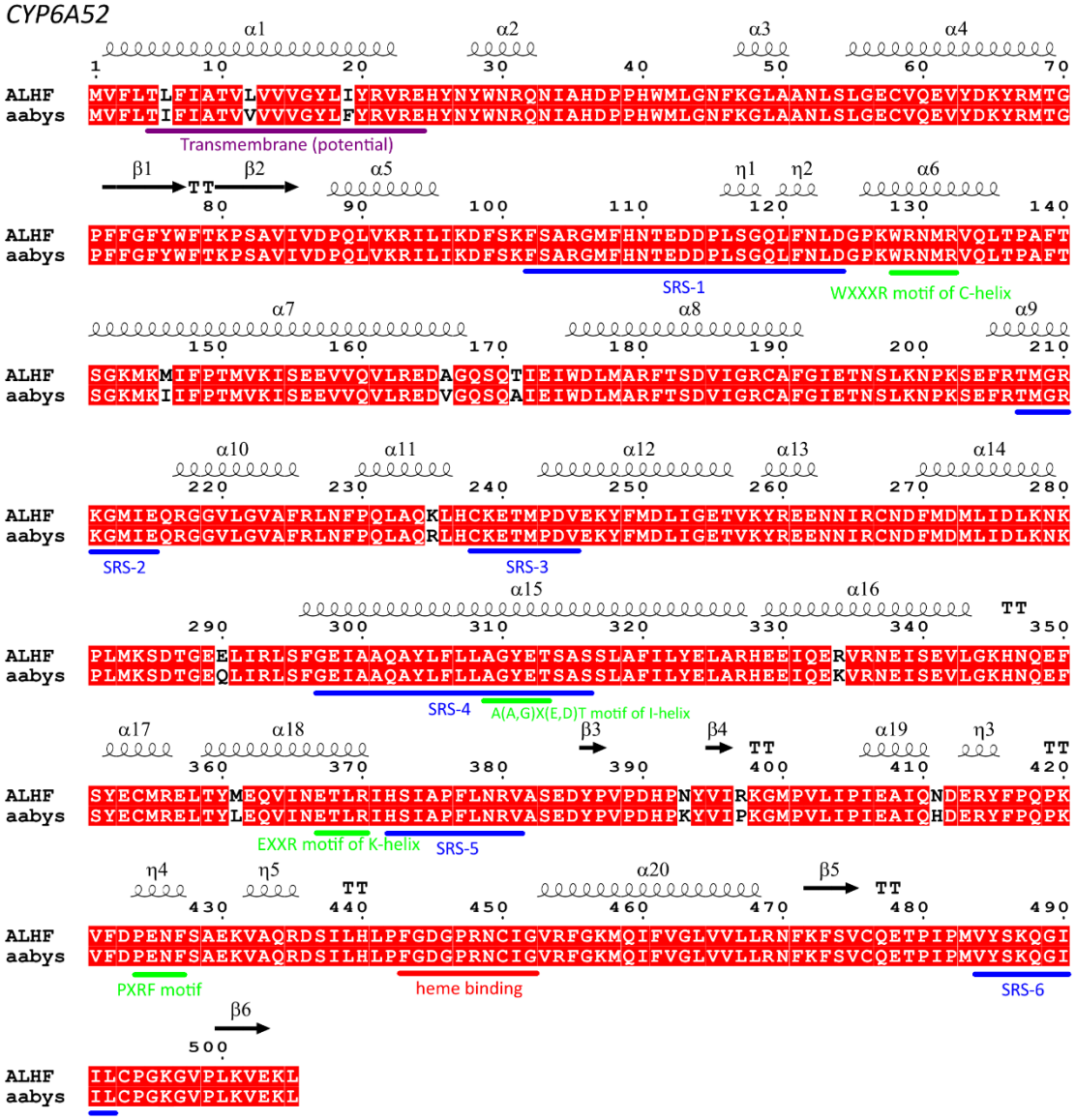


Figure S4.1 RT-PCR amplification of the transgenic *Drosophila melanogaster* lines using gene specific primers. The lanes, within gene, indicated with a "-" represent the non-transgenic empty vector control *D. melanogaster* line. "+" represent the transgenic *D. melanogaster* lines containing the house fly P450 genes. *The GelPilot 1Kb (+) ladder (Qiagen Inc, Valencia, CA) was used as the molecular size reference, with the numbers on the figure indicating the DNA band size in bp.


CYP4S24
ALHF MDLLTINSLSLMGLAVLALFLRYLTTILEFIRLIKWSSTVPGPSIGELMENAKKGQILNWFTELREKYGP
aabys MDLLTINSLSLMGLAVLALFLRYLTTILEFIRLIKWSSTVPGPSIGELMENAKKGQILNWFTELREKYGP
Transmembrane (potential)
ALHF TFRIWMGRDLLVFFSDPEDVKQLLSNNSLLHKSRNYSSLEPWLGKGLLTSSGEAWHRRRKLLTPAFHFRI
aabys TFRIWMGRDLLVFFSDPEDVKQLLSNNSLLHKSRNYSSLEPWLGKGLLTSSGEAWHRRRKLLTPAFHFRI
SRS-1
WXXXR motif of C-helix
ALHF LSEFKEPMEENCKILASRLREKCNGEEFDIYPYITLFALDVICETAMGIKKNAQMQSQSEYVRAVQRICH
aabys LSEFKEPMEENCKILASRLREKCNGEEFDIYPYITLFALDVICETAMGIKKNAQMQSQSEYVRAVQRICH
SRS-2
ALHF ILHTKAFSFWQRTNPFFKYTEPGRERDACLRILHDETNRVIKLRRKMLEDAKIQKMADTENLLDNIGKKK
aabys ILHTKAFSFWQRTNPFFKYTEPGRERDACLRILHDETNRVIKLRRKMLEDAKIQKMADTENLLDNIGKKK
SRS-3
ALHF RFAFLDMLLISQMEGVDLTDREIREEVDTFMFEGHDTTSSAIGFAIYLLSQHSEVQQEAFEECMILEGNE
aabys RFAFLDMLLISQMEGVDLTDREIREEVDTFMFEGHDTTSSAIGFAIYLLSQHSEVQQEAFEECMILEGNE
SRS-4
A(A,G)X(E,D)T motif of I-helix
ALHF KENMPYLEAVIKETLRLYPSVPTYARYITEDLQVGKLTLPKGSNISIMTYLTQRDERYFPEPEKFDPQRF
aabys KENMPYLEAVIKETLRLYPSVPTYARYITEDLQVGKLTLPKGSNISIMTYLTQRDERYFPEPEKFDPQRF
SRS-5
PXRF motif
ALHF LQKDKEQHPFSFVAFSAGPRNCIGQKFAMLELKCSLSYLLRNFEFLPVKDFKPNALPELVTKSGNGIKIR
aabys LQKDKEQHPFSFVAFSAGPRNCIGQKFAMLELKCSLSYLLRNFEFLPVKDFKPNALPELVTKSGNGIKIR
heme binding
SRS-6
ALHF IKSRGN
aabys IKSRGN


CYP6A36
ALHF MIALTILLSVLLILIAYIGRLLSEHFNYWRHQGVACEPPNWLLGNLRGLSTSRSFNQIIREYYEKYRNSA
aabys MIALAILLSVLLILIAYIGRLLSEHFNYWRHQGVACEPPNWLLGNLRGLSTSRSFNQIIREYYEKYRNSA
Transmembrane (potential)
ALHF GPFAGFYWLHKKAVFVLDPELIKQILIKDFNKFTDRGFFSNEEDDPLSGQLINLSGNKWRNMRNKLSPAF
aabys GPFAGFYWLHKKAVFVLDPELIKQILIKDFNKFTDRGFFSNEEDDPLSGQLINLSGNKWRNMRNKLSPAF
SRS-1
WXXXR motif of C-helix
ALHF TSGRMKAMFPLVIKLGYDLVEVVGERLKKHDDVVEVRDLAAFFTSDVIGTCAFGLNMNCMKNSDAEFLKM
aabys TSGRMKAMFPLVIKLGYDLVEVVGERLKNHDDVVEVRDLAAFFTSDVIGTCAFGLNMNCMKNSDAEFLKM
ALHF GRKALNEQRYGTLGFVLRFSFPDLCRRLHMKETLDDVEKYFMQIVQETVDYRERENVKRNDFMDMLIDLK
aabys GRKALNEQRYGTLGFVLRFSFPDLCRRLHMKETLDDVEKYFMQIVQETVDYRERENVKRNDFMDMLIDLK
SRS-2
SRS-3
ALHF NNKLIKDESGEEFINLTFGQIAAQAFVFLLAGFETSSTTMSFALYELAQHLEVQQRSREEVENVLKAHNG
aabys NNKLIKDESGEEFINLTFGQIAAQAFVFLLAGFETSSTTMSFALYELAQHLEVQQRAREEVGNVLKAHNG
SRS-4
A(A,G)X(E,D)T motif of I-helix
ALHF TFDYECLKEMVYLEQVIQETLRFYTTIPTINRLASEDYVVGDNPKYVIKKGMAVVIPAAALHRDERYYPQ
aabys TFDYECLKEMVYLEQVIQETLRFYTTIPTINRLASEDYVVGDNPKYVIKKGMAVVIPAAALHRDERYYPQ
EXXR motif of K-helix
SRS-5
ALHF PDVFNPEHFAASQVAERDSVLNLSFGDGPRNCIGMRFGKMQAMIGLALLLKNFRFTVCEKTQIPLKLDKK
aabys PDVFNPEHFAASQVAERDSVLNLSFGDGPRNCIGMRFGKMQAMIGLALLLKNFRFTVCEKTQIPLKLDKK
PXRF motif
heme binding
SRS-6
ALHF SIIVAPESGIYLRVERI
aabys SIIVAPESGIYLRVERI

CYP6A52


α1
α2
α3
α4
1 10 20 30 40 50 60 70
ALHF MVFLTLFIATVLVVVGYLIYRVREHYNYWNRQNIAHDPPHWMLGNFKGLAANLSLGECVQEVYDKYRMTG
aabys MVFLTIFIATVVVVGYLFYRVREHYNYWNRQNIAHDPPHWMLGNFKGLAANLSLGECVQEVYDKYRMTG
Transmembrane (potential)
β1
β2
α5
η1
η2
α6
80 90 100 110 120 130 140
ALHF PFFGFYWFTKPSAVIVDPQLVKRILIKDFSKFSARGMFHNTEDDPLSGQLFNLDGPKWRNMRVQLTPAFT
aabys PFFGFYWFTKPSAVIVDPQLVKRILIKDFSKFSARGMFHNTEDDPLSGQLFNLDGPKWRNMRVQLTPAFT
SRS-1
WXXXR motif of C-helix
α7
α8
α9
150 160 170 180 190 200 210
ALHF SGKMKMIFPTMVKISEEVVQVLREDAGQSQTIEIWDLMARFTSDVIGRCAFGIETNSLKNPKSEFRTMGR
aabys SGKMKIIFPTMVKISEEVVQVLREDVGQSQAIEIWDLMARFTSDVIGRCAFGIETNSLKNPKSEFRTMGR
α10
α11
α12
α13
α14
220 230 240 250 260 270 280
ALHF KGMIEQRGGVLGVAFRLNFPQLAQKLHCKETMPDVEKYFMDLIGETVKYREENNIRCNDFMDMLIDLKNK
aabys KGMIEQRGGVLGVAFRLNFPQLAQRLHCKETMPDVEKYFMDLIGETVKYREENNIRCNDFMDMLIDLKNK
SRS-2
SRS-3
α15
α16
290 300 310 320 330 340 350
ALHF PLMKSDTGEELIRLSFGEIAAQAYLFLLAGYETSASSLAFILYELARHEEIQERVRNEISEVLGKHNQEF
aabys PLMKSDTGEQLIRLSFGEIAAQAYLFLLAGYETSASSLAFILYELARHEEIQEKVRNEISEVLGKHNQEF
SRS-4
A(A,G)X(E,D)T motif of I-helix
α17
α18
β3
β4
α19
η3
360 370 380 390 400 410 420
ALHF SYECMRELTYMEQVINETLRIHSIAPFLNRVASEDYPVPDHPNYVIRKGMPVLIPIEAIQNDERYFPQPK
aabys SYECMRELTYLEQVINETLRIHSIAPFLNRVASEDYPVPDHPKYVIPKGMPVLIPIEAIQHDERYFPQPK
EXXR motif of K-helix
SRS-5
η4
η5
α20
β5
430 440 450 460 470 480 490
ALHF VFDPENFSAEKVAQRDSILHLPFGDGPRNCIGVRFGKMQIFVGLVVLLRNFKFSVCQETPIPMVYSKQGI
aabys VFDPENFSAEKVAQRDSILHLPFGDGPRNCIGVRFGKMQIFVGLVVLLRNFKFSVCQETPIPMVYSKQGI
PXRF motif
heme binding
SRS-6
β6
500
ALHF ILCPGKGVPLKVEKL
aabys ILCPGKGVPLKVEKL

