A Thesis Submitted for the Degree of PhD at the University of Warwick

Permanent WRAP URL:

http://wrap.warwick.ac.uk/78850

Copyright and reuse:

This thesis is made available online and is protected by original copyright.

Please scroll down to view the document itself.

Please refer to the repository record for this item for information to help you to cite it.

Our policy information is available from the repository home page.

For more information, please contact the WRAP Team at: wrap@warwick.ac.uk

Identification of novel transcripts involved in insecticide resistance in African malaria vectors.

Victoria Anne Ingham
BA Hons (Oxon), MSc

Submitted for the degree of Doctor of Philosophy

Systems Biology DTC, University of Warwick, in collaboration with the Vector Biology Department at Liverpool School of Tropical Medicine

November 2015

Table of Contents

List of Figures

List of Tables

Acknowledgements

Firstly, I would like to thank my three supervisors Professor Hilary Ranson, Dr Jonathan Moore and Dr Simon Wagstaff for the chance to complete my PhD and all of the help and assistance afforded to me throughout my 3 years. I have been extremely fortunate to have three friendly and approachable supervisors who I have always been able to go to for guidance and support.

Secondly, thank you to the entire Vector Biology group, members past and present, who have assisted me at various points and made working across two departments a pleasure. I would like to particularly offer my thanks and appreciation to Patricia Pignatelli who has been fantastic in lending me a hand whenever I need it.

Thank you also, to the members of the Systems Biology DTC who have helped me when I've needed it and made me welcome on my visits to the department.

Finally, thank you to my support at home, to both my parents and my boyfriend, who have made living and working in both Coventry and Liverpool easy by visiting and supporting me, enabling me to move easily between the two cities.

Declaration

This thesis is submitted to the University of Warwick in support of my application for the degree of Doctor of Philosophy. I declare that all work contained within the thesis was written solely by myself, and all work contained herein is my own, unless otherwise stated. No part(s) of this thesis have been submitted to any other institution or for any other reward.

Summary

Insecticide resistance in malaria vectors is increasing across Sub-Saharan Africa, due, at least in part, due to the scale up of vector control programmes implemented to overcome morbidity and mortality associated with malaria. Four classes of public health insecticides are widely utilised in vector control strategies such as indoor residual spraying and insecticide treated nets; pyrethroids in particular are heavily used being the only class licensed for bed net usage. Despite the increase in resistance to these insecticides, and the implications this has for malaria control, resistance mechanisms remain incompletely understood. This study identifies transcripts associated with resistance to pyrethroid insecticides in African malaria vectors.

Study one describes the specific localisation of resistance-associated transcripts in different body parts of the mosquito. Both known and novel candidates are identified and their enrichments in two major detoxification structures (the midgut and malpighian tubules), the abdomen integument and the remaining tissues determined. Study two lays the foundation for the remaining studies by using a meta-analysis based approach to analyse all available pyrethroid resistant versus susceptible microarray datasets at LSTM performed on *Anopheles gambiae* s.l. Studies three through five explore three major candidates arising from this meta-analysis *SAP2*, a chemosensory protein; *Maf-S* a transcription factor with known involvement in redox response in *Drosophila* and *Met* a second transcription factor known to be involved in resistance to pyriproxyfen.

Abbreviations

ABCs: ATP-binding cassette transporters

ace-1: Acetylcholinesterase-1

AHR: Aryl- hydrocarbon

ANOVA: Analysis of variance

ARE: antioxidant response element

bHLH-PAS: Basic helix-loop-helix-PER-ARNT-SIM

CCEs: Carboxyl/cholinesterase

cDNA: Complementary DNA

CHC: Cuticular hydrocarbons

COEs: Carboxylesterases

cRNA: Complementary RNA

CSPs: Chemosensory proteins

DAVID: Database for Annotation, Visualization and Integrated Discovery

DDT: Dichlorodiphenyltrichloroethane

DE: Differential expression

DNA: Deoxyribonucleic acid

ER: Endoplasmic reticulum

FDR: False discovery rate

GaGa: Gamma-Gamma Model

GFP: Green fluorescent protein

GO: Gene ontology

GSTs: Glutathione-S-transferases

IRS: Indoor residual spraying

ITNs: Insecticide treated nets

kdr: Knockdown resistance

KEGG: Kyoto encyclopaedia of genes and genomes

LLINs: Long lasting insecticidal nets

LSTM: Liverpool School of Tropical Medicine

Met: *Methoprene tolerant*

MeV: Multi-experiment viewer

MCMC: Mixed chain Monte Carlo

mRNA: Messenger RNA

MS: Mean square

NADPH: Nicotinamide adenine dinucleotide phosphate

P450s: Cytochrome p450s

PBS: Phosphate-buffered saline

PSA: Pyrethroid survivor arrays

qPCR: Quantitative PCR

RNA: Ribonucleic acid

RNAi: RNA interference

R vs S: Resistant versus Susceptible

SS: Sum of squares

tRNA: Transfer RNA

UDP: Uridine disphosphate

UGT: UDP-glucuronosyltransferase

WHO: World Health Organization

XRE: Xenobiotic response element

1. INTRODUCTION

Malaria

Overview.

In 2013 there were an estimated 198 million cases of malaria, with 584 000
deaths, 78% of these being in children under five (WHO, 2014) (Figure 1.1).
Although these statistics are alarming, malaria cases have actually fallen
dramatically since 2000, with a 47% decrease in mortality rates and number of
infections decreasing by 26% (WHO, 2014). Currently, 97 countries and
territories have on-going malaria transmission and seven countries are in
prevention of reintroduction stage, equating, in total, to 104 endemic countries.
The greatest burden of malaria is found in Africa, with 80% of cases and 90% of
all malaria associated deaths (WHO, 2014). Several problems confound the
reporting of malaria, for example, people with acquired immunity show mild or
no symptoms after infection. Moreover, the symptoms of malaria are not readily
distinguishable from several other diseases and both self-treating and
misdiagnosis are common due to poor medical infrastructures (Hay *et al.*, 2010a;
Mharakurwa *et al.*, 2013)

Causative agent and transmission.
Malaria is caused by five species of the protozoan apicomplexan parasite
Plasmodium, with *Plasmodium falciparum* causing around half a million juvenile
deaths from malaria, mostly in Sub-Saharan Africa (Fidock, 2013). Although
much research is focused on *P. falciparum* and *P. vivax*, multiple species of the
malarial parasite can be found in sympatry, with co-infection of both human and
vector hosts common (Bruce *et al.*, 2000; Gnémé *et al.*, 2013). All human malarial
parasites are vectored by mosquitoes of the *Anopheline* genus, with at least 41

species known to transmit the disease worldwide (Sinka *et al.*, 2010b). Many competent vectors are found within Africa, a factor in the high prevalence of malaria across the continent (Sinka *et al.*, 2010b). The most highly efficient vectors are part of the *Anopheles gambiae* complex, which is composed of the dominant vector species *An. gambiae*, *An. coluzzii, An. arabiensis*, *An. merus* and *An. melas* with the latter two being restricted to coastal regions (Hay *et al.*, 2010b; Coetzee *et al.*, 2013). In addition to the *An. gambiae* complex, Africa has several other competent vectors including *An. funestus*, *An. nili* and *An. moucheti*. Dominant vector species are described as such due to the high impact on public health, largely due to their anthrophilic biting behaviour (Sinka *et al.*, 2010b).

Vector ecology is complex, with variations across seasons and between urban and rural areas (Hay *et al.*, 2010b). *An. gambiae* has a very large geographical range (Sinka *et al.*, 2010a; Sinka *et al.*, 2010b; Hay *et al.*, 2010b) with five chromosomal forms (Coluzzi *et al.*, 1985) and two molecular forms (della Torre *et al.*, 2001). Improvements in molecular systematics since the early 21st century have allowed more detailed exploration of the *An. gambiae* complex (Coetzee *et al.*, 2013). These data have shown that the two molecular forms - 'M' and 'S', have very little gene flow across West Africa (<1%) (della Torre *et al.*, 2005) and individual countries (Coetzee *et al.*, 2013). Although reproductive isolation of the two forms is incomplete (Weetman *et al.*, 2012), there are pre-existing premating barriers. An example of this is that M and S forms have distinct ecologies, with M form associated with permanent bodies of water, whereas S form is linked with temporary sites that are often rain dependent, such as puddles (della Torre *et al.*, 2005). Incipient speciation due to the lack of gene

flow was the accepted consensus; however, in 2013 Coetzee et al. (Coetzee *et al.*, 2013) described two new species within the *An. gambiae* complex, renaming the 'M' form as *An. coluzzii* whilst retaining *An. gambiae* for the 'S' form and describing an Ethiopian species: *An. amharicus* (Coetzee *et al.*, 2013)*.

With 3.3 billion people at risk of malaria (WHO, 2014), equating to over half the world's population, interventions have been put into place, with the aim of lowering malaria related morbidity and mortality and eventual eradication. Although artemisinin based therapy is an effective treatment for the parasite, the majority of these interventions focused on vector control (WHO, 2014), with insecticide treated nets (ITNs) being by far the most important intervention, accounting for around half of the decline in *P. falciparum* transmission (Bhatt *et al.*, 2015).
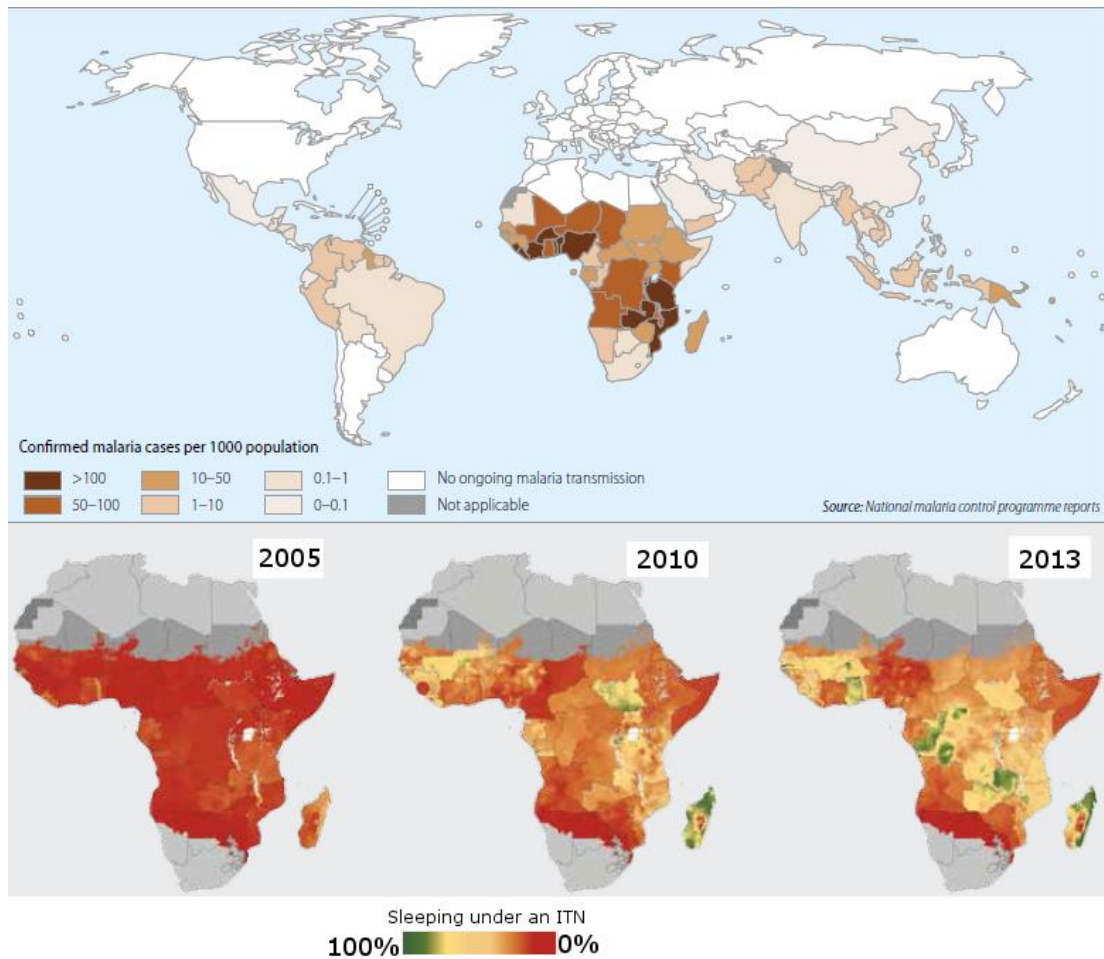
Figure 1.1: Malaria transmission and bed net usage. (a) Countries with on-going malaria transmission. (b) Proportion of households with bed nets from 2005 - 2014. Modified from the World Malaria Report (WHO, 2014).

Vector Control

Overview.

The total international disbursement for malaria has risen rapidly in the last 13 years, from US$100 million in 2000 to an estimated US$2.7 billion in 2013; however, US$5.1 billion is estimated to be needed to achieve universal coverage of malaria interventions (WHO, 2014). The malaria intervention strategies currently utilised for vector control include indoor residual spraying (IRS) and insecticide treated nets (ITNs). Almost half the at risk population (49%) had

access to at least one ITN within a single household due a projected 214 million nets to be distributed by the end of 2014 (WHO, 2014). With only one class of insecticide licensed for bed net use, the pyrethroids (Davies *et al.*, 2007), selection pressure for resistance against this insecticide class is high.

Pyrethroids are one class out of four currently used in IRS programmes, with the others being the organochlorides, organophosphates and carbamates. The most infamous insecticide, DDT, an organochloride, is still utilised for IRS. Pyrethroids were the most commonly used insecticide class in IRS programmes until recently. This, in addition to the use of this class in agriculture and domestic applications such as aerosols and coils, creates several major exposure routes for the mosquitoes (Davies *et al.*, 2007; Ranson *et al.*, 2011). More recently, pyrethroids have been replaced by carbamates and organophosphates in many IRS programmes due to problems of resistance (Mnzava *et al.*, 2015).

To be successful, malaria intervention strategies must take into account numerous factors, including complex vector dynamics observed locally and continentally across Africa. *Anopheles* mosquitoes respond to vector control through evolution of resistance, changes to behaviour and/or changes in population density and composition, all of which impact vector control. Behavioural changes in response to ITNs have been demonstrated in female mosquitoes, showing changes in biting behaviours across Africa, including Benin (Moiroux *et al.*, 2012), Senegal (Sougoufara *et al.*, 2014), Tanzania (Russell *et al.*, 2011) and Equatorial Guinea (Reddy *et al.*, 2011). Changes in vector composition are also seen, with declines in the *Anopheles gambiae* complex and increase in

other dominant vector species (Bayoh *et al.*, 2010; Derua *et al.*, 2012; Jones *et al.*, 2012b; Mwangangi *et al.*, 2013; Lwetoijera *et al.*, 2014; McCann *et al.*, 2014).

Vector control: Insecticide treated nets.

The percentage of household bed nets has risen from 3% in 2000 to 49% in 2013 (WHO, 2014) (Figure 1.1). Every year it is necessary to distribute 150 million bed nets to protect at-risk people throughout the African continent, which equates to one LLIN for every 1.8 person. These LLINs are generally provided free of charge to all at risk, with a mass distribution every three years. The scale up of global interventions has significantly reduced the mortality and morbidity associated with malaria infection (WHO, 2014).

The increased use of insecticides in agriculture, IRS and ITNs exerts a large selection pressure on the exposed mosquito populations and has resulted in resistance to the major classes of insecticide (Fanello *et al.*, 2000; Awolola *et al.*, 2005; Liu *et al.*, 2006; Davies *et al.*, 2007; N'Guessan *et al.*, 2007; Müller *et al.*, 2008a; Müller *et al.*, 2008b; Awolola *et al.*, 2009; Kawada *et al.*, 2011; Ranson *et al.*, 2011; Namountougou *et al.*, 2012; Edi *et al.*, 2012; Badolo *et al.*, 2012; Temu *et al.*, 2012; Jones *et al.*, 2012b; Bigoga *et al.*, 2012; Kwiatkowska *et al.*, 2013; Protopopoff *et al.*, 2013; Jones *et al.*, 2013; Mawejje *et al.*, 2013; Ilboudo-Sanogo *et al.*, 2013; Corbel & N'Guessan, 2013; Koffi *et al.*, 2013; Nardini *et al.*, 2013; Riveron *et al.*, 2013; Shetty *et al.*, 2013; Namountougou *et al.*, 2013; Fossog Tene *et al.*, 2013; Nkya *et al.*, 2014; Abdalla *et al.*, 2014; Edi *et al.*, 2014a; Ibrahim *et al.*, 2014; Djègbè *et al.*, 2014; Mitchell *et al.*, 2014; Ndiath *et al.*, 2014; Thomsen *et al.*, 2014; Edi *et al.*, 2014b; Toé *et al.*, 2015). Resistance has now been identified in at

least 64 countries, with 77 countries actively monitoring insecticide resistance (WHO, 2014). A lack of investment in the discovery of new public health insecticides has resulted in no new active ingredients being produced in the last 30 years (Davies *et al.*, 2007; Ranson *et al.*, 2011) making the issue of resistance management and mitigation critical.

Common resistance mechanisms employed by *Anopheline* vectors that have been studied extensively include target site mutations, resulting in a loss of sensitivity to specific insecticides (Martinez‐Torres *et al.*, 1998; Verhaeghen *et al.*, 2006a; Davies *et al.*, 2007; Donnelly *et al.*, 2009; Namountougou *et al.*, 2013) and metabolic resistance, through five key detoxification families (Newcomb *et al.*, 1997; Hemingway & Karunaratne, 1998; Amenya *et al.*, 2008; Djouaka *et al.*, 2008; Müller *et al.*, 2008b; Stevenson *et al.*, 2011; Ranson *et al.*, 2011; Ahn *et al.*, 2012; Mitchell *et al.*, 2012; Jones *et al.*, 2012b; Chandor-Proust *et al.*, 2013; Dermauw & Van Leeuwen, 2013; Fossog Tene *et al.*, 2013; Riveron *et al.*, 2014a; Aravindan V, Muthukumaravel S, 2014; Djègbè *et al.*, 2014; Mitchell *et al.*, 2014; Epis *et al.*, 2014; Liu *et al.*, 2015). Other mechanisms that are less well studied include putatively reduced uptake of the insecticide due to changes in the cuticle, present in both *Helicoverpa armigera* (Gunning *et al.*, 1995; Ahmad *et al.*, 2006) and *Triatoma infestans* (Pedrini *et al.*, 2009) and potential behavioural changes mentioned above.

Insecticide Resistance Mechanisms.

Overview.
Two main evolutionary forms of resistance have been studied and characterised

in *Anopheles* species: metabolic resistance and target site resistance. Target site

resistance leads to a decreased response to the insecticide due to mutations in

the specific area targeted by each class of insecticides (Martinez‐Torres *et al.*,

1998), whereas metabolic resistance acts through increased clearance and

detoxification of insecticides (Xu *et al.*, 2005).

Target site resistance.
Out of all the resistance mechanisms in the *An. gambiae* species complex, the

most thoroughly understood is knock-down resistance or *kdr*. *kdr* is a target site

mutation at the *para* voltage gated sodium channel, which provides resistance to

the two insecticides that target this; DDT and pyrethroids. These insecticides

work by preventing the sodium channels from shutting, resulting in prolonged

nervous impulses, eventually causing paralysis and death (Davies *et al.*, 2007;

Ranson *et al.*, 2011). There are two main mutations in the sodium channel, a

substitution of a leucine allele with a phenylalanine or serine residue at position

1014 (Davies *et al.*, 2007; Ranson *et al.*, 2011). The two substitutions were

originally termed *kdr* East (Ramphul *et al.*, 2009) and *kdr* West (Dabire *et al.*,

2009) reflecting their presumed origin in the African continent; however, both

mutations are now widely distributed (Djègbè *et al.*, 2011; Ndiath *et al.*, 2012;

Mawejje *et al.*, 2013; Nwane *et al.*, 2013; Ibrahim *et al.*, 2014; Dabiré *et al.*, 2014).

kdr resistance is found in *An. gambiae* (Martinez‐Torres *et al.*, 1998; Ranson *et

al.*, 2000; Fanello *et al.*, 2000; Dabire *et al.*, 2009; Namountougou *et al.*, 2013;

Kabula *et al.*, 2014)*, An. coluzzii* (Djègbè *et al.*, 2011; Kwiatkowska *et al.*, 2013;

Ibrahim *et al.*, 2014; Edi *et al.*, 2014b) and *An. arabiensis* (Djègbè *et al.*, 2011;

Kawada *et al.*, 2011; Jones *et al.*, 2012b) but has not been seen in *An. funestus*

(Coetzee & Koekemoer, 2013)*.*

A second target site mutation conferring insecticide resistance affects a

carboxylesterase, *ace-1.* The *ace-1* mutation causes acetylcholine esterase to be

insensitive to organophosphate and carbamate classes of insecticide.

Acetylcholine esterase hydrolyses the neurotransmitter acetylcholine, leading to

termination of nerve signals. Both organophosphates and carbamates bind to

acetylcholine esterase, inhibiting it, leading to paralysis and death. *ace-1*

resistance in Dipterans is conferred due to three separate amino acid

substitutions (Alout & Weill, 2008), but only a single point mutation in the ace-1

gene at position 119, resulting in a glycine to serine substitution, has been

described in *An. gambiae* s.l (Weill *et al.*, 2003; Weill *et al.*, 2004; Weetman *et al.*,

2015).

Metabolic resistance.
A second method of resistance is based around a reduced exposure to the

insecticide through several methods: (i) preventing uptake (ii) sequestration (iii)

metabolism and (iv) excretion (Xu *et al.*, 2005). Both sequestration and

metabolism are important in arthropod insecticide resistance and can be

subdivided into phase I-III. Phase I entails enzymes such as cytochrome p450s or

carboxylesterases activating the compounds, making them more reactive and

water soluble allowing phase II conjugation enzymes such as GSTs and UDP-

glycosyltransferases to add specific moieties. Lastly, in phase III, the polar or

conjugated compounds can be transported out of the cells by cellular

transporters such as ABC transporters (Xu *et al.*, 2005).

Metabolic resistance has been shown to play a role in resistance to insecticides,

although the true extent and genetic involvement is unknown (Ranson *et al.*,

2011). Five classes of detoxification proteins have been linked extensively with

removal or sequestration of the active insecticide before it reaches its target site

(Ranson *et al.*, 2002; Ranson *et al.*, 2011). These classes are: cytochrome P450s,

carboxylesterases (COEs), ABC transporters, glutathione-S-transferases (GSTs)

and glucuronosyltransferases (UDPs) (Ranson *et al.*, 2002; Ahn *et al.*, 2012).

Detoxification enzymes are encoded by large gene families in *An. gambiae* and

An. coluzzii, with 111 Cytochrome P450s, 31 GSTs, 51 COEs (Ranson *et al.*, 2002),

55 ABC transporters (Roth *et al.*, 2003, Ranson and Pignatelli, unpublished daa)

and 26 UDPs (Ahn *et al.*, 2012).

Cytochrome p450s are a superfamily of mono-oxygenases, acting by using a

heme group to oxidise molecules, often resulting in hydrophilic secondary

products for clearance (Figure 1.2). Cytochrome P450s are the most well studied

of the five detoxification families and have been shown to directly metabolise

three of the four commonly used insecticide classes (pyrethroids,

organophosphates and DDT) (Müller *et al.*, 2008b; Stevenson *et al.*, 2011;

Mitchell *et al.*, 2012; Chandor-Proust *et al.*, 2013) with a probable role in

carbamate metabolism (Edi et al, 2014). Of the Cytochrome P450 sub families,

the CYP6 group is most widely implicated in the metabolism of insecticides in *An.*

gambiae with *CYP6P3* directly involved with pyrethroid and organophosphate

metabolism (Müller *et al.*, 2008b, Paine and Yunta Yanes, unpublished data), *CYP6M2* involved in DDT and pyrethroid metabolism (Stevenson *et al.*, 2011; Mitchell *et al.*, 2012) and the *CYP6Z* family implicated in pyrethroid metabolite binding in *Aedes* mosquitoes (Chandor-Proust *et al.*, 2013).
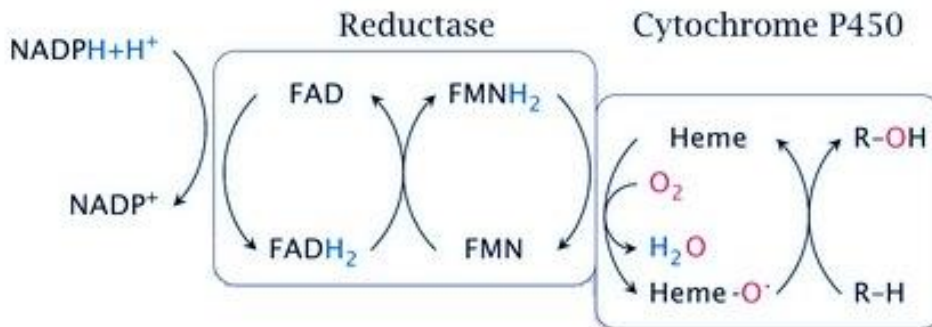
Figure 1.2: Mechanism of action of Cytochrome P450s. Incomplete diagram showing: (1) Substrate binds in proximity to the heme group, which results in a change of conformation, displacing H_2O from the heme iron, changing its state. (2) Substance binding induces electron transfer from NADPH via a reductase. (3) Oxygen binds ferrous heme. (4) A second electron is transferred from the reductase, reducing $Fe-O_2$. (5) The peroxide group in (4) is protonated twice, releasing H_2O, forming a highly active P450 compound. (6) After catalysis, the system returns to the original state. (Meunier *et al.*, 2004).

Many insecticides contain ester bonds and so are amenable to hydrolysis via esterases (Sogorb & Vilanova, 2002) (Figure 1.3). Esterases have been linked to insecticide sequestration, being the primary resistance mechanism for organophosphates and a secondary mechanism for carbamates (Hemingway & Karunaratne, 1998). Esterases have been implicated in organophosphate

resistance in *Aedes aegyptii* and *Culex pipiens* by ubiquitous over-expression in resistant strains (Guillemaud *et al.*, 1997; Bisset *et al.*, 2011; Poupardin *et al.*, 2014; Grigoraki *et al.*, 2015) and pyrethroid resistance in *Musca domestica* (Zhang *et al.*, 2007). Carboxylesterases have also been found up-regulated in pyrethroid resistant populations in a number of *An. gambiae, coluzzii* and *arabiensis* microarray experiments (Witzig, 2012; Jones *et al.*, 2012b; Kwiatkowska *et al.*, 2013; Mawejje *et al.*, 2013; Jones *et al.*, 2013; Fossog Tene *et al.*, 2013; Hemingway *et al.*, 2013; Abdalla *et al.*, 2014; Thomsen *et al.*, 2014; Toé *et al.*, 2015) although a direct role in insecticide metabolism has not yet been demonstrated for *Anopheles* carboxylesterases .
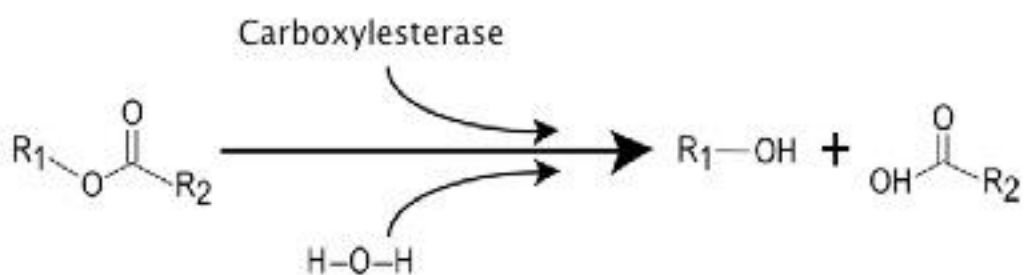
Figure 1.3: Mechanism of action of carboxylesterases. Carboxylesterases hydrolyse ester groups, resulting in the formation of an alcohol and an acid metabolite (Hemingway & Karunaratne, 1998).

Glutathione-S-transferases are a large family of enzymes that act as detoxifiers by conjugating glutathione, dehydrochlorination, glutathione peroxidase activity or passive binding of/to toxic compounds (Hayes & Wolf, 1988; Mannervik *et al.*, 1988; Yang *et al.*, 2001) (Figure 1.4). GSTs have been linked to detoxification of public health insecticides in the common housefly, *Musca domestica* (Clark & Shamaan, 1984; Wei *et al.*, 2001; Kristensen, 2005), *Drosophila melanogaster*

(Tang & Tu, 1994; Pedra *et al.*, 2004; Le Goff *et al.*, 2006; Low *et al.*, 2010), the planthopper, *Nilaparvata lugens* (Vontas *et al.*, 2001; Vontas *et al.*, 2002) and a number of mosquito species (Ranson *et al.*, 2001; Ortelli *et al.*, 2003; Enayati *et al.*, 2005; Lumjuan *et al.*, 2007; Lumjuan *et al.*, 2011; Aravindan *et al.*, 2014). GSTs have been experimentally shown to conjugate glutathione to organophosphate (Huang *et al.*, 1998) as a secondary detoxification step (Hemingway *et al.*, 1991). Organochlorides are dehydrochlorinated by GSTs; this is a particularly important resistance mechanism for DDT breakdown (Ranson *et al.*, 2001; Ortelli *et al.*, 2003; Lumjuan *et al.*, 2007; Low *et al.*, 2010; Lumjuan *et al.*, 2011; Aravindan *et al.*, 2014). Evidence exists for a role of GSTs in pyrethroid resistance (Vontas *et al.*, 2001; Lumjuan *et al.*, 2011), which may be due to the protective activity against lipid peroxidation due to oxidative stress (Yang *et al.*, 2001). Furthermore, a single amino acid change, L119F in an up-regulated GST (*GSTE2* in *An. funestus*) has been shown to enlarge the DDT binding cavity, allowing more efficient metabolism of DDT (Riveron *et al.*, 2014b). *GSTE2* has also been demonstrated to directly metabolise the pyrethroid permethrin (Riveron *et al.*, 2014b).

Although GSTs are found ubiquitously across all organisms, insects contain two classes of GSTs that are not found in mammalian organisms, the delta and epsilon classes, both of which are found in *An. gambiae*, with seven and 12 members respectively (Ranson *et al.*, 2002). The epsilon class has been linked to insecticide resistance with *GSTE2* demonstrating the highest level of dehydrochlorinase activity (Ortelli *et al.*, 2003). GSTE2 in *An. funestus* has been shown to directly metabolise both permethrin and DDT (Riveron *et al.*, 2014b);

further to this, GSTs have been found upregulated in microarray experiments on resistant mosquitoes (Witzig, 2012; Jones *et al.*, 2012b; Kwiatkowska *et al.*, 2013; Mawejje *et al.*, 2013; Jones *et al.*, 2013; Fossog Tene *et al.*, 2013; Hemingway *et al.*, 2013; Abdalla *et al.*, 2014; Thomsen *et al.*, 2014; Toé *et al.*, 2015).
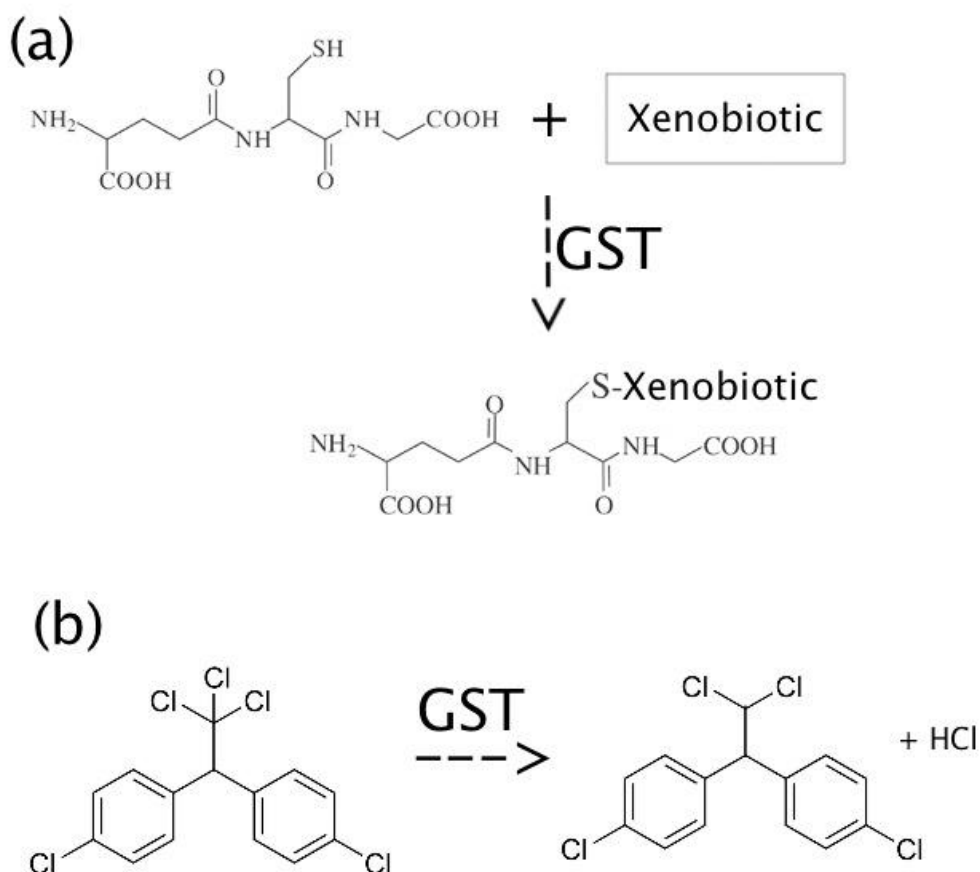
Figure 1.4: Glutathione-S-transferase mechanisms. (a) Glutathione conjugation to an unspecified xenobiotic and resultant product. (b) DDT dehydrochlorination through GST action (Armstrong, 1991).

UDP-glycosyltransferases (UGTs) catalyse the addition of sugars donated by a UDP-glycoside to small lipophilic compounds, resulting in hydrophilic compounds which can then be easily excreted (Ahn *et al.*, 2012; Bozzolan *et al.*,

2014) (Figure 1.5). UGTs are found in the endoplasmic reticulum membrane and have a conserved structure, with an N-terminal binding domain and C-terminal UDP-glycoside binding domain (Magdalou *et al.*, 2010). Although little is known about arthropod UGTs, activity has been observed against a variety of plant allelochemicals (Huang *et al.*, 2008b; Ahn *et al.*, 2011). Insect UGTs have also been implicated in resistance towards several classes of insecticide (Pedra *et al.*, 2004; Vontas *et al.*, 2005; Silva *et al.*, 2012; Bozzolan *et al.*, 2014), a hypothesis strengthened by their localisation to the three primary detoxification structures; the midgut, malpighian tubules and fat bodies (Ahn *et al.*, 2012). Furthermore, UGTs have been detected as differentially expressed in a variety of microarray experiments on resistant mosquitoes (Witzig, 2012; Jones *et al.*, 2012b; Kwiatkowska *et al.*, 2013; Mawejje *et al.*, 2013; Jones *et al.*, 2013; Fossog Tene *et al.*, 2013; Hemingway *et al.*, 2013; Abdalla *et al.*, 2014; Thomsen *et al.*, 2014; Toé *et al.*, 2015).
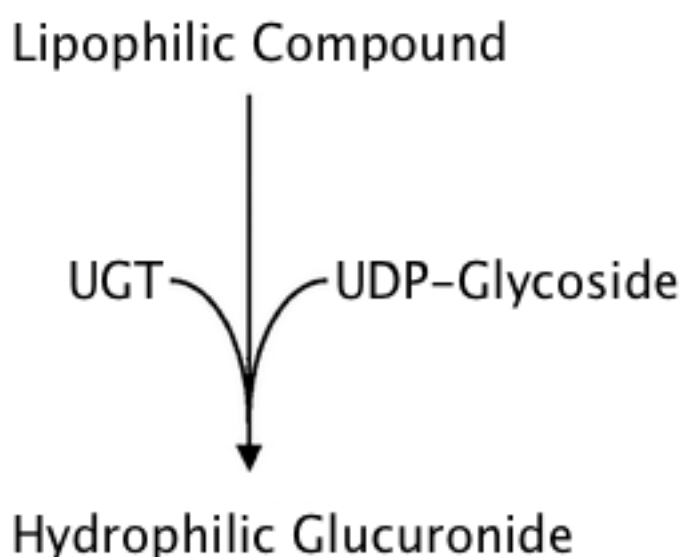
Figure 1.5: General UDP-glycosyltransferases mechanism. Lipophilic compounds have glycoside groups conjugated to them through the UGT enzyme to produce hydrophilic glucuronide (Tukey & Strassburg, 2000).

Although mammalian ABC transporters are very well studied due to their roles in drug resistance, very little is known of the arthropod ABCs (Dermauw & Van Leeuwen, 2013). ABC-transporters act as a pump removing xenobiotics and potentially GST-conjugates found in cells in an energy-dependent manner, thereby lowering the concentration to a sub-lethal amount (Cole *et al.*, 1994; Schinkel *et al.*, 1995; Germann & Chambers, 1998). Mammalian ABC transporters have been shown to work in conjunction with phase II conjugating enzymes such as GSTs and UDP-glycosyltransferases, further strengthening their putative role in xenobiotic transport (Dermauw & Van Leeuwen, 2013). Although direct links between insecticide resistance and ABCs are missing, work in *Drosophila* have linked ABCs to resistance through both insecticide transport (Buss & Callaghan, 2008) and a putative 'detoxification network' (Shah *et al.*, 2012). ABC transporters are regularly found as upregulated in microarray experiments of insecticide resistant mosquitoes (Witzig, 2012; Jones *et al.*, 2012b; Kwiatkowska *et al.*, 2013; Mawejje *et al.*, 2013; Jones *et al.*, 2013; Fossog Tene *et al.*, 2013; Hemingway *et al.*, 2013; Abdalla *et al.*, 2014; Thomsen *et al.*, 2014; Toé *et al.*, 2015).

Cuticular resistance.
The mode of application of adult-targeted insecticides in malaria control requires the insecticide to penetrate the cuticle of the mosquito's appendages. In

other species, reductions in the rate of cuticular penetration have been associated with resistance. In *An. gambiae* both genes encoding structural proteins within the cuticle and those encoding enzymes which catalyse key steps in the biosynthetic pathway of the outer layer of the cuticle (cuticular hydrocarbons) have been found up-regulated in multiple pyrethroid resistant populations (Awolola *et al.*, 2009; Jones *et al.*, 2013). In addition, cytochrome p450s are also involved in *Drosophila* cuticular hydrocarbon synthesis, with *CYP4G16* and *CYP4G17* homologs (Qiu *et al.*, 2012). Thickening of the cuticle is a putative mechanism to limit uptake of the insecticide through the abdomen of the mosquito. Ongoing studies are investigating the rate of uptake of insecticides in susceptible and resistant populations and comparing their CHC composition.

Transcriptional control of xenobiotic response.
Very little is known about the transcriptional control of xenobiotic detoxification in Dipteran species. In mammalian systems, *Aryl- hydrocarbon receptor* (*AHR*) and *NRF2* have been shown to be key regulators of drug-metabolising enzymes through binding xenobiotic and antioxidant response elements (XRE/ARE) (Miao *et al.*, 2005). In black swallowtail caterpillars, *Papilio polyxenes* and corn earworms, *Helicoverpa zea* the expression of Cytochrome P450s have been shown to be regulated by the action of XRE binding in response to toxic compounds (Brown *et al.*, 2005; Zhang *et al.*, 2010). In the case of *P. polyxenes* the binding response was shown to be through *AHR* in a homologous manner to mammalian systems (Brown *et al.*, 2005). Despite these links in Lepidopteran species, the *Drosophila* homologue *spineless* appears to play only a developmental role (Burgess & Duncan, 1990; Emmons *et al.*, 1999; Kim *et al.*,

2006). Just two transcription factors have been linked to insecticide resistance in *Drosophila*: *Maf-S* to DDT (Misra *et al.*, 2011; Misra *et al.*, 2013) and *Methoprene-tolerant* to methoprene and pyriproxyfen (Charles *et al.*, 2011; Zhang *et al.*, 2011; Zou *et al.*, 2013). These are described in more detail in subsequent chapters.

Microarray Analysis.

Microarray and enrichment analysis.

No gold standard currently exists for the analysis of microarray data, with different groups employing different methods, resulting in variation in the differentially expressed genes identified. Many tools are currently available for analysis, including several packages for R (Smyth, 2004; Rossell, 2009) and graphical user interface based tools (Mehta & Rani, 2011), such as MeV (Pan, 2002). Linear models are often applied, which are a generalisation of linear regression, taking into account multiple variables; the limma (Smyth, 2004) package in the statistical software R and variations of ANOVAs most often found in the literature. All methods available within R require pre-processed data, which in itself produces variability in the data, with different methods of background correction (Kooperberg *et al.*, 2002; Ritchie *et al.*, 2007), within and between array normalisation (Yang *et al.*, 2002; Bolstad *et al.*, 2003; Smyth & Speed, 2003; Yang & Thorne, 2003; Irizarry *et al.*, 2003; McGee & Chen, 2006).

The above analytical approaches are suitable for a direct comparison approach, producing a p-value based on component A vs component B on a single array.

Comparing across multiple arrays, with different experimental methods leads to additional challenges. These comparisons can be analysed using a venn-diagram like approach, using the GaGa package in R (Rossell, 2009). The GaGa analysis technique differs from the traditional approach by attempting to discern patterns across several independent microarray datasets, resulting in a fold change that is a direct multiplier between the two original fold changes. GaGa works by modelling the gene expression data by fitting a gamma-gamma distribution (Rossell, 2009) and resulting in specified groups that show a similar pattern across genes or a different expression pattern between one or several groups.

Microarray data analysis problems are further confounded by the long gene list outputs, with no obvious biological meaning (Koschmieder *et al.*, 2012). These long gene lists require interpretation, with several popular methods available, including DAVID analysis (Huang *et al.*, 2007; Huang *et al.*, 2008a), KEGG pathway analysis (Kanehisa & Goto, 2000) and GO term enrichment (Ashburner *et al.*, 2000). Some of these tools are available for *Anopheles* specific analysis, though the data is often based on *Drosophila* homologs and is thus, less complete (Consortium, 2003; Lyne *et al.*, 2007). A dedicated *Anopheles* online database is available but currently, without any enrichment analysis tools (Lawson *et al.*, 2009).

Pathway analysis is an important step from a gene centric approach to understanding the holistic changes that occur within an organism under varying conditions (Khatri *et al.*, 2012). Pathway analysis of *An. gambiae* is currently

hampered by the lack of knowledge of complete pathways within insects. KEGG (Kanehisa & Goto, 2000) and Biocarta (Nishimura, 2001) are the most complete pathway databases; however, neither contain any fly specific pathways, being based around human and *E. coli* metabolic processes. Furthermore, pathway analysis currently suffers from the failure to take into account any form of pathway hierarchy, with genes higher up in a given pathway likely to have greater affect than those at the terminus (Tarca *et al.*, 2009). Numerous methods have been developed to improve pathway analysis, with different statistical techniques and completely new methods being employed (John Tomfohr, Jun Lu, 2005; Werner, 2008; Tarca *et al.*, 2009; Khatri *et al.*, 2012; Huang & Lin, 2013).

In addition to large genes lists, microarrays offer a rich source of data that is difficult to interpret. Using correlation networks, it has previously been demonstrated that biological patterns can be inferred (Stuart *et al.*, 2003; Mutwil *et al.*, 2010). A variety of methods exist for the production of clusters within correlation networks, such as cliques/factions, MCMC models and Bayesian methods (Stuart *et al.*, 2003; Mutwil *et al.*, 2010). By using a strict correlation cut off and *a priori* knowledge, these networks can give rise to clusters showing similar patterns of expression patterns. These data can allow inference of pathways or functionality of the clusters.

In order to ensure that maximal information is gained from microarray experiments, whole organism arrays should be supplanted with analysis of tissue and organ specific expression (Dorman, 2007; Johnson *et al.*, 2013). When this is not possible, using meta-analysis techniques across many microarrays

with similar experimental design can add confidence to candidate selection and allow identification of relationships between transcripts that were missed in stand-alone experiments (Tseng *et al.*, 2012).

This study sought to use alternative approaches for the analysis of multiple microarray data sets to identify transcripts that were commonly associated with pyrethroid resistance in multiple populations of *An. gambiae* sl.

Hypotheses

1) Structure specific microarray experiments will reveal patterns of transcript expression and novel insecticide resistance candidate transcripts missed using whole organism microarrays.

2) A meta-analysis based approach will reveal relationships between arrays and transcripts and identify novel insecticide resistance candidate transcripts.

3) Candidates for transcriptional regulation of xenobiotic response can be identified through literature searches and existing microarray data.

2. MATERIALS AND METHODS

Study One: Analysis of gene expression in different body parts.

Mosquito rearing conditions

The *An. gambiae* used in these experiments were all reared under standard insectary conditions at 27°C and 70-80% humidity under a 16:8 hour photoperiod. The N'Gousso strain is originally from Cameroon and is susceptible to all classes of insecticide (Harris *et al.*, 2010). N'Guosso is the M molecular form of *An. gambia*e, recently re-classified as a separate species, *An. coluzzi* (Coetzee *et al.*, 2013). In contrast, the Tiassalé strain from Côte D'Ivoire is resistant to all classes of insecticide (Edi *et al.*, 2012; Edi *et al.*, 2014a). This strain was colonised from the field site in 2012 and is a mixture of the M and S molecular forms. At the time of the study, the LD_{50} for the Tiassalé strain was 68 and 81 fold higher than the corresponding value for the N'Gousso strain for permethrin and deltamethrin respectively. Further details of the resistance profile of this strain are contained within the references (Edi *et al.*, 2012; Edi *et al.*, 2014a).

Microarray experiments

In this section, P. Pignatelli performed all preparation for cRNA, C. Jones, V. Ingham and P. Pignatelli performed the hybridisations and subsequent scanning. RNA was extracted from three dissected body parts: the malpighian tubules, the midgut and the abdomen integument (containing the fat body but also epidermal, neuronal, muscle and oenocyte cells) with the remaining undissected body parts forming a fourth sample group. Mosquitoes were collected between the hours of 8am and 2pm and dissected immediately on a CO_2 block. Post dissection, each body part was added to extraction buffer from the PicoPure RNA extraction kit, heated for 30 minutes at 42°c and frozen at -80°c as per manufacturer's instructions. Each biological replicate for each strain consisted of RNA, extracted

using PicoPure RNA Isolation kit (Arcturus), from 12 three-five day old non-blood fed, presumed mated females. The quantity and quality of the RNA was assessed using a nanodrop spectrophotometer (Nanodrop Technologies UK) and Bioanalyser (Agilent) respectively. Four biological replicates were prepared for each body part per strain. RNA from the four dissections was pooled according to the proportion of RNA extracted from each body part to reconstitute the 'whole organism' (7%, 6%, 24% and 63% RNA from abdomen integument, malpighian tubules, midgut, and remaining material respectively). The use of a reconstituted reference sample minimised potential sources of bias that could have arisen from circadian changes in gene expression and changes in the proportion of the M or S molecular form in the different biological replicates. 100ng of RNA was amplified and labelled with Cy3 and Cy5, using the Two colour low input Quick Amp labelling kit (Agilent) following the manufacturer's instructions. Samples were then purified (Qiagen) with the cRNA yield and quality assessed using the nano-drop and Bioanalyser respectively. RNA from each Tiassalé body part was competitively hybridised with the respective N'Gousso body part, as well as each body part from the resistant and susceptible strain being compared to the re-constituted whole organism. Dye swaps were performed on two out of four technical replicates for each array, to correct for dye bias. Labelled cRNAs were hybridised to the whole genome 8x15k *Anopheles gambiae* array (ArrayExpress accession number A-MEXP-2211). Microarray hybridisation, washing and scanning were performed according to previously described protocols (Mitchell *et al.*, 2012).

Microarray analysis

The resulting data were analysed using R. Within-array normalisation was carried out by Loess, and between array normalisation by Aquantile both found within the limma package (Smyth, 2004). Signals were corrected for dye by performing and correcting for dye swaps. The limma package (Smyth, 2004) was used to fit linear models to normalised corrected signals to assess differential expression using a design matrix to infer contrast matrices where necessary. All parameters used were default. A bespoke pipeline using the GaGa package (Rossell, 2009) was used to fit gamma-gamma models of variation to normalised corrected signals, in order to assign probes to one of two patterns of expression: X equals Y or X does not equal Y, where X represents the resistant population arrays and Y the susceptible arrays. These data were subsequently used to assess enrichment in each expression pattern, through GO term analysis using the TopGO package (Alexa & Rahnenfuhrer, 2010). A standard FDR adjusted p value cut off of $p \leq 0.05$ was applied to all data describing localisation of detoxification candidates. A second stringent selection method was used to reduce the probe list based on previously published methodology (Morey *et al.*, 2006), requiring that the following criteria were met: adjusted p-value ≤ 0.001, raw fluorescence intensity > median, and Tiassalé vs. N'Gousso ± 0.485 Log_2 fold change between the strains. All candidates selected also demonstrated a positive GaGa analysis fold change, indicative of higher transcript levels and hence localisation in the resistant tissue to the susceptible.

qPCR

RNA (4μg) from each biological replicate was reverse transcribed using Oligo dT (Invitrogen) and Superscript III (Invitrogen) according to manufacturer's instructions. Quantitative real-time PCR was performed using SYBR Green Supermix III (Applied Biosystems) using an MX3005 and the associated MxPro software (Agilent). Primer Blast (NCBI) (Ye *et al.*, 2012) was used to design primer pairs (Table 2.1). Where possible, primers were designed to span an exon junction but this was not achievable for six of the P450 genes (*CYP325A1*, *CYP6P3*, *CYP4G17*, *CYP6Z3*, *CYP12F2* and *CYP6Z2*) due to the high degree of sequence similarity. Each 20μl reaction contained 10 μl SYBR Green Supermix, 0.3 μM of each primer and 1μl of 1:10 diluted cDNA. Standard curves were produced using whole N'Gousso cDNA, in 1, 1:5, 1:25, 1:125 dilutions. qPCR was performed with the following conditions: 3 minutes at 95°C, with 40 cycles of 10 seconds at 95°C and 10 seconds at 60°C. All amplification efficiencies of designed primers were within acceptable range (90-120%), following MIQE guidelines (Bustin *et al.*, 2009).

