Molecular mechanisms underlying insecticide resistance phenotypes in the major malaria vector *Anopheles gambiae sensu stricto*

Thesis submitted in accordance with the requirements of the University of Liverpool for the degree of Doctor in Philosophy

by

Sara Nicola Mitchell

January 2011

Acknowledgements

Firstly I would like to thank my supervisors Martin Donnelly, Hilary Ranson and Emma Warr for their guidance, encouragement and support throughout this project. I also must thank Pie Muller for his patience, help and advice on all things microarray related (those R scripts would have been double-dutch otherwise!).

Invaluable advice and help in the lab has been courtesy of Craig Wilding, Dave Weetman and Naomi Dyer. Special thanks must be given to Craig for help with microarray and protein expression work as well as for supplying much needed lab banter.

I am indebted to the protein people for their time and patience explaining techniques to me and for being available for help and advice at any time. Thanks to Andrew Dowd, Bradley Stevenson and Mark Paine, with special mention to Andrew for his friendship, availability for pub excursions and endless supply of sarcasm.

For her infinite patience, excellent advice on all manner of issues (work related and not) as well as friendship and for often being the only other female in the pub on Friday evenings, Amy Lynd.

I am also thankful for the friendship and help in the lab supplied by Shavanthi, Amanda and Pat, with a special mention for Shav for always being available for essential shopping excursions.

I must also thank Gwen Finnegan for being able to sort any situation and for getting me to and from all manner of geographical locations with such ease.

For help during my field work I must thank Alex Egir-Yawson, Mike Osai, Alessi, Biney, and Osei my driver for always raising a smile.

Lastly I'm hugely thankful for the belief and support of Dave, my family and my friends, especially through the manic times!

Abstract

Insecticide resistance may compromise current insecticide based malaria control efforts. In the absence of new insecticidal compounds, understanding the mechanisms behind resistance to current insecticides is essential for prolonging control through resistance monitoring and management. Moreover, information on pathways implicated in resistance will be fundamental for informed insecticide design.

Previous work on insecticide resistance in malaria vectors has taken a candidate gene approach focusing on genes and gene families previously associated with the resistance phenotype. Though progress has been made, this approach may have overlooked novel mechanisms of resistance. This study aimed to take a more objective approach by employing whole genome expression profiling to examine resistance related gene expression in wild caught *Anopheles gambiae sensu stricto*. Two whole genome microarrays were designed (4x44K and 8x15K arrays) and employed to profile expression in DDT and carbamate (bendiocarb) resistant *An. gambiae s.s.* collected from the Accra region of Ghana, West Africa.

A number of genes from detoxifying gene families previously associated with resistance were identified including cytochrome P450s (*CYP325C1-3*, *CYP6M2*), glutathione-S-transferases (*GSTS1_1*, *GSTD3*) and esterases (*COEJHE2E*), which in part validated previous candidate gene approaches. However a number of novel resistance associated mechanisms were also discovered. Chemosensory pathways were implicated in both resistance to bendiocarb and DDT with a number of 'odorant binding proteins' (OBPs) and a putative OBP receptor over-expressed. *OBP3* expression was elevated in bendiocarb resistant samples when compared to non-exposed field controls. In subsequent recombinant protein assays this OBP was shown to interact with bendiocarb *in vitro* suggesting a potential role in detection/transport of this compound *in vivo*. Furthermore, OBPs were found to be over-expressed in DDT resistant field samples when compared to laboratory controls. Cuticular proteins were also over-expression in both DDT (*CPR9*) and bendiocarb (*CPLC4, CPLC17*) resistant samples, which may support a role for cuticular modification in the resistance phenotype; a mechanism which has received scant attention to date.

A second objective of this work was to investigate the role of allelic variation in resistance to insecticides. Through sequencing of a glutathione-S-transferase gene linked to DDT metabolic resistance (*GSTe2*), a number of single nucleotide polymorphisms (SNPs) were identified which segregated with DDT resistance phenotype in two laboratory colonies of *An. gambiae s.s.* Two of these SNPs conferred amino acid alterations which were proximal to the putative DDT binding domain. Recombinant expression of variant enzymes from the two strains of mosquito was pursued and DDTase assays performed with recombinant protein *in vitro*. The *GSTe2* allele from the resistant strain of *An. gambiae s.s.* (ZAN/U) exhibited a significantly higher maximum enzyme velocity (V_{max}) with substrate DDT compared to alleles from the susceptible strain (Kisumu), which exhibited substrate inhibition at high DDT concentrations. These results support a role for allelic variation in the resistance phenotype.

This study was the first to apply whole genome transcriptional profiling to insecticide resistant mosquito populations from the field. The identification of novel pathways and mechanisms potentially related to the resistance phenotype highlights the importance of taking an objective approach when investigating resistance related gene expression. Moreover, work on the enzyme *GSTe2* supports the potential role of allelic variation in conferring high levels of insecticide resistance, and emphasises the need to screen for genetic polymorphisms as well as elevated gene expression levels. Results from this work provide candidate markers for resistance screening in the field as well as potentially novel mechanisms of resistance, which will aid the development of new insecticidal compounds.

Table of Contents

List of Figures

List of Tables.

List of Abbreviations

An.	*Anopheles*
Ae.	*Aedes*
Cx.	*Culex*
s.s.	*sensu stricto*
s.l.	*sensu lato*
Bti	*Bacillus thuringiensis israelensis*
LLIN	long lasting insecticidal net
ITN	insecticide treated net
IRS	indoor residual spraying
WHO	World Health Organisation
WHOPES	WHO pesticide evaluation scheme
PBO	piperonyl butoxide
DDT	dichlorodiphenyltrichloroethane
DDE	dichlorodiphenyldichloroethylene
OP	organophosphate
GST	glutathione-S-transferase
GSH	glutathione
P450	cytochrome P450
iAChE	insensitive acetyl cholinesterase
ace-1	acetylcholinesterase
kdr	knockdown resistance
OBP	odorant binding protein
RFLP	restriction fragment length polymorphism
BLAST	basic local alignment and search tool
SNP	single nucleotide polymorphism
gDNA	genomic DNA
LD	linkage disequilibrium
PCR	polymerase chain reaction
NGS	next generation sequencing
LSTM	Liverpool School of Tropical Medicine
IVCC	Innovative Vector Control Consortium
LT_{50}	Median lethal time
IPTG	isopropyl β-D-1-thiogalactopyranoside
XGAL	bromo-chloro-indolyl-galactopyranoside
LB	Luria-Bertani

Chapter 1: Introduction

1.1 Vector control

In the absence of effective vaccines and with burgeoning resistance to expensive anti-malarial drugs, vector control remains one of the most widely employed methods for malaria control (Ramirez *et al.*, 2009).

The current strategies advocated by the World Health Organisation (WHO) (2005a) for the control of malaria focus on mosquito control, early diagnosis and prompt case management. They are categorised into four areas:

1. Preventative intervention strategies aimed at controlling the mosquito vector with particular focus on scaling up the provision of insecticide treated bednets.
2. Access to prompt and effective treatment.
3. Prevention and control of malaria epidemics.
4. Monitoring and evaluation of control.

WHO promote an integrated approach to vector management incorporating several techniques including the use of insecticide treated nets (ITNs), indoor residual spraying (IRS), biological control of larvae and environmental management to reduce breeding grounds.

Larvicidal approaches and the use of entomopathic fungi have received more interest of late, prompted by successes in field trails and emerging resistance to current insecticidal compounds. Microbial larvicidal toxins such as *Bacillus thuringiensis* var. *israelensis (Bti)* have been successfully employed in Kenyan trails where vector reduction and enhanced malaria control was reported when toxins were used alone or in combination with ITNs (Fillinger and Lindsay, 2006; Fillinger *et al.*, 2009). While new evidence suggests that adult mosquitoes can be employed for targeted and efficient dissemination of larvicides in the field (Devine and Killeen, 2010; Devine *et al.*, 2009).

Entomopathic fungi such as *Metarhizium anisopliae* have potential to act as life-shortening control agents resulting in fewer mosquitoes reaching infectious age (Farenhorst and Knols, 2007). While the multiple forms of fungal attack *in vivo* reduce the risk of resistance development (Knols *et al.*, 2010). Infection with fungi species *Beauveria bassiana* was also shown to block *Plasmodium* development and transmission in *Anopheles* (Blanford *et al.*, 2005); although a rodent malaria system was employed. Currently there is a deficit of large scale field trails examining the efficacy of fungal mosquito control; this may be partly attributed to the lack of long-lasting fungal formulas meeting WHO approval (Knols *et al.*, 2010). Although a small scale trial in Tanzania demonstrated the potential for mosquito control using a fungal based approach (Scholte *et al.*, 2005).

However, by far the most extensively employed methods of vector control involve the use of insecticidal compounds, most of which act upon targets within the insect

nervous system. Four classes of insecticide are widely used: the pyrethroids *e.g.* permethrin and deltamethrin; the organochlorines such as Dichloro-Diphenyl-Trichloroethane (DDT); the organophosphates including malathion and temephos, and the carbamates such as bendiocarb (WHO, 2006b). The last major class of insecticides discovered were the pyrethroids in the 1970s with little notable advance in insecticide development for public health since.

1.1.1 Insecticide treated nets (ITNs)

The preference of major malaria vectors to feed on humans in their dwellings (endophagic) between dusk and dawn (Service, 2004) means that bed-nets are an effective barrier to malaria transmission (Ross, 1910). However damaged nets provide little protection (Lines *et al.*, 1987; Mwangi *et al.*, 2003; Curtis *et al.*, 1996), which led to the treatment of bed-nets with an insecticidal chemical barrier (Curtis *et al.*, 2006; Snow *et al.*, 1988). Currently, pyrethroids are the only class of insecticide cleared for use on ITNs by the WHO pesticide evaluation scheme (WHOPES) due to their rapid knockdown effects, high insecticidal potency and most importantly low mammalian toxicity (WHO, 2007b). Pyrethroids, a synthetic derivative of plant extract pyrethrum, target the voltage gated sodium channels on nerve membranes, a target that is shared with the organochloride DDT (Vijverberg *et al.*, 1982). By delaying the closure of a small percentage of sodium channels, pyrethroids prolong the inward current of sodium ions into the nerve axon resulting in repetitive firing of the nervous system (Corbett *et al.*, 1984).

Insecticide treated netting has proven to be a successful control option with a meta analysis of 22 trails concluding ITN use effectively reduced malaria associated childhood morbidity and mortality as well as reducing infection during pregnancy (Lengeler, 2004). D'Alessandro *et al.* (1995) reported a 25% reduction in mortality and decreased rates of malaria parasitaemia in children aged 1-9 years, twelve months post ITN distribution in Gambia. While a randomised controlled trial of ITNs in an area of Kenya with intense malaria transmission (Asembo Bay) reported a 74% reduction in malaria attack rates in infancy (Ter Kuile *et al.*, 2003b) and a 38% reduction in the incidence of parasitaemia in pregnancy (Ter Kuile *et al.*, 2003a). If coverage is high in areas of intense transmission, ITNs can also provide community wide protection by significantly reducing the infective mosquito population; a protective effect was witnessed in Asembo Bay compounds without ITNs located within 300 metres of compounds supplied with nets (Hawley *et al.*, 2003).

A Cochrane review highlighted that if full ITN coverage was achieved in Africa, 370,000 child deaths per year would be prevented (Lengeler, 2004). Convincing evidence of efficacy in the field led WHO to strongly advocate the use of ITNs in their Roll Back Malaria campaign (Nabarro, 1999). In 2000 the Abuja Malaria summit set a target of 60% of children and pregnant women in Africa to be protected from malaria using ITNs with an aim of halving malaria mortality in Africa by 2010 (WHO, 2000). The Presidents Malaria Initiative (PMI), subsequently launched by President Bush and the US government in 2005, aimed to expand access to free and highly subsidised long

lasting nets to the poorest and most vulnerable groups in Africa as well as stimulating the market and demand for ITNs (WHO, 2005a). Initially, pyrethroid treated nets required re-treatment after ~three washes or at least once a year to maintain efficacy (WHO, 2005a). However long-lasting insecticidal nets (LLINs), which incorporate insecticide into or around the net fibres, have largely replaced nets that require retreatment and are now advocated by the WHO Global Malaria Programme (WHO, 2007b). By WHO definition LLINs must remain effective without re-treatment for at least 20 standard washes or three years deployment under field conditions.

ITN coverage has increased in many African countries since 2004 due to a combination of social marketing and free distribution. In Kenya a 60% increase in ITN coverage for children under five was reported after two years of ITN deployment through commercial marketing, provision of subsidized nets and free distribution (Noor *et al.*, 2007); it was also concluded that for social equity free distribution was the best option (WHO, 2006a). A sentiment shared by others (Curtis *et al.*, 2003; Cohen and Dupas, 2010). However to date, the Abuja targets have not been met with free ITN access still not a reality for those most vulnerable throughout Africa (Snow and Marsh, 2010). Moreover there are worrying discrepancies between ITN possession and use (Korenromp *et al.*, 2003; Macintyre *et al.*, 2006), which must be addressed through education.

The use of non-pyrethroid insecticides for ITNs is a future option in light of emerging pyrethroid resistance (Malcom, 1988; Elissa *et al.*, 1993; Chandre *et al.*, 1999; Ranson *et al.*, 2010); however insecticide formulas must achieve WHOPES approval. Combining two insecticides of different classes on a single net as a mixture or mosaic may render nets more effective and slow the development of resistance (USAID, 2007). Guillet *et al.* (2001) investigated the efficacy of ITNs combining a carbamate insecticide with pyrethroids under field conditions where wild mosquito populations were reportedly resistant to both classes. Marginally lower bloodfeeding rates were recorded with the dual-treated net compared to a pyrethroid-treated net. However there was no significant difference in *An. gambiae* mortality between the two types of net and some users (~20%) reported adverse side-effects from sleeping under carbamate treated netting. A recently developed commercial mosaic net incorporates the synergist piperonyl butoxide (PBO) into a deltamethrin treated LLIN. PBO inhibits insecticide metabolising enzymes and is mainly used to potentiate pyrethroid activity (Jones, 1998). In field comparisons PBO combined LLINS killed more pyrethroid resistant *An. gambiae* (52%) compared with nets treated with deltamethrin alone (44%); however results were not significant (P=0.036) and the authors concluded that PBO synergistic effects were lower than expected (N'Guessan *et al.*, 2010). Moreover, it was noted that PBO-combined nets contained twice the insecticide concentration compared with the deltamethrin only LLINs, meaning that the effects of PBO were indistinguishable (Corbel *et al.*, 2010).

1.1.2 Indoor residual spraying (IRS)

Pyrethroid insecticides are also employed in indoor residual spraying, although it has been argued that ITNs are a more affordable and acceptable insecticidal option (Curtis *et al.*, 1998). Indeed, comparative assessment of five ITN and two IRS programmes in Africa found IRS 6-10 times more costly per disability adjusted life year (DALY) compared to LLINs (Yukich *et al.*, 2007). While a recent Cochrane review suggested ITNs provide better protection in areas of unstable malaria transmission; however data was limited and it was concluded that more trials were required to provide adequate comparison (Pluess *et al.*, 2010).

Others would argue that a more effective option is an integrated vector management approach combining interventions such as ITN provision and IRS with regulatory and operational measures to ensure cost-effective and sustainable control (van den Berg and Takken, 2010). Such an approach may also delay resistance development if different insecticide classes are employed. From household surveys and a review of previous literature Kleinschmidt *et al.* (2009b) concluded that combined ITN and IRS intervention reduced the risk of malaria infection compared to one method alone, although it was not possible to control for all potential confounding factors. In addition, when the effect of multiple control methods including IRS, free LLIN distribution and provision of combined therapy, was assessed over a period of four years on Bioko Island, large reductions in infectivity rates as well as a 64% reduction in child mortality were reported (Kleinschmidt *et al.*, 2009a). However employing both ITNs and IRS is potentially prohibitively expensive, especially in combination with free drug provision.

One major advantage of indoor residual spraying over ITN intervention is less restriction on insecticides permitted with all four classes of insecticide approved by WHOPES for IRS purposes (Table 1.1). This may make IRS a more flexible control option, which is important in light of increasing insecticide resistance.

The period between 1935 and 1940 witnessed the first large-scale control of malaria vectors using indoor spraying with pyrethrum insecticide (Gilles, 2002). However the discovery of the residual insecticidal activity of DDT by Müller and Wiesmann in the late 1930's, and the later development of other synthetic insecticides revolutionised malaria control (Beales and Gilles, 2002).

DDT is an organochlorine insecticide, a group which also includes hexachlorocyclohexane (HCH), dieldrin and endosulfan. However DDT is the only insecticide from this class currently approved by WHO for IRS with nearly 547 tonnes employed annually in Africa (Zaim and Jambulingam, 2007). As mentioned, DDT shares its target, the voltage gated sodium channels of insect nerve membranes, with pyrethroid insecticides (Soderlund and Bloomquist, 1989), which has implications for cross resistance. DDT has many properties which favour its use in IRS including stability on impervious surfaces and residual activity of up to a year making DDT a more cost effective option (WHO, 2007c). Moreover, the use of DDT is now restricted to public health meaning reduced selection pressure from other applications. While

DDTs potential spatial repellency effects on mosquitoes may limit human-vector contact (Grieco *et al.*, 2007).

Indoor spraying with DDT was a fundamental part of The WHO malaria eradication campaign during the 1960s, which resulted in spectacular reductions in malaria incidence and mortality (Carter and Mendis, 2002). However concerns over bioaccumulation in mammalian tissues via the food chain (Carson, 1963) and emerging resistance has lead to a decline in the use of DDT and the compound being banned in many countries. Other insecticides have largely replaced DDT for IRS, particularly pyrethroids (Zaim *et al.*, 2000). However WHO have recently advocated the use of DDT for IRS which has led to a number of control programmes employing this organochlorine once more (WHO, 2006c; WHO, 2007c; Coleman *et al.*, 2008).

Table 1.1. WHO (2009) recommended insecticides for indoor residual treatment against malaria mosquito vectors.

(http://www.who.int/whopes/Insecticides_IRS_Malaria_09.pdf). (1) CS-capsule suspension, EC–emulsifiable concentrate, SC–suspension concentrate, WG–water dispersible granule, WP–wettable powder. (2) OC–organochlorines, OP–organophosphates, C–carbamates, PY–pyrethroids.

Insecticide compounds and formulations (1)	*Class group (2)*	*Dosage (g a.i./m²)*	*Mode of action*	*Duration of effective action (months)*
DDT WP	OC	1-2	contact	>6
Malathion WP	OP	2	contact	2-3
Fenitrothion WP	OP	2	contact & airborne	3-6
Pirimiphos-methyl WP & EC	OP	1-2	contact & airborne	2-3
Bendiocarb WP	C	0.1-0.4	contact & airborne	2-6
Propoxur WP	C	1-2	contact & airborne	3-6
Alpha-cypermethrin WP & SC	PY	0.02-0.03	contact	4-6
Bifenthrin WP	PY	0.025-0.05	contact	3-6
Cyfluthrin WP	PY	0.02-0.05	contact	3-6
Deltamethrin WP, WG	PY	0.02-0.025	contact	3-6
Etofenprox WP	PY	0.1-0.3	contact	3-6
Lambda-cyhalothrin WP, CS	PY	0.02-0.03	contact	3-6

There is currently a lack of definitive studies on the effect of IRS on malaria transmission/morbidity/mortality. A recent Cochrane review concluded that IRS reduced malaria incidence in unstable transmission settings, though results were mixed in areas of stable transmission (Pluess *et al.*, 2010). Data from randomized trials were also too few to quantify the effect size of IRS intervention.

Retrospective assessment of IRS programmes and analysis of malariological indices pre and post intervention provide some indication of IRS success in the field. One of the first large scale reviews collated parasitological and entomological observations pre and post IRS from 36 programmes in various epidemiological zones throughout Africa (Kouznetsov, 1977). It was concluded that spraying in endemic areas may not achieve

the reduction in transmission required for malaria eradication; however IRS made "considerable contribution to the general improvement of health and socio-economic well being".

Mabaso *et al.* (2004) conducted a historical review assessing the malaria situation in southern Africa before and after IRS intervention. Pre-spraying, malaria was predominantly hyper-endemic and relatively stable with intense seasonal transmission. Widespread IRS with pyrethrum began in 1931 followed by DDT spraying in the 1950-60s, with pyrethroids and carbamates replacing DDT during the late 1990s. Immediately post IRS dramatic reductions in malaria transmission and vector densities were recorded and subsequently epidemics became less frequent; although it was not possible to control for the effects of anti-malarial drug use and potential improvements in the standard of living.

The ability to switch insecticides for IRS may help to overcome the problems of emerging resistance. When populations of *An. funestus* became resistant to the pyrethroids used for IRS in South Africa (Hargreaves *et al.*, 2000) the spraying campaign switched to the carbamate, bendiocarb and subsequently DDT (Coleman *et al.*, 2008). While in Bioko Island, Equatorial Guinea, insecticide classes were also switched during an IRS programme after the first round of pyrethroid spraying failed to reduce the *An. gambiae* s.s. population. A switch to a carbamate during the second round of spraying was successful in reducing *An. gambiae s.s.* numbers (Sharp *et al.*, 2007).

Comprehensive resistance surveillance is essential for successful control. Before implementing an IRS programme or making the decision to switch insecticides all mosquito populations implicated in disease transmission should be assessed for resistance to all potential control compounds.