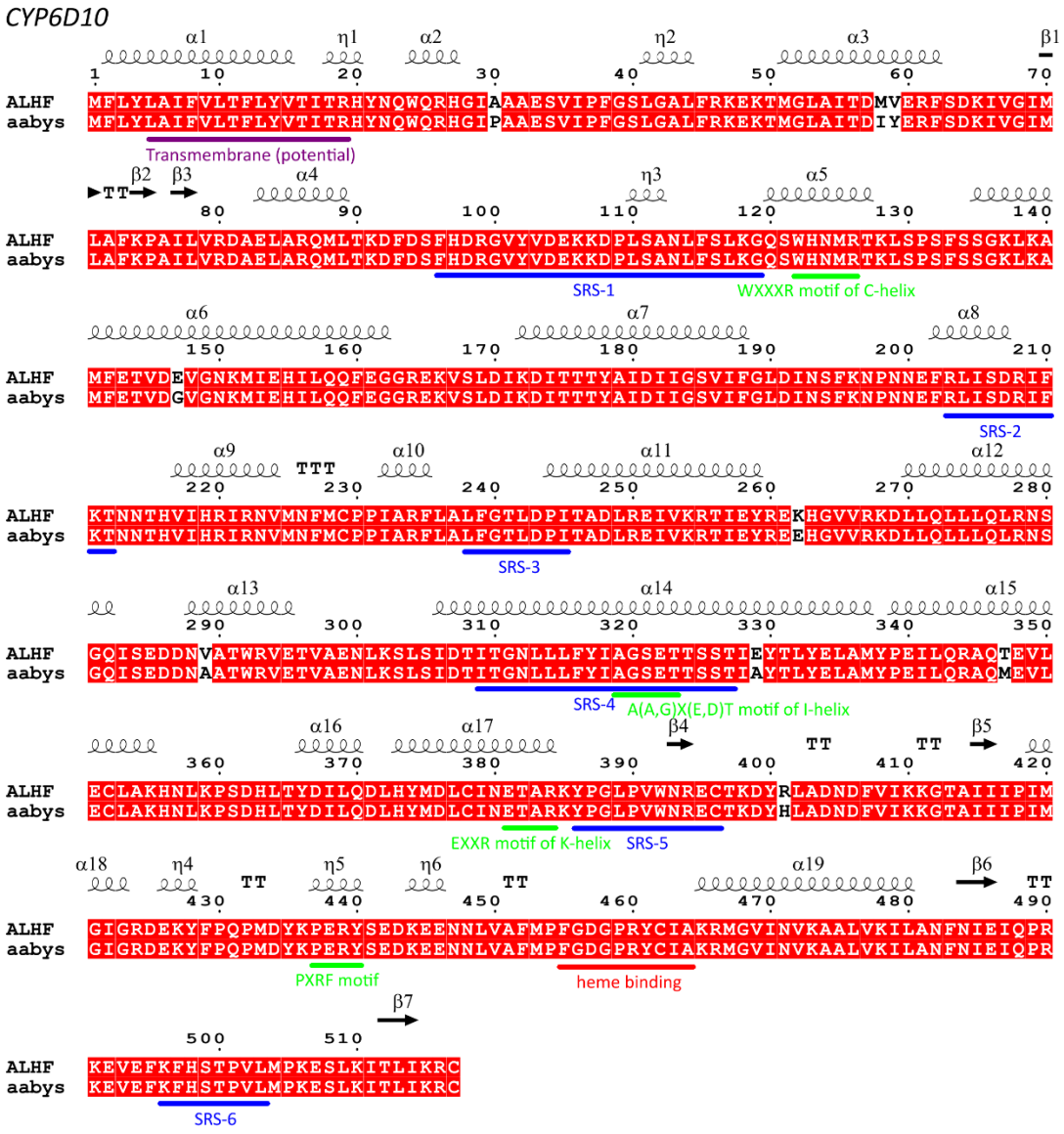


Figure S4.2-5. Deduced amino acid sequence alignment of P450s. The sequence analysis was conducted by T-COFFEE (http://tcoffee.crg.cat/) and ESPript 3.0 (http://espript.ibcp.fr/ESPript/ESPript/) (Gouet et al. 2003). Alph-helices, eta-helices, beta strands and strict beta turns are marked by α, η, β and TT respectively.

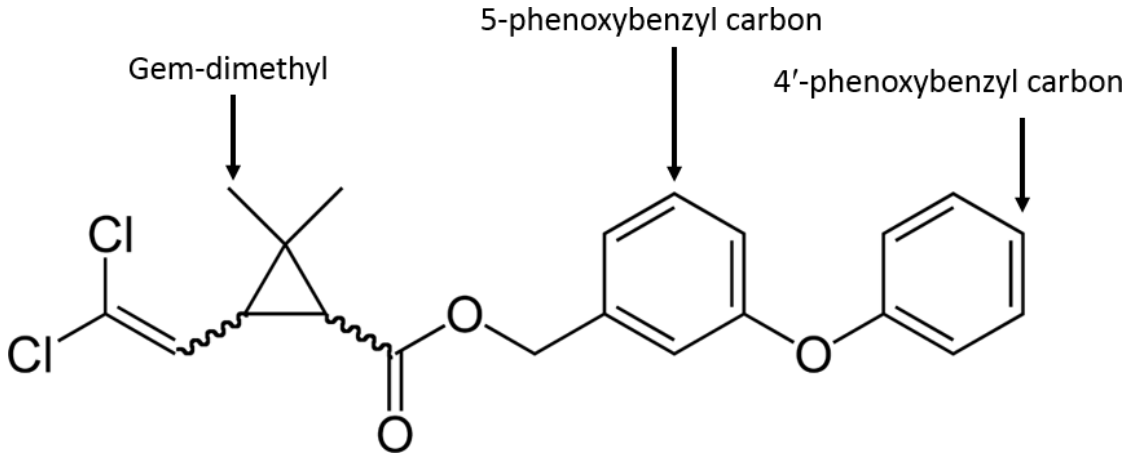


Figure S4.6 Putative metabolic site of permethrin. Arrows point to putative metabolic sites of P450s.

Chapter 5: Molecular and Functional Characterization of Carboxylesterases from the House Fly, *Musca domestica*

5.1 Abstract

To characterize the carboxylesterase genes that play important roles in the pyrethroid resistance of house flies, 26 carboxylesterase genes were selected from the house fly based on our whole transcriptome analysis, to conduct the expression profile analysis in different house fly strains with different levels of permethrin resistance and autosome combinations. Our study showed that 2 carboxylesterase genes, MdαE7 and MdβE2 were co-up-regulated in insecticide-resistant house flies compared to -susceptible house flies, and the expression of these genes was regulated by *cis* or *trans* regulatory factors/genes, which were mainly on autosomes 1, 2 and 5. Transgenic expression analysis of MdβE2 in *Drosophila melanogaster* demonstrated that elevated expression of this gene confers resistance to permethrin in the transgenic *Drosophila*. Homology modeling and

permethrin docking analysis further suggested the potential abilities of MdαE7 and MdβE2 to metabolize permethrin.

5.2 Introduction

The development of insecticide resistance in insect pests is becoming a global challenge and threatening the sustainable use of insecticides. Efforts to characterize the molecular mechanisms involved in insecticide resistance are becoming increasingly important for practical applications, such as the design of novel strategies to prevent or minimize the spread and evolution of resistance development and the control of insect pests (Roush et al. 1990).

Increased metabolic detoxification, which can facilitate insecticide excretion, is one of the significant mechanisms involved in pyrethroid resistance in insects (Zhang et al. 2010). One of the major routes of pyrethroid metabolism is via carboxylesterase-mediated hydrolysis (Huang et al. 2005b). Studies have shown that carboxylesterase gene amplification, up-regulation, coding sequence mutations, or a combination of these mechanisms, are the predominant molecular basis of the esterase-mediated resistance to pyrethroids (Li et al. 2007; Zhang et al. 2010).

Interestingly, a recent study of mosquitoes showed that cytochrome P450 monooxygenases can metabolize PBAlc (3-phenoxybenzoic alcohol) and PBAld (3-phenoxybenzaldehyde), common pyrethroid metabolites produced by carboxylesterases, producing PBA (3-phenoxybenzoic acid). In addition, transcription of these P450 genes was induced by PBAlc, PBAld and PBA (Alexia et al. 2013), indicating potential interaction between carboxylesterases and P450s. Previous studies also showed that expression of some important P450 genes involved in insecticide metabolism were regulated by *trans*- or *cis*-regulatory factors/genes (Carino et al. 1994; Liu and Scott 1996; Maitra et al. 2000; Maitra et al. 1996; Zhu et al. 2008a), suggesting that insecticide resistance development of insects is not only conferred via multiple metabolism genes, but also mediated through the interaction of regulatory factors and metabolism genes (Li et al. 2013). Although no regulatory factors involved in insecticide resistance have yet been identified, these studies inferred that studies of carboxylesterases would be an important intermediary step in characterizing the regulatory genes involved in insecticide resistance. Recent advances in genome/whole transcriptome sequencing technology have provided opportunities for discovering and studying all of the carboxylesterases in one single organism.

The house fly, *Musca domestica*, is an important cosmopolitan pest that causes more than 100 human and animal intestinal diseases, such as cholera, typhoid fever, salmonellosis and polio (Keiding 1986). It is also an effective vector of *Escherichia coli* O157:H7 among cattle, and from cattle to humans (Ahmad et al. 2007; Sasaki et al. 2000). The major barrier to house fly control is their remarkable ability to develop insecticide resistance. Because of their rapid ability to develop resistance and cross-resistance to insecticides, their well described linkage map for five autosomes and two sex chromosomes (X and Y) (Hiroyosh 1960; Hiroyosh 1977; Milani et al. 1967; Nickel and Wagoner 1974; Tsukamoto et al. 1961) and their relatively well studied biochemistry and genetics of insecticide resistance, the house fly has demonstrated to be a useful model to study and predict resistance in not only themselves but also other insect species.

With the availability of the first adult house fly transcriptome and genome database (Li et al. 2013), 26 carboxylesterase genes were identified in the house fly. From this, we conducted an expression profile study of five carboxylesterase genes in different house fly strains with different combinations of autosomes from a resistant house fly strain, ALHF, explored the function of one carboxylesterase gene in insecticide resistance using *Drosophila* transgenic techniques, and conducted modeling and permethrin docking

analyses to investigate the roles of carboxylesterases involved in pyrethroid resistance in house flies.

5.3 Materials and Methods

5.3.1 House fly strains and lines

Two house fly parental strains were used in this study. ALHF is a wild-type insecticide-resistant strain collected from a poultry farm in Alabama in 1998. This strain was further selected with permethrin for six generations after collection to reach a high level of resistance, and has been maintained under biannual selection with permethrin (Liu and Yue 2000; Tian et al. 2011). aabys is an insecticide-susceptible strain with recessive morphological markers ali-curve (*ac*), aristapedia (*ar*), brown body (*bwb*), yellow eyes (*ye*), and snipped wings (*snp*) on autosomes 1, 2, 3, 4, and 5, respectively.

A cross of ALHF female and aabys male was performed with each of ~400 flies, and the F1 males (~400 flies) were then backcrossed with aabys female. Five back-cross (BC_1) lines with the following genotypes were isolated: *ac/ac*, +/*ar*, +/*bwb*, +/*ye*, +/*snp*; +/*ac*, *ar/ar*, +/*bwb*, +/*ye*, +/*snp*; +/*ac*, +/*ar*, *bwb/bwb*, +/*ye*, +/*snp*; +/*ac*, +/*ar*, +/*bwb*, *ye/ye*, +/*snp*; and +/*ac*, +/*ar*, +/*bwb*, +/*ye*, *snp/snp* (Li et al. 2013). Homozygous lines

ac/*ac*, +/+, +/+, +/+, +/+ (A2345); +/+, *ar* /*ar*, +/+, +/+,+/+ (A1345); (+/+, +/+, *bwb*/*bwb*, +/+, +/+ (A1245); +/+, + /+, +/+, *ye*/*ye*, +/+ (A1235) and +/+, +/+, +/+, +/+, *snp*/*snp* (A1234) were accomplished by sorting for appropriate phenotypic markers and selecting with permethrin at a corresponding dose that caused ~70% mortality for each of lines for three generations. One hundred single-pair crossing of each of lines for the desired phenotype and genotype were then set up (Liu and Yue 2000; Tian et al. 2011). The name of each line indicates which of its autosomes bear wild-type markers from ALHF. For instance, the A2345 strain has wild-type markers on autosomes 2, 3, 4 and 5 from ALHF and the mutant marker on autosome 1 from aabys.

5.3.2 Phylogenetic analysis of carboxylesterase genes in *M. domestica*

To classify carboxylesterase genes that identified from the house fly based on our transcriptome and genome analysis, phylogenetic tree was constructed with the carboxylesterases that from the *Drosophila melanogaster* (http://flybase.org/) and *Anopheles gambiae* (https://www.vectorbase.org/). Briefly, the amino acid sequences were analyzed using ClustalW alignment through Molecular Evolutionary Genetic Analysis software version 6 (MEGA 6) (http://www.megasoftware.net/) (Tamura et al.

2013). Subsequently, the alignment result was submitted to construct the phylogenetic tree with neighbor-joining algorithm. A total of 1,000 bootstrap replications were used to test of phylogeny. The house fly carboxylesterases were named based on the major clades that they were clustered.

5.3.3 RNA extraction, cDNA preparation and gene expression detection

A total of 20 3-day old adult female house flies from each of three house fly strains (ALHF, aabys and CS) and five house fly lines (A2345, A1345, A1245, A1235, A1234) were flash frozen on dry ice and immediately processed for RNA extraction. Total RNA (0.5 μg/sample) from each house fly sample was reverse-transcribed using SuperScript II reverse transcriptase (Stratagene) in a total volume of 20 μl. The quantity of cDNAs was measured using a spectrophotometer prior to qRT-PCR, which was performed with the SYBR Green master mix Kit and ABI 7500 Real Time PCR system (AppliedBiosystems). Each qRT-PCR reaction (15 μl final volumes) contained 1× SYBR Green master mix, 1 μl of cDNA, and a gene specific primer pair at a final concentration of 0.3-0.5 μM (Table S5. 1). A 'no-template' negative control and all samples were performed in triplicate. Relative expression levels of specific genes were calculated by the $2^{-\Delta\Delta Ct}$ method using

SDS RQ software (Livak and Schmittgen 2001). The β-actin gene, an endogenous control, was used to normalize expression of target genes (Zhu et al. 2008a). Each experiment was repeated three times with different preparations of RNA samples. The statistical significance of the gene expressions was calculated using Student's t-test for all pairwise sample comparisons and a one-way analysis of variance (ANOVA) for multiple sample comparisons (SPSS v19.0), a value of $P < 0.05$ was considered statistically significant.

5.3.4 Autosomal mapping of selected carboxylesterase genes in *M. domestica*

Five house fly BC_1 lines were used to determine genetic linkage of up-regulated P450 genes. Allele specific PCR was conducted using the cDNA from 5 BC_1 lines (Liu et al. 1995). The ALHF allele specific primer pair was designed based on the specific sequence of the genes from ALHF by placing a specific nucleotide polymorphism at the 3' end of each primer to permit preferential amplification of specific alleles from ALHF. Two rounds of PCR were conducted. For the first PCR reaction, the allele-independent primer pairs (Table S5.1) were used for generating P450 cDNA fragments, respectively. The first PCR solution with cDNA template and a primer pair were heated to 95°C for 3

min, followed by 35 cycles of 95°C for 30 s, 60°C for 30 s, and 72°C for 1 min, then 72°C for 10 min. The second PCR was employed with 0.5 µl of the first round PCR reaction solution and the allele specific primer pair (Table S5.1). The second PCR reaction was heated to 95°C for 3 min, followed by 35 cycles of 95°C for 30 s, 62°C for 30 s, and 72°C for 30 s, then 72°C for 10 min. One of each allele specific primer pair was designed based on the specific sequence of the genes from ALHF by placing a specific nucleotide polymorphism at the 3’ end of the primer to permit preferential amplification of the allele from ALHF. Each experiment was repeated three times with different mRNAs, and the PCR products were sequenced at least once for each gene.