DESCRIPTION	ID	FORWARD	REVERSE
CYP6Z3	AGAP008218	CCACGCAATTGCATTGGTCT	TTCTACGCGCATGGGGAAAC
CYP6Z2	AGAP008217	ACTATCCGTTCGGAGCTGGT	TGTGGCCTGGAACTTGAAC
CYP6P3	AGAP002865	TGTGATTGACGAAACCCTTCGGAAG	ATAGTCCACAGACGGTACGCGGG
CYP12F2	AGAP008020	GCAACAAAGGTCGTTACGCA	AATGTCAGGTGGGCTGTACG
CYP305A1	AGAP005656	TGAACGAGGTACACCGGAAC	GTAACACCCAGTCGTGTCCC
CYP4G17	AGAP000877	TGACGGTGGACATTCTGCTC	GTCACACATTTTCATGACAGCCA
Diminutive	AGAP000646	CGTCCGATTCCGATGAAGAAA	CGAGGCGGTGCGTATCTTG
Unknown	AGAP001717	TTTGCCATTGTCCCGTTTGC	AACTGTTTCGGTTCGGTGGT
Xanthine Dehydrogenase	AGAP006226	CTAATCTCGGCTGACACGC	GCTGTTAGCTTTGTGCACCTT
LRIM8B	AGAP007456	ACGATGACGATCACACGGATT	CGTCACGCACTAGACAGGTT
Semialdehyde Synthase	AGAP008632	GTTTATGCACGTTGGACCCG	TCCGATCGATTTGGGCATCA
Unknown	AGAP010047	TAGCTGCTACGACGAATCGC	ACTGCCACATCCAGCAACAT
Short-Chain Dehydrogenase/Reductase	AGAP011852	CGAACTCGTTCCACAATGCG	ACAATGACCGCCGGATTGAC
S7	AGAP010592	AGAACCAGCAGACCACCATC	GCTGCAAACTTCGGCTATTC

EF	AGAP005128	GGCAAGAGGCATAACGATCAATGCG	GTCCATCTGCGACGCTCCGG

Table 2.1: qPCR primer list. Forward and reverse primers used for all qPCR reactions. All primer products are between 80 and 150 base pairs and follow MIQE guidelines.

Preparation of antibodies

All antibody staining and confocal microscopy was performed by J Vontas et al, University of Crete.

Fragments encoding unique peptides for *CYP6Z1* and *CYP4G17* were cloned into the pET 16b vector. Upon expression, the resulting His-tagged peptide was purified to homogeneity by Ni-NTA affinity chromatography and used to raise rabbit polyclonal antibodies. The CYP6Z1 peptide sequence was: VALRDLNNPDSFINNIRTAGVFLCPGLLKFTGINSLSPPMKKFTTEVISSHLHQRETGQ VTRKDFIQMLTDLRRKAGSSGEETLTDA and the CYP4G17 peptide: KRQLKIHLRLDPLFNLTGVKKEQERLLQIIHGLTRKVVREKKQLYERQMAEGKMPSPS LTEIIGKEEKPGEGQLGGSPAFISQ. The antibody for CYP6Z2 was a gift from Dr Mark Paine (LSTM, UK). Rabbit antibodies to CYP6Z2 were prepared to the C terminal peptide sequence MRIDHRK by Moravian Biotechnology, Brno, Czech Republic.

Immunofluorescence and microscopy

Female mosquitoes (three-five days old) were fixed in cold solution of 4% paraformaldehyde (PFA) (methanol free, Thermo scientific) in phosphate-buffered saline (PBS) for four hours and then were cryo-protected in 30% sucrose/PBS at 4^0 C for 12 h. Finally, mosquitoes were immobilized in optimum cutting temperature (O.C.T.) (Tissue-Tek, SAKURA) and stored at -80^0C until use.

Immunofluorescent analysis, followed by confocal microscopy, was performed on longitudinal sections of frozen pre-fixed mosquito specimens. More detailed, 10um sections, obtained in Leitz kryostat 1720 digital, were washed (three x five minutes) with 0.05% Tween in PBS and blocked for three hours in blocking solution (1% Fetal Bovine Serum, biosera, in 0.05% Triton/PBS). Then, the sections were stained with rabbit primary antibodies in 1/500 dilution, followed by goat anti-rabbit (Alexa Fluor 488, Molecular Probes) (1/1000) that gave the green fluorescence. Also To-PRO 3-Iodide (Molecular Probes), which stains DNA specifically (red fluorescence), was used, after RNAse A treatment. As controls, pre-immune serums (in 1/500 dilutions) and anti-rabbit (Alexa Fluor 488, 1/1000) were tested, in parallel with a-P450's to check specificity of each primary antibody. Finally images were obtained on Leica TCS-NT Laser Scanning microscope using the 40x-objective.

Study Two: Meta-analysis

Microarray analysis

Microarray datasets, and associated metadata were provided by members of the Department of Vector Biology at LSTM. All microarrays were analysed using the limma R package, applying linear models to correct and normalise data, inferring differential transcript expression. Data was normalised using affycoretools (MacDonald, 2008). Both within (loess), and between (aquantile) array normalisation, in addition to background correction (mle) were performed. Dye swap correction was used where necessary.

The data had been obtained using two different experimental designs. For some

experiments, RNA extracted from one resistant strain was compared directly

with RNA extracted from a susceptible strain, in other instances, complete loop

designs were used, comparing each extracted RNA with multiple strains. Both R

vs S and complete loop methodologies were represented in the arrays. In the

case of complete loop designs, a design vector with susceptible population as the

reference was used to calculate the contrast matrix. False discovery rate testing

was used for multiple test correction. All other parameters were kept as default.

(Code available at:

https://github.com/VAI3064/PhD-Code/blob/master/limma%20function.R,

https://github.com/VAI3064/PhD-Code/blob/master/MyNormalizeBetweenArrays.R,

https://github.com/VAI3064/PhD-Code/blob/master/MyNormalizeWithinArrays.R,

https://github.com/VAI3064/PhD-Code/blob/master/myMA.RG.R,

https://github.com/VAI3064/PhD-Code/blob/master/myMA2.RG.R)

Metadata collection
Metadata associated with each array was collected, including insecticide

exposure, geographical location, elevation, species and resistant and susceptible

strain. A single metric for geographical distance was calculated in two ways;

firstly, inputting longitude and latitude into a PCA analysis and taking

component one as the measure or secondly, by producing generalised distances

by drawing five equally sized blocks across the African continent. In order to

determine the importance of these factors, the variance associated with each was

calculated using η^2 and ω^2. ω^2 is an unbiased estimator and was therefore used

for all graphical plots.

$$\eta^2 \; = \; \frac{SS_{TREATMENT}}{SS_{TOTAL}}$$

$$\omega^2 \; = \; \frac{SS_{TREATMENT} \; - \; (\alpha - 1)\, MS_{ERROR}}{S_{TOTAL} \; + \; MS_{ERROR}}$$

SS= Sum of squares, MS = Mean square, $(\alpha - 1)$ = degrees of freedom.

Clustering

Data to be clustered was split into data subsets, dependent upon the number of arrays (n) each transcript was significant in (p ≤ 0.05). For all available arrays: n > one, n >10 and n > 20. A smaller subset of arrays was also used to reduce the experimental variables; in this instance: *An. coluzzii* arrays exposed to a pyrethroid and compared to N'Gousso susceptible reference and *An. arabiensis* array exposed to a pyrethroid. In the case of multiple arrays for the same experimental subset, only a unique resistant population was included. For the reduced subset: n > one, n > five and n > eight. Hierarchical clustering and factor analysis were performed on the relative fold changes of the datasets, with hierarchical clustering using Euclidean distance as the preferred distance metric and following psych guidelines (Revelle, 2015 1.5.6) for factor analysis. To allow visualisation of the correlation or anti-correlation of the arrays, factor analysis was also used. Correlation matrices used in factor analysis and visualisations were performed using default settings in R. The optimal number of clusters for each subset was determined using the nFactor R package, which uses Kaiser rule, Cattell's scree test and Horn's parallel analysis to calculate. (Code available at: https://github.com/VAI3064/PhD-Code/blob/master/Clustering.R)

Enrichment

Each cluster was split into up and down regulation across each group of arrays -

An. coluzzii arrays and *An. arabiensis* arrays. The resultant gene lists were

searched for detoxification gene family members (cytochrome p450s, GSTs,

UDPs, COE/CCEs and ABCs), GO term, domain, reactome and pathway

enrichments. Enrichments were performed using DAVID and Flymine, which also

display publication enrichment. The phyper function in R was used to determine

enrichment where stated. Significance was determined using $p \leq 0.05$.

GO term attribution

GO terms were assigned using Blast2Go software. Peptide sequences were

extracted from BioMart for the significant transcripts ($p \leq 0.001$) in each array.

These peptides were used to BLASTp, InterPro scan, map, annotate and produce

annotation statistics. Annotation statistics were outputted for each of the three

GO term categories and concatenated into one file after running GOSlim

(http://geneontology.org/page/go-slim-and-subset-guide).

GaGa analysis

A bespoke GaGa (Rossell, 2009) pipeline was used to assign all transcripts into

predefined patterns, using a gamma-gamma distribution model; in this case *An.*

arabiensis and *An. coluzzii* following either (a) the same pattern (b) a different

pattern. As with the limma analysis, the same corrections were performed for

array normalisation. (Code available at:

https://github.com/VAI3064/PhD-Code/blob/master/GaGaFunction.R,

https://github.com/VAI3064/PhD-Code/blob/master/MyNormalizeBetweenArrays.R,

https://github.com/VAI3064/PhD-Code/blob/master/MyNormalizeWithinArrays.R,

https://github.com/VAI3064/PhD-Code/blob/master/myMA.RG.R,

https://github.com/VAI3064/PhD-Code/blob/master/myMA2.RG.R,

https://github.com/VAI3064/PhD-Code/blob/master/FCs.R,

https://github.com/VAI3064/PhD-Code/blob/master/GO.R,

https://github.com/VAI3064/PhD-Code/blob/master/GO_test2.R,

https://github.com/VAI3064/PhD-Code/blob/master/all_orthologs_search.R,

https://github.com/VAI3064/PhD-Code/blob/master/anopheles_descriptions2.R,

https://github.com/VAI3064/PhD-Code/blob/master/colouring.R,

https://github.com/VAI3064/PhD-Code/blob/master/colourmatrix.R,

https://github.com/VAI3064/PhD-Code/blob/master/compare_lists.R,

https://github.com/VAI3064/PhD-Code/blob/master/culex_with_anopheles2.R,

https://github.com/VAI3064/PhD-Code/blob/master/fold_changes2.R,

https://github.com/VAI3064/PhD-Code/blob/master/full_descriptions2.R,

https://github.com/VAI3064/PhD-Code/blob/master/gene_means.R,

https://github.com/VAI3064/PhD-Code/blob/master/heatmaps.R,

https://github.com/VAI3064/PhD-Code/blob/master/heatmap_production.R,

https://github.com/VAI3064/PhD-Code/blob/master/match_drosophila2.R,

https://github.com/VAI3064/PhD-Code/blob/master/orthologs2.R,

https://github.com/VAI3064/PhD-Code/blob/master/orthologs_with_anopheles2.R,

https://github.com/VAI3064/PhD-Code/blob/master/pathway_test.R,

https://github.com/VAI3064/PhD-Code/blob/master/pathways2.R,

https://github.com/VAI3064/PhD-Code/blob/master/reactome.R,

https://github.com/VAI3064/PhD-Code/blob/master/reactome_test.R,

https://github.com/VAI3064/PhD-Code/blob/master/split_for_drosophila.R)

Transcription factor identification
A transcription factor database, FlyTF.org, exists for *Drosophila*, which includes a

list of curated putative transcription factors (Adryan & Teichmann, 2006). These

data were downloaded and converted to *Anopheles* homologs using FlyMine

(Lyne *et al.*, 2007). These homologs were used to search the transcript list

outputs from GaGa to identify transcription factors. The resultant lists were then

searched for transcription factors significant in at least eight out of 10 of the

arrays.

Study Three: Deltamethrin candidate search

Mosquito rearing conditions
Mosquito rearing conditions for Tiassalé and N'Gousso strains are described in

Study One. The strain VK7 is an *An. coluzzii* population originating from Valle de

Kou, Burkina Faso. VK7 is resistant to both DDT and deltamethrin but is

susceptible to bendiocarb, with an LD_{50} compared to N'Gousso at the time of the

study of 11-fold for permethrin and 13-fold for deltamethrin.

qPCR
qPCR was performed on both whole organism extractions and dissected body

parts. The qPCR methodology and primers are described in Study One and

Tables 2.1 and 2.2.

Transcript ID	Forward Sequence	Reverse Sequence
SAP2 (AGAP008052)	GCAGCTTGAGAGCGTCTTCT	GAAAGCGATGGCGACGAACA
EF-like protein (AGAP002603)	AAAACCATCCCAGCAAACGTG	TCGTAATCATCGTAGACAGCG
CYP6Z3 (AGAP008217)	CCACGCAATTGCATTGGTCT	TTCTACGCGCATGGGGAAAC
UGT302A1 (1) (AGAP006222)	TGGCTGGGGTATCGGAGTTA	GCAGTATAGCTTGGGTCACCG
PT3 (AGAP004262)	AAAGGAACTGAAGCCTGCGA	TCTTCCAAATGTTTCTTAGTTCCCT
Acyl-CoA Dehydrogenase (AGAP008769)	TTCTTCGAGGAGGACGGCTAC	CATCAGGCCCATCTCCCACA
ABCB4 (AGAP006364)	AGAAAAGCAACGGGTAGCGA	ATTGTCTGTCGCTCTTGCCA

α-Crystallin Chain B (AGAP007160)	CGGGCATCAGTATCGCTAGG	GGGCTTCAACACCGATGGAA

Table 2.2: Primer sequences for qPCR. Sequences of primers used in all qPCR reactions.

RNAi

PCR was performed using Phusion® High-Fidelity DNA Polymerase (Thermo Scientific) following manufacturer's instructions and primer sets with a T7 promoter sequence at the 5' end of both the sense and antisense primers (Table 2.3). Primers were designed to have an asymmetric product sequence with a length of 300-600bp, a GC content of 20-50% and no more than three consecutive nucleotides. PCR was performed with the following cycle: three minutes 98°C, 35 cycles of seven seconds at 98°C and 10 seconds at 72°C, with a final hold at 72°C for seven minutes. PCR products were resolved on 1.5% agarose gels for 45 minutes and the correct length amplifications identified. The PCR products were purified using a Qiagen QIAquick PCR Purification kit following manufacturer's instructions. dsRNA was synthesised using a Megascript® T7 Transcription (Ambion) kit, with a 16 hour 37°C incubation, following manufacturer's instructions. The dsRNA was then cleaned using a MegaClear® Transcription Clear Up (Ambion) kit, with DEPC water, twice heated at 65°C for 10 minutes, to elute the sample. The resultant dsRNA product was analysed using a nanodrop spectrometer (Nanodrop Technologies, UK) and subsequently concentrated to 3µg/µl using a vacuum centrifuge at 35°C. Injections were then carried out using a nanoinjector with 69nl of product injected directly into the thorax, between the cuticle plates of the abdomen, underneath the wing. Injections were carried out on 100 three-five day old,

presumed mated, non-blood fed females that were immobilised on a CO_2 block.

As a control non-endogenous *GFP* dsRNA was injected at the same amount and

concentration.

Transcript ID	Forward Sequence	Reverse Sequence
SAP2 (AGAP008052)	taatacgactcactataggg TTCTCGTTCCGGTTGCTTCA	taatacgactcactataggg TAGTAGACCCCATTCCCCACTT
SAP2 (2) (AGAP008052)	taatacgactcactataggg TTCAAGTGCCTGATGGACGA	taatacgactcactataggg AGTATCACAAAGCGACCACCAT
EF-like protein (AGAP002603)	taatacgactcactataggg ATGTCTCGCCACCGAAATGT	taatacgactcactataggg AAGCTTCTTCTCCGAGCGTT
CYP6Z3 (AGAP008217)	taatacgactcactataggg GACACTGCAGCGTTTGACTA	taatacgactcactataggg GCAGCTTTGCCATTGCTTCT
UGT302A1 (1) (AGAP006222)	taatacgactcactataggg TGGGTTCCATGTCAACCGTAA	taatacgactcactataggg CATTCGACGGGAAATCGTCT
UGT302A1 (2) (AGAP006222)	taatacgactcactataggg TCTGCTGAGTTGCACCGAAT	taatacgactcactataggg ATGAGATGCTTTGCGCCATC

Table 2.3: Primer sequences for dsRNA. Sequences of primers used for dsRNA

synthesis, each sequence appended to the T7 binding site (shown in lower case).

Localisation of candidate genes
Antennae, heads and legs were dissected and whole bodies taken from three-five

day old presumed mated females. Dissected structures were taken in three

biological replicates with pools of 40-50 mosquitoes. Whole bodies were also

taken in three biological replicates from pools of seven-10 females.

Induction
SAP2 induction assays were carried out on three-five day old, presumed mated, non-blood fed Tiassalé females at three time points post deltamethrin bioassay tube exposure: 24 hours, 48 hours and 72 hours. These time points were compared to an unexposed pool of female mosquitoes after qPCR.

Bioassays
WHO bioassay tube test kits (World Health Organization, 2013) were used under standard conditions for one-hour exposures to 0.05% deltamethrin and 30 minute exposures to 0.01% bendiocarb impregnated papers. Mosquitoes were then left in a control tube, under insectary conditions for 24 hours and mortality recorded.

Survival analysis
72 hours post injection three pools of 25 injected and 25 uninjected adult female mosquitoes were transferred into a cup and fed twice daily with 1:10 glucose solution. Numbers of dead mosquitoes were recorded every 24 hours as a proxy for fitness. Statistical analysis and graphical representation was produced using the survival package (Therneau & Grambsch, 2000) in R.

Localisation Enrichment
MozAtlas (Chintapalli *et al.*, 2007) and data generated in Chapter 1 (Ingham *et al.*, 2014) were used to identify body parts in which chemosensory transcripts were expressed in *An gambiae.*

Study Four: *Keap1-Maf-S-cnc* pathway

Probability of a transcript being differentially regulated in n arrays by chance
A binomial test was carried out to determine which of the transcripts were significant in more or less arrays than expected by chance. The probability of

success, in this case significance, was calculated using the average number of all significant transcripts in all arrays. The probability of success, over all arrays was 0.4692 and that of failure, 0.5308.

Enrichment tests

Enrichments were performed using DAVID functional annotation for transcript lists that were longer than 100 transcripts. Smaller transcript lists were assessed for specific enrichments using a hypergeometric test. In both cases, significance was determined as $p \leq 0.05$.

Mosquito rearing conditions

See Study One: Mosquito rearing conditions for details.

RNAi

RNAi was performed as in Study Three: RNAi with primers from Table 2.4.

RNAi Transcript	Forward Sequence	Reverse Sequence
Maf-S (1) (AGAP01405-RA)	taatacgactcactatagggAGGT GTGTGCTGTGCAAGAT	taatacgactcactatagggAGGT GTGTGCTGTGCAAGAT
Maf-S (2) (AGAP01405-RA)	taatacgactcactatagggACG ATGAGCTGGTGTCGATT	aatacgactcactatagggAGTG AAACATTCGGCACGGT
GFP	taatacgactcactatagggAGAAC GTAAACGGCCACAAGTTC	taatacgactcactatagggAGACTT GTACAGCTCGTCCATGCC

Table 2.4: RNAi primers. RNAi primers used in the synthesis of dsRNA, each with a T7 promoter region (shown in lower case).

qPCR

qPCR was performed on whole post knockdown Tiassalé cDNA using four transcripts: AGAP008212-RA, AGAP004382-RA, AGAP008358-RA and AGAP007504-RA. For qPCR methodology see Study One: qPCR, primers are shown in Table 2.5, housekeeping controls used were standard *EF* and *S7* (Table 2.1)..

Transcript ID	Forward Sequence	Reverse Sequence
CYP6M2 (AGAP008212-RA)	TACGATGACAACAAGGGCAAG	GCGATCGTGGAAGTACTGG
GSTD3 (AGAP004382-RA)	CCGCACATTAAGGGATGGGT	CGCTTTGGTAGCTTCCTCCA
CYP4H17 (AGAP008358-RA)	TACTGACCAGTGCCTTGCTG	GTGCCATTCATTTCCATGTCCT
ABCA3 (AGAP007504-RA)	CACGGACATCGTTGGTGGTA	TCGGTGGGATCTTCCTCCAT

Table 2.5: Primer sequences for *Maf-S* knockdown qPCR. Sequences of primers used in all qPCR validation.

Microarrays

A whole-genome microarray approach was used to determine the effect of *Maf-S* knockdown on transcriptional profiles. The transcriptional profiles of *Maf-S* knockdowns were compared against a GFP injected control. All mosquitoes used in the microarray were adult females from the Tiassalé strain, five to eight days old. RNA was extracted as above 72 hours post injection. The protocol used for the microarray is as in above, Study One: Microarray experiments. Submitted to ArrayExpress, accession E-MTAB-4042.

Microarray analysis

For microarray analysis methods, see Study Two: Microarray analysis.

Pathway analysis

KEGG pathway (Kanehisa & Goto, 2000) was explored using both the colour pathway web interface and the Cytoscape 3.1 plugin, CytoKEGG.

Correlation networks

Correlation networks were produced using a correlation matrix with a Euclidean distance metric, with all transcripts from all 27 available microarrays [Chapter

4]. To identify only arrays with a strong correlation or anti-correlation, only those transcripts with a correlation of ±0.8 were used. These data were extracted and visualised using $\log_2$ fold change and array index.

Study Five: *Met* involvement with public health insecticides

Mosquito rearing conditions
See Study One: Mosquito rearing conditions for details.

RNAi
RNAi was performed as in Study Three: RNAi with the primers in Table 2.6.

RNAi Transcript	Forward Sequence	Reverse Sequence
Met (AGAP006022-RA)	taatacgactcactatagggAACGAGAATGGGCTTTAGGAA	taatacgactcactatagggTGGTCCGAAATGGTGTAAGG
Met (AGAP006022-RA) (2)	taatacgactcactatagggTTAGCGCGTGTGATGAAGGT	taatacgactcactatagggTGCACCACTTTGGTATCGCT
GFP	taatacgactcactatagggAGAACGTAAACGGCCACAAGTTC	taatacgactcactatagggAGACTTGTACAGCTCGTCCATGCC

Table 2.6: RNAi *Met* primers. Primer list used in dsRNA construction for *Met*.

qPCR

qPCR was performed as in Study One: qPCR using primers listed in Table 2.7, housekeeping controls used were standard *EF* and *S7* (Table 2.1).

Transcript	Forward Sequence	Reverse Sequence
Met (AGAP006022-RA)	TGCGGGTGGACTATGTGTTT	ACGTCACCGTGAGCAGAAAT

Table 2.7: *Met* qPCR primers.

Microarrays
A whole-genome microarray approach was used to determine the effect of *Met* knockdown on transcriptional profiles. The transcriptional profiles of *Met-RNAi*

knockdowns were compared against a GFP injected control. The protocol used

for the microarray is as in above, Study One: Microarray experiments. Submitted

to ArrayExpress, accession number: E-MTAB-4043.

Microarray analysis
For microarray analysis methods, see Study Two: Microarray analysis.

Enrichments and Pathway Analysis
Sets of differentially expressed transcripts were analysed using DAVID (Huang *et*

al., 2007) to identify significantly enriched GO terms, and InterPro ID clusters.

Transcript lists were uploaded using Ensembl-transcript IDs. KEGG pathway

(Kanehisa & Goto, 2000) was used to colour the pathways in the web interface

and the Cytoscape 3.1 plugin, CytoKEGG.

Bioassays
Refer to Study three: Bioassays for protocol.

Correlation networks
Correlation networks were performed as in Study Four: Correlation networks.

Yeast two-hybrid
Drosophila yeast two-hybrid data were used to identify experimentally validated

transcription factor interactors. A plugin for Cytoscape was used to visualise the

data available in DroID (Yu *et al.*, 2008; Murali *et al.*, 2011). Candidate

transcription factors were searched for experimental availability. Where found,

the interactors three steps out were explored for the *a priori* detoxification

families.

Motif identification
Motifs were retrieved from both the JASPAR Core Insect database (Bryne *et al.*,

2008) for *Met* and an experimentally validated *Met* motif (Zou *et al.*, 2013).

These data were used to search 1000bp upstream from the coding sequences of

all transcripts significant in the array. Two R packages, MotifDb (Shannon *et al.*, 2013) and Biostrings (Pages *et al.*, n.d.) were used to iteratively search the upstream regions for the corresponding motifs; the number of occurrences were then recorded and summed (https://github.com/VAI3064/PhD-Code/blob/master/Motif%20Search.R). Using random selections of non-significant transcripts a null model was produced by iterating through the search 10 times and averaging the motif occurrences across each nucleotide position.

3. DISSECTING THE ORGAN SPECIFICITY OF INSECTICIDE RESISTANCE CANDIDATE GENES IN *ANOPHELES GAMBIAE*: KNOWN AND NOVEL CANDIDATE GENES.

V A Ingham, C M Jones, P Pignatelli, V Balabanidou, J Vontas, S C Wagstaff, J D Moore and H Ranson.

Introduction

Insecticides play a vital role in reducing malaria transmission in Africa. An escalation in the use of two very effective tools, indoor residual spraying with insecticides and insecticide treated bednets, has led to impressive reductions in malaria with child death rates halved and more than 3.3 million lives saved since 2000 (WHO, 2014). Inevitably, as insecticide use has intensified, malaria vectors have developed resistance to these chemicals (Ranson *et al.*, 2011; Edi *et al.*, 2012; Edi *et al.*, 2014a). With just four classes of insecticides available for public health and only the pyrethroids approved for bednet treatment, this poses a major challenge to sustaining and extending recent achievements in malaria reduction.

Numerous studies have attempted to identify the genes responsible for insecticide resistance in the major malaria vectors. One of the most potent mechanisms identified to date is increased activity of enzymes that detoxify insecticides (Ranson *et al.*, 2002; Liu *et al.*, 2006; Ranson *et al.*, 2011). Four enzyme families are known to be associated with insecticide metabolism (carboxylesterases (CCEs), glutathione transferases (GSTs), UDP glucornyltransferases (UGTs) and cytochrome P450s (P450s)) and a number of individual enzymes, most notably from the cytochrome P450 family, have been implicated in conferring resistance to one or more insecticide classes (Chiu *et al.*, 2008; Müller *et al.*, 2008b; Mitchell *et al.*, 2012; Edi *et al.*, 2014a). More recently, the importance of interactions between different enzymes and transporters in the insecticide detoxification pathway has been recognised (Ismail *et al.*, 2013; Chandor-Proust *et al.*, 2013). To dissect these pathways further, and to distinguish members of these large gene families with housekeeping functions from those more likely to have

detoxification roles, further information on their sites of expression is required. The first objective of this study was to characterise expression patterns of the key gene families associated with insecticide resistance across the major organs linked to xenobiotic detoxification in insects, with a particular focus on the P450s. Although transcriptomes of many of the key tissues in *An. gambiae* have already been described (Marinotti *et al.*, 2006; Koutsos *et al.*, 2007; Neira Oviedo *et al.*, 2008; Baker *et al.*, 2011), this study used material from a highly insecticide resistant strain and from a susceptible strain to identify genes whose tissue-specific enrichment might be linked to the resistance phenotype.

To date, all comparisons of the transcriptome between insecticide resistant and susceptible malaria vectors have compared gene expression in the whole organism. This approach has the potential to miss candidates. If, for example, expression of a gene is restricted to an organ that contributes only a small proportion of mRNA to the total RNA pool, or differential expression occurs in only one tissue, even large differences in expression between a resistant and susceptible population may not be detectable (Johnson *et al.*, 2013). Thus the second objective of the study was to compare gene expression in key body parts between an insecticide resistant and susceptible strain of mosquito to identify candidates not immediately apparent in whole organism microarray studies. Adult mosquitoes were dissected into body parts that could be readily separated with minimal risk of contamination, and are suspected to be involved in metabolic resistance.

Aims

1. Identify body part specific expression for known detoxification transcripts.

2. Explore transcript expression for all detoxification family members and their respective localisations.

3. Discover new candidate insecticide resistance candidates and their body part enrichments.

4. Compare the efficacy for candidate discovery of whole organism arrays and the body part dissection arrays.

Results

RNA was extracted from three dissected body parts: the Malpighian tubules, the midgut and the abdomen integument (including the fat body but also epidermal, neuronal, muscle and oenocyte cells) with the remaining undissected portions forming a fourth group (including head, antennae, thorax, legs and wings). Each biological replicate consisted of 15-20 adult female mosquitoes from the major malaria vector *An. gambiae*. Dissections were performed on a lab susceptible population, N'Gousso, originally from Cameroon and a population resistant to all four major classes of insecticide, Tiassalé, originally from Côte D'Ivoire.

Transcription in each body part was compared (i) directly to the corresponding whole organism for both Tiassalé and N'Gousso and (ii) resistant against the corresponding susceptible body part (Figure 3.1).
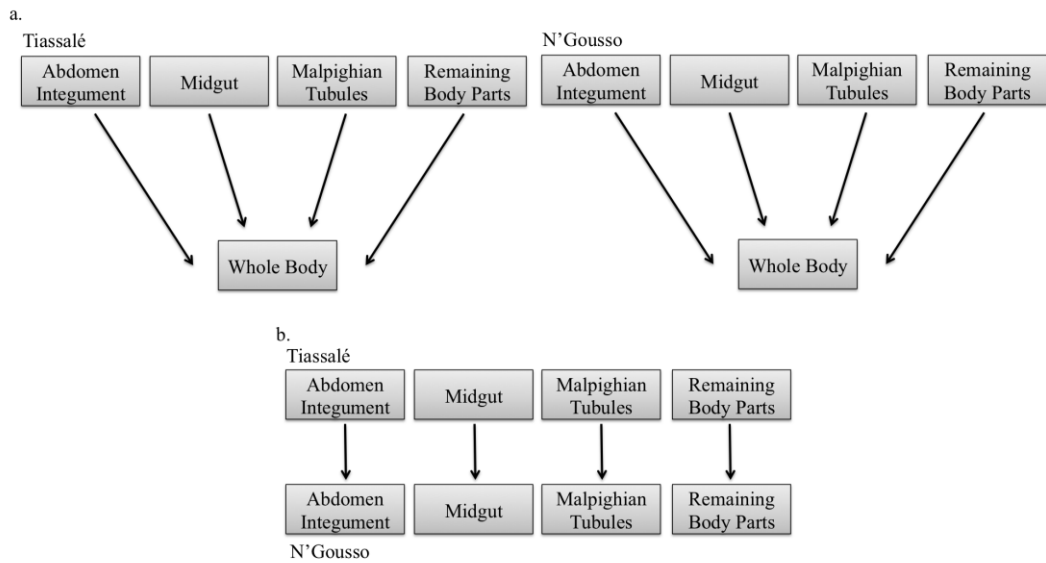
Figure 3.1: Schematic of design of microarray experiment. a. Sample vs reference design, for each of the susceptible lab population N'Gousso and the resistant population Tiassalé b. Resistant (T) vs Susceptible (N) design, for each individual tissue.

Body Part Specific Transcript Enrichment: Overview.
Probes showing enriched expression in each of the four body parts were determined using a multiple test correction significance cut-off of $p \leq 0.05$ for both the resistant and susceptible strains. As expected, a clear positive correlation can be observed for local gene expression between the resistant and susceptible strains in each body part. This is illustrated in Figure 3.2, where the vast majority of probes follow a y=x trend.
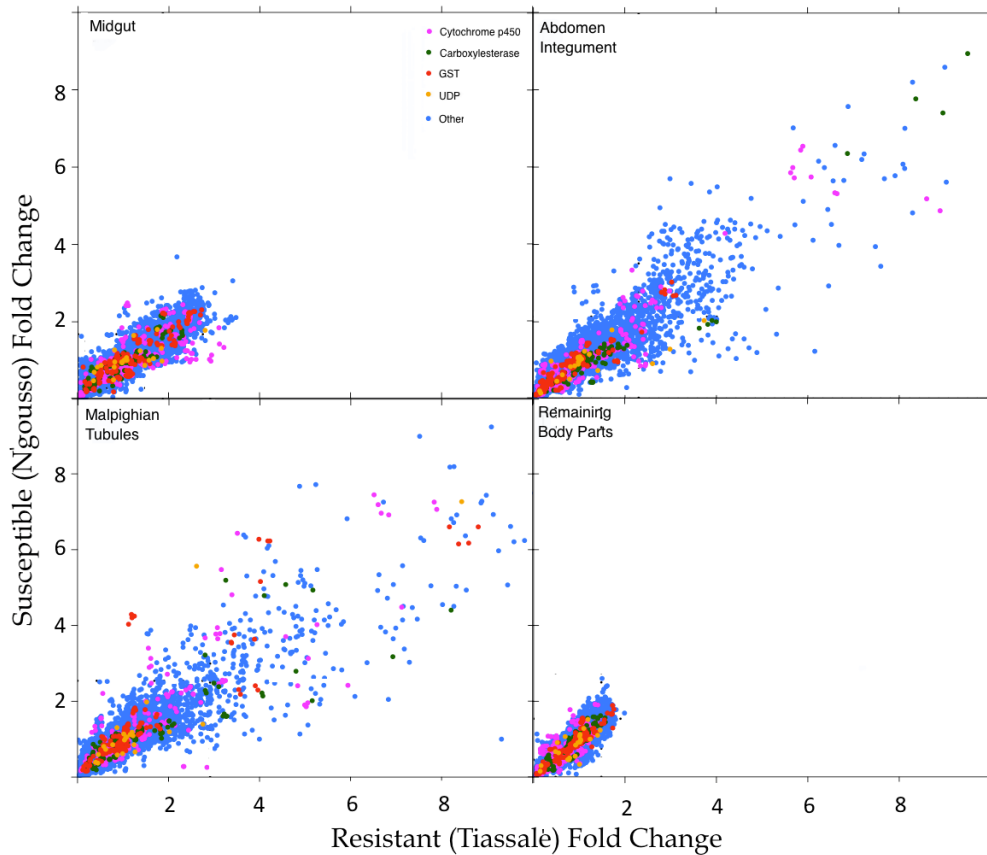
Figure 3.2: Gene enrichment in individual body parts in insecticide susceptible and resistant mosquitoes. Fold change in transcription of all probes in individual tissues against the reconstituted whole organism, plotted for Tiassalé (resistant, x) and N'Gousso (susceptible, y). Probes for the four detoxification families are indicated in pink (cytochrome p450s, 393 probes), green (carboxylesterases, 168 probes), red (GSTs, 152 probes) and yellow (UDPs, 26 probes).

In all body parts, similar numbers of probes showed enriched or depleted expression compared to the whole organism, with a range of 1.4% to 4.4% of the total probes (Table 3.1). The magnitude of change in expression of individual probes in the midgut and remaining body part is relatively low with no probes exceeding $\log_2$ fold change of four, compared to the abdomen integument and

malpighian tubules where eight and 61 genes respectively are above eight-fold enriched (Figure 3.2).

	Number Detoxification probes	Percentage Detoxification	Enrichment p-value`
Abdomen Integument			
Overexpressed in Tiassalé	16	7.58	0.0012
Overexpressed in N'Gousso	9	4.27	0.12
Overexpressed in both strains	7	3.32	0.16
Underexpressed in Tiassalé	0	0.00	
Underexpressed in N'Gousso	0	0.00	
Underexpressed in both strains	0	0.00	
Non-differential	179	84.83	
Midgut			
Overexpressed in Tiassalé	29	13.74	1.80E-09
Overexpressed in N'Gousso	25	11.85	8.97E-09
Overexpressed in both strains	19	9.00	1.41E-06
Underexpressed in Tiassalé	2	0.95	0.52
Underexpressed in N'Gousso	8	3.79	0.94
Underexpressed in both strains	1	0.47	0.94
Non-differential	127	60.19	
Malpighian Tubules			
Overexpressed in Tiassalé	18	8.53	2.66E-07
Overexpressed in N'Gousso	19	9.00	1.49E-08
Overexpressed in both strains	14	6.64	2.25E-06
Underexpressed in Tiassalé	0	0.00	
Underexpressed in N'Gousso	0	0.00	
Underexpressed in both strains	0	0.00	

Non-differential	160	75.83	
Remaining body parts			
Overexpressed in Tiassalé	4	1.90	0.65
Overexpressed in N'Gousso	8	3.79	0.21
Overexpressed in both strains	3	1.42	0.32
Underexpressed in Tiassalé	39	18.48	1.82E-14
Underexpressed in N'Gousso	39	18.48	2.22E-15
Underexpressed in both strains	31	14.69	2.85E-12
Non-differential	87	41.23	

Table 3.1: Overview of probes over- or under-transcribed in each body part for both the resistant and susceptible strains when compared to the whole organism. The local transcription of probes in each of the two mosquito strains is expressed as total number of probes over- or under-transcribed in a particular body part, compared to the reconstituted whole organism, as a percentage of the total probes on the array. For genes that were represented by multiple probes, an average of all the probes was used. `p-value determined by hypergeometric test, significance p ≤ 0.05.

Body Part Specific Transcript Enrichment: Detoxification Genes.
The cytochrome P450 family has been most strongly linked with insecticide resistance in *Anopheles* mosquitoes, with several enzymes capable of detoxifying insecticides from more than one class. Identifying the primary sites of expression of this enzyme family will aid prediction of function and help identify the key organs largely responsible for insecticide detoxification in resistant mosquitoes. Within the P450 gene family, body part enrichment shows some relationship with the gene tree clustering (Figure 3.3) with the *CYP9J*, *6A* and *6P* families largely

enriched in the midgut, the *CYP4G*s, *6Y*s and *325C*s enriched in the abdomen integument and the *CYP6Z* family enriched in the malpighian tubules. The diversity in enrichment patterns within this gene family led us to look specifically at thirteen cytochrome P450s that had been implicated in insecticide resistance to see if these were enriched in a particular body part (Figure 3.4). The criteria for inclusion of these P450s as candidate insecticide resistant genes was that they had been found to be significantly over-expressed in pyrethroid and/or DDT resistant *An. gambiae* populations in more than one independent study. Four of these were significantly (p ≤ 0.05) enriched in the malpighian tubules (*CYP6M3, CYP6Z1, CYP6Z2* and *CYP6Z3*) compared to whole organism and two were enriched in each of the midgut and abdomen integument (*CYP9J5/CYP4H24* and *CYP4G16/CYP4G17* respectively). The remaining five showed no significant tissue enrichment.
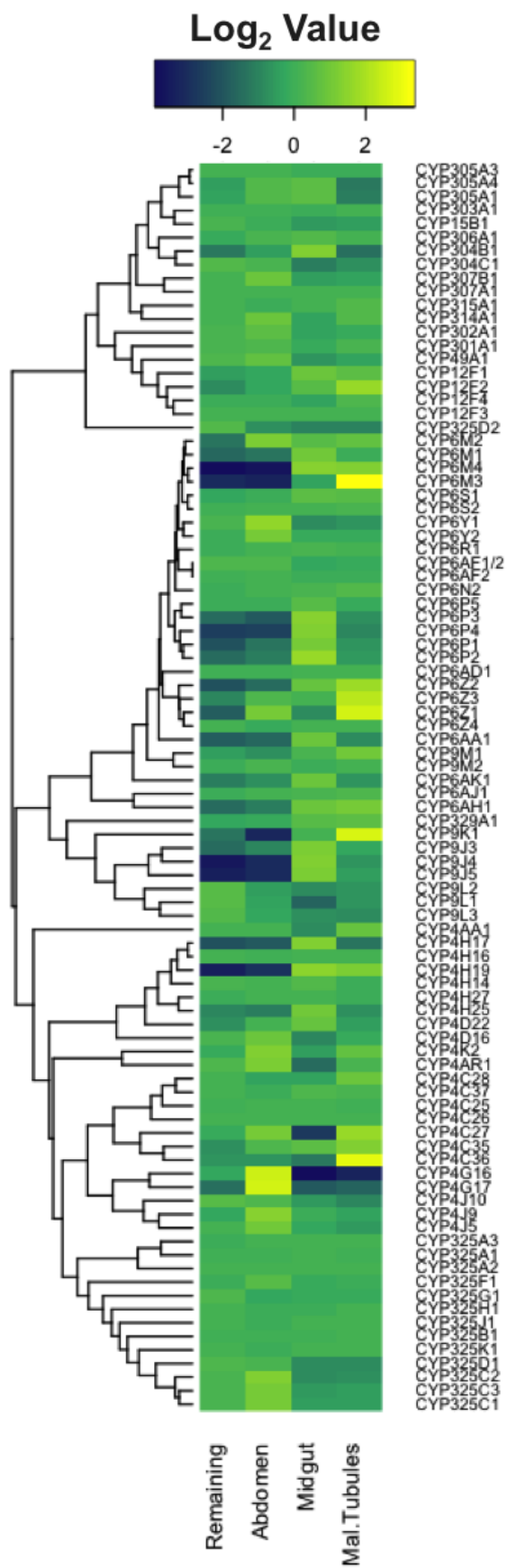
Figure 3.3: Local expression of all cytochrome p450s, following a phylogenetic dendrogam. Full length protein sequence alignment and neighbour joining tree as computed on MEGA5 decorated with local $\log_2$ transcription of Tiassalé cytochrome p450s.
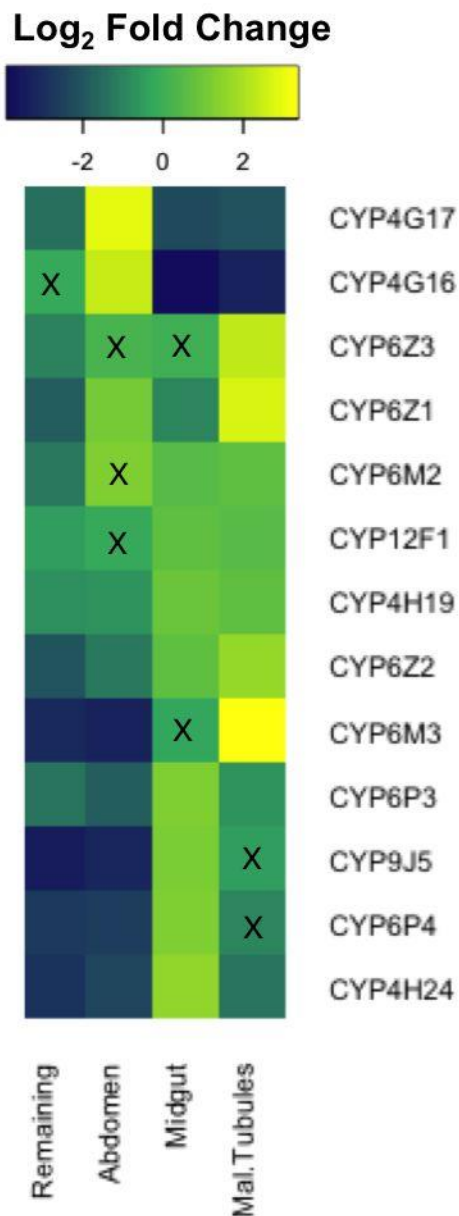
Figure 3.4: Local expression of cytochrome P450s linked to insecticide resistance. Heatmap showing the $\log_2$ fold change of a subset of cytochrome

p450s (implicated in insecticide resistance in previous studies) in different body parts of the Tiassalé strain. Crosses indicate non-significance (p >0.05).

The enrichment of a subset of the cytochrome P450s in particular body parts was confirmed by qPCR (Figure 3.5). In addition, antibodies were available for three cytochrome P450s, which were used to confirm the major sites of expression within the abdomen integument of the Tiassalé resistant strain. In agreement with microarray expression data, *CYP6Z1* and *CYP6Z2* were detected in the malpighian tubules of resistant mosquitoes, whilst *CYP4G17*, identified as enriched in the abdomen integument by both microarray and qPCR, was found only in the oenocytes, a cellular layer located under the abdomen cuticle (Figure 3.6). A *CYP4G* from *Drosophila melanogaster, CYP4G1*, has also been shown to be highly enriched in oenocyte cells; this *Drosophila* enzyme catalyses a key step in the formation of cuticular hydrocarbons (Qiu *et al.*, 2012). The potential role of over-expression of *4G17*, and its paralogue *4G16*, in altering the cuticular structure in insecticide resistant mosquitoes is currently under investigation.
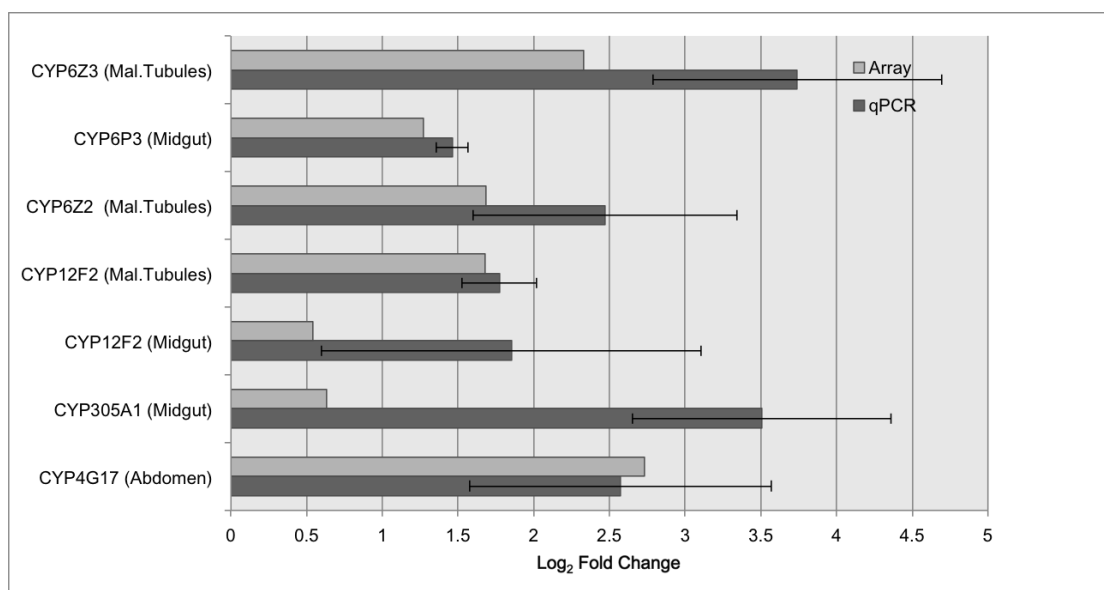
Figure 3.5: qPCR validation of body part enrichment of cytochrome P450s.

Differences in transcription levels in Tiassalé RNA from individual dissected body parts, compared to the reconstituted whole, were measured by qPCR for six P450 genes and compared to the data obtained from the microarray. Data represent $\log_2$ fold change with associated standard error bars.
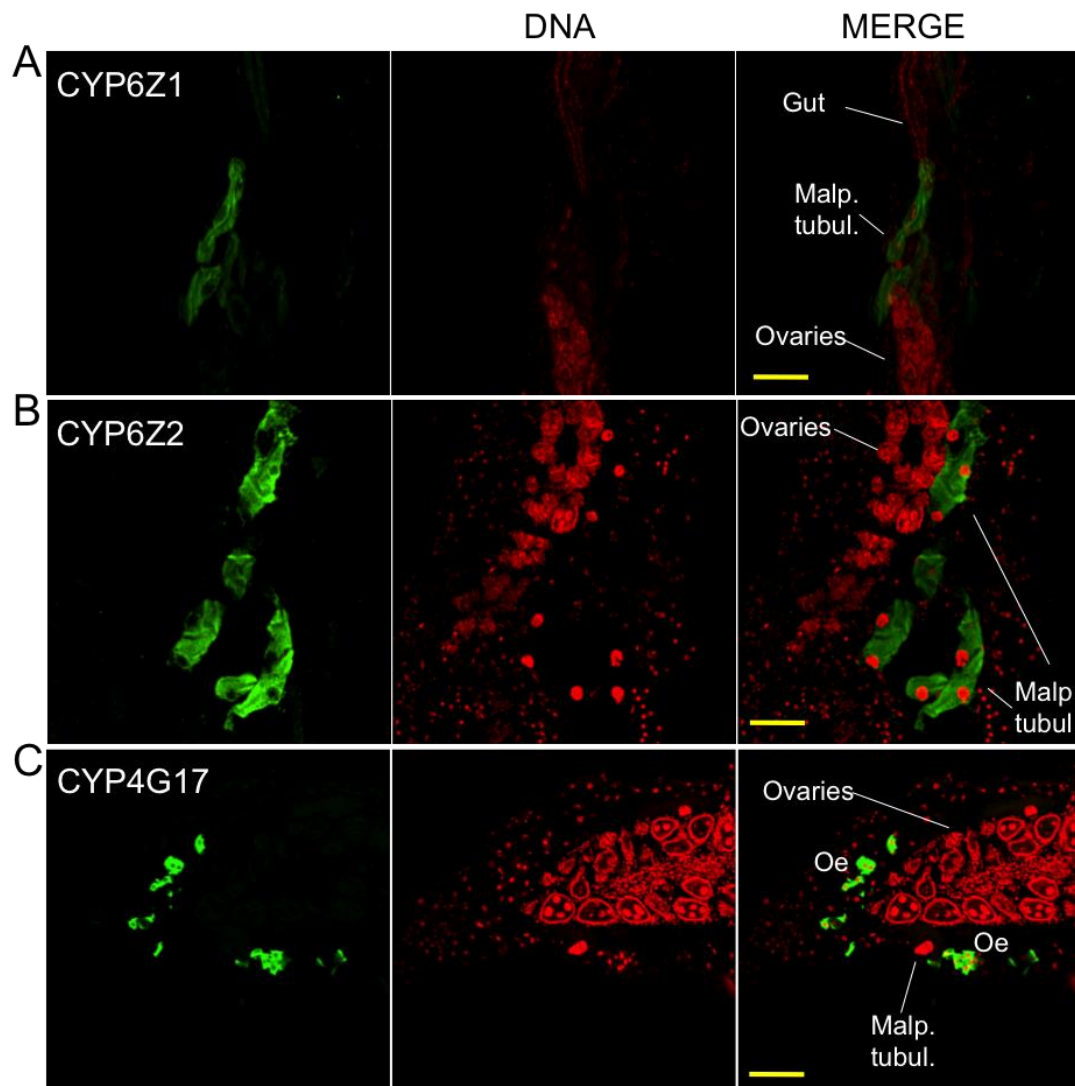
Figure 3.6: Immunohistochemical stainings of cytochrome P450s associated with pyrethroid resistance. Longitudinal sections from resistant mosquito (Tiassalé) specimens immunostained with A) a-CYP6Z1, B) a-CYP6Z2 and C) a-CYP4G17 specific antibodies (left panel, green color). The middle panel shows the same sections stained red using the nucleic acid stain TOPRO. The

merged immunohistochemical staining (P450 stainings and nuclei) appears in the right column. Bar scale (yellow line): 100 nm. Performed by J. Vontas group, University of Crete.

Body Part Specific Transcript Enrichment: Novel Genes.

The two strains used in this study originate from sites separated by approximately 2,500km, so substantial variation in gene transcription between them is to be expected, regardless of the difference in their insecticide resistance profile. Thus caution must be applied when correlating gene transcription levels with the resistance phenotype. However, with a goal of identifying further candidate insecticide resistance genes from the direct comparisons between dissected body parts from the resistant and susceptible strain for further functional validation, we applied a stringent selection process to derive a new gene list. This involved selecting only genes detected as enriched in body part from Tiassalé versus N'Gousso via both the limma and GaGa methods with an adjusted p value ≤ 0.001 and setting a cut off of 1.4-fold differential expression (see methods section). This identified a list of 134 transcripts, representing 105 genes transcribed at higher levels in the Tiassalé strain and 16 genes with higher transcription in the susceptible N'Gousso strain (Table 3.2). Eleven of these transcripts were selected for qPCR validation through having the highest averaged fold change, yielding a positive correlation with the array (Pearson's correlation (r=0.870)) (Figure 3.7).