For instance WHO have recently advocated the use of DDT for IRS, after phasing out its use almost thirty years ago, claiming the insecticide is the most effective for this purpose and poses no health risk when use correctly (WHO, 2007c). However the successful reduction in malaria incidence and vector populations attributed to switching back to DDT in South Africa (Maharaj *et al.*, 2005) was only possible due to the lack of target site resistance (*kdr*) in the *An. funestus* population. This success may not be replicable in many parts of Africa where the *kdr* mutation is at high frequencies. In addition metabolic resistance to DDT conferred by detoxifying enzymes has emerged in populations of major malaria vector *An. arabiensis* in South Africa (Hargreaves *et al.*, 2003), which may compromise the success of DDT for IRS in the area.

Even in light of emerging resistance, IRS remains one of the most widely used and effective methods of vector control. Indoor spraying with insecticides is the vector control method of choice during malaria epidemics (WHO, 2005a) and can also be beneficial in refugee camps where the insides of tents are sprayed to reduce mosquito densities and malaria transmission (WHO, 2006b).

1.2 Insecticide resistance

The WHO (1957) definition of insecticide resistance is "the ability in a strain of some organism to tolerate doses of a toxicant that would prove lethal to a majority of individuals in a normal population of the same species". Insecticide resistance mechanisms can be placed into four categories (IRAC, 2009):

1. Behavioural – physical avoidance of insecticide treated areas.
2. Cuticular – reduced insecticide absorbance due to alterations of the cuticle.
3. Target site – mutations within the insecticide target.
4. Metabolic - increased detoxification of insecticidal compounds.

At present, evidence for the former two resistance mechanisms is inadequate or inconclusive. Some studies suggest a role for behavioural modification in the resistance phenotype (Silverman and Ross, 1994) (Lockwood *et al.*, 1985; Mathenge *et al.*, 2001); however this type of behaviour can often be difficult to characterise in an experimental system with many potential confounding effects (Sparks *et al.*, 1989) (Beales and Gilles, 2002). While early studies in *Musca domestica* implicated reduced cuticular penetration in insecticide resistance (DeVries and Georghiou, 1981; Plapp and Hoyer, 1968), with more recent studies exploring this phenomenon in other insect species through the use of radio-labelled insecticides (Puinean *et al.*, 2010; Ahmad *et al.*, 2006) and electron microscopy (Wood *et al.*, 2010; Pedrini *et al.*, 2009).

The majority of resistance characterisation has focused on the two major physiological mechanisms; reduced sensitivity of insecticide targets (target site resistance) and enhanced enzymatic detoxification of insecticides (metabolic resistance). Listed below (Table 1.2) are the major mechanisms of resistance to insecticides of public health use in malaria vectors, highlighting modes of cross-resistance.

Table 1.2. Major biochemical insecticide resistance mechanisms in mosquitoes.

Dot size is indicative of the relative impact of resistance mechanism (Table taken from the Insecticide Resistance Action Committee website http://www.irac-online.org). GSH – glutathione; *kdr* – knockdown resistance associated point mutations within the joint DDT and pyrethroid target, the voltage gated sodium channels; MACE – mutated acetylcholinesterase gene encoding the target site of both organophosphate and carbamate insecticides.

Biochemical mechanism of resistance					
	Metabolic			Target-site	
	Esterases	Monoxygenases	GSH S-transferases	*kdr*	MACE
Pyrethroids	●	●		○	
DDT		●	●	○	
Carbamates	●				○
Organophosphates	●	●			○

Cross-resistance to insecticides can develop when two or more insecticides share a target-site within the insect or a common detoxification pathway. Possession of a mutual target sites has lead to cross resistance between DDT and pyrethroids as well as between carbamates and organophosphate compounds. For instance mutations within the voltage gated sodium channels of nerve membranes confers knock down resistance (*kdr*) to both DDT and pyrethroids (Willamson *et al.*, 1993). While detoxifying enzyme, *CYP6G1,* is associated with resistance to three classes of insecticide in *Drosophila* (Joußen *et al.*, 2008; Le Goff *et al.*, 2003; Daborn *et al.*, 2002).

Exposing insects to a number of difference insecticides can lead to the development of multiple resistance to several compounds (Najera and Zaim, 2001), leaving few alternatives for insecticidal control. Though paradoxically the employment of insecticides of different classes as mixtures or in rotation is also key to the management of insecticide resistance (Corbel *et al.*, 2002; Rodriguez *et al.*, 2006). Numerous cases of multiple resistance have been reported in mosquito populations involving both target site mutations and elevated detoxifying enzymes (Corbel *et al.*, 2007; Dabire *et al.*, 2008; Perera *et al.*, 2008), which may threaten the efficacy of insecticide based control.

1.2.1 Target site resistance

Genetic mutations within genes encoding insecticide target sites can result in reduced sensitivity to compounds. A number of target site mutations have been recorded in insects.

1.2.1.1 GABA receptor

The receptors for neurotransmitter γ–aminobutyric acid (GABA) are the target site for the phenylpyrazole fipronil and chlorinated cyclodiene insecticides such as dieldrin. Mutations within the *Rdl* subunit of the GABA receptor have been associated with cross resistance to fipronil and cyclodiene insecticides. Resistance conferring mutations within the *Rdl* subunit were first documented in *Drosophila melanogaster* where an alanine to serine point mutation at position 302 (*A320S*) conferred resistance to cyclodienes and GABA receptor antagonist picrotoxinin (ffrench-Constant *et al.*, 1993). *Rdl* mutations have subsequently been associated with fipronil and cyclodiene resistance in the German cockroach (*Blattella germanica*) (Kristensen *et al.*, 2005) and *Drosophila simulans* (Le Goff *et al.*, 2005).

In *Anopheles* an alanine to serine mutation at position 296 within the *Rdl* locus was also found to confer dieldrin resistance in *An. arabiensis,* while an *A296G* substitution was associated with resistance in *An. gambiae s.s.* suggesting independent evolution of dieldrin resistance in these two malaria vectors (Du *et al.*, 2005). However more recently the M and S molecular forms of *An. gambiae s.s.* were found to carry different substitutions, with the *A296S* mutation reported in M forms and *A296G* found in S forms, suggesting independent selective sweeps within *An. gambiae s.s.* (Lawniczak *et al.*, 2010). *Rdl* mutations have also been associated with resistance to phenylpyrazole fipronil in *Anopheles stephensi* (Kolaczinski and Curtis, 2001) and *An. gambiae* (Brooke *et al.*, 2000).

1.2.1.2 Acetylcholinesterase (AChE)

Both organophosphate and carbamate insecticides share the enzyme acetylcholinesterase (AChE) as their target site in insects. By inhibiting this enzyme, the neurotransmitter acetylcholine accumulates in the synapses between cells impairing proper nerve function resulting in convulsions, paralysis then death (Corbett *et al.*, 1984). Organophosphates (OP) cause an almost irreversible inhibition of AChE while the action of carbamate insecticides upon this enzyme is considered to be more reversible (Becker *et al.*, 2003).

Mutations occurring within the *ace* locus, encoding the insect acetylcholinesterase, confer cross resistance to carbamates and organophosphates. Five point mutations within the *ace*-2 gene have been found in a resistant strain of *Drosophila*, with multiple mutations conferring a greater resistance phenotype (Mutero *et al.*, 1994).

A single mutation, *G119S*, on exon 5 of gene *ace-1* is associated with high levels of resistance in *Culex pipiens* and *An. gambiae* (Weill *et al.*, 2003). First reported in Cote

d'Ivoire, West Africa (N'Guessan *et al.*, 2003), this target site resistance to OPs and carbamates appears to be spreading to other West African populations, with reports of the *ace-1*^{R} mutation at frequencies between 0.04 and 0.66 in *An. gambiae* from neighbouring Burkina Faso (Djogbenou *et al.*, 2008b; Dabire *et al.*, 2009). A departure of *G119S* frequency from Hardy Weinberg equilibrium was also noted in Burkina Faso (Djogbenou *et al.*, 2008b) and in *An. gambiae* populations from Côte d'Ivoire (Djogbenou *et al.*, 2009). This departure resulted in an excess of heterozygotes and led to the hypothesis of an *ace-1* gene duplication creating permanent "heterozygotes" carrying a mutated and a wild-type gene (Djogbenou *et al.*, 2008a). This *ace-1* duplication phenomenon was first described in *Culex pipiens* through heterozygote excess (Lenormand *et al.*, 1998). Functioning acetylcholinesterase is essential for nerve signalling, therefore fitness costs are likely when insects are homozygous for the *ace*-1^R mutation (Djogbenou *et al.*, 2010; Bourguet *et al.*, 2004); gene duplication may help to overcome these costs.

1.2.1.3 Voltage gated sodium channels (*kdr*)

A trait termed *knock down resistance* (*kdr*) conferring resistance to the paralysing and lethal effects of pyrethroids, DDT and pyrethrins was first identified in the house fly, *Musca domestica* (Busvine, 1951) and was later found to be a central nervous system site-insensitivity factor (Miller *et al.*, 1979) mapping to the sodium channel gene locus (Willamson *et al.*, 1993). A second, more resistant phenotype has also been recorded in the house fly and was designated *super-kdr* (Sawicki, 1978).

Point mutations within the sodium channel gene were found to be associated with these knockdown phenotypes. In *Musca domestica* a *L1014F* substitution in the IIS6 transmembrane segment of the sodium channel correlated with the *kdr* trait, and an additional *M918T* replacement in segment IIS4-S5 was recorded in *super-kdr* flies only (Willamson *et al.*, 1996). The *L1014F* mutation in the *para*-orthologous sodium channel gene has subsequently been recorded in several insect species as well as numerous novel mutations which are comprehensively reviewed by Soderlund & Knipple (2003).

In *Anopheles gambiae, kdr* resistance was initially found in West African populations of the S molecular form associated with a single nucleotide change conferring the *L1014F* substitution (Martinez-Torres *et al.*, 1998). A second mutation (leucine-serine) was later implicated in *DDT* resistant *An. gambiae* originating in Kenya, East Africa (Ranson *et al.*, 2000a). However more recently both West and East African mutations have been recorded in S-form *An. gambiae s.s.* collected in Gabon (Pinto *et al.*, 2006), Cameroon (Etang *et al.*, 2006) and Uganda (Verhaeghen *et al.*, 2006), suggesting high levels of knockdown resistance.

The 'West African' *L1014F* mutation is now distributed throughout West and Central Africa, while the *L1014S* mutation remains mainly in Eastern regions (Pinto *et al.*, 2007). Differences have also been recorded in the frequency of the *L1014F* mutation between the two molecular forms of *An. gambiae s.s.*, with the allele occurring more

commonly in the S form (della Torre *et al.*, 2005). The *L1014F* reached M form populations via introgression from the S form, supported by the absence of tightly linked upstream intronic polymorphisms in wild type M forms (Weill *et al.*, 2000). This apparent gene flow between the two forms raises the risk of resistance rapidly spreading between molecular forms. Through analysis of the intron upstream of the *kdr* locus Pinto *et al.* (2007) concluded that rather than just two independent origins (East and West) at least four independent mutational events have led to *kdr* resistance alleles in Afrotropical *An. gambiae s.s.* of the S molecular form. While in addition to S form introgression, a *de novo* M form mutation is believed to have occurred in *An. gambiae s.s.* from Bioko Island, as linkage between *kdr* and upstream polymorphisms found in the S form was absent (Reimer *et al.*, 2005).

1.2.2 Metabolic Resistance

Metabolic resistance to insecticides commonly involves increased activity of proteins from large enzymatic gene families, namely, glutathione-s-transferases (GSTs), esterases and cytochrome P450 monooxygenases. These detoxifying enzymes convert insecticides to less toxic compounds through hydrolysis, oxidation and conjugation. Within the *An. gambiae* genome there are approximately 175 of these metabolic genes and ~ 235 in *Ae. aegypti*, with the cytochrome P450s showing the greatest diversity in both vectors (Table 1.3). Four main genetic alterations are believed to confer metabolic resistance either acting alone or in combination:

1. Over-expression
2. Gene duplication
3. Methylation
4. Allelic variation

Table 1.3. Classification of detoxification genes in *Drosophila melanogaster*, *Anopheles gambiae* and *Aedes aegypti*.

Table reproduced from Strode *et al.* (2008).

	D. melanogaster	*An. gambiae*	*Ae. aegypti*
Glutathione transferases			
Delta	11	12	8
Epsilon	14	8	8
Omega	5	1	1
Sigma	1	1	1
Theta	4	2	4
Zeta	2	1	1
Others	0	3	3
Total	**37**	**28**	**26**
Cytochrome P450s			
CYP4 clade	32	46	57
CYP3 clade			
CYP6	22	30	44
CYP9	5	9	37
Others	9	1	1
CYP2 clade	7	10	12
Mitochondrial clade	11	9	9
Total	**86**	**105**	**160**
Carboxy/cholinesterases			
Alpha esterases	13	16	22
Hormone processing			
Beta esterases	3	5	2
Juvenile hormone	2	4	6
esterases			
Others	3	4	7
Glutactin	4	9	10
Acetylcholinesterases	1	2	2
Total	**26**	**40**	**49**

1.2.2.1 Over-expression

Elevated expression of detoxifying enzymes is often linked to insecticide resistance. Among the most frequently implicated genes are members of the cytochrome P450 family which are associated with resistance to all major insecticidal classes. This complex superfamily of heme-containing enzymes has important roles in biosynthesis as well as detoxification; with P450s linked to hormone biosynthesis and degradation in insects (Niwa *et al.*, 2004; Helvig *et al.*, 2004). Notably, insect P450s also activate OP

insecticides, converting the phosphorothionate to the more toxic oxon form (Hemingway *et al.*, 2004), which is similar to the activation route of certain drugs in humans (Patterson and Murray, 2002).

The P450 CYP6 clade is most commonly associated with insecticide resistance via over-expression (Hemingway *et al.*, 2004). The multi-resistant Rutgers strain of *Musca domestica*, in addition to six other resistant strains, constitutively over-express the P450 gene *CYP6A1*, with expression at least ten times higher in the Rutgers strains compared with a reference strain (Carino *et al.*, 1994). The *CYP6A1* protein was subsequently found to be metabolically active against diazinon and cyclodiene insecticides (Feyereisen, 1999).

While in mosquitoes, over-expression of P450s is most commonly associated with pyrethroid resistance (Hargreaves *et al.*, 2000; Muller *et al.*, 2008b; Djouaka *et al.*, 2008; Shen *et al.*, 2003). Gene *CYP6M2* was found to be over-expressed via microarray analysis in a highly pyrethroid resistant Ghanaian colony (Muller *et al.*, 2007), as well as resistant field samples from Cameroon (P. Muller pers. comm.) and Benin (Djouaka *et al.*, 2008). HPLC evidence suggests this enzyme is capable of pyrethroid metabolism *in vitro* (M. Paine pers. comm.). More conclusive evidence for P450 involvement in pyrethroid resistance was provided by a study which detected over-expression of *CYP6P3* in resistant *An. gambiae* field isolates, subsequently demonstrating *CYP6P3* mediated pyrethroid metabolism in recombinant protein assays (Muller *et al.*, 2008b).

In *Culex* species elevated expression of CYP4 genes has been associated with pyrethroid resistance. Five CYP4 genes were found to be over-expressed via microarray and Northern blot in *Culex pipiens pallens* exhibiting deltamethrin resistance (Shen *et al.*, 2003). While in *Culex quinquefasciatus CYP4H34* as well as *CYP9M10* were over-expressed in a highly permethrin resistant strain (Komagata *et al.*, 2010). To date the role of these *Culex* P450 genes in insecticide metabolism has yet to be confirmed by recombinant protein analysis (or other mechanism).

P450 over-expression in *Drosophila* has been linked to DDT resistance with gene *CYP6G1* over-transcribed in twenty DDT resistant strains; transgenic over-expression of *CYP6G1* in susceptible strains was necessary and sufficient for DDT resistance (Daborn *et al.*, 2002).

Elevated expression of glutathione-S-transferase genes is also associated with DDT resistance in numerous insect species. GSTs facilitate the reductive dehydrochlorination of DDT producing easily excretable metabolites (Enayati *et al.*, 2005). Clark & Shamaan provided the first evidence of GST involvement in DDT resistance in houseflies (1984), while *GSTD1* was subsequently implicated in DDT metabolism in Drosophila (Tang and Tu, 1994).

Elevated activity of epsilon class GST, *GSTe2*, is linked to DDT resistance in mosquitoes. Prapanthadara *et al.* (1993) first reported quantitative and qualitative changes in GST conferring DDT resistance in the East African ZAN/U strain of *An. gambiae*. Evidence was later provided for the involvement of gene *GSTe2* with mRNA

levels ~5 fold higher in the ZAN/U strain compared to a susceptible strain, and recombinant *GSTe2* exhibiting high DDT dehydrochlorinase activity (Ranson *et al.*, 2001). Elevated *GSTe2* expression in *An. gambiae* has been associated with putative *cis* regulatory elements in the promoter region (Ding *et al.*, 2005). The *Ae. aegypti GSTe2* gene is also associated with DDT resistance with over-expression in resistant strains and DDT metabolising activity of recombinant protein (Lumjuan *et al.*, 2005); suggesting a parallel evolution of resistance in the two species.

GSTs are also implicated in organophosphate resistance via two distinct pathways; O-dearylation where glutathione reacts with the leaving group and O-dealkylation in which glutathione is conjugated to the alkyl portion of the insecticide (Enayati *et al.*, 2005). Elevated GST expression has been linked to OP resistance in the housefly (Wang *et al.*, 1991; Fournier *et al.*, 1992), with the O-dealkylation pathway implicated (Oppenoorth *et al.*, 1979). While in the diamondback moth (*Plutella xylostella*) over-expression of *GST-3* is linked to OP resistance via O-dearylation (Huang *et al.*, 1998). Recombinant GSTs from the housefly and diamondback moth are able to bind and metabolise organophosphate insecticides *in vitro*, further supporting their role in resistance (Huang *et al.*, 1998; Wei *et al.*, 2001).

Glutathione-S-transferases are also involved in protection against oxidative stress by aiding the removal of electrophilic and lipophilic compounds (reviewed by Salinas & Wong (1999)). Elevated expression of GSTs has been linked to pyrethroid resistance through antioxidant defence against products of pyrethroid induced lipid peroxidation (Vontas *et al.*, 2001).

1.2.2.2 Gene duplication

Gene duplication or amplification resulting in multiple gene copies within a cell can help satisfy demand for a particular gene product; this mechanism has been associated with OP, carbamate and pyrethroid insecticide resistance (Devonshire and Field, 1991).

Elevated esterases activity via gene duplication can confer broad-spectrum insecticide resistance through rapid insecticide binding and slow enzymatic turn-over (sequestration). In aphid *Myzus persicae*, two amplified esterases have been documented, *E4* and *FE4*, which give a broad spectrum of resistance to OP, carbamate and pyrethroid insecticides, accounting for as much as one percent of total body protein (Field *et al.*, 1988; Devonshire *et al.*, 1999).

Esterase gene amplification is also associated with broad-spectrum OP resistance in *Culex* species. Two non-specific carboxylesterases, *estα2*^{1} and *estβ2*^{1}, are duplicated in an OP resistant strain of *Culex quinquefasciatus* (Vaughan *et al.*, 1997). While both gene duplication and over-expression were implicated in over-production of esterases situated in two closely linked loci (*Est-2* and *Est-3*) in OP resistant *Culex pipiens* (Raymond *et al.*, 1998). To date, eleven alleles conferring OP resistance have been identified at these two loci which are commonly known as the *Ester* superlocus due to tight linkage between genes (Cui *et al.*, 2007).

More recently *CYP6P9* was found to be over-expressed in a pyrethroid resistant South African strain of *An. funestus* (Amenya *et al.*, 2008), with a QTL locus associated with resistance containing a cluster of P450 genes (Wondji *et al.*, 2007). Subsequently two over-expressed P450s within this cluster were found to be tandemly duplicated in the genome through analysis of BAC clones as well as field and laboratory isolates (Wondji *et al.*, 2009). Quantitative PCR revealed *CYP6P9* and *CYP6P4* to be 25 and 51 fold over-expressed respectively in the resistant strain; however the contribution of each gene copy in elevated expression was not resolved. Similar duplication of P450s associated with insecticide resistance is witnessed in *Drosophila* (Emerson *et al.*, 2008).

1.2.2.3 Methylation

DNA Methylation, specifically the addition of a methyl group to the 5 position of the cytosine pyrimidine ring, is another phenomenon associated with insecticide resistance; though to date has only been reported in Aphid species. Amplified copies of esterase *E4* in peach-potato aphid *M. persicae* were found to be highly methylated in resistant clones (Field *et al.*, 1989). Loss of methylation in amplified esterase sequences was subsequently found to be associated with a reversion of resistance in the absence of OP selection pressure, though when resistance was re-selected elevated esterases were no longer methylated (Hick *et al.*, 1996). A similar pattern of amplification and methylation of an OP resistance associated esterase was later reported in the greenbug aphid *Schizaphis graminum* (Ono *et al.*, 1999). The association between methylation and increased gene expression is in opposition with vertebrate systems in which methylation is predominantly correlated with gene silencing (Baylin, 2005; Turker, 2002).