5.3.5 Transgenic expression of candidate carboxylesterase genes in *Drosophila melanogaster* and toxicity of permethrin to the transgenic lines

The full length of candidate carboxylesterase genes were amplified from cDNA of ALHF house fly using Platinum *Taq* DNA polymerase High Fidelity (Invitrogen) with specific primer pairs (Table S5.1) based on the 5' and 3' end sequences of the genes. PCR products were purified using a QIAquick Gel Extraction Kit (Qiagen). The purified PCR products were ligated into pCR 2.1 vector using the Original TA Cloning kit (Invitrogen)

as described by the manufacturer. The full lengths of P450 genes were cloned in One Shot TOPO 10F' cells using the One Shot TOP10F' Chemically Competent *E. coli* kit (Invitrogen). Cloning and sequence analyses were repeated at least three times and three TA clones from each replication were verified by sequencing. The clones were then sub-cloned into a pUASTattB vector (a gift from Dr. Johannes Bischof, University of Zurich). The plasmid of each pUASTattB-up-regulated P450 gene was transformed into the germ line of the M{vas-int.Dm}ZH-2A, M{3xP3-RFP.attP'}ZH-58A strain of *D. melanogaster* (Bloomington stock #24484), resulting in site specific integration on chromosome 2R (Rainbow Transgenic Flies Inc.). Flies were then reciprocally-crossed against a W^{1118} strain to obtain a transgenic line with the orange eye phenotype, then balanced against the *D. melanogaster* balancer strain w[1118]/Dp(1;Y)y[+]; sna[Sco]/CyO, P{ry[+t7.2]=sevRas1.V12}FK1 (Bloomington stock #6312) to generate a homozygous line containing the Cytochrome P450 transgene. The insertion of the up-regulated P450 genes in the transgenic fly lines were further confirmed using RT-PCR. The homozygous lines were crossed with the GAL4-expressing *D. melanogaster* strain P{Act5C-GAL4}17bFO1 (Bloomington stock #3954) which expresses GAL4 under control of the Act5C promoter, resulting in ubiquitous non-tissue-specific expression. The F1

generation of these crosses expressed GAL4 and contained a single copy of the Cytochrome P450 transgene which was under control of the UAS enhancer. The expression of the transgenes in transgenic *Drosophila* flies was confirmed using qRT-RCR. The ribosomal protein L11 (RPL11) of *D. melanogaster* was used as an endogenous control to normalize expression of target genes.

Permethrin toxicity bioassays were then conducted on 3-day posteclosion female *D. melanogaster* of F1 UAS-GAL4 crosses to examine the toxicity of permethrin to transgenic flies. Briefly, a serial concentrations of permethrin solution in acetone, ranging from 3 ng/µL to 150 ng/µL that gave >0 and <100% mortality to the tested flies were prepared, two hundred microliter of each permethrin solution were evenly coated on the inside of individual 20 mL glass scintillation vials. Twenty female flies were transferred to each of the prepared vials, which were plugged with cotton balls soaked with 15% sucrose. The vials for the control groups were coated with acetone alone and plugged with identical 15% sucrose-soaked cotton balls. The mortality was scored after 24 h exposure to permethrin. Flies that did not move were scored as dead. Each bioassay was independently replicated three times using only flies that exhibited the correct morphological marker eyes. The *D. melanogaster* line (Bollmington stock #24484),

which containing the empty pUAST vector donated insert, but no transgene from *M. domestica* were used as the control reference line. Bioassay data were pooled and probit analysis was conducted. Significant difference in the resistance levels of the *D. melanogaster* lines were determined based on non-overlap of 95% confidence intervals (CI). All *D. melanogaster* were reared on Jazz-Mix *D. melanogaster* food (Fisher Scientific, Kansas City, MO) at 25 ± 2 ^{o}C under a photoperiod of 12:12 (L:D) h.

5.3.6 Homology modeling and permethrin docking analysis of selected carboxylesterase

Structural modeling was performed by the I-TASSER server with the combined methods (Roy et al. 2010; Zhang 2008). Multiple models were predicted by the I-TASSER for each carboxylesterase. The top scoring model was submitted to the FG-MD server for fragment guided molecular dynamics structure refinement (Zhang et al. 2011). Model quality was controlled by Ramachandran plots generated with Procheck (http://services.mbi.ucla.edu/SAVES/) (Laskowski et al. 1993) and ProSA-web (https://prosa.services.came.sbg.ac.at/prosa.php) (Sippl 1993; Wiederstein and Sippl 2007). The volume of the substrate binding cavity was characterized by VOIDOO with a

1.4 Å probe (Kleywegt et al. 2001). Proteins and ligands were prepared for docking with Autodock Tools v1.5.6 (http://mgltools.scripps.edu/downloads). Molecular docking was performed by Autodock 4.2. (Morris et al. 2009). Ligand permethrin structures were retrieved from the ZINC database (Irwin et al. 2012). For all dockings, a search space with a grid box of 60 x 60 x 60 Å, was centered at the serine of the catalytic triad of carboxylesterases. The figures were produced by Pymol (http://www.pymol.org/) (DeLano 2002) and Ligplot (Laskowski and Swindells 2011).

5.4 Results

5.4.1 Identification and phylogenetic analysis of carboxylesterase genes in *M. domestica*

A total of 26 putative carboxylesterases was identified from the *M. domestica* based on our transcriptome and genome analysis. The tree showed that a total of 11 clades within three groups of dietary/detoxification, hormone/semiochemical processing, and neuro/developmental functions were characterized (Figure 5.1). Among the 26 identified carboxylesterase genes of the house fly, thirteen were in the dietary detoxification group (α-esterase clade). Eight of the carboxylesterase genes were involved in the

hormone/semiochemical processing, 1 juvenile hormone esterase, 2 β-esterases and 5 integument esterases. Five carboxylesterase genes were located in neuro/developmental functions group, with 2 genes were in the glutactin-like esterase clade and 3 in uncharacterized putative neuroreceptor, acetylcholinesterase and neuroligin clades, respectively. None of the carboxylesterase genes of house flies that belong to the gliotactin, uncharacterized group and neurotactin clades were detected in our study. The accession numbers of these genes are shown in Table S5.1. Again, the names were based on the major clades where they were clustered.

5.4.2 Autosomal location and expression profile of selected carboxylesterase genes in ALHF and aabys *M. domestica* strains

Carboxylesterases in clades of α-esterase and β-esterase have been found to be involved in insecticide resistance in different insects. Based on our transcriptome analysis of house flies, five carboxylesterase genes from these clades were selected for investigation in autosomal location and examination of the relative expression profile in ALHF and aabys house fly strains. Our autosome mapping results showed that the ALHF allele-specific primer sets for these carboxylesterase genes amplified specific DNA

fragments only in flies with the autosome 2 wild-type marker from ALHF (Figure 5.2), which demonstrated that all of these carboxylesterase genes were on autosome 2. Meanwhile, the expression levels of 2 carboxylesterase genes (MdαE7 and MdβE2) were found significantly ($P < 0.05$) up-regulated in the ALHF house flies and MdαE10 was found significantly down-regulated in ALHF house flies compared to aabys house flies (Figure 5.3). No significant difference in the transcriptional level of MdαE1 and MdαE5 was observed between the ALHF and aabys house flies.

5.4.3 Autosome co-regulation of up-regulated carboxylesterase gene expression in insecticide-resistant (ALHF) house flies

We next examined the autosomal linkage of factors from different autosomes to determine the effects of the co-regulation on the expression of the up-regulated MdαE7 and MdβE2 genes among five house fly homozygous lines of A2345, A1345, A1245, A1235 and A1234. Analyzing the gene expression changes resulting from autosome replacement in ALHF house flies enabled us to evaluate the role of genes or factors on each autosome play in carboxylesterase gene overexpression in ALHF house flies. The results showed that when autosomes 2 and 5 of ALHF house flies was replaced by the

corresponding autosome from aabys, respectively (i.e., line A1345 and A1234), the expression of MdαE7 was down-regulated compared to ALHF house flies (Figure 5.4). Meanwhile, when autosomes 1, 2 and 5 of ALHF house flies were replaced by the corresponding autosome from aabys respectively (i.e., line A2345, A1345 and A1234), the expression of MdβE2 was down-regulated compared to ALHF house flies. Suggesting that factors on autosomes 2 and 5 were involved in the expression regulation of MdαE7 genes, and factors on autosomes 1, 2 and 5 were involved in the expression regulation of MdβE2 genes in ALHF house flies.

5.4.4 Functional study of MdβE2 in the transgenic *D. melanogaster*

Overexpression of MdαE7 has been reported to be associated with pyrethroid resistance in a Chinese house fly strain (Zhang et al. 2010). To investigate whether MdβE2 can independently confer resistance to permethrin, transgenic study of MdβE2 using the GAL4-UAS enhancer trap system of *D. melanogaster* was conducted. We first determined the presence of the MdβE2 genes in transgenic lines of *D. melanogaster* by RT-PCR; the result showed that the transgenic line has obtained the target transgene (Figure S5.1). To test whether the GAL4-UAS expressing system of *D. melanogaster* can

increase the expression of the target MdβE2 gene, qRT-PCR was employed to detect the difference in the expression level of three lines: Control (Bloomington stock #24484 *D. melanogaster* line containing the empty pUAST vector), GAL4 (ubiquitous Act5C driver line) and Control + GAL4 (the F1 progeny from the cross between the control females and the GAL4 males), which were not transformed with the MdβE2 recombinant plasmid, and two transgene lines, MdβE2 (MdβE2 homozygous transgene line) and MdβE2+GAL4 (the F1 progeny from the cross between the MdβE2 homozygous transgene line females and the GAL4 males). Our results showed that the expression of MdβE2 was not detected in the non-transgenic *D. melanogaster* lines (Control, GAL4 and Control+GAL4), but was detected in transgenic lines (MdβE2 and MdβE2+GAL4), and the expression of MdβE2 was enhanced by the GAL4-UAS expressing system of *D. melanogaster* (MdβE2+GAL4 > P450) (Figure 5.5).

We next characterized the sensitivities of the non-transgenic and transgenic *D. melanogaster* lines to permethrin. The bioassay results showed that there was no significant difference of permethrin toxicity among the three non-transgenic lines based on the overlapping 95% confidence intervals (Figure 5.5). However, after the MdβE2 was highly expressed in *D. melanogaster*, the permethrin resistance level was significantly

increased compared to non-transgenic lines (Figure 5.5 and S5.2). These results indicated that MdβE2 is capable of conferring permethrin tolerance in *D. melanogaster*, further suggesting that MdβE2 plays important roles in permethrin resistance in *M. domestica*.

5.4.5 Homology modeling and permethrin docking analysis

Homology modeling and permethrin docking studies were conducted to investigate the potential permethrin metabolic detoxification differences between MdαE7 and MdβE2. The deduced amino acid sequence of MdαE7 showed a 27.3% identity with MdβE2. In contrast to the low overall sequence identity, the three dimensional structure shows a high degree (90.6%) of conservation (Figure 5.7). Both MdαE7 and MdβE2 contained conserved carboxylesterase characteristics, such as amphipathic α-helix in the N-terminus that acts as transmembrane anchor, canonical catalytic triad (Serine, glutamate and glycine) and oxyanion hole (Alanine and glycine) (Figure 5.6, S5.2, S5.3). The overall structure of MdαE7 and MdβE2 were both composed of a catalytic domain, a αβ domain and a regulatory domain, respectively. The main differences between MdαE7 and MdβE2 were the channel opening to the catalytic site and the size of the catalytic pocket. Although MdβE2 has a wider substrates channel opening at the entrance of the

catalytic cavity due to the short amino acid chain on the N-terminus, it has a smaller catalytic pocket compared to MdαE7. The substrate binding cavities of these two esterases were extremely asymmetrical, with a small and a large pocket responsible for depositing alkyl group and early/aerial groups of substrate, respectively.

Docking analysis showed that permethrin (1*R*, *trans*) is, as expected, a suitable substrate for MdαE7 and MdβE2. Snugly fitting between the permethrin and catalytic cavities of both MdαE7 and MdβE2 suggested that the active sites of MdαE7 and MdβE2 were ideally shaped for permethrin (Figure 5.7). However, a higher binding affinity with permethrin (-10.94 Kcal/mol) and shorter distance (3.8 Å) from the carbon atom of the ester group of permehtrin to the hydroxyl function serine in MdβE2 was detected compared to that in MdαE7 (-10.1 Kcal/mol, 4.6 Å), indicating a stronger nucleophilic attack and permethrin metabolism ability in MdβE2.

Since permethrin is a hydrophobic molecule, it interacts almost entirely through van der Waals contacts with carboxylesterases. The only exception is the hydrogen bond formed between the carboxyl group of permethrin and arginine in MdαE7 and histidine in MdβE2. Hydrogen bonds with these atoms may stabilize the intermediate that was

formed during hydrolysis (del Loandos et al. 2012; Newcomb et al. 1997; Satoh and Hosokawa 2006).

5.5 Discussion

The importance and interest in carboxylesterases are evident by the increasing number of studies and citations. In this study, 26 putative carboxylesterase genes were identified from the house fly based on our transcriptome and genome database. They were divided into three major subgroups based on catalytic ability and cell/subcellular location. An expansion of the carboxylesterase gene members in the dietary/detoxification group was observed in the house fly. This expansion implies a clue to study the ecological and/or physiological environments adaption strategy of ALHF house flies, such as the house fly may have benefited from this expansion by obtaining ability to metabolize insecticides.

In insects, esterase is one of the major metabolic enzymes that detoxify insecticides in the first phase of metabolism (Huang et al. 2005a; Soderlund 1992). Up-regulation of carboxylesterase gene expression leads to increased levels of total carboxylesterase that can significantly affect the disposition of xenobiotics or endogenous compounds (Bass and Field 2011; Cao et al. 2008; Li et al. 2007). In the current study, we found the

expression of MdαE7 and MdβE2 in the ALHF house fly strain was significantly higher than that in aabys strain. In addition, the overexpression of MdαE7 has also been reported in a beta-cypermethrin resistant house fly strain in China and has been tentatively linked to beta-cypermethrin resistance (Zhang et al. 2010), strongly suggesting a common feature and the important role of these genes in pyrethroid resistance in *M. domestica*. Meanwhile, we also found the expression of MdαE10 was down-regulated in ALHF house flies compared to aabys house flies. The result may not only reveal equally dynamic changes in abundance for both the increased and decreased carboxylesterase expression in resistant house flies, but may also indicate an important feature of the gene regulation system of the house fly in response to environmental challenges.