ID	Description	Body Part	Whole Array Data	Limma	GaGa
AGAP000163-RA	GSTMS2	AI	1.07	0.91	1.21
AGAP000179-RA	cg2867	AI	1.22	0.66	1.61
AGAP000179-RB	amidophosphoribosyltransferase	AI	1.19	0.73	1.67
AGAP000179-RC	amidophosphoribosyltransferase	AI	0.90	0.71	1.64

AGAP000646-RA	diminutive	AI	1.22	0.89	2.08
AGAP000708-RA	phosphorylase b kinase	AI	-0.19	0.64	0.64
AGAP001076-RA	CYP4G16	AI	1.37	0.01	1.46
AGAP001116-RA	d-amino acid oxidase	AI	1.26	0.09	1.29
AGAP001468-RA	niemann-pick c1	AI	0.82	1.05	1.67
AGAP001582-RA	transient receptor potential cation channel protein	AI	-1.55	1.33	4.01
AGAP001649-RA	glucosylceramidase	AI	NA	1.45	1.11
AGAP001717-RA	AGAP001717-PA	AI	1.84	1.04	3.79
AGAP001769-RA	beat protein	AI	NA	1.49	2.27
AGAP001769-RD	beat protein	AI	NA	1.49	2.23
AGAP002154-RB	stretchin- isoform g	AI	0.54	0.87	1.46
AGAP002194-RB	inositol-trisphosphate 3-kinase	AI	2.19	0.86	3.24
AGAP002416-RA	CYP4K2	AI	0.66	0.84	1.62
AGAP002517-RA	AGAP002517-PA	AI	NA	0.76	0.95
AGAP002588-RA	AGAP002588-PA	AI	0.85	1.42	2.82
AGAP002726-RB	cuticular protein isoform a	AI	2.97	1.59	2.54
AGAP002736-RA	isoform b	AI	1.75	0.43	1.65
AGAP002736-RB	isoform b	AI	2.41	0.46	1.94
AGAP002914-RA	galactokinase 1	AI	1.79	1.02	3.03
AGAP002914-RB	galactokinase 1	AI	1.76	1.09	3.04
AGAP002914-RC	galactokinase 1	AI	1.74	1.01	2.95
AGAP003162-RA	protein unc-13 homolog d	AI	1.95	1.26	2.48
AGAP003341-RB	ecotropic viral integration site	AI	0.43	0.76	0.73
AGAP003515-RA	nitrilase member 2	AI	1.00	0.91	3.27
AGAP003734-RA	cg3823 cg3823-pa	AI	0.44	1.04	0.92
AGAP003755-RA	isoform f	AI	1.12	0.50	0.81
AGAP003814-RA	ptpla domain protein	AI	0.46	0.81	1.95
AGAP003826-RA	proteasome assembly chaperone 2	AI	-0.26	1.01	0.63
AGAP004099-RA	isoform a	AI	NA	0.59	0.61
AGAP004248-RA	GPXH3	AI	0.67	1.81	1.84
AGAP004402-RA	sodium-dependent phosphate transporter	AI	0.30	0.65	1.39
AGAP004620-RA	AGAP004620-PA	AI	0.45	1.13	1.08

AGAP005563-RA	sugar transporter	AI	NA	0.24	1.91
AGAP005651-RA	cytoplasmic trna 2-thiolation protein 2	AI	NA	0.66	0.80
AGAP005986-RA	fatty acyl- reductase 1	AI	0.92	0.65	1.17
AGAP006358-RA	ttc27 protein	AI	NA	0.68	1.05
AGAP006610-RA	glutaminyl-peptide cyclotransferase	AI	0.30	0.56	1.31
AGAP006727-RA	alpha-esterase	AI	0.73	0.54	1.59
AGAP006741-RA	isoform d	AI	NA	0.92	2.33
AGAP006769-RA	AGAP006769-PA	AI	2.54	1.83	3.07
AGAP006788-RA	arylalkylamine n-acetyltransferase	AI	0.65	2.08	2.83
AGAP006921-RA	ced-6	AI	0.33	0.30	0.72
AGAP006921-RB	ced-6	AI	0.36	0.35	0.69
AGAP006934-RA	AGAP006934-PA	AI	0.14	0.62	0.69
AGAP007029-RA	glucosyl glucuronosyl transferases	AI	1.77	1.46	2.01
AGAP007045-RA	carboxypeptidase n subunit 2	AI	2.34	1.52	3.01
AGAP007319-RA	AGAP007319-PA	AI	0.78	2.09	2.73
AGAP007456-RA	**membrane glycoprotein lig-1**	**AI**	**2.33**	**2.32**	**4.55**
AGAP007490-RA	isoform a	AI	0.24	0.81	0.82
AGAP007691-RB	serpin 18 non-inhibitory serine protease inhibitor	AI	NA	0.70	0.88
AGAP007691-RC	serpin 18 non-inhibitory serine protease inhibitor	AI	NA	0.86	0.92
AGAP008584-RA	lysosomal alpha-mannosidase	AI	1.74	0.33	0.55
AGAP008889-RA	abc transporter	AI	0.54	1.00	1.41
AGAP009110-RA	AGAP009110-PA	AI	2.43	0.35	3.05
AGAP009214-RA	serine protease	AI	0.79	1.02	0.77
AGAP009521-RA	ankyrin erythrocytic	AI	NA	1.29	1.86
AGAP009790-RA	chitin binding peritrophin-	AI	1.33	0.95	1.86
AGAP009842-RA	ribonuclease t2	AI	1.09	0.76	3.03
AGAP010047-RA	**salivary cys-rich secreted peptide**	**AI**	**1.08**	**1.54**	**3.76**
AGAP010681-RA	cytosolic 5 -nucleotidase iii	AI	0.50	1.05	1.45
AGAP010928-RA	tetratricopeptide repeat protein 36	AI	0.89	1.26	3.00
AGAP010935-RA	delta-aminolevulinic acid dehydratase	AI	0.80	1.39	1.92
AGAP010962-RA	AGAP010962-PA	AI	1.24	0.76	1.30
AGAP011507-RA	COE13O	AI	1.37	1.13	1.84

AGAP013219-RA	elongase	AI	NA	0.81	1.70
AGAP000985-RA	actin-related protein 2	MT	-0.43	0.63	1.51
AGAP001563-RA	glycerol kinase	MT	0.41	0.84	1.69
AGAP002323-RA	isoform b	MT	0.44	0.91	2.91
AGAP002867-RA	CYP6P4	MT	3.65	0.91	3.13
AGAP002869-RA	CYP6P2	MT	1.92	1.00	1.39
AGAP004382-RA	GSTD3	MT	1.17	0.19	0.88
AGAP005749-RA	GSTO1	MT	0.73	0.55	1.85
AGAP006226-RA	**aldehyde oxidase**	**MT**	**1.18**	**0.57**	**3.61**
AGAP006576-RB	malate dehydrogenase	MT	-1.19	0.95	1.00
AGAP007301-RA	isoform a	MT	NA	1.03	1.61
AGAP007793-RA	anterior fat body protein	MT	0.26	0.52	3.10
AGAP008218-RA	**CYP6Z2**	**MT**	**3.26**	**0.80**	**4.90**
AGAP008632-RA	**alpha-aminoadipic semialdehyde mitochondrial**	**MT**	**1.70**	**1.06**	**5.81**
AGAP009946-RA	GSTMS3	MT	0.59	0.26	0.94
AGAP010911-RA	COE09895	MT	-0.43	0.43	1.31
AGAP011852-RA	**short-chain dehydrogenase**	**MT**	**2.16**	**0.81**	**4.48**
AGAP000717-RA	monocarboxylate transporter	Midgut	NA	0.49	1.95
AGAP000985-RA	actin-related protein 2	Midgut	-0.43	0.63	1.51
AGAP002323-RA	isoform b	Midgut	0.44	0.91	2.91
AGAP002865-RA	**CYP6P3**	**Midgut**	**2.87**	**1.31**	**2.75**
AGAP002869-RA	CYP6P2	Midgut	1.92	1.20	1.38
AGAP003493-RB	sugar transporter	Midgut	0.50	1.06	2.03
AGAP003493-RC	sugar transporter	Midgut	0.34	1.00	1.75
AGAP003493-RD	sugar transporter	Midgut	0.31	1.03	1.83
AGAP003499-RA	AGAP003499-PA	Midgut	0.21	0.77	1.10
AGAP004808-RA	protease m1 zinc metalloprotease	Midgut	0.59	0.70	1.05
AGAP005749-RA	GSTO1	Midgut	0.73	0.62	1.85
AGAP006222-RA	UGT302A1	Midgut	1.73	0.90	3.11
AGAP006226-RA	aldehyde oxidase	Midgut	1.18	0.65	3.97
AGAP006710-RA	chymotrypsin 1	Midgut	NA	0.55	7.20
AGAP006792-RA	AGAP006792-PA	Midgut	-0.84	0.99	2.10

AGAP007301-RA	isoform a	Midgut	NA	1.03	1.61
AGAP007552-RA	isoform a	Midgut	-0.46	0.80	5.20
AGAP007650-RA	gadd45 cg11086-pa	Midgut	NA	0.84	2.15
AGAP007879-RB	steroid dehydrogenase	Midgut	NA	0.75	1.71
AGAP008218-RA	**CYP6Z2**	**Midgut**	**3.26**	**0.93**	**4.41**
AGAP008358-RA	CYP4H17	Midgut	1.91	0.81	2.05
AGAP008632-RA	alpha-aminoadipic semialdehyde mitochondrial	Midgut	1.70	1.06	5.82
AGAP008931-RA	sodium-dependent phosphate transporter	Midgut	1.09	0.57	2.49
AGAP010328-RA	triacylglycerol lipase	Midgut	1.86	1.28	5.97
AGAP010911-RA	COE09895	Midgut	-0.43	0.50	1.37
AGAP011434-RA	zinc carboxypeptidase a 1	Midgut	0.35	0.98	2.34
AGAP011984-RA	n-acetyl galactosaminyl transferase 6	Midgut	0.73	0.96	2.46
AGAP013036-RA	ov-16 antigen	Midgut	0.90	0.54	2.41
AGAP013386-RA	dna polymerase zeta catalytic subunit	Midgut	-0.16	0.78	1.67

Table 3.2: Probe list from stringent analysis of direct transcriptome comparisons of dissected body parts from susceptible and resistant strains. AGAP identifier, description, body part that the probe fits the stringent selection criteria, whole organism microarray experiment $\log_2$ fold change; resistant vs susceptible body part microarray $\log_2$ fold change and GaGa $\log_2$ fold change for given body part. qPCR validation has been performed on several candidates (bold). Cells are coloured with a gradient dependent upon the directionality of the fold change, down regulated transcripts are indicated in red and up regulated in green. AI = Abdomen Integument and MT = Malpighian Tubules.

All of the 9 cytochrome P450s (*CYP6P2, P3* and *P4, CYP6Z2, Z3, CYP4H17, CYP4K2, CYP4C35* and *CYP4G16*) on this candidate list were also detected in the direct

comparisons of the whole organism transcriptomes between these strains. Indeed the majority of the 134 transcripts on our candidate list derived from comparisons of the dissected body parts were also detected in comparisons of the whole transcriptome between the two strains. Nevertheless, 22 (16.4%) would have been missed using whole organism arrays and 23 (17.2%) of the genes are regulated in the opposite direction between the whole organism and dissected body part comparisons. Of the 22 transcripts not detected in strain comparisons at the whole organism level, the majority (13) were detected from the abdomen integument (Table 3.3). Abdomen integuments from other resistant strains of *An. gambiae* are being dissected to search for further supporting evidence for a role of these transcripts in conferring resistance prior to follow-up functional analysis.
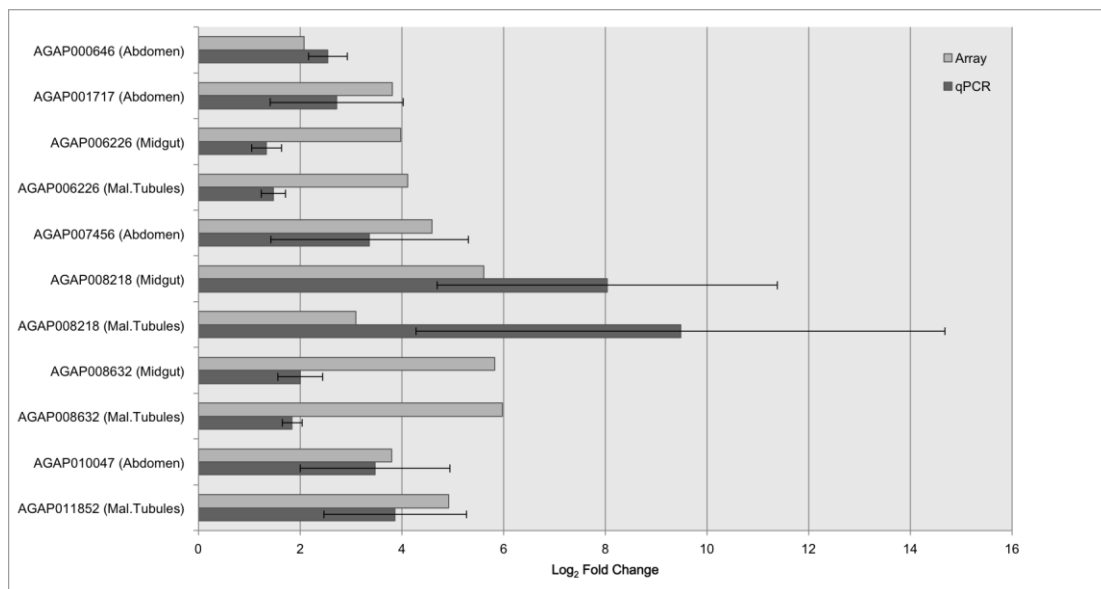
Figure 3.7: Validation of selection of transcripts from the candidate list by qPCR. Log_2 fold change from qPCR data and array data of the direct comparison between the resistant and susceptible populations. Genes were selected from the

'stringent candidate gene list' consisting of probes enriched in the resistant versus susceptible strain in one or more dissected body parts. Standard error bars are shown.

AGAP Identifier	Description	Tissue	Log$_2$ Fold Change
AGAP006710	Chymotrypsin-1	Midgut	7.20
AGAP006741	Putative Tight Junction Protein	Abdomen Integument	2.33
AGAP001769	Beat Protein	Abdomen Integument	2.27
AGAP007650	GADD45	Abdomen Integument	2.15
AGAP000717	Monocarboxylate Transporter	Midgut	1.52
AGAP005563	Sugar Transporter	Abdomen Integument	1.91
AGAP009521	Ankyrin Erythrocytic	Abdomen Integument	1.86
AGAP007879	Steroid Dehydrogenase	Midgut	1.74
AGAP013219	Elongase	Abdomen Integument	1.70
AGAP007301	Conserved hypothetical protein	Malpighian Tubules and Midgut	1.61
AGAP001649	Glucosylceramidase	Abdomen Integument	1.11
AGAP006358	TTC27	Abdomen Integument	1.05
AGAP002517	Unknown	Abdomen Integument	0.95
AGAP007691	Serpin 18	Abdomen Integument	0.92(-RC) 0.87(-RB)
AGAP005651	Cytoplasmic tRNA 2-thiolation	Abdomen Integument	0.80
AGAP004099	Conserved Hypothetical Protein	Abdomen Integument	0.60
AGAP005334	C-Type Lectin	Midgut	-3.62
AGAP003271	Anexin B10B	Midgut	-1.87
AGAP002752	DNAJ Homolog	Remaining Body Parts	-1.10

Table 3.3: Highest/Lowest expressed transcripts between resistant and susceptible populations. AGAP identifier, description, localisation of enrichment/depletion and fold change for the highest/lowest expression changes between the resistant and susceptible tissues that were non-significant in whole

transcriptome comparisons. Splice variants, as predicted by VectorBase, are labelled with variations of -Rx

Discussion

Microarrays are widely used to identify insecticide resistance mechanisms in mosquito populations (Müller *et al.*, 2008b; Edi *et al.*, 2012; Jones *et al.*, 2012b; Tene *et al.*, 2013). However, all of these studies have used whole-organism RNA as a template, which may dilute or distort the final level of transcript expression detectable. By dissecting some of the major body parts involved in xenobiotic detoxification from two strains of *An. gambiae*, differing in their resistance phenotype we have been able to simultaneously identify the local expression profiles of known insecticide resistance candidates and compare transcription between the two strains within individual body parts.

This rich data set will be useful for establishing pathways of detoxification as genes catalysing the three classic classes of drug metabolism (oxidation, conjugation and excretion) (Gibson & Skett, 2001) would be expected to be co-regulated. Pyrethroid mimetic activity–based probes, used to detect pyrethroid metabolising enzymes in the rat liver, identified a potential network of drug metabolising enzymes from multiple families involved in pyrethroid metabolism (Ismail *et al.*, 2013). Applying the same approach to insecticide resistant mosquitoes and using the data on local transcription from the current data set, will help unravel the pathways of insecticide metabolism selected for by intensive use of pyrethroids.

Although potential new insecticide resistance candidates have emerged from this study, it is encouraging that the majority of candidate insecticide resistant transcripts identified from direct comparison of the transcriptomes of dissected body parts were also detected in the whole organism comparisons. No single body part emerged as the key site of over-transcription of putative insecticide resistance genes in this study and it is therefore recommended that, unless resources enable a more comprehensive study design involving multiple dissected tissues, transcriptional approaches to identify candidate insecticide resistance transcripts continue to use the whole body transcriptome. Nevertheless this data set on local sites of transcription should be consulted when designing follow-up qPCR validation steps, or for screening known candidates in field populations.

4. EXPLORATORY META-ANALYSIS OF *ANOPHELES* MICROARRAY EXPERIMENTS

Unpublished

V A Ingham, S C Wagstaff, J D Moore and H Ranson.

Introduction

Resistance to public health insecticides is present across Sub-Saharan Africa

(Diabate *et al.*, 2004; Awolola *et al.*, 2005; Verhaeghen *et al.*, 2006b; Dabire *et al.*,

2009; Awolola *et al.*, 2009; Djègbè *et al.*, 2011; Namountougou *et al.*, 2013;

Kabula *et al.*, 2014; Toé *et al.*, 2015), driven, in part, by malaria intervention

programmes aimed at the vector populations of *Anopheles* mosquitoes (WHO,

2014). The intense selection pressure applied by usage of all four classes of

public health insecticides (pyrethroids, carbamates, organophosphates and

organochlorides) is leading to an ever-increasing number of resistant

populations (Ranson *et al.*, 2011; WHO, 2014) and decreased efficacy of key

intervention strategies (N'Guessan *et al.*, 2007; Kelly-Hope *et al.*, 2008; Strode *et

al.*, 2014). In order to discern the transcript-based changes in resistant

mosquitoes, many microarray studies have been carried out comparing various

populations of resistant *Anopheles* species to insecticide susceptible populations

(Edi *et al.*, 2012; Jones *et al.*, 2012b; Jones *et al.*, 2013; Kwiatkowska *et al.*, 2013;

Abdalla *et al.*, 2014). These data represent an invaluable source of transcriptome

changes and yet no studies have been performed that holistically examine the

data. Furthermore, analysis has often been prejudiced towards describing

candidates from gene families previously associated with insecticide resistance.

The increase in availability of large datasets following similar research

hypotheses has paved the way for large-scale meta-analysis based approaches,

combining a compendium of statistical methods to increase sensitivity and

output valid conclusions. Meta-analyses performed on microarray data fall under

several categories, including differential gene expression, enrichment analysis,

network and co-expression analysis and inter-study prediction analysis, all of which have positives (Tseng *et al.*, 2012). The aim of this study is to produce a comprehensive meta-analysis using a variety of the techniques above on microarray data comparing resistant and susceptible populations of (i) *An. coluzzii*, (ii) *An. gambiae* or (iii) *An. arabiensis.* Meta-data collected that is associated with each array was tested for importance in explaining the variability between the arrays. A variety of clustering techniques will be applied both at the differential transcript level and the enrichment level to determine whether an underlying mechanism is responsible for the groupings of arrays. Further, hypothesis driven approaches will be explored: (a) in an attempt to determine new candidates for a role in insecticide resistance and (b) to determine the transcriptional regulation of resistance.

Aims

1. Analyse all suitable microarray data available by fitting suitable linear models.

2. Use exploratory analysis methods to determine the relationships between the analysed array studies.

3. Use hypothesis-driven analysis to elucidate candidate transcripts involved in insecticide resistance.

4. Elucidate the transcriptional regulation of insecticide resistance through hypothesis led analysis.

Results

Microarray meta-analysis: Revealing relationships.
27 microarray studies, each representing a resistant versus susceptible population of *An gambiae* (1), *An coluzzii* (16) or *An arabiensis* (10) (Table 4.1),

were analysed using the limma package for R (Smyth, 2004) and a significance cut-off applied (adjusted p ≤ 0.05). These data encompassed a large geographical area, including Eastern and Western Africa, (Figure 4.1) and were collected between 2009 and 2012. All arrays demonstrated adequate normalisation with comparable ranges of fold change, as assessed through visual means following previously published methodology (Quackenbush, 2001), in addition to principal component analysis. Although meta-data was collected for each array, these data explained less than 4% of the variance across the datasets and were therefore discarded from further analyses (Table 4.2).

Population	Population ID	Exposure Status	Country	Species	Susceptible Population	Reference
Bioko	***coluzzii.1***	***Deltamethrin***	***Equatorial Guinea***	***An. coluzzii***	***N'Gousso***	(Hemingway *et al.,* 2013)
Bioko	*coluzzii.2*	*Unexposed*	*Equatorial Guinea*	*An. coluzzii*	*N'Gousso*	(Hemingway *et al.,* 2013)
DSSBA	arabiensis.1	DDT	Burkina Faso	*An. arabiensis*	Moz	(Jones *et al.,* 2012b)
DSSBA	arabiensis.2	Unexposed	Burkina Faso	*An. arabiensis*	Moz	(Jones *et al.,* 2012b)
VK7 2011	*coluzzii.3*	*Unexposed*	*Burkina Faso*	*An. coluzzii*	*Mali*	(Toé *et al.,* 2015)
VK7 2011	*coluzzii.4*	*Deltamethrin*	*Burkina Faso*	*An. coluzzii*	*Mali*	(Toé *et al.,* 2015)
Yaoundé	coluzzii.5¶	Unexposed	Cameroon	*An. coluzzii*	N'Gousso	(Tene *et al.,* 2013)
Kitwe	gambiae.1¶	Deltamethrin	Zambia	*An. gambiae*	Kisumu	(Thomsen *et al.,* 2014)

Bouake	*coluzzii.6¶*	*Deltamethrin*	*Côte D'Ivoire*	*An. coluzzii*	*N'Gousso*	(Toé *et al.*, 2015)
M'Be	*coluzzii.7¶*	*Deltamethrin*	*Côte D'Ivoire*	*An. coluzzii*	*N'Gousso*	(Toé *et al.*, 2015)
Tiassalé	*coluzzii.8¶*	*Deltamethrin*	*Côte D'Ivoire*	*An. coluzzii*	*N'Gousso*	(Toé *et al.*, 2015)
VK6	*coluzzii.9¶*	*Unexposed*	*Burkina Faso*	*An. coluzzii*	*N'Gousso*	(Kwiatkowska *et al.*, 2013)
VK7	*coluzzii.10¶*	*Unexposed*	*Burkina Faso*	*An. coluzzii*	*N'Gousso*	*Toé, Unpublished data*
Youandé	coluzzii.11*¶	Deltamethrin	Cameroon	*An. coluzzii*	N'Gousso	(Tene *et al.*, 2013)
Youandé	**coluzzii.12¶**	**Deltamethrin**	**Cameroon**	***An. coluzzii***	**N'Gousso**	(Tene *et al.*, 2013)
Jinja	arabiensis.3	Unexposed	Uganda	*An. arabiensis*	Dongola	(Mawejje *et al.*, 2013)
Jinja	**arabiensis.4**	**Permethrin**	**Uganda**	***An. arabiensis***	**Dongola**	(Mawejje *et al.*, 2013)
Jinja	arabiensis.5	Unexposed	Uganda	*An. arabiensis*	Moz	(Mawejje *et al.*, 2013)

Jinja	arabiensis.6	Permethrin	Uganda	*An. arabiensis*	Moz	(Mawejje *et al.*, 2013)
Pemba	arabiensis.7	Unexposed	Tanzania	*An. arabiensis*	Dar	(Jones *et al.*, 2013)
Pemba	**arabiensis.8**	**Lambda-Cyhalothrin**	**Tanzania**	**An. arabiensis**	**Dar**	(Jones *et al.*, 2013)
VK7 2012	*coluzzii.13*	*Deltamethrin*	*Burkina Faso*	*An. coluzzii*	*N'Gousso*	(Toé *et al.*, 2015)
Tengrela	*coluzzii.14*	*Unexposed*	*Burkina Faso*	*An. coluzzii*	*N'Gousso*	(Toé *et al.*, 2015)
VK7 2012	*coluzzii.15*	*Deltamethrin*	*Burkina Faso*	*An. coluzzii*	*Mali*	(Toé *et al.*, 2015)
Tengrela	*coluzzii.16*	*Unexposed*	*Burkina Faso*	*An. coluzzii*	*Mali*	(Toé *et al.*, 2015)
Wad Medani	**arabiensis.9¶**	**Permethrin**	**Sudan**	**An. arabiensis**	**Dongola**	(Abdalla *et al.*, 2014)
Ndja	**arabiensis.10¶**	**Permethrin**	**Chad**	**An. arabiensis**	**Moz**	(Witzig, 2012)

*Mosquitoes were collected from polluted environments

¶ = R vs S design, all others being complete loop designs

Table 4.1: Study information. Population name represents the area from which the strain was sampled in addition to insecticide exposure, where relevant. Population ID is the array ID used throughout, unique for a specific array set. Country of origin, species, susceptible reference strain and paper reference are also given. Bold experiments indicate subsection used in focused, smaller meta-analysis, with only *An. coluzzii* compared to N'Gousso and all *An. arabiensis* exposed to pyrethroid. Italics represent experiments used in comparing pyrethroid survivors to the associated pyrethroid unexposed experiments. All insecticides are pyrethroids except for DDT, which is an organochloride.

Figure 4.1: Study distribution map. Map showing the locations of all collection

sites used for the microarrays in this study. Red points show *An. coluzzii,* the

green point shows *An. gambiae,* whilst blue points show *An. arabiensis.*

Data Used	Number of Transcripts (Transcripts/Probes on Array)	ω^2
A *priori* Transcripts	236/807	2.52%
All Transcripts	14106/15164	0.28%
Transcripts Significant in at Least 1 Study	12099/13222	0.2%
Transcripts Significant in Less Than 3 Studies	742/753	0.4%
Transcripts Significant in Less Than 5 Studies	3249/3319	0.24%
Transcripts Significant in Greater Than 9 Studies	2715/3325	0.53%
Transcripts Significant in Greater Than 14 Studies	287/852	2.19%
Transcripts Significant in Greater Than 19 Studies	166/679	3.43%
Transcripts Significant in Greater Than 24 Studies	138/567	3.95%

Table 4.2: Summary of metadata variance. Data description, number of transcripts / number of probes and the ω^2 Statistic showing total variance explained by the factors included in the meta-analysis.

Hierarchical clustering and factor analysis were performed using transcripts significantly differentially expressed, irrespective of directionality, in: one or more study (n = 13,221), 10 or more studies (n = 2,714) and 20 or more studies (n = 165); this exploratory analysis allowed robustness of clustering groupings to be assessed. Although distinct clusters were observed (Figure 4.2), the groups of clusters were not consistent across the three transcript subsets used. Three small consistent clusters were seen across all three data subsets (Figure 4.2), these were: (i) arabiensis.7/8 and coluzzii.3/4/15/16, (ii) coluzzii.1/2/10/11/13/14 and (iii) coluzzii.6/7/8. Cluster (i) represents two *An. arabiensis* studiess, along with three VK7 populations and a Tengrela population compared to the Mali susceptible population. Cluster (ii) is composed of several *An. coluzzii* populations, each of which is compared to the susceptible population N'Gousso. Although numerous *An. coluzzii* arrays with the susceptible N'Gousso are absent from this second cluster, it is apparent that the susceptible strain's transcriptomic background plays a large role in determining the differential transcript list; this is further exemplified when it is considered that the two VK7 (2012) and Tengrela (2012) arrays are split between the aforementioned clusters. The final cluster, (iii), is represented by three Côte D'Ivoire populations, Tiassalé, M'Be and Bouaké, collected at the same time and compared to the same susceptible population, N'Gousso.

The overall lack of continuity or obvious hypothesis behind these groupings points to a large disparity in transcript changes amongst the available studies. It is likely that the numerous different factors such as susceptible reference, insecticide exposure status and resistance profiles are confounding the analysis. It is further possible that different resistance mechanisms or transcriptomic backgrounds are present across the disparate datasets which is not unexpected given that the resistant populations were collected from across the continent over a period of three years.

Figure 4.2: Hierarchical clustering plots of all microarray studies.

Hierarchical clustering plots with each node representing a respective study.

Data subsets used to generate the matrices were probes significant in: a. one or

more study b. 10 or more studies c. 20 or more studies.

GO Term clustering shows low discriminatory groupings.

The number of significant transcripts varied largely across the data sets, which may influence the clustering (Figure 4.3). As an alternative to clustering based on fold-change, GO term enrichments were used as a proxy for clustering to reduce noise in the datasets. 5941 GO terms were present across all significant transcripts (p ≤ 0.001). GO term enrichment was determined using a hypergeometric test for each study and comparing the term to those found in a whole genome search. 47 unique GO terms represented the top 10 terms for each study, demonstrating a largely overlapping landscape of ontologies. Many of the GO terms represented are related to G-protein coupled receptor signalling, changes in transcription, ion binding, chemosensory and neuronal mechanisms. Interestingly, there were few terms related to xenobiotic metabolism, such as oxidation/reduction processes; however, chemosensory mechanisms may represent an altered response to insecticide detection.
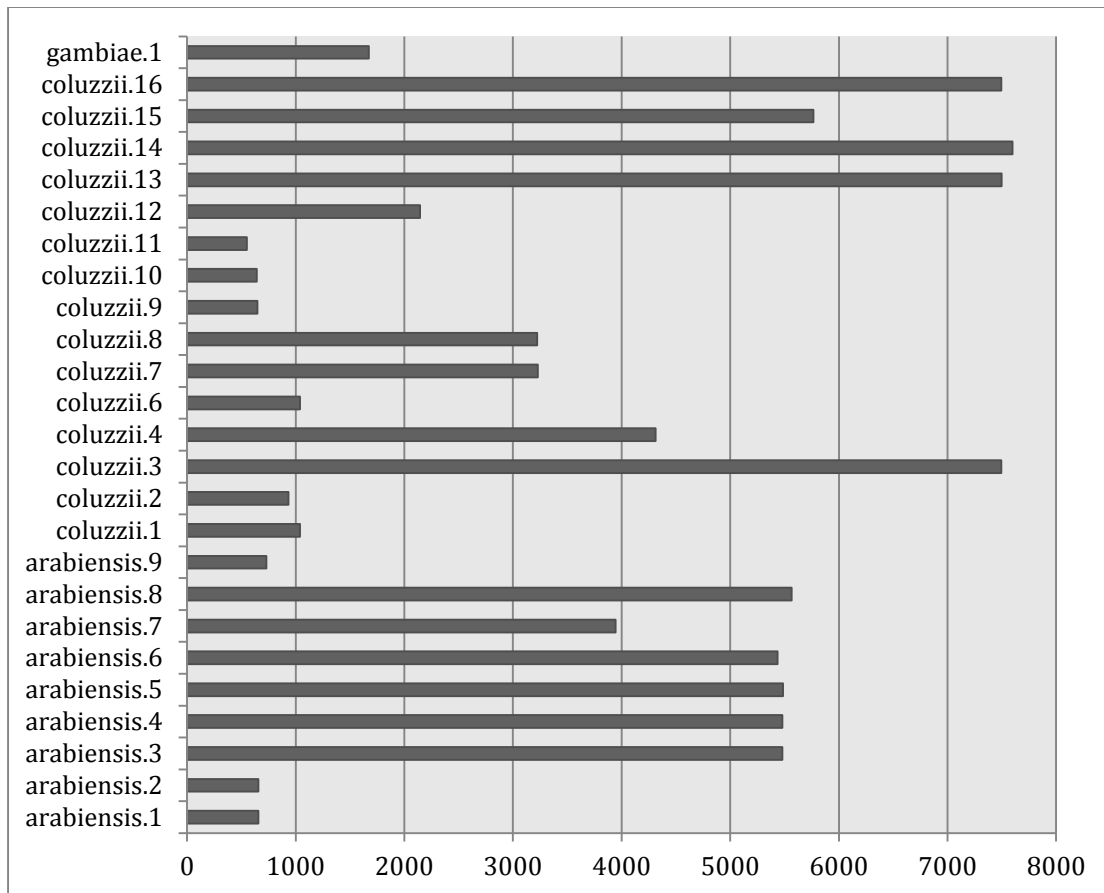
Figure 4.3: Total number of significant (p ≤ 0.001) transcripts for each study. coluzzii.5 and arabiensis.10 had no significant transcripts after a cut-off of adjusted p ≤ 0.001 was applied. The range of remaining studies was 551-7597, with a mean of 3459.

As with earlier clustering, hierarchical clustering was performed on a correlation matrix of the summated GO terms as determined by Blast2Go (Conesa *et al.*, 2005) after applying GO Slim (Consortium, 2004) (Figure 4.4). Although these data show shorter branch lengths and hence, lower distances between the arrays than previous clustering, reaffirming commonality in GO term identity, there is again no clear mechanism apparent. Exposure status and species identity does not appear to play a role in the structuring of the clusters, nor does the identity

of the resistant population, as VK7 is found in three disparate clusters (Figure 4.4).
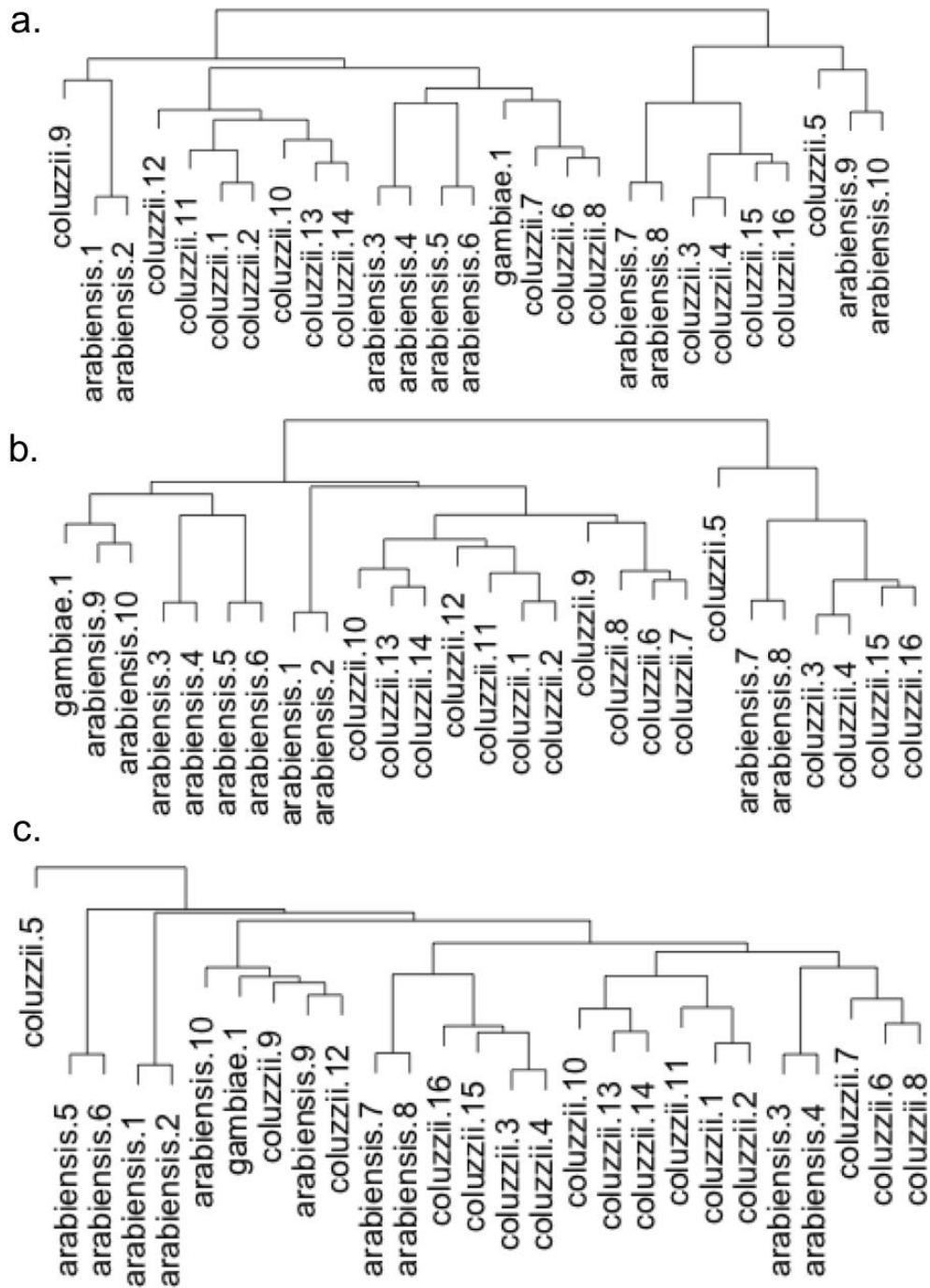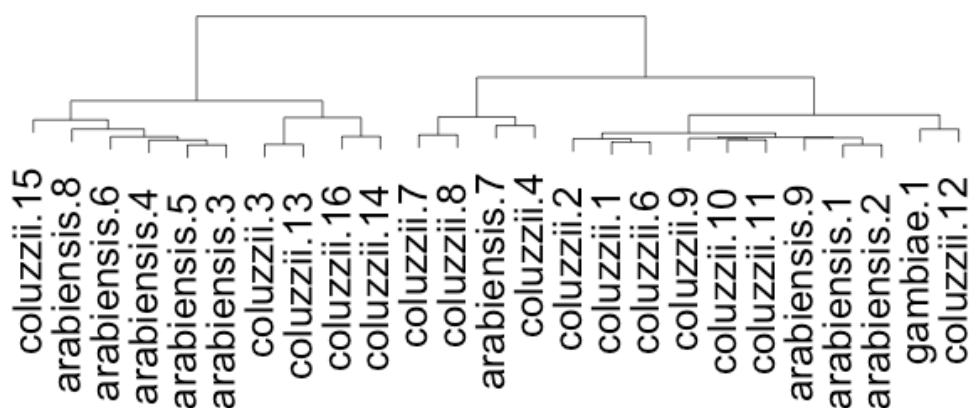

Figure 4.4: Hierarchical clustering plot of GO term number in each study. Hierarchical clustering plot with each node representing a respective study; branch length shows inferred distances, computed from a Euclidean distance matrix.

An. coluzzii and An. arabiensis show unrelated transcriptomes.

Due to the limitations associated with meta-analysis of the entire data set, the studies used were reduced to only pyrethroid exposed populations of *An. arabiensis* and *An. coluzzii*. In addition, only *An. coluzzii* populations compared to the susceptible population N'Gousso were selected, leaving 10 pyrethroid survivor studies (PSA) (bold, Table 4.1). Again, three data subsets were used in clustering, these were transcripts significant in: one or more study (n = 12,179), five or more studies (n = 5,827) and eight or more studies (n = 383). The PSA clustering was much more robust (Figure 4.5) for these smaller subsets than the clustering for all studies (Figure 4.2). The former two data subsets (n = 12,179 and n= 5, 827) clustered with optimally four clusters and the latter clustering

with three, as determined by nFactor (Raiche & Magis, n.d.). In each data subset, two of the clusters represented *An. coluzzii* groupings separating into those from Côte D'Ivoire and those that are not and *An. arabiensis* cluster(s).
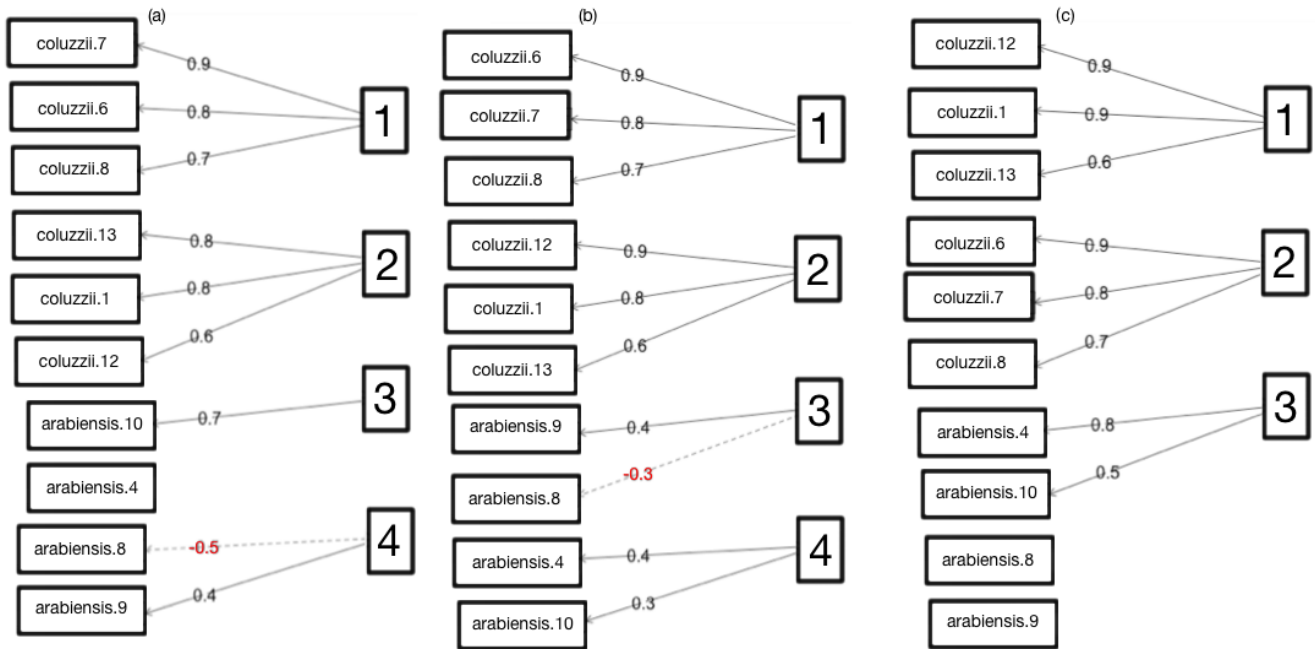
Figure 4.5: Factor analysis using a smaller subset of studies. Factor analysis performed on probes significant in (a) One or more arrays (b) Five or more studies (c) Eight or more studies. Optimum number of clusters was calculated for each probe list using nFactor (Raiche & Magis, n.d.), factor analysis was performed using psych (Revelle, 20151.5.6). Numbers above the lines represent the degree of relatedness between the studies. One through four represent the individual clusters.

The lack of clear grouping within the *An. arabiensis* studies was explored using a heatmap visualisation of the study correlations (Figure 4.6); this data demonstrated a much higher variability within the *An. arabiensis* studies when compared with the *An. coluzzii* arrays. The greater variability in *An. arabiensis*

may be due to the large geographical separations of the populations compared to the *An. coluzzii* experiments (~2080 miles in a straight line compared to ~1050 miles). The variability within the *An. arabiensis* group was further exemplified when probes significantly differential in (i) only *An. coluzzii* and (ii) only *An. arabiensis* studies were split into two lists containing probes showing the same directionality of fold change within groups (i) and (ii). The number of probes up-regulated in the resistant compared to the susceptible strain were 455 in *An. coluzzii* and just 3 in *An. arabiensis*. Similarly, probes down-regulated were 327 and 2 in *An. coluzzii* and *An. arabiensis* respectively.
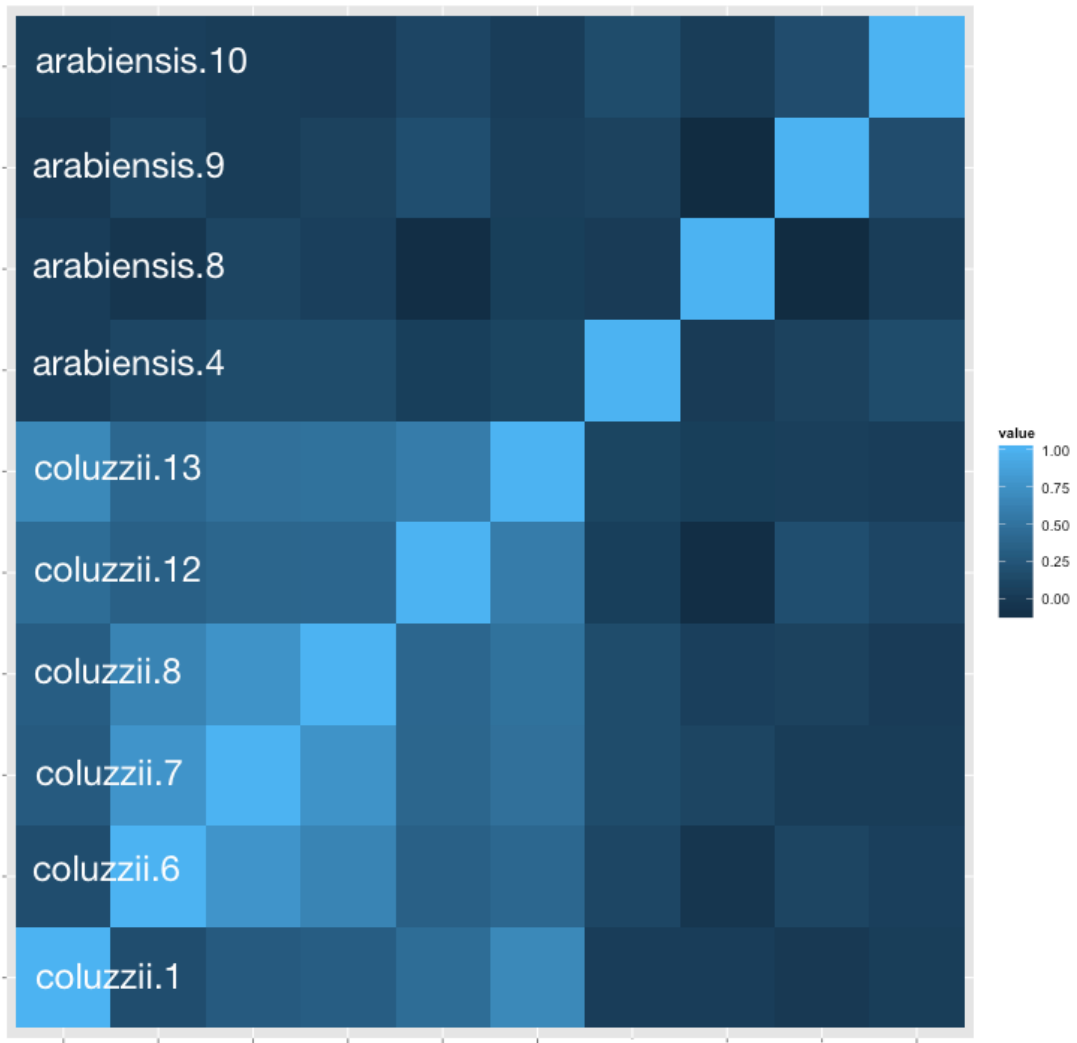
Figure 4.6: Correlation plot of transcripts significant in one or more studies.

Studies are represented once in each row/column, with the corresponding correlation shown as a block of colour. Correlation is represented on a gradient from dark blue (no correlation) to turquoise (complete correlation). Data for both up- and down-regulation is included in the figure.

Overall, there are 13 transcripts (Table 4.3) that are differentially expressed across all *An. coluzzii* and *An. arabiensis* arrays; however, only one of these showed the same fold change directionality, a ribosome production factor

(AGAP003048-RA) was overexpressed in all studies. AGAP003048-RA is a

homolog of *Drosophila non3*, which has been linked with several developmental

processes (Guertin *et al.*, 2006; Blanco *et al.*, 2010). The remaining transcripts

contain 4 cytochrome p450s, 3 transcripts of unknown function and two serine

proteases. AGAP003444-RA only has homologs in other *Anopheles* species,

suggesting a lineage specific gene. Similarly, AGAP011785-RA is only found in

hematophagous insects. *ANXB9* has a close (83% identity) homolog in *Drosophila*

that demonstrates a response to stress, including cold stress and changes in

oxygen conditions (Telonis-Scott *et al.*, 2009; Yin *et al.*, 2013), which may be

indicative of a response to insecticide-imposed stresses.

Identifier	Description
AGAP000088-RA	CYP4H19
AGAP001240-RA	SP11372
AGAP003048-RA	Ribosome production factor 2
AGAP003444-RA	Unknown
AGAP003790-RB	ANXB9
AGAP007220-RA	DNA-directed RNA polymerase III subunit RPC9
AGAP008018-RA	CYP12F4
AGAP009053-RA	Lipopolysaccharide-induced TNF-alpha factor
AGAP009374-RA	CYP9M1
AGAP009753-RA	Unknown
AGAP011785-RA	CLIPE6
AGAP013036-RA	Unknown

AGAP013490-RA	CYP4H24

Table 4.3: Transcripts significant across all ten studies. Transcripts that are significantly differentially expressed across all *An. coluzzii* and *An. arabiensis* studies, showing ID and descriptions of each of the transcripts.

Up- and down-regulated transcripts in *An. coluzzii* show different enrichments. Those transcripts up-regulated are enriched in various detoxification and histone protein domains, including ABC transporters, carboxylesterases, UDP-glucosyltransferase, histone core and histone H2B. Down-regulated transcripts show domain enrichment for peptidase activity and a large number of transcripts (109) with unknown function. Seven ABC transporter, seven cytochrome p450s and two UDPs/GSTs transcripts are present in the up-regulated gene list.