1.2.2.4 Allelic variation

Mutations occurring within enzymes linked to resistance have the potential to alter metabolic activity, conferring higher levels of insecticide resistance. Mutations within esterase genes are associated with organophosphate resistance in the sheep blow fly (*Lucilia cuprina*) and house fly (*Musca domestica*). A tryptophan to leucine mutation at position 251 in the *E3* esterase from *Lucilia cuprina* confers resistance to malathion, while a second mutation in *E3* (*G137D*) results in a broad spectrum resistance to OP insecticides, excluding malathion (Newcomb *et al.*, 1997; Campbell *et al.*, 1998). The same *G137D* substitution in the orthologous esterase confers OP resistance in the house fly (Claudianos *et al.*, 1999).

Allelic variation within P450 gene *CYP6A2* has also been implicated in DDT resistant *Drosophila melanogaster*. Three amino-acid changes were associated with the *CYP6A2* allele from a DDT resistance strain and a recombinant mutant containing all three amino-acid substitutions had enhanced DDTase activity compared to the wild-type protein (Amichot *et al.*, 2004).

1.2.3 Impact of insecticide resistance on vector control

Despite insecticide resistance being documented in malaria vectors since the 1950s (Elliott and Ramakrishna, 1956; Ramakrishna and Elliott, 1959; Peffly, 1959), evidence for resistance linked control failure in the field is limited and potentially conflicting.

The case of pyrethroid resistance in South African malaria vectors suggests insecticide resistance has the potential to compromise control (Hargreaves *et al.*, 2000). In 1996 after ~40 years of use, DDT was replaced by pyrethroid deltamethrin for IRS in the KwaZulu-Natal region due to environmental concerns; by 1999 malaria incidence had increased more than 6 fold with annual cases rising from approximately 4000 to 27,000. This malaria upsurge was attributed to pyrethroid resistance in vector *An. funestus* (Hargreaves *et al.*, 2000; Brooke *et al.*, 2001). *An. funestus* were retrieved from houses recently sprayed with deltamethrin and remained susceptible to DDT suggesting metabolic rather than *kdr* resistance (Hargreaves *et al.*, 2000), which was supported by biochemical assays indicating over-expression of P450 monooxygenases (Brooke *et al.*, 2001). However sample sizes for resistance characterisation from the field were small (~40 wild caught mosquitoes) (Hargreaves *et al.*, 2000). In addition, factors such as climate fluctuations, population movement, *Plasmodium* drug resistance and HIV prevalence are also likely to have contributed to the increase in malaria incidence (Craig *et al.*, 2004). Subsequent re-introduction of DDT in 2000, in combination with increased availability of effective antimalarials, resulted in a reduction in malaria cases to <10,000; however DDT resistance was subsequently detected in populations of malaria vector *An. arabiensis* in the KwaZulu-Natal district (Hargreaves *et al.*, 2003).

Pyrethroid resistance was also implicated in the failure of IRS to control *An. gambiae s.s* populations on Bioko Island, Equatorial Guinea (Sharp *et al.*, 2007). The first of three spraying rounds employed synthetic pyrethroids and resulted in significant reductions of both the *An. melas* and *An. funestus* populations; however *An. gambiae s.s.* numbers were not reduced. This was attributed to the presence of target site resistance mutation, *kdr*, in the *An. gambiae s.s.* M form population, with *kdr* frequency increasing from 50 to 78% post spraying. Evidence of increased *kdr* frequency in *An. gambiae s.s.* post intervention has been reported elsewhere with the *L1014S* mutation doubling after ITN deployment in Kenya (Stump *et al.*, 2004), and increasing from 0.01-0.86 over 6 years post ITN and pyrethroid IRS intervention in Burundi (Protopopoff *et al.*, 2008). On Bioko Island, a switch to carbamate insecticides for IRS in subsequent spray rounds reduced the *An. gambiae s.s.* population and maintained control of the two other malaria vectors (Sharp *et al.*, 2007). However *An. gambiae s.s.* sporozoite rates were reduced after the initial pyrethroid spray round despite an increase in mosquito numbers; this was attributed to a change in the age structure of *An. gambiae s.s* after pyrethroid exposure and suggests that malaria control would have been effective even in the presence of insecticide resistance. Although reduced sporozoite rates may have been confounded by improved case management and intermittent

preventive therapy instigated as part of the integrated control programme (Kleinschmidt *et al.*, 2009a).

As pyrethroids are the only class cleared for use on ITNs, the emergence and spread of resistance is a concern. An experimental hut trial in Benin employing deliberately holed lambdacyhalothrin (pyrethroid) treated nets found evidence of treated nets failing to control *An. gambiae* with high *kdr* frequency (83%) (N'Guessan *et al.*, 2007); no reduction in bloodfeeding and only ~30% mortality was reported in the *kdr* resistant population when treated nets were compared to untreated nets, while ~99% mortality and a reduction in feeding of ~96% was recorded in a susceptible population.

However *kdr* resistance may not compromise current control methods according to some authors as the diminished sensitivity to the irritant/repellent effects of insecticides associated with *knock down resistance* can result in mosquitoes being killed through prolonged insecticide contact (Darriet *et al.*, 2000; Chandre *et al.*, 2000; Corbel *et al.*, 2004). In experimental hut and baited tunnel tests, holed pyrethroid treated ITNs provided "good levels" of protection against *An. gambiae* s.s. strains homozygous for the *kdr* mutation; in hut trials only ~ 3-8% of resistant strains were able to blood feed and 45-54% mortality was recorded, however higher mortality rates were reported in the fully susceptible comparison strain (~91-95%) and bloodfeeding rates were lower (0-1%) (Chandre *et al.*, 2000). While Darriet *et al.* (2000) recorded higher mortality rates in an area of Côte d'Ivoire where *An. gambiae s.s.* are resistant to pyrethroids (56.4%) compared to a susceptible area (44.3%) after experimental trials with deltamethrin treated ITNs; this was attributed to resistant mosquitoes being less repelled so a larger proportion acquired a lethal dose of insecticide. Henry *et al.* (2005) reported 56% protective efficacy in a case control study in which villages were provided with pyrethroid treated ITNs despite a *kdr* frequency of ~90% in the main malaria vector *An. gambiae s.s.*; although *kdr* frequencies and entomological indices were not directly investigated.

The reports of ITN efficacy in the presence of *kdr* resistance are encouraging and suggest pyrethroid treated nets may provide protection even in regions where *kdr* frequency is high. However the influence of metabolic resistance on ITN efficacy has yet to be explored, though it may be argued that the potential repellent effects of pyrethroids (Achee *et al.*, 2009; Mongkalangoon *et al.*, 2009) may reduce house entry and feeding rates even in the presence of metabolic resistance. Of greater concern is the potential for *kdr* resistance to be combined with metabolic mechanisms, as suspected in Ghana (Muller *et al.*, 2008b) and Benin (Corbel *et al.*, 2007), as both the repellent and lethal effects of pyrethroids may be compromised.

1.2.4 Detecting insecticide resistance mechanisms

Monitoring resistance in the field is essential for management and ensuring the longevity of current insecticide based vector control. Conventionally, resistance is suspected in field populations through observation of reduced insecticide efficacy and/or suggested by simple bioassay procedures on wild caught insects. Once discovered, genes or genetic alterations conferring resistance and the frequency, dominance and penetrance within a population can be assessed as these factors will determine the emergence and subsequent spread of resistance.

The monitoring of vector resistance must be globally standardised to ensure comparability of data from different locations (WHO, 2006b). The WHO has published standardised test procedures (bioassays) for detecting and monitoring insecticide resistance in malaria vectors (WHO, 1998b) as well as a standardised manual outlining biochemical and molecular techniques for detecting the mechanisms involved in resistance (WHO, 1998a); although detecting these mechanisms in the field can often be problematic.

Assays for detection of known point mutations conferring target site resistance are also widely available. PCR based approaches have been developed for the detection of both *kdr* associated mutations in *An. gambiae* (Martinez-Torres *et al.*, 1998; Ranson *et al.*, 2000a) as well as insensitive *AChE* via PCR-RFLP (Weill *et al.*, 2004). With field application in mind, more 'low-tech' techniques such as Hot Ligation Oligonucleotide Assays (HOLA) (Lynd *et al.*, 2005; Rajatileka *et al.*, 2008) have also been developed. While recently, high through-put and potentially more reliable techniques for detection involving fluorescent Real-Time PCR assays such as TaqMan® have become available (Bass *et al.*, 2010; Bass *et al.*, 2007). However prior knowledge of the mutations involved is essential for this type of detection assay.

Screening for metabolic resistance in the field is potentially more problematic due to a lack of information on the exact molecular mechanisms involved and the resultant absence of quick diagnostic tests for modified enzyme activity. Current biochemical procedures rely on microtitre plate assays monitoring the enzymatic conversion of a model substrate to product, often with a resultant colour change. The use of synergists such as piperonyl butoxide (PBO) can suggest monooxygenases based resistance through P450 inhibition (Jones, 1998), and S,S,S-tributylphosphorotrithionate can be employed as an inhibitor/indicator of esterase and glutathione S-transferase resistance mechanisms. These techniques provide an indication of enzyme groups potentially involved, however they lack specificity and sensitivity.

Current molecular and biochemical detection methods fail to address the potential for novel mechanisms of resistance. Exploratory studies employing transcriptional profiling, quantitative trail loci (QTL) and association mapping have provided insight into the genes and pathways involved in insecticide resistance. Expression analysis using microarray platforms to compare resistant and susceptible mosquito populations have been employed extensively in *An. gambiae* (David *et al.*, 2005; Muller *et al.*, 2008b; Muller *et al.*, 2007; Djouaka *et al.*, 2008), as well as in *An. funesus* (Wondji

unpublished), *An. arabiensis* (Muller *et al.*, 2008a), *An. stephensi* (Vontas *et al.*, 2007), *Ae. aegypti* (Strode *et al.*, 2006; Poupardin *et al.*, 2008), and *Culex* species (Wu *et al.*, 2004; Komagata *et al.*, 2010; Liu *et al.*, 2007). The majority of these experiments have taken a candidate gene approach employing gene families previously implicated in metabolic resistance, based on an initial Detox microarray design (David *et al.*, 2005). A number of genes have been identified from these profiling studies, with recombinant protein assays supporting an involvement in resistance in some instances (Muller *et al.*, 2008b; Chiu *et al.*, 2008). However, a candidate gene approach may miss potentially novel resistance mechanisms involving genes not previously implicated in resistance.

Microarray technology also fails to address the issue of allelic variation, providing no information on polymorphisms within genes which may alter protein behaviour and associated resistance phenotype. For detection of mutated genes, techniques such as transcriptional profiling using next generation sequencing technologies (Schuster, 2008) are more appropriate, though at present more costly.

Information on regions of the genome that are important in conferring resistance can also be gathered from quantitative trait loci mapping which has been used to successfully identify chromosomal regions linked to a resistance phenotype such as P450 (Ranson *et al.*, 2004; Wondji *et al.*, 2007) and GST gene clusters (Ranson *et al.*, 2000b). This approach involves identification of genetic markers which are statistically measured for association to a particular quantitative phenotype through genetic crosses (The Complex Trait Consortium, 2003). Although the use of colonised inbred material required for most QTL approaches may not be representative of resistance in the field, and fine scale mapping of loci implicated in a phenotype can be a laborious task.

The non-specific nature of biochemical assays and the time consuming and costly nature of microarray and QTL analysis means that at present, high through-put screening of vector populations for metabolic resistance is unavailable.

One area of research which may address this problem involves single nucleotide polymorphism (SNP) identification and association mapping. Through genomic screening of resistant and susceptible vector populations, polymorphic markers (SNPs) statistically associated with resistance can be identified. One such approach was taken by Weetman *et al.* (2010) in which SNPs were first identified via sequencing of wild caught African *An. gambiae s.s.* of diverse geographical origin, concentrating on genes putatively associated with resistance (Wilding *et al.*, 2009). Approximately 1500 of the SNPs identified were then successfully employed in an association mapping study, identifying resistance linked polymorphisms through screening ~1500 field samples phenotyped for pyrethroid resistance (Weetman *et al.*, 2010). However a number of issues reduced the power of this study including low levels of linkage disequilibrium (LD) and population stratification related to co-occurrence of the M and S molecular forms.

Once resistance associated SNPs are identified, simple screening assays can be developed for use in monitoring metabolic resistance in field populations. Sequencing

of two pyrethroid resistance associated P450 genes in resistant and susceptible strains of *An. funestus* identified SNPs which segregated with phenotype (Wondji *et al.*, 2009). These polymorphisms may provide markers for DNA based diagnostic tests if this association is also present in wild populations. Single nucleotide polymorphism approaches do, however, require initial identification of SNPs which can be costly and tedious. However new techniques such as Restriction site Associated DNA (RAD) or next generation sequencing may provide methods for simultaneous SNP discovery and typing, reducing both time and cost (Mardis, 2008; Baird *et al.*, 2008).

Polymorphisms within regulatory regions rather than in up-regulated genes may be fundamental in conferring certain resistance phenotypes. Detection of important regulatory regions may rely on expression QTL (eQTL) mapping which identifies genomic regions associated with elevated expression levels (Gilad *et al.*, 2008). Ruden *et al.* (2009) took an eQTL approach to identify regulatory elements behind neurotoxin induced gene expression in *Drosophila*, identifying both *cis* and *trans* regulatory factors using inbred insect lines. Alternatively, *cis/trans* regulation can be inferred by comparing abundance and ratio of species specific transcripts in parental and F1 hybrids of strains showing differential expression; this approach was employed by Wittkopp *et al.* (2004) in *Drosophila*. Although again, both approaches are labour intensive and employ inbred colony material.

1.2.5 Strategies to overcome insecticide resistance

Resistance conferring mutations usually come with a fitness cost to the animal; for instance mutations within gene *ace-1,* associated with OP and carbamate resistance, have been linked to pupal mortality, longer development times and shorter wing length in mosquitoes (Djogbenou *et al.*, 2010; Bourguet *et al.*, 2004). While resistance to *Bti* was associated with reduced survival, slower development and lower insect mass in a number of insect species (Gassmann *et al.*, 2009). Therefore in the absence of selection pressure the mutations and resistance should be lost.

It is this reversion to susceptibility which is fundamental in resistance management and control strategies which can rely on the rotation/mosaic of insecticides possessing different target sites (Rodriguez *et al.*, 2006; Hougard *et al.*, 2003) or the selected application of insecticides in peak insect seasons (Lenormand *et al.*, 1999). The confinement of OP application to summer months in Southern France resulted in a reduced *ace-1*^{R} frequency in *Culex pipiens* during the winter (Lenormand *et al.*, 1999). While a switch to carbamate compounds for IRS on Bioko Island controlled the *An. gambiae s.s.* population after IRS with pyrethroids potentiated target site resistance and failed to reduce numbers (Sharp *et al.*, 2007). It is feasible that if resistance to carbamate insecticides becomes an issue, a switch back to pyrethroids on Bioko Island may provide control if the *kdr* frequencies have reduced sufficiently. A large scale field trail in Mexico between 1996-2002 evaluated insecticide rotation, mosaic and single compound use for control of malaria vector *An. albimanus* via IRS. High level resistance development was reduced and resistance kept at low levels through rotations and mosaic applications compared to single insecticide use (Rodriguez *et al.*,

2006); however DDT resistance levels did not decline overtime when DDT selection pressure was reduced.

Insecticide resistance reversion can be slow especially if, as suspected in Mexico, the negative selection associated with resistance has been counterbalanced in the population through other genetic alterations. For instance DDT resistance in *Drosophila melanogaster* has almost reached fixation globally despite DDT being withdrawn for control (Catania *et al.*, 2004); although it must be noted that the DDT-R locus does confer resistance to other insecticides currently in use (Daborn *et al.*, 2002; Daborn *et al.*, 2001). An investigation into the fitness cost of the DDT-R locus revealed that when maternally inherited the resistance loci actually *increased* fitness through fecundity and developmental advantages (McCart *et al.*, 2005); however a closely linked gene conferring fitness advantages could not be ruled out.

Moreover, FFrench-Constant argues that as many resistance associated mutations predate insecticide use, for instance the *W251L* esterase mutation associated with malathion resistance in the sheep blow fly (Hartley *et al.*, 2006), they are unlikely to carry large fitness costs and so would be maintained for longer in the population when selection pressure is relieved (Ffrench and Richard, 2007). Resistance associated fitness costs may also be overcome by gene duplication providing a functional version of a mutated gene, as suspected for the *ace-1* gene in *An. gambiae s.s* (Djogbenou *et al.*, 2008a) and *Culex pipiens* (Bourguet *et al.*, 1996; Lenormand *et al.*, 1998).

In light of increasing resistance, one way of ensuring the future of insecticide based control is to focus on improving and developing insecticidal compounds. The production of novel and improved insecticidal formulations is a major objective of the Innovative Vector Control Consortium (IVCC), a public-private enterprise which aims to reduce mosquito borne disease transmission through effective vector control (Hemingway *et al.*, 2006). As information on the mechanisms of resistance becomes available there is opportunity for rational design of new insecticides and re-design of current compounds, ensuring new target sites and reduced susceptibility to enzymatic breakdown. For instance pyrethroid structures could be altered so that they less susceptible to P450 mediated metabolism. Indeed, IVCC insecticide development at LSTM involves testing new compounds against resistant mosquito strains as well as recombinant enzymes linked to resistance (M. Paine pers. comm.). This approach is similar to that taken during human drug development in which drugs are tested with human P450s to prevent drug-interaction and unwanted metabolism *in vivo* (Ansede and Thakker, 2004). However the production of a single new insecticidal compound is very costly, estimated at ~$70 million, which is almost half the annual amount spent on insecticides for public health use (Hemingway *et al.*, 2006). This limits investment from the commercial sector.

A final strategy to overcome insecticide resistance is to seek non-insecticidal vector control. A promising alternative to chemical control is the use of genetically altered mosquitoes potentially refractory to parasite infection (Kokoza *et al.*, 2010; Yoshida *et al.*, 2007) or modified insects which help to reduce mosquito densities in the wild

through induced mortality or sterility on mating (Catteruccia *et al.*, 2009; Fu *et al.*, 2010; Phuc *et al.*, 2007). Recently a proof of principle trial for a RIDL® technique has taken place on Grand Cayman employing transgenic *Ae. aegypti* which carry a repressible dominant lethal gene causing mortality in offspring on mating with wild females (OXITEC, 2010). Approximately 60-80% population suppression was recorded in a 15 hectare area of the island from the release of ~5000 transgenic males/week/hectare over 6 months (A. Harris pers. comm.). However the costs and logistics of such techniques are a major concern as well as public acceptance of 'GM insects'. Moreover, RIDL® techniques will be much more difficult to transfer to malaria vectors for a number of reasons including the complexity of species and molecular forms involved in transmission, the inability to desiccate eggs for storage and transport, and the current lack of suitable transgenic lines. Additionally these techniques are perhaps more suited to isolated populations such as those found on islands, where high levels of wild type mosquito migration is less of an issue.

Ultimately to control resistance and maintain effective vector control we must understand the mechanisms behind the phenotype and develop high through-put monitoring systems for detection in the field. Furthermore, with better understanding of the molecular causes of resistance the development of new chemical based control can be rational and informed.

1.3 Study aims/strategies

Investigation of insecticide resistance in mosquitoes to date has predominantly taken a candidate gene(s) strategy concentrating on genes and gene families previously associate with the phenotype (Muller *et al.*, 2008b; Muller *et al.*, 2007; Strode *et al.*, 2006; Djouaka *et al.*, 2008; Marcombe *et al.*, 2009; David *et al.*, 2005). Though notable success has been achieved using this focused approach, novel mechanisms and pathways of resistance may have been overlooked. This study aimed to take a more objective approach to characterise the genetic causes of insecticide resistance in the major malaria vector *An. gambiae s.s.*.

To ensure that this work was relevant to vector control in the field, wild caught mosquitoes from a malaria endemic country were employed. In Ghana, West Africa, high levels of pyrethroid resistance have already been documented in *An. gambiae s.s.* populations (Muller *et al.*, 2008b; Muller *et al.*, 2007; Adasi and Hemingway, 2008; Coetzee *et al.*, 2006; Klinkenberg *et al.*, 2008); therefore resistance to two candidate insecticides, DDT and the carbamate bendiocarb, likely to replace pyrethroids for indoor residual spraying was investigated.

The primary objective of this work was to enhance the understanding of insecticide resistance through profiling gene expression associated with resistance phenotypes. Identifying the genetic causes of resistance is necessary for informed insecticide development and successful resistance management. The study design was as follows:

1. Field collected *An. gambiae s.s* from the urban Accra region of Ghana were characterised for resistance to bendiocarb and DDT. The molecular form and presence/absence of known target site mutations were also assessed as these factors may influence the resistance phenotype. (Chapter 2)
2. Whole genome microarrays were designed and employed to establish the gene expression profiles of both DDT and bendiocarb resistant *An. gambiae s.s.* mosquitoes. (Chapters 3, and 5)
3. The function of a selection of candidate genes was validated using alternative expression analysis and recombinant protein assays. (Chapters 4 and 5).

The secondary objective of the study was to address the neglected role of allelic variation in the insecticide resistance phenotype. Polymorphisms influencing insecticide resistance will provide markers for resistance detection in the field.