Characterizing the regulatory genes that regulate the expression of detoxification genes (such as carboxylesterases and cytochrome P450s) presents challenges even today. However, this area of research attracts a great deal of effort since characterization of regulatory genes will provide novel strategies to prevent or minimize the spread and evolution of resistance development thus increasing the control of insect pests. Interestingly, recent studies have found that P450 genes that are involved in pyrethroid resistance development of house flies were on autosome 1, 2, 3 and 5, and the expression

of these genes was regulated by *cis*- or *trans*- factors on autosome 2 (Li et al. 2013; Liu and Scott 1996; Liu et al. 1995; Zhu et al. 2008a; Zhu et al. 2008b; Zhu and Liu 2008). In addition, a study of mosquitoes has found that the transcription of cytochrome P450 monooxygenases related to pyrethroid resistance was induced by PBAlc and PBAld, common pyrethroid metabolites produced by carboxylesterases (Alexia et al. 2013). In our study, we found all of the selected carboxylesterases genes were on autosome 2, and overexpression of MdαE7 in ALHF house flies was regulated by factors on autosomes 2 and 5, overexpression of MdβE2 in ALHF house flies was regulated by factors on autosomes 1, 2 and 5. This information implied there was a potential interaction between these carboxylesterases and P450s in house flies. What is more interesting is expression of both carboxylesterases and P450s have been reported to be regulated by the pregnane X receptor (Goodwin et al. 2002; Xu et al. 2009) or nuclear factor-4 alpha in humans and mice (Furihata et al. 2006; Jover et al. 2001). Taken together, these results indicate a complicated regulatory network exists both between and among insecticide resistance genes and regulatory genes. Characterizing carboxylesterases genes will be an important step in understanding the regulatory network and pathways involved in insecticide resistance development in house flies.

A study has shown that *Drosophila melanogaster* is useful for functionally studying the role of metabolic enzymes, such as carboxylesterase genes, in conferring metabolism-based insecticide resistance (Durban et al. 2012). In our study, a transgene *Drosophila* line with the house fly MdβE2 gene was generated successfully. The important roles played by MdβE2 in permethrin resistance of ALHF house flies were well supported by the transgenic expression of this gene in *D. melanogaster*. Demonstrating the ability of this gene to confer permethrin resistance in ALHF house flies.

Commercial permethrin is formulated as a blend of two pairs of diastereomers (1*R* *trans*-permethrin and 1*S* *trans*-permethrin, 1*R* *cis*-permethrin and 1*S* *cis*-permethrin) in a ratio of approximately 75:25 (Chavasse and Yap 1997; Ross et al. 2006). Insect esterses generally show a selectivity favoring trans isomers of pyrethroids (Heidari et al. 2005). In our study, only the 1R trans-permethrin binding model was observed, suggesting 1*S* *trans*-permethrin, 1*R* *cis*-permethrin and 1*S* *cis*-permethrin may not be the substrates for MdαE7 and MdβE2. Failure of MdαE7 and MdβE2 to effectively metabolize these permethrin isomers may be due to disruption in the oxyanion hole formation (Hemmert and Redinbo 2010).

In conclusion, 26 carboxylesterase genes were identified from *M. domestica* based on our transcriptome and genome analysis. These genes were clustered into different clades in our phylogenetic analysis. Further study showed that 5 selected carboxylesterase genes (MdαE1, MdαE5, MdαE7, MdαE10 and MdβE2) were all on autosome 2. The expression of MdαE7 and MdβE2 was up-regulated in the ALHF house files compared to aabys house flies, and the expression of these two genes was regulated by *trans*- and/or *cis*-acting factors. Functional characterization of MdβE2 in the transgenic *D. melanogaster*, modeling and permethrin docking analyses have lifted the veil on the pivotal roles of MdαE7 and MdβE2 in permethrin resistance of house flies, and opened the avenue for MdαE7 and MdβE2 *in vitro* study. Future study including the identification and characterization of the promoter regions and regulatory factors that are involved in the insecticide resistance development will open the door for new pest control strategies.

5.6 Reference

Ahmad, A, Nagaraja, T and Zurek, L (2007) Transmission of *Escherichia coli* O157: H7 to cattle by house flies. Preventive Veterinary Medicine 80: 74-81.

Alexia, C-P, Jaclyn, B, Myriam, R-K, Jessica, R, Emilie, G-C, Rodolphe, P,

Muhammad, AR, Mark, P, Chantal, D-V and Stephane, R (2013) The central role of mosquito cytochrome P450 CYP6Zs in insecticide detoxification revealed by functional expression and structural modelling. Biochemical Journal 455: 75-85.

Bass, C and Field, LM (2011) Gene amplification and insecticide resistance. Pest Management Science 67: 886-890.

Cao, C-W, Zhang, J, Gao, X-W, Liang, P and Guo, H-L (2008) Overexpression of carboxylesterase gene associated with organophosphorous insecticide resistance in cotton aphids, *Aphis gossypii* (Glover). Pesticide Biochemistry and Physiology 90: 175-180.

Carino, FA, Koener, JF, Plapp, FW, Jr. and Feyereisen, R (1994) Constitutive overexpression of the cytochrome P450 gene CYP6A1 in a house fly strain with metabolic resistance to insecticides. Insect Biochemistry and Molecular Biology 24: 411-418.

Chavasse, DC and Yap, HH (1997) Chemical methods for the control of vectors and pests of public health importance. Chemical methods for the control of vectors and pests of public health importance. 129 pp.

Daborn, PJ, Lumb, C, Harrop, TW, Blasetti, A, Pasricha, S, Morin, S, Mitchell, SN, Donnelly, M, Müller, P and Batterham, P (2012) Using *Drosophila melanogaster* to

validate metabolism-based insecticide resistance from insect pests. Insect Biochemistry and Molecular Biology 42: 918-924.

del Loandos, M, Muro, AC, Villecco, MB, Masman, MF, Luiten, P, Andujar, SA, Suvirec, F and Enriz, RD (2012) Catalytic and molecular properties of rabbit liver carboxylesterase acting on 1, 8-cineole derivatives. Natural Product Communications 7: 1117-1122.

DeLano, WL (2002) The PyMOL molecular graphics system. CA. USA

Furihata, T, Hosokawa, M, Masuda, M, Satoh, T and Chiba, K (2006) Hepatocyte nuclear factor-4alpha plays pivotal roles in the regulation of mouse carboxylesterase 2 gene transcription in mouse liver. Arch Biochem Biophys 447: 107-117.

Goodwin, B, Redinbo, MR and Kliewer, SA (2002) Regulation of CYP3A Gene Transcription by the Pregnane X Receptor. Annual review of pharmacology and toxicology 42: 1-23.

Heidari, R, Devonshire, A, Campbell, B, Dorrian, S, Oakeshott, J and Russell, R (2005) Hydrolysis of pyrethroids by carboxylesterases from Lucilia cuprina and *Drosophila melanogaster* with active sites modified by in vitro mutagenesis. Insect Biochemistry and Molecular Biology 35: 597-609.

Hemmert, AC and Redinbo, MR (2010) A structural examination of agrochemical processing by human carboxylesterase 1. Journal of Pesticide Science 35: 250-256.

Hiroyosh, T (1960) Some New Mutants and Linkage Groups of the House Fly1. Journal of Economic Entomology 53: 985-990.

Hiroyosh, T (1977) Some new mutants and revised linkage maps of the housefly, *Musca domestica* L. Japanese Journal of Genetics 52: 275-288.

Huang, H, Fleming, CD, Nishi, K, Redinbo, MR and Hammock, BD (2005a) Stereoselective hydrolysis of pyrethroid-like fluorescent substrates by human and other mammalian liver carboxylesterases. Chemical Research in Toxicology 18: 1371-1377.

Huang, H, Stok, JE, Stoutamire, DW, Gee, SJ and Hammock, BD (2005b) Development of optically pure pyrethroid-like fluorescent substrates for carboxylesterases. Chemical Research in Toxicology 18: 516-527.

Irwin, JJ, Sterling, T, Mysinger, MM, Bolstad, ES and Coleman, RG (2012) ZINC: a free tool to discover chemistry for biology. Journal of Chemical Information and Modeling 52: 1757-1768.

Jover, R, Bort, R, Gómez-Lechón, MJ and Castell, JV (2001) Cytochrome P450 regulation by hepatocyte nuclear factor 4 in human hepatocytes: A study using

adenovirus-mediated antisense targeting. Hepatology 33: 668-675.

Keiding, J (1986) House-fly; biology and control. OMS.

Kleywegt, G, Zou, J-Y, Kjeldgaard, M and Jones, T (2001) Around O. International Tables for Crystallography Volume F: Crystallography of Biological Macromolecules: 353-356.

Laskowski, RA, MacArthur, MW, Moss, DS and Thornton, JM (1993) PROCHECK: a program to check the stereochemical quality of protein structures. Journal of Applied Crystallography 26: 283-291.

Laskowski, RA and Swindells, MB (2011) LigPlot+: multiple ligand–protein interaction diagrams for drug discovery. Journal of Chemical Information and Modeling 51: 2778-2786.

Li, M, Reid, WR, Zhang, L, Scott, JG, Gao, X, Kristensen, M and Liu, N (2013) A whole transcriptomal linkage analysis of gene co-regulation in insecticide resistant house flies, *Musca domestica*. BMC Genomics 14: 803.

Li, X, Schuler, MA and Berenbaum, MR (2007) Molecular mechanisms of metabolic resistance to synthetic and natural xenobiotics. Annual Review of Entomology 52: 231-253.

Liu, N and Scott, JG (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*. Biochemical Genetics 34: 133-148.

Liu, N, Tomita, T and Scott, JG (1995) Allele-specific PCR reveals that CYP6D1 is on chromosome 1 in the house fly, *Musca domestica*. Cellular and Molecular Life Sciences 51: 164-167.

Liu, N and Yue, X (2000) Insecticide resistance and cross-resistance in the house fly (Diptera: Muscidae). Journal of Economic Entomology 93: 1269-1275.

Livak, KJ and Schmittgen, TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25: 402-408.

Maitra, S, Dombrowski, SM, Basu, M, Raustol, O, Waters, LC and Ganguly, R (2000) Factors on the third chromosome affect the level of Cyp6a2 and Cyp6a8 expression in *Drosophila melanogaster*. Gene 248: 147-156.

Maitra, S, Dombrowski, SM, Waters, LC and Ganguly, R (1996) Three second chromosome-linked clustered Cyp6 genes show differential constitutive and barbital-

induced expression in DDT-resistant and susceptible strains of *Drosophila melanogaster*. Gene 180: 165-171.

Milani, R, Rubini, P and Franco, M (1967) Sex determination in the housefly. Genetica Agraria 21: 385-411.

Morris, GM, Huey, R, Lindstrom, W, Sanner, MF, Belew, RK, Goodsell, DS and Olson, AJ (2009) AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. Journal of Computational Chemistry 30: 2785-2791.

Newcomb, R, Campbell, P, Ollis, D, Cheah, E, Russell, R and Oakeshott, J (1997) A single amino acid substitution converts a carboxylesterase to an organophosphorus hydrolase and confers insecticide resistance on a blowfly. Proceedings of the National Academy of Sciences 94: 7464-7468.

Nickel, C and Wagoner, D (1974) Mutants on linkage groups 3 and 4 of the house fly. Annals of the Entomological Society of America 67: 775-776.

Ross, MK, Borazjani, A, Edwards, CC and Potter, PM (2006) Hydrolytic metabolism of pyrethroids by human and other mammalian carboxylesterases. Biochemical pharmacology 71: 657-669.

Roush, RT, Hoy, CW, Ferro, DN and Tingey, WM (1990) Insecticide Resistance in

the Colorado Potato Beetle (Coleoptera: Chrysomelidae): Influence of Crop Rotation and Insecticide Use. Journal of Economic Entomology 83: 315-319.

Roy, A, Kucukural, A and Zhang, Y (2010) I-TASSER: a unified platform for automated protein structure and function prediction. Nature protocols 5: 725-738.

Sasaki, T, Kobayashi, M and Agui, N (2000) Epidemiological potential of excretion and regurgitation by *Musca domestica* (Diptera: Muscidae) in the dissemination of *Escherichia coli* O157: H7 to food. Journal of Medical Entomology 37: 945-949.

Satoh, T and Hosokawa, M (2006) Structure, function and regulation of carboxylesterases. Chemico-Biological Interactions 162: 195-211.

Sippl, MJ (1993) Recognition of errors in three-dimensional structures of proteins. Proteins: Structure, Function, and Bioinformatics 17: 355-362.

Soderlund, D (1992) Metabolic considerations in pyrethroid design. Xenobiotica 22: 1185-1194.

Tamura, K, Stecher, G, Peterson, D, Filipski, A and Kumar, S (2013) MEGA6: molecular evolutionary genetics analysis version 6.0. Molecular Biology and Evolution 30: 2725-2729.

Tian, L, Cao, C, He, L, Li, M, Zhang, L, Liu, H and Liu, N (2011) Autosomal

interactions and mechanisms of pyrethroid resistance in house flies, *Musca domestica*. International Journal of Biological Sciences 7: 902-911.

Tsukamoto, M, Baba, Y and Hiraga, S (1961) Mutations and linkage groups in Japanese strains of the housefly. Japanese Journal of Genetics 36: 168-174.

Wiederstein, M and Sippl, MJ (2007) ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Researchearch 35: W407-W410.

Xu, C, Wang, X and Staudinger, JL (2009) Regulation of tissue-specific carboxylesterase expression by pregnane x receptor and constitutive androstane receptor. Drug Metabolism and Disposition 37: 1539-1547.

Zhang, J, Liang, Y and Zhang, Y (2011) Atomic-level protein structure refinement using fragment-guided molecular dynamics conformation sampling. Structure 19: 1784-1795.