Candidate Selection: Transcription factors.
The PSA subset of studies were explored using GaGa analysis, a bespoke pipeline written in R to group transcripts with similar patterns of expression across given arrays. GaGa was used in an attempt to find transcripts that followed the same pattern of expression across all 10 studies. GaGa analysis results in very large transcript lists due to the requirement that each transcript be present in one user defined group; this is shown with 6215 transcripts demonstrating the same pattern across ten studies. In this instance, the large size of the transcript list aided in searching for transcription factors. In total, 264 putative transcription factors were identified as showing the same expression pattern across all 10 studies. To narrow down the number of transcription factors, transcripts that were significantly ($p \leq 0.05$) differentially expressed in at least eight of the 10

studies were extracted, resulting in eight transcription factor candidates (Figure 4.7). Eight out of 10 studies was selected as a cut-off to produce more than one transcriptions factor, as only *Met* is significant across all 10 arrays. Consistency in significance was expected for strong insecticide resistance candidates, putatively playing a role in resistance across species, geographical locality and time.
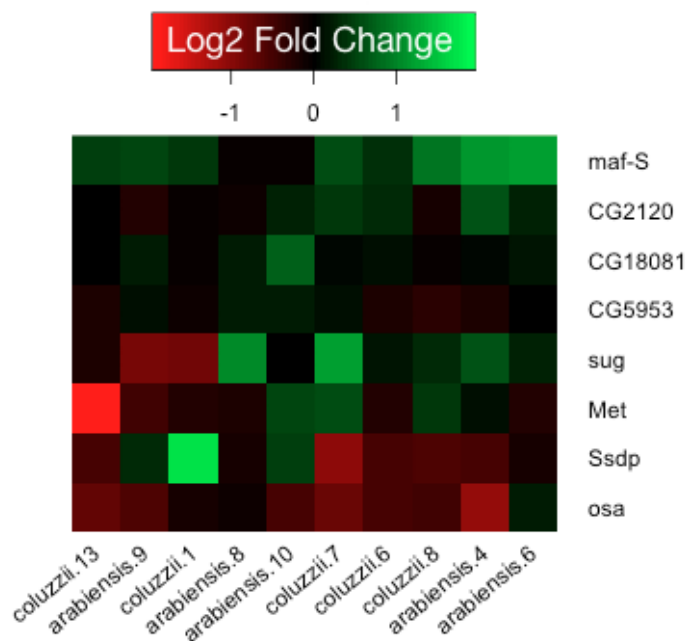
Figure 4.7: Transcription Factor Candidates. Heatmap showing the $\log_2$ fold change of eight candidate transcription factors that are differential ($p \leq 0.05$) across at least eight of the 10 pyrethroid exposed studies. Naming conventions are taken from *Drosophila* gene names on FlyTF (Adryan & Teichmann, 2006). (Code available: https://github.com/VAI3064/PhD-Code/blob/master/heatmap_function.R)

Of these transcripts both *Met* and *Maf-S* have been implicated in resistance to at least one insecticide (Li *et al.*, 2007; Misra *et al.*, 2011; Misra *et al.*, 2013). Of the other six transcripts, research in *Drosophila* has shown that: *osa* is involved in

chromatin re-modeling (Monribot-Villanueva *et al.*, 2013), stem cell commitment (Zeng *et al.*, 2013) and cell-cycle regulation (Terriente-Félix & de Celis, 2009); *sug* is involved in certain stress responses (Farhadian *et al.*, 2012) and *Ssdp* is involved with development (Chen *et al.*, 2002).

Candidate Selection: Transcriptomic response to a core of pyrethroid resistant populations.

Recognising that, even in areas where the majority of mosquitoes are resistant to pyrethroids, field collections of mosquitoes will contain a heterogeneous mix of individuals differing in their level of resistance, some studies pre-exposed field populations to pyrethroids prior to RNA extraction. Mosquitoes were exposed to pyrethroid impregnated papers in WHO susceptibility assays and RNA extracted from survivors 48 hours after the exposure (to reduce the chance of detecting genes induced by insecticide exposure rather than constitutively overexpressed in the most resistant subset of the population, as previously published (Jones *et al.*, 2012b)).

In order to generate a candidate list from the meta-analysis, a hypothesis was formed postulating that survivors of pyrethroid exposure would form a truly resistant subset of transcriptomes and would differ from the unexposed population's transcriptomes. Surviving pyrethroid exposure was therefore used as a proxy for a completely resistant sub-population of the resistant strains, as complete recovery is expected after 48 hours (Jones *et al.*, 2012b). Three datasets where both the surviving and unexposed subsets were compared to the same susceptible strains (N'Gousso and Mali) were available (Figure 4.8a). Two

of these were from VK7 in Burkina Faso, collected in subsequent years, and one was from Bioko Island. In order to compare survivor and unexposed arrays, six array experiments involving *An. coluzzii* unexposed populations were utilised. These array sets represented four populations (Bioko, VK6, VK7 and Tengrela), which were compared to the same susceptible populations as the exposed studies (N'Gousso and Mali). In this case, both VK7 and Tengrela were represented in duplicate, collected in the years 2011 and 2012.

To identify suitable candidates, a minimum selection criterion that the transcripts must be significant in all of the survivor studies was set (n = 969 transcripts). In addition, the list was searched for transcripts showing consistent up-regulation and down-regulation (Figure 4.8).
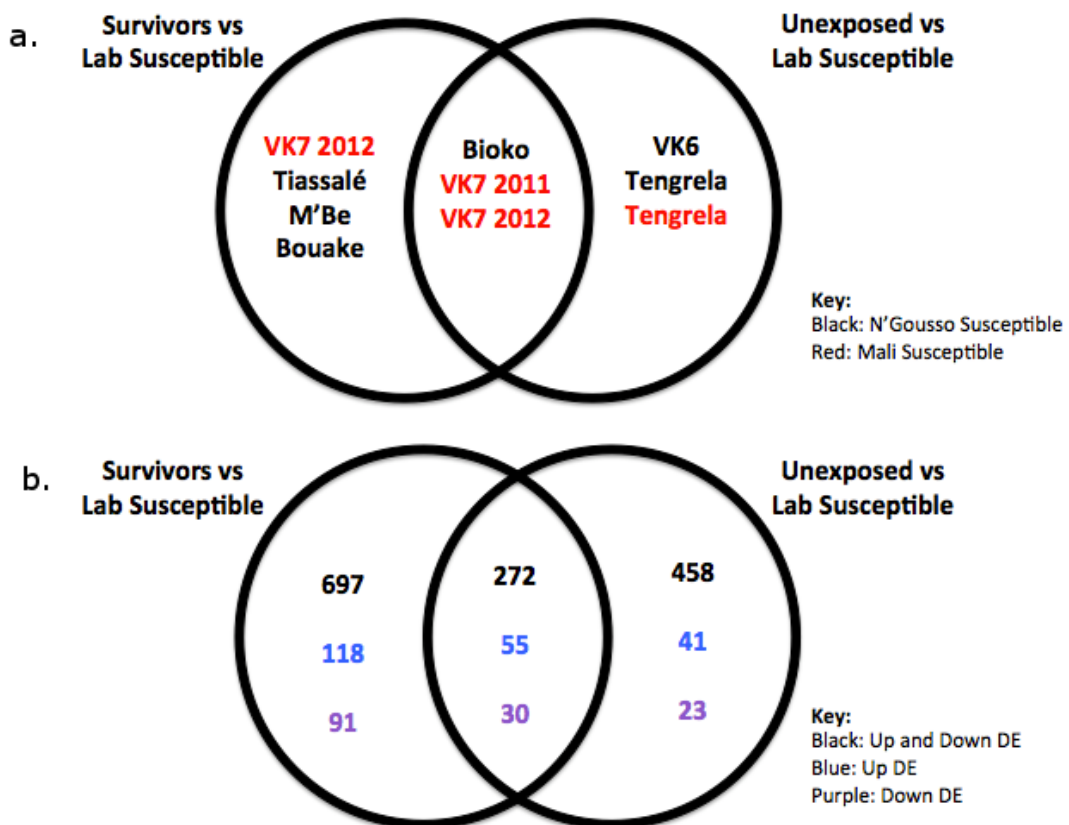
Figure 4.8: Summary of limma analysis for pyrethroid studies. Venn diagrams showing (a) studies used for survivors of pyrethroid exposure and the companion unexposed studies, with either N'Gousso (black) or Mali (red) susceptible populations. (b) Number of differentially expressed (DE) transcripts for each group of studies, showing numbers of all DE transcripts (black), up DE transcripts in all studies (blue) and down DE transcripts in all studies (purple).

The 969 significant transcripts significant in all survivor studies were then compared to the significant transcripts from the 6 unexposed *An. coluzzii* studies (italics, Table 4.1) using a Mann-Whitney U test with a significance cut-off $p \leq$ 0.05 (n = 11 transcripts) (Table 4.4, Figure 4.9).
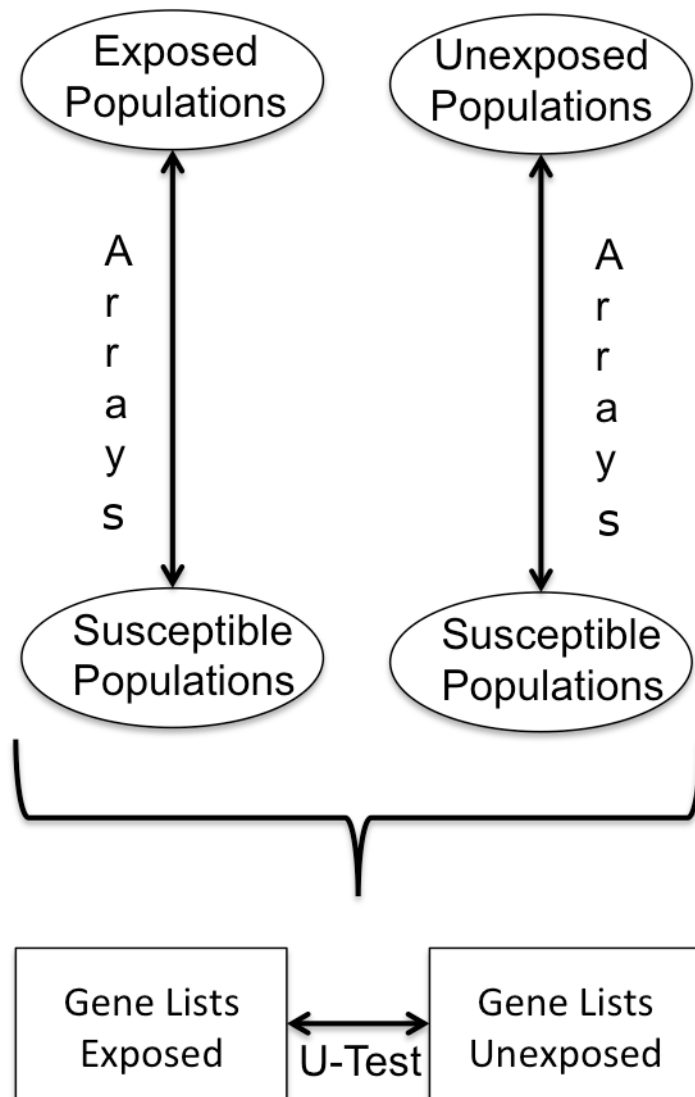
Figure 4.9: Schematic of experimental design. Array experiments composed of deltamethrin exposed and unexposed *An. coluzzii* populations compared to a susceptible population were analysed using limma. The resultant gene lists were taken and Mann-Whitney U tests performed. Significance cut-off of p$\leq$ 0.05 used in both instances.

AGAP Identifier	Description

AGAP008052-RA	SAP2, Chemosensory Protein.
AGAP007918-RA	Xanthine Dehydrogenase
AGAP006364-RC	ABC-Transporter (B family)
AGAP005114-RB	Unknown
AGAP005252-RA	MYD, Toll Signalling
AGAP012742-RA	Unknown
AGAP000536-RA	PGRPS1
AGAP003917-RA	Alpha-Acetylytransferase
AGAP005573-RA	Unknown
AGAP009374-RA	Unknown
AGAP013374-RA	Heme Peroxidase

Table 4.4: Transcripts showing significant differential expression, both up and down, between deltamethrin exposed and unexposed *An. coluzzii* populations. For each transcript, AGAP identifier and associated description are shown.

Gene Enrichment: Survivor Studies.

969 transcripts were differentially expressed across all survivor studies; these transcripts may demonstrate a core mechanism necessary for survival to pyrethroid exposure or a generalised response to stress; both are likely important for insecticide resistance. The resistant populations were collected from geographically disparate localities (Figure 4.1, Figure 4.8a, Table 4.1) including Côte D'Ivoire in 2009 (Bouake, M'Be and Tiassalé), Equitorial Guinea in 2009 (Bioko) and Vallé de Kou in Burkina Faso in Year 2011 and 2012 (VK7). All resistant populations were compared to the lab susceptible N'Gousso, originally

from Cameroon, and two VK7 populations collected in 2011 and 2012 were also compared to the susceptible strain from Mali. When split into up-regulated and down-regulated transcripts, enrichments were explored to identify similarities and differences between the up- and down-regulated subsets. The up-regulated transcript list was significantly enriched for detoxification family genes, including cytochrome P450s, GSTs, UDPs, ABC transporters and carboxylesterases ($p \leq 0.001$), with 12 unique transcripts (Table 4.5). Included in these transcripts are two previously identified insecticide resistance candidate genes, *CYP6P3* and *CYP6Z3* (Djouaka *et al.*, 2008; Müller *et al.*, 2008b; Chandor-Proust *et al.*, 2013; Djègbè *et al.*, 2014). Over a quarter (51) of the up-regulated transcripts were of unknown function. Similarly, almost a quarter (25) of the down-regulated transcripts were of unknown function; however, there is no enrichment (p=0.18) of detoxification transcripts in this list, instead there is enrichment for processes such as translation, tRNA, and ligase activity.

Identifier	Description
AGAP006364-RA	ABCB4
AGAP006364-RB	ABCB4
AGAP006364-RC	ABCB4
AGAP006726-RA	COEAE5G
AGAP000818-RA	CYP9K1
AGAP012293-RA	CYP9L3
AGAP002419-RA	CYP4D22
AGAP002865-RA	CYP6P3
AGAP007480-RA	CYP6AH1

AGAP008217-RA	CYP6Z3
AGAP006222-RA	UGT302A1
AGAP007589-RA	UGT306A2
AGAP004164-RD	GSTD1
AGAP009946-RA	GSTMS3

Table 4.5: Identities of detoxification family transcripts up-regulated in survivor studies. AGAP identifiers and gene names of detoxification family members identified in the list of transcripts up-regulated across all survivor arrays. Splice variants are identified with -RX.

Discussion

Twenty-seven microarray studies comparing resistant *Anopheles* populations to susceptible populations were analysed and subsequently used in a meta-analysis approach to determine whether there was a clear relationship between experiments. The approach yielded no clear hypothesis, despite the utilisation of both a differential expression analysis and an enrichment analysis approach. The differential expression analysis took into account each individual fold change of transcripts, maximising the information available; however, due to large variability in the datasets, an enrichment method was also explored to reduce the noise by grouping transcripts into functional groups. It is clear, however, that approaches taken here do not identify factors that may determine the overall clustering of array sets. Given this information, it may be that there is not one universal mechanism for resistance to insecticides or that the resistance mechanism itself is not a key determinant in the transcriptome clustering due to masking by other transcriptomic differences. Nevertheless, there are clear commonalities across the studies as shown by the limited gene ontology

landscape, with 47 unique terms representing the top 10 GO terms for each study. These GO terms were largely devoid of obvious xenobiotic response related terms, despite all study populations demonstrating resistance to at least one class of insecticide and in many cases, multiple classes.

An important confounding factor in clustering the studies appears to be the large disparity seen between different *An. arabiensis* studies, when compared to the *An. coluzzii* studies. It is likely that this is due to the much larger geographical range of the population origins, in addition to a wide variety of susceptible backgrounds. The pyrethroid exposed *An. arabiensis* studies have only five transcripts showing the same directionality across four array sets; this is in stark contrast to the *An. coluzzii* studies which show 782 transcripts commonly differentially expressed across six array sets. Only one transcript AGAP003048-RA is significant across all experiments, including both *An. arabiensis* and *An. coluzzii*, demonstrating the same directionality of fold change. AGAP003048-RA is ribosome production factor 2, the closest *Drosophila* ortholog is *novel nucleolar protein 3*, which has been shown to regulate the stability and translation of mRNAs (Lu *et al.*, 2014). Due to the function of this transcript, it is likely that changes to translation may be a general response to chemical stress, with the insect responding by changing levels of mRNA expression as needed.

Using hypothesis driven clustering, it has been possible to extract manageable transcript lists, which will be taken forward to qPCR and phenotyping. These lists have allowed identification of putatively important candidates, showing differential expression over numerous geographical locations, species and

temporal separation. Comparing survivors of pyrethroid exposure to unexposed populations, it is clear that some transcripts show greater differential expression in a core resistant subset of mosquitoes. Furthermore, 12 detoxification gene members have been shown to be up-regulated across all survivor populations and may represent a core detoxification resistance mechanism. Similarly, searching for transcription factors gave rise to just seven transcripts, which may represent transcription factors implicated in the transcriptional control of processes involved in insecticide resistance. Two of these have previously been implicated in insecticide resistance, *Met* and *Maf-S* (Misra *et al.*, 2011; Charles *et al.*, 2011; Misra *et al.*, 2013; Zou *et al.*, 2013), further supporting the validity of this methodology.

In conclusion, although there are clear similarities in the ontologies displayed by the significant transcripts, there is no obvious hypothesis that explains the variability amongst the array sets. It is likely that there is some underlying mechanism that is not intuitively obvious leading to the clusters, which may be due to large variations in resistance mechanisms employed by various mosquito populations, sampling times or other experimental factors. However, by deriving a hypothesis and utilising select array sets, it is possible to produce manageable transcript lists, which may aide in further study of the control and mechanisms of insecticide resistance in *Anopheles* species.

5. FUNCTIONAL CHARACTERISATION OF CANDIDATE TRANSCRIPTS.

Unpublished

V A Ingham, J Bagit, S C Wagstaff, J D Moore and H Ranson.

V.A.I., performed all *in silico* analysis. V.A.I., J.D.M., H.R. and S.C.W. designed the study. V.A.I. performed all experiments. J.B. assisted in bioassays and mosquito rearing.

Introduction

In Chapter 4 a number of novel candidate insecticide resistance transcripts were described. In this chapter a subset of these were validated by qPCR and then functionally assessed by RNAi. This led to the identification of a chemosensory protein that appears critical for mediating resistance to pyrethroids in the Tiassalé strain.

Aims

1. Confirm differential expression of deltamethrin resistance candidates in laboratory colonies

2. Use both RNAi and bioassays to test the association between gene expression and the resistance phenotype.

Results

Candidate Identification

A hypothesis driven meta-analysis approach was used to identify transcripts showing differential expression in mosquitoes surviving deltamethrin exposure and those significantly different in matched unexposed populations [Chapter 4]. The approach gave rise to a number of transcripts showing consistent up-regulation in survivor populations and several transcripts demonstrating significantly differential expression between survivors and unexposed populations. In order to select candidates with the highest chance of confirmation, a number of transcripts from each list, two candidates with the lowest p-value were selected from the 11-transcript list and five with the largest average fold change across the arrays were selected from the 173 transcript list, in addition to a UGT, which had the highest fold change for detoxification transcripts outside the top five. All transcripts demonstrated consistent up-

regulation. (Table 5.1).

Transcript ID	Description	Fold Change	List
AGAP008052-RA	SAP2	12.15	Survivor vs Unexposed
AGAP006364-RC	ABC-B	2.01	Survivor vs Unexposed
AGAP006222-RA	UGT302A1	4.08	Detoxification/Up in all
AGAP004262-RA	Protein Take-Out 3	9.18	Up in all
AGAP008769-RA	Acyl-CoA dehydrogenase	2.52	Up in all
AGAP007160-RA	α-crystallin chain B	8.75	Up in all
AGAP002603-RA	EF1-like protein	7.46	Up in all
AGAP008217-RA	CYP6Z3	7.40	Detoxification/Up in all

Table 5.1: Candidate transcripts from deltamethrin survivors. Eight

transcripts selected for showing (a) consistent up-regulation across all survivor

studies (up in all) or (b) significant differential expression between deltamethrin

survivors and matched unexposed arrays (Survivor vs Unexposed).

Detoxification family transcripts are also noted (detoxification). Transcript ID,

with splice variant (-RX), description and average fold change across

deltamethrin survivors are shown.

qPCR validation of candidates
The differential expression of the eight selected candidate transcripts was

validated via qPCR (Figure 5.1). Two resistant populations, Tiassalé and VK7

were compared to the lab susceptible strain, N'Gousso. Both strains are resistant

to pyrethroids and DDT and Tiassalé is also resistant to organophosphates and

carbamates. All but one transcript (AGAP007160) demonstrated the expected

fold change directionality in both resistant strains when compared to the lab

susceptible. Of the eight transcripts tested, four (AGAP008052, AGAP008217,

AGAP006222 and AGAP002603) demonstrated fold changes greater than five

and three (AGAP008052, AGAP008217 and AGAP006222) greater than 20 fold

up regulation.
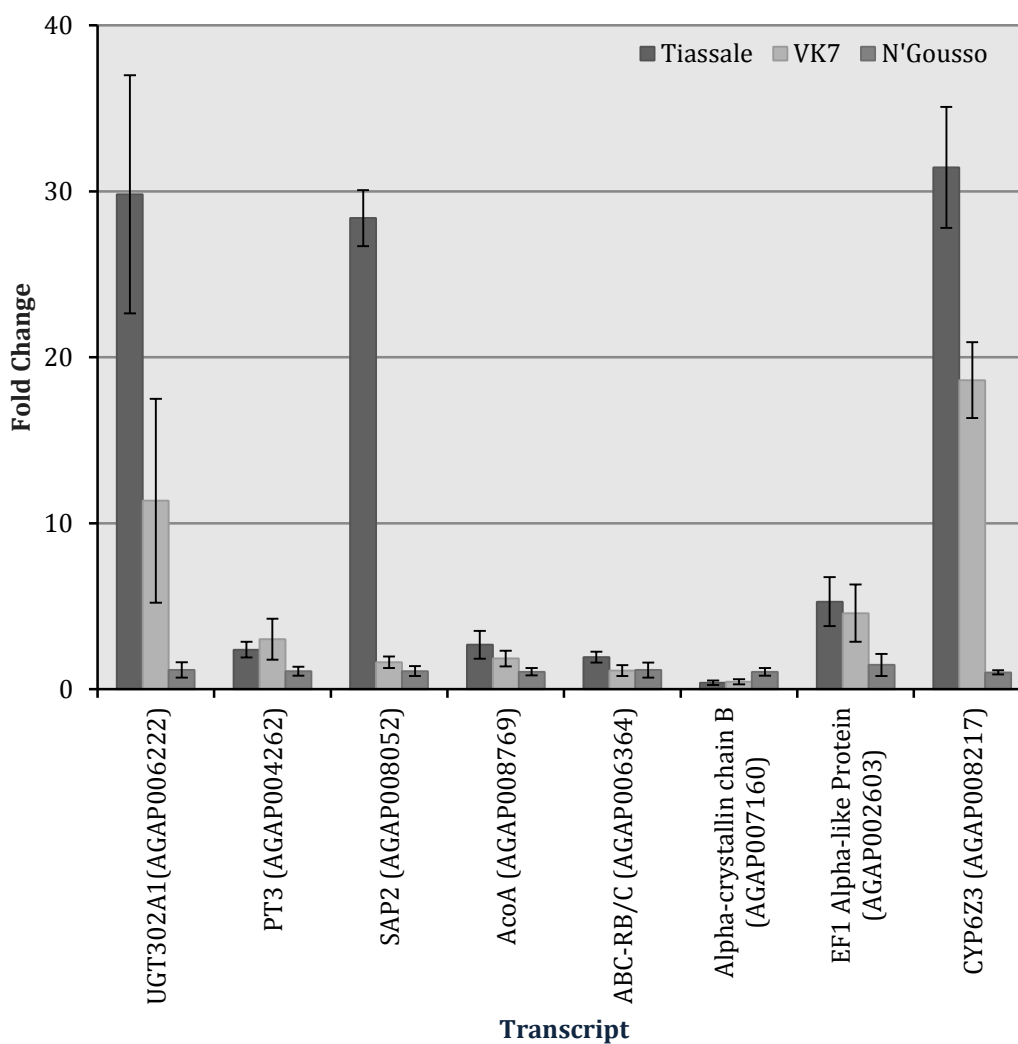
Figure 5.1: qPCR validation of eight candidate transcripts. Fold changes from

eight candidate transcripts were determined by qPCR. Two resistant

populations, Tiassalé and VK7 were compared to the lab susceptible population N'Gousso.

RNAi: Phenotypic response to insecticides.
Out of the eight candidates validated by qPCR the four most highly differentially expressed transcripts in Tiassalé were further assessed to determine if they played any role in resistance to deltamethrin. dsRNA of *UGT302A1*, *SAP2*, *Elongation factor 1-like protein* and *CYP6Z3* was synthesised and injected into three-five day old Tiassalé female mosquitoes. In each case, the level of gene silencing was calculated by comparison of the transcript levels in a subset of the injected mosquitoes with levels in uninjected and *GFP* injected controls 72 hours after injection. (Figure 5.2). Of the four candidates only two (*SAP2* and *EF1-like protein*) demonstrated any consistent silencing, with 58% and 92% reduction in transcript levels when compared to uninjected controls as determined by ΔΔct analysis. In the case of *UGT302A1* an increase of expression was seen after dsRNA injection; as this was one of the strongest candidates from the qPCR, this was repeated with a second dsRNA construct, with the same result (data not shown). Similarly in *CY6Z3* knockdown was not compared, as qPCR analysis revealed up-regulation after dsRNA injection compared to both *GFP* injected and control cDNA.
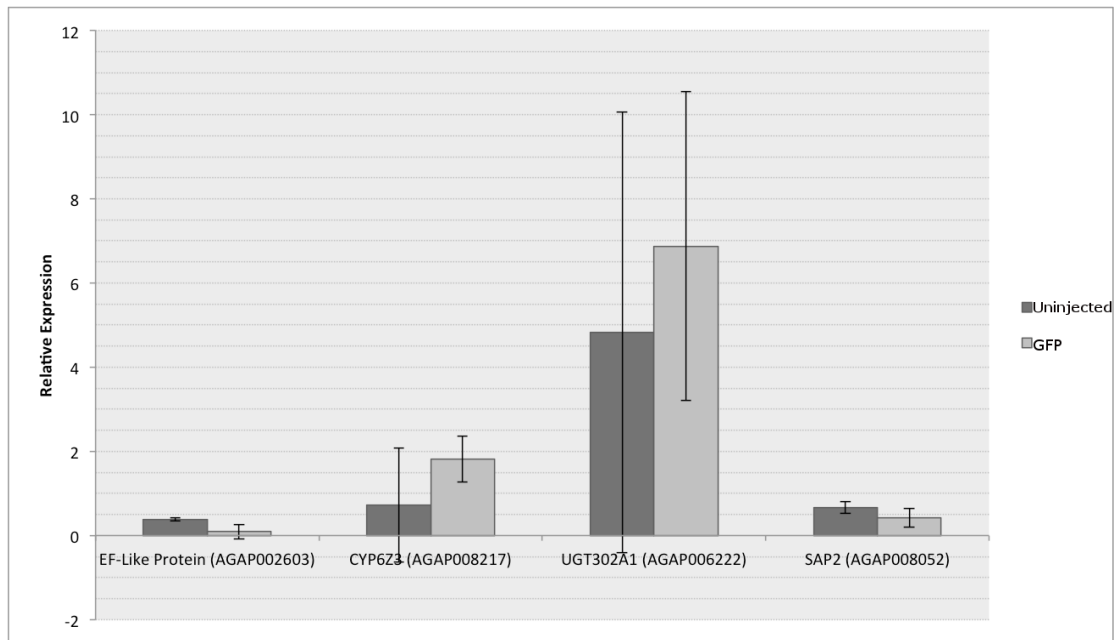
Figure 5.2: Transcript expression levels 72 hours post dsRNA injection.

Expression of four candidate insecticide resistance transcripts determined 72 hours post dsRNA injection comparing dsRNA injected to both uninjected controls and GFP injected controls. Error bars are shown using standard errors.

dsRNA injected survivors of *CYP6Z3, SAP2* and *EF-like protein* were exposed to 0.05 % deltamethrin in WHO susceptibility tubes for one hour and mortality recorded 24 hours after exposure (Figure 5.3).
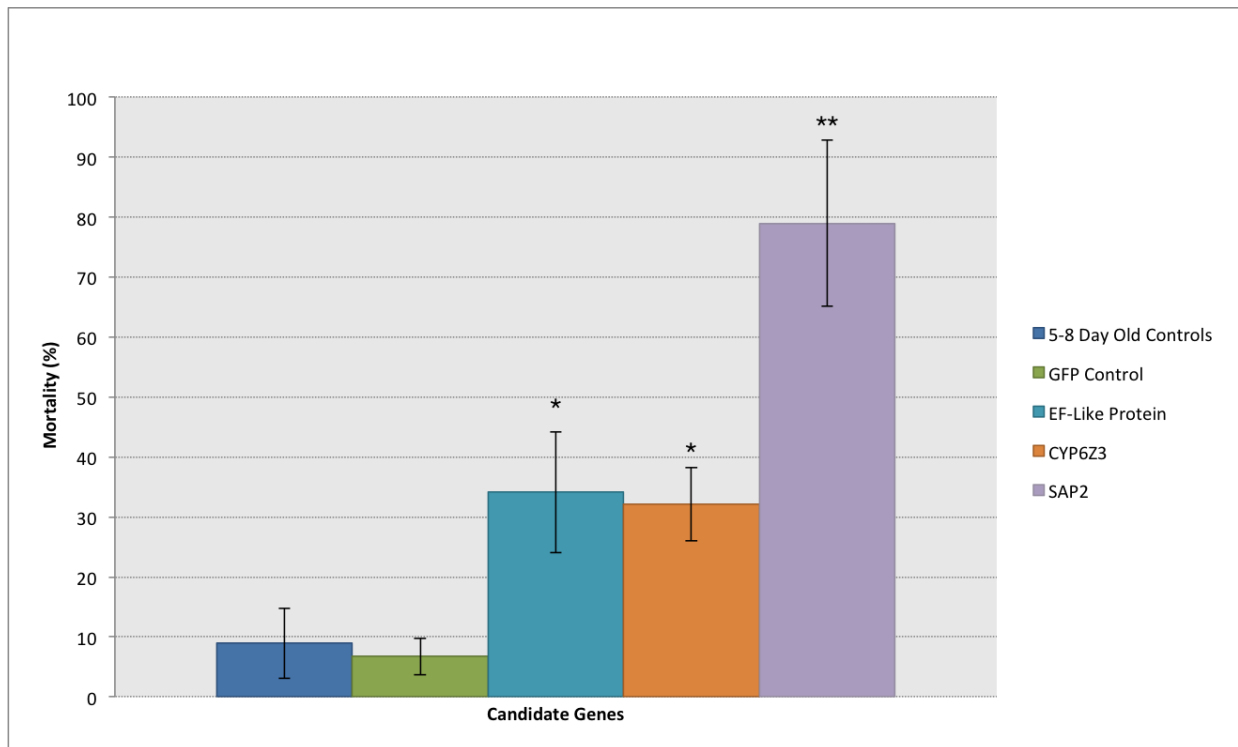
Figure 5.3: Mortality levels of dsRNA injected mosquitoes after exposure to deltamethrin. Mortality levels for each knock-down line 24 hours post exposure to 0.05 % deltamethrin for one hour in a WHO bioassay tube. 5-8 day old controls were used to determine that the resistant profile is consistent. Significance when compared to uninjected control is indicated by ** p ≤ 0.001 and * p ≤ 0.05.

Silencing of either *EF-like protein* or *SAP2* significantly reduced the level of resistance to deltamethrin in the Tiassalé strain. The greatest impact was observed in the *SAP2* silenced mosquitoes where the proportion being killed by the insecticide increased from 8.9% in the controls to 78.9% in the *SAP2* knock-downs. The *CYP6Z3* RNAi lines were also more susceptible to deltamethrin than controls. This is surprising given that qPCR showed little silencing of this gene

but may be explained by effects on closely related paralogues which warrant further investigation.

To determine whether the striking phenotype in the *SAP2* RNAi lines was specific to deltamethrin, a second population of dsRNA injected Tiassalé mosquitoes were exposed to the carbamate bendiocarb for 30 minutes in the WHO tube assay. Here, the *SAP2* silenced mosquitoes were actually more resistant to insecticide exposure. A 30-minute bendiocarb exposure resulted in 66% mortality in the controls but only 26.3% when *SAP2* was silenced (Figure 5.4). In this experiment, there was a difference in age between the uninjected and injected mosquitoes (three-five days versus five-eight days) but as resistance to bendiocarb has been shown to decrease with mosquito age (Jones *et al.*, 2012a) this is unlikely to explain the results.
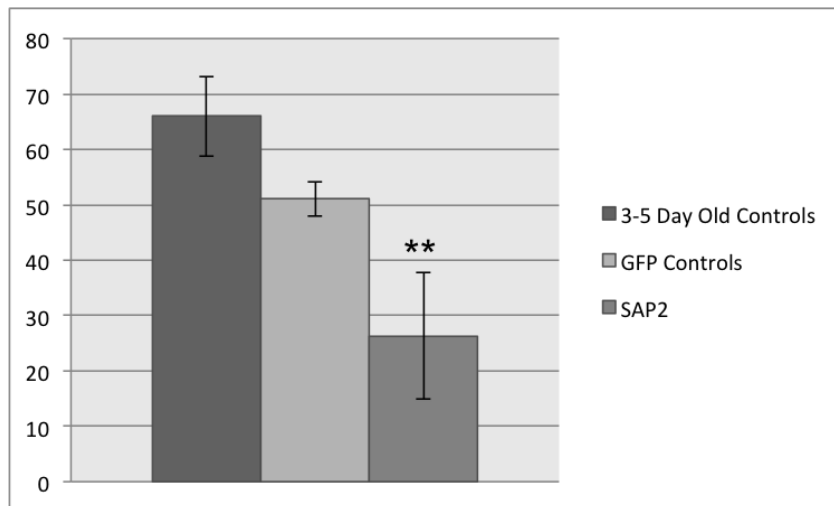
Figure 5.4: Mortality 24 hours post bendiocarb exposure. *SAP2* mortality 24 hours post 30-minute exposure to bendiocarb in a WHO tube bioassay compared to *GFP* injected and uninjected controls. Significance when compared to uninjected control is indicated by ** p ≤ 0.001. Although GFP control mosquitoes

were five-eight days old, there is no significant difference in mortality between uninjected three-five day old and five-eight day old mosquitoes.

SAP2: Validating the phenotype.

The partial restoration of susceptibility to deltamethrin observed in Tiassalé mosquitoes in which expression of *SAP2* has been significantly reduced suggests that *SAP2* overexpression may be contributing to pyrethroid resistance. To confirm the phenotype observed, a second dsRNA construct of *SAP2* was made, RNAi performed and phenotype determined (Figure 5.5). Although there is apparent overlap due to the small size of the transcript, Figure 5.5b shows that the second *SAP2* construct also has efficient knockdown (0.26 ± 0.18 compared to uninjected and 0.41 ± 0.28 compared to *GFP* injected controls) and demonstrates the same high mortality phenotype ($67.65\% \pm 18.60$) after exposure to deltamethrin tube assays

(a) GTGC<mark>TTCAAACGCAGCTTGAGAGC</mark>GTCTTCTCCACAGCAATCGTCCGCACTTTCATTTCGCGTCTTTAG
TCGCACCAGTAAACTATCACTCTTTGCCACATCATGAAACTGTTCGTCGCCATCGCTTTCGCCCTGCTG
GCCCTTGCCGCCGCCCAGGAGCAGTACACCACCAAGTACGACGGTATCGATCTGGACGAGATCCTGAAG
TCGGACCGTCTGTTCAACAACTAC<mark>TTCAAGTGCCTGATGGACGA</mark>GGGCCGCTGCACCCCGGACGGTAAC
GAGCTGAAGAAGATCCTGCCGGAAGCGCTGCAGACCAACTGTGAAAAGTGCAGCGAGAAGCAGCGCAGC
GGTGCGATCAAGGTGATCAACTACGTGATCGAGAACCGCAAGGAGCAG<mark>TGGGATGCTCTGCAGAAGAA</mark>G
TACGATCCGGAGAATCTGTACGTCGAGAAGTACCGCGAGGAGGCCAAGAAGGAGGGCATCAAGCTGGAA
TAAGCTTGATTAGCGGGGGTGATGTAAGCGATCGAT<mark>ATGGTGGTCGCTTTGTGATACT</mark>TGTTTGTTGAT
GTTATCTCAAAAATAAATGTGATAAAAATTTGGCTCTAAAATA
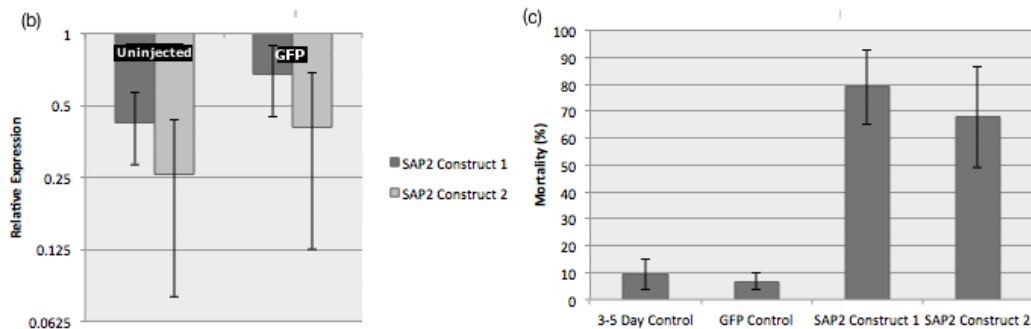
Figure 5.5: Two dsRNA SAP2 constructs. To confirm the phenotype observed by *SAP2* RNAi knockdown a second dsRNA product was made. (a) cDNA sequence of *SAP2* (AGAP008052-RA) with first (pink) and second (yellow) primers mapped. (b) qPCR knockdown of each *SAP2* construct compared to uninjected and GFP injected controls. (c) Percentage mortality 24 hours post 1 hour WHO deltamethrin tube exposure for each *SAP2* construct and both uninjected and GFP injected controls. Both show significance of $p < 0.001$.

SAP2: Localisation and induction.

To investigate the location of expression of *SAP2* qPCR was performed on cDNA extracted from both dissected appendages (heads, antennae, legs) and compared to whole body extractions.

SAP2 showed strong enrichment in both the antennae and the legs, with some enrichment in the head (Figure 5.6). It is likely the enrichment in the head is an artifact due to the difficulty in removing the whole antennae from the head and presence of maxillary palps. *SAP2* expression was enriched 321-fold in the antennae and 32-fold enrichment was observed in the legs compared to the whole body.
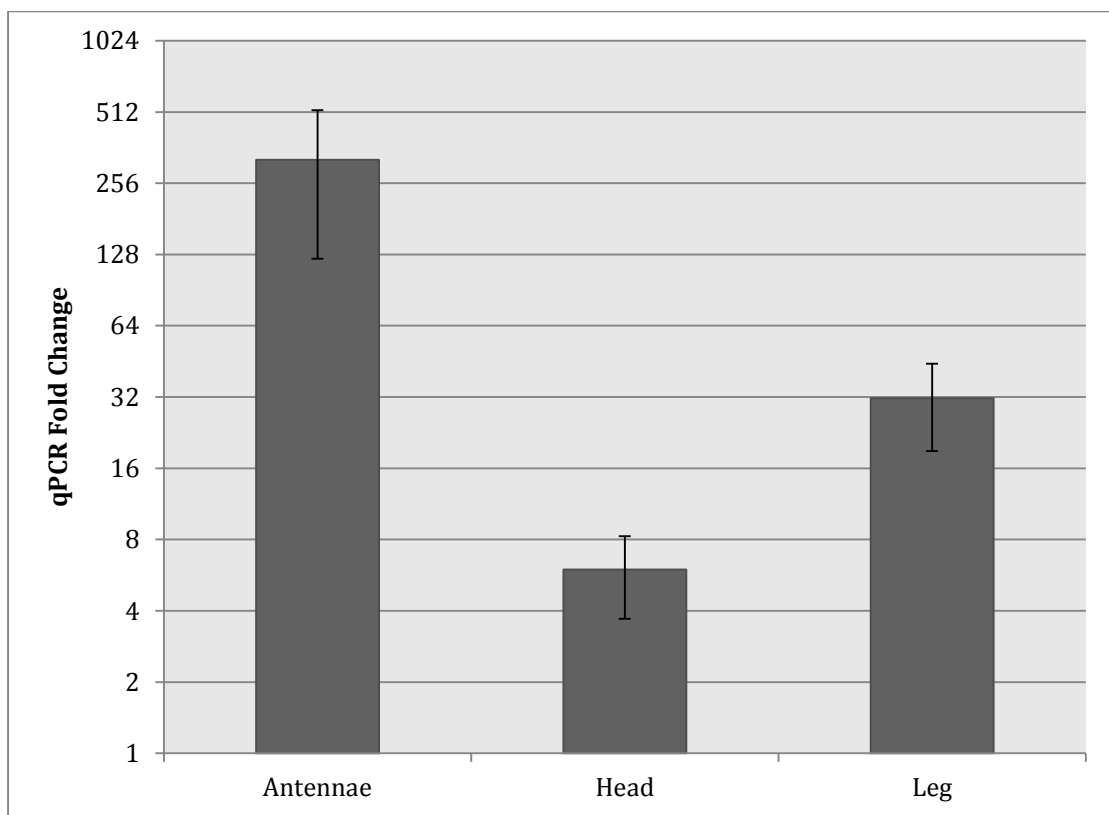
Figure 5.6: SAP2 localisation. qPCR of *SAP2* on dissected structures from resistant Tiassalé mosquitoes. Antennae, heads and legs were dissected from 40-50 adult 3-5 day old female Tiassalé mosquitoes in three biological replicates, whilst whole body cDNA was extracted from 7-10 adults in three biological replicates. ΔΔct was calculated by comparing each dissected body part to the whole organism.

CSPs are induced in response to thiamethoxam (a neonicotinoid) exposure in the whitefly, *Bemisia tobaci* (Gnankiné *et al.*, n.d.). To test for induction of *SAP2* after exposure to deltamethrin, pools of mosquitoes were tested across 4 time points: 0 hours (unexposed), 24 hours, 48 hours and 72 hours. qPCR was performed, comparing each exposed time point to unexposed mosquitoes. Figure 5.7 shows that *SAP2* appears to be expressed constitutively, with no obvious induction after insecticide exposure; furthermore there is evidence for depletion of *SAP2* 48 hours (p = 0.033) after insecticide exposure, the time at which the arrays were performed, demonstrating *SAP2* may be even more highly expressed in resistant mosquitoes than seen on the arrays.
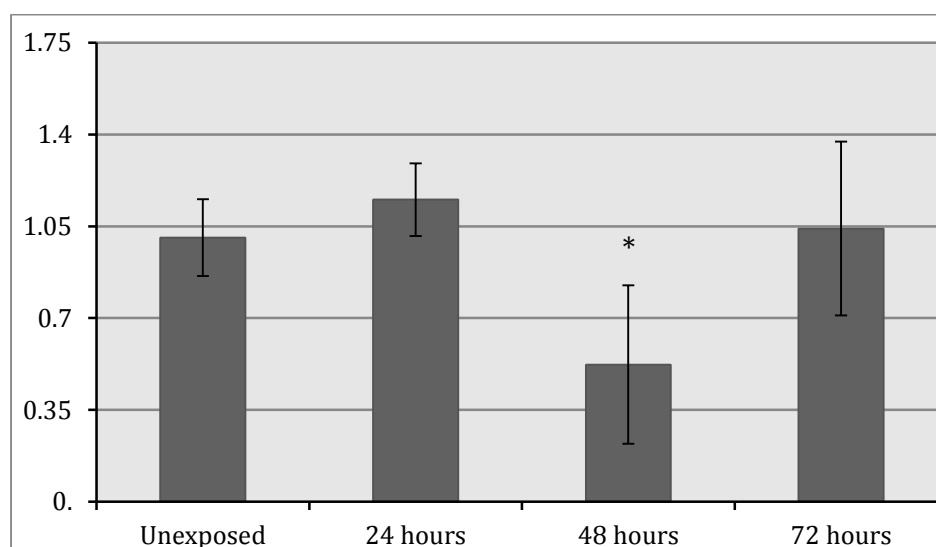
Figure 5.7: SAP2 transcript expression post deltamethrin exposure. qPCR of SAP2 on Tiassalé mosquitoes 24, 48 and 72 hours post exposure to one hour 0.05% deltamethrin papers, using a standard WHO bioassay, compared to unexposed controls, significance (p ≤ 0.05) marked by *.

To investigate whether the knockdown of *SAP2* caused a fitness cost that contributed to the increased susceptibility to pyrethroids, a survival analysis was carried out comparing *SAP2* injected mosquitoes with uninjected controls (Figure 5.8). No significant difference in survival was recorded over two weeks of observation, suggesting no major fitness cost in the *SAP2* silenced mosquitoes under the laboratory conditions tested.
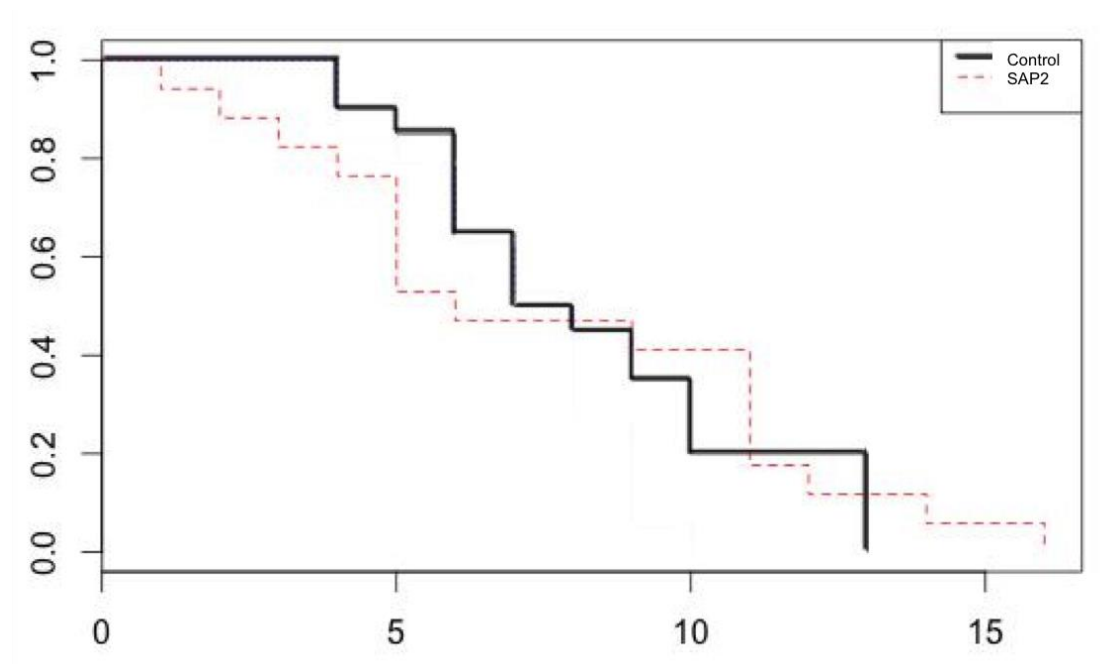
Figure 5.8: Kaplan-Meier plot of survival on injected and uninjected mosquitoes. Mortality of *SAP2* injected non-blood fed females compared to uninjected controls over two weeks (x). The mosquitoes were provided continual access to sugar solution throughout the experiment. No significant (p = 0.20) difference was seen between the test and control variables (test determined using survival package in R (Therneau & Grambsch, 2000)).

Discussion

Seven of the eight transcripts selected from the candidate gene lists produced in Chapter 4 were confirmed to be overexpressed in resistant populations by qPCR with four (AGAP008052, AGAP008217, AGAP006222 and AGAP002603)

showing over five times up-regulation in insecticide resistant populations when compared to a susceptible population. These four transcripts have varying functions: AGAP008052 (*SAP2*) is a chemosensory protein that has previously been shown to bind medium-sized aromatic compounds (Iovinella *et al.*, 2013), AGAP008217 (*CYP6Z3*) a cytochrome p450 that has been shown to metabolise insecticides (Chandor-Proust *et al.*, 2013), AGAP006222 (*UGT302A1*) a member of the UDP-glucosyltransferases, which represent key members in the third stage of xenobiotic detoxification (Xu *et al.*, 2005) and AGAP002603 (*elongation factor alpha-like protein*) whose *Drosophila* homolog (*HSB1*) has a role in non-stop mRNA decay through activation of an ATP-binding cassette, degrading mRNA lacking STOP codons (Kashima *et al.*, 2014). Of these, only *the EF-like protein* and *SAP2* were successfully silenced through dsRNA injections and mosquitoes with reduced expression of either of these genes were significantly more susceptible to deltamethrin. Silencing *SAP2* almost completely restored susceptibility to deltamethrin in an extremely pyrethroid resistant *Anopheles* population, Tiassalé. Surprisingly, silencing of *SAP2* had an opposite effect on carbamate resistance, with the gene knock-down mosquitoes showing increased bendiocarb associated mortality. The differing phenotype observed between deltamethrin and bendiocarb is surprising and hint at a more complex role in response to insecticide than simply sequestration.

Chemoreception is a key mechanism widely utilised by insect species for exploring and interacting with their environments (Krieger & Ross, 2002; Matsuo *et al.*, 2007; Asahina *et al.*, 2008). Insects are exposed to a vast compendium of chemicals, including insecticides, to which they must respond

appropriately in order to survive (Suh *et al.*, 2014). Hematophagous insects in particular must use chemoreception to identify a host for a blood meal (Mboera *et al.*, 2000; Guerenstein & Lazzari, 2009). These responses are controlled by large families of chemosensory receptors located in the sensilla, which are distributed across appendages such as legs, antennae and maxillary palps (reviewed: Guidobaldi *et al.*, 2014). In order to be sensed, the odorants must diffuse through pores, across the sensillum lymph and finally to the chemosensory proteins found on dendrite membranes of olfactory receptor neurones (Steinbrecht, 1997; Steinbrecht & Stankiewicz, 1999).

Two classes of chemosensory receptors have been documented in insects, chemosensory proteins (CSPs) and odorant binding proteins (OBPs) (Pelosi *et al.*, 2006). Both are small soluble polypeptides found in high concentration around olfactory receptor neurons and split into families based on sequence and structural homology (Pelosi *et al.*, 2006). Both CSPs and OBPs are thought to solubilise hydrophobic odorants and transport them across the sensillum lymph through a variety of mechanisms (Leal *et al.*, 1999; Kaissling, 2001; Suh *et al.*, 2014). OBPs are present in high numbers in mosquito genomes, with 69, 109 and 111 identified in *An. gambiae, Culex quinquefasciatus* and *Aedes aegypti* respectively (Manoharan *et al.*, 2013). In contrast, just seven CSPs have been identified in *An. gambiae* (Iovinella *et al.*, 2013) and 27 in *Cx. quinquefasciatus* (Pelletier & Leal, 2011).