4. The implications of allelic variation within an enzyme previously linked to DDT resistance, *GSTe2*, were investigated through recombinant protein expression and enzymatic assay. (Chapter 6)

Chapter 2: Field collections and resistance phenotyping of *An. gambiae s.s.* from Ghana, West Africa

2.1 Abstract

An. gambiae s.s. populations from southern Ghana, Africa are known to be highly resistant to pyrethroid insecticides; the class of insecticide used for bednets and increasingly for indoor residual spraying. This resistance may, in due course, necessitate a switch to other insecticidal classes so in the present study we describe the patterns of resistance to two likely candidate compounds, bendiocarb and DDT. Resistance to both compounds was recorded in wild caught _An. gambiae s.s._ from the Accra region of Ghana. After one hour exposure to 0.1% bendiocarb an LT_{50} of ~1 hour was recorded while DDT resistance was exceptionally high with only 33% mortality after 6 hours exposure to 4% DDT. Molecular level analysis revealed both the M and S molecular forms of _An. gambiae s.s._ and the presence of target site mutations _kdr_ and _ace-1^R_, associated with DDT/pyrethroid and carbamate/organophosphate resistance respectively. Both target site mutations were significantly associated with molecular form, with mutations most commonly found in the S forms. The presence of _ace-1^R_ was also significantly associated with bendiocarb phenotype, suggesting a role in resistance. However the presence of _kdr_ was not associated with phenotype, which perhaps reflects low levels of variation in _kdr_ frequency. To investigate the molecular basis of this resistance, samples were selected based on phenotype and genetic background for whole genome transcriptional profiling.

2.2 Introduction

The Republic of Ghana in West Africa is bordered by Cote d'Ivoire to the west, Togo to the east and Burkina Faso to the north. Ghana has an estimated population of 23,837,000 (United-Nations, 2009) of which over 18% live in the greater metropolitan area around the capital, Accra (GhanaStatisticalService, 2010). Endemic mosquito borne diseases include malaria, lymphatic filariasis and yellow fever (CDC, 2010) with *Anopheline* transmitted malaria accounting for over 11% of total deaths (WHO, 2002). The most predominant malaria vector in the region is *An. gambiae s.s.* (Klinkenberg *et al.*, 2008; de Souza *et al.*, 2010) with other potential vectors of the *An. gambiae s.l.* complex (*An. melas and An. arabiensis*) as well as *An. funestus* also present (Appawu *et al.*, 2004; Appawu *et al.*, 1994; Coetzee *et al.*, 2006).

The M and S molecular forms of *An. gambiae s.s.* are found sympatrically throughout Ghana (Yawson *et al.*, 2004; de Souza *et al.*, 2010). The M form are reported to predominate in the northern savanna and coastal regions, such as Accra, while the S form are predominant in the central regions (de Souza *et al.*, 2010). These differences

may reflect preferences in breeding habitats. M forms are thought to favour more permanent sites, such as irrigation systems and rice growing areas while S forms are associated with temporal aquatic habitats and are therefore more reliant upon rainfall (Diabate *et al.*, 2003; Toure *et al.*, 1998).

A number of insecticide-based vector control efforts have been implemented in Ghana; these include insecticide treated net (ITN) distribution both via public-private partnerships (NetMark, 2010) and as an adjunct to national childhood immunisation programmes (WHO, 2005b). Indoor residual spraying (IRS) using the organophosphate pirimiphos-methyl (WHO, 2007a) and the pyrethroid alpha-cypermethrin have also been performed (RTI, 2008). Ghana is a member of the Presidents Malaria Initiative (PMI) which in collaboration with the Ghanaian National Malaria Control Programme advocates ITNs distribution and IRS with monitoring and surveillance systems in place (PMI, 2010). In addition to government led programmes, private sector providers offer IRS in urban settings (E. Klinkenburg pers. comm.) and residents themselves employ control through mosquito coils and sprays (Klinkenberg *et al.*, 2008). Mosquito vectors may also be exposed to a variety of insecticides employed to control crop pests in both rural areas and sites of urban agriculture (Ntow *et al.*, 2008; Ntow *et al.*, 2006).

This widespread use of insecticides has resulted in the development of resistance to a number of compounds used routinely in vector control. The numerous reports of resistance to pyrethroids, the only class recommended by the WHO Pesticide Evaluation Scheme (WHOPES) for use on ITNs, in Ghanaian *Anopheles* populations is of particular concern. Coetzee *et al.*(2006) reported that wild caught adult *An. gambiae s.s.* S forms from the south central Ashanti region of Ghana exhibited WHO defined resistance to various pyrethroid insecticides including two type II compounds, deltamethrin and lambda-cyhalothrin. Resistance to permethrin, a type I pyrethroid, has been found in southern Ghanaian S forms (Muller *et al.*, 2008b; Klinkenberg, 2006) with mortality rates after 1 hour exposure as low as 16.8% in females. This resistance phenotype was subsequently linked to a combination of target site and metabolic resistance (Muller *et al.*, 2008b; Muller *et al.*, 2007). Although in earlier phenotypic assays, southern Ghanaian *An. gambiae s.s.* did not exhibit such high levels of resistance. Mosquitoes were reportedly susceptible to permethrin and deltamethrin (Kristan *et al.*, 2003); however molecular form was not resolved. Whether this discrepancy reflects rapid evolution of resistance in the intervening years, marked spatial heterogeneity in resistance, or differences in molecular form composition is unknown.

Some resistance to other insecticides has also been documented with Coetzee *et al* (2006) reporting only 30.8% mortality following standard DDT bioassays in S forms, though sample sizes were small. In contrast Kristan *et al.* (2003) did not find evidence of DDT resistance in south western *An. gambiae s.s.*. Carbamate and dieldrin resistance have also been reported in Ghanaian S forms (Coetzee *et al.*, 2006; Brooke *et al.*, 2006).

The potential for highly, and indeed multiple, insecticide resistant vector populations is a concern for any vector control programme. Ideally insecticide susceptibility studies should precede interventions as was the case in Ghana where alpha-cypermethrin was chosen for IRS only after efficacy studies had been performed (PMI, 2010). With few compounds available for use in vector control new insecticides must be sought to manage resistant mosquito populations. In short, existing insecticides must be used correctly which means tailoring their use according to vector resistance and rotating or combining insecticides to prevent long periods of selection with a single compound. Sharp *et al.*(2007) illustrated that in the presence of DDT/pyrethroid target site resistance (*kdr*) changing from pyrethroid to carbamate based IRS brought the vector population under control in Bioko, Equatorial Guinea. Similarly 'two-in one' bednets containing a combination of pyrethroid and carbamate insecticides were found to be more effective against *An. gambiae* compared with pyrethroids alone in field trials (Guillet *et al.*, 2001). This combination principle has been developed commercially with Vestergaard Frandsen recently launching a combination ITN (Permanet 3.0) containing the pyrethroid deltamethrin and a synergist piperonyl butoxide (PBO) in an attempt to combat metabolic resistance to pyrethroids.

To enable the development of new insecticides as well as systems for monitoring resistance in the field it is essential to understand the mechanisms behind the phenotype. By uncovering the genes and pathways involved in resistance one can rationally design compounds which will not succumb to metabolic breakdown or which target new sites within the mosquito. Molecular markers of resistance are also necessary for field monitoring to provide 'early warning systems' for resistance development so insecticides can be switched before control efficacy is compromised.

With this in mind, it was decided to phenotype wild caught mosquitoes from the urban Accra region of Ghana for resistance to two insecticides, the organochlorine DDT and the carbamate bendiocarb. These two compounds are likely candidates to replace pyrethroids for IRS in the face of increasing pyrethroid resistance in Ghana. To date there are no published reports of resistance to these compounds in this urban region of Ghana and inconsistencies in documented resistance country wide (Coetzee *et al.*, 2006; Kristan *et al.*, 2003). Once phenotyped, individuals will be subject to target site resistance mechanism screens and expression profiling to identify key genes/pathways associated with the resistance phenotype.

Expression profiling of resistant mosquitoes in the past has predominantly relied on laboratory colonies selected with insecticides for high levels of resistance (David *et al.*, 2005; Vontas *et al.*, 2007; Muller *et al.*, 2007). The use of lab colonies, though convenient, is not ideal and may not represent the full spectrum of resistance mechanisms circulating in wild populations. For this reason wild caught mosquitoes from a malaria endemic region where resistance is emerging were selected for this study as they are representative of populations transmitting *Plasmodium* in the field. Furthermore, phenotypic data could aid decision making for future control.

2.3 Methods

2.3.1 Field collections

Mosquito larval collections were performed between 9th May – 30th May 2008 using the standard dipping method (Service, 1993). Eight breeding sites in the Greater Accra region were located by searching for water bodies in areas known to harbour *An. gambiae* mosquitoes from a previous study (Klinkenberg, 2006). The eight individual collection sites were subsequently grouped into four main collection areas based upon proximity. At each site GPS coordinates were recorded for map construction (Figure 2.1). Details of the collection sites together with illustrative photographs are given in Table 2.1 and Figure 2.2-Figure 2.11.

Figure 2.1. Map of the Ghanaian study area.

Collection sites are grouped into four areas and the position of the insectaries and laboratory at the Atomic Agency is also shown. Major water ways are indicated in blue, main roads in grey and areas of urban agriculture in green. The lower black line represents the coast, (Original GPS scaffold courtesy of E. Klinkenburg).

2.3.1.1 Collection site information

Table 2.1. The location and characteristics of the larval sites sampled in the Accra region of Ghana.

Area	Site	GPS coordinates	Description	Collection dates
ONE	Alajo	05.60327°N, 000.21218°W	Dirt track close to brick factory; clean shallow pools and hoof prints mainly harbouring early instar larvae.	15.05.08, 30.05.08
	Kotobabi	05.59508°N, 000.20879°W	Residential/retail area; collected large numbers of all *An. gambiae* stages from roadside gutters where *Culex* species were also present. Water contained organic contaminants, possibly pollutants from adjacent businesses (car garage) and housing.	15.05.08, 19.05.08, 28.05.08
	Roman Ridge	05.60833°N, 000.19593°W	Area of urban agriculture; collected a few larvae from pools of clean water containing grassy vegetation.	19.05.08, 30.05.08
	Dzorwulu	05.61640°N, 000.20206°W	Back road behind a busy highway; collected many mixed age larvae from clean shallow pools with some vegetation present.	30.05.08
TWO	Labadi	05.57304°N, 000.16089°W	Site of a refuse tip, very dirty foul smelling water; collected lots of pupae and 1st instars implying a fast pace of development.	22.05.08, 28.05.08
THREE	Korle Bu	05.53804°N, 000.23168°W	Teaching hospital site; collected from shallow clean road side pools containing vegetation and from man-made water holes in the residential area.	20.05.08, 28.05.08
FOUR	Madina	05.67352°N, 000.17754°W	Residential area; collected from shallow foul smelling pools containing rotten refuse. Mainly *An. gambiae* although some *Culex* present.	12.05.08
	Redco	05.68168°N, 000.17992°W	Residential area; collected many *An. gambiae* all stages from pools and trenches in a building site.	12.05.08, 14.05.08

2.3.1.2 Collection site photos

Area one: Alajo; Kotobabi; Roman Ridge; Dzorwulu

Figure 2.2. Collection site at Alajo.

Collections were made from shallow temporal pools adjacent to a brick factory (right in picture). Water was relatively clean containing a mixture of *An. gambiae* life stages.

Figure 2.3. Collection site at Kotobabi.

Mosquitoes collected from gutters containing run-off from house and businesses. Water was heavily polluted with refuse and organic matter and contained many *An. gambiae* and *Culex*.

Figure 2.4. Typical breeding site at Roman Ridge.

Collections were made at a site of urban agriculture from water collecting in irrigation channels, small ponds and disused building equipment (pictured).

Figure 2.5. Collection site at Dzorwulu.

Mosquitoes were collected from clean pools which were possibly run off from nearby streams. The site was situated on a back road behind a busy main highway.

Area two (Labadi)

Figure 2.6. Typical collection site in Labadi.

Shallow pools in a residential refuse dump were sampled. Water was heavily contaminated containing with rotten garbage and human waste. Mainly young *An. gambiae* instars and pupae collected.

Area three (Korle Bu)

Figure 2.7. Breeding site in Korle Bu area.

An example of a clean pond containing a mixture of aquatic insect species from which *An. gambiae* were sampled.

Figure 2.8. Typical man-made water holes in Korle Bu area.

These water holes contained an abundance of *An. gambiae*.

Area Four (Madina; Zongo)

Figure 2.9. Collection site at Madina.

An. gambiae mosquitoes were collected from shallow pools formed in disused land in a residential area where refuse had been dumped.

Figure 2.10. Polluted collection site in Madina containing *An. gambiae*.

Figure 2.11. Collection site at Redco.

Deep pool on a building site containing clean water from which *An. gambiae* were collected.

2.3.2 Mosquito rearing in field insectaries

Larvae were transported from collection sites in plastic containers to insectary facilities at the Ghana Atomic Energy Commission (GAEC) where they were split into plastic trays containing purified water and maintained on flaked fish food (Tetra). Pupae were picked daily and placed into plastic cages (BugDorm,Taiwan) and allowed to emerge.

To standardise the age of adults for testing, pupae harvested from a maximum of three days were placed in a single cage. Upon emergence adults of both sexes were maintained on a 10% sugar solution until 3-5 days post eclosion. Mosquitoes collected from all sites were predominantly *Anopheles gambiae s.l.* based upon morphology (Gillies and Coetzee, 1987).

2.3.3 Insecticide resistance phenotyping

2.3.3.1 Selection time determination

Mosquitoes from initial collections were used to construct a time response curve to achieve a discriminant, resistant *vs* susceptible, exposure time for DDT and bendiocarb. Non-bloodfed female and male mosquitoes aged 3-5 days were aspirated from cages and placed into WHO resting tubes lined with filter paper. Batches of 20-30 mosquitoes were then exposed to insecticide treated papers, obtained from WHO (WHO, 2001), containing either 4% DDT or 0.1% bendiocarb for a range of times (Figure 2.12 and Figure 2.13). Control assays were performed with papers treated with the appropriate insecticide carrier. Abbott's correction was employed if control mortality was > 5% (Abbott, 1925).

2.3.3.2 Selections

The initial experiments revealed extremely high levels of DDT resistance which prevented calculation of a complete time response curve. So a maximum exposure time, which could be accommodated within a standard working day, was decided upon and mosquitoes were exposed to 4% DDT for six hours. From initial bioassays, 6 hours exposure gave an average mortality of 40% [18-67%, 95% CI] in females and 91% [57-100%, 95%CI] in males. Construction of a time-response curve was also problematic for bendiocarb as there was marked fluctuation in mortality. A one hour standard WHO exposure time was decided upon to provide at least an LT_{50}. Average mortality in females after 1 hour exposure was 73% [56-85%, 95%CI] and in males 63% [44-80%, 95%CI].

Insecticide selections were performed on non-bloodfed adult mosquitoes aged 3-5 days. Mosquitoes were either exposed to 0.1% bendiocarb papers for one hour or 4% DDT for six hours. A separate set of mosquitoes were also exposed to the corresponding control papers for each insecticide (WHO, 2001). Papers were re-used a maximum of 6 times and stored at 4°C between assays. Once exposed, mosquitoes were transferred into resting tubes and provided with cotton wool soaked in 10% sugar solution. Male and female mosquito mortality was recorded 24 hours after exposure. Dead mosquitoes were then removed from the resting tubes using forceps and transferred into individual numbered PCR tubes (punctured to allow air circulation) before placing on silica gel for preservation. Alive mosquitoes were briefly chilled at 4°C to prevent escape during transfer to tubes. Individual females were selected and a hind leg removed using forceps. The body of the mosquito was then submerged in RNA preservative solution, RNAlater® (Ambion), in a 2ml tube labelled with a unique

number, date of the bioassay, collection site and treatment (test or control). The leg was placed into a punctured PCR tube labelled with the corresponding unique identifier and placed on silica gel. RNAlater® samples were chilled overnight at 4°C to allow the solution to penetrate the material before being transferred to -20°C freezer according to manufacturer's instructions.

2.3.4 Molecular characterisation

2.3.4.1 DNA extraction from legs

Each leg removed from female mosquitoes was transferred to a 96 well plate for DNA extraction using a DNeasy Blood & Tissue kit (Qiagen). Individual legs were incubated for three hours in a shaking incubator at 56°C in 200µl of lysis buffer with 20µl of proteinase K. An optional RNaseA step was performed before transferring to a plate of 96 purification columns followed by washing, and elution of DNA in 100µl of elution buffer. Plates were then stored at 4°C for immediate use and later at -20°C for long term storage.

2.3.4.2 Species identification & molecular form determination

A standard PCR protocol exploiting sequence differences in ribosomal DNA was used to identify members of the *Anopheles gambiae s.l.* complex (Scott *et al.*, 1993). Each 25µl PCR contained 1µl of DNA solution from leg extraction or positive control DNA for the members of the species complex (*An. arabiensis; An. melas/merus; An. quadriannulatus; An. gambiae s.s.*). Cycling conditions were 95°C for 15 minutes, 30 cycles of 95°C for 30 seconds, 50°C for 30 seconds, 72°C for 30 seconds, then final extension of 10 minutes at 72°C. Five microlitres of product was loaded onto a 1.5% agarose gel containing 0.5µg/ml ethidium bromide and visualised on a gel imaging system. For specimens identified as *An. gambiae s.s.* an RFLP protocol (Fanello *et al.*, 2002) was used to determine molecular form (M or S). Eight microlitres of the species ID PCR product were digested with 5 units of restriction enzyme *CfoI* (Promega) for 3 hours at 37°C. Post digest 5µl of product were analysed on a 1.5% agarose gel (0.5µg/ml ethidium bromide). The presence of *CfoI* restriction site GCG^C in S forms produces two fragments at 257 and 110bp. The restriction site is absent in M forms and a single 367bp band is observed.

2.3.4.3 Target site mutation characterisation

Samples were screened for the presence of insecticide target site mutations in the voltage-gated sodium channel, *kdr* West (*L1014F*) & East (*L1014S*), and acetylcholinesterase mutation *G119S* using custom TaqMan® assays (*Assay-by-Design* ABI) based on primers designed by Bass *et al.*(2007; 2010) . Each reaction contained 1µl of DNA template, 0.5µl of custom primer and probe mix (ABI), 10µl of SensiMix Probe 2x reaction buffer containing DNA polymerase, dNTPs, $MgCl_2$ (6mM) (Quantace) and nuclease free water up to 20µl. Cycling and fluorescence measurement were performed on a MiniOpticon (BioRad) with 10 minutes at 95°C,

followed by 40 cycles of 10 seconds at 95°C and 45 seconds at 60°C. Positive homozygous, heterozygous and wild-type genomic controls were run alongside samples to aid genotyping, (see Appendix 9 for full assay conditions)

2.4 Results

2.4.1 Selection time determination

2.4.1.1 Bendiocarb

A total of 312 mosquitoes of both sexes were exposed to 0.1% bendiocarb papers for 15, 30, 45 or 60 minutes (a breakdown of exposure time and mosquito numbers can be found in Appendix 1). Mortality curves for male and female mosquitoes according to exposure time were constructed using percentage mortality with 95% confidence intervals (Figure 2.12). Fluctuations in mortalities from initial bioassays produced a mortality curve approaching asymptote before 100%. Rather than using a specific LT (lethal time) we therefore decided upon an exposure time of 60 minutes for selections. This equated to an LT of approximately 70% in females and 65% in males which produced good segregation between resistant and susceptible mosquitoes whilst still providing adequate numbers of resistant females for future work.

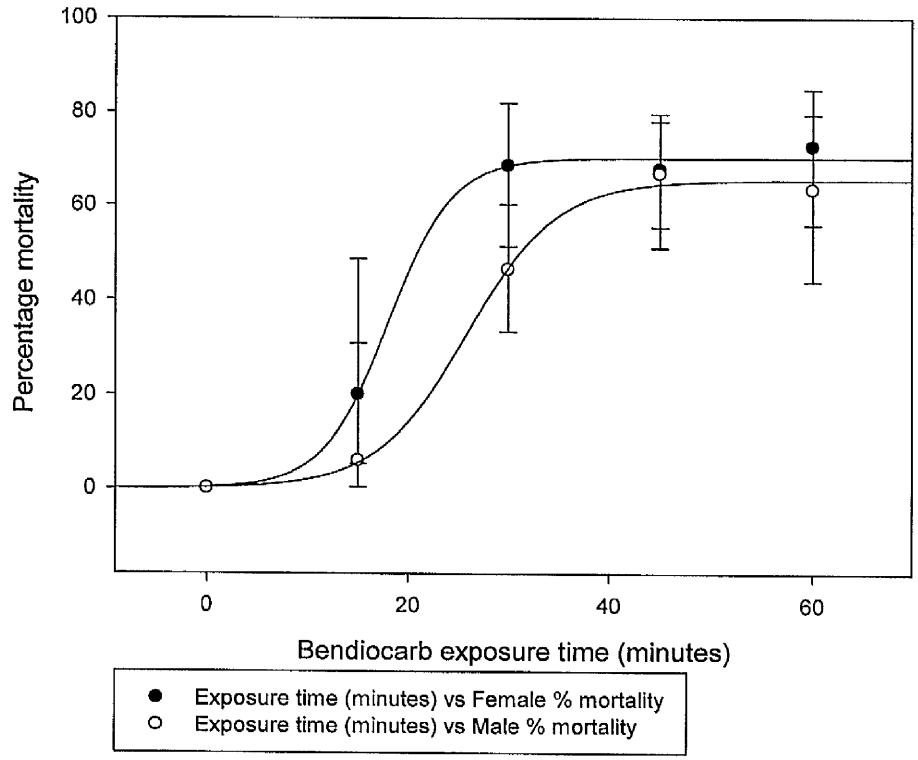

Figure 2.12. Bendiocarb time response curves

Twenty four hour mortality post exposure to 0.1% bendiocarb in male and female *An. gambiae* collected from Madina, Ghana 2008. Error bars indicate 95% confidence intervals (CI) (Newcomb, 1998) calculated in VassarStats (http://faculty.vassar.edu/lowry/prop1.html) with continuity correction. See Appendix 1 for raw bioassay data.