Zhang, L, Shi, J, Shi, X, Liang, P, Gao, J and Gao, X (2010) Quantitative and qualitative changes of the carboxylesterase associated with beta-cypermethrin resistance in the housefly, *Musca domestica* (Diptera: Muscidae). Comparative biochemistry and physiology. Part B, Biochemistry & Molecular Biology 156: 6-11.

Zhang, Y (2008) I-TASSER server for protein 3D structure prediction. BMC Bioinformatics 9: 40.

Zhu, F, Feng, JN, Zhang, L and Liu, N (2008a) Characterization of two novel cytochrome P450 genes in insecticide-resistant house-flies. Insect Molecular Biology 17: 27-37.

Zhu, F, Li, T, Zhang, L and Liu, N (2008b) Co-up-regulation of three P450 genes in response to permethrin exposure in permethrin resistant house flies, *Musca domestica*. BMC Physiol 8: 18.

Zhu, F and Liu, N (2008) Differential expression of CYP6A5 and CYP6A5v2 in pyrethroid-resistant house flies, *Musca domestica*. Arch Insect Biochem Physiol 67: 107-119.

Table S5.1 List and sequences of the primers used.

Gene	Forward primer (5'-3')	Reverse primer (5'-3')
Primers used for quantitative real time PCR		
MdβE2	GGCTTTTAGTACATCTGCTCGG	GCACTGGGTAGTGGAAGATTAG
MdαE7	CCACTGTAGAACCATACCAGACA	CCAACTCCCAAGGCACATA
MdαE5	AAGACACGTCACTGCTAACC	CCAGGCCGAAGAAAGGTATT
MdαE10	CCTCTACACATTTCCTAACAGCCA	GCTTGTCCTCGTCATTGAGAA
MdαE1	GCATTCCATTGCTCATTGG	TGCCGTGCCAGAAGTATTTA
Actin	ATGAGGCTCAGAGCAAACGTGGTA	AGTCATCTTCTCGCGATTGGCCTT
Dm RPL11	CGATCCCTCCATCGGTATCT	AACCACTTCATGGCATCCTC
Primers used for autosome mapping		
MdβE2	F1:TTGAAATGTCCCAATTTGGA	AGAGCATATCCCAAACTATAATC
	F2:TTATGACTCGGCATCCAAGA	
MdαE7	F1:GTTTGGGTGTGTTGGGTTTC	TTCACTATGGCAGCCCTTTC
	F2:CGGTAATTCCATGTGCTCATT	
MdαE5	F1:ACTATTCGGAGAGAGTGCCGG	AATAATAATAGGTGGGTGTT
	F2:AAGAAGAACAATTCAATCATCTA	
MdαE10	F1:CCCGGCAATGCTGGTATCAAAGA	AACCTCGACATCCTTATTTGC
	F2:AATGGGTTAAGCAATACATC	
MdαE1	F1:AAACATCTTCTCCGGTCTGTG	TTGACGGCAATTCGCATTTGAT
	F2:CTGACCGGTCGGTCACAT	
Primers used for functional study		
MdβE2	CCGGAATTCCAAAATGAAATTTAAACTTTTCG	CTAGCTCGAGGACAACAATGGGTTTACA
Primers used for full length amplification		
MdβE2	ATGAAATTTAAACTTTTCGTATTGGG	TTAGACAACAATGGGTTTACAA

Table S5.2 List of carboxylesterase genes in *Musca domestica*

Clade	Accession number	Gene name
α-esterase	AFP60546	MdαE1
	AFP63016	MdαE2
	AFP62560	MdαE3
	AFP63126	MdαE4
	AFP62322	MdαE5
	AFP62813	MdαE6
	AFP62162	MdαE7*
	AFP63133	MdαE8
	AFP62337	MdαE9
	AFP62145	MdαE10
	AFP63619	MdαE11
	XP_005175450	MdαE12
	XP_005174776	MdαE13
Juvenile hormone esterase	XP_005181511	MdJhe1
β-esterase	XP_005181018	MdβE1
	XP_005183940	MdβE2
Integument esterase	XP_005180750	MdIntE1
	XP_005180749	MdIntE2
	XP_005180748	MdIntE3
	XP_005180752	MdIntE4
	XP_005177449	MdIntE5
Glutactin	AFP62929	MdGluE1
	AFP64249	MdGluE2
Uncharacterized putative neuroreceptor	XP_0051885078	MdUucE1
Acetylcholinesterase	AF281161	MdAceE1*
Neuroligin	XP_005185445	MdNeuE1

* Indicated gene has been reported in other studies.

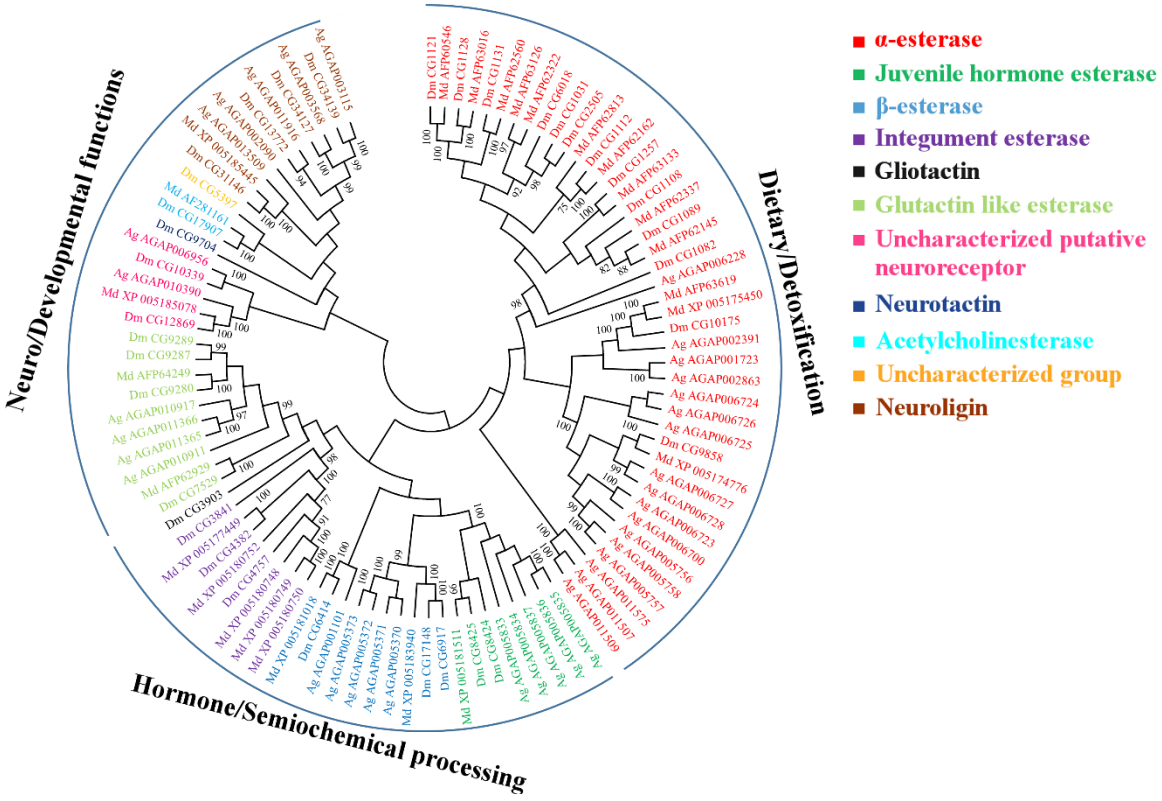


Figure 5.1 Neighbor-joining consensus trees of carboxylesterases. MUSCLE software was used to perform multiple sequence alignment (Edgar, 2004). The phylogenetic tree was generated by MEGA 6.0 using the amino acid sequences from *Drosophila melanogaster* (Dm) and *Anopheles gambiae* (Ag). Sequences of *Drosophila melanogaster* were from flybase (http://flybase.org/), and sequences of *Anopheles gambiae* were from Vectorbase (https://www.vectorbase.org/). Distance bootstrap values >70% (1000 replicates) are shown at the tree branches. The nomenclatures of the clades were according to Oakeshott et al.

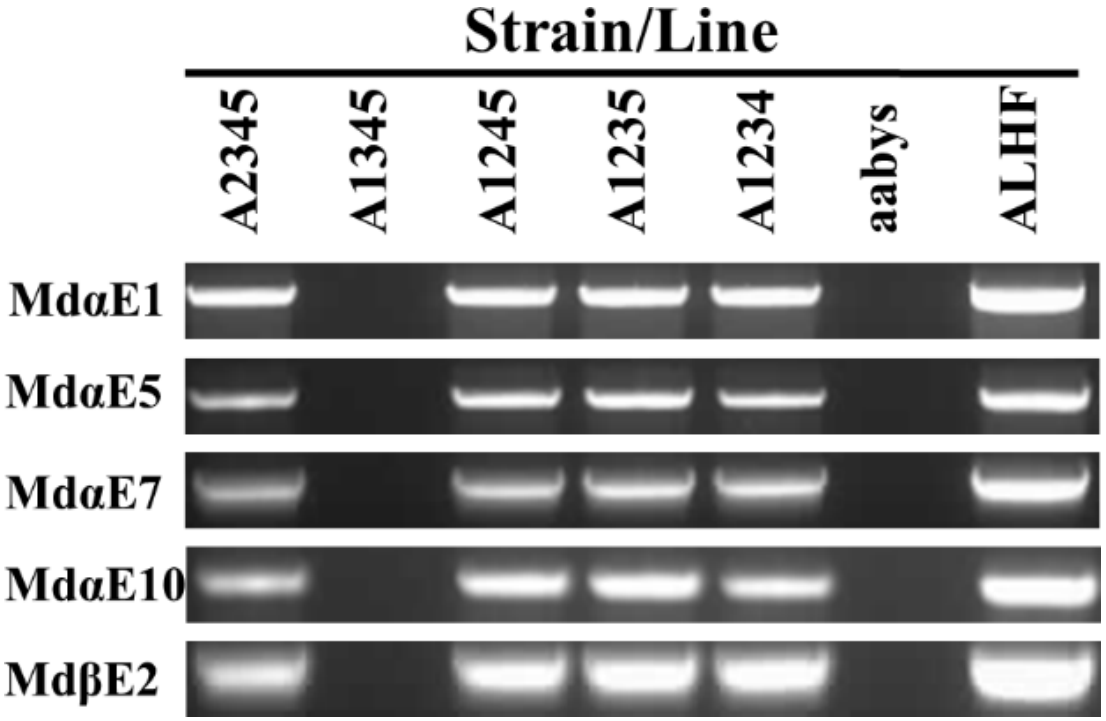


Figure 5.2 Allele-specific RT-PCR autosomal mapping of the *Musca domestica* carboxylesterase genes. PCR fragments were generated using the allele-specific primer set according to the sequence of each gene from ALHF. The absence of a PCR product in a house fly line indicated that the gene was located on the corresponding autosome from aabys (i.e. the absence of a band in the A1345 line indicates that the gene was present on autosome 2).

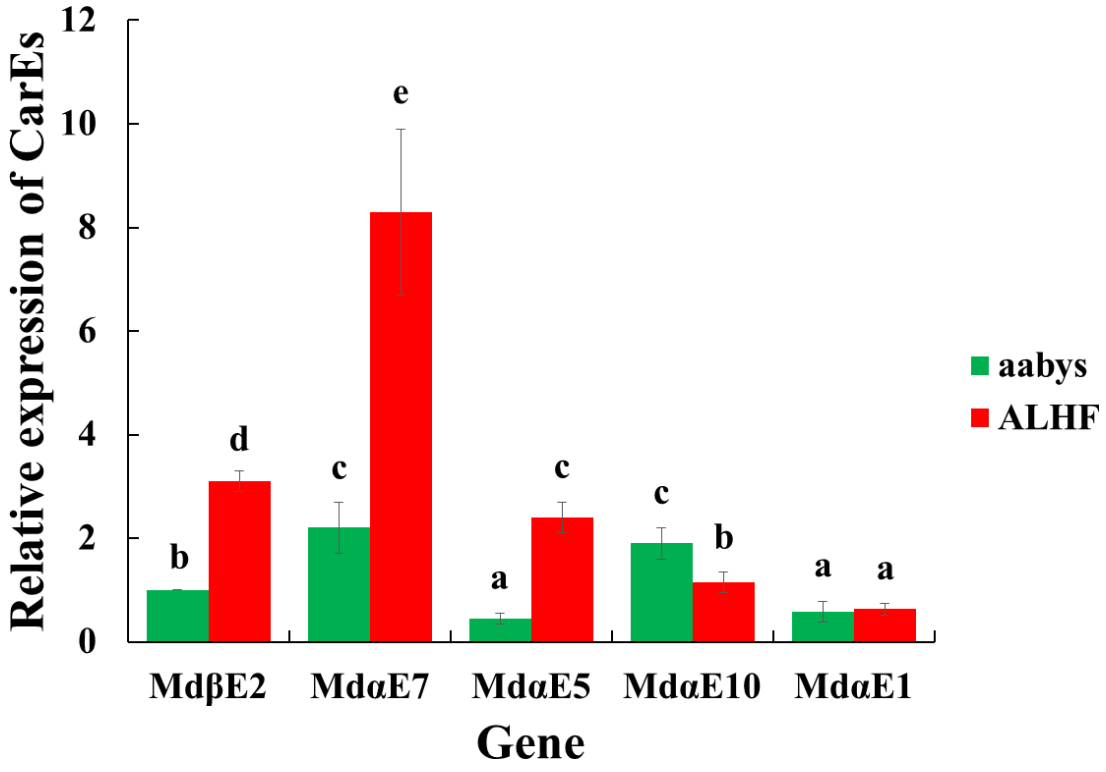


Figure 5.3 Relative expression of selected carboxylesterases in aabys and ALHF *Musca domestica* strains. The relative gene expression ratios were calculated by comparing the expression of MdβE2 in aabys strain. Data were shown as the mean ± SEM (n = 3). There was no significant difference in the expression level designated with the same letter ($P < 0.05$).