SAP2 is one of the seven chemosensory proteins found in the genome of *An. gambiae,* with an eighth defined as a precursor (Iovinella *et al.*, 2013). CSPs are

believed to play a primary role in detection of chemicals and are expressed in sensory organs, such as the antennae, legs and maxillary palps (Guidobaldi *et al.*, 2014; Suh *et al.*, 2014). Although ligands for chemosensory proteins are not well studied, binding activity to medium-sized aromatic compounds has been demonstrated for four of the *An. gambiae* CSPs (Iovinella *et al.*, 2013). Database searches, and analysis of the microarray data comparing expression in different body parts described in Chapter 3 (Ingham *et al.*, 2014) and MozAtlas (Chintapalli *et al.*, 2007) showed that four of the CSP genes were enriched in the head, carcass or abdomen integument (Table 5.2) consistent with available literature (Guidobaldi *et al.*, 2014), whereas others were enriched in the midgut and/or malpighian tubule. Although this appears unusual it is well documented that CSPs play a role outside of chemosensation (Nomura *et al.*, 1992; Maleszka *et al.*, 2007; Guo *et al.*, 2011). Localisation in *An. gambiae* was determined by qPCR with >200 fold enrichment observed in the antennae, whilst legs and heads also demonstrated enrichment compared to the whole body. The presence of *SAP2* in both the antennae and legs may indicate a role for transportation of insecticides through the sensillum lymph and to olfactory neurons in these appendages, triggering a behavioural or physiological response. *SAP2* may also act as a sequestration mechanism, by binding to and hence lowering the concentration of free deltamethrin in the mosquito; this could be explored using a competitive binding assay with a second compound with a higher affinity for *SAP2*.

ID	Name	MozAtlas Enrichment	Chapter Three (Body parts in which gene is enriched)	Chapter 3 (Body parts in which gene is depleted)

AGAP008051	*SAP1*	Carcass/Head	Abdomen integument, midgut and Malpighian tubules	Remaining undissected structures
AGAP008052	*SAP2*	N/A	Abdomen integument, midgut and Malpighian tubules	Remaining undissected structures
AGAP008054	*SAP3*	Carcass/Head	Abdomen integument and midgut	Remaining undissected structures
AGAP008055	*CSP3*	Head	Abdomen integument, midgut and Malpighian tubules	Remaining undissected structures
AGAP008058	*CSP5*	No enrichment	Midgut and Malpighian tubules	Abdomen Integument
AGAP008059	*CSP1*	Carcass	Midgut and Malpighian tubules	Abdomen Integument
AGAP008062	*CSP4*	No enrichment	Midgut	None

Table 5.2: Enrichment of *An. coluzzii* CSPs. Demonstrates current knowledge on the expression of the seven confirmed Anopheline CSPs, based both on MozAtlas (male and female dissections on: head, midgut, malpighian tubules, testis, ovaries, carcass, salivary glands and whole body) (Baker *et al.*, 2011) and Chapter 3 (Ingham *et al.*, 2014).

In conclusion, *SAP2* appears to be a very promising and novel candidate for pyrethroid resistance in *An. gambiae* mosquitoes, almost completely restoring susceptibility to deltamethrin after dsRNA-induced knockdown. Although the

mechanism has not been studied here, *SAP2* was shown to be enriched in sensory appendages supporting a hypothesis that *SAP2* may detect and bind to pyrethroid insecticides. Unlike previous reports of induction after exposure to insecticides, *SAP2* appears to be constitutively expressed in resistant *An. gambiae* mosquitoes, putatively pointing to a crucial mechanism of resistance that has not yet been documented.

6. TRANSCRIPTIONAL REGULATION OF XENOBIOTIC RESPONSE: *MAF-S*

Unpublished

V A Ingham, P Pignatelli, S C Wagstaff, J D Moore and H Ranson.

V.A.I., performed all *in silico* analysis. V.A.I., J.D.M., H.R. and S.C.W. designed the study. V.A.I. and P.P. performed the microarray experiments, with V.A.I. performing all other experiments.

Introduction

Transcriptomic responses to insecticide selection in *Anopheles* species have been well documented in the form of microarray data, comparing resistant and susceptible populations (Jones *et al.*, 2013; Mawejje *et al.*, 2013; Hemingway *et al.*, 2013; Abdalla *et al.*, 2014; Toé *et al.*, 2015). A large number of transcripts whose expression is associated with insecticide resistance have been identified, including cytochrome p450s, GSTs, carboxylesterases and UDP-transferases; major families in the insect detoxification system (Xu *et al.*, 2005). However, little to no information exists on the regulatory processes governing expression of these genes. Of the few transcription factors explored for a role in co-ordinating xenobiotic response, *Maf-S* has been experimentally shown to influence responses to xenobiotic compounds in *Drosophila*, including the insecticides DDT and malathion by constitutive over-expression (Misra *et al.*, 2011; Misra *et al.*, 2013).

Maf-S is a small nuclear-located transcription factor that heterodimerises with *Cap 'n' Collar* (*cnc*) in order to bind to antioxidant response elements in the genome (Figure 6.1). The binding of *Maf-S-cnc* is controlled by the cytoplasmic location of *cnc*, which is bound by an actin binding ubiquitin ligase, *keap1*, in the absence of electrophiles and reactive oxygen species (Veraksa *et al.*, 2000). *Keap1* is present in *Anopheles* in a one-to-two homology; both homologs are annotated as Kelch-like proteins showing relatively sequence coverage (<40%) and identity of 73% and 75%; indicative of shared domains. The *Maf-cnc-Keap1* pathway is conserved from humans to *Drosophila* and appears to show the same regulatory interactions (Sykiotis & Bohmann, 2008; Sykiotis & Bohmann, 2010).

Interestingly, *Maf-S-cnc* was shown to be necessary for the induction of the

Drosophila cytochrome *CYP6A2* (Misra *et al.*, 2011). Several members of the

CYP6 family of P450s in *Anopheles* have been previously implicated in resistance

to pyrethroids and DDT (Chiu *et al.*, 2008; Müller *et al.*, 2008b; Chandor-Proust *et*

al., 2013). Furthermore, the midgut and malpighian tubules have previously

been shown to be hubs of detoxification gene up-regulation in *Anopheles*

(Ingham *et al.*, 2014), consistent with the main localisation of *Maf-S-cnc* in

Drosophila (Misra *et al.*, 2013) and in *Anopheles* (Baker *et al.*, 2011; Ingham *et al.*,

2014) to these structures. This evidence taken together, indicate that the *Maf-S-*

cnc-Keap1 pathway may be a key regulator in xenobiotic response in *Anopheles*

mosquitoes.

In this chapter, the role of the *Maf-s-cnc* pathway in regulating detoxification

family genes was investigated by RNAi attenuation of *Maf-S.* These data give

strong indication that *Maf-S* is involved in the transcriptional regulation of

xenobiotic response in the malaria vector species, *Anopheles.*
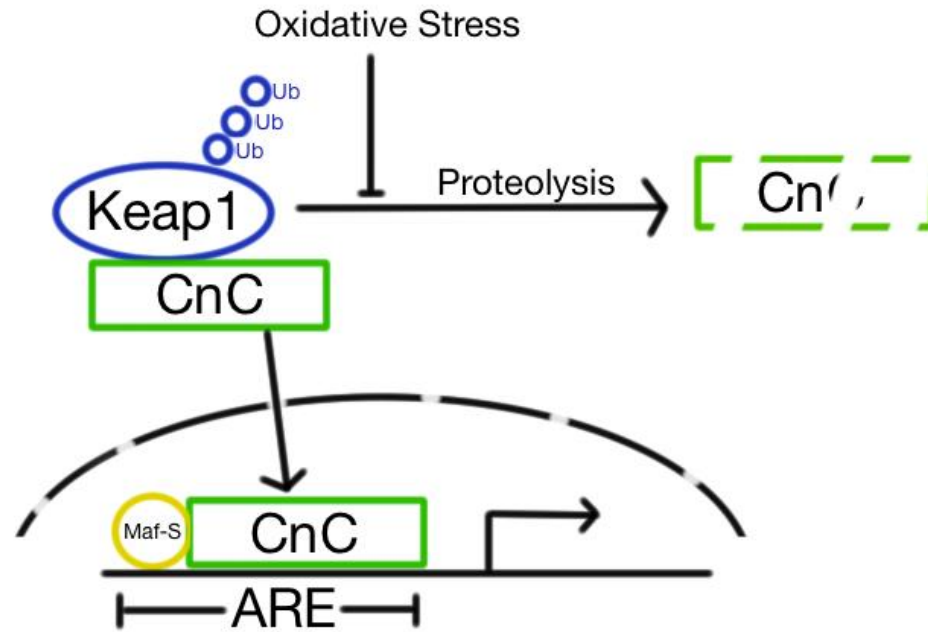
Figure 6.1: *Maf-S-cnc-Keap1* **pathway.** *Keap1* is a ubiquitin ligase that is bound

to *CnC* and breaks it down via proteasomal degradation in the absence of

oxidative stress. When the cell is put under oxidative stress, proteolysis is

blocked and *CnC* is released by *Keap1*, at which point *CnC* enters the nucleus,

binds freely available *Maf-S* and hence antioxidant response elements in the

genome, initiating transcription.

Aims

1. Identify a role for the *Maf-S-cnc-Keap1* pathway using previously analysed

microarray data comparing resistant and susceptible mosquito populations.

2. Use RNAi knockdowns to determine whether the above pathway results in

changes to detoxification transcripts.

3. Perform a microarray experiment to validate the role of the above pathway in

involvement in xenobiotic response in *Anopheles* mosquitoes.

4. Use the microarray data to identify other pathways and roles for the above

pathway in *Anopheles* mosquitoes.

Results

Twenty-seven microarray studies, representing a resistant compared to a susceptible population of *An. coluzzii*, *An. gambiae* or *An. arabiensis* (Table 4.1) were analysed using limma analysis. These data were used to examine the expression of the *Maf-S-cnc-Keap1* pathway in *Anopheles* mosquitoes. Both *Maf-S* and *cnc* were differentially expressed in 20 experiments, whilst one of the *Keap1* homologs was differential in 18, significantly greater than expected by chance (p = 0.0028, p = 0.0028 and p = 0.019) (Figure 6.2). *Maf-S* is up-regulated in 18 out of the 20 experiments, whereas both *Keap1* and *cnc* show both up- (11 experiments for each) and down-regulation across the experiments. In cases with different directionality of *cnc* and *Maf-S* it may be that these transcription factors perform differing functions depending on heterodimer formation, as with mammalian *Maf*, the homolog of *Maf-S* (Hegde *et al.*, 1998; Melloul *et al.*, 2002; Kannan *et al.*, 2012).
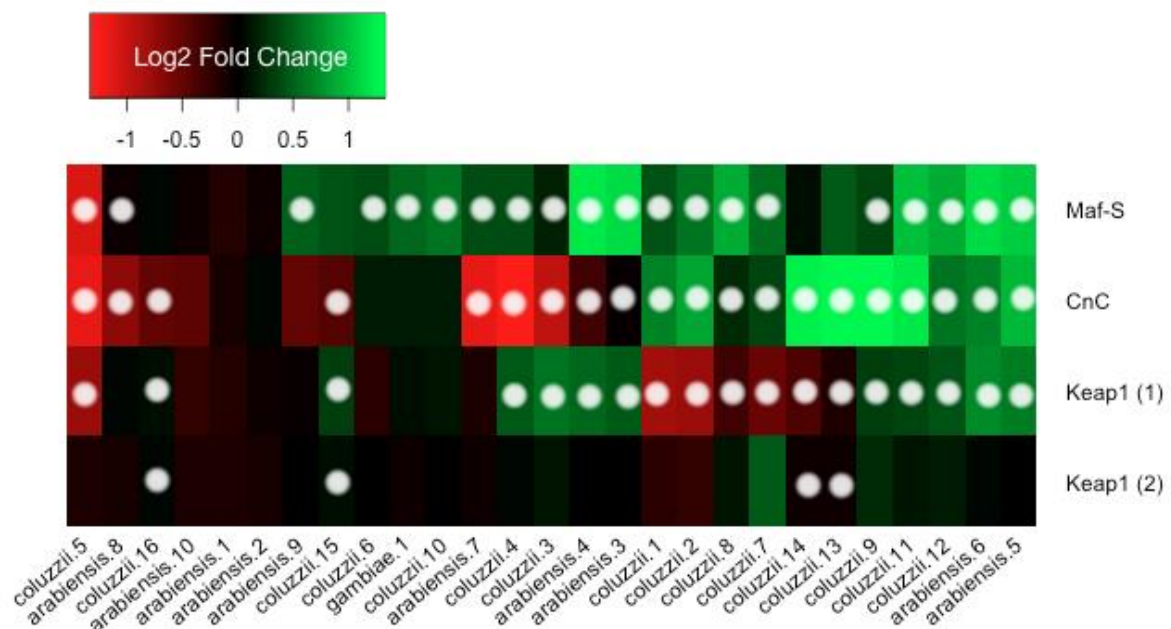
Figure 6.2: Log$_2$ fold change values for the *Maf-S-cnc-Keap1* pathway. Log$_2$ fold changes used to cluster the 27 experiments based on expression of *Maf-S*

(AGAP010405), *cnc* (AGAP005300) and two *Keap1* homologs ((1) AGAP003645 and (2) AGAP012550). White circles represent a significance of p ≤ 0.05.

Correlation networks reveal putative transcriptional interactions
To explore potential functions of the pathway, correlation networks were produced for *Maf-S*, *cnc*, and the *Keap1* homologs using the 27 studies previously analysed [Chapter 4]. These correlation networks allow identification of transcripts whose expression is co-ordinately seen across multiple array experiments, each comparing insecticide resistant and susceptible *Anopheles* mosquitoes. The correlation networks were stringent and took into account both correlation and anti-correlation by applying a cut off of ±0.8. *cnc* has strong correlations with 254 other transcripts; GO term enrichments of these transcripts demonstrated roles in signalling, development and regulation and were also significant for transcriptional cofactor activity, all of which are consistent with a role in transcriptional activation (p ≤0.05). Both *Keap1* homologs had few correlated transcripts with just three for AGAP003645, including *TPX4*, a juvenile hormone inducible protein and a GTPase and one for AGAP012550 with unknown function. *Maf-S* however, had 14 transcripts that demonstrated strong correlations, 6 of which represented members of *a priori* detoxification families (Table 6.1, Figure 6.3), a significant enrichment in detoxification families (p < 0.001). Amongst the detoxification transcripts is *CYP6M2,* a proven pyrethroid metaboliser (Djègbè *et al.*, 2014). The large difference in number of co-correlated transcripts could be explained by the nature of *cnc*'s interaction with *Keap1*, with under- or over-expression of this transcript leading to (i) all *cnc* bound to *Keap1* or (ii) saturation of *Keap1* and

free *cnc* respectively; whilst *Maf* over-expression may result in a faster

heterodimerisation with *cnc.*

Identifier	Correlation	Description
AGAP011911-RA	0.800	F-box protein 39
AGAP007504-RA	0.861	ATP-binding cassette sub-family A member 3
AGAP003146-RA	-0.841	RalA-binding protein 1
AGAP011892-RA	0.817	Required for meiotic nuclear division 5 homolog A
AGAP001297-RA	0.833	Solute carrier family 25
AGAP008212-RA	0.816	Cytochrome P450 - CYP6M2
AGAP004164-RD	0.801	Glutathione S-transferase - GSTD1_6
AGAP006662-RA	-0.818	Unknown
AGAP004382-RA	0.829	Glutathione S-transferase - GSTD3
AGAP007589-RA	0.835	Glucosyl/glucuronosyl transferases - UGT306A2
AGAP008358-RA	0.851	Cytochrome P450 - CYP4H17
AGAP011749-RA	0.803	Unknown
AGAP001995-RA	0.805	20S proteasome subunit alpha 2
AGAP003320-RA	0.824	Unknown

Table 6.1: *Maf-S* correlated transcripts. AGAP identifier, correlation value and

description of *Maf-S* correlated probes.
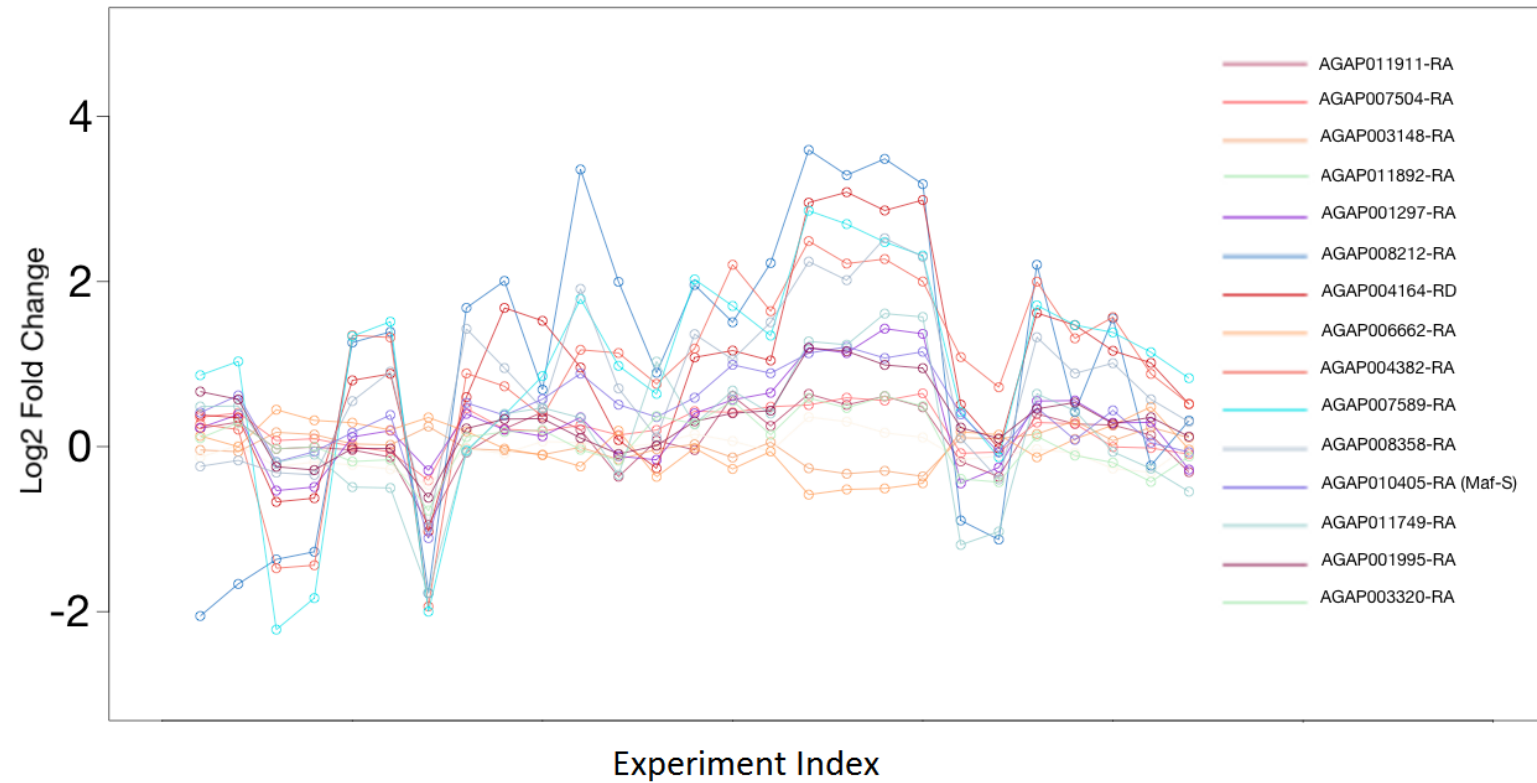
Figure 6.3: *Maf-S* correlation network. Log$_2$ fold change (y) of *Maf-S* and co-correlated transcripts across 27 analysed studies [Chapter 4] (x).

Maf-S knockdown and influence on putative detoxification interactors

To test whether _Maf-S_ was regulating expression of other genes in the correlation network, the expression of four of the detoxification transcripts in _Maf-S_ RNAi knockdowns was measured and compared to _GFP_-injected controls. The _Maf-S_ RNAi knockdown level was determined by qPCR, with the creation of two constructs (Figure 6.4, Figure 6.7). Construct two showed the greatest reduction in expression at 85.9% compared to uninjected control and so was used on all further experiments. Suppressing expression of _Maf-S_ resulted in reduced expression of two P450s (_CYP6M2_ and _CYP4H17_) and a glutathione transferase (_GSTD3_) as predicted by the correlation network. In contrast, _ABCA3_ whose expression was anti-correlated with _Maf-S_ in the arrays, showed elevated expression when _Maf-S_ was silenced suggesting that this gene is repressed by _Maf-S_ (Figure 6.5).
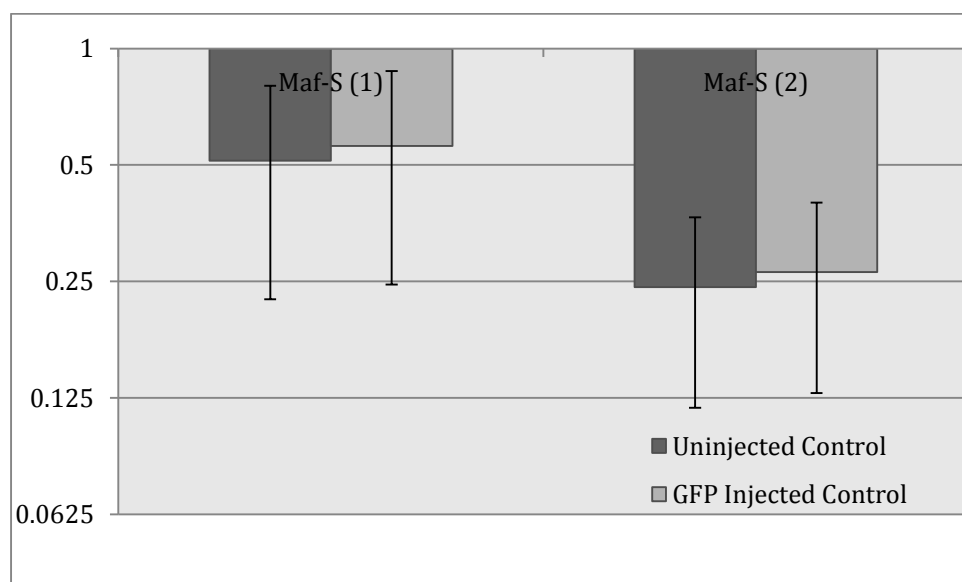
Figure 6.4: Knockdown levels of two _Maf-S_ constructs. qPCR confirmation of two _Maf-S_ knockdowns, using two different constructs compared to uninjected and GFP injected controls. Log_2 $\Delta\Delta CT$ values (y) and two constructs (x).
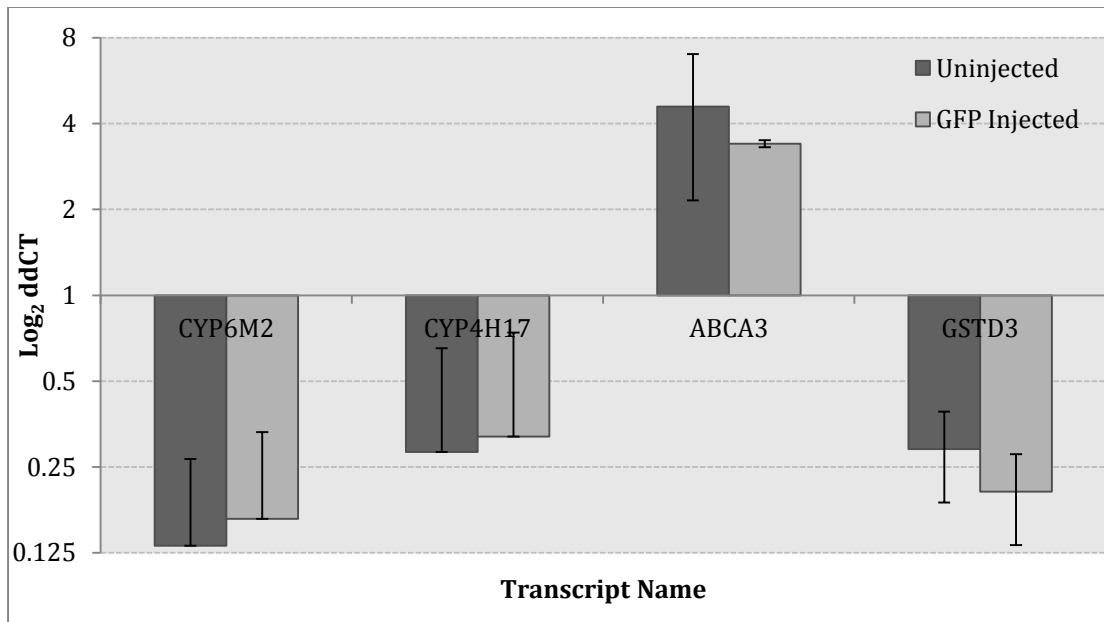
Figure 6.5: *Maf-S* influence on correlated transcripts. qPCR of four transcripts represented in the *Maf-S* correlation network, performed on *Maf-S* knockdown cDNA normalised against expression in a GFP-injected control and an uninjected control. 3 biological replicates were used, with cDNA extracted from 7-8 day old females in injected samples; this represented 72 hours post injection.

Transcriptional response to Maf-S knockdown
The above qPCR experiments suggested that *An. gambiae Maf-S* plays a similar role to its ortholog in *Drosophila melanogaster* in regulating the expression of detoxification gene families. To further investigate the genes putatively regulated by this transcription factor, microarrays were performed comparing *Maf-S* knockdown and *GFP*-injected control dsRNA. Three biological replicates were used in each instance with each replicate consisting of seven to 10 adult females. Females were injected at three-five days old, left for 72 hours and RNA extracted. RNA was compared between *Maf-S* injected and *GFP*-injected controls, to account for transcript changes due to dsRNA injection. Six arrays were

produced, using the three biological replicates and three dye swaps; one was
discarded due to poor quality scanning (Figure 6.6).
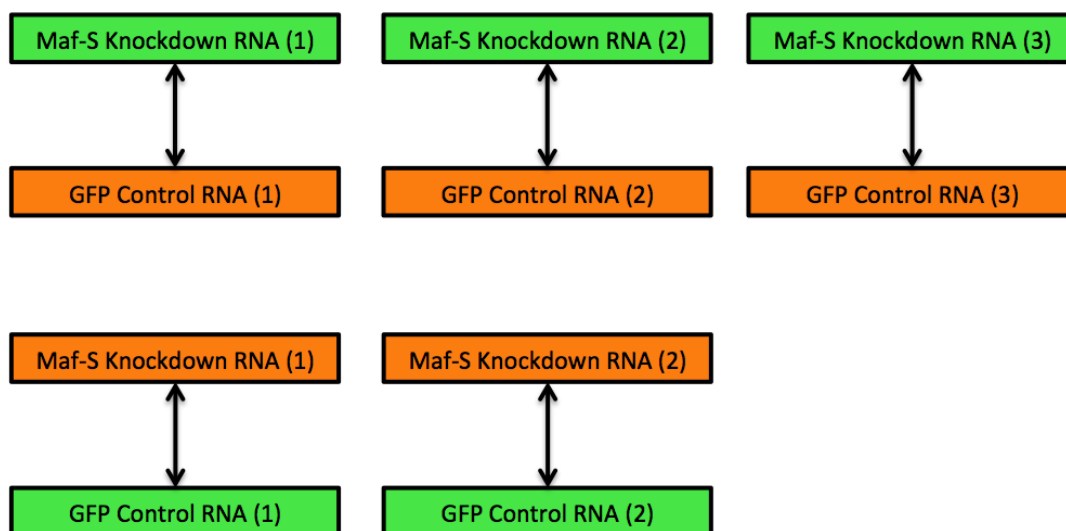
Figure 6.6: Microarray design. Designs used for microarrays comparing *Maf-S*
knockdown RNA with GFP injected control RNA. Green boxes represent RNA
tagged with *cy3* and red those tagged with *cy5*.

3401 transcripts were significantly differentially expressed ($p \leq 0.05$), of these
1703 were down-regulated and 1698 were up-regulated. The up-regulated
transcript list is highly enriched in ion transport, ATP and purine processes and
ligand channel activity (Table 6.2). In contrast, the down-regulated transcript list
shows enrichment in key processes such a RNA binding and ribonucleoprotein
(Table 6.2).

Given the role of *Maf-S-cnc-Keap1* in xenobiotic response, the numbers of
detoxification gene family transcripts were measured. In total, 64 transcripts
representing detoxification genes were significantly differentially expressed, 31

in the up-regulated list and 33 in the down-regulated list. 17 cytochrome p450s, two ABCs, eight GSTs, two UGTs and three COEs were present in the down-regulated transcript list (Table 6.3). Earlier qPCR experiments using the same RNA had shown *CYP6M2* and *CYP4H17* to be down regulated in Maf-S silenced mosquitoes and this was confirmed in the microarray: both genes are significantly down-regulated (FC = 0.41, 0.69). A further five cytochrome p450s belonging to the CYP6 subfamily were also down-regulated. Up-regulated detoxification transcripts include five ABCs, six COEs, 15 cytochrome p450s, three GSTs and twos UGTs. In addition to detoxification transcripts, 24 cuticular proteins are differentially expressed after *Maf-S* knockdown, 19 of which are up-regulated. Of the 14 transcripts showing correlation with *Maf-S* across the 27 microarray experiments comparing resistant and susceptible mosquitoes, four are also significant on the array and are the only transcripts to show the predicted fold change directionality. The low number of transcripts may be because the stringency for the correlation was too low (it was set at ±0.8 but >±0.9 may be required for accurate identification of co-correlated transcripts). *Maf-S* itself is not significantly differently expressed on the array; this may be due to the different location of the array probe compared to the construct used (Figure 6.7) or due to the incomplete knockdown of *Maf-S*.

GO Term	Adjusted p-value	Transcript List
Ion transport	4.70E-03	Up
Generation of precursor metabolites and energy	8.20E-03	Up
ATP metabolic process	3.00E-02	Up
ATP biosynthetic process	3.00E-02	Up
Purine nucleoside triphosphate metabolic process	4.80E-02	Up
Ribonucleoside triphosphate biosynthetic process	4.80E-02	Up
Purine ribonucleoside triphosphate metabolic process	4.80E-02	Up

Purine nucleoside triphosphate biosynthetic process	4.80E-02	Up
Ribonucleoside triphosphate metabolic process	4.80E-02	Up
Purine ribonucleoside triphosphate biosynthetic process	4.80E-02	Up
Nucleoside triphosphate biosynthetic process	4.80E-02	Up
Transmembrane transport	4.50E-02	Up
Nucleoside triphosphate metabolic process	4.10E-02	Up
Calcium ion binding	4.00E-04	Up
Inorganic cation transmembrane transporter activity	1.50E-02	Up
Gated channel activity	1.60E-02	Up
Hydrogen ion transmembrane transporter activity	2.60E-02	Up
Ligand-gated ion channel activity	2.60E-02	Up
Monovalent inorganic cation transmembrane transporter activity	2.60E-02	Up
Ligand-gated channel activity	2.60E-02	Up
Substrate specific channel activity	3.20E-02	Up
Ion channel activity	3.20E-02	Up
Passive transmembrane transporter activity	3.10E-02	Up
Channel activity	3.10E-02	Up
Membrane-enclosed lumen	1.50E-03	Down
Cytosol	2.00E-03	Down
Chromosome	5.50E-03	Down
Ribonucleoprotein complex	6.90E-03	Down
Intracellular organelle lumen	7.00E-03	Down
Organelle lumen	7.00E-03	Down
Nuclear lumen	7.20E-03	Down
Intracellular	1.00E-02	Down
Non-membrane-bounded organelle	1.00E-02	Down
Nucleoplasm	1.10E-02	Down
Chromosomal part	1.10E-02	Down
Nucleoplasm part	3.50E-02	Down
RNA binding	4.80E-02	Down

Table 6.2: Significant GO terms for significantly up- and down-regulated transcripts. Significant GO terms for transcript lists significantly up- (green) or down- (blue) regulated after *Maf-S* knockdown. GO terms, Benjamini adjusted p-values and transcript membership shown as calculated by DAVID (Huang *et al.,* 2007).

Systematic ID	Gene Name	Fold Change	Fold Change Directionality
AGAP004383-RA	GSTD10	3.42	Up

AGAP002090-RA	COE '16738'	1.95	Up
AGAP013490-RA	CYP4H24	1.88	Up
AGAP000506-RA	ABC-G	1.85	Up
AGAP000088-RA	CYP4H19	1.82	Up
AGAP003066-RA	CYP304B1	1.67	Up
AGAP006227-RA	COEAE1F	1.65	Up
AGAP006726-RA	COEAE5G	1.64	Up
AGAP008436-RA	ABC-C	1.61	Up
AGAP012291-RA	CYP9J3	1.61	Up
AGAP008214-RA	CYP6M4	1.59	Up
AGAP006775-RA	UGT308D1	1.59	Up
AGAP006723-RA	COEAE2G	1.57	Up
AGAP002418-RA	CYP4D16	1.57	Up
AGAP005660-RA	CYP305A4	1.51	Up
AGAP006727-RA	COEAE6G	1.51	Up
AGAP010961-RA	CYP6AK1	1.51	Up
AGAP008022-RA	CYP12F1	1.49	Up
AGAP008401-RA	UGT301E2	1.46	Up
AGAP003608-RA	CYP4AA1	1.43	Up
AGAP000818-RA	CYP9K1	1.43	Up
AGAP008018-RA	CYP12F4	1.42	Up
AGAP008437-RA	ABC-C	1.42	Up
AGAP009799-RA	ABC-C	1.42	Up
AGAP002071-RA	ABC-D	1.36	Up
AGAP005657-RA	CYP305A3	1.34	Up
AGAP013128-RA	CYP6AA2	1.32	Up
AGAP003257-RA	GSTU2	1.29	Up
AGAP002894-RA	CYP6Z4	1.29	Up
AGAP006228-RA	COEAE2F	1.24	Up
AGAP008217-RA	CYP6Z3	0.86	Down
AGAP002898-RC	GSTZ1	0.84	Down
AGAP002429-RA	CYP314A1	0.84	Down
AGAP005837-RA	COEJHE5E	0.81	Down
AGAP006047-RA	CYP4J9	0.81	Down
AGAP009190-RA	GSTE8	0.80	Down
AGAP005752-RA	UGT308A2	0.79	Down
AGAP001076-RA	CYP4G16	0.79	Down
AGAP010404-PC	GSTS1_1	0.79	Down
AGAP008207-RA	CYP6Y2	0.78	Down
AGAP004172-RA	GSTD9	0.78	Down
AGAP002417-RA	CYP4AR1	0.78	Down
AGAP000877-RA	CYP4G17	0.75	Down
AGAP009192-RA	GSTE5	0.75	Down
AGAP008208-RA	CYP6Y1	0.75	Down
AGAP009191-RA	GSTE6	0.72	Down
AGAP008358-RA	CYP4H17	0.69	Down

AGAP002416-RA	CYP4K2	0.68	Down
AGAP009463-RA	ABC-G	0.68	Down
AGAP005372-RA	COEBE3C	0.67	Down
AGAP012294-RA	CYP9L2	0.67	Down
AGAP001333-RA	ABC-G	0.66	Down
AGAP005750-RA	UGT308B2	0.66	Down
AGAP008404-RA	UGT301C1	0.65	Down
AGAP008218-RA	CYP6Z2	0.64	Down
AGAP005749-RA	GSTO1	0.64	Down
AGAP005992-RA	CYP302A1	0.64	Down
AGAP003067-RA	CYP304C1	0.63	Down
AGAP009246-RA	CYP4C27	0.60	Down
AGAP011507-RA	COE13O	0.56	Down
AGAP007374-RA	UGT49A3	0.51	Down
AGAP002867-RA	CYP6P4	0.44	Down
AGAP008212-RA	CYP6M2	0.41	Down
AGAP004164-RA	GSTD1_6	0.35	Down

Table 6.3: Significant detoxification transcripts. 64 transcripts representing detoxification family members that show significant differential expression after knockdown of *Maf-S*. Green highlighting represents transcripts showing an increased expression after *Maf-S* knockdown and blue reduced expression.

>AGAP010405|AGAP010405-RA
CCTTGATCGGGTGAATTTCAAAGAATGGAATAAGGAAAATCGTGCAGAATTCATAAT
AACTTTGAACATCATTCTACAAAAGTTTAATTAAACTTTATAGTTGGTGCAGATCGT
ATATAATAATATTTCAAATCTCAAAAGTGTCGAGAGACTACGAAAGCGTTCGTCATT
GGTTCTGGTTGCAGTACAGCTCTATAGGGCTCAGGGAACACATTCGGTTCGATTCGA
TCGATAACTGAATAGATGTGCGTCTTCTAGAAAGTGCCGAGCAGGTGTGTGCTGTGC
AAGATCCAAACCACGTAAAATGCCGCAGGAAATTAAACGGGAACGTAAAACTGCTCA
GTCAATGATGCGCCCGAAATCTATTCACTGGAATGAAGCACCACTATCCCCCTGTCC
CATCCCCGATATTAGTGACGATGAGCTGGTGTCGATTACGGTACGCGATCTCAATCG
TACGCTTAAAATGAGAGGTCTAACCCGCGAGGAAATTGTGCGCATGAAGCAGCGGCG
TCGTACGCTGAAAAATCGAGGATATGCCGCCAGCTGCCGCATAAAGCGCATCGAGCA
GAAGGATGAACTGGACGCGGAAAAATCACAAGAATGGCGGGATATGGAGCTGATGCA
CGAGGAAACCGGTCGATTGCAGGAAGAGAACGATTCGCTCCGTAACAAATACGAGGC
GTTGCGAAAGTTTGCTCTTTCGAAAAAGATTCCACTACCGCCGGAGCTTGACGTGCT
GTGAAGTGAGGTCAAGACGTGGCCAATGACCGTGCCGAATGTTTCACTGCTGTAACT
AACTTGTTTATATTCCATCTCAGTACTTCGATTAGTATAATTGTATATTGAACATTA
AAATAATAAAGTAACATTTTATATTCTACACTCGCTTGCACGTGGCCATGCTTCGAC
ACTAGGTGCAATTGTAGATCAATTTGAATATTTG

>AGAP010405|AGAP010405-RB
GTGGAACAACAATTAGCTAGCCTTGATCGGGTGAATTTCAAAGAATGGAATAAGGAA
AATCGTGCAGAATTCATAATAACTTTGAACATCATTCTACAAAAGTTTAATTAAACT
TTATAGTTGGTGCAGATCGTATATAATAATATTTCAAATCTCAAAAGTGTCGAGAGA
CTACGAAAGCGTTCGTCATTGGTTCTGGTTGCAGTACAGCTCTATAGGGCTCAGGCA
CTATCTGGTTTGTGCCCTTTTTTCATTGCATGGCGTCACCTAAAACATTCTACCCAT
AGGAACACATTCGGTTCGATTCGATCGATAACTGAATAGATGTGCGTCTTCTAGAAA
GTGCCGAGCAGGTGTGTGCTGTGCAAGATCCAAACCACGTAAAATGCCGCAGGAAAT
TAAACGGGAACGTAAAACTGCTCAGGCACCACTATCCCCCTGTCCCATCCCCGATAT
TAGTGACGATGAGCTGGTGTCGATTACGGTACGCGATCTCAATCGTACGCTTAAAAT
GAGAGGTCTAACCCGCGAGGAAATTGTGCGCATGAAGCAGCGGCGTCGTACGCTGAA
AAATCGAGGATATGCCGCCAGCTGCCGCATAAAGCGCATCGAGCAGAAGGATGAACT
GGAGACGGAAAAATCACAAGAATGGCGGGATATGGAGCTGATGCACGAGGAAACCGG
TCGATTGCAGGAAGAGAACGATTCGCTCCGTAACAAATACGAGGCGTTGCGAAAGTT
TGCTCTTTCGAAAAAGATTCCACTACCGCCGGAGCTTGACGTGCTGTGAAGTGAGGT
CAAGACGTGGCCAATGACCGTGCCGAATGTTTCACTGCTGTAACTAACTTGTTTATA
TTCCATCTCAGTACTTCGATTAGTATAATTGTATATTGAACATTAAAATAATAAAGT
AACATTTTATATTCTACACTCGCTTGCACGTGGCCATGCTTCGACACTAGGTGCAAT
TGTAGATCAATTTGAATATTTG

Figure 6.7: *Maf-S* **construct mapping.** cDNA sequence of two *Maf-S* splice variants, showing the position of both primer set one (purple) and two (yellow) as well as the array probe positions (red). AGAP010405-RA and AGAP010405-RB differ in the splicing of a small second exon and changes to the third and fourth exon, which will knock-down both splice variants due to primer positioning. Exon boundaries are marked in green.

Pathway changes after knockdown of Maf-S

KEGG pathway was used to determine pathways that are putatively altered after *Maf-S* knockdown. Although *Anopheles* pathways are available, they are largely based on human and mouse homology and so are mostly incomplete. Nevertheless, it is possible to determine whether pathways show changes and the directionality of these changes for key metabolic pathways. Using CytoKEGG (https://code.google.com/p/cytokegg/), a KEGG plugin for Cytoscape, pathways specific for *Anopheles* species were used. The top two pathways were represented by purine metabolism and pyrimidine metabolism, followed by the spliceosome structure, with 49, 42 and 38 differential transcripts respectively representing 40.2%, 65.6% and 48.1% of the steps in the pathways (Figure 6.8). The large scale down-regulation seen in both the purine and pyrimidine pathways could represent sizeable changes in nucleotide production due to reduced transcriptional activity. Similarly, the spliceosome may be down-regulated due to a reduction in overall mRNA production as a result of altered transcription.

(a)

(b)
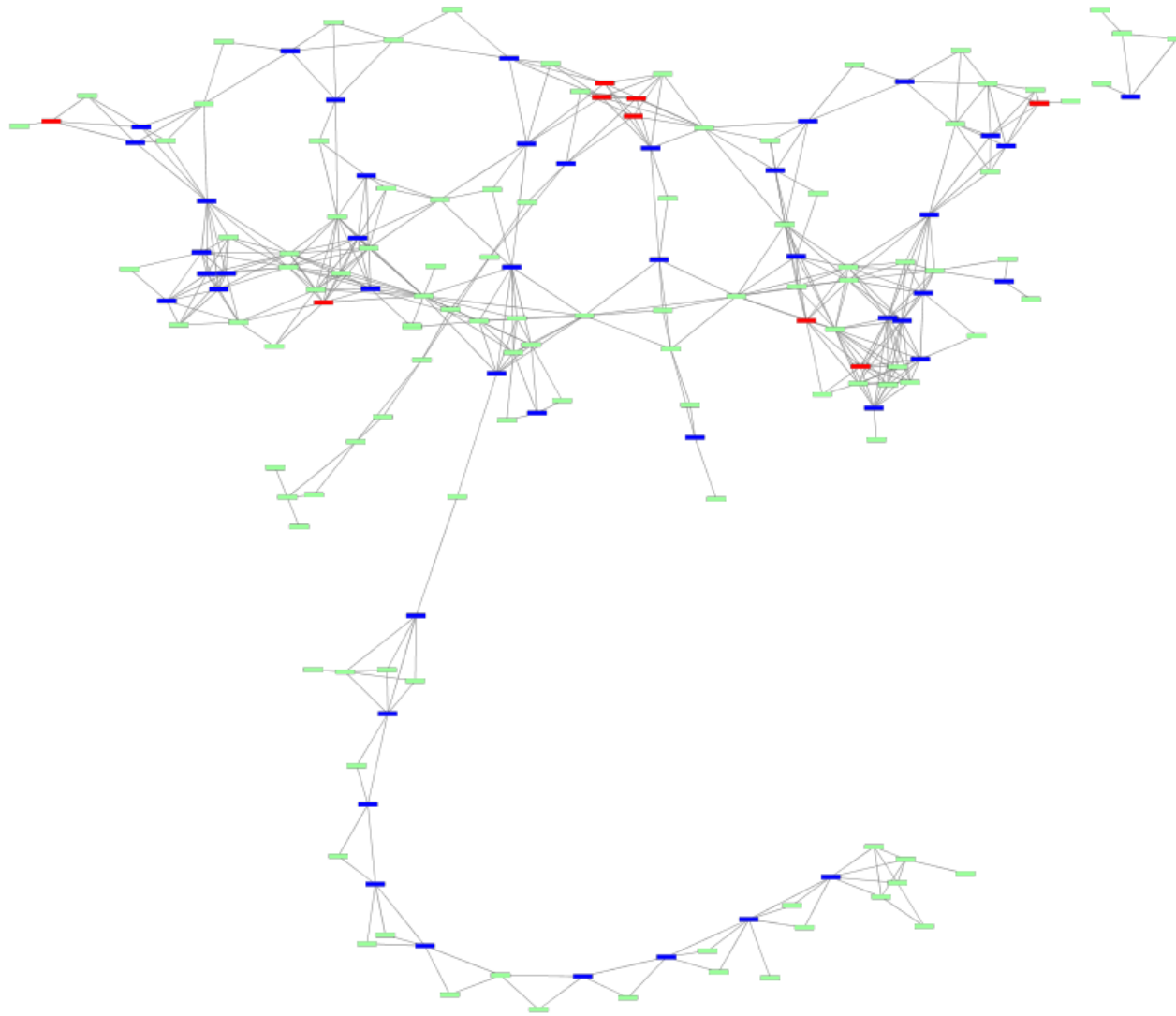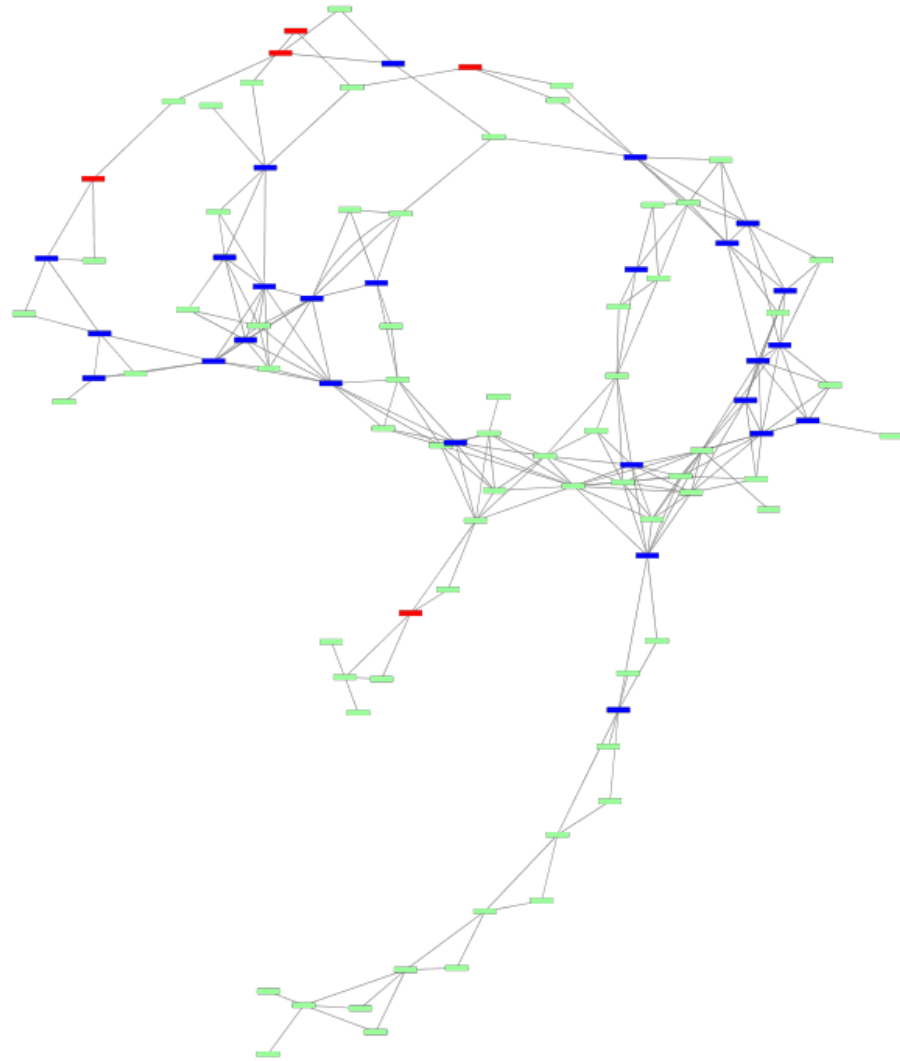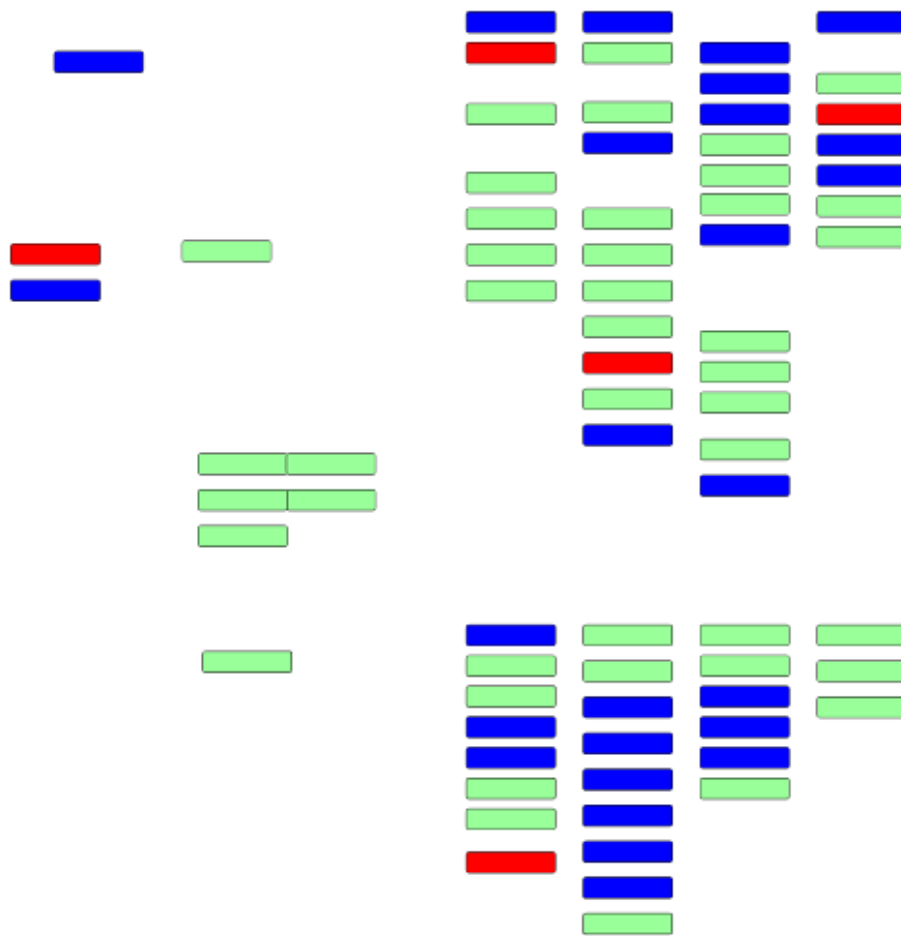
Figure 6.8: Purine and pyrimidine metabolic pathways. KEGG pathways for
An. gambiae: (a) purine metabolism (aga00230), (b) pyrimidine metabolism
(aga00240) and (c) spliceosome (aga03040) with significant transcripts mapped
from the *Maf-S-RNAi* array. In each case blue represents down-regulation in *Maf-
S* silenced mosquitoes, red up-regulation and green is non-differential
transcripts/compounds.