2.4.1.2 DDT

Eighty seven mosquitoes of both sexes were exposed to DDT for either 60, 90 or 360 minutes (Figure 2.13, Appendix 2). Fewer mosquitoes and exposure times were used for DDT selection time determination as it became clear that resistance levels were so high that a maximum time of exposure feasible in a day (360 minutes) would be necessary. Mean mortality after 6 hours exposure was 40% [18-67%, 95% CI] in females and 91% [57-100%, 95%CI] in males.

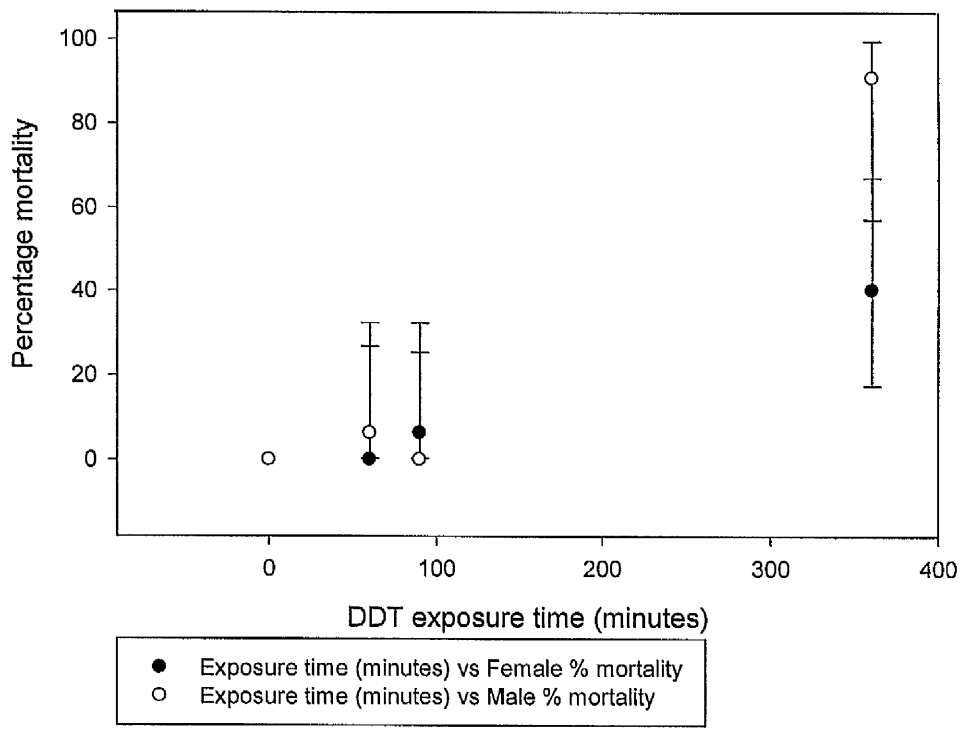

Figure 2.13. DDT time response data.

Twenty four hour mortality post exposure to 4% DDT in male and female *An. gambiae* collected from Madina, Ghana 2008. Error bars indicate 95% confidence intervals (CI) (Newcomb, 1998) calculated in VassarStats (http://faculty.vassar.edu/lowry/prop1.html) with continuity correction. See Appendix 2 for raw bioassay data.

Selections

Hence forth, as only female *An. gambiae s.s.* are vectors of malaria, the female resistance phenotype will be the main focus of discussion; however male phenotype will be commented on for reasons of general interest.

2.4.2 Bendiocarb selections

A total of 765 mosquitoes (560 females, 205 males) were selected by exposure to 0.1% bendiocarb for one hour. Percentage mortality calculated from all four collection areas was 60% [56-64%, 95% CI] in females and 70% [63-76%, 95%CI] in males.

Area two gave the lowest female mortality at 39% [33-46%, 95%CI] while area three exhibited the highest mortality at over 98% in both sexes (See Appendix 3 for breakdown of mortality according to sex and area).

2.4.3 DDT selections

A total of 582 mosquitoes underwent DDT selection (exposure to 4% papers for six hours), comprising of 446 females and 136 males. Overall percentage mortality in females was 24% [20-28%, 95%CI] while male mean mortality was 61% [55-69%, 95%CI], although again results from each of the four areas were variable. Lowest female mortality, 18% [12-27%, 95%CI], was recorded at area three while area four had the highest female mortality (40% [18-67%, 95%CI]). However sample size for this area was low reflected in the wide confidence interval. (See Appendix 4 for breakdown of mortality according to sex and area).

2.4.4 Sample recording and storage

Only female mosquitoes surviving insecticide or control tube exposure were stored for RNA extraction 24 hours post exposure (n=1052, Table 2.2) with all dead mosquitoes from bioassays stored on silica gel.

Table 2.2. Total numbers of bendiocarb and DDT phenotyped *An. gambiae s.l.* females stored for gene expression analyses.

Specimens were stored in RNAlater® (Ambion) and grouped by collection area. Test samples are females surviving exposure to the discriminant dose for each insecticide; 1 hour exposure to 0.1% bendiocarb or 6 hours exposure to 4% DDT. Controls are age-matched female mosquitoes exposed to corresponding control papers for the same duration as the test females.

AREA	Bendiocarb		DDT	
	Test	Control	Test	Control
ONE	68	44	160	106
TWO	143	32	81	90
THREE	1	120	91	65
FOUR	11	31	9	0
Total	223	227	341	261

2.4.5 Species identification and molecular form determination

In total, 669 DNA extractions were performed from individual legs from female mosquitoes recorded in Table 2.2 to provide adequate genotyped material for subsequent microarray experiments. All samples were *An. gambiae sensu stricto* and an example of a diagnostic agarose gel is shown in Figure 2.14.

Of the 669 females identified as *An. gambiae s.s.*, 310 individuals were M form, 356 were S form and 3 were suspected M/S hybrids. The putative hybrids have yet to be confirmed via sequencing. An example of an M/S diagnostic gel is shown in Figure 2.15. Various proportions of M and S were recorded from each site with area two collections being predominantly S form (93%, [88.0-96.2, 95%CI]), while site three collections were 92% M form [85.0-96.0, 95%CI] (see Appendix 5 for a breakdown of molecular forms by collection site).

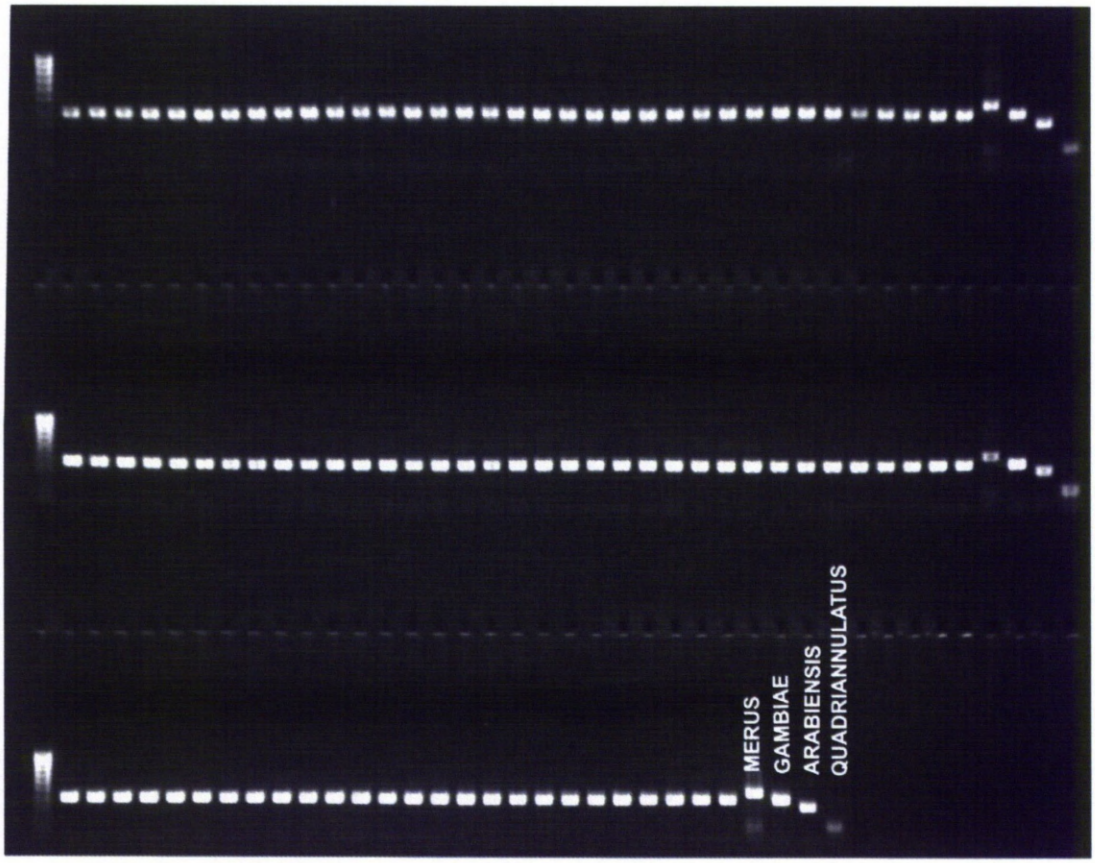

Figure 2.14. Species identification of *An. gambiae s.l.* mosquitoes from Accra, Ghana using diagnostic PCR.

Five microlitres of species diagnostic PCR based on the method of Scott *et al.* (1993) were loaded onto a 1.5% agarose gel (0.5µg/ml ethidium bromide). A 1Kb Hyperladder IV (Bioline) was loaded in the left hand lane of each row and positive controls (*An. merus; An. gambiae; An. arabiensis & An. quadriannulatus)* run in the last four wells of each row.

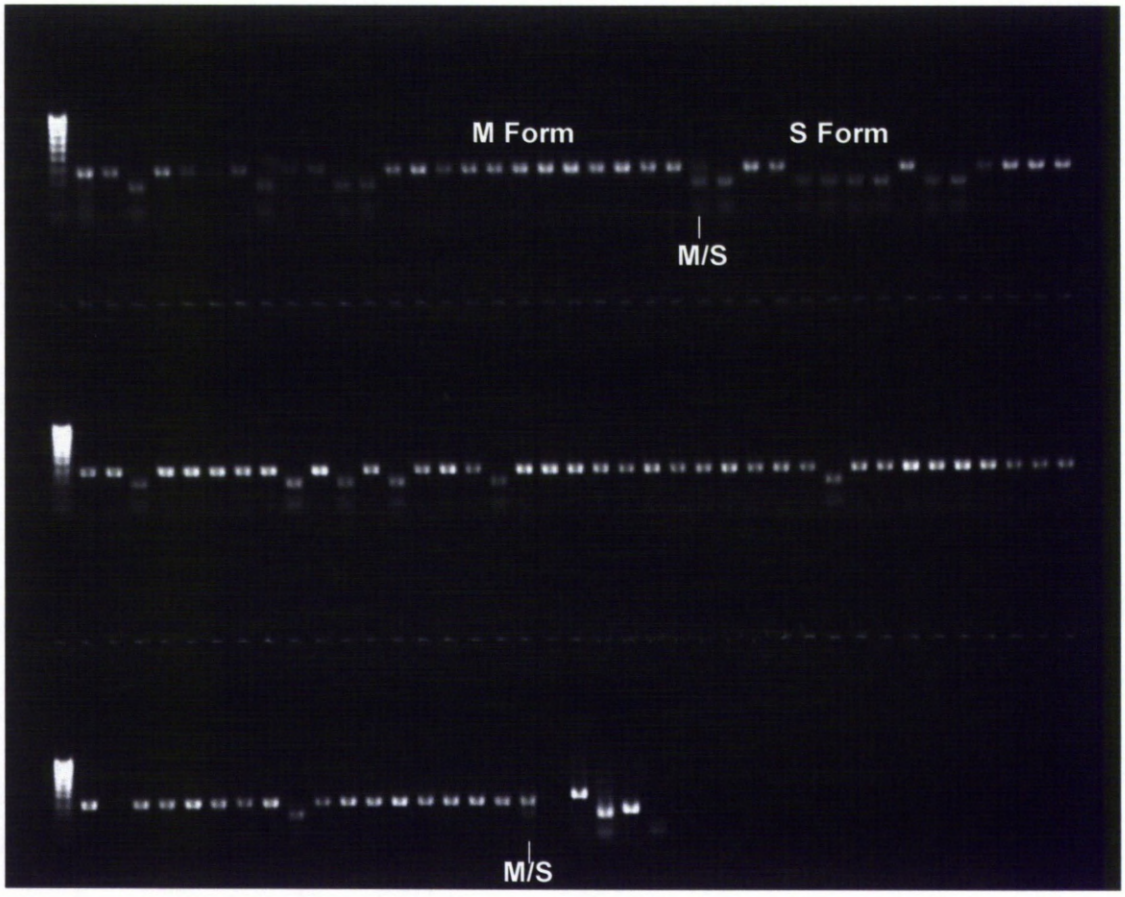

Figure 2.15. Molecular form determination of *An. gambiae s.s.* using a restriction fragment length polymorphism (RFLP) protocol (Fanello *et al.*, 2002).

Five microlitres of a *CfoI* digested PCR product was loaded onto a 1.5% agarose gel (0.5µg/ml ethidium bromide). M forms were characterised by a single band at 367bp, while S forms yielded two bands at 257 and 110bp and putative hybrids contained all three bands. A 1Kb ladder (Hyperladder IV, Bioline) was run in left hand lane.

2.4.6 Target site mutation characterisation

2.4.6.1 Insensitive *ace*-1

In total three hundred and thirty-three bendiocarb phenotyped or control female mosquitoes were characterised for the presence of the *G119S* mutation in the acetylcholinesterase gene (*ace1*) using a TaqMan assay (Bass *et al.*, 2010). A subset of individual *ace1* genotypes (81) were confirmed via sequencing and 100% concordance with the TaqMan® assay recorded (D. Birks, pers. comm.). Distribution of the *G119S* target site mutation varied with both molecular form and phenotype (

Table 2.3; Figure 2.16).

Table 2.3. The frequency of the *G119S* mutation in the *ace-1* gene of *An. gambiae s.s.* females collected from Ghana, West Africa.

Genotypes were determined using the TaqMan assay of Bass *et al.* (2010) and are shown according to molecular form (M or S) and bendiocarb phenotype (Test or Control).

	S-FORM		M-FORM		
G119S genotype	**Control**	**Test**	**Control**	**Test**	**Total**
RR	10	48	1	1	60
SR	30	88	24	3	145
SS	33	0	94	1	128
Total	73	136	119	5	333

To investigate potential associations between genotype and both molecular form and resistance phenotype, contingency table analysis was performed in RxC (http://www.marksgeneticsoftware.net/rxc.htm) employing the Metropolis algorithm (Metropolis *et al.*, 1953). A significant association between *An. gambiae s.s.* molecular form and the presence of the *G119S* mutation was observed in un-exposed mosquitoes with the serine allele at a higher frequency in the S forms (0.34) compared with M forms (0.11) ($p<0.001$).

Due to an association between molecular form and genotype, the M and S forms were split before investigating links between *G119S* and phenotype to avoid confounding results. In both S forms ($p< 0.001$) and M forms ($p<0.01$) a significant relationship between *G119* genotype and bendiocarb phenotype was observed. Higher proportions of RR and SR individuals were recorded in tested mosquitoes compared with controls (Figure 2.16); although the total number of M form test mosquitoes was low (n=5).

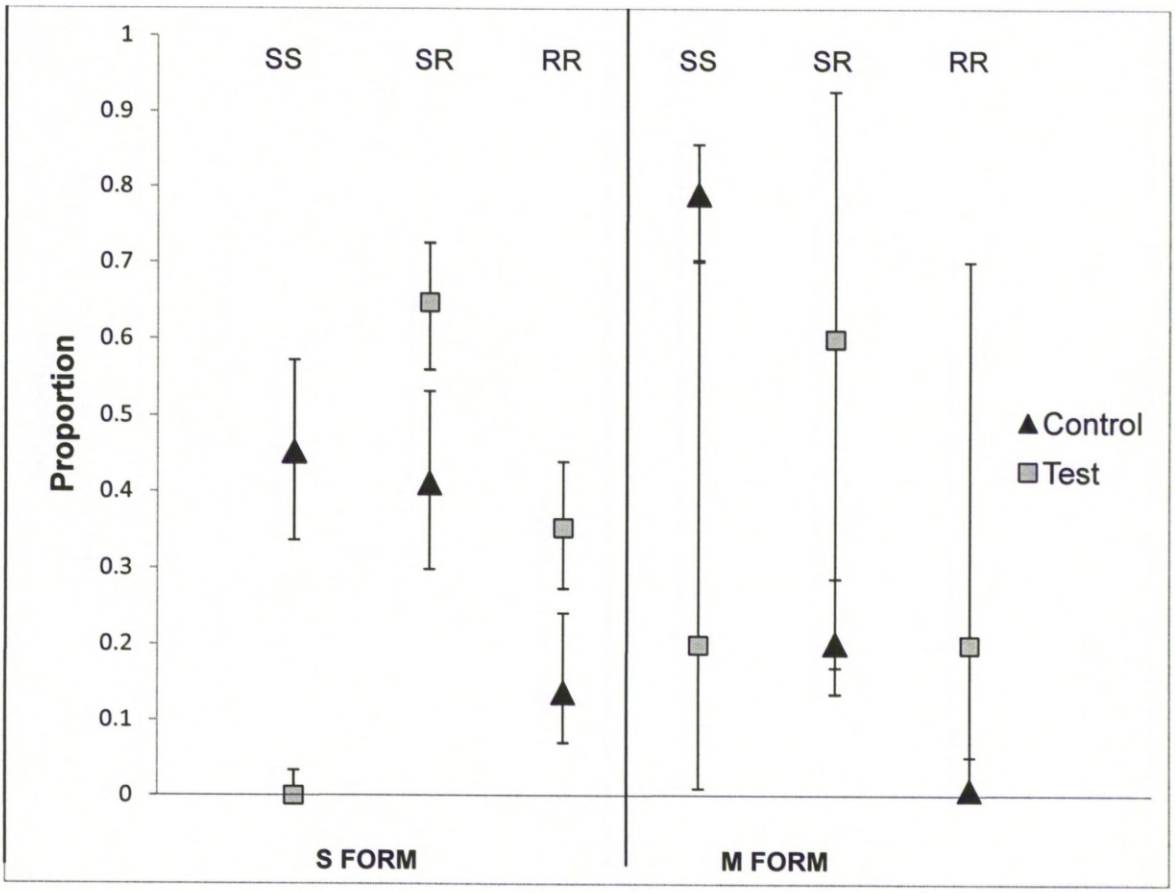

Figure 2.16. The association between *ace-1* genotype and bendiocarb susceptibility in female *An. gambiae s.s.* of the S and M molecular form.

Genotypes at the 116 codon of the *ace-1* gene are given at the top of the figure with SS representing wildtype mosquitoes, RR representing *G119S* resistant homozygotes and RS heterozygotes. Test individuals are those that survived a 1 hour exposure to 0.1% bendiocarb, control individuals were exposed to control tubes for one hour. Proportions are plotted with 95% confidence intervals (Vassarstats http://faculty.vassar.edu/lowry/VassarStats.html). S form controls n=73, S form test n=136, M form control n=119, M form test n=5.

Interim conclusion

Only five M-form individuals survived a 1 hour exposure to 0.1% bendiocarb. Consequently the gene expression studies detailed in Chapters 3 and 4 focused solely upon S-form *An. gambiae s.s.*. As individuals heterozygous for *G119S* make up the largest proportion of S forms phenotyped (~56%), test and controls from this genotype were pursued. By standardising the target-site genotypes of test and control groups we hope to remove this potential confounder as in an earlier study (Muller *et al.*, 2008b).

2.4.6.2 *KDR* West

Two hundred and eight DDT selected and control females were characterised for the presence of both *L1014S* and *L1014F kdr* mutations using a TaqMan assay (Bass *et al.*, 2010), with a subset of these confirmed via Sanger sequencing. Overall genotypic frequencies are shown in Table 2.4. The *kdr* East allele, *L1014S*, was initially detected in 11 individuals however this was not confirmed via sequencing. To establish whether the *L1014F* mutation was associated with molecular form, contingency table analysis was performed in RxC. The *kdr* West (*L1014F*) mutation was significantly associated ($p<0.01$) with *An. gambiae s.s.* molecular form with higher proportions of homozygous (RR) individuals and an absence of wild-type genotypes (SS) in the S form (Table 2.4, Figure 2.17).

Molecular forms were separated before investigating potential association between DDT phenotype and genotype due to the link between form and the *L1014F* mutation. No association between phenotype (test or control) was uncovered in either M ($p= 0.400$) or S forms ($p=1.00$), with similar proportions of each genotype in test and control groups (Figure 2.17).

Table 2.4. The frequency of the *kdr* West *(L1014F)* mutation in *An. gambiae s.s.* females collected from Ghana, West Africa.

Genotypes were determined using the TaqMan assay of Bass *et al.* (2010) and are shown according to molecular form (M or S) and DDT phenotype (Test or Control).