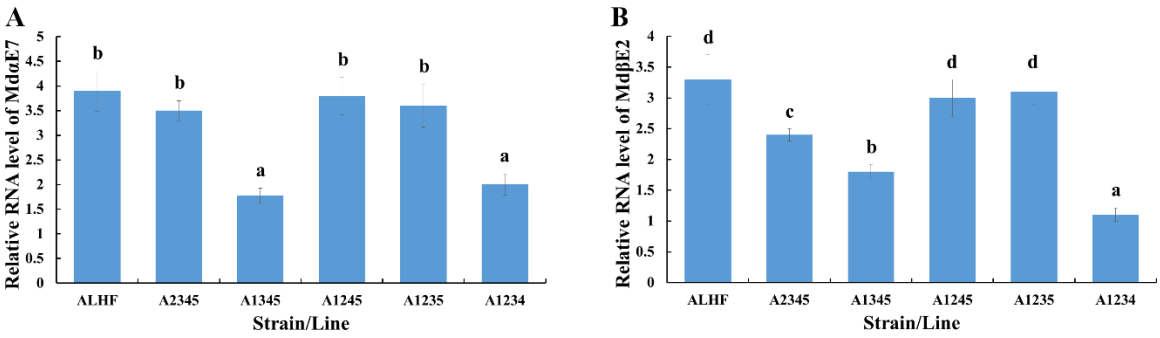


Figure 5.4 Relative expression of MdαE7 and MdβE2 in ALHF and five house fly homozygous lines. The relative expression levels of MdαE7 are shown as a ratio in comparison with that in aabys strain. The results are shown as the mean ± SEM (n = 3). There was no significant difference in the expression level among samples designated with the same letter ($P < 0.05$). (A) MdαE7. (B) MdβE2.

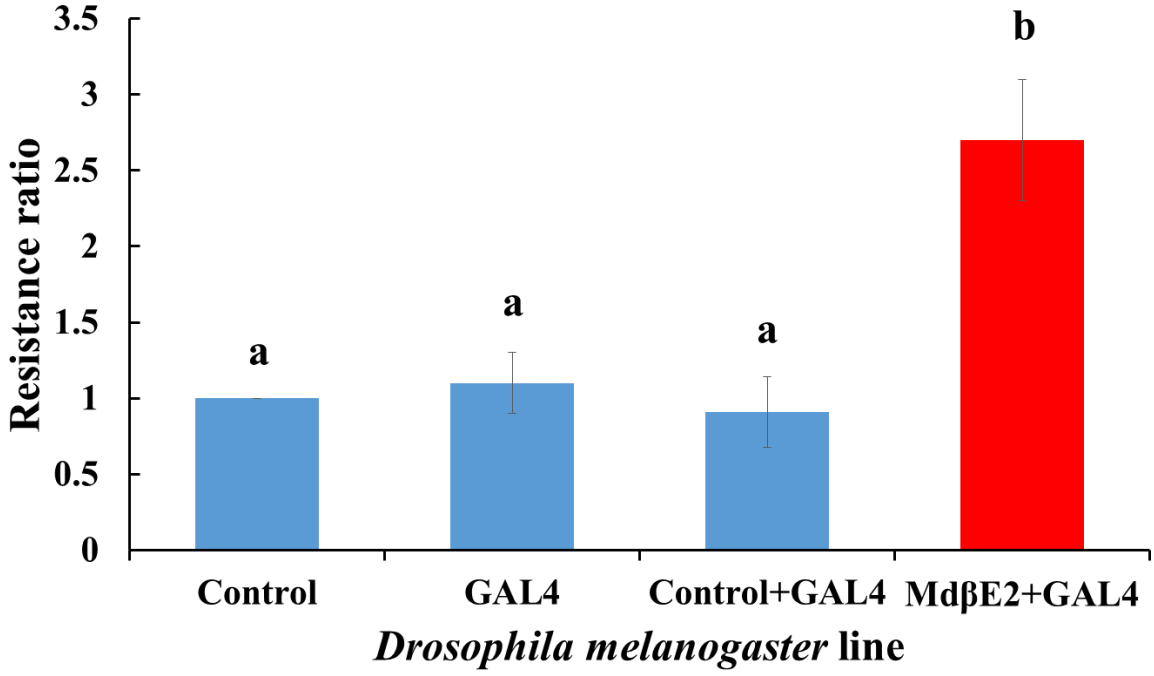


Figure 5.5 Toxicity of permethrin to non-transgenic and transgenic *Drosophila melanogaster* lines. Resistance ratios = LC_{50} of *Drosophila melanogaster* lines/LC_{50} of Control line. There was no significant difference in the resistance level among samples designated with the same ($P < 0.05$).

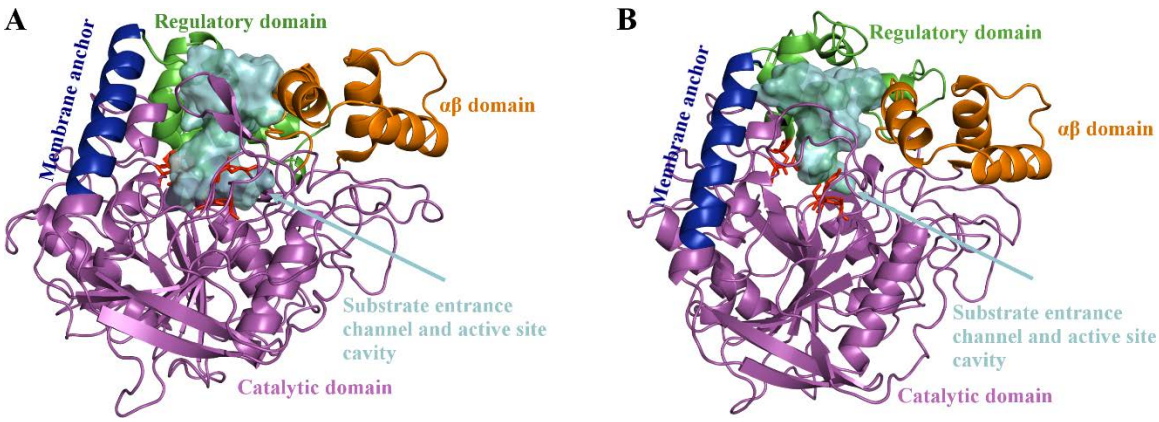


Figure 5.6 Overall structure of MdαE7 and MdβE2. Membrane anchor, Regulatory domain, αβ domain, catalytic domain, substrate entrance channel and active site are colored blue, green, orange, purple and cyan, respectively. Residues of the catalytic triad (serine, glutamate and histidine) and oxyanion hole (alanine and glycine) are shown as red sticks.

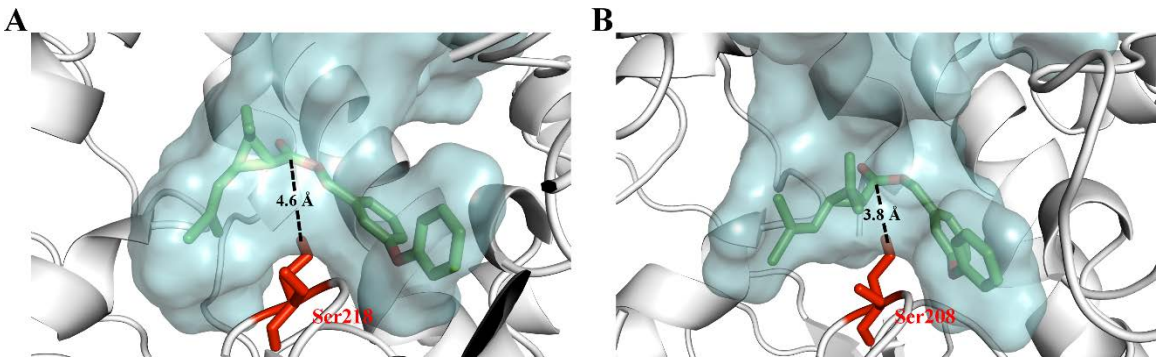


Figure 5.7 Stereo view of permethrin bound within the active site cavity of MdαE7 and MdβE2. The active site cavity and substrate channel are represented in cyan. Serine is showed in sticks and labeled in red. Permethrin is showed in green sticks. The dashed line represents the distance between the metabolic site of permethrin and the Serine of CarEs. (A) Permehtrin bound within the active site cavity of MdαE7. (B) Permehtrin bound within the active site cavity of MdβE2.

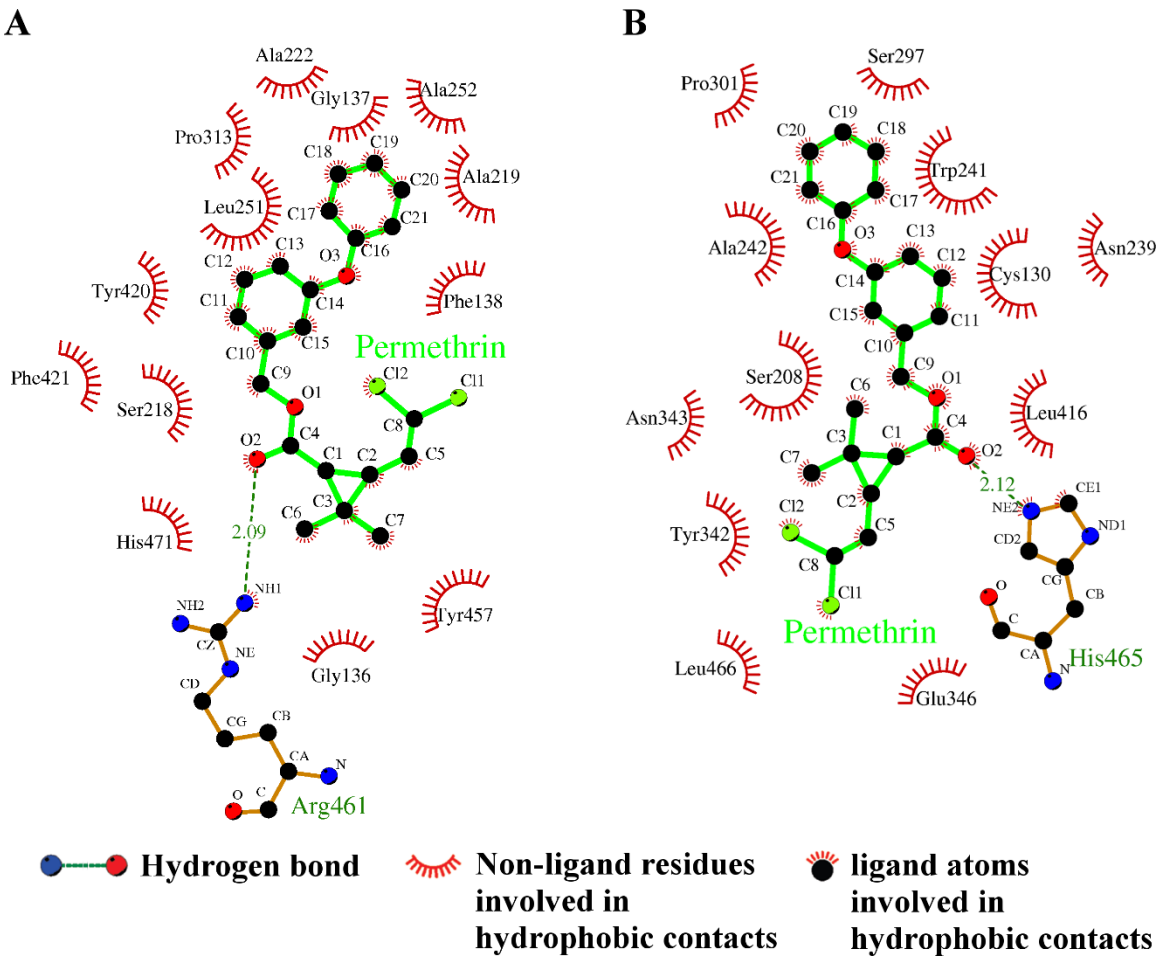


Figure 5.8 Ligplot representation of the contacts between permethrin and carboxylesterases. (A) MdαE7. (B) MdβE2.

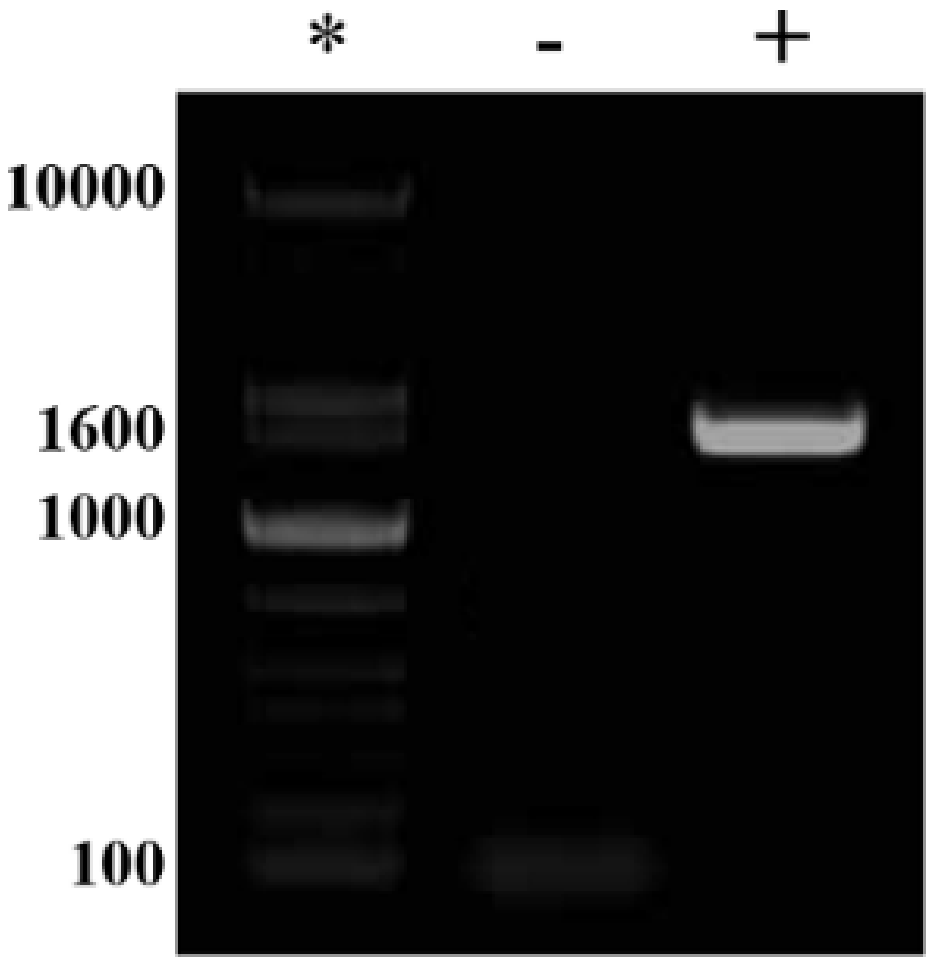


Figure S5.1 RT-PCR amplification of the transgenic *Drosophila melanogaster* lines using gene specific primers. The lane indicated with "-" represent the non-transgenic empty vector control *D. melanogaster* line. "+" represent the transgenic *D. melanogaster* line containing the MdβE2 gene. * The GelPilot 1Kb (+) ladder (Qiagen Inc, Valencia, CA) was used as the molecular size reference, with the numbers on the figure indicating the DNA band size in bp.