In addition to the two pathways and a biological structure described above, the *Maf-S-RNAi* array also shows many large changes over a broad range of pathways and structures, including oxidative phosphorylation, RNA transport, carbon metabolism, FoxO signalling and ubiquitin mediated proteolysis. The oxidative phosphorylation pathway subunits show very clear up-regulation along each complex in the chain, with several components of each complex showing altered expression levels after *Maf-S* knockdown. Arginine and Proline metabolism demonstrates both up- and down-regulation at different points in the pathway. The resultant effect on the pathway is the down-regulation of alanine production from CMP, L-arginisuccinate, L-citrulline and L-aspartate and the up regulation from ammonia, L-glutamate, L-glutamine and L-glutamate semialdehyde. The glycine, serine and threonine metabolism pathway demonstrates down-regulation of numerous transcripts in one clear subsection of the pathway. Here, the production of L-serine and glycine are heavily down-regulated, resulting in lower product for the entry points into purine, sphingolipid, cyanoamino acid, methane and sulphur metabolism pathways. FoxO signalling shows very clear changes to the terminal signalling, after integration of all signalling events by FoxO. Oxidative stress resistance and repair and muscle atrophy demonstrates up-regulation, whilst metabolism shows down-regulation; the two further branch points, cell cycle and autophagy show a mixture of fold change directionalities (Figure 6.9)

(a)

149
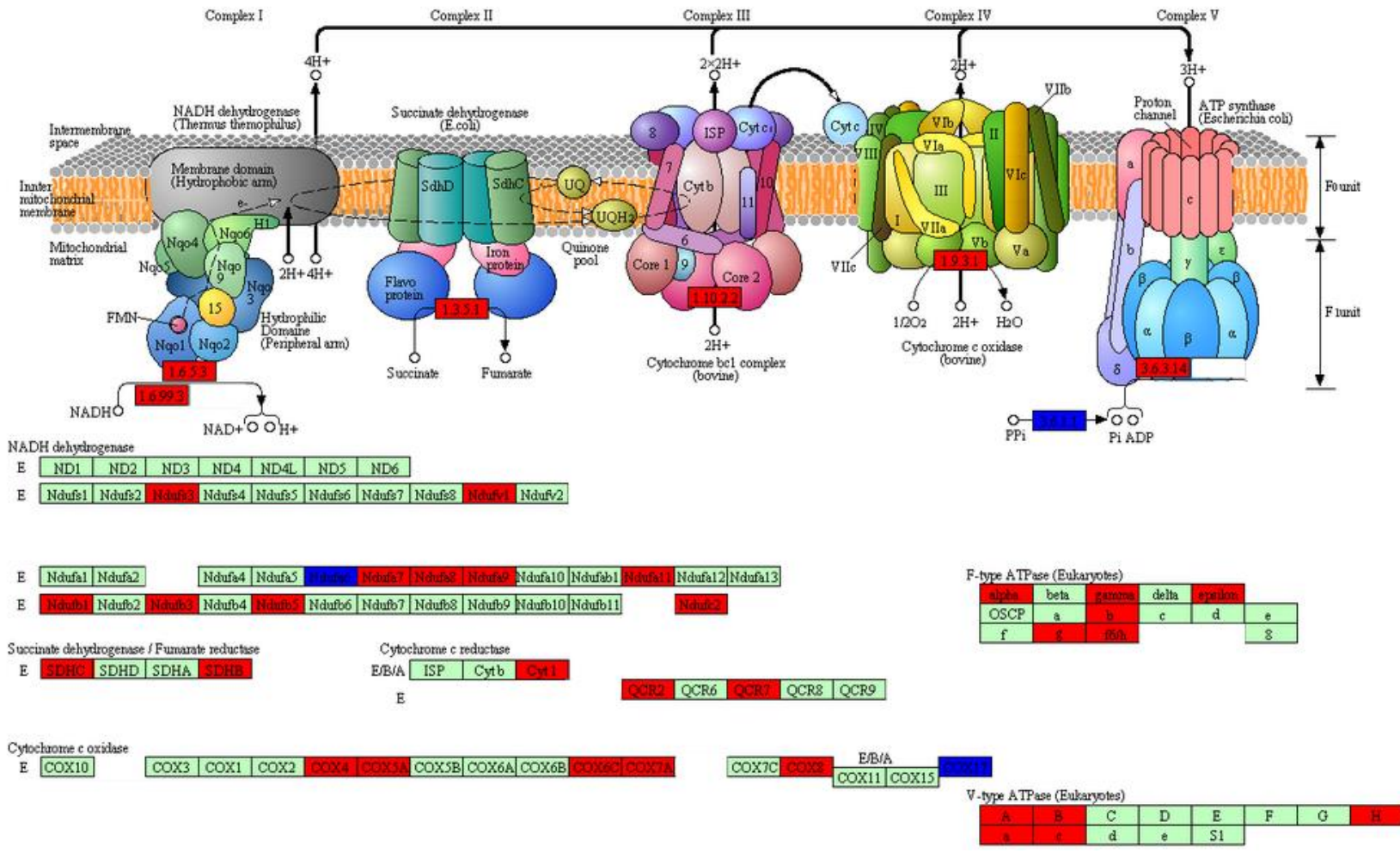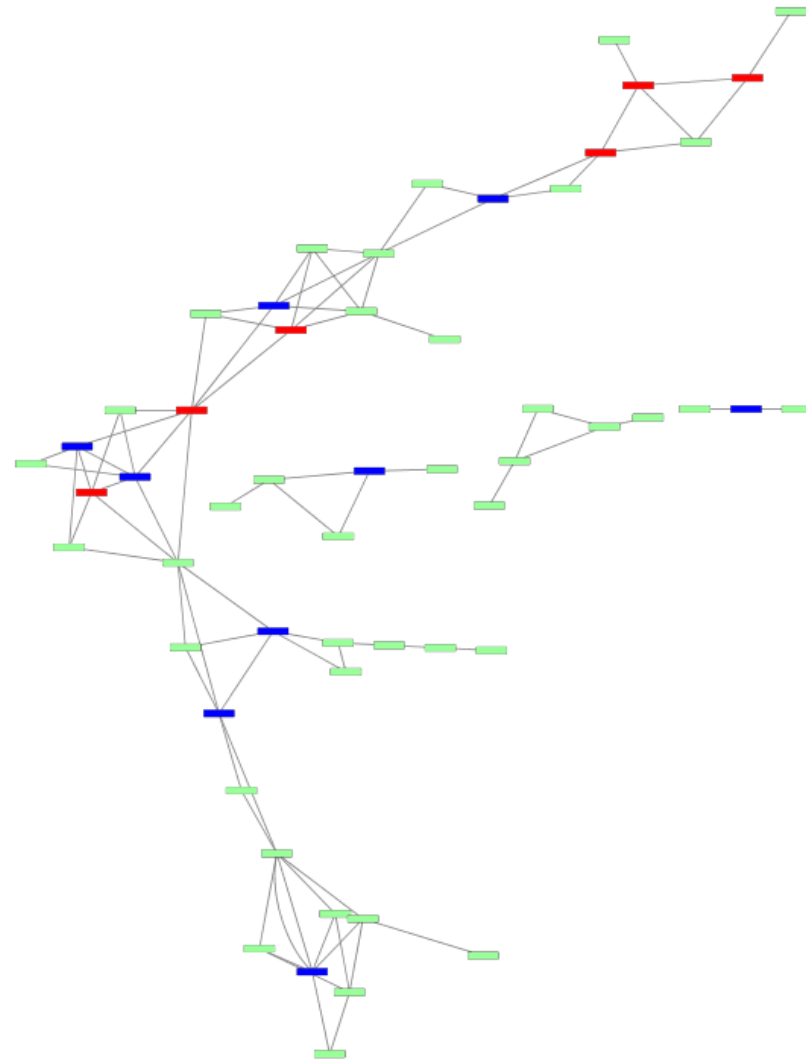
(c)
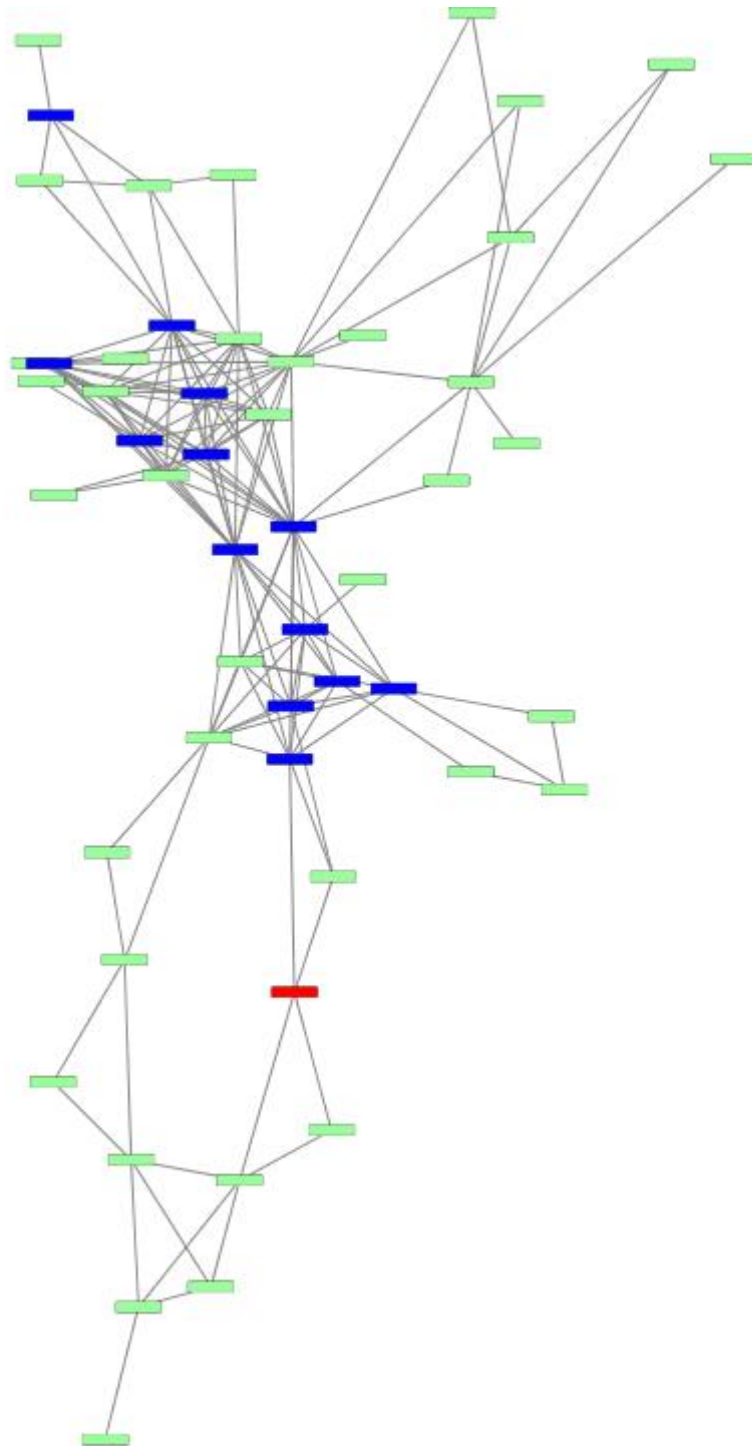
(d)
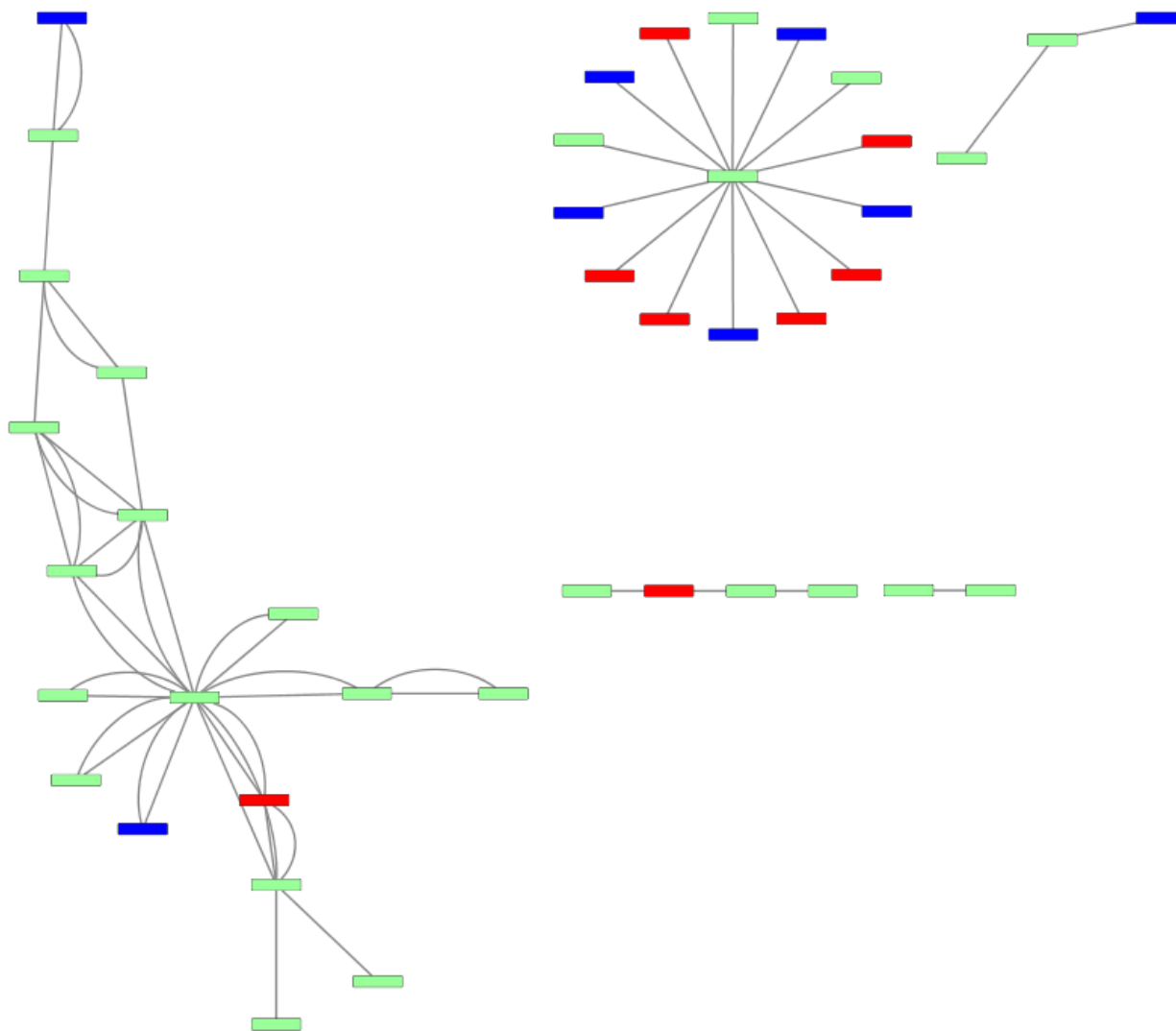
Figure 6.9: Pathways showing significant transcript changes in *Maf-S-RNAi*. (a) Oxidative phosphorylation pathway/subunits (aga00190) (b) Arginine and Proline Metabolism (aga00030) (c) Glycine, Serine and Threonine Metabolism (aga00260) and (d) FoxO Signalling (aga04068) from KEGG *Anopheles gambiae* repository; blue shows down-regulation of transcripts, red up-regulation and green shows no significant difference in transcript levels/compounds.

Discussion

Maf-S was explored as a putative candidate for the transcriptional regulation of xenobiotic response in *Anopheles* mosquitoes after being shown to be involved in the regulation of response to DDT, malathion and phenobarbital in *Drosophila* (Sykiotis & Bohmann, 2008; Misra *et al.*, 2011). The *cnc* and *Maf-S* genes appear to be conserved in *Anopheles* mosquitoes; however, the two putative homologs of *Drosophila Keap1* show low query coverage (<40%) but high identity (>70%) and so may not represent true homologs; however conservation of motifs and domains give weight to the annotation. Of the 27 datasets analysed comparing resistant and susceptible populations, both *cnc* and *Maf-S* were significantly differentially regulated in more studies that expected by chance. In some instances *Maf-S* and *cnc* demonstrated different directionalities of fold change, which may indicate that the pathway is not important in resistance to those populations. A confounding factor in the use of these studies is that they are left for 48 hours post exposure, which may remove evidence for fast transcriptional responses and hence give unexpected transcript enrichments. Both qPCR and bioassay results indicate that *Maf-S* itself is not important in resistance in Tiassalé mosquitoes; these data may be due to a lack of involvement in deltamethrin resistance (Figure 6.10) and more targeted involvement to organochlorides and organophosphates, as in *Drosophila* (Misra *et al.*, 2011; Misra *et al.*, 2013). To determine any role in insecticide resistance, both DDT and Malathion resistance should be explored in *Anopheles*.

Despite initial evidence pointing to a lack of involvement of *Maf-S* in response to insecticides, *Maf-S*-RNAi knockdown shows evidence of involvement in key detoxification transcripts, including *CYP6M2*, previously shown to metabolise insecticides (Djègbè *et al.*, 2014). Further to this, 64 detoxification transcripts demonstrated significant differential expression on *Maf-S*-RNAi arrays. Included in these transcripts are representative members of the CYP6Z and P family in addition to *GSTD1*, all implicated in insecticide metabolism (Müller *et al.*, 2008b; Low *et al.*, 2010; Chandor-Proust *et al.*, 2013). Extensive motif searching in 1000bp regions upstream of significant transcripts revealed the presence of putative antioxidant response elements (Sykiotis & Bohmann, 2008) in these regions, however, it has been difficult to determine significance of these motifs due to the length and degeneracy of the sequence. KEGG pathways also demonstrated significant changes in a number of key metabolic pathways after *Maf-S* knockdown. Included in these pathways is pyrimidine and purine metabolism, which may be due to large-scale changes in transcription after knockdown of a transcription factor. These pathway searches do not include any detoxification pathways relevant for insects and so limit the presence of *a priori* detoxification candidates in the searches.

In conclusion, there is evidence of a role in the regulation of important detoxification transcripts by *Maf-S* and hence the *Maf-S-cnc-Keap1* pathway in *Anopheles.* However, the pathway does not seem to be important in the response to deltamethrin in Tiassalé mosquitoes. Further supporting evidence for a role of this pathway in xenobiotic response may be identified

through use of organochloride and organophosphate insecticides in bioassays and use of populations that demonstrate strong resistance to these insecticides.
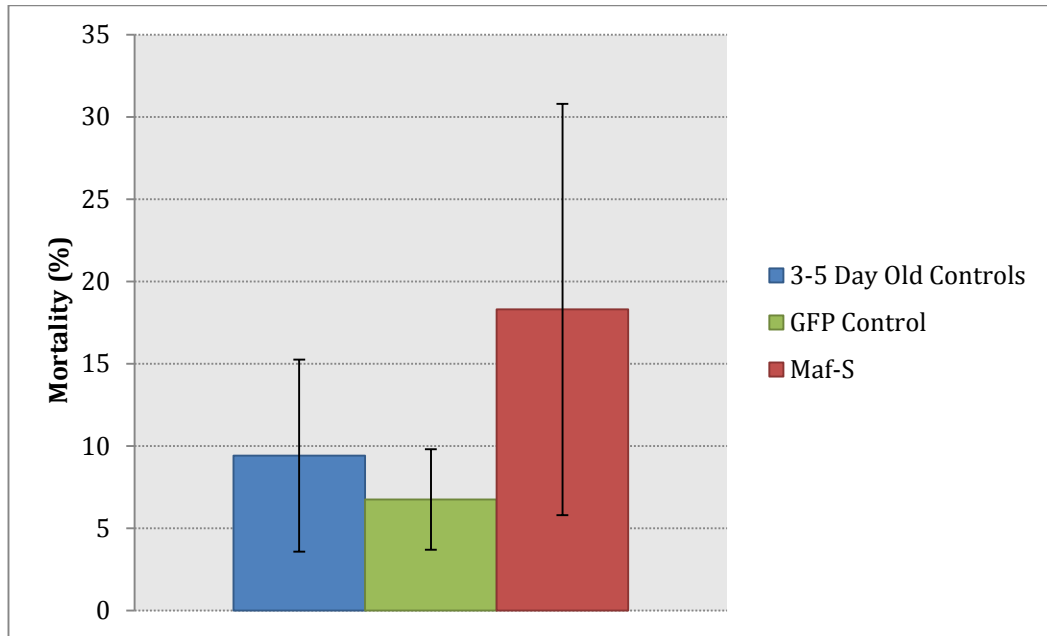
Figure 6.10: Mortality after exposure to deltamethrin. Percentage mortality (y) recorded for Tiassalé mosquitoes 24 hours after WHO tube assay exposure for one hour deltamethrin exposure, 72 hours post-injection. Both *GFP* (green) and uninjected (blue) controls were compared to *Maf-S-RNAi* mosquitoes (purple).

7. TRANSCRIPTIONAL REGULATION OF XENOBIOTIC RESPONSE: *METHOPRENE TOLERANT*

Unpublished

V A Ingham, P Pignatelli, S C Wagstaff, J D Moore and H Ranson.

Introduction

Juvenile hormone (JH) mimics are used as insecticides by acting as insect growth regulators. They are arthropod specific and, unlike conventional insecticides, non-neurotoxic, unless exposed in extremely high dose (Miyamoto *et al.*, 1993). JH mimics act as a larval control by causing errant transcription of genes at the last instar larvae, resulting in death before the pupae stage (Invest & Lucas, 2008). Additionally, JH mimics cause sterilisation of adult mosquitoes by affecting ovarian development and reducing the number of eggs produced (Invest & Lucas, 2008; Koama *et al.*, 2015). The two most commercially successful JH mimic chemicals are methoprene (Henrick *et al.*, 1973) and pyriproxyfen (Hirano *et al.*, 1998), which is now being examined for use on insecticide treated nets (Ohashi *et al.*, 2012; Ngufor *et al.*, 2014; Kawada *et al.*, 2014).

Studies have identified the transcription factor *Methoprene tolerant* (*Met*) as the JH receptor (Ashok *et al.*, 1998). Experiments in *Drosophila* have shown that *Met* null mutants have high levels of resistance to the effects of JH and the mimics methoprene (Wilson & Ashok, 1998) and pyriproxyfen (Charles *et al.*, 2011). *Met* is part of the basic helix-loop-helix-PAS transcription factor family, which has been shown to require the formation of a heterodimer in the presence of JH to bind to and transcriptionally regulate DNA (Li *et al.*, 2011).

The increasing use of JH mimics in malaria control programmes, leads to concerns over the development of resistance. The central role of *Met* in controlling the response to these JH mimics makes it a good candidate for further study. In this study the transcriptomic response to *Met* knockdown has been examined by comparing RNAi knockdown RNA to GFP injected control RNA on a microarray platform. In addition the study aimed to verify any consensus motifs, which have been identified in both *Drosophila* and *Aedes aegypti*, with both showing binding upstream of transcripts directly regulated by *Met* (Bryne *et al.*, 2008; Li *et al.*, 2011; Shin *et al.*, 2012; Zou *et al.*, 2013)*.* Changes to enrichments of different GO terms and pathways have been explored, in addition to insecticide phenotype and changes in expression of detoxification gene family members. These changes have then been compared with a second transcription factor knockdown, *Maf-S* [Chapter 6], which has been implicated in xenobiotic response (Misra *et al.*, 2011; Misra *et al.*, 2013). The results demonstrate changes to pathways, enrichments and detoxification genes as well as a sizeable overlap of transcriptome changes between the two transcription factors.

Aims
1. Identify potential transcripts that are regulated by *Methoprene Tolerant* (*Met*).

2. Explore putative pathways and enriched transcripts that may be transcriptionally regulated by *Met.*

3. Determine whether there is a role for *Met* in the response (and hence resistance) to insecticides.

4. Identify the canonical binding motif for *Met* in *Anopheles* mosquitoes.

5. Compare the overall effect of transcription factor knockdown using both *Met* and *Maf-S* RNAi arrays.

Results

Transcriptomic changes after RNAi knockdown

RNA was extracted from *Met* dsRNA injected, uninjected and GFP injected six-eight day old female mosquitoes. Each biological replicate contained eight-12 insecticide resistant (JH susceptible) Tiassalé mosquitoes, originating in Côte D'Ivoire, from the major malaria vector *An. gambiae*. Knockdown levels were confirmed using qPCR comparing *Met-RNAi* cDNA to both GFP and uninjected controls. Two separate RNAi constructs were used (Figure 7.2). The knock-down levels compared to uninjected and GFP-injected controls were 0.0414 (±0.0454) and 0.0241 (±0.0264) for *Met* construct 2 versus GFP injected and uninjected controls respectively, with no clear knockdown for construct 1. These data are indicative of knockdown levels of over 85%. *Met* construct 2 was therefore used in all further experiments. *Met-RNAi* was then compared with GFP injected control (Figure 7.1) on a microarray platform.

Figure 7.1: Microarray design. Designs used for microarrays comparing

Met knockdown RNA with GFP injected control RNA. Green boxes represent

RNA tagged with *cy3* and orange those tagged with *cy5*. Dye swaps are

shown on the second row.

Figure 7.2: *Met* construct mapping. FASTA cDNA sequence of

AGAP006662-RA, *Met*. Purple shows the position of primers for *Met*
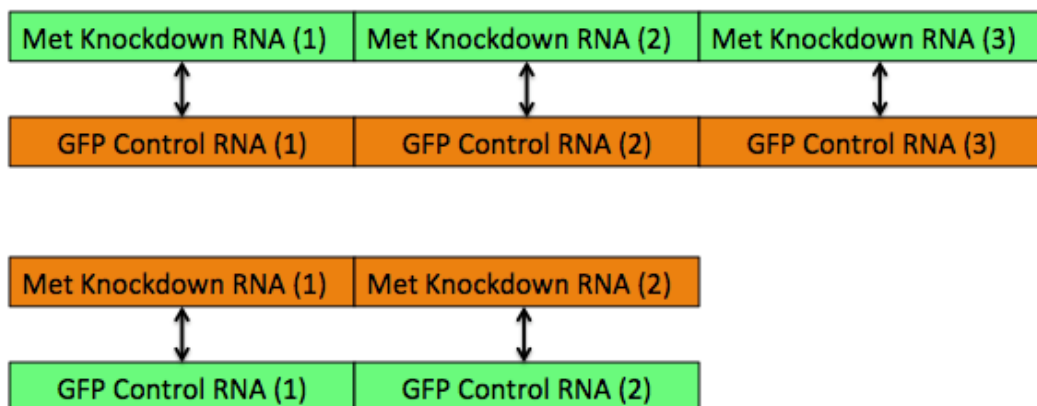
construct 1; yellow for construct 2, red shows the microarray probe location and green exon start sites.

Transcripts differentially expressed in the *Met-RNAi* arrays compared to GFP injected controls were determined by limma analysis (Smyth, 2004), using a multiple test correction cut-off of p ≤ 0.05. In total 1349 transcripts were differentially expressed, 671 of which were up-regulated and 678 down-regulated. Up-regulated transcripts were significantly enriched for Gene Ontology (GO) terms related to system processes and various binding activities; these differ from the down-regulated transcripts which are enriched for oxidation reduction processes, structural assembly, heme binding and transferase activity (Table 7.1).

GO ID	p-value	Transcript List
Synapse	1.90E-02	Up
Regulation of systems process	2.70E-03	Up
Actin binding	5.40E-04	Up
Cytoskeletal protein binding	1.30E-03	Up
Nucleotide binding	3.20E-03	Up
mRNA binding	1.80E-02	Up
oxidation reduction	6.30E-06	Down
Arginine metabolic process	5.20E-03	Down
Chromatin assembly	1.30E-02	Down
Nucleosome assembly	1.30E-02	Down
Macromolecular complex assembly	1.30E-02	Down
Protein-DNA complex assembly	1.50E-02	Down
Cellular macromolecular complex assembly	1.90E-02	Down
Nucleosome organization	1.90E-02	Down
Macromolecular complex subunit organization	2.70E-02	Down
Glutamine family amino acid metabolic process	2.80E-02	Down
Translation	4.20E-02	Down
Steroid biosynthetic process	4.20E-02	Down
Cellular macromolecular complex subunit organization	4.40E-02	Down

Lipid localization	4.40E-02	Down
Lipid transport	4.40E-02	Down
Chromatin assembly or disassembly	4.40E-02	Down
RNA polymerase II transcription factor activity	9.80E-02	Down
RNA polymerase activity	8.50E-02	Down
DNA-directed RNA polymerase activity	8.50E-02	Down
Inositol or phosphatidylinositol phosphatase activity	6.30E-02	Down
Oxygen transporter activity	6.30E-02	Down
Oxidoreductase activity	4.20E-02	Down
Nucleotidyltransferase activity	2.90E-02	Down
Lipid transporter activity	2.30E-02	Down
Ribosome	1.10E-04	Down
Ribonucleoprotein complex	3.20E-04	Down
Nucleosome	3.30E-03	Down
Mitochondrion	4.00E-03	Down
Protein-DNA complex	8.40E-03	Down
Non-membrane-bounded organelle	8.40E-03	Down
Intracellular non-membrane-bounded organelle	1.40E-02	Down
Membrane-enclosed lumen	4.00E-02	Down
Prefoldin complex	4.00E-02	Down
Chromatin	4.60E-02	Down

Table 7.1: Significant Gene Ontology terms for transcripts DE in *Met* knockdown arrays. GO term ID, p-value and direction of differential expression after *Met* knockdown, up (blue) and down (green).

Pathway changes after Met knockdown

KEGG pathway (Kanehisa & Goto, 2000) was used to search transcripts that were significantly differentially expressed after knockdown of *Met*. KEGG pathway maps *Anopheles* genes onto pre-existing human pathways and so is not always complete and lacks species and insect specific pathways. To overcome redundant pathways being analysed, CytoKEGG was used, which loads only pathways and transcripts relevant for the selected organism, removing redundancy (https://code.google.com/p/cytokegg/).

Both purine and pyrimidine metabolism demonstrate changes throughout

the pathway, which may be indicative of changes in levels of transcription

after *Met* knockdown. *Met-RNAi* also demonstrates small changes (<30%

transcripts) in a number of pathways and structures including: alanine,

glutamate and aspartate metabolism, with 20 transcripts showing

differential regulation; endocytosis with 16 transcripts; and RNA transport

with 14 transcripts. In addition, *Met-RNAi* also alters the expression of

transcripts associated with the structures spliceosome and phagosome, with

down-regulation to starch metabolism and peroxisome targeting sequence

transcripts related to amino acid metabolism (Figure 7.3). The majority of

these changes do not show clear directionality across an entire pathway,

instead showing localised changes within a subsection of a pathway; this

may be due to the relatively small number of differential transcripts in the

Met-RNAi. It may also be indicative of changes to pathways not shown in

KEGG, as *Met* is known to be involved in both sexual development of

mosquitoes and larval development (Wilson *et al.*, 2006; Charles *et al.*, 2011;

Zou *et al.*, 2013), two key areas missing in the KEGG database.

(a)
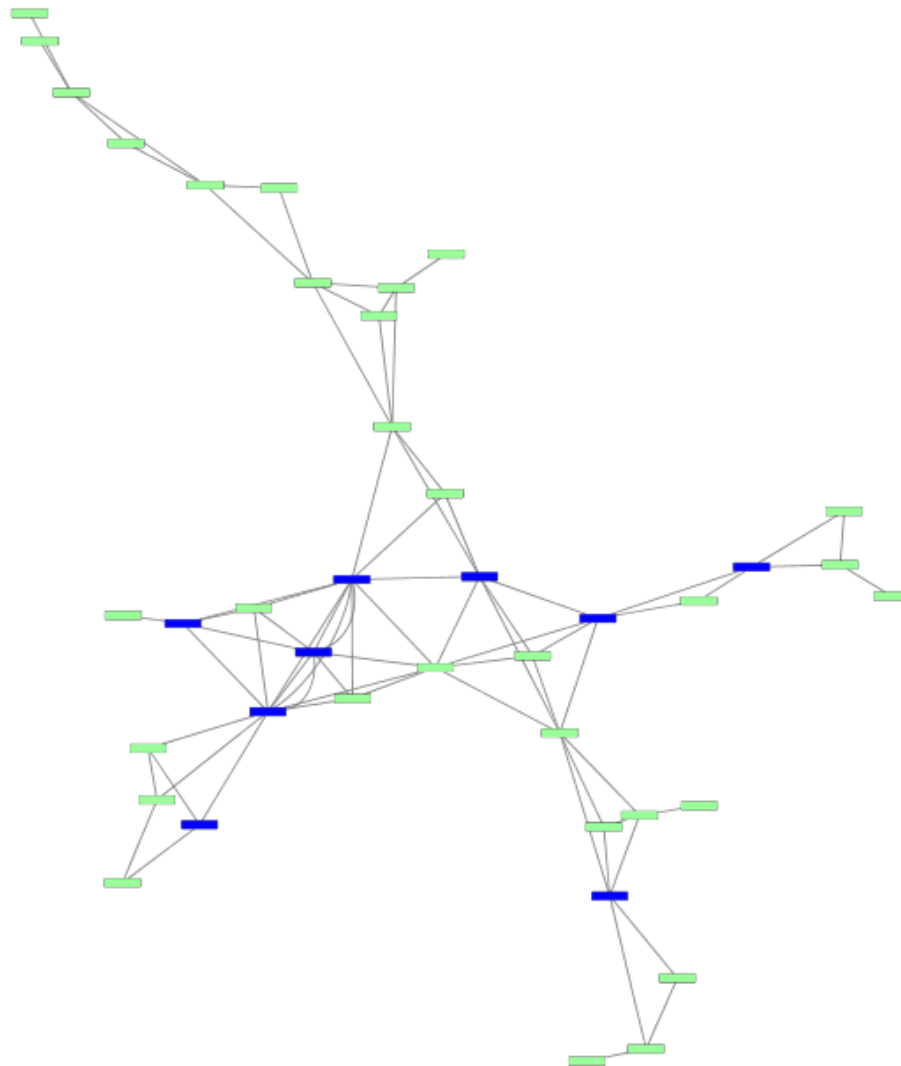
(b)

amino acid metabolism (PTS1 type)

AGT	DAO	DDO	IDH
		HMGCL	HAO
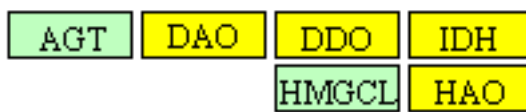
Figure 7.3: Changes due to *Met-RNAi*. KEGG pathway for *Anopheles gambiae* (a) starch and sucrose metabolism (aga00500) and (b) subsection of peroxisome structure (aga04146) specifically peroxisome targeting signal (PTS1) type proteins. Results are shown for differentially expressed transcripts in *Met-RNAi*; blue nodes indicate down-regulation of transcripts and yellow up-regulation. Green nodes show no differential expression/compounds.

Phenotyping of Met-RNAi

Previous work has shown that *Met* is differentially regulated across 27 microarray experiments comparing resistant and susceptible *Anopheles* species from disparate geographical localities [Chapter 4]. Given the nature of the microarray data used in this study, *Met* knock-downs were not explored for effect due to JH mimics and instead explored for a phenotype to public health insecticides. In order to determine whether *Met* played any role in resistance to common insecticides, phenotyping was carried out using WHO tube bioassays, using both 0.05% deltamethrin and 0.1% bendiocarb hour/half-hour long exposures on at least 75 female six-eight day old Tiassalé mosquitoes, 72 hours post injection (Figure 7.4). Both bendiocarb and deltamethrin exposures showed significant increases in mortality in the *Met-RNAi* mosquitoes compared to both uninjected and GFP injected controls, indicative of a role for *Met* in resistance to these insecticides.
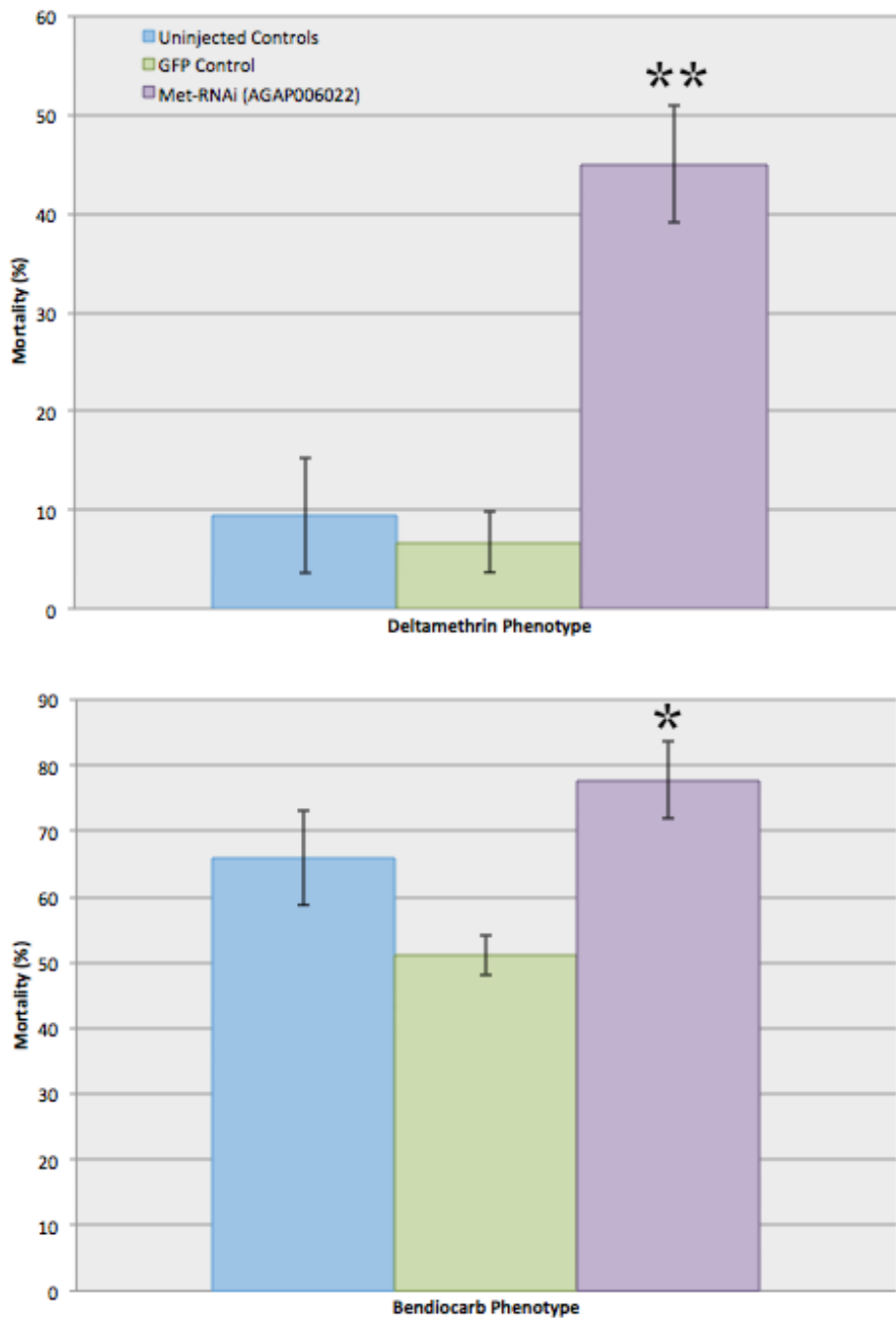
Figure 7.4: Mortality associated with exposure to insecticides.

Percentage mortality (y) recorded for Tiassalé mosquitoes 24 hours after
WHO tube assay exposure for (a) one hour deltamethrin exposure and (b)
30 minute bendiocarb exposure. Both *GFP* (green) and uninjected (blue)
controls were compared to *Met-RNAi* mosquitoes (purple). Significance is

shown as p ≤ 0.05 (*) and p ≤ 0.001 (**) for both GFP and uninjected

controls.

To explore the link between *Met-RNAi* and resistance to deltamethrin and

bendiocarb, detoxification gene enrichment was explored in significant

transcript lists. In total six detoxification transcripts are present in the up-

regulated transcript list, showing no significant enrichment of these gene

families. Conversely, the down-regulated list is significantly enriched for

detoxification family members ($p = 9.73e-7$), with 27 representative

transcripts (Table 7.2), of these 15 are cytochrome p450s and the list

includes the insecticide metaboliser *CYP6M2* (Stevenson *et al.*, 2011;

Mitchell *et al.*, 2012), consistent with *Met* having an activating role in

transcription of these detoxification genes.

Gene ID	Gene Name	Fold Change
AGAP011507-RA	COE13O	0.443
AGAP006724-RA	COEAE3G	0.714
AGAP005372-RA	COEBE3C	0.678
AGAP005658-RA	CYP15B1	0.776
AGAP005992-RA	CYP302A1	0.580
AGAP008682-RA	CYP307B1	0.769
AGAP002429-RA	CYP314A1	0.796
AGAP009246-RA	CYP4C27	0.579
AGAP009241-RA	CYP4C36	0.706
AGAP002418-RA	CYP4D16	0.731
AGAP000877-RA	CYP4G17	0.646

AGAP001864-RA	CYP4H15	0.609
AGAP002416-RA	CYP4K2	0.546
AGAP008212-RA	CYP6M2	0.375
AGAP008213-RA	CYP6M3	0.332
AGAP008208-RA	CYP6Y1	0.559
AGAP008207-RA	CYP6Y2	0.584
AGAP012294-RA	CYP9L2	0.642
AGAP004172-RA	GSTD9	0.717
AGAP009195-RA	GSTE1	0.761
AGAP009197-RA	GSTE3	0.608
AGAP009191-RA	GSTE6	0.794
AGAP000165-RA	GSTMS1	0.779
AGAP008404-RA	UGT301C1	0.645
AGAP007028-RA	UGT302J1	0.652
AGAP005750-RA	UGT308B2	0.546
AGAP007374-RA	UGT49A3	0.526

Table 7.2: Detoxification family members down-regulated in *Met-RNAi*.

Gene ID, detoxification family name and associated fold change for

detoxification genes present in the significantly down-regulated transcript

list.

Met Xenobiotic Response: Evidence
To further examine the evidence of *Met*'s involvement in transcriptional

control of detoxification families, a correlation network was produced for

Met, with available microarray data comparing resistant and susceptible

mosquitos [Chapter 4]. There were no obvious enrichments of transcripts

related to detoxification. However, from the 176 transcripts showing co-

correlation with *Met*, GO term enrichment for translational activity,

response to chemical stimulus and metabolic processes was observed, in

addition to changes in the Hedgehog signalling pathway, concurrent with

the available literature (Li *et al.*, 2007; Charles *et al.*, 2011; Villarreal *et al.*,

2015).

Met Xenobiotic Response: Yeast Two-Hybrid

Further exploration of putative *Met* interactors was possible due to the

availability of *Drosophila* yeast two-hybrid data (Figure 7.5). Yeast two-

hybrid data identifies physical interactions between proteins, in contrast to

the genetic interactions and so does not identify transcriptional regulation

and hence, allows exploration of *Met* effects using a different kind of

network. However, the available data demonstrates some relationship of

Met to detoxification candidate families, in agreement with the data from *Ae.*

aegyptii microarrays (Zou *et al.*, 2013) and the microarray data presented

here. Furthermore, these data demonstrate a direct interaction of *Met* with

a second transcription factor, *Myb* (AGAP008160-RA). *Myb* has no prior

implication in detoxification of xenobiotics; instead being involved in key

cell cycle control points (Bonke *et al.*, 2013; DeBruhl *et al.*, 2013; Moutinho-

Pereira *et al.*, 2013).
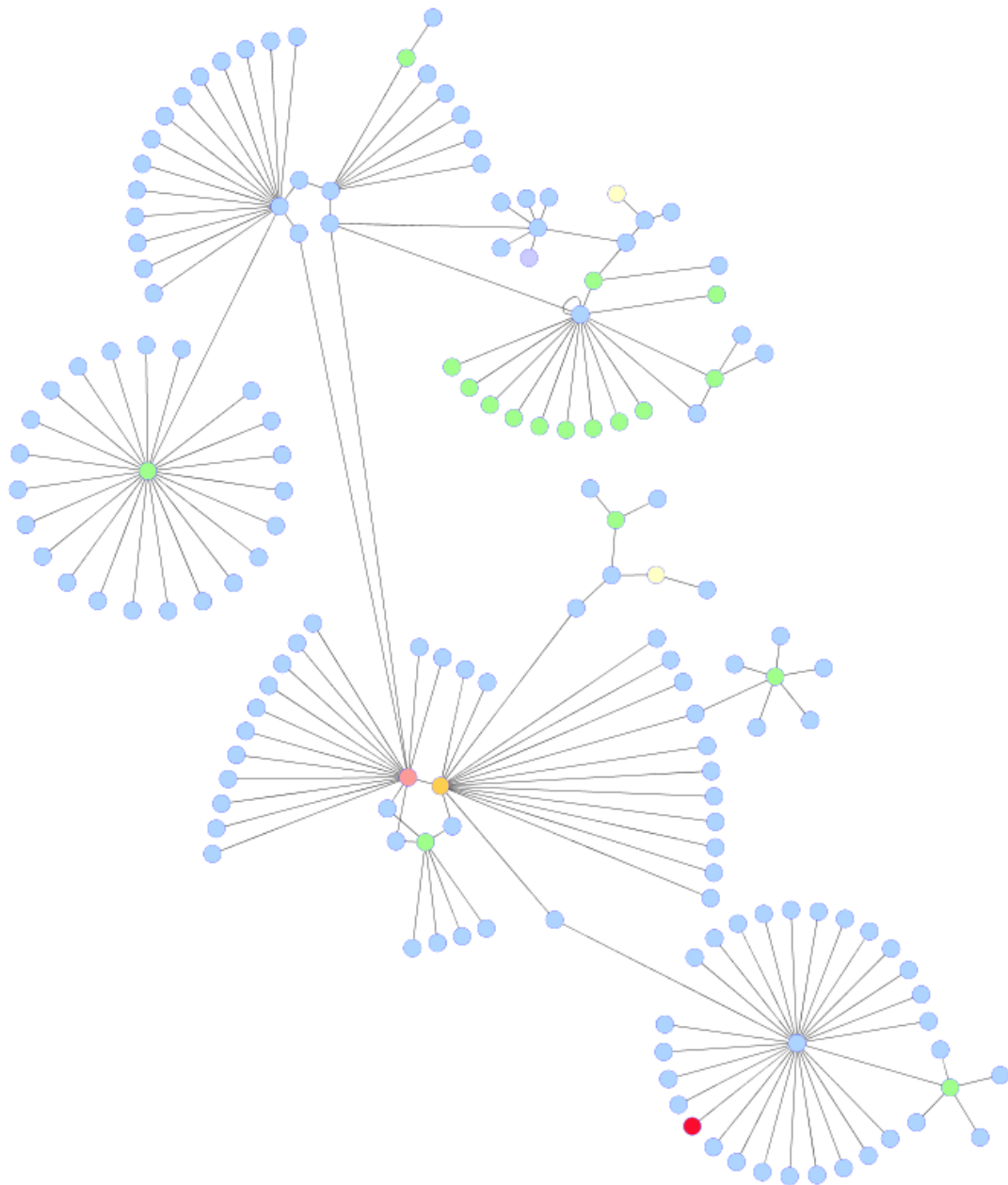
Figure 7.5: *Drosophila* **yeast-two hybrid interactome for** *Met*.

Met (orange) and *Myb* (pink) are shown as direct protein-protein interactors

in the *Drosophila* interactome produced from DroID (Murali *et al.*, 2011)

within Cytoscape (Shannon *et al.*, 2003). Cytochrome p450s, GSTs, a

carboxylesterase and a cuticular protein are highlighted in green, yellow,

purple and red respectively. Heterodimerisation with the transcription

factor *Myb* was observed as expected (Li *et al.*, 2011).

Transcription factor knockdowns: general or specific response?
Due to the generalist transcription and translation pathways showing

changes after knockdown of *Met,* a hypothesis was formed that knockdown

of a transcription factor has both generalist and specific effects. To test this

hypothesis results from RNAi with a second transcription factor, *Maf-S* were

compared. Previous publications have shown that the transcription factor

Maf-S is implicated in resistance to DDT and malathion, through constitutive

activation of the *Maf-S-cnc-keap1* pathway in *Drosophila* (Misra *et al.*, 2011;

Misra *et al.*, 2013). *Maf-S-RNAi* microarray data of the same experimental

design as the current *Met-RNAi* is available [Chapter 6] and can be used to

compare the transcriptomic response after knockdown of different

transcription factors. 3401 and 1349 respective transcripts for each of *Maf-*

S-RNAi and *Met-RNAi* were significantly differentially expressed in each

experiment. In the respective arrays a similar proportion of transcripts were

up- and down-regulated compared to *GFP* controls (Figure 7.8). Although a

large number of transcripts were unique to each RNAi subset, there was also

a sizeable overlap in transcripts between the two arrays, which may be

accounted for due to large-scale transcriptional alterations in both the

knockdowns caused by changes to transcription through transcription

factor knockdown. The majority of the transcripts demonstrated relatively

small changes in expression (fold change of 0.6-2), with only one and three

transcripts for *Met-RNAi* and *Maf-S-RNAi* respectively showing a fold change greater than five-fold.
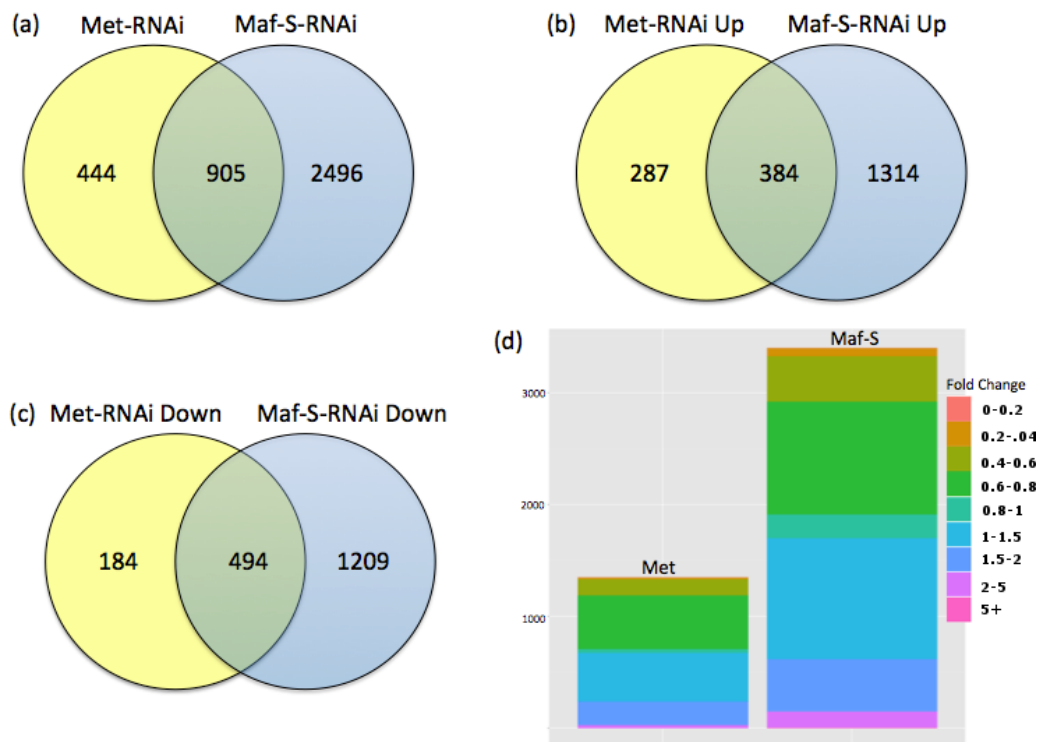
Figure 7.6: Summary data for both *Met-RNAi* and *Maf-S-RNAi* arrays. (a) Significant probes (b) Significantly up-regulated probes (c) Significantly down-regulated probes between the *Met* knockdown and *Maf-S* knockdown and (d) frequency (y) of each range of significant fold change for both *Met* and *Maf-S* knockdown arrays (x).