	S-FORM		M-FORM		
***L1014F* genotype**	**Control**	**Test**	**Control**	**Test**	**Total**
RR	18	20	11	9	58
SR	19	20	41	63	143
SS	0	0	2	5	7
Total	37	40	54	77	208

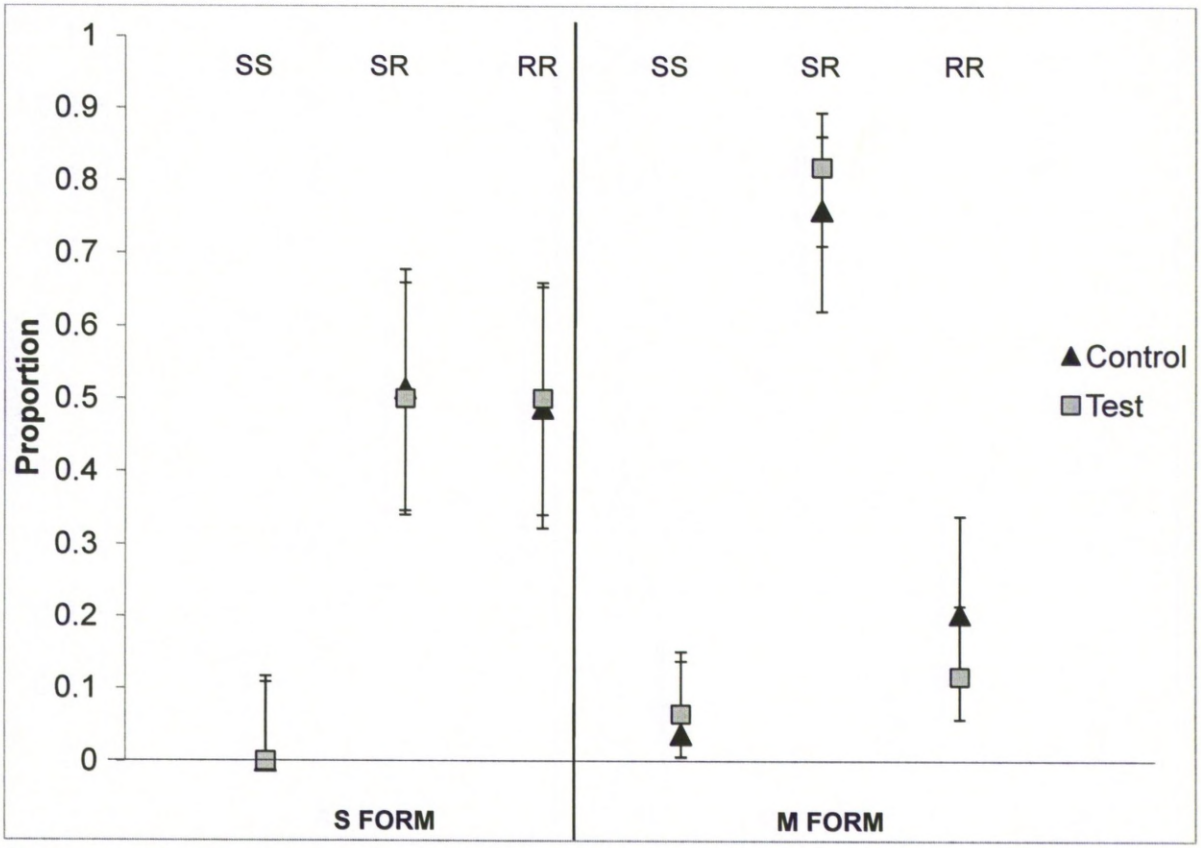

Figure 2.17. The association between *kdr* West *(L1014F)* genotype and DDT susceptibility in female *An. gambiae s.s.* of the S and M molecular form.

Genotypes at the 1014 codon of the sodium channel gene are given at the top of the figure with SS representing wildtype mosquitoes, RR representing *L1014F* resistant homozygotes and RS heterozygotes. Test individuals are those that survived a 6 hours exposure to 4% DDT, control individuals were exposed to control tubes for 6 hours. Proportions are plotted with 95% confidence intervals (Vassarstats http://faculty.vassar.edu/lowry/VassarStats.html). S form controls n= 37, S form test n= 40, M form control n= 54, M form test n= 77.

Interim conclusion

M form mosquitoes were the larger proportion of those phenotyped (63%) and were utilised for the gene expression experiments detailed in Chapter 5. *L1014F* heterozygotes which were the predominant genotype in the M form were preferentially selected for expression analyses to attempt to control for potential confounding effects.

2.5 Discussion

Mosquitoes collected in the Accra region of Ghana, May 2008, were all identified as *An. gambiae sensu stricto* with a mixture of M and S molecular forms present. Although documented in the area (de Souza *et al.*, 2010) other members of the *An. gambiae sensu lato* species complex were absent from collections.

WHO define insecticide resistance as <80% mortality within 24 hours of a one hour exposure to a predetermined dose (WHO, 1998a). This definition is perhaps rather arbitrary as there is little evidence that the presence of WHO defined resistance correlates with insecticidal control failure. The agrochemical industry employs more precise and practical definitions of resistance, with the Insecticide Resistance Action Committee (IRAC) outlining criteria for insecticide resistance (Tomlin, 1997): The compound to which resistance has developed must have a recommended dose against the pest which in the past has been successful; failure for reasons of poor storage, application or environmental factors must be ruled out; the recommended dose must fail to reduce pests below a threshold which in agricultural terms is economic; failure of control must be caused by a heritable change in susceptibility in the population.

Therefore IRAC stipulate that the term resistance is only applicable to field failure of the compound. Though aware of their limitations we employ WHO cut-offs for insecticide resistance in order to be comparable to other studies of resistance in mosquito vectors. Thus according to WHO definition mosquitoes phenotyped for resistance in this study exhibited high levels of resistance to both DDT and bendiocarb insecticides. *An. gambiae* from this area are also known to exhibit resistance to pyrethroid insecticides (Adasi and Hemingway, 2008; Muller *et al.*, 2008a). However the impact of this insecticide resistance on vector control efficacy is unknown. The exact causes of the multiple and high levels of insecticide resistance in these *An. gambiae s.s.* populations remain elusive, however we can identify some potential selection pressures.

Firstly as mentioned previously, Ghana has advocated the use of pyrethroid treated nets (ITNs) and Indoor Residual Spraying (IRS) for a number of years (WHO, 2005b; WHO, 2007a; PMI, 2010). IRS and ITN use have exposed vector populations to pyrethroid and organophosphate (OP) compounds (WHO, 2007a; RTI, 2008). Target site resistance (*kdr*) to pyrethroids confers cross resistance to DDT while point mutations in the organophosphate target, *ace-1*, confers cross resistance to carbamates; it is feasible that selection pressures from pyrethroid and OP use have selected for target site resistance to DDT and the carbamate, bendiocarb. Indeed the percentage of *An. gambiae* reportedly carrying the *kdr* mutation increased from 1% to 85% in 5 years post IRS with alpha-cypermethrin in Burundi (Protopopoff *et al.*, 2008), while increased frequencies of *L1014F* were also recorded in Niger after nationwide ITN distribution (Czeher *et al.*, 2008).

Secondly, in the Accra region intensive urban agriculture which relies on heavy pesticide usage is common, especially around the collection sites in area one (Klinkenberg, 2006), with some known areas of agriculture indicated in Figure 2.1.

Such agricultural practises may expose vectors breeding in those areas to a cocktail of compounds. Indeed Achonduh *et al.* (2008) found high levels of toxicity in soil and water run-off from agricultural areas in Accra, likely to be pyrethroid and organophosphate residues. However no correlation between contamination level and resistance in *An. gambiae* was found; although any relationship was potentially masked by failure to take molecular form distribution into account. Yadouleton *et al.* (2009) report on the rapid increase in urban vegetable cultivation in Benin and the resultant improper use of insecticides putting selection pressure on *An. gambiae* in the area. Insecticide resistance was reportedly higher in mosquitoes from areas of urban agriculture compared with rural rice growing areas where insecticide use is lower (Corbel *et al.*, 2007). Many previous authors have also made connections between the use of pesticides in agriculture and the development of resistance in vector populations (Diabate *et al.*, 2002; Muller *et al.*, 2008a; Lines, 1988). Without an extensive comparison of resistance in mosquitoes from areas of urban agriculture and urban areas without crop cultivation, the link between agricultural pesticide use and high levels or resistance found here is speculative; however the evidence from previous studies supports a role for this type of selection.

Lastly it was noted that many of the urban breeding sites sampled contained heavily polluted water. For instance at Kotobabi (area one) collections were taken from a gutter running along a main road with a car repair garage located opposite (Figure 2.3). The water was putrid containing rotten garbage and run-off from local businesses and homes. While the Labadi site (area two) was located in a refuse dump with water again containing rotten garbage and human waste (Figure 2.6) with similar polluted sites observed in Madina (four) where water contained domestic waste (Figure 2.9 and Figure 2.10). Although no quantitative measurements of water quality were made it was clear from observations that the water was heavily contaminated; exposure to high levels of toxins during mosquito development may select for individuals who are able to withstand these conditions via detoxification or other means. A number of gene families linked to general detoxification of xenobiotics such as cytochrome P450s and glutathione-S-transferases are also linked to insecticide resistance, so the possibility of general breeding site toxicants causing cross resistance to insecticides must be considered. Indeed Poupardin *et al.* (2008) report on the cross induction of detoxification genes by environmental xenobiotics and insecticides in *Aedes aegypti*. Increased tolerance to permethrin was reported after larvae were exposed to copper and hydrocarbon fluoranthene (a constituent of road surfaces likely to be present in urban breeding sites) associated with an induction of cytochrome P450s and GSTs. While Djouaka *et al.* (2007) suggest that Nigerian mosquito populations breeding in areas polluted with petrol show higher levels of pyrethroid resistance to those breeding in cleaner waters. However results from this study are inconclusive due to an absence of mortality confidence intervals and the use of arbitrary WHO resistance definitions which meant differences in mortality of just 2-3% separated some 'resistant' and 'susceptible' populations. *An. gambiae* are known to prefer clean water for breeding with more polluted waters usually associated with *Culex* and some *Aedes* species

(Service, 1993). This could be an example of phenotypic plasticity in urban *An. gambiae* populations whereby mosquitoes are adjusting to cope with polluted breeding sites in the absence of clean water bodies. Indeed *An. gambiae s.l* breeding in polluted areas has previously been described in the Accra region (Klinkenberg, 2006; Chinery, 1984) as well as in Kenya, East Africa (Keating *et al.*, 2004; Mireji *et al.*, 2008). This is an interesting research area which should be pursued further but is outside the realms of this study due to a lack of empirical data on water quality.

2.5.1 Molecular forms

Both M and S molecular forms were found sympatrically in accordance with previous findings (Yawson *et al.*, 2004; de Souza *et al.*, 2010). Two hundred and five M form controls were recorded compared with lower numbers of S form controls (115) (Appendix 6 and 7). Differences in the insecticide resistance profile between molecular forms meant proportions were potentially skewed in tested groups. Yawson *et al.* reported that S forms predominate in coastal regions of Ghana (2004). However extensive distribution analysis collating published and current data contradicts this, with M form *An. gambiae s.s* found to be dominating Ghanaian coastal areas with S forms found mainly in central regions (de Souza *et al.*, 2010). A number of factors may influence the numbers of each molecular form sampled, for instance S forms may be preferentially sampled during the rainy season (April-October) as breeding in this form is associated with rainfall (Diabate *et al.*, 2003; Caputo *et al.*, 2008). While the types of breeding sites sampled may affect collections with M forms preferring more permanent breeding sites (Toure *et al.*, 1994; Toure *et al.*, 1998).

The proportions of M and S collected from each sampled area also varied with some site collections skewed towards one molecular form. Area three (Korle Bu) control mosquitoes were 92% [85-96%, 95% CI] M forms (Appendix 6), while area two (Labadi) phenotyped samples were overall 93% S form; however a lower percentage of S forms (69% [50-83%, 95% CI]) were recorded in the control group alone which suggests a bias introduced by insecticide exposure. These differences may reflect the type of breeding site available in each area and preferences of the two molecular forms (Caputo *et al.*, 2008; Diabate *et al.*, 2003). For instance the water bodies sampled in Labadi were transient shallow water pools (Figure 2.6) likely fed by rainwater while collection sites in Korle Bu were predominantly man-made concrete water-holes (Figure 2.7 and Figure 2.8) which provide more permanent breeding pools. However caution needs to be given when interpreting results of molecular form distribution as this study was not designed to be an extensive entomological survey rather sampling of a specific species (*An. gambiae s.s.*) in these areas over a relatively short period of time (one month).

Insecticide resistance is unevenly distributed between molecular forms of *An. gambiae s.s.* in Sub-Saharan Africa. Ranson *et al.* (2009) report on extensive resistance testing from three Africa countries where the distribution of molecular form post bioassay differed significantly from the population as a whole. The majority of DDT and permethrin survivors and all bendiocarb and fenitrothion survivors were of the S form

suggesting higher levels of resistance in comparison with the M form. The majority of resistance documented in Ghana has involved the S molecular forms of *An. gambiae s.s.* when the two molecular forms are known to occur sympatrically (Coetzee *et al.*, 2006; Muller *et al.*, 2008b; Brooke *et al.*, 2006). However Anto *et al.* (2009) report on the beginnings of permethrin and DDT resistance in M forms originating in Northern Ghana, which may be partially explained by increasing frequency of target site mutation *L1014F* in Ghanaian M forms (Lynd *et al.*, 2010). Results from this study support an unequal distribution of resistance within forms with the highest levels of bendiocarb resistance recorded in the predominantly S form samples from area two, while no WHO defined bendiocarb resistance was found in area 3 which were 92% [85-96%, 95% CI] M form (Figure 2.18, below). A similar absence of bendiocarb resistance in M forms from area four (Madina) was recorded in 2010 with resistance only present in S forms (D. Weetman pers. comm.).

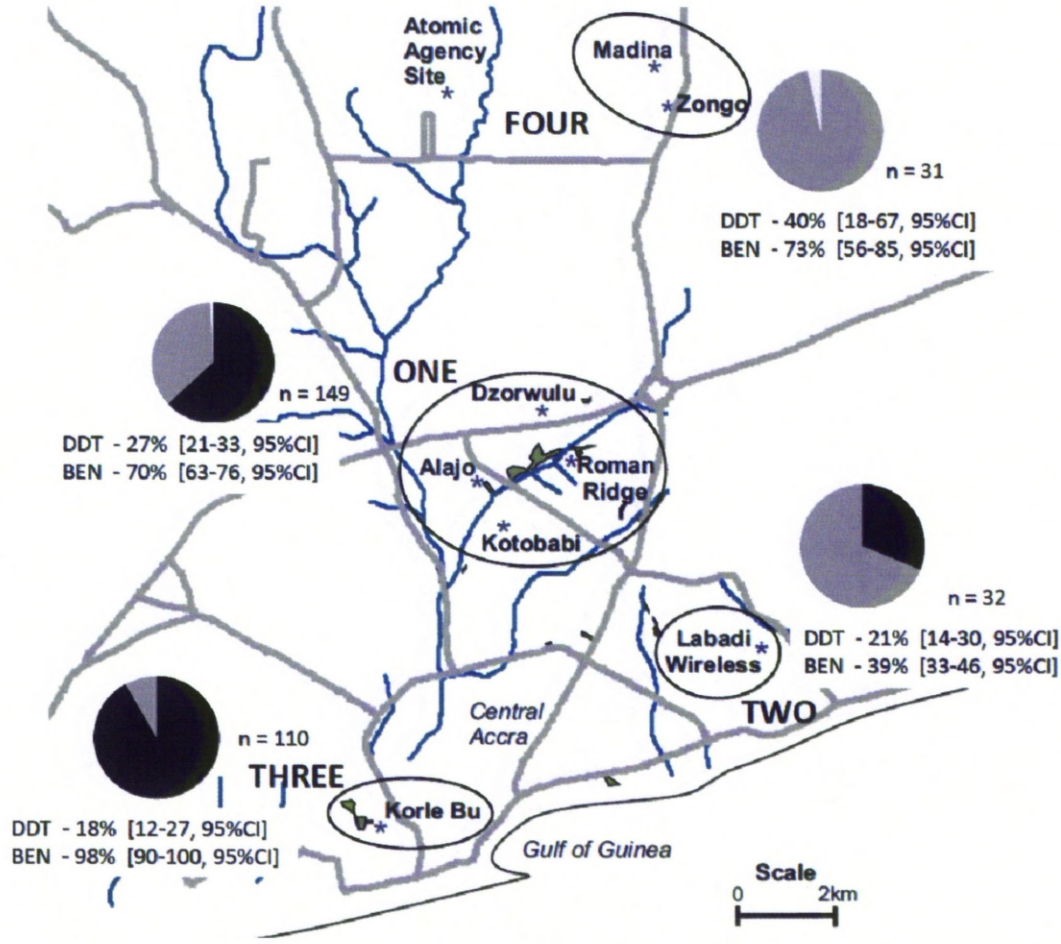

Figure 2.18. Map indicating molecular form composition of non-exposed female *An. gambiae* s.s and percentage mortality of females post insecticide selection with DDT or bendiocarb at four collection areas in the Accra region of Ghana, West Africa.

Pie charts indicate M forms in black, S forms in grey and suspected M/S hybrids in white. Percentage mortality of females exposed to either 4 % DDT for 6 hours or 0.1% bendiocarb (BEN) for 1 hour are show with 95% confidence intervals (Newcomb, 1998) .

The differences in resistance phenotype may reflect, at least in part, differences in the frequencies of target site mutations in the two forms as a significant association between molecular form and presence of target site mutations *L1014F* and *G119S* was uncovered in unselected (control) mosquitoes using contingency table analysis. The frequency of the *G119S* allele in unselected S form *An. gambiae s.s.* was over three times that of the M form controls. Similar frequencies of the *G119S* mutation have been reported in sympatric populations in Burkina Faso with S form allelic frequencies between 0.25-0.5 and M forms 0.03-0.13 (Dabire *et al.*, 2008; Djogbenou *et al.*, 2008b). However it is suspected from subsequent 2010 collections that the frequency of *G119S* in Ghanaian M forms is increasing and this is currently under investigation (D. Weetman pers. comm.). The *G119S* mutation is likely to have been a unique event in one molecular form and introgressed into the other (Djogbenou *et al.*, 2008a). To date

there is no conclusive evidence for which form the mutation first arose in, however as the frequency is higher in S form populations it is likely to have subsequently moved into M forms via a rare introgression event; a phenomenon already documented for the *kdr* mutation (Weill *et al.*, 2000).

A similar relationship was witnessed for the *kdr* West allele, *L1014F*, whereby the frequency was higher in the S forms controls compared with M forms; although the difference was not as marked as the *G119S* frequency. Higher frequencies of the *kdr* allele are commonly witnessed in the S molecular form (reviewed by Santolamazza *et al.* (2008)) with previous reporting of *L1014F* at almost fixation in Ghanaian S forms (98-100% allele frequency), while M form frequency was considerably lower than reported here at 3.38% (Yawson *et al.*, 2004). The *kdr* mutation is believed to have reached M forms through introgression from S forms with two polymorphic markers associated with the *L1014F* mutation in both molecular forms only found in S forms wild-types (Weill *et al.*, 2000). Although a suspected M form de novo *kdr* mutation was reported on Bioko Island (Reimer *et al.*, 2005). Substantial introgression in Ghana is also supported by a lack of genetic differentiation between forms in a microsatellite study (Yawson *et al.*, 2007). The frequency of *L1014F* in Ghanaian M forms has continued to rise with an increase in allele frequency from 0.03 to 0.54 reported in 5 years between 2002-2007 (Lynd *et al.*, 2010) (Figure 2.19, below). Increasing *kdr* frequency in M forms has also been reported in Burkina Faso, West Africa (Diabate *et al.*, 2003).

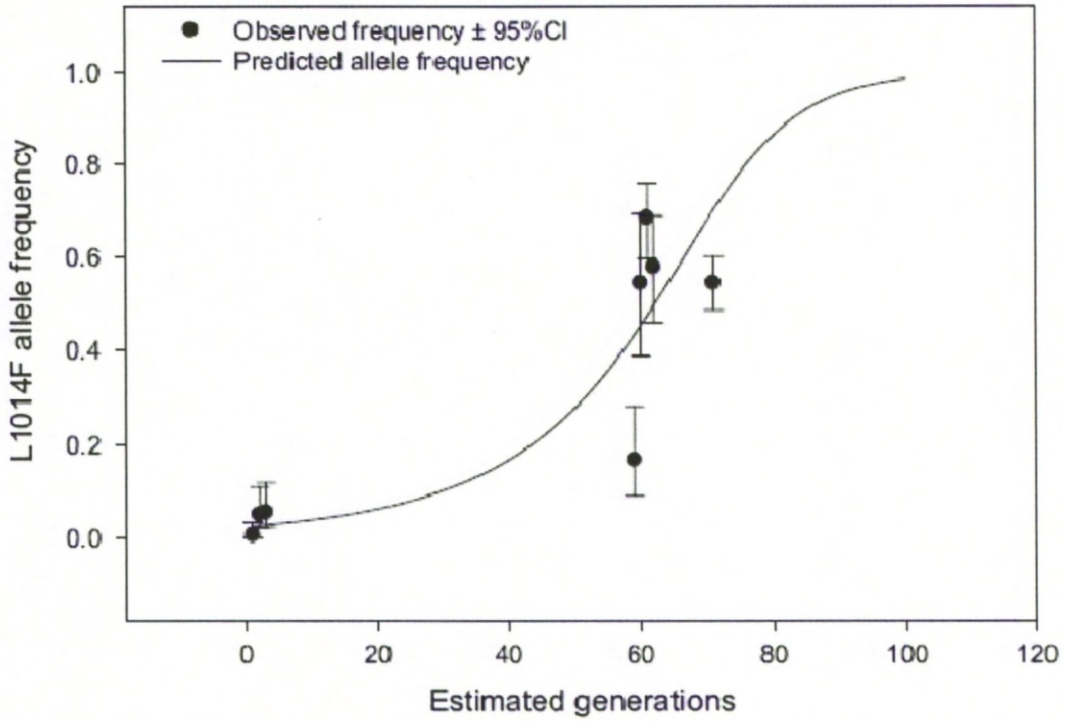

Figure 2.19. Observed and predicted changes in *L1014F* allele frequency in *An. gambiae* M-form populations from southern Ghana.