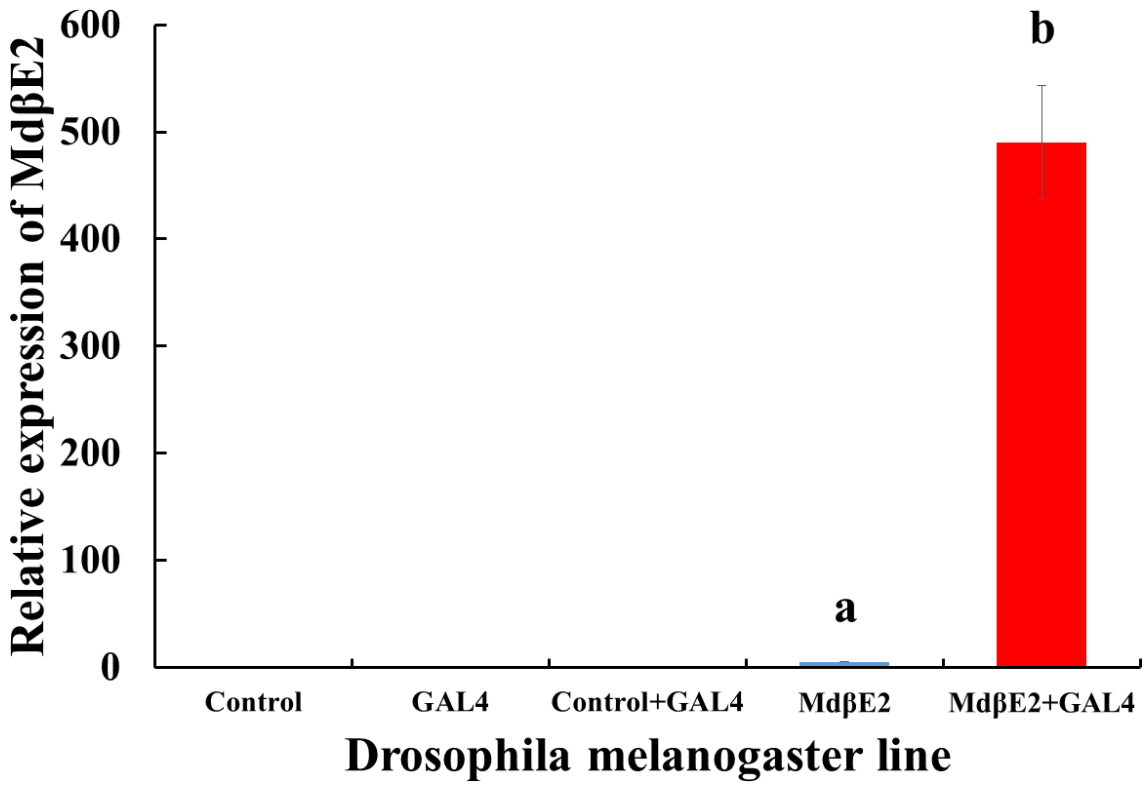
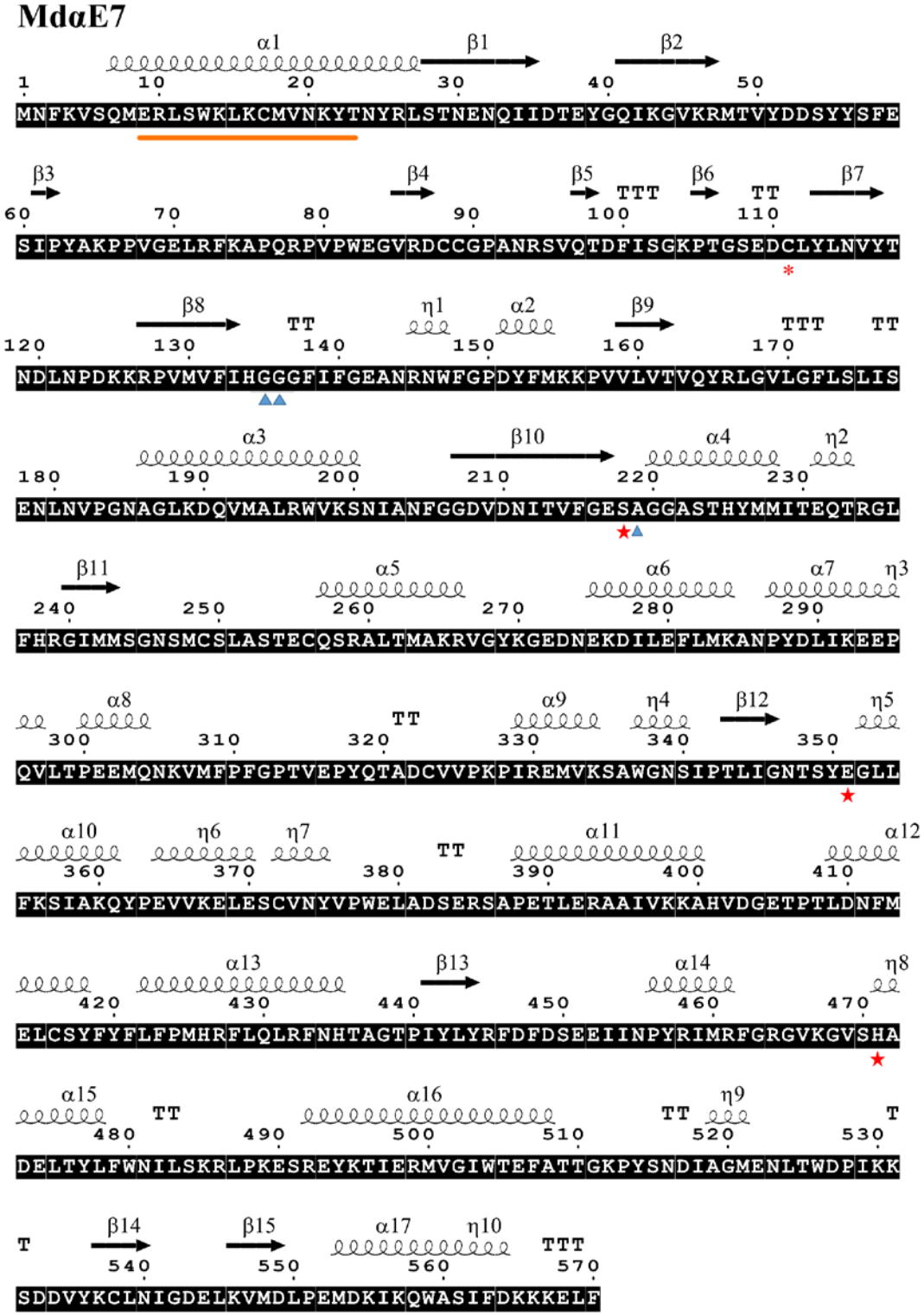


Figure S5.2 Transgenic expression of MdβE2 in *Drosophila melanogaster*. The relative expression of the transgene was quantified by qRT-PCR. "MdβE2" represents the homozygous transgenic *Drosophila* line with house fly MdβE2 genes, "MdβE2+GAL" represents the F1 generation of homozygous transgenic *Drosophila* line crossed with GAL4 driver line. Data shown are mean + SEM (n = 3). Different letters above the bars indicate significant differences among different lines ($P < 0.05$).


MdαE7
α1 β1 β2
1 10 20 30 40 50
MNFKVSQMERLSWKLKCMVNKYTNYRLSTNENQIIDTEYGQIKGVKRMTVYDDSYYSFE
β3 β4 β5 TTT β6 TT β7
60 70 80 90 100 110
SIPYAKPPVGELRFKAPQRPVPWEGVRDCCGPANRSVQTDFISGKPTGSEDCLYLNVYT
β8 TT η1 α2 β9 TTT TT
120 130 140 150 160 170
NDLNPDKKRPVMVFIHGGGFIFGEANRNWFGPDYFMKKPVVLVTVQYRLGVLGFLSLIS
α3 β10 α4 η2
180 190 200 210 220 230
ENLNVPGNAGLKDQVMALRWVKSNIANFGGDVDNITVFGESAGGASTHYMMITEQTRGL
β11 α5 α6 α7 η3
240 250 260 270 280 290
FHRGIMMSGNSMCSLASTECQSRALTMAKRVGYKGEDNEKDILEFLMKANPYDLIKEEP
α8 TT α9 η4 β12 η5
300 310 320 330 340 350
QVLTPEEMQNKVMFPFGPTVEPYQTADCVVPKPIREMVKSAWGNSIPTLIGNTSYEGLL
α10 η6 η7 TT α11 α12
360 370 380 390 400 410
FKSIAKQYPEVVKELESCVNYVPWELADSERSAPETLERAAIVKKAHVDGETPTLDNFM
α13 β13 α14 η8
420 430 440 450 460 470
ELCSYFYFLFPMHRFLQLRFNHTAGTPIYLYRFDFDSEEIINPYRIMRFGRGVKGVSHA
α15 TT α16 TT η9 T
480 490 500 510 520 530
DELTYLFWNILSKRLPKESREYKTIERMVGIWTEFATTGKPYSNDIAGMENLTWDPIKK
T β14 β15 α17 η10 TTT
540 550 560 570
SDDVYKCLNIGDELKVMDLPEMDKIKQWASIFDKKKELF

MdβE2

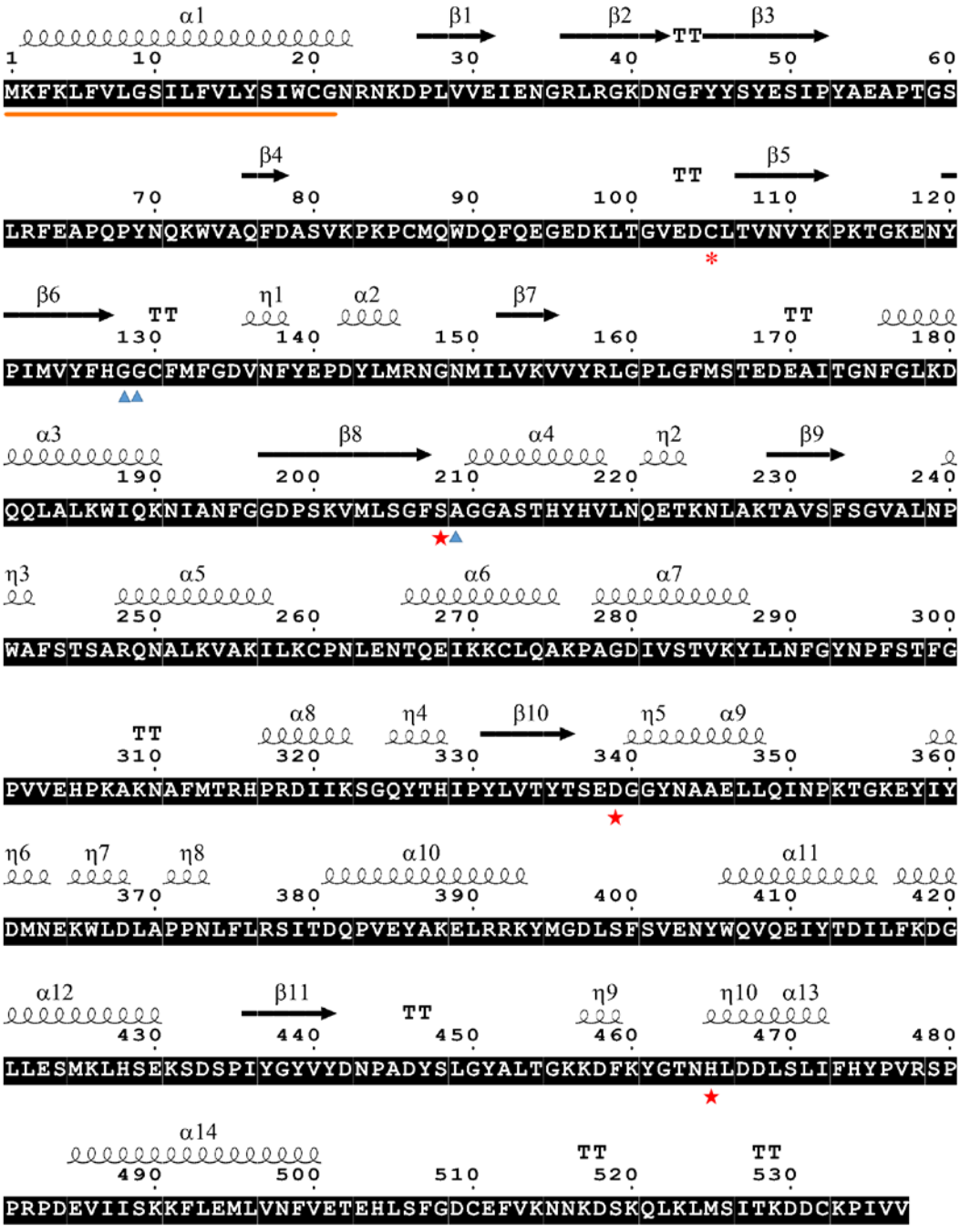

α1 β1 β2 TT β3
1 10 20 30 40 50 60
MKFKLFVLGSILFVLYSIWCGNRNKDPLVVEIENGRLRGKDNGFYYSYESIPYAEAPTGS
β4 TT β5
70 80 90 100 110 120
LRFEAPQPYNQKWVAQFDASVKPKPCMQWDQFQEGEDKLTGVEDCLTVNVYKPKTGKENY
β6 TT η1 α2 β7 TT
130 140 150 160 170 180
PIMVYFHGGCFMFGDVNFYEPDYLMRNGNMILVKVVYRLGPLGFMSTEDEAITGNFGLKD
α3 β8 α4 η2 β9
190 200 210 220 230 240
QQLALKWIQKNIANFGGDPSKVMLSGFSAGGASTHYHVLNQETKNLAKTAVSFSGVALNP
η3 α5 α6 α7
250 260 270 280 290 300
WAFSTSARQNALKVAKILKCPNLENTQEIKKCLQAKPAGDIVSTVKYLLNFGYNPFSTFG
TT α8 η4 β10 η5 α9
310 320 330 340 350 360
PVVEHPKAKNAFMTRHPRDIIKSGQYTHIPYLVTYTSEDGGYNAAELLQINPKTGKEYIY
η6 η7 η8 α10 α11
370 380 390 400 410 420
DMNEKWLDLAPPNLFLRSITDQPVEYAKELRRKYMGDLSFSVENYWQVQEIYTDILFKDG
α12 β11 TT η9 η10 α13
430 440 450 460 470 480
LLESMKLHSEKSDSPIYGYVYDNPADYSLGYALTGKKDFKYGTNHLDDLSLIFHYPVRSP
α14 TT TT
490 500 510 520 530
PRPDEVIISKKFLEMLVNFVETEHLSFGDCEFVKNNKDSKQLKLMSITKDDCKPIVV

Figure S5.3-4 Deduced amino acid sequence of MdαE7 and MdβE2. The sequence analysis was conducted by T-COFFEE (http://tcoffee.crg.cat/) and ESPript 3.0 (http://espript.ibcp.fr/ESPript/ESPript/) (Gouet et al. 2003). Alph-helices, eta-helices, beta strands and strict beta turns are marked by α, η, β and TT respectively. * represents N-terminal conserved cysteine for disulfide bond. The catalytic triads are marked with red star. Residues for oxyanion hole are marked with blue triangle. The predicted signal peptide is underlined in orange. Signal peptide of each protein was predicted by using SignalP4.1 web tools (http://www.cbs.dtu.dk/services/SignalP/).