Table 7.3 shows the top five GO enrichment and InterPro terms for each array. *Met-RNAi* shows multiple enrichments for catalytic activity, including oxidoreductase, hydrolase and phosphotransferase activity in relation to the cytoplasm and non-membrane bound organelles. InterPro enrichments show calponin-like actin binding, nitrilase, histone and cytochrome p450 domains, in agreement with the GO term enrichments. Similarly, *Maf-S-RNAi*

shows enrichments for detoxification related activity, including oxidation-reduction process, antioxidant activity, as well as multiple enrichment demonstrating changes to RNA and DNA-related activity. The cellular component terms differ from those for *Met-RNAi*, with 18 terms demonstrating different significant locations. The significant InterPro terms further agree with the GO terms, with cytochrome p450 domain, EF-hand, aldehyde dehydrogenase, heat shock protein, GST and peptidase domains being enriched.

Term	*Maf-S-RNAi*	*Met-RNAi*
Molecular Function	1. Nucleotidyltransferase activity. 2. DNA-directed RNA polymerase activity. 3. RNA polymerase activity. 4. Aminoacyl-tRNA ligase activity. 5. Ligase activity, forming C-O bonds.	1. Phosphotransferase activity. 2. Oxioreductase activity. 3. Disulphide oxioreductase activity. 4. Tertapyrrole binding. 5. Heme binding.
Biological Process	1. Oxidation-Reduction. 2. tRNA metabolic process. 3. Nucleosome assembly. 4. Chromatin assembly. 5. Protein-DNA complex assembly.	1. Chromatin assembly. 2. Nucleosome assembly. 3. Protein-DNA complex assembly. 4. Oxidation-Reduction. 5. Regulation of system process
Cellular Component	1. Nucleosome. 2. Chromosome. 3. Mitochondrion. 4. Protein-DNA complex. 5. Cytosol.	1. Non-membrane-bound organelle. 2. Intracellular non-membrane-bound organelle. 3. Protein-DNA complex. 4. Nucleosome. 5. Prefoldin complex.
InterPro	1. Calcium-binding EF-hand. 2. EF-hand.	1. Calponin-like actin-binding.

3. <mark>Histone-Fold.</mark>		2. Nitrilase/cyanide hydratase and
4. <mark>Histone Core.</mark>		apolipoprotein N-acyltransferase.
5. Alpha crystallin/Heat shock protein.		3. <mark>Histone-Fold.</mark>
		4. <mark>Histone Core.</mark>
		5. Actinin-type, actin-binding.

Table 7.3: Enrichment terms for *Maf-S-RNAi* and *Met-RNAi*. Top five InterPro terms and GO terms for the three respective categories for each experiment. Terms common to both arrays are highlighted in yellow.

As both transcription factors have been linked to resistance to insecticides and demonstrate significant enrichment of detoxification transcripts, the overlap in differential detoxification transcripts was explored (Wilson & Ashok, 1998; Charles *et al.*, 2011; Misra *et al.*, 2011; Misra *et al.*, 2013). Of 64 detoxification transcripts significant in *Maf-S* knockdown and 65 in *Met* knockdown, 20 were significant on both arrays and 18 demonstrated the same fold change directionality (Table 7.4). Importantly *CYP6M2,* a key insecticide metaboliser (Djègbè *et al.*, 2014), is present in both significant transcript lists from each knockdown, appearing to be activated by both *Maf-S* and *Met* in a similar way. Just two transcripts appear to show transcriptional repression by the transcription factors, with activation being far more frequent.

Transcript ID	Gene Name	*Maf-S* Fold Change	*Met* Fold Change	Same Directionality
AGAP002090-RA	COE '16738'	1.95	1.52	TRUE
AGAP005837-RA	COEJHE5E	0.81	1.26	FALSE
AGAP011507-RA	COE13O	0.56	0.44	TRUE
AGAP005372-RA	COEBE3C	0.67	0.68	TRUE

AGAP000877-RA	CYP4G17	0.75	0.65	TRUE
AGAP002416-RA	CYP4K2	0.68	0.55	TRUE
AGAP002418-RA	CYP4D16	1.57	0.73	FALSE
AGAP002429-RA	CYP314A1	0.84	0.80	TRUE
AGAP005992-RA	CYP302A1	0.64	0.58	TRUE
AGAP008207-RA	CYP6Y2	0.78	0.58	TRUE
AGAP008208-RA	CYP6Y1	0.75	0.56	TRUE
AGAP008212-RA	CYP6M2	0.41	0.38	TRUE
AGAP009246-RA	CYP4C27	0.60	0.58	TRUE
AGAP010961-RA	CYP6AK1	1.51	1.41	TRUE
AGAP012294-RA	CYP9L2	0.67	0.64	TRUE
AGAP005750-RA	UGT308B2	0.66	0.55	TRUE
AGAP007374-RA	UGT49A3	0.51	0.53	TRUE
AGAP009191-RA	GSTE6	0.72	0.79	TRUE
AGAP004172-RA	GSTD9	0.78	0.72	TRUE
AGAP008404-RA	UGT301C1	0.65	0.65	TRUE

Table 7.4: Detoxification transcripts significantly differential in both *Met* and *Maf-S* knockdowns. ID, gene name, fold changes for both *Maf-S* and *Met* knockdowns and TRUE/FALSE stating whether the directionality of fold change is consistent across the two experiments.

Discussion

Methoprene tolerant has been shown to play an important role in the resistance to juvenile hormone homologs methoprene (Wilson & Ashok, 1998) and pyriproxyfen (Charles *et al.*, 2011) in *Drosophila*. To determine whether these phenotypes are true for *Anopheles*, and hence, may impact malaria control efforts, the functional consequences of *Met* disruption were explored. In this study RNAi knock-downs of *Met* performed on *Anopheles* mosquitoes were compared to *GFP* injected control mosquitoes and the effects explored. *Met* appears to influence 1349 transcripts, with an almost exact 50-50 split in transcription activation and suppression. GO terms indicate that the transcripts showing transcriptional activation by *Met* may play some role in detoxification activity and numerous key metabolic

processes; this is further evidenced by the changes to metabolic pathways and important structures. Due to the known role of *Met* in juvenile hormone binding and reproductive development (Invest & Lucas, 2008; Koama *et al.*, 2015), it is likely that enrichment terms and pathways will fail to support these literature based roles due a lack of insect-specific terms and pathways. The lack of insect-specific pathways available in KEGG is further highlighted by a publication studying the transcriptomic changes in the fat body during the gonadotrophic cycle (48hrs-72hrs) related to *Met* in *Aedes aegypti* (Roy *et al.*, 2015)*,* which demonstrated 392 transcripts associated with reproductive development; a role which would be missed purely through enrichment searches.

Met insecticide resistance phenotyping was determined by WHO tube bioassays for both bendiocarb and deltamethrin, performed on adult females 72 hours post injection. Both results show a significant increase in mortality after the knockdown of *Met*, indicating a potential role in resistance to these insecticides. This hypothesis was further supported by the presence of 27 detoxification family members down-regulated after knockdown of *Met*. Included in these detoxification transcripts is *CYP6M2* a proven insecticide metaboliser (Djègbè *et al.*, 2014). The role of *Met* in detecting a response to chemical stimuli by activation of detoxification transcripts is further supported by the enrichments of co-correlated transcripts and protein-protein interactions in *Drosophila*. Despite this evidence, no role has been reported for *Met* in resistance to conventional public health insecticides; varying fold change directionalities in analysed

resistant versus susceptible microarrays further supports this. It may be that the change in mortality displayed after RNAi knockdown is a secondary affect due to the influence of *Met* on over 10% of the known detoxification transcripts.

Previous transcription factor knockdowns have been performed on *Maf-S* [Chapter 6], a transcription factor putatively involved in xenobiotic response (Misra *et al.*, 2011; Misra *et al.*, 2013). It was hypothesised that knocking down transcription factors would have an overlapping response, due the reduced levels of transcription and hence, translation. Comparing the transcripts differentially expressed in each of the knockdown arrays, it is clear there is a large overlap in differential transcripts, but there are also clear unique transcripts to each array. *Maf-S* knockdown influenced 2.5x the number of transcripts of *Met* knockdown. Despite the sizeable overlap in the differential transcripts, GO terms and InterPro clusters demonstrated differences in molecular functions, cellular components, biological functions and InterPro domains. Oxidation-reduction processes are significant in each, as are changes to chromatin and nucleosome assembly and enrichments in histone domains. The role in response to toxic compounds is furthered by the significant overlap in 20 detoxification transcripts, including *CYP6M2,* an important insecticide metaboliser (Djègbè *et al.*, 2014).

In conclusion, this study provides preliminary evidence that *Met* is involved in the insecticidal activity of bendiocarb and deltamethrin as well as supporting the role of this transcription factor in regulating important

metabolic processes. However, enrichments and pathway terms are missing crucial insect-specific terms, potentially removing evidence for known roles of *Met*, nevertheless this approach yielded valuable insights on global changes due to transcription factor knock-down. Global changes after transcription factor knockdown are evident in changes to histone proteins and general translation, in addition to potential roles of both transcription factors in oxidation-reduction processes. However, it is obvious that there are unique responses to each transcription factor, most likely related to their specific functions (Wilson & Ashok, 1998; Misra *et al.*, 2011; Charles *et al.*, 2011; Misra *et al.*, 2013).

8. GENERAL DISCUSSION

Resistance to insecticides in *Anopheles* species is caused, in part, by both
metabolic resistance and target site changes; however, this is far from the
full picture (Ranson *et al.*, 2011). The aim of this study was threefold: (i) to
expand understanding of the localisation of known detoxification
transcripts; (ii) to identify novel transcripts involved both directly in
resistance and in the transcriptional control of resistance through means of
a meta-analysis on existing microarray data and (ii) explore the data
generated to characterise the role of several transcripts.

In the first instance, microarrays were performed on dissected structures
(midgut, malpighian tubules, abdomen integument and remaining tissues)
for: (i) Resistant population Tiassalé structures compared to Tiassalé whole;
(ii) Susceptible population N'Gousso structures compared to N'Gousso
whole and (iii) Resistant Tiassalé structures compared to the corresponding
susceptible N'Gousso structures. These data were used to reveal the
localisation of all detoxification family members and to identify novel
candidates. Unsurprisingly and in agreement with available literature, the
malpighian tubules appeared to be a hotbed of detoxification activity (Brun
et al., 1996; Yang *et al.*, 2007; Chahine & O'Donnell, 2011) with the largest
range of fold change, along with the abdomen integument. Perhaps more
surprisingly there was clear segregation of cytochrome p450 subfamilies by
enrichment in different body parts; this was further demonstrated when
three members of the *CYP6Z* family and two members of the *CYP4G* family
were confirmed to be enriched in the malpighian tubules and abdomen

integuments respectively by qPCR. The presence of both *CYP4G16* and

CYP4G17 in the abdomen integument is in agreement with previous

publications on the *Drosophila* homolog *CYP4G1*, which is implicated in

cuticular hydrocarbon synthesis (Qiu *et al.*, 2012).

In addition to exploring the structural enrichment of detoxification

candidates, the data were also explored for novel insecticide resistance

candidates. 134 transcripts were identified as enriched in the resistant

Tiassalé strain when compared to the susceptible N'Gousso; 105 of which

were up-regulated in the resistant strain when compared to the susceptible.

To determine the suitability of the selection criteria, 11 candidates were

validated by qPCR and demonstrated a high positive correlation with the

arrays (Pearson's r = 0.870). Interestingly, all nine of the cytochrome p450s

present on this list, in addition to >80% of the transcripts were detected in

equivalent whole organism arrays. These data suggest that using body part

microarrays adds very little sensitivity in the detection of insecticide

resistance candidates, which is in disagreement with previous findings

exploring RNAseq data (Johnson *et al.*, 2013); however, these data offer a

rich data set for structural specificity and the identification of networks of

detoxification family members.

The second aim of this study was to identify candidates that are both

involved directly with resistance to insecticides and those that may

transcriptionally control response to insecticides. In order to fulfil these

aims 27 data sets available at LSTM, comparing a resistant and susceptible

An. gambiae, *An. coluzzii* or *An. arabiensis* were analysed. These data were then used in a global meta-analysis approach in order to identify key determinants of resistance, including novel candidates that may have been missed in the traditional 'cherry-picking' type approach employed by many microarray studies. The initial hypothesis that clear clustering of the microarrays would be seen based upon obvious variables such as geographical location; species and insecticides exposure status was soon rejected due to lack of clear consistent clusters. It is likely that this *ad hoc* based approach was unable to overcome large variance introduced by factors such as large geographical distances between population collections, experimental design and use of different reference susceptible strains. Indeed, *An. arabiensis* arrays demonstrated considerably more inter-species variance than did the *An. coluzzii* arrays, which had samples that were collected from much less disparate sites. Despite these initial problems, the use of GO terms in lieu of expression data allowed identification of clear overlaps in significant ontologies across the arrays. These data included terms related to signalling, transcription, chemo-sensation, neuronal mechanisms and ion binding, interestingly excluding traditional xenobiotic detoxification terms. Although GO term data is useful for summarising the overall processes and functions enriched in a gene list, data is lost in their use. Despite this, the approach here allowed identification of broad categories of change present across several, or indeed singular data sets.

Although the above data allowed for identification of over-riding ontology terms significant in resistant populations, the variance in number of GO

terms was extremely high. To overcome some of the noise in the dataset a smaller selection of arrays were utilised, including only pyrethroid exposed *An. coluzzii* arrays that were compared to the susceptible N'Gousso and similarly exposed *An. arabiensis* arrays. These data further demonstrated the huge disparity in *An. arabiensis* populations but did allow identification of few transcripts that show significantly differential expression across all populations, only *An. coluzzii* and only *An. arabiensis*, in addition to clear clustering in the *An. coluzzii* arrays. 13 transcripts demonstrated changes across all arrays, with one transcript (AGAP003048-RA) being upregulated across all this subset of arrays. It appears that AGAP003048, a ribosomal production factor plays a role in developmental processes (Guertin *et al.*, 2006; Blanco *et al.*, 2010) and may represent a change to global transcription. Interestingly, many of the remaining transcripts are largely Dipteran specific, with *ANXB9* potentially representing a change to the mosquito's reaction to stress (Telonis-Scott *et al.*, 2009; Yin *et al.*, 2013). These data should be explored with caution due to the limitation of one susceptible genetic background in the *An. coluzzii* arrays reducing confidence in these data sets.

To identify transcripts with either a role in transcriptional regulation of response to insecticides or direct involvement in pyrethroid resistance, hypothesis-driven analysis was used on the microarray data. Identification of putative transcriptional control was done using a pattern-matching algorithm on pyrethroid survivor arrays. 6215 transcripts demonstrated the same pattern of expression across all ten arrays used, of which 264 were

homologs to *Drosophila* transcription factors. A significance cut-off was applied to reduce the number of transcription factors to a manageable number; eight were identified. Of these eight, two had been linked to resistance to insecticide in *Drosophila*: *Met* and *Maf-S* (Li *et al.*, 2007; Misra *et al.*, 2011; Misra *et al.*, 2013).

Transcripts involved directly in pyrethroid resistance were found by requiring that they be: (i) Significant in all pyrethroid survivor populations and/or (ii) Significantly differentially expressed between survivors and complementary insecticide unexposed mosquitoes. These data resulted in (i) 969 and (ii) 12 transcripts that show putative involvement in pyrethroid resistance. These data showed the usefulness of hypothesis-driven analysis and provided a basis on which to answer several questions in the remaining studies:

1. Do any of the identified resistance associated transcripts show an effect on the pyrethroid resistance phenotype?

2. Is the *Maf-S* pathway involved in pyrethroid resistance?

3. Does *Met* play a role in resistance to insecticides classes besides the JH mimics?

Exploring these questions to the second layer of the project was accomplished in studies three through five, identifying three putatively important candidates with various roles in insecticide resistance.

Study three involved the exploration of transcripts that appear to play a role in resistance to pyrethroids. Eight candidates were validated by qPCR in two resistant populations, VK7 and Tiassalé. Of the eight candidates, seven demonstrated up-regulation in both resistant populations; of these, four showed up regulation of greater than five-fold and three up-regulation of greater than 20-fold. The four candidates showing high levels of up-regulation included two detoxification transcripts: a UGT (*UGT302A1*) and a cytochrome p450 (*CYP6Z3*), which is an ortholog of an enzyme that can metabolise insecticide detoxification products (Chandor-Proust *et al.*, 2013) lending weight to the methodology applied. The further two transcripts were a chemosensory protein (*SAP2*) and an elongation factor-like protein (*EF1-like*). Knockdown of each of *SAP2* and *EF-like* was successful, with *SAP2* knockdown almost completely restoring deltamethrin susceptibility in an extremely pyrethroid resistant population, Tiassalé. Surprisingly, the opposite effect was seen after bendiocarb exposure, hinting to complex involvement in cross-resistance to carbamates. Interestingly, *SAP2* has previously been shown to bind mid-sized aromatic compounds with similar structure to both carbamates and pyrethroids (Iovinella *et al.*, 2013); however, no insecticide binding activity has ever been tested.

Chemosensory proteins play a key role in the sensing of environmental cues (Krieger & Ross, 2002; Matsuo *et al.*, 2007; Asahina *et al.*, 2008), but have also been shown to be involved in processes outside of chemosensation (Nomura *et al.*, 1992; Maleszka *et al.*, 2007; Guo *et al.*, 2011). In agreement with current literature (Guidobaldi *et al.*, 2014) and database searching, *SAP2* appears to be enriched in the antennae (>200x) and legs (>30x); a sensory system and an area which will come into contact with insecticide. These data indicate that *SAP2* plays some role in chaperoning insecticide after contact or approach to insecticides. It is possible that sequestration or signalling behavioural changes have led to the vastly differing phenotype after knockdown. To determine whether *SAP2* expression increases on exposure to insecticide, as in *B. tabaci* (Liu *et al.*, 2014), qPCR was performed 24, 48 and 72 hours post exposure, with no evidence of induction; this likely means that *SAP2* is constitutively overexpressed in Tiassalé. These data, taken together are indicative of a novel mechanism of insecticide resistance in *Anopheles* mosquitoes and one that may be important in malaria control due to the extreme restoration of susceptibility to deltamethrin.

Transcriptional control of detoxification family members due to increased oxidative stress (Veraksa *et al.*, 2000) has been shown in *Drosophila* particularly in reference to resistance to DDT and malathion through constitutive activation of the *cnc-Maf-Keap1* pathway (Misra *et al.*, 2011; Misra *et al.*, 2013). Data from the 27 arrays analysed in Chapter 4 show that *cnc*, *Maf-S* and one *Keap1* homolog are significant in more arrays than

expected by chance, indicative of some involvement in resistance in *Anopheles* species. Correlation networks of *Maf-S* in particular further reinforce this hypothesis, with six out of the 14 co-correlated transcripts representing detoxification candidates. qPCR on cDNA produced from Tiassalé after knockdown of *Maf-S* demonstrated that four of these detoxification family members showed altered expression post *Maf-S* knockdown. A microarray was performed on the Tiassalé knockdown RNA and compared to GFP injected control, in total, 64 transcripts representing detoxification candidates had significantly altered expression. 31 and 33 detoxification transcripts were up- and down-regulated respectively, including *CYP6M2*, a known pyrethroid metaboliser (Stevenson *et al.*, 2011; Mitchell *et al.*, 2012), *GSTD1* and the CYP6Z and 6P family, all previously implicated in resistance (Müller *et al.*, 2008b; Low *et al.*, 2010; Chandor-Proust *et al.*, 2013). These data taken together are indicative of a similar role in *Anopheles* to that of the pathway in *Drosophila*, in the protection against oxidative stress. However, perturbation of both *cnc* and *Keap1* will be necessary to determine if these changes are due to the pathway or isolated to only *Maf-S*.

In addition to changes to the detoxifiers, large-scale changes in metabolic pathways were observed. These changes could be due to knock on changes to transcriptional activity, specifically when it is considered that purine and pyrimidine metabolism were the most affected pathways. As well as changes to these major metabolic pathways, changes to FoxO signalling, RNA transport and proteolysis were observed. Unfortunately, specific pathways

related to oxidative stress and xenobiotic metabolism are not well mapped

for insects and so would be missed in a pathway analysis approach, leading

to a large caveat in this particular aspect of the analysis.

The second transcription factor identified as being putatively involved in

insecticide resistance is *Methoprene-tolerant*, the juvenile hormone receptor

(Ashok *et al.*, 1998) implicated in resistance to the JH mimics pyriproxyfen

(Charles *et al.*, 2011) and methoprene (Wilson & Ashok, 1998). A microarray

was performed comparing *Met* knockdown RNA to GFP injected control

RNA. A relatively small transcriptomic change was observed, with 1349

significantly differentially expressed transcripts. GO term analysis revealed

that oxidation reduction processes were enriched in the down-regulated

transcripts, indicative of some role in oxidative stress response. The

pyrethroid and bendiocarb resistance profiles were determined after *Met*

knockdown, with a significant increase in mortality post-knockdown. To

further explore this phenotype, detoxification transcript enrichment was

determined using the microarray data. Only six detoxifiers were present in

the up-regulated transcript list, whereas 27 were present in the down-

regulated list - a significant enrichment. Again, *CYP6M2*, previously

implicated in insecticide metabolism (Stevenson *et al.*, 2011; Mitchell *et al.*,

2012), was present, indicating some role in transcriptional control of

resistance.

Due to the similarity in pathways affected by the knockdowns of *Maf-S* and

Met, a comparison was performed on the altered transcripts to determine if

the response to transcription factor knockdown was a general or specific response. Both knockdowns resulted in a large number of significantly differential candidates, with a similar proportion up- and down-regulated in each. The overlap between the two transcript lists was sizeable, indicative of a generalist response to knockdown or some overlapping function of the transcription factors. These data may be explained by (i) a generalised change to transcription after knockdown of wide-ranging transcription factors, (ii) an overlapping role in response to insecticide exposure and/or (iii) a batch affect. Indeed, the GO terms show substantial overlaps, including oxidation-reduction, various genome related assemblies and histone changes, showing aspects of both hypothesis (i) and (ii). This overlap of transcripts extends to the significantly differentially expressed detoxifiers, with 20 significant on both arrays.

Overall, these studies fulfil the objectives of the project and reveal three very interesting and novel insecticide resistance candidates in the form of *SAP2*, *Maf-S* and *Met*. These three transcripts represent not only putative direct involvement in resistance to a variety of insecticides but also potential transcriptional control points; an area that is currently underexplored in vector biology.

9. APPENDIX

See CD for all relevant tables.

Bibliography

ABDALLA, H., WILDING, C.S., NARDINI, L., PIGNATELLI, P., KOEKEMOER, L.L., RANSON, H. & COETZEE, M. 2014. Insecticide resistance in Anopheles arabiensis in Sudan: temporal trends and underlying mechanisms. *Parasites & Vectors*, **7**, 213.

ADRYAN, B. & TEICHMANN, S.A. 2006. FlyTF: a systematic review of site-specific transcription factors in the fruit fly Drosophila melanogaster. *Bioinformatics*, **22**, 1532–1533.

AHMAD, M., DENHOLM, I. & BROMILOW, R.H. 2006. Delayed cuticular penetration and enhanced metabolism of deltamethrin in pyrethroid-resistant strains of Helicoverpa armigera from China and Pakistan. *Pest management science*, **62**, 805–810.

AHN, S., BADENES-PÉREZ, F.R., REICHELT, M., SVATOŠ, A., SCHNEIDER, B., GERSHENZON, J. & HECKEL, D.G. 2011. Metabolic detoxification of capsaicin by UDP-glycosyltransferase in three Helicoverpa species. *Archives of insect biochemistry and physiology*, **78**, 104–118.

AHN, S.-J., VOGEL, H. & HECKEL, D.G. 2012. Comparative analysis of the UDP-glycosyltransferase multigene family in insects. *Insect biochemistry and molecular biology*, **42**, 133–147.

ALEXA, A. & RAHNENFUHRER, J. 2010. topGO: enrichment analysis for gene ontology. *R package version 2.8*.

ALOUT, H. & WEILL, M. 2008. Amino-acid substitutions in acetylcholinesterase 1 involved in insecticide resistance in mosquitoes. *Chemico-biological interactions*, **175**, 138–141.

AMENYA, D.A., NAGURAN, R., LO, T., RANSON, H., SPILLINGS, B.L., WOOD, O.R., BROOKE, B.D., COETZEE, M. & KOEKEMOER, L.L. 2008. Over expression of a cytochrome P450 (CYP6P9) in a major African malaria vector, Anopheles funestus, resistant to pyrethroids. *Insect molecular biology*, **17**, 19–25.

ARAVINDAN V, MUTHUKUMARAVEL S, G.K. 2014. Interaction affinity of Delta and Epsilon class glutathione-s-transferases (GSTs) to bind with DDT for detoxification and conferring resistance in Anopheles gambiae, a malaria vector. *Journal of Vector Borne Disease*, **51**, 8–15.

ARAVINDAN, V., MUTHUKUMARAVEL, S. & GUNASEKARAN, K. 2014. Interaction affinity of Delta and Epsilon class glutathione-s-transferases (GSTs) to bind with DDT for detoxification and conferring resistance in Anopheles gambiae, a malaria vector. *Journal of vector borne diseases*, **51**, 8.

ARMSTRONG, R.N. 1991. Glutathione S-transferases: reaction mechanism, structure, and function. *Chemical research in toxicology*, **4**, 131–140.

ASAHINA, K., PAVLENKOVICH, V. & VOSSHALL, L.B. 2008. The survival advantage of olfaction in a competitive environment. *Current Biology*, **18**, 1153–

1155.

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S. & Eppig, J.T. 2000. Gene Ontology: tool for the unification of biology. *Nature genetics*, **25**, 25–29.

Ashok, M., Turner, C. & Wilson, T.G. 1998. Insect juvenile hormone resistance gene homology with the bHLH-PAS family of transcriptional regulators. *Proceedings of the National Academy of Sciences*, **95**, 2761–2766.

Awolola, T.S., Oduola, O.A., Strode, C., Koekemoer, L.L., Brooke, B. & Ranson, H. 2009. Evidence of multiple pyrethroid resistance mechanisms in the malaria vector Anopheles gambiae sensu stricto from Nigeria. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, **103**, 1139–1145.

Awolola, T.S., Oyewole, I.O., Amajoh, C.N., Idowu, E.T., Ajayi, M.B., Oduola, A., Manafa, O.U., Ibrahim, K., Koekemoer, L.L. & Coetzee, M. 2005. Distribution of the molecular forms of Anopheles gambiae and pyrethroid knock down resistance gene in Nigeria. *Acta tropica*, **95**, 204–209.

Badolo, A., Traore, A., Jones, C.M., Sanou, A., Flood, L., Guelbeogo, W.M., Ranson, H. & Sagnon, N. 2012. Three years of insecticide resistance monitoring in Anopheles gambiae in Burkina Faso: resistance on the rise. *Malar J*, **11**, e232.

Baker, D.A., Nolan, T., Fischer, B., Pinder, A., Crisanti, A. & Russell, S. 2011. A comprehensive gene expression atlas of sex-and tissue-specificity in the malaria vector, Anopheles gambiae. *BMC genomics*, **12**, 296.

Bayoh, M.N., Mathias, D.K., Odiere, M.R., Mutuku, F.M., Kamau, L., Gimnig, J.E., Vulule, J.M., Hawley, W.A., Hamel, M.J. & Walker, E.D. 2010. Anopheles gambiae: historical population decline associated with regional distribution of insecticide-treated bed nets in western Nyanza Province, Kenya. *Malar J*, **9**, 62.

Bhatt, S., Weiss, D.J., Cameron, E., Bisanzio, D., Mappin, B., Dalrymple, U., Battle, K.E., Moyes, C.L., Henry, A. & Eckhoff, P.A. 2015. The effect of malaria control on Plasmodium falciparum in Africa between 2000 and 2015. *Nature*, **526**, 207–211.

Bigoga, J.D., Ndangoh, D.N., Awono-Ambene, P.H., Patchoke, S., Fondjo, E. & Leke, R.G.F. 2012. Pyrethroid resistance in Anopheles gambiae from the rubber cultivated area of Niete, South Region of Cameroon. *Acta Tropica*, **124**, 210–214 Available at: http://www.sciencedirect.com/science/article/pii/S0001706X12002823.

Bisset, J.A., Rodríguez, M.M., Ricardo, Y., Ranson, H., Perez, O., Moya, M. & Vazquez, A. 2011. Temephos resistance and esterase activity in the mosquito Aedes aegypti in Havana, Cuba increased dramatically between 2006 and 2008. *Medical and veterinary entomology*, **25**, 233–

239.

BLANCO, E., RUIZ-ROMERO, M., BELTRAN, S., BOSCH, M., PUNSET, A., SERRAS, F. & COROMINAS, M. 2010. Gene expression following induction of regeneration in Drosophila wing imaginal discs. Expression profile of regenerating wing discs. *BMC developmental biology*, **10**, 94.

BOLSTAD, B.M., IRIZARRY, R.A., ÅSTRAND, M. & SPEED, T.P. 2003. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics*, **19**, 185–193.

BONKE, M., TURUNEN, M., SOKOLOVA, M., VÄHÄRAUTIO, A., KIVIOJA, T., TAIPALE, M., BJÖRKLUND, M. & TAIPALE, J. 2013. Transcriptional networks controlling the cell cycle. *G3: Genes| Genomes| Genetics*, **3**, 75–90.

BOZZOLAN, F., SIAUSSAT, D., MARIA, A., DURAND, N., POTTIER, M., CHERTEMPS, T. & MAÏBÈCHE-COISNE, M. 2014. Antennal uridine diphosphate (UDP)-glycosyltransferases in a pest insect: diversity and putative function in odorant and xenobiotics clearance. *Insect molecular biology*.

BROWN, R.P., MCDONNELL, C.M., BERENBAUM, M.R. & SCHULER, M.A. 2005. Regulation of an insect cytochrome P450 monooxygenase gene (CYP6B1) by aryl hydrocarbon and xanthotoxin response cascades. *Gene*, **358**, 39–52.

BRUCE, M.C., DONNELLY, C.A., ALPERS, M.P., GALINSKI, M.R., BARNWELL, J.W., WALLIKER, D. & DAY, K.P. 2000. Cross-Species Interactions Between Malaria Parasites in Humans. *Science* , **287** , 845–848, 10.1126/science.287.5454.845.

BRUN, A., CUANY, A., LE MOUEL, T., BERGE, J. & 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.

BRYNE, J.C., VALEN, E., TANG, M.-H.E., MARSTRAND, T., WINTHER, O., DA PIEDADE, I., KROGH, A., LENHARD, B. & SANDELIN, A. 2008. JASPAR, the open access database of transcription factor-binding profiles: new content and tools in the 2008 update. *Nucleic acids research*, **36**, D102–D106.

BURGESS, E.A. & DUNCAN, I. 1990. Direct control of antennal identity by the spineless-aristapedia gene of Drosophila. *Molecular and General Genetics MGG*, **221**, 347–352.

BUSS, D.S. & CALLAGHAN, A. 2008. Interaction of pesticides with p-glycoprotein and other ABC proteins: A survey of the possible importance to insecticide, herbicide and fungicide resistance. *Pesticide Biochemistry and Physiology*, **90**, 141–153.

BUSTIN, S.A., BENES, V., GARSON, J.A., HELLEMANS, J., HUGGETT, J., KUBISTA, M., MUELLER, R., NOLAN, T., PFAFFL, M.W. & SHIPLEY, G.L. 2009. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clinical chemistry*, **55**, 611–622.

CHAHINE, S. & O'DONNELL, M.J. 2011. Interactions between

detoxification mechanisms and excretion in Malpighian tubules of Drosophila melanogaster. *Journal of Experimental Biology*, **214**, 462–468 Available at: http://jeb.biologists.org/content/214/3/462.abstract.

CHANDOR-PROUST, A., BIBBY, J., REGENT-KLOECKNER, M., ROUX, J., GUITTARD-CRILAT, E., POUPARDIN, R., ASAM RIAZ, M., PAINE, M., DAUPHIN-VILLEMANT, C. & REYNAUD, S. 2013. The central role of mosquito cytochrome P450 CYP6Zs in insecticide detoxification revealed by functional expression and structural modelling. *Biochemical Journal*.

CHARLES, J.-P., IWEMA, T., EPA, V.C., TAKAKI, K., RYNES, J. & JINDRA, M. 2011. Ligand-binding properties of a juvenile hormone receptor, Methoprene-tolerant. *Proceedings of the National Academy of Sciences*, **108**, 21128–21133.

CHEN, L., SEGAL, D., HUKRIEDE, N.A., PODTELEJNIKOV, A. V, BAYARSAIHAN, D., KENNISON, J.A., OGRYZKO, V. V, DAWID, I.B. & WESTPHAL, H. 2002. Ssdp proteins interact with the LIM-domain-binding protein Ldb1 to regulate development. *Proceedings of the National Academy of Sciences*, **99**, 14320–14325.

CHINTAPALLI, V.R., WANG, J. & DOW, J.A.T. 2007. Using FlyAtlas to Identify Better Drosophila melanogaster Models of Human Disease. *Nat Genet*, **39**, 715–720 Available at: http://dx.doi.org/10.1038/ng2049.

CHIU, T.-L., WEN, Z., RUPASINGHE, S.G. & SCHULER, M.A. 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.

CLARK, A.G. & SHAMAAN, N.A. 1984. Evidence that DDT-dehydrochlorinase from the house fly is a glutathione S-transferase. *Pesticide Biochemistry and Physiology*, **22**, 249–261.

COETZEE, M., HUNT, R.H., WILKERSON, R., DELLA TORRE, A., COULIBALY, M.B. & BESANSKY, N.J. 2013. Anopheles coluzzii and Anopheles amharicus, new members of the Anopheles gambiae complex. *Zootaxa*, **3619**, 246–274.

COETZEE, M. & KOEKEMOER, L.L. 2013. Molecular Systematics and Insecticide Resistance in the Major African Malaria Vector Anopheles funestus. *Annual Review of Entomology*, **58**, 393–412, 10.1146/annurev-ento-120811-153628.

COLE, S.P.C., SPARKS, K.E., FRASER, K., LOE, D.W., GRANT, C.E., WILSON, G.M. & DEELEY, R.G. 1994. Pharmacological characterization of multidrug resistant MRP-transfected human tumor cells. *Cancer Research*, **54**, 5902–5910.

COLUZZI, M., PETRARCA, V. & DI DECO, M.A. 1985. Chromosomal inversion intergradation and incipient speciation in Anopheles gambiae. *Italian Journal of Zoology*, **52**, 45–63.

CONESA, A., GÖTZ, S., GARCÍA-GÓMEZ, J.M., TEROL, J., TALÓN, M. & ROBLES, M. 2005. Blast2GO: a universal tool for annotation, visualization and analysis in

functional genomics research. *Bioinformatics*, **21**, 3674–3676.

CONSORTIUM, F. 2003. The FlyBase database of the Drosophila genome projects and community literature. *Nucleic Acids Research*, **31**, 172–175.

CONSORTIUM, G.O. 2004. The Gene Ontology (GO) database and informatics resource. *Nucleic acids research*, **32**, D258–D261.

CORBEL, V. & N'GUESSAN, R. 2013. Distribution, Mechanisms, Impact and Management of Insecticide Resistance in Malaria Vectors: A Pragmatic Review.

DABIRE, K.R., DIABATE, A., NAMOUNTOUGOU, M., TOE, K.H., OUARI, A., KENGNE, P., BASS, C. & BALDET, T. 2009. Distribution of pyrethroid and DDT resistance and the L1014F kdr mutation in Anopheles gambiae sl from Burkina Faso (West Africa). *Transactions of the Royal Society of Tropical Medicine and Hygiene*, **103**, 1113–1120.

DABIRÉ, R.K., NAMOUNTOUGOU, M., DIABATÉ, A., SOMA, D.D., BADO, J., TOÉ, H.K., BASS, C. & COMBARY, P. 2014. Distribution and Frequency of <italic>kdr</italic> Mutations within <italic>Anopheles gambiae</italic> s.l. Populations and First Report of the <italic>Ace.1</italic>G119S Mutation in <italic>Anopheles arabiensis</italic> from Burkina Faso (West Africa). *PLoS ONE*, **9**, e101484 Available at: http://dx.doi.org/10.1371%2Fjournal.pone.0101484.

DAVIES, T.G.E., FIELD, L.M., USHERWOOD, P.N.R. & WILLIAMSON, M.S. 2007. DDT, pyrethrins, pyrethroids and insect sodium channels. *IUBMB life*, **59**, 151–162.

DEBRUHL, H., WEN, H. & LIPSICK, J.S. 2013. The complex containing Drosophila Myb and RB/E2F2 regulates cytokinesis in a histone H2Av-dependent manner. *Molecular and cellular biology*, **33**, 1809–1818.

DELLA TORRE, A., FANELLO, C., AKOGBETO, M., DOSSOU-YOVO, J., FAVIA, G., PETRARCA, V. & COLUZZI, M. 2001. Molecular evidence of incipient speciation within Anopheles gambiae ss in West Africa. *Insect molecular biology*, **10**, 9–18.

DELLA TORRE, A., TU, Z. & PETRARCA, V. 2005. On the distribution and genetic differentiation of Anopheles gambiae ss molecular forms. *Insect biochemistry and molecular biology*, **35**, 755–769.

DERMAUW, W. & VAN LEEUWEN, T. 2013. The ABC gene family in arthropods: comparative genomics and role in insecticide transport and resistance. *Insect biochemistry and molecular biology*.

DERUA, Y.A., ALIFRANGIS, M., HOSEA, K.M., MEYROWITSCH, D.W., MAGESA, S.M., PEDERSEN, E.M. & SIMONSEN, P.E. 2012. Change in composition of the Anopheles gambiae complex and its possible implications for the transmission of malaria and lymphatic filariasis in north-eastern Tanzania. *Malar J*, **11**, 10–1186.

DIABATE, A., BRENGUES, C., BALDET, T., DABIRÉ, K.R., HOUGARD, J.M., AKOGBETO, M.,

KENGNE, P., ... CHANDRE, F. 2004. The Spread of the Leu-Phe kdr Mutation Through Anopheles gambiae Complex in Burkina Faso: Genetic Introgression and de novo Phenomena. *Tropical Medicine & International Health*, **9**, 1267–1273 Available at: http://dx.doi.org/10.1111/j.1365-3156.2004.01336.x.

DJÈGBÈ, I., AGOSSA, F.R., JONES, C.M., POUPARDIN, R., CORNELIE, S., AKOGBÉTO, M., RANSON, H. & CORBEL, V. 2014. Molecular characterization of DDT resistance in Anopheles gambiae from Benin. *Parasites & vectors*, **7**, 409, 10.1186/1756-3305-7-409.

DJÈGBÈ, I., BOUSSARI, O., SIDICK, A., MARTIN, T., RANSON, H., CHANDRE, F., AKOGBÉTO, M. & CORBEL, V. 2011. Dynamics of insecticide resistance in malaria vectors in Benin: first evidence of the presence of L1014S kdr mutation in Anopheles gambiae from West Africa. *Malar J*, **10**, 261.

DJOUAKA, R., BAKARE, A., COULIBALY, O., AKOGBETO, M., RANSON, H., HEMINGWAY, J. & 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 C7 - 538*, **9**, 1–10 LA – English, 10.1186/1471-2164-9-538.

DONNELLY, M.J., CORBEL, V., WEETMAN, D., WILDING, C.S., WILLIAMSON, M.S. & BLACK IV, W.C. 2009. Does< i> kdr</i> genotype predict insecticide-resistance phenotype in mosquitoes? *Trends in parasitology*, **25**, 213–219.

DORMAN, N. 2007. Microarray-based gene expression profiles in multiple tissues of the domesticated silkworm, Bombyx mori. *Genome Biol.*, **8**, 162, 10.2144/000113917.

EDI, C. V, DJOGBÉNOU, L., JENKINS, A.M., REGNA, K., MUSKAVITCH, M.A.T., POUPARDIN, R., JONES, C.M., ESSANDOH, J., KÉTOH, G.K. & PAINE, M.J.I. 2014a. CYP6 P450 Enzymes and ACE-1 Duplication Produce Extreme and Multiple Insecticide Resistance in the Malaria Mosquito Anopheles gambiae. *PLoS genetics*, **10**, e1004236.

EDI, C.A. V, KOUDOU, B.G., BELLAI, L., ADJA, A.M., CHOUAIBOU, M., BONFOH, B., BARRY, S.J.E., ... WEETMAN, D. 2014b. Long-term trends in Anopheles gambiae insecticide resistance in Côte d'Ivoire. *Parasites & Vectors*, **7**, 500, 10.1186/s13071-014-0500-z.

EDI, C. V., BENJAMIN, K.G., JONES, C.M., WEETMAN, D. & RANSON, H. 2012. Multiple-Insecticide Resistance in Anopheles gambiae Mosquitoes, Southern Côte d'Ivoire. *Emerging Infectious Diseases*, **18**, 1508–1511.

EMMONS, R.B., DUNCAN, D., ESTES, P.A., KIEFEL, P., MOSHER, J.T., SONNENFELD, M., WARD, M.P., DUNCAN, I. & CREWS, S.T. 1999. The spineless-aristapedia and tango bHLH-PAS proteins interact to control antennal and tarsal development in Drosophila. *Development*, **126**, 3937–3945.

ENAYATI, A.A., RANSON, H. & HEMINGWAY, J. 2005. Insect glutathione transferases and insecticide resistance. *Insect Molecular Biology*, **14**, 3–

8.

EPIS, S., PORRETTA, D., MASTRANTONIO, V., URBANELLI, S., SASSERA, D., DE MARCO, L., MEREGHETTI, V., … BANDI, C. 2014. Temporal dynamics of the ABC transporter response to insecticide treatment: insights from the malaria vector Anopheles stephensi. *Sci. Rep.*, **4** Available at: http://dx.doi.org/10.1038/srep07435.

FANELLO, C., AKOGBETO, M. & DELLA TORRE, A. 2000. Distribution of the pyrethroid knockdown resistance gene (kdr) in Anopheles gambiae sl from Benin. *Trans R Soc Trop Med Hyg*, **94**, 132.

FARHADIAN, S.F., SUÁREZ-FARIÑAS, M., CHO, C.E., PELLEGRINO, M. & VOSSHALL, L.B. 2012. Post-fasting olfactory, transcriptional, and feeding responses in Drosophila. *Physiology & behavior*, **105**, 544–553.

FIDOCK, D.A. 2013. Eliminating Malaria. *Science* , **340** , 1531–1533, 10.1126/science.1240539.

FOSSOG TENE, B., POUPARDIN, R., COSTANTINI, C., AWONO-AMBENE, P., WONDJI, C.S., RANSON, H. & ANTONIO-NKONDJIO, C. 2013. Resistance to DDT in an Urban Setting: Common Mechanisms Implicated in Both M and S Forms of Anopheles gambiae in the City of Yaoundé Cameroon. *PLoS ONE*, **8**, e61408 Available at: http://dx.doi.org/10.1371/journal.pone.0061408.

GERMANN, U.A. & CHAMBERS, T.C. 1998. Molecular analysis of the multidrug transporter, P-glycoprotein. *Cytotechnology*, **27**, 31–60.

GIBSON, G.G. & SKETT, P. 2001. Introduction to drug metabolism. *Introduction to drug metabolism*. Nelson Thornes.

GNANKINÉ, O., BASSOLÉ, I.H.N., CHANDRE, F., GLITHO, I., AKOGBETO, M., DABIRÉ, R.K. & MARTIN, T. Insecticide resistance in Bemisia tabaci Gennadius (Homoptera: Aleyrodidae) and Anopheles gambiae Giles (Diptera: Culicidae) could compromise the sustainability of malaria vector control strategies in West Africa. *Acta Tropica*, http://dx.doi.org/10.1016/j.actatropica.2013.06.004.

GNÉMÉ, A., GUELBÉOGO, W., RIEHLE, M., TIONO, A., DIARRA, A., KABRÉ, G., SAGNON, N. & VERNICK, K. 2013. Plasmodium species occurrence, temporal distribution and interaction in a child-aged population in rural Burkina Faso. *Malaria Journal*, **12**, 1–9, 10.1186/1475-2875-12-67.

GRIGORAKI, L., LAGNEL, J., KIOULOS, I., KAMPOURAKI, A., MOROU, E., LABBÉ, P., WEILL, M. & VONTAS, J. 2015. Transcriptome Profiling and Genetic Study Reveal Amplified Carboxylesterase Genes Implicated in Temephos Resistance, in the Asian Tiger Mosquito Aedes albopictus.

GUERENSTEIN, P.G. & LAZZARI, C.R. 2009. Host-seeking: how triatomines acquire and make use of information to find blood. *Acta tropica*, **110**, 148–158.

GUERTIN, D.A., GUNTUR, K.V.P., BELL, G.W., THOREEN, C.C. & SABATINI, D.M. 2006. Functional genomics identifies TOR-regulated genes that control growth and division. *Current Biology*, **16**, 958–970.

GUIDOBALDI, F., MAY-CONCHA, I.J. & GUERENSTEIN, P.G. 2014. Morphology and physiology of the olfactory system of blood-feeding insects. *Journal of Physiology-Paris*, **108**, 96–111.

GUILLEMAUD, T., MAKATE, N., RAYMOND, M., HIRST, B. & CALLAGHAN, A. 1997. Esterase gene amplification in Culex pipiens. *Insect Molecular Biology*, **6**, 319–327.

GUNNING, R. V, DEVONSHIRE, A.L. & MOORES, G.D. 1995. Metabolism of esfenvalerate by pyrethroid-susceptible and-resistant Australian Helicoverpa armigera (Lepidoptera: Noctuidae). *Pesticide Biochemistry and Physiology*, **51**, 205–213.

GUO, W., WANG, X., MA, Z., XUE, L., HAN, J., YU, D. & KANG, L. 2011. CSP and takeout genes modulate the switch between attraction and repulsion during behavioral phase change in the migratory locust. *PLoS Genet*, **7**, e1001291.

HARRIS, C., LAMBRECHTS, L., ROUSSET, F., ABATE, L., NSANGO, S.E., FONTENILLE, D., MORLAIS, I. & COHUET, A. 2010. Polymorphisms in Anopheles gambiae immune genes associated with natural resistance to Plasmodium falciparum. *PLoS pathogens*, **6**, e1001112.

HAY, S.I., OKIRO, E.A., GETHING, P.W., PATIL, A.P., TATEM, A.J., GUERRA, C.A. & SNOW, R.W. 2010a. Estimating the Global Clinical Burden of Plasmodium falciparum Malaria in 2007. *PLoS Med*, **7**, e1000290 Available at: http://dx.doi.org/10.1371/journal.pmed.1000290.

HAY, S.I., SINKA, M.E., OKARA, R.M., KABARIA, C.W., MBITHI, P.M., TAGO, C.C., BENZ, D., … GODFRAY, H.C.J. 2010b. Developing Global Maps of the Dominant Anopheles Vectors of Human Malaria. *PLoS Med*, **7**, e1000209 Available at: http://dx.doi.org/10.1371/journal.pmed.1000209.

HAYES, J.D. & WOLF, C.R. 1988. Role of glutathione transferase in drug resistance. *Glutathione Conjugation: Mechanisms and Biological Significance*, 3150–3355.

HEGDE, S.P., KUMAR, A., KURSCHNER, C. & SHAPIRO, L.H. 1998. c-Maf interacts with c-Myb to regulate transcription of an early myeloid gene during differentiation. *Molecular and cellular biology*, **18**, 2729–2737.

HEMINGWAY, J. & KARUNARATNE, S.H. 1998. Mosquito carboxylesterases: a review of the molecular biology and biochemistry of a major insecticide resistance mechanism. *Medical and Veterinary Entomology*, **12**, 1–12.

HEMINGWAY, J., MIYAMOTO, J. & HERATH, P.R.J. 1991. A possible novel link between organophosphorus and DDT insecticide resistance genes in Anopheles: Supporting evidence from fenitrothion metabolism studies. *Pesticide Biochemistry and Physiology*, **39**, 49–56.

HEMINGWAY, J., VONTAS, J., POUPARDIN, R., RAMAN, J., LINES, J., SCHWABE, C., MATIAS, A. & KLEINSCHMIDT, I. 2013. Country-level operational implementation of the Global Plan for Insecticide Resistance Management. *Proceedings of the National Academy of Sciences* , **110** , 9397–9402, 10.1073/pnas.1307656110.

HENRICK, C.A., STAAL, G.B. & SIDDALL, J.B. 1973. Alkyl 3, 7, 11-trimethyl-2, 4-dodecadienoates, a new class of potent insect growth regulators with juvenile hormone activity. *Journal of agricultural and food chemistry*, **21**, 354–359.

HIRANO, M., HATAKOSHI, M., KAWADA, H. & TAKIMOTO, Y. 1998. Pyriproxyfen and other juvenile hormone analogues. *Reviews in Toxicology*, **2**, 357–394.

HUANG, D., SHERMAN, B., TAN, Q., COLLINS, J., ALVORD, G., ROAYAEI, J., STEPHENS, R., BASELER, M., LANE, C. & LEMPICKI, R. 2007. The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists. *Genome Biology*, **8**, R183, doi: 10.1186/gb-2007-8-9-r183.

HUANG, D.W., SHERMAN, B.T. & LEMPICKI, R.A. 2008a. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protocols*, **4**, 44–57 Available at: http://dx.doi.org/10.1038/nprot.2008.211.

HUANG, F.-F., CHAI, C.-L., ZHANG, Z., LIU, Z.-H., DAI, F.-Y., LU, C. & XIANG, Z.-H. 2008b. The UDP-glucosyltransferase multigene family in Bombyx mori. *Bmc Genomics*, **9**, 563.

HUANG, H.-S., HU, N.-T., YAO, Y.-E., WU, C.-Y., CHIANG, S.-W. & SUN, C.-N. 1998. Molecular cloning and heterologous expression of a glutathione S-transferase involved in insecticide resistance from the diamondback moth, Plutella xylostella. *Insect biochemistry and molecular biology*, **28**, 651–658.

HUANG, Y.-T. & LIN, X. 2013. Gene set analysis using variance component tests. *BMC bioinformatics*, **14**, 210.

IBRAHIM, S.S., MANU, Y.A., TUKUR, Z., IRVING, H. & WONDJI, C.S. 2014. High frequency of kdr L1014F is associated with pyrethroid resistance in Anopheles coluzzii in Sudan savannah of northern Nigeria. *BMC infectious diseases*, **14**, 441.