Observed data obtained from surveys conducted in 2002, 2007, and 2008 (this study). First collection point (Generation 1) was June 2002. Data from 2002, first three data points, are taken from Yawson *et al.* (2004) all other data are novel. One generation per month is assumed following Lehmann *et al.* (2009). The 95% CIs for each observed data-point were calculated according to Newcomb (1998). Expected data generated from simultaneous maximum likelihood estimates of initial frequency and selection and dominance coefficients. Figure and adapted from Lynd *et al.* (2010).

2.5.2 Bendiocarb resistance

Carbamate insecticides are promising candidates for IRS and insecticide treated materials in areas, such as Ghana, where resistance to pyrethroids is extensively documented (Guillet *et al.*, 2001; Djenontin *et al.*, 2009). There is a lack of previous data on the presence of bendiocarb resistance in Ghana with only one study reporting resistance with a mortality rate of 56% [40-72%, 95%CI] post WHO bioassay; however, the sample size was small (n=39) (Coetzee *et al.*, 2006).

In this study, overall male mortality after 1 hour exposure to 0.1% bendiocarb was marginally higher compared with females, although confidence intervals overlap.

Variability in mortality from the four collection areas was noted with female mortality ranging from 39-98%, which is likely to reflect the distribution of molecular forms (Figure 2.18) and hence target site resistance attributed to *G119S* mutation. For instance the highest bendiocarb resistance was recorded in site two (Labadi) where female percentage mortality was 39.1% and the control group predominantly S form (69% [50-83%, 95% CI]); the molecular form showing higher *G119S* allele frequency.

In contrast, female mortality at site three (Korle Bu) was 98.3% and control mosquitoes were 91.8% [84.6-96.0, 95% CI] M form. A later study in the same Accra region (Weetman *et al.* unpublished data) found higher bendiocarb resistance in the S molecular form with ~60% mortality post WHO standard bioassay and no survivors recorded in the M forms after the same exposure regime. To date this is the first reporting of *G119S* frequencies in Ghanaian *An. gambiae* populations.

Departure of *G119S* frequencies from Hardy Weinberg equilibrium causing heterozygote excess has been observed in wild populations of *An. gambiae* from Burkina Faso (Djogbenou *et al.*, 2008b) and Côte d'Ivoire (Djogbenou *et al.*, 2009). The *ace -1* gene encodes for an essential neurotransmitter, acetylcholinesterase, and for this reason individuals homozygous for the *G119S* mutation potentially incur fitness costs. In laboratory experiments *An. gambiae* homozygous for the *G119S* mutation experienced significantly higher mortality during pupation compared to a susceptible strain, suggesting marked effects on survivorship (Djogbenou *et al.*, 2010). While in previous field studies longer developmental time and reduced adult size were found to be associated with the *G119S* mutation in *Cx. pipiens* (Bourguet *et al.*, 2004). It has been postulated that the *ace-1* gene has duplicated in some *An. gambiae* populations creating "permanent heterozygotes" which may help to overcome the fitness costs of the mutated gene (Djogbenou *et al.*, 2008a). This phenomenon was previously reported in *Culex pipiens* mosquitoes (Bourguet *et al.*, 1996). *Ace-1* duplication may be a "serious challenge" for malaria control as insecticide rotation relies on the fitness costs of resistance being disadvantageous in the absence of insecticidal selection pressure (Djogbenou *et al.*, 2009). In this study, a departure from Hardy-Weinberg equilibrium was not detected in S form controls although sample size was small (n=73). The most common genotype in S form controls was wildtype (SS) accounting for ~45%, followed by heterozygotes with homozygotes the least common genotype at ~ 13%. At present the *ace-1* duplication is only inferred via departure from Hardy-Weinberg Equilibrium and cannot be detected on an individual basis due to a lack of diagnostic features within the duplicated gene (Djogbenou *et al.*, 2009).

Significant association between bendiocarb phenotype (test/control) and *G119S* genotype was found in both molecular forms, with mosquitoes surviving bendiocarb selection displaying higher *G119S* allele frequencies compared to controls. This is indicative of an association between the mutation and bendiocarb resistance in the area. If time permitted the dead mosquitoes from selection could be genotyped to establish whether a bias towards *ace-1* wild-type in susceptible mosquitoes is found.

2.5.3 DDT resistance

DDT use during the WHO lead Global Malaria Eradication Campaign of 1955 saw the risk of malaria eliminated in areas such as Europe, North America, the Caribbean and parts of Asia and South-Central America (Carter and Mendis, 2002). Subsequently DDT became out of favour for insecticide control due to bioaccumulation and safety fears (Carson, 1963). However DDT is currently seeing a resurgence in use for mosquito control in Africa supported by the World Health Organization's 2006 advocacy for its use in IRS (WHO, 2006b; WHO, 2007c) and the 2001 Stockholm Convention on Persistent Organic Pollutants permitting use of DDT for vector control. The organochlorine compound is inexpensive and highly effective with one of the longest residual activities of all insecticides (WHO, 2007c). In addition to toxic effects, DDT may cause irritancy and spatial repellency in mosquitoes (WHO, 2007c; Achee *et al.*, 2009) with some suggesting that these effects could provide protection against DDT resistant mosquito populations. Sharma *et al.* reported that IRS with DDT effectively controlled malaria transmission and reduced numbers of a DDT resistant vector population in India (2005). However this study failed to address environmental and temporal confounders and lacked an insecticide free control. Grieco *et al.* (2007) found DDT capable of preventing resistant strains of *Ae. aegypti* from entering huts in lab and field trials arguing that repellency is the primary mode of DDT action; although this mosquito species is not a vector of malaria and we cannot infer all species will respond similarly. In 2007 over 10 African countries were using DDT routinely in IRS regimes (WHO, 2007a).

Concerns over DDT resistance remain, especially in light of a cross resistance to pyrethroids caused by mutations in the shared target, the voltage gated sodium channels of the insect nervous system (Rashatwar and Matsumura, 1985). Resistance to DDT can also be caused by metabolic processes within the insect with enzyme glutathione-S-transferase-epsilon 2 (*GSTe2*) linked to DDT resistance in *An. gambiae* (Ortelli *et al.*, 2003) and *Ae. aegypti* (Lumjuan *et al.*, 2005).

DDT resistance has been previously reported in Ghanaian *An. gambiae* with high levels of resistance recorded in a south-central S form population (31% mortality [15-52, 95% CI]) after one hours exposure to 4% papers, although sample size was low (n=26) (Coetzee *et al.*, 2006). However in south-western Ghana resistance levels were reportedly much lower with mortality rates of 94-100% after WHO bioassays recorded in *An. gambiae s.s.,* although molecular form was not resolved (Kristan *et al.*, 2003). The mechanisms behind suspected DDT resistance in Ghana have not been fully investigated although the *kdr* mutation *L1014F* is reported to be present at high frequencies in the S molecular form (Yawson *et al.*, 2004). In light of inconclusive evidence for resistance and potential reintroduction of DDT for control, this insecticide was chosen for phenotypic and genotypic analysis of *An. gambiae s.s* originating in the Accra region of Ghana.

It became apparent during initial bioassays that resistance to DDT in the area was extremely high with 90 minutes exposure to 4% papers resulting in only ~3% mortality

overall in both sexes of *An. gambiae s.s* (Appendix 2). For selections a 6 hour exposure time was chosen as this was the maximum exposure achievable within a working day. If time and collection numbers had permitted more intermediate time points and also longer exposure times could have been added to bioassay data to achieve a mortality curve similar to that produced for bendiocarb. However, natural mortality may have become a confounder if exposure time was greatly extended. An alternative method would be to vary DDT concentration to achieve an LC (Lethal Concentration) curve; though this option was not available in the field laboratory. Moreover, at concentrations above 4%, DDT solutions become difficult to impregnate onto test papers.

Overall mortality post selection was extremely low with a combined male and female mortality of ~32% recorded after 6 hours exposure. This may result in difficulties identifying resistance markers when comparing selected mosquitoes to non-exposed (control) material as resistance levels are likely to be high in all individuals. Ideally, exposure time should discriminate the highly resistant from susceptible. However the experimental design was constrained by the need to use control mosquitoes rather than those which did not survive DDT exposure as post-mortem RNA degradation may occur in this latter group.

Mortality between the sexes post selection varied with males more susceptible compared with females. Males also appeared marginally more susceptible to bendiocarb in this study. From an evolutionary perspective it seems likely that females would become more resistant to insecticides used in control as it is only females who host seek and thus contact ITNs and insecticide treated surfaces. Indeed higher levels of pyrethroid resistance were reported in female *An. gambiae* in Ghana (Muller *et al.*, 2008b), while male *An. arabiensis* were consistently more susceptible to DDT compared to females from the same colony at various ages (Matambo *et al.*, 2007); though data in the latter study lacked error measurement. However any selection pressures during aquatic stages would also be a selective force on male mosquitoes. In this study higher resistance levels in female adults would suggest insecticide exposure during adult stages via ITNs and IRS was the major selective force, however confounding factors such as size differences between sexes must also be considered.

After contingency table association analysis there appeared to be no connection between DDT phenotype and the presence of the *L1014F* target site mutation in either M or S molecular forms. This lack of association is likely to result from low levels of variation in *kdr* mutation frequency within these mosquito populations (Donnelly *et al.*, 2009). Overall *L1014F* allele frequency was high in both M and S form controls. However five M-form mosquitoes without *kdr* managed to survive 6 hours of DDT exposure, while there were greater numbers of *L1014F* heterozygotes (63) compared with homozygotes (9) in M form tested mosquitoes with numbers equally matched in the S form test group (Table 2.4). These results in combination with extremely high levels of DDT resistance may implicate mechanisms other than target site mutation in the resistance phenotype. Other authors also conclude that *kdr* does not fully explain DDT resistance in mosquitoes (reviewed by Brooke *et al.* (2008)).

A significant affiliation between the *L1014F* mutation and molecular form was discovered when comparing non-exposed (control) individuals of each form ($p < 0.01$). Higher allele frequency was recorded in S forms controls compared with M forms in addition to a complete absence of wild-type individuals in the S form group. This higher level of *kdr* in S forms is supported by the work of Yawson *et al.* (2004) who reported *L1014F* frequency almost at fixation (98-100%) in Ghanaian S forms. Although, as discussed, a rapid increase in the frequency of *L1014F* in Ghanaian M forms is taking place (Lynd *et al.*, 2010); with data from this study contributing to this publication (see publications in support).

Limited initial *kdr* screening via TaqMan assay suggested the presence of the *kdr* East allele (*L1014S*) in 11 individuals, though this was not confirmed through sequencing of the sodium channel. However both mutations have been reported in *An. gambiae s.s.* populations in Uganda (Verhaeghen *et al.*, 2006), Gabon (Pinto *et al.*, 2006) and Cameroon (Etang *et al.*, 2006). Reimer *et al.* (2008) investigated the relationship between *kdr* frequency and resistance to DDT and pyrethroids in a Cameroonian population containing both alleles. Both East and West mutations were associated with DDT resistance however *L1014F* (West) homozygotes provided greater resistance to permethrin compared with insects carrying both mutations (*L1014F/S* heterozygotes) and it was concluded that the movement of the West allele into Eastern populations was of greater concern. However *de novo* mutation rather than just migration of alleles may be important in the spread of *kdr* resistance. Indeed at least four independent mutation events have been detected for *kdr* alleles in African S forms (Pinto *et al.*, 2007) while an independent M form mutation is suspected to have occurred in a population from Bioko Island (Reimer *et al.*, 2005). It is essential to screen for both East and West mutations in African mosquito populations. However mutations at position 1014 in the sodium channel may not provide the full story of knock down resistance with many additional mutations noted in other insect species (reviewed in Davies *et al.* (2008)).

Such high levels of DDT resistance and *L1014F* frequency in combination with reported levels of pyrethroid resistance may imply these two compounds are no longer appropriate for control of *An. gambiae s.s* in Ghana. In support of this N'Guessan reported on reduced efficacy of ITNs and IRS with pyrethroid lambdacyhalothrin in an area of Benin with high *kdr* frequency in *An. gambiae* (2007). However some researchers suggest DDT's repellent action may provide protection against resistant mosquitoes (Grieco *et al.*, 2007), and pyrethroid treated ITNs reportedly remain effective in populations with high frequencies of *kdr* (Chandre *et al.*, 2000; Asidi *et al.*, 2005). Overall though, there is a lack of evidence that DDT can provide effective malaria control in the presence of highly resistant vector populations.

Kdr alone may not compromise insecticidal control with DDT or pyrethroids; however if present in combination with metabolic resistance mechanisms the consequences for control may be considerable.

2.5.4 Conclusions

2.5.4.1 Multiple and highly resistant population

Genetic diversity and substrate plasticity allow the large enzyme families associated with metabolic detoxification to cause broad spectrum insecticide resistance. In contrast, target site mutation mechanisms produce resistance to compounds targeting a single site; although when two insecticides share the same target, as is the case with DDT and pyrethroids, cross resistance can occur. However there is a restriction on the number of mutations which can accumulate in a target gene before normal physiological activities of the gene product are compromised. One way to overcome these restrictions is to carry a duplicate target gene as is suspected with *ace-1* in *An. gambiae* (Djogbenou *et al.*, 2008a).

When both target site and metabolic mechanisms occur concomitantly in a vector population the impact on control could be marked. Multiple resistance in mosquito population has been previously reported with *Anopheles* species in Sri Lanka exhibiting resistance to DDT, carbamates, organophosphates and up to 6 pyrethroids attributed to high insensitive acetylcholinesterase frequencies and elevated metabolic enzyme levels (Perera *et al.*, 2008). Corbel *et al.* (2007) documented elevated levels of resistance to DDT, pyrethroid, dieldrin and beginnings of carbamate resistance in *An. gambiae* and *Cx. quinquefaciatus* in Benin associated with a high frequency of *kdr* mutation *L1014F* and increased metabolic enzyme levels.

Resistance to multiple compounds was detected in the mosquito populations sampled in this study with high levels of both DDT and bendiocarb resistance recorded in samples from areas one, two and four (Appendix 3 and 4). In addition a small set of pyrethroid WHO bioassays suggested permethrin resistance in area four (19% mortality after 1 hour exposure) and deltamethrin resistance in area three (76% mortality) (Appendix 8). This resistance may be due in part to mosquitoes possessing target site mutations with 29% of samples screened for both mutations (n=182) positive for *L1014F* and *G119S*. However there is a likely role for metabolic processes in this multiple resistance phenotype.

2.5.4.2 Understanding resistance

It is essential that field populations are screened for resistance to all classes of insecticide before deciding on an appropriate intervention. A limited number of compounds are available for control due to a lack of novel insecticides and strict rules on toxicology. An understanding of the complexities of resistance through investigation of the molecular biology underpinning this trait is essential for the development of new insecticides. In addition field monitoring of resistance indicators is vital to enable decision making for tailored insecticide use and will rely on identifying molecular markers of resistance.

Chapter 3: Microarray analysis of bendiocarb resistant *An. gambiae s.s.* from Ghana, West Africa

3.1 Abstract

Bendiocarb insecticide resistance in field caught *An. gambiae s.s.* from Ghana, West Africa, was investigated using whole genome transcriptional profiling. To date, studies of differential expression in insecticide resistant mosquitoes have relied on small candidate gene arrays which do not objectively screen all known transcripts and may miss novel resistance mechanisms. Therefore a new whole genome microarray was designed to compare resistant females (those which survived 1 hours exposure to 0.1% bendiocarb) with non-exposed controls from the same area. Members of the three major enzyme families previously linked to resistance were identified as over-expressed in resistant mosquitoes including *CYP325C1-3*, *COEJHE2E* and *GSTS1_1*. However a number of novel candidates such as odorant binding proteins, odorant receptors and genes linked to cuticular development were also identified.

3.2 Introduction

A number of approaches are available when investigating a trait of interest on a molecular level. Genotype-phenotype association studies such as genome-wide single nucleotide polymorphism analysis (Craig and Stephan, 2005; Gibbs and Singleton, 2006) or quantitative trait locus mapping (The Complex Trait Consortium, 2003) have been used to uncover genes or allelic variants linked to diseases such as Crohn's (Fransen *et al.*, 2010) and traits such as grain width and height in rice (Song *et al.*, 2007). Similarly, expression studies can uncover genes associated with a phenotype via modified levels of messenger RNA or protein and thus provides a link between genotype and phenotype.

Expression studies may involve characterising the proteins (proteomics) or the genes transcribed (transcriptomics) in individuals from a population exhibiting a certain trait. Both these technologies have the capacity to characterise expression patterns for all genes/proteins and thus have potential to provide an unbiased profile.

Proteomic approaches include mass spectrometry, antibody capture and electrophoresis based methods such as 2D PAGE (current approaches reviewed by Chaerkady *et al.* (2008)). Protein based techniques provide representation of translated genes *in vivo*, however proteins maybe difficult to identify (Bertone and Snyder, 2005) and techniques tend to be less high through-put compared to profiling the messenger RNA.

3.2.1 Transcriptome analysis

Numerous techniques are employed in transcriptome analysis (reviewed by Huestis *et al.* (2009)) including PCR bases methods such as amplified fragment length polymorphism (AFLP) which involves cDNA digestion, amplification and gel separation, with expression related banding patterns interrogated through sequencing of excised bands (Vuylsteke *et al.*, 2007). Serial analysis of gene expression (SAGE) is another restriction enzyme based approach in which cDNA is digested and tagged before cloning and sequencing; the frequency of each tag is indicative of the transcript abundance (Ial *et al.*, 1999).

Techniques based on the amplification or cloning of cDNA followed by sequencing address both allelic and expression level differences. More recently high through-put sequencing techniques, which involve sequencing tagged cDNA using platforms such as 454 pyrosequencing and SOLiD™ have become available to the transcriptomics field (Mardis, 2008). These next-generation sequencing (NGS) technologies create short fragment reads (35-250bp) which can be pieced together to re-create the transcriptome at costs much lower than traditional Sanger sequencing (reviewed by Hudson *et al.* (2008)). However the technology still remains expensive and the analytical approaches for expression profiling using these platforms are in their infancy. Moreover, very short fragment reads such as those provided by Solexa and SOLiD™ may be difficult to map and analyse in absence of a well characterised genome scaffold. Microarray technology currently provides an option for high through-put, extensive characterisation of the transcriptome at an affordable cost. Robust protocols and analysis pipelines mean microarrays are often the method of choice for transcription profiling in organisms where a well annotated genome is available (Baginsky *et al.*, 2010).

3.2.2 Microarray transcriptome analysis

Microarray technology was employed for gene expression analysis in this study, with potential candidates proteomically validated where possible. Microarrays were first developed in 1995 as a high capacity means for profiling many genes in parallel (Schena *et al.*, 1995). Since then array technology has been applied to many molecular fields including SNP detection, micro-RNA characterisation, alternative splicing, histone modification (ChIP), DNA methylation and protein profiling. A series of *Nature Genetics* supplementary issues ("The Chipping Forecast", I, II, III) have been dedicated to microarray technology and its applications (1999 vol. 21, 2002 vol. 32, 2005 vol. 37).

Expression arrays involve the isolation and fluorescent labelling of messenger RNA (target) followed by hybridization to an array containing short sequences (probes) derived from transcript information. Two colour array systems employ competitive hybridisation between two samples labelled with different dyes, routinely cyanine 3 and 5, to establish fold differences in gene expression between samples on laser excitation and scanning (Duggan *et al.*, 1999).

Advances in array technology and competition between commercial companies has led to affordable, high density arrays capable of profiling over a million probes in a single sample. However microarrays are potentially sensitive to cross-hybridisation of probes to non-target sequences (false positives) when genes are closely related or share a section of sequence similarity (Wren *et al.*, 2002). While SNP differences between targets and sequences employed during probe design may result in a loss of signal (false negatives). Both of these issues may be overcome, at least in part, by employing multiple probes per gene to provide greater coverage of the transcript (Roh *et al.*, 2010).

Expression profiling with microarrays has been used for over a decade in a wide range of fields with some of the earliest array experiments concerned with profiling differences in cancerous and normal tissues in humans (Kononen *et al.*, 1998; Khan *et al.*, 1999). Subsequent array deployment has inevitably followed whole genome sequence availability, with genome-wide profiling in *E.coli* (Richmond *et al.*, 1999), expression analysis during embryogenesis in *Drosophila* (Tomancak *et al.*, 2002) and transcript profiling of the humans stages of malaria parasite *Plasmodium falciparum* (Bozdech *et al.*, 2003) rapidly following genome publication.