Chapter 6: Research Summary and Future Studies

6.1 Research summary

My doctoral research projects were mainly focused on the mechanisms that involve in the pyrethroid insecticide resistance development in the house fly, *Musca domestica*. Specifically, my work focused on the investigation and characterization of metabolic genes, such as cytochrome P450 genes and carboxylesterase genes that were involved in the development of insecticide resistance and their expressional regulation in response to insecticide stimulation in the house fly, *Musca domestica*. I conducted the transcriptome analysis of the house fly, generated the first adult house fly transcriptome database. This database contained 14488 contigs, and 43% of the contigs contained coding regions, among which 1316 genes were identified as being co-up-regulated in insecticide resistant house fry strain (ALHF) in comparison to two insecticide susceptible strains (aabys and CS). The majority of these up-regulated genes fell within the SCOP categories of metabolism, general, intra-cellular processes, and regulation, and covered three key

detailed functional categories: redox detailed function category in metabolism, signal transduction and kinases/phosphatases in regulation, and proteases in intra-cellular processes. The redox group contained detoxification gene superfamilies, including cytochrome P450s, glutathione S-transferases, and esterases. The signal transduction and kinases/phosphatases groups contained gene families of rhodopsin-like GPCRs, adenylate and guanylate cyclases, protein kinases and phosphatases. The proteases group contained genes with digestive, catalytic, and proteinase activities. Genetic linkage analysis with house fly lines comparing different autosomal combinations from ALHF revealed that the up-regulation of gene expression in the three key SCOP detailed function categories occurred mainly through the co-regulation of factors among multiple autosomes, especially between autosomes 2 and 5, suggesting that signaling transduction cascades controlled by GPCRs, protein kinase/phosphates and proteases may be involved in the regulation of resistance P450 gene regulation. Based on these findings, I then concentrated my research on two important mechanisms, which are cytochrome P450-mediated and carboxylesterase-mediated detoxification. I examined the interaction of the resistant key P450 genes and factors on different autosomes through house fly lines with different combination of autosomes from a resistant house fly strain, ALHF; exploring

the function of key P450 genes in insecticide resistance using *Drosophila* transgenic techniques; and analyzing P450 modeling and permethrin docking to investigate the roles of P450s that are involved in pyrethroid resistance in house flies. My findings demonstrated that P450 constitutive over-expression or induction responsible for permethrin resistance in house flies, and these P450 genes are regulated by *trans*- and/or *cis*-acting factors. Our results also confirm and support the co-overexpression of multiple P450 genes as likely to be key factors enhancing permethrin resistance in house flies. Suggesting that beside detoxification, other roles of P450 may also involve in permethrin resistance in house flies. Meanwhile, a total of 26 carboxylesterase genes was identified from the house fly. I conducted the expression profile study of five carboxylesterase genes in different house fly strains with different combination of autosomes from a resistant house fly strain, ALHF; explored the function of one carboxylesterase gene in insecticide resistance using *Drosophila* transgenic techniques; and conducted modeling and permethrin docking analyses to investigate the roles of carboxylesterases that are involved in pyrethroid resistance in house flies. The results showed that The expression of two carboxylesterase gene was up-regulated in ALHF house files compared to aabys house flies, and the expression of these two genes was regulated by *trans*- and/or *cis*-

acting factors. Functional characterization of one gene in the transgenic *D. melanogaster*, modeling and permethrin docking analyses indicated that this gene is capable of conferring permethrin tolerance in *M. domestica*. These results from my dissertation research clearly demonstrated that multiply mechanisms co-confer insecticide resistance, and up-regulation of P450 genes and carboxylesterase genes play important roles in insecticide detoxification. A complicated regulatory network between insecticide resistance genes and regulatory genes, among resistance genes or regulatory genes that refer to insecticide resistance were exist in house flies. My research lifted the veil on the pivotal roles of P450 and carboxylesterase genes in permethrin resistance of house flies, and opened the avenue for *in vitro* study.

6.2 Future studies

The detoxification roles of P450s and carboxylesterases in the pyrethroid resistance development of house flies have been revealed in my studies. However, the results of my studies also implied that beside detoxification, other roles of P450s may also involve in permethrin resistance in house flies. Based on these results, our future studies will include two aspects.

6.2.1 In vitro expression and the metabolism study of house fly key P450 genes on

permethrin insecticide

We will select key p450 genes and characterize their metabolisms on permethrin insecticide. Briefly, we will express the key p450 genes and cytochrome b5 reductase in vitro using Bac to Bac baculovirus expression system (Figure 6.1), and then characterize the activities of P450 proteins through measure the catalytic rate of 7-ethoxyresorufin *O*-deethylation (EROD) and 7-methoxyresorufin O-demethylation (MROD). The product of reactions, resorufin can be detected using the fluorimetric assay or high performance liquid chromatography (HPLC). At last, the metabolisms study of the expression P450s with permethrin will be conducted. Each reaction contained the P450/reductase (3:1), 1,2-didodecanoylrac- glycero-3-phosphocholine, NADPH regenerating system, and 60 uM permethrin in a total volume of 250 ul. Briefly, the reconstitution reaction will be initiated by addition of NADPH regenerating system, and incubated at 30 °C with shaking at 1,000 × *g*. Reactions will be stopped with 0.1 mL of methanol and incubated with shaking (1,000 × *g*) for 5 min at 30 °C to dissolve all available permethrin. Samples will be then centrifuged at 16,000 × *g* for 10 min at 4 °C, 150 μL supernatant will be transferred to HPLC vials. The quantity of pyrethroid remaining in the samples will be determined by reverse-phase HPLC with a monitoring absorbance wavelength of 232 nm.

We predict that key p450 genes and reductase will be expressed in vitro by using Bac to Bac baculovirus expression system, and recombinant P450 proteins will metabolize permethrin.

6.2.2 Characterization of the molecular mechanisms of P450 and carboxylesterase expression regulation in ALHF house flies

The regulatory elements of up-regulated P450 and carboxylesterase genes, which are mostly found upstream from the promoter regions and are responsible for binding regulatory proteins are still unknown in house flies. I will focus on characterization of the molecular mechanisms that involved in the regulation of P450 and carboxylesterase gene expression associated to insecticide resistance in house flies. Electrophorethc mobility shift assay (EMSA), also known as gel shift assay will be used for the regulatory elements of the target genes investigation.

First, I will find the promoter region sequence differences of the target genes between insecticide susceptible and resistant house flies. These regions are considered as candidate regulatory elements, and will be selected for further investigation of their roles in the interaction with transcriptional regulatory factors. I will design and synthesize the

radioactive probe DNA. After incubation for binding reaction of radioactive probe and extracted nuclear protein, electrophoresis and autoradiograph development will be conducted. The DNA-protein complex will be resolved by electrophoresis on a non-denaturing polyacrylamide gel and visualized by autoradiography by exposure to X-ray film with intensifying screens (Figure 6.2). Based on the mobility of the molecules and signal appearing on the film after autoradiography, the interaction between the regulatory factors at the promoter region of target genes and the transcriptional regulatory factor will be investigated eventually.

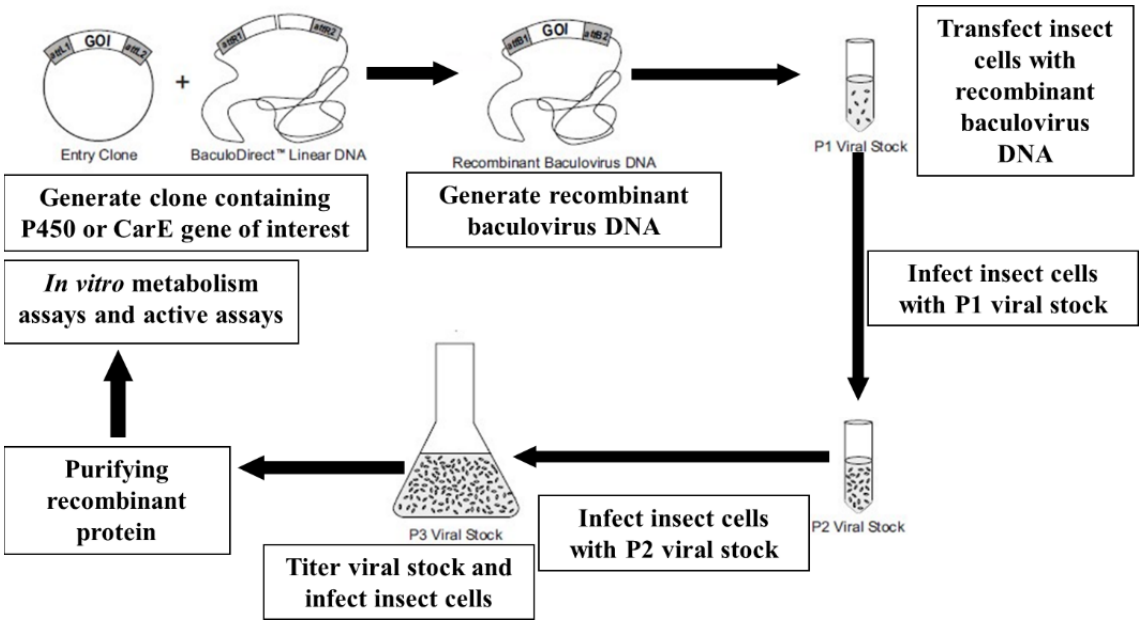


Figure 6.1 Schematic for the express target genes of interest using the BaculoDirect™ Baculovirus Expression System.

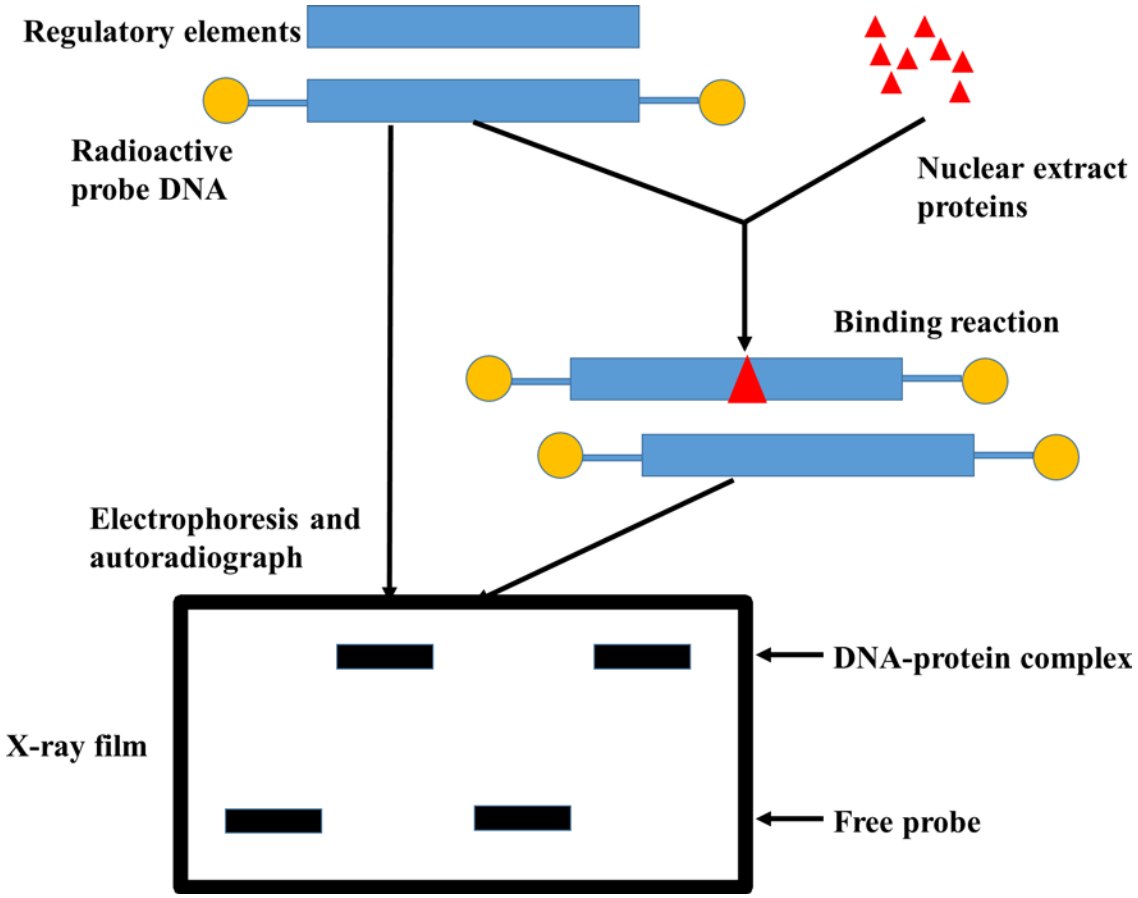


Figure 6.2 The scheme of electrophoretic mobility shift assay (EMSA)