ILBOUDO-SANOGO, E., BADOLO, A., GUELBEOGO, W.M., SAGNON, N. & SIRIMA, S.B. 2013. Insecticide Resistance in Malaria Vectors in Areas of Seasonal Malaria Transmission in Burkina Faso: Knock-Down Resistance Gene Distribution.

INGHAM, V.A., JONES, C.M., PIGNATELLI, P., BALABANIDOU, V., VONTAS, J., WAGSTAFF, S.C., MOORE, J.D. & RANSON, H. 2014. Dissecting the organ specificity of insecticide resistance candidate genes in Anopheles gambiae: known and novel candidate genes. *BMC Genomics*, **15**, 1018, 10.1186/1471-2164-15-1018.

INVEST, J.F. & LUCAS, J.R. 2008. Pyriproxyfen as a mosquito larvicide. *In Proceedings of the Sixth International Conference on Urban Pests Veszprem, Hungary*. 239–245.

IOVINELLA, I., BOZZA, F., CAPUTO, B., DELLA TORRE, A. & PELOSI, P. 2013. Ligand-binding study of Anopheles gambiae chemosensory proteins. *Chemical senses*, **38**, 409–419.

IRIZARRY, R.A., HOBBS, B., COLLIN, F., BEAZER-BARCLAY, Y.D., ANTONELLIS, K.J., SCHERF, U. & SPEED, T.P. 2003. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* , **4** , 249–264, 10.1093/biostatistics/4.2.249.

ISMAIL, H.M., O'NEILL, P.M., HONG, D.W., FINN, R.D., HENDERSON, C.J., WRIGHT, A.T., CRAVATT, B.F., HEMINGWAY, J. & PAINE, M.J.I. 2013. Pyrethroid activity-based probes for profiling cytochrome P450 activities associated with insecticide interactions. *Proceedings of the National Academy of Sciences*, 201320185.

JOHN TOMFOHR, JUN LU, T.B.K. 2005. Pathway level analysis of gene expression using singular value decomposition. *BMC Bioinformatics*, **6** Available at: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1261155/pdf/1471-2105-6-225.pdf.

JOHNSON, B.R., ATALLAH, J. & PLACHETZKI, D.C. 2013. The importance of tissue specificity for RNA-seq: highlighting the errors of composite structure extractions. *BMC genomics*, **14**, 586.

JONES, C.M., HAJI, K.A., KHATIB, B.O., BAGI, J., MCHA, J., DEVINE, G.J., DALEY, M., KABULA, B., ALI, A.S. & MAJAMBERE, S. 2013. The dynamics of pyrethroid resistance in Anopheles arabiensis from Zanzibar and an assessment of the underlying genetic basis. *Parasites & vectors*, **6**, 343.

JONES, C.M., SANOU, A., GUELBEOGO, W.M., SAGNON, N., JOHNSON, P.C.D. & RANSON, H. 2012a. Aging partially restores the efficacy of malaria vector control in insecticide-resistant populations of Anopheles gambiae s.l. from Burkina Faso. *Malaria Journal*, **11**, 24, 10.1186/1475-2875-11-24.

JONES, C.M., TOÉ, H.K., SANOU, A., NAMOUNTOUGOU, M., HUGHES, A., DIABATÉ, A., DABIRÉ, R., SIMARD, F. & RANSON, H. 2012b. Additional Selection for Insecticide Resistance in Urban Malaria Vectors: DDT Resistance in Anopheles arabiensis from Bobo-Dioulasso, Burkina Faso. *PLoS ONE*, **7**, e45995 Available at: http://dx.doi.org/10.1371/journal.pone.0045995.

KABULA, B., KISINZA, W., TUNGU, P., NDEGE, C., BATENGANA, B., KOLLO, D., MALIMA, R., KAFUKO, J., MOHAMED, M. & MAGESA, S. 2014. Co-occurrence and distribution of East (L1014S) and West (L1014F) African knock-down resistance in Anopheles gambiae sensu lato population of Tanzania. *Tropical Medicine & International Health*, n/a – n/a, 10.1111/tmi.12248.

KAISSLING, K.-E. 2001. Olfactory perireceptor and receptor events in moths: a kinetic model. *Chemical Senses*, **26**, 125–150.

KANEHISA, M. & GOTO, S. 2000. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic acids research*, **28**, 27–30.

KANNAN, M.B., SOLOVIEVA, V. & BLANK, V. 2012. The small MAF transcription factors MAFF, MAFG and MAFK: current knowledge and perspectives. *Biochimica et Biophysica Acta (BBA)-Molecular Cell Research*, **1823**, 1841–1846.

KASHIMA, I., TAKAHASHI, M., HASHIMOTO, Y., SAKOTA, E., NAKAMURA, Y. & INADA, T.

2014. A functional involvement of ABCE1, eukaryotic ribosome recycling factor, in nonstop mRNA decay in Drosophila melanogaster cells. *Biochimie*, **106**, 10–16.

KAWADA, H., DIDA, G.O., OHASHI, K., KAWASHIMA, E., SONYE, G., NJENGA, S.M., MWANDAWIRO, C. & MINAKAWA, N. 2014. A Small-Scale Field Trial of Pyriproxyfen-Impregnated Bed Nets against Pyrethroid-Resistant <italic>Anopheles gambiae</italic> s.s. in Western Kenya. *PLoS ONE*, **9**, e111195 Available at: http://dx.doi.org/10.1371%2Fjournal.pone.0111195.

KAWADA, H., FUTAMI, K., KOMAGATA, O., KASAI, S., TOMITA, T., SONYE, G., MWATELE, C., NJENGA, S.M., MWANDAWIRO, C. & MINAKAWA, N. 2011. Distribution of a knockdown resistance mutation (L1014S) in Anopheles gambiae ss and Anopheles arabiensis in Western and Southern Kenya. *PLoS One*, **6**, e24323.

KELLY-HOPE, L., RANSON, H. & HEMINGWAY, J. 2008. Lessons from the past: managing insecticide resistance in malaria control and eradication programmes. *The Lancet infectious diseases*, **8**, 387–389.

KHATRI, P., SIROTA, M. & BUTTE, A.J. 2012. Ten Years of Pathway Analysis: Current Approaches and Outstanding Challenges. *PLoS Comput Biol*, **8**, e1002375 Available at: http://dx.doi.org/10.1371/journal.pcbi.1002375.

KIM, M.D., JAN, L.Y. & JAN, Y.N. 2006. The bHLH-PAS protein Spineless is necessary for the diversification of dendrite morphology of Drosophila dendritic arborization neurons. *Genes & development*, **20**, 2806–2819.

KOAMA, B., NAMOUNTOUGOU, M., SANOU, R., NDO, S., OUATTARA, A., DABIRÉ, R.K., MALONE, D. & DIABATÉ, A. 2015. The sterilizing effect of pyriproxyfen on the malaria vector Anopheles gambiae: physiological impact on ovaries development. *Malaria journal*, **14**, 101.

KOFFI, A.A., ALOU, L.P.A., ADJA, M.A., CHANDRE, F. & PENNETIER, C. 2013. Insecticide resistance status of Anopheles gambiae ss population from M'Be: a WHOPES-labelled experimental hut station, 10 years after the political crisis in Cote d'Ivoire. *Malaria journal*, **12**, 151.

KOOPERBERG, C., FAZZIO, T.G., DELROW, J.J. & TSUKIYAMA, T. 2002. Improved background correction for spotted DNA microarrays. *Journal of Computational Biology*, **9**, 55–66.

KOSCHMIEDER, A., ZIMMERMANN, K., TRIßL, S., STOLTMANN, T. & LESER, U. 2012. Tools for Managing and Analyzing Microarray Data. *Briefings in Bioinformatics*, **13**, 46–60, 10.1093/bib/bbr010.

KOUTSOS, A.C., BLASS, C., MEISTER, S., SCHMIDT, S., MACCALLUM, R.M., SOARES, M.B., COLLINS, F.H., ... CHRISTOPHIDES, G.K. 2007. Life cycle transcriptome of the malaria mosquito Anopheles gambiae and comparison with the fruitfly Drosophila melanogaster . *Proceedings of the National Academy of Sciences* , **104** , 11304–11309, 10.1073/pnas.0703988104.

KRIEGER, M.J.B. & ROSS, K.G. 2002. Identification of a major gene regulating

complex social behavior. *Science*, **295**, 328–332.

KRISTENSEN, M. 2005. Glutathione S-transferase and insecticide resistance in laboratory strains and field populations of Musca domestica. *Journal of economic entomology*, **98**, 1341–1348.

KWIATKOWSKA, R.M., PLATT, N., POUPARDIN, R., IRVING, H., DABIRE, R.K., MITCHELL, S., JONES, C.M., DIABATÉ, A., RANSON, H. & WONDJI, C.S. 2013. Dissecting the mechanisms responsible for the multiple insecticide resistance phenotype in Anopheles gambiae ss, M form, from Vallée de Kou, Burkina Faso. *Gene.*

LAWSON, D., ARENSBURGER, P., ATKINSON, P., BESANSKY, N.J., BRUGGNER, R. V, BUTLER, R., CAMPBELL, K.S., CHRISTOPHIDES, G.K., CHRISTLEY, S. & DIALYNAS, E. 2009. VectorBase: a data resource for invertebrate vector genomics. *Nucleic acids research*, **37**, D583–D587.

LE GOFF, G., HILLIOU, F., SIEGFRIED, B.D., BOUNDY, S., WAJNBERG, E., SOFER, L., AUDANT, P. & FEYEREISEN, R. 2006. Xenobiotic response in Drosophila melanogaster: sex dependence of P450 and GST gene induction. *Insect biochemistry and molecular biology*, **36**, 674–682.

LEAL, W.S., NIKONOVA, L. & PENG, G. 1999. Disulfide structure of the pheromone binding protein from the silkworm moth, Bombyx mori. *FEBS letters*, **464**, 85–90.

LI, M., MEAD, E.A. & ZHU, J. 2011. Heterodimer of two bHLH-PAS proteins mediates juvenile hormone-induced gene expression. *Proceedings of the National Academy of Sciences*, **108**, 638–643.

LI, Y., ZHANG, Z., ROBINSON, G.E. & PALLI, S.R. 2007. Identification and characterization of a juvenile hormone response element and its binding proteins. *Journal of Biological Chemistry*, **282**, 37605–37617.

LIU, G.X., XUAN, N., CHU, D., XIE, H.Y., FAN, Z.X., BI, Y.P., PICIMBON, J., QIN, Y.C., ZHONG, S.T. & FA LI, Y. 2014. Biotype expression and insecticide response of Bemisia tabaci chemosensory protein-1. *Archives of insect biochemistry and physiology*, **85**, 137–151.

LIU, N., LI, M., GONG, Y., LIU, F. & LI, T. 2015. Cytochrome p450s–their expression, regulation, and role in insecticide resistance. *Pesticide Biochemistry and Physiology*.

LIU, N., XU, Q., ZHU, F. & ZHANG, L. 2006. Pyrethroid resistance in mosquitoes. *Insect Science*, **13**, 159–166.

LOW, W.Y., FEIL, S.C., NG, H.L., GORMAN, M.A., MORTON, C.J., PYKE, J., MCCONVILLE, M.J., BIERI, M., MOK, Y.-F. & ROBIN, C. 2010. Recognition and detoxification of the insecticide DDT by Drosophila melanogaster glutathione S-transferase D1. *Journal of molecular biology*, **399**, 358–366.

LU, Z., GUAN, X., SCHMIDT, C.A. & MATERA, A.G. 2014. RIP-seq analysis of eukaryotic Sm proteins identifies three major categories of Sm-containing ribonucleoproteins. *Genome Biol*, **15**, R7.

LUMJUAN, N., RAJATILEKA, S., CHANGSOM, D., WICHEER, J., LEELAPAT, P., PRAPANTHADARA, L., SOMBOON, P., LYCETT, G. & RANSON, H. 2011. The role of the Aedes aegypti Epsilon glutathione transferases in conferring resistance to DDT and pyrethroid insecticides. *Insect biochemistry and molecular biology*, **41**, 203–209.

LUMJUAN, N., STEVENSON, B.J., PRAPANTHADARA, L., SOMBOON, P., BROPHY, P.M., LOFTUS, B.J., SEVERSON, D.W. & RANSON, H. 2007. The Aedes aegypti glutathione transferase family. *Insect biochemistry and molecular biology*, **37**, 1026–1035.

LWETOIJERA, D., HARRIS, C., KIWARE, S., DONGUS, S., DEVINE, G., McCALL, P. & MAJAMBERE, S. 2014. Increasing role of Anopheles funestus and Anopheles arabiensis in malaria transmission in the Kilombero Valley, Tanzania. *Malaria Journal*, **13**, 1–10, 10.1186/1475-2875-13-331.

LYNE, R., SMITH, R., RUTHERFORD, K., WAKELING, M., VARLEY, A., GUILLIER, F., JANSSENS, H., … MICKLEM, G. 2007. FlyMine: An Integrated Database for Drosophila and Anopheles Genomics. *Genome Biology C7 - R129*, **8**, 1–16 LA – English, 10.1186/gb-2007-8-7-r129.

MACDONALD, J.W. 2008. Affycoretools: Functions useful for those doing repetitive analyses with Affymetrix GeneChips. *R package, version*, **1**.

MAGDALOU, J., FOURNEL-GIGLEUX, S. & OUZZINE, M. 2010. Insights on membrane topology and structure/function of UDP-glucuronosyltransferases. *Drug metabolism reviews*, **42**, 159–166.

MALESZKA, J., FORÊT, S., SAINT, R. & MALESZKA, R. 2007. RNAi-induced phenotypes suggest a novel role for a chemosensory protein CSP5 in the development of embryonic integument in the honeybee (Apis mellifera). *Development genes and evolution*, **217**, 189–196.

MANNERVIK, B., HELENA DANIELSON, U. & KETTERER, B. 1988. Glutathione transferases: structure and catalytic activity. *CRC critical reviews in biochemistry*, **23**, 283–337.

MANOHARAN, M., CHONG, M.N.F., VAÏTINADAPOULÉ, A., FRUMENCE, E., SOWDHAMINI, R. & OFFMANN, B. 2013. Comparative genomics of odorant binding proteins in Anopheles gambiae, Aedes aegypti, and Culex quinquefasciatus. *Genome biology and evolution*, **5**, 163–180.

MARINOTTI, O., CALVO, E., NGUYEN, Q.K., DISSANAYAKE, S., RIBEIRO, J.M.C. & JAMES, A.A. 2006. Genome-wide analysis of gene expression in adult Anopheles gambiae. *Insect molecular biology*, **15**, 1–12.

MARTINEZ-TORRES, D., CHANDRE, F., WILLIAMSON, M.S., DARRIET, F., BERGE, J.B., DEVONSHIRE, A.L., GUILLET, P., PASTEUR, N. & PAURON, D. 1998. Molecular characterization of pyrethroid knockdown resistance (kdr) in the major malaria vector Anopheles gambiae ss. *Insect molecular biology*, **7**, 179–184.

MATSUO, T., SUGAYA, S., YASUKAWA, J., AIGAKI, T. & FUYAMA, Y. 2007. Odorant-binding proteins OBP57d and OBP57e affect taste perception and host-plant preference in Drosophila sechellia. *PLoS Biol*, **5**, e118.

MAWEJJE, H.D., WILDING, C.S., RIPPON, E.J., HUGHES, A., WEETMAN, D. & DONNELLY, M.J. 2013. Insecticide resistance monitoring of field-collected Anopheles gambiae sl populations from Jinja, eastern Uganda, identifies high levels of pyrethroid resistance. *Medical and veterinary entomology*, **27**, 276–283.

MBOERA, L.E.G., TAKKEN, W., MDIRA, K.Y., CHUWA, G.J. & PICKETT, J.A. 2000. Oviposition and behavioral responses of Culex quinquefasciatus to skatole and synthetic oviposition pheromone in Tanzania. *Journal of Chemical Ecology*, **26**, 1193–1203.

MCCANN, R.S., OCHOMO, E., BAYOH, M.N., VULULE, J.M., HAMEL, M.J., GIMNIG, J.E., HAWLEY, W.A. & WALKER, E.D. 2014. Reemergence of Anopheles funestus as a vector of Plasmodium falciparum in western Kenya after long-term implementation of insecticide-treated bed nets. *The American journal of tropical medicine and hygiene*, **90**, 597–604.

MCGEE, M. & CHEN, Z. 2006. Parameter estimation for the exponential-normal convolution model for background correction of Affymetrix GeneChip data. *Statistical applications in genetics and molecular biology*, **5**.

MEHTA, J.P. & RANI, S. 2011. Software and Tools for Microarray Data Analysis. *In Gene Expression Profiling*. Springer, 41–53.

MELLOUL, D., MARSHAK, S. & CERASI, E. 2002. Regulation of insulin gene transcription. *Diabetologia*, **45**, 309–326.

MEUNIER, B., DE VISSER, S.P. & SHAIK, S. 2004. Mechanism of oxidation reactions catalyzed by cytochrome P450 enzymes. *Chemical reviews*, **104**, 3947–3980.

MHARAKURWA, S., MUTAMBU, S.L., MBERIKUNASHE, J., THUMA, P.E., MOSS, W.J. & MASON, P.R. 2013. Changes in the burden of malaria following scale-up of malaria control interventions in Mutasa District, Zimbabwe. *Malaria journal*, **12**, 223.

MIAO, W., HU, L., SCRIVENS, P.J. & BATIST, G. 2005. Transcriptional Regulation of NF-E2 p45-related Factor (NRF2) Expression by the Aryl Hydrocarbon Receptor-Xenobiotic Response Element Signaling Pathway DIRECT CROSS-TALK BETWEEN PHASE I AND II DRUG-METABOLIZING ENZYMES. *Journal of Biological Chemistry*, **280**, 20340–20348.

MISRA, J.R., HORNER, M.A., LAM, G. & THUMMEL, C.S. 2011. Transcriptional regulation of xenobiotic detoxification in Drosophila. *Genes & development*, **25**, 1796–1806.

MISRA, J.R., LAM, G. & THUMMEL, C.S. 2013. Constitutive activation of the Nrf2/Keap1 pathway in insecticide-resistant strains of Drosophila. *Insect Biochemistry and Molecular Biology*, **43**, 1116–1124, http://dx.doi.org/10.1016/j.ibmb.2013.09.005.

MITCHELL, S.N., RIGDEN, D.J., DOWD, A.J., LU, F., WILDING, C.S., WEETMAN, D., DADZIE, S., … DONNELLY, M.J. 2014. Metabolic and Target-Site Mechanisms Combine to Confer Strong DDT Resistance in <italic>Anopheles gambiae</italic>. *PLoS ONE*, **9**, e92662 Available at:

http://dx.doi.org/10.1371/journal.pone.0092662.

MITCHELL, S.N., STEVENSON, B.J., MÜLLER, P., WILDING, C.S., EGYIR-YAWSON, A., FIELD, S.G., HEMINGWAY, J., PAINE, M.J.I., RANSON, H. & DONNELLY, M.J. 2012. Identification and Validation of a Gene Causing Cross-Resistance Between Insecticide Classes in Anopheles gambiae From Ghana. *Proceedings of the National Academy of Sciences*, 10.1073/pnas.1203452109.

MIYAMOTO, J., HIRANO, M., TAKIMOTO, Y. & HATAKOSHI, M. 1993. Insect growth regulators for pest control, with emphasis on juvenile hormone analogs. Present status and future prospects. *In ACS Symposium series American Chemical Society*.

MNZAVA, A.P., KNOX, T.B., TEMU, E.A., TRETT, A., FORNADEL, C., HEMINGWAY, J. & RENSHAW, M. 2015. Implementation of the global plan for insecticide resistance management in malaria vectors: progress, challenges and the way forward. *Malaria Journal*, **14**, 173, 10.1186/s12936-015-0693-4.

MOIROUX, N., GOMEZ, M.B., PENNETIER, C., ELANGA, E., DJÈNONTIN, A., CHANDRE, F., DJÈGBÉ, I., GUIS, H. & CORBEL, V. 2012. Changes in Anopheles funestus biting behaviour following universal coverage of long-lasting insecticidal nets in Benin. *Journal of Infectious Diseases*, 10.1093/infdis/jis565.

MONRIBOT-VILLANUEVA, J., JUÁREZ-URIBE, R.A., PALOMERA-SÁNCHEZ, Z., GUTIÉRREZ-AGUIAR, L., ZURITA, M., KENNISON, J.A. & VÁZQUEZ, M. 2013. TnaA, an SP-RING Protein, Interacts with Osa, a Subunit of the Chromatin Remodeling Complex BRAHMA and with the SUMOylation Pathway in Drosophila melanogaster. *PloS one*, **8**.

MOREY, J.S., RYAN, J.C. & VAN DOLAH, F.M. 2006. Microarray validation: factors influencing correlation between oligonucleotide microarrays and real-time PCR. *Biological procedures online*, **8**, 175–193.

MOUTINHO-PEREIRA, S., STUURMAN, N., AFONSO, O., HORNSVELD, M., AGUIAR, P., GOSHIMA, G., VALE, R.D. & MAIATO, H. 2013. Genes involved in centrosome-independent mitotic spindle assembly in Drosophila S2 cells. *Proceedings of the National Academy of Sciences*, **110**, 19808–19813.

MÜLLER, P., CHOUAIBOU, M., PIGNATELLI, P., ETANG, J., WALKER, E., DONNELLY, M., SIMARD, F. & RANSON, H. 2008a. 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 Available at: http://dx.doi.org/10.1111/j.1365-294X.2007.03617.x.

MÜLLER, P., WARR, E., STEVENSON, B.J., PIGNATELLI, P.M., MORGAN, J.C., STEVEN, A., YAWSON, A.E., … DONNELLY, M.J. 2008b. Field-Caught Permethrin-Resistant Anopheles gambiae Overexpress CYP6P3, a P450 That Metabolises Pyrethroids. *PLoS Genet*, **4**, e1000286 Available at: http://dx.doi.org/10.1371/journal.pgen.1000286.

MURALI, T., PACIFICO, S., YU, J., GUEST, S., ROBERTS, G.G. & FINLEY, R.L. 2011. DroID 2011: a comprehensive, integrated resource for protein, transcription factor, RNA and gene interactions for Drosophila. *Nucleic acids research*, **39**, D736–D743.

MUTWIL, M., USADEL, B., SCHÜTTE, M., LORAINE, A., EBENHÖH, O. & PERSSON, S. 2010. Assembly of an interactive correlation network for the Arabidopsis genome using a novel heuristic clustering algorithm. *Plant Physiology*, **152**, 29–43.

MWANGANGI, J., MBOGO, C., ORINDI, B., MUTURI, E., MIDEGA, J., NZOVU, J., GATAKAA, H., … BEIER, J. 2013. Shifts in malaria vector species composition and transmission dynamics along the Kenyan coast over the past 20 years. *Malaria Journal C7 - 13*, **12**, 1–9 LA – English, 10.1186/1475-2875-12-13.

N'GUESSAN, R., CORBEL, V., AKOGBÉTO, M. & ROWLAND, M. 2007. Reduced efficacy of insecticide-treated nets and indoor residual spraying for malaria control in pyrethroid resistance area, Benin. *Emerging infectious diseases*, **13**, 199.

NAMOUNTOUGOU, M., DIABATÉ, A., ETANG, J., BASS, C., SAWADOGO, S.P., GNANKINIÉ, O., BALDET, T., … DABIRÉ, R.K. 2013. First report of the L1014S kdr mutation in wild populations of Anopheles gambiae M and S molecular forms in Burkina Faso (West Africa). *Acta Tropica*, **125**, 123–127, http://dx.doi.org/10.1016/j.actatropica.2012.10.012.

NAMOUNTOUGOU, M., SIMARD, F., BALDET, T., DIABATÉ, A., OUÉDRAOGO, J.B., MARTIN, T. & DABIRÉ, R.K. 2012. Multiple insecticide resistance in Anopheles gambiae sl populations from Burkina Faso, West Africa. *PLoS One*, **7**, e48412.

NARDINI, L., CHRISTIAN, R., COETZER, N. & KOEKEMOER, L. 2013. DDT and pyrethroid resistance in Anopheles arabiensis from South Africa. *Parasites & Vectors*, **6**, 1–9, 10.1186/1756-3305-6-229.

NDIATH, M.O., MAZENOT, C., SOKHNA, C. & TRAPE, J.-F. 2014. How the Malaria Vector Anopheles gambiae Adapts to the Use of Insecticide-Treated Nets by African Populations. *PLoS ONE*, **9**, e97700 Available at: http://dx.doi.org/10.1371/journal.pone.0097700.

NDIATH, M.O., SOUGOUFARA, S., GAYE, A., MAZENOT, C., KONATE, L., FAYE, O., SOKHNA, C. & TRAPE, J.-F. 2012. Resistance to DDT and Pyrethroids and Increased kdr Mutation Frequency in <italic>An. gambiae</italic> after the Implementation of Permethrin-Treated Nets in Senegal. *PLoS ONE*, **7**, e31943 Available at: http://dx.doi.org/10.1371%2Fjournal.pone.0031943.

NEIRA OVIEDO, M., VANEKERIS, L., CORENA-MCLEOD, M.D.P. & LINSER, P.J. 2008. A microarray-based analysis of transcriptional compartmentalization in the alimentary canal of Anopheles gambiae (Diptera: Culicidae) larvae. *Insect molecular biology*, **17**, 61–72.

NEWCOMB, R.D., CAMPBELL, P.M., OLLIS, D.L., CHEAH, E., RUSSELL, R.J. &

Oakeshott, J.G. 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.

Ngufor, C., N'Guessan, R., Fagbohoun, J., Odjo, A., Malone, D., Akogbeto, M. & Rowland, M. 2014. Olyset Duo (R)(a pyriproxyfen and permethrin mixture net): an experimental hut trial against pyrethroid resistant Anopheles gambiae and Culex quinquefasciatus in Southern Benin. *PloS one*, **9**, e93603.

Nishimura, D. 2001. BioCarta. *Biotech Software & Internet Report: The Computer Software Journal for Scient*, **2**, 117–120.

Nkya, T.E., Akhouayri, I., Poupardin, R., Batengana, B., Mosha, F., Magesa, S., Kisinza, W. & David, J.-P. 2014. Insecticide resistance mechanisms associated with different environments in the malaria vector Anopheles gambiae: a case study in Tanzania. *Malaria journal*, **13**, 28.

Nomura, A., Kawasaki, K., Kubo, T. & Natori, S. 1992. Purification and localization of p10, a novel protein that increases in nymphal regenerating legs of Periplaneta americana (American cockroach). *The International journal of developmental biology*, **36**, 391–398.

Nwane, P., Etang, J., Chouaïbou, M., Toto, J.C., Koffi, A., Mimpfoundi, R. & Simard, F. 2013. Multiple insecticide resistance mechanisms in Anopheles gambiae s.l. populations from Cameroon, Central Africa. *Parasites & Vectors*, **6**, 41, 10.1186/1756-3305-6-41.

Ohashi, K., Nakada, K., Ishiwatari, T., Shono, Y., Lucas, J.R. & Mito, N. 2012. Efficacy of pyriproxyfen-treated nets in sterilizing and shortening the longevity of Anopheles gambiae (Diptera: Culicidae). *Journal of medical entomology*, **49**, 1052–1058.

Ortelli, F., Rossiter, L., Vontas, J., Ranson, H. & Hemingway, J. 2003. Heterologous expression of four glutathione transferase genes genetically linked to a major insecticide-resistance locus from the malaria vector Anopheles gambiae. *Biochem. J*, **373**, 957–963.

Pages, H., Gentleman, R., Aboyoun, P. & DebRoy, S. Biostrings: String objects representing biological sequences, and matching algorithms, 2008. *R package version*, **2**, 160.

Pan, W. 2002. A comparative review of statistical methods for discovering differentially expressed genes in replicated microarray experiments. *Bioinformatics*, **18**, 546–554.

Pedra, J.H.F., McIntyre, L.M., Scharf, M.E. & Pittendrigh, B.R. 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.

Pedrini, N., Mijailovsky, S.J., Girotti, J.R., Stariolo, R., Cardozo, R.M., Gentile, A. & Juárez, M.P. 2009. Control of pyrethroid-resistant Chagas disease

vectors with entomopathogenic fungi. *PLoS Negl Trop Dis*, **3**, e434.

PELLETIER, J. & LEAL, W.S. 2011. Characterization of olfactory genes in the antennae of the Southern house mosquito, Culex quinquefasciatus. *Journal of insect physiology*, **57**, 915–929.

PELOSI, P., ZHOU, J.-J., BAN, L.P. & CALVELLO, M. 2006. Soluble proteins in insect chemical communication. *Cellular and Molecular Life Sciences CMLS*, **63**, 1658–1676.

POUPARDIN, R., SRISUKONTARAT, W., YUNTA, C. & RANSON, H. 2014. Identification of Carboxylesterase Genes Implicated in Temephos Resistance in the Dengue Vector <italic>Aedes aegypti</italic>. *PLoS Negl Trop Dis*, **8**, e2743 Available at: http://dx.doi.org/10.1371/journal.pntd.0002743.

PROTOPOPOFF, N., MATOWO, J., MALIMA, R., KAVISHE, R., KAAYA, R., WRIGHT, A., WEST, P.A., KLEINSCHMIDT, I., KISINZA, W. & MOSHA, F.W. 2013. High level of resistance in the mosquito Anopheles gambiae to pyrethroid insecticides and reduced susceptibility to bendiocarb in north-western Tanzania. *Malaria journal*, **12**, 149.

QIU, Y., TITTIGER, C., WICKER-THOMAS, C., LE GOFF, G., YOUNG, S., WAJNBERG, E., FRICAUX, T., TAQUET, N., BLOMQUIST, G.J. & FEYEREISEN, R. 2012. An insect-specific P450 oxidative decarbonylase for cuticular hydrocarbon biosynthesis. *Proceedings of the National Academy of Sciences*, **109**, 14858–14863.

QUACKENBUSH, J. 2001. Computational analysis of microarray data. *Nature reviews genetics*, **2**, 418–427.

RAICHE, G. & MAGIS, D. nFactors: Parallel analysis and non graphical solutions to the Cattell Scree Test.

RAMPHUL, U., BOASE, T., BASS, C., OKEDI, L.M., DONNELLY, M.J. & MÜLLER, P. 2009. Insecticide resistance and its association with target-site mutations in natural populations of Anopheles gambiae from eastern Uganda. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, **103**, 1121–1126.

RANSON, H., CLAUDIANOS, C., ORTELLI, F., ABGRALL, C., HEMINGWAY, J., SHARAKHOVA, M. V, UNGER, M.F., COLLINS, F.H. & FEYEREISEN, R. 2002. Evolution of supergene families associated with insecticide resistance. *Science*, **298**, 179–181.

RANSON, H., JENSEN, B., VULULE, J.M., WANG, X., HEMINGWAY, J. & COLLINS, F.H. 2000. Identification of a point mutation in the voltage-gated sodium channel gene of Kenyan Anopheles gambiae associated with resistance to DDT and pyrethroids. *Insect molecular biology*, **9**, 491–497.

RANSON, H., N'GUESSAN, R., LINES, J., MOIROUX, N., NKUNI, Z. & CORBEL, V. 2011. Pyrethroid Resistance in African Anopheline Mosquitoes: What Are The Implications For Malaria Control? *Trends in Parasitology*, **27**, 91–98 Available at: http://www.sciencedirect.com/science/article/pii/S14714922100017 53.

RANSON, H., ROSSITER, L., ORTELLI, F., JENSEN, B., WANG, X., ROTH, C., COLLINS, F. & HEMINGWAY, J. 2001. Identification of a novel class of insect glutathione S-transferases involved in resistance to DDT in the malaria vector Anopheles gambiae. *Biochem. J*, **359**, 295–304.

REDDY, M.R., OVERGAARD, H.J., ABAGA, S., REDDY, V.P., CACCONE, A., KISZEWSKI, A.E. & SLOTMAN, M.A. 2011. Outdoor host seeking behaviour of Anopheles gambiae mosquitoes following initiation of malaria vector control on Bioko Island, Equatorial Guinea. *Malar J*, **10**, 184.

REVELLE, W. 2015. psych.

RITCHIE, M.E., SILVER, J., OSHLACK, A., HOLMES, M., DIYAGAMA, D., HOLLOWAY, A. & SMYTH, G.K. 2007. A Comparison of Background Correction Methods for Two-Colour Microarrays. *Bioinformatics*, **23**, 2700–2707, 10.1093/bioinformatics/btm412.

RIVERON, J.M., IBRAHIM, S.S., CHANDA, E., MZILAHOWA, T., CUAMBA, N., IRVING, H., BARNES, K.G., NDULA, M. & WONDJI, C.S. 2014a. The highly polymorphic CYP6M7 cytochrome P450 gene partners with the directionally selected CYP6P9a and CYP6P9b genes to expand the pyrethroid resistance front in the malaria vector Anopheles funestus in Africa. *BMC genomics*, **15**, 817.

RIVERON, J.M., IRVING, H., NDULA, M., BARNES, K.G., IBRAHIM, S.S., PAINE, M.J.I. & WONDJI, C.S. 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, 10.1073/pnas.1216705110.

RIVERON, J.M., YUNTA, C., IBRAHIM, S.S., DJOUAKA, R., IRVING, H., MENZE, B.D., ISMAIL, H.M., HEMINGWAY, J., RANSON, H. & ALBERT, A. 2014b. A single mutation in the GSTe2 gene allows tracking of metabolically based insecticide resistance in a major malaria vector. *Genome Biol*, **15**, R27.

ROSSELL, D. 2009. GaGa: A Parsimonious and Flexible Model for Differential Expression Analysis. *Annals of Applied Statistics*, **3**, 1035–1051 Available at: http://arxiv.org/pdf/0910.1479.pdf.

ROTH, C.W., HOLM, I., GRAILLE, M., DEHOUX, P., RZHETSKYT, A., WINCKER, P., WEISSENBACH, J. & BREY, P.T. 2003. Identification of the Anopheles gambiae ATP-binding Cassette Transporter Superfamily Genes. *Molecules & Cells*, **15**, 150 Available at: http://search.ebscohost.com.ezproxy.liv.ac.uk/login.aspx?direct=true&db=a9h&AN=10242181&site=ehost-live&scope=site.

ROY, S., SAHA, T.T., JOHNSON, L., ZHAO, B., HA, J., WHITE, K.P., GIRKE, T., ZOU, Z. & RAIKHEL, A.S. 2015. Regulation of Gene Expression Patterns in Mosquito Reproduction. *PLoS Genet*, **11**, e1005450.

RUSSELL, T.L., GOVELLA, N.J., AZIZI, S., DRAKELEY, C.J., KACHUR, S.P. & KILLEEN, G.F. 2011. Increased proportions of outdoor feeding among residual malaria vector populations following increased use of insecticide-treated nets in rural Tanzania. *Malar J*, **10**, 80.

SCHINKEL, A.H., MOL, C., WAGENAAR, E., VAN DEEMTER, L., SMIT, J.J.M. & BORST, P. 1995. Multidrug resistance and the role of P-glycoprotein knockout mice. *European Journal of Cancer*, **31**, 1295–1298.

SHAH, S., YARROW, C., DUNNING, R., CHEEK, B., VASS, S., WINDASS, J. & HADFIELD, S. 2012. Insecticide detoxification indicator strains as tools for enhancing chemical discovery screens. *Pest management science*, **68**, 38–48.

SHANNON, P., MARKIEL, A., OZIER, O., BALIGA, N.S., WANG, J.T., RAMAGE, D., AMIN, N., SCHWIKOWSKI, B. & IDEKER, T. 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome research*, **13**, 2498–2504.

SHANNON, P., SHANNON, M.P., RUNIT, S. & BIOCVIEWS GENOMICSEQUENCE, M. 2013. Package "MotifDb."

SHETTY, V., SANIL, D. & SHETTY, N.J. 2013. Insecticide susceptibility status in three medically important species of mosquitoes, Anopheles stephensi, Aedes aegypti and Culex quinquefasciatus, from Bruhat Bengaluru Mahanagara Palike, Karnataka, India. *Pest management science*, **69**, 257–267 Available at: http://europepmc.org/abstract/MED/22926921.

SHIN, S.W., ZOU, Z., SAHA, T.T. & RAIKHEL, A.S. 2012. bHLH-PAS heterodimer of methoprene-tolerant and Cycle mediates circadian expression of juvenile hormone-induced mosquito genes. *Proceedings of the National Academy of Sciences*, **109**, 16576–16581.

SILVA, A.X., JANDER, G., SAMANIEGO, H., RAMSEY, J.S. & FIGUEROA, C.C. 2012. Insecticide resistance mechanisms in the green peach aphid Myzus persicae (Hemiptera: Aphididae) I: a transcriptomic survey. *PloS one*, **7**, e36366.

SINKA, M.E., RUBIO-PALIS, Y., MANGUIN, S., PATIL, A.P., TEMPERLEY, W.H., GETHING, P.W., VAN BOECKEL, T., KABARIA, C.W., HARBACH, R.E. & HAY, S.I. 2010a. The dominant Anopheles vectors of human malaria in the Americas: occurrence data, distribution maps and bionomic précis. *Parasites & vectors*, **3**, 72.

SINKA, M.H., RUBIO-PALIS, Y., MANGUIN, S., PATIL, A.P., TEMPERLEY, W.H., GETHING, P.W., VAN BOECKEL, T., KABARIA, C.W., HARBACH, R. & HAY, S. 2010b. The Dominant Anopheles Vectors of Human Malaria in Africa, Europe and the Middle East: Occurrence Data, Distribution Maps and Bionomic Précis. *Parasites & Vectors*, **3**, 117–152, 10.1186/1756-3305-3-72.

SMYTH, G.K. 2004. Linear models and empirical bayes methods for assessing differential expression in microarray experiments. *Stat Appl Genet Mol Biol*, **3**, 3.

SMYTH, G.K. & SPEED, T.P. 2003. Normalization of cDNA Microarray Data. *Methods*, **31**, 265–273 Available at: http://www.bioinformatics.plantbio.ohiou.edu/reference/Smyth.pdf.

SOGORB, M.A. & VILANOVA, E. 2002. Enzymes involved in the detoxification of organophosphorus, carbamate and pyrethroid insecticides through

hydrolysis. *Toxicology letters*, **128**, 215–228.

Sougoufara, S., Diédhiou, S.M., Doucouré, S., Diagne, N., Sembène, P.M., Harry, M., Trape, J.-F., Sokhna, C. & Ndiath, M.O. 2014. Biting by Anopheles funestus in broad daylight after use of long-lasting insecticidal nets: a new challenge to malaria elimination. *Malaria journal*, **13**, 125.

Steinbrecht, R.A. 1997. Pore structures in insect olfactory sensilla: a review of data and concepts. *International Journal of Insect Morphology and Embryology*, **26**, 229–245.

Steinbrecht, R.A. & Stankiewicz, B.A. 1999. Molecular composition of the wall of insect olfactory sensilla—the chitin question. *Journal of insect physiology*, **45**, 785–790.

Stevenson, B.J., Bibby, J., Pignatelli, P., Muangnoicharoen, S., O'Neill, P.M., Lian, L.-Y., Müller, P., Nikou, D., Steven, A. & Hemingway, J. 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.

Strode, C., Donegan, S., Garner, P., Enayati, A.A. & Hemingway, J. 2014. The impact of pyrethroid resistance on the efficacy of insecticide-treated bed nets against African anopheline mosquitoes: systematic review and meta-analysis. *PLoS medicine*, **11**, e1001619.

Stuart, J.M., Segal, E., Koller, D. & Kim, S.K. 2003. A Gene-Coexpression Network for Global Discovery of Conserved Genetic Modules. *Science* , **302** , 249–255, 10.1126/science.1087447.

Suh, E., Bohbot, J.D. & Zwiebel, L.J. 2014. Peripheral olfactory signaling in insects. *Current Opinion in Insect Science*, **6**, 86–92, http://dx.doi.org/10.1016/j.cois.2014.10.006.

Sykiotis, G.P. & Bohmann, D. 2008. Keap1/Nrf2 signaling regulates oxidative stress tolerance and lifespan in Drosophila. *Developmental cell*, **14**, 76–85.

Sykiotis, G.P. & Bohmann, D. 2010. Stress-activated cap'n'collar transcription factors in aging and human disease. *Science signaling*, **3**, re3.

Tang, A.H. & Tu, C.P. 1994. Biochemical characterization of Drosophila glutathione S-transferases D1 and D21. *Journal of Biological Chemistry*, **269**, 27876–27884.

Tarca, A.L., Draghici, S., Khatri, P., Hassan, S.S., Mittal, P., Kim, J., Kim, C.J., Kusanovic, J.P. & Romero, R. 2009. A novel signaling pathway impact analysis. *Bioinformatics* , **25** , 75–82 Available at: http://bioinformatics.oxfordjournals.org/content/25/1/75.abstract.

Telonis-Scott, M., Hallas, R., McKechnie, S.W., Wee, C.W. & Hoffmann, A.A. 2009. Selection for cold resistance alters gene transcript levels in Drosophila melanogaster. *Journal of Insect Physiology*, **55**, 549–555.

Temu, E.A., Maxwell, C., Munyekenye, G., Howard, A.F. V, Munga, S., Avicor, S.W., Poupardin, R., ... Ranson, H. 2012. Pyrethroid Resistance in

Anopheles gambiae, in Bomi County, Liberia, Compromises Malaria Vector Control. *PLoS ONE*, **7**, e44986 Available at: http://dx.doi.org/10.1371/journal.pone.0044986.

TENE, B.F., POUPARDIN, R., COSTANTINI, C., AWONO-AMBENE, P., WONDJI, C.S., RANSON, H. & ANTONIO-NKONDJIO, C. 2013. Resistance to DDT in an Urban Setting: Common Mechanisms Implicated in Both M and S Forms of Anopheles gambiae in the City of Yaoundé Cameroon. *PloS one*, **8**, e61408.

TERRIENTE-FÉLIX, A. & DE CELIS, J.F. 2009. Osa, a subunit of the BAP chromatin-remodelling complex, participates in the regulation of gene expression in response to EGFR signalling in the Drosophila wing. *Developmental biology*, **329**, 350–361.

THERNEAU, T.M. & GRAMBSCH, P.M. 2000. Modeling survival data: extending the Cox model. *Modeling survival data: extending the Cox model*. Springer Science & Business Media.

THOMSEN, E.K., STRODE, C., HEMMINGS, K., HUGHES, A.J., CHANDA, E., MUSAPA, M., KAMULIWO, M., ... COLEMAN, M. 2014. Underpinning Sustainable Vector Control through Informed Insecticide Resistance Management. *PLoS ONE*, **9**, e99822 Available at: http://dx.doi.org/10.1371%2Fjournal.pone.0099822.

TOÉ, K.H., N'FALÉ, S., DABIRÉ, R.K., RANSON, H. & JONES, C.M. 2015. The recent escalation in strength of pyrethroid resistance in Anopheles coluzzi in West Africa is linked to increased expression of multiple gene families. *BMC genomics*, **16**, 146.

TSENG, G.C., GHOSH, D. & FEINGOLD, E. 2012. Comprehensive literature review and statistical considerations for microarray meta-analysis. *Nucleic acids research*, gkr1265.

TUKEY, R.H. & STRASSBURG, C.P. 2000. Human UDP-glucuronosyltransferases: metabolism, expression, and disease. *Annual review of pharmacology and toxicology*, **40**, 581–616.

VERAKSA, A., MCGINNIS, N., LI, X., MOHLER, J. & MCGINNIS, W. 2000. Cap "n"collar B cooperates with a small Maf subunit to specify pharyngeal development and suppress deformed homeotic function in the Drosophila head. *Development*, **127**, 4023–4037.

VERHAEGHEN, K., VAN BORTEL, W., ROELANTS, P., BACKELJAU, T. & COOSEMANS, M. 2006a. Detection of the East and West African kdr mutation in Anopheles gambiae and Anopheles arabiensis from Uganda using a new assay based on FRET/Melt Curve analysis. *Malaria Journal*, **5**, 16.

VERHAEGHEN, K., VAN BORTEL, W., ROELANTS, P., BACKELJAU, T. & COOSEMANS, M. 2006b. Detection of the East and West African kdr mutation in Anopheles gambiae and Anopheles arabiensis from Uganda using a new assay based on FRET/Melt Curve analysis. *Malaria journal*, **5**, 16, 10.1186/1475-2875-5-16.

VILLARREAL, C.M., DARAKANANDA, K., WANG, V.R., JAYAPRAKASH, P.M. & SUZUKI, Y.

2015. Hedgehog signaling regulates imaginal cell differentiation in a basally branching holometabolous insect. *Developmental Biology*.

VONTAS, J., BLASS, C., KOUTSOS, A.C., DAVID, J., KAFATOS, F.C., LOUIS, C., HEMINGWAY, J., CHRISTOPHIDES, G.K. & 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, J., SMALL, G. & HEMINGWAY, J. 2001. Glutathione S-transferases as antioxidant defence agents confer pyrethroid resistance in Nilaparvata lugens. *Biochem. J*, **357**, 65–72.

VONTAS, J., SMALL, G., NIKOU, D., RANSON, H. & 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. *Biochem. J*, **362**, 329–337.

WEETMAN, D., MITCHELL, S.N., WILDING, C.S., BIRKS, D.P., YAWSON, A.E., ESSANDOH, J., MAWEJJE, H.D., DJOGBENOU, L.S., STEEN, K. & RIPPON, E.J. 2015. Contemporary evolution of resistance at the major insecticide target site gene Ace-1 by mutation and copy number variation in the malaria mosquito Anopheles gambiae. *Molecular Ecology*.

WEETMAN, D., WILDING, C.S., STEEN, K., PINTO, J. & DONNELLY, M.J. 2012. Gene flow–dependent genomic divergence between anopheles gambiae M and S forms. *Molecular biology and evolution*, **29**, 279–291.

WEI, S.H., CLARK, A.G. & 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.

WEILL, M., LUTFALLA, G., MOGENSEN, K., CHANDRE, F., BERTHOMIEU, A., BERTICAT, C., PASTEUR, N., PHILIPS, A., FORT, P. & RAYMOND, M. 2003. Comparative genomics: Insecticide resistance in mosquito vectors. *Nature*, **423**, 136–137.

WEILL, M., MALCOLM, C., CHANDRE, F., MOGENSEN, K., BERTHOMIEU, A., MARQUINE, M. & RAYMOND, M. 2004. The unique mutation in ace-1 giving high insecticide resistance is easily detectable in mosquito vectors. *Insect Molecular Biology*, **13**, 1–7, 10.1111/j.1365-2583.2004.00452.x.

WERNER, T. 2008. Bioinformatics Applications for Pathway Analysis of Microarray Data. *Current Opinion in Biotechnology*, **19**, 50–54 Available at: http://www.sciencedirect.com/science/article/pii/S0958166907001474.

WHO. 2014. World Malaria Report. *World Malaria Report*. Geneva Available at: http://apps.who.int/iris/bitstream/10665/144852/2/9789241564830_eng.pdf?ua=1.

WILSON, T.G. & ASHOK, M. 1998. Insecticide resistance resulting from an absence of target-site gene product. *Proceedings of the National*

Academy of Sciences, **95**, 14040–14044.

WILSON, T.G., YERUSHALMI, Y., DONNELL, D.M. & RESTIFO, L.L. 2006. Interaction between hormonal signaling pathways in Drosophila melanogaster as revealed by genetic interaction between Methoprene-tolerant and Broad-Complex. *Genetics*, **172**, 253–264.

WITZIG, C. 2012. Investigating the genetic basis of pyrethroid resistance in two members of the Anopheles gambiae complex. *Investigating the genetic basis of pyrethroid resistance in two members of the Anopheles gambiae complex*. University of Liverpool.

WORLD HEALTH ORGANIZATION. 2013. Test procedures for insecticide resistance monitoring in malaria vector mosquitoes.

XU, C., LI, C.Y.-T. & KONG, A.-N.T. 2005. Induction of phase I, II and III drug metabolism/transport by xenobiotics. *Archives of pharmacal research*, **28**, 249–268.

YANG, J., MCCART, C., WOODS, D.J., TERHZAZ, S., GREENWOOD, K.G., FFRENCH-CONSTANT, R.H. & DOW, J.A.T. 2007. A Drosophila systems approach to xenobiotic metabolism. *Physiological Genomics*, **30**, 223–231 Available at: http://physiolgenomics.physiology.org/content/30/3/223.abstract.

YANG, Y., CHENG, J.-Z., SINGHAL, S.S., SAINI, M., PANDYA, U., AWASTHI, S. & AWASTHI, Y.C. 2001. Role of glutathione S-transferases in protection against lipid peroxidation overexpression of HGSTA2-2 in K562 cells protects against hydrogen peroxide-induced apoptosis and inhibits JNK and caspase 3 activation. *Journal of Biological Chemistry*, **276**, 19220–19230.

YANG, Y.H., DUDOIT, S., LUU, P., LIN, D.M., PENG, V., NGAI, J. & SPEED, T.P. 2002. Normalization for cDNA microarray data: a robust composite method addressing single and multiple slide systematic variation. *Nucleic acids research*, **30**, e15–e15.

YANG, Y.H. & THORNE, N.P. 2003. Normalization for two-color cDNA microarray data. *Lecture Notes-Monograph Series*, 403–418.

YE, J., COULOURIS, G., ZARETSKAYA, I., CUTCUTACHE, I., ROZEN, S. & MADDEN, T.L. 2012. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC bioinformatics*, **13**, 134.

YIN, S., XUE, J., SUN, H., WEN, B., WANG, Q., PERKINS, G., ZHAO, H.W., ELLISMAN, M.H., HSIAO, Y. & YIN, L. 2013. Quantitative Evaluation of the Mitochondrial Proteomes of Drosophila melanogaster Adapted to Extreme Oxygen Conditions.

YU, J., PACIFICO, S., LIU, G. & FINLEY, R.L. 2008. DroID: the Drosophila Interactions Database, a comprehensive resource for annotated gene and protein interactions. *BMC genomics*, **9**, 461.

ZENG, X., LIN, X. & HOU, S.X. 2013. The Osa-containing SWI/SNF chromatin-remodeling complex regulates stem cell commitment in the adult Drosophila intestine. *Development*, **140**, 3532–3540.

ZHANG, C., LUO, X., NI, X., ZHANG, Y. & LI, X. 2010. Functional characterization of cis-acting elements mediating flavone-inducible expression of CYP321A1. *Insect biochemistry and molecular biology*, **40**, 898–908.

ZHANG, L., GAO, X. & 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, Z., XU, J., SHENG, Z., SUI, Y. & PALLI, S.R. 2011. Steroid receptor co-activator is required for juvenile hormone signal transduction through a bHLH-PAS transcription factor, methoprene tolerant. *Journal of Biological Chemistry*, **286**, 8437–8447.

ZOU, Z., SAHA, T.T., ROY, S., SHIN, S.W., BACKMAN, T.W.H., GIRKE, T., WHITE, K.P. & RAIKHEL, A.S. 2013. Juvenile hormone and its receptor, methoprene-tolerant, control the dynamics of mosquito gene expression. *Proceedings of the National Academy of Sciences*, **110**, E2173–E2181.