3.2.3 Microarray studies in *Anopheles*

The first microarray experiment in *An. gambiae* preceded the availability of the genome sequence and employed probes designed against ~6,000 expressed sequence tags (ESTs) (Dimopoulos *et al.*, 2000) to investigate gene expression associated with bacterial challenge, oxidative stress, injury and *Plasmodium berghei* infection (Dimopoulos *et al.*, 2002). This array was subsequently employed to study *P. berghei* melanisation in a refractory strain of *An. gambiae* in which elevated expression of redox related genes implicated a chronic state of oxidative stress in the refractory phenotype (Kumar *et al.*, 2003).

Since these early experiments the application of microarray expression profiling in *Anopheles* has been used to explore topics including differential expression between sexes (Justice *et al.*, 2003; Geng *et al.*, 2009), life-stages (Strode *et al.*, 2006), molecular forms (White *et al.*, 2010; Turner *et al.*, 2005) and parasite infected and non-infected insects (Felix *et al.*, 2010; Dong *et al.*, 2006). More recently microarray analysis has been employed in an attempt to accurately age grade mosquitoes by identifying gene whose expression profiles 'delineate chronological age' (Wang *et al.*, 2010a; Cook and Sinkins, 2010). Age distribution of female *Anopheles* in the wild is a critical component of vectorial capacity and disease transmission related to the malaria parasites extrinsic incubation period within the mosquito.

Some investigators opt for a whole genome profiling technique using large scale microarrays (Erickson *et al.*, 2009; Dong *et al.*, 2006), while others have taken a candidate gene approach employing gene families thought to be associated with a trait of interest (Strode *et al.*, 2006; Felix *et al.*, 2010). The combination of two genomes on a single array has also been exploited for investigations into parasite-host interactions;

Xu *et al.* (2005) took a candidate gene approach and employed an array containing ESTs expressed during the insect stages of *P. berghei* and *An. stephensi* genes known to be induced *by Plasmodium* infection. Furthermore commercial arrays (GeneChip® Plasmodium/Anopheles Genome Array – Affymetrix) are now available that combine the whole genome of *P. falciparum* with that of its host *An. gambiae*.

More recently transcriptome sequencing techniques have been applied to the field of mosquito molecular biology. Deep sequencing has been used in an attempt to resolve the phylogenies of 10 mosquito species (Hittinger *et al.*, 2010) and to establish early zygotic (Biedler and Tu, 2010) and microRNA gene expression patterns (Skalsky *et al.*, 2010) in *Aedes* and *Culex* species.

3.2.4 Analysis of insecticide resistance related gene expression

Discovering the genetic causes of insecticide resistance is a major area of research to which microarray technology has been applied. One of the first microarray experiments to explore the resistance phenotype used a candidate gene approach employing an array consisting of 132 metabolic genes, focusing on 90 cytochrome P450s, to investigate DDT resistance in *Drosophila* (Le Goff *et al.*, 2003). Subsequently a similar candidate gene approach was applied to *An. gambiae* insecticide resistance research with an array (the detoxification chip) consisting of 230 genes putatively involved in insecticide metabolism (David *et al.*, 2005). Genes present on the chip included those encoding cytochrome P450s, GSTs and carboxylesterases, all members of families previously associated with metabolic break down of insecticides. Initial investigations on laboratory colonies uncovered five genes up-regulated in a DDT resistant strain of *An. gambiae* including *GSTe2,* and up-regulation of P450 gene *CYP325A3* in a permethrin tolerant strain.

Subsequently this *An. gambiae* candidate gene array was employed to profile transcription in pyrethroid resistant field populations (Muller *et al.*, 2007; Muller *et al.*, 2008b). One candidate P450, *CYP6P3*, up-regulated in resistant mosquitoes, was capable of metabolising both alpha-cyano and non-alpha-cyano pyrethroids in a recombinant protein system (Muller *et al.*, 2008b).

A similar 'detoxification chip' has been designed for dengue and yellow fever vector *Ae. aegypti* (Strode *et al.*, 2008). This array has been employed to investigate pyrethroid and organophosphate resistance in an *Ae. aegypti* colony isolated from the island of Martinique (Marcombe *et al.*, 2009). Biochemical assays implicated elevated detoxifying enzymes in the resistance phenotype which was supported by significant up-regulation of 31 'detox genes', predominantly P450s from the CYP6 and CYP9 class, through microarray analysis. However the Martinique resistant colony was compared to a susceptible colony isolated from Bora Bora, French Polynesia, meaning geographical differences may have confounded results.

The potential cross-induce of detoxification genes via sub-lethal exposure to environmental xenobiotics and insecticides permethrin and temephos was also explored using the *Aedes* 'detox-chip' (Poupardin *et al.*, 2008). *Ae. aegypti* larvae

exhibited elevated expression of 12 detox genes, of which 11 were cytochrome P450s, directly after a 24 hour exposure to a sub-lethal dose of xenobiotics or insecticides; although qPCR failed to validate over-expression of four of these P450s. Two class 9 P450s (*CYP9M8, CYP9M9*) were found to be induced by both permethrin and herbicide atrazine via microarray and qPCR, suggesting cross induction. Currently the potential role for these P450s in pyrethroid resistance is unknown. Interestingly, a moderate increase in permethrin tolerance was also reported post exposure to xenobionts fluoranthene (a road surface constituent) and copper, which may suggest that environmental toxins have the capacity to influence insecticide tolerance. More recently this group employed NGS to investigated gene expression in the same *Ae. aegypti* strain directly after 48 hours exposure to sub-lethal doses of the same three pollutants (atrazine, fluoranthene and copper) and insecticides permethrin, imidacloprid and proproxur (David *et al.*, 2010). In the first application of NGS with respect to insecticide resistance, longSAGE and Solexa deep sequencing of RNA revealed up regulation of cuticular proteins, transporters and enzymes in xenobiotic and insecticide exposed mosquitoes compared to non-exposed controls. However, the two P450 genes induced by 24 hours exposure to permethrin and atrazine in the previous study (Poupardin *et al.*, 2008) were not found to be up-regulated post exposure via this sequencing approach.

The NGS study also identified a large number of transcripts which could not be mapped to the current *Ae. aegypti* genome. Vontas *et al.* (2010), in commentary, suggest that these novel transcripts may be a result of alternative splicing or novel transcriptional units, which NGS has the advantage over microarray in detecting. However 'novel transcripts' may also reflect issues with the quality of the current genome annotation. The authors also highlight cost and data-processing issues which mean it is unlikely next-generation sequencing will completely replace microarray transcriptome analysis as a 'routine tool' in insecticide resistance characterisation.

Gene induction upon exposure to insecticides has also been explored in *An. gambiae*. Vontas *et al.* (2005) employed a genome wide approach to investigate the potential for induced gene expression after a 15minutes exposure to sub-lethal doses of permethrin in a strain exhibiting resistance to pyrethroids. Mosquitoes were sampled at 2, 6, 10 and 24 hours post exposure and compared with heat-treated (45°C) controls using an array containing 20,000 *An. gambiae* ESTs. Only minor changes in gene expression were recorded 2 and 24 hours post exposure, however 76 ESTs were differentially transcribed in permethrin exposed insects 6 and 10 hours post exposure compared to heat treated controls. The detoxification, mitochondrial/redox and stress response functional classes were represented by 24 of these differentially expressed genes including carboxylesterase *COEAE2F*, putative nitrilases and P450 electron donor cytochrome B_5. No P450s were reportedly induced; however the authors noted a deficit of detoxification associated genes on this microarray

In *Drosophila* a four hour exposure to 6 different insecticides including organophosphate diazinon and organochlorine DDT, induced 'minimal' gene induction when compared to unexposed files using a candidate gene array containing 186

detoxification related genes (Willoughby *et al.*, 2006). Exposure to DDT elicited low level induction of two genes; *CYP12D1*, a gene not previously associated with resistance; and *GSTD2* which previously exhibited no DDTase activity *in vitro* (Tang and Tu, 1994). Conversely, exposure to xenobiotics caffeine and Phenobarbital elicited significant induction of a number of P450 and GST genes. The authors concluded that constitutive gene expression was more important to the resistance phenotype; however as the array contained only 186 genes, the induction of other genes may have been missed.

A candidate gene microarray approach has the advantage of being focused, generating manageable amounts of expression data. However this technique is only beneficial when the investigator has previous knowledge of the genes involved in a trait and there is also the potential to miss novel candidates. The whole genome approach is more 'open' and less biased providing a fuller picture of the transcriptome; however microarrays of this nature can be more expensive to produce, are dependent on good annotation, and the large data sets generated difficult to interpret.

To date, transcription profiling of insecticide resistant insect populations has provided significant insight into the genetic basis of the resistance phenotype. Ultimately data generated from these experiments may lead to the identification of metabolic targets for new insecticides/synergists and the development of simple molecular tools for the detection of resistant alleles in field populations (David *et al.*, 2005).

3.3 Aims

This study aimed to profile the whole transcriptome of a bendiocarb resistant field population of *An. gambiae s.s.* using microarray technology. The resistant mosquitoes are those females which survived a one hour exposure to 0.1% bendiocarb (test), and were compared to non-exposed females from the same area (control). A custom whole genome microarray was employed to profile transcripts exhibiting differential expression between the test and control groups. This was the first time whole genome transcription profiling had been applied to insecticide resistant wild caught *An. gambiae* populations.

Subsequently differential expression was validated using reverse-transcription quantitative PCR and promising candidates pursued using recombinant protein technology.

3.4 Methods

3.4.1 Microarray design: 4x44K 'AGAMIMMUNODETOX'

Previously at LSTM, microarray analysis had been performed using an in-house oligo-spotted microarray containing a subset of genes putatively linked to insecticide resistance, with notable success (David *et al.*, 2005; Muller *et al.*, 2007; Strode *et al.*, 2008; Muller *et al.*, 2008b). However a candidate gene approach may result in novel

mechanisms of resistance being missed. To profile the genome in a less biased fashion, a new whole genome array design was sought.

The new microarray design employed the latest annotation of the *An. gambiae s.s.* genome (Ensembl 49, 2008) on a commercial array platform (Agilent). This platform enabled high density spot coverage and consistent quality, both of which had been issues with previous in-house arrays. Agilent provide a number of 'off the shelf' catalogue arrays as well as custom array design using online package eArray (https://earray.chem.agilent.com/earray/).

An *Anopheles gambiae* 4x44K whole genome array was available from the Agilent catalogue which used various public databases for probe design namely: Ensembl (49); RefSeq (27); Unigene (build10) containing known genes from GenBank and ESTs from dbEST; and USCS anoGam1 from a 10x whole genome shotgun assembly. However after performing a stand-alone BLAST search (StandAlone WWW BLAST Server ftp://ftp.ncbi.nih.gov/blast/ (Altschul *et al.*, 1990)) against the probes on this array using a set of 282 detoxification genes from an in-house database, some genes were not represented. This may be a result of poor database deposition of a number of the genes from the 'Detox-Chip' array, which had been fully sequenced in-house. Given that this set of detoxification genes represent gene families known to be involved in insecticide resistance it was imperative that they were represented on the new array. In addition, it was noted that the commercial array contained varying numbers of redundant replicate probes which were used to 'fill-up space' on the array.

It was therefore decided to amend the commercial 4x44K array (1,417 Agilent controls and 43,803 custom probes) using the online design tool eArray. The new array, the *'AGAMIMMUNODETOX'*, was a collaborative venture with colleagues from Imperial College London (G. Christophides and A. Jackson) and was designed as follows:

In total 19,744 Agilent designed whole genome probes were employed. Firstly, duplicate probes from the catalogue array were filtered and removed and probes sorted according to ID. One unique probe per single transcript (cDNA/EST) was selected on the basis of an Agilent probe scoring system (BC score). Probes were then duplicated so two copies of one unique probe per unique transcript were included on the array. To ensure coverage of detoxification candidates, additional probes were designed in eArray for the 282 *Anopheles gambiae* 'Detox' genes. These genes were taken from the *An. gambiae* 'Detox-Chip' design (David *et al.*, 2005) and along with some additional candidate cDNAs were uploaded into eArray and three unique 60mer (60bp) probes designed per gene using a base composition method with vector and repeat masking (for definitions see Appendix 10). Probes were designed in a sense orientation with a 3' sequence bias employing the 'Best Probe' methodology (Appendix 10), with the *An. gambiae* transcriptome used as the background system. Each of the three probes was duplicated once on the array creating a maximum of 6 probes per gene.

As this array was a collaborative project with a group working on insect immunity at Imperial College London, probes were designed for 574 immunity related genes

provided by A. Jackson. These probes were designed with the same parameters as the Detox probes. Two probes per gene were designed in eArray, which were then duplicated on the microarray.

The remaining ~350 probe spaces were employed for 'Coefficient of Variance' (CV) gene probes which are used for quality control across the array. These gene probes (35 in total) were selected from a list provided by Agilent and replicated ten times as recommended (eArray). The array platform also contained an Agilent control grid harbouring positive and negative control spots as well as manufacturing controls to assess quality during production.

3.4.2 Sample pooling

From phenotypic and genotypic characterisation of bendiocarb resistance in field collected mosquitoes (Chapter 2) it was apparently that the vast majority (97%) of *An. gambiae* females able to survive 1 hour exposure to 0.1% bendiocarb were of the S molecular form. Fifty six percent of S-form mosquitoes phenotyped for resistance to bendiocarb were also heterozygous for the *G119S ace-1* mutation. For this reason test and control samples were drawn from S-form females heterozygous for the *G119S* mutation. By standardising the molecular form and target-site genotypes within groups to be compared, we hoped to reduce potential confounding as in an earlier study (Muller *et al.*, 2008b)."

Data on all S form individuals that were heterozygous at the *iAChE* (*G119S*) locus were tabulated. Females that survived bendiocarb exposure were selected as 'test' animals. These females were drawn from collection area two (Labadi) where resistance was most marked (Figure 2.1, Chapter 2). Control animals were selected from all collection areas sampled.

3.4.3 Experimental design

Two common designs for microarray experiments are straight dye swap and reference designs. The former involves a pair-wise comparison with reciprocal dye labelling of samples, while reference designs compare all samples of interest to a known reference RNA pool and may also employ dye swaps (Figure 3.1). However these designs are not always optimal. For instance with a straight dye swap two arrays will ultimately contain the same information from a single pair-wise comparison, while half the information from each array in a reference design will concern the reference RNA sample which is of little or no interest.

For these reasons a looped microarray design was chosen for this experiment as each array contains a unique and informative comparison, while each sample is labelled with both dyes addressing potential dye related confounding (Figure 3.1). The loop design was first proposed by Kerr and Churchill (2001) who advocated the design for providing twice the data of a reference design using the same number of arrays whilst balancing for dye effects without direct dye swaps. Experimental evaluation of the loop versus reference design concluded that looped microarray designs were more precise and

able to detect a greater number of differentially expressed genes (Vinciotti *et al.*, 2005). Moreover, an interwoven looped design stabilises the experiment so that information for a sample is unlikely to be lost due to array failure, as each sample is involved in up to four separate comparisons (Kerr, 2001; Vinciotti *et al.*, 2005).

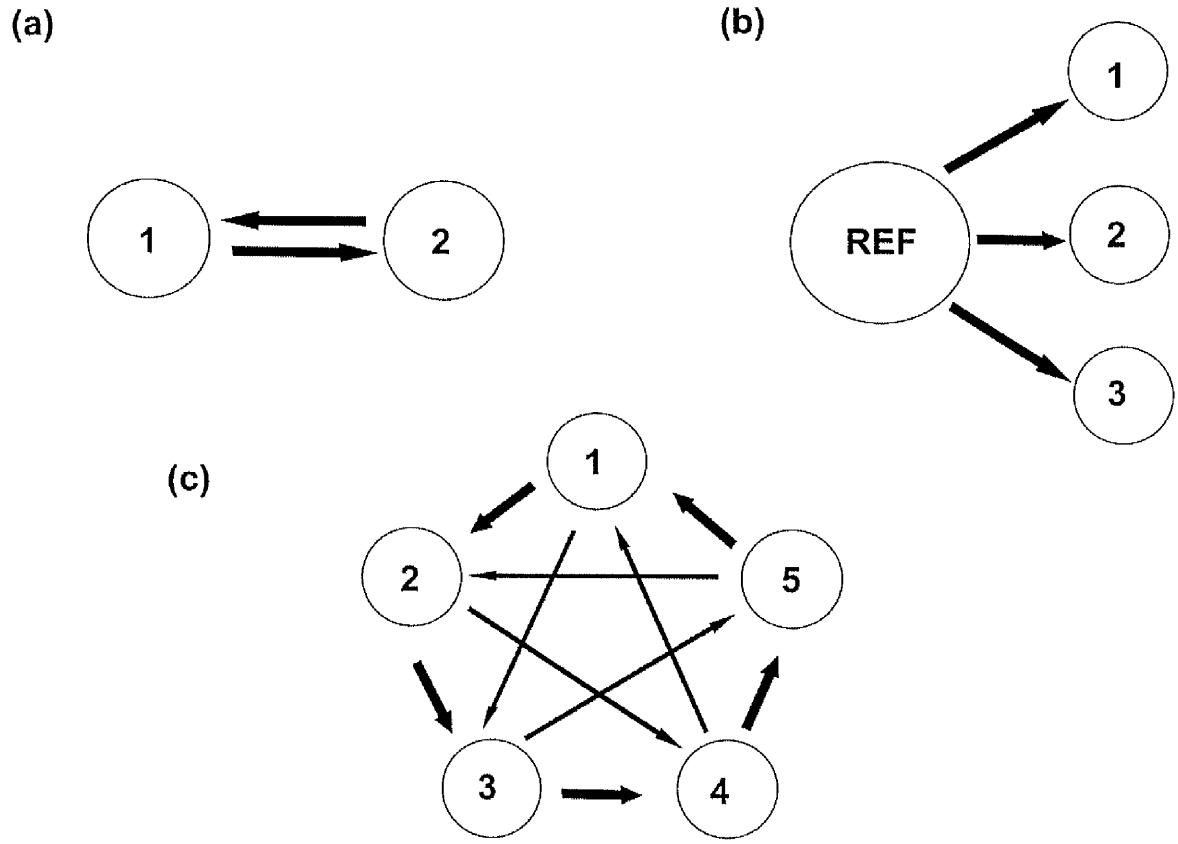

Figure 3.1. Examples of microarray experimental designs:

(a) straight dye swap for two samples involving two microarrays (black arrows) with reverse labelling in one array; (b) reference design for three samples which are compared to a reference pool (REF), dye swaps may also be used; (c) interwoven loop design for five samples, each array involves a unique comparison and each sample is labelled with both dyes.

Five test pools of RNA each extracted from five female *An. gambiae s.s.* S forms collected from area two (Labadi) were selected for the experiment according to RNA quality (Appendix 11). These test females had survived a one hour exposure to bendiocarb (0.1%), and by WHO definition, were resistant to this compound. The five comparative control pools were selected from areas one, two and four, and each contained RNA from five female *An. gambiae s.s* S forms which had been exposed to bendiocarb control papers for one hour (Appendix 11). These controls were representative of the sampled S form population without prior insecticide selection, and so were likely to contain a mixture of resistant and susceptible insects. This experimental design has been successful employed to investigate resistance related

expressed in a previous study (Muller *et al.*, 2008b); though the potential for fold changes to be constrained in a test and control comparison was highlighted. RNA samples were arranged in an interwoven loop design (Figure 3.2) for microarray analysis using a customised 4x44K microarray (*AGAMIMMUNODETOX*).

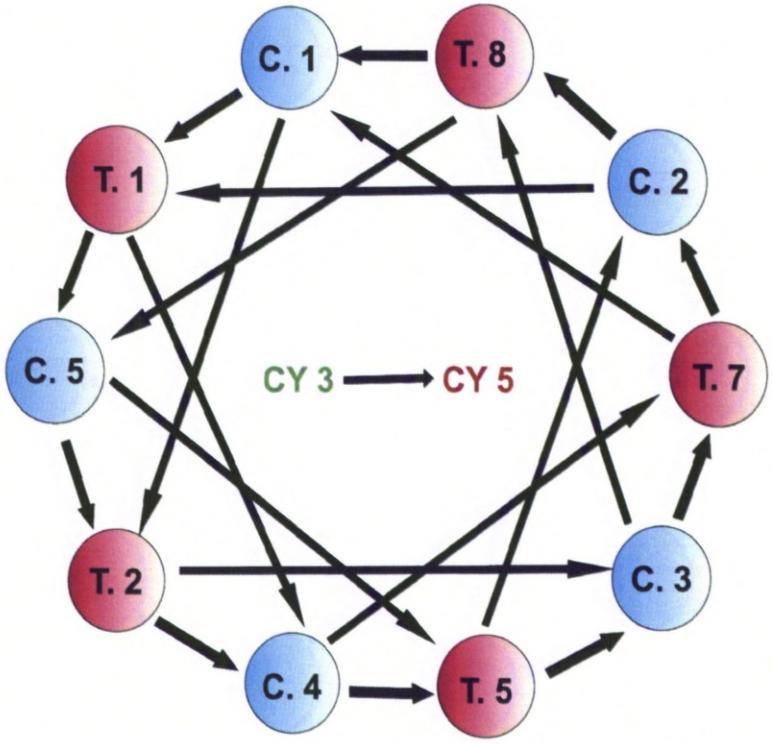

Figure 3.2. Interwoven loop design for a microarray experiment comparing bendiocarb selected and control *An. gambiae s.s.* S forms from the Accra region of Ghana, West Africa.

Five pools of RNA from test females (T - red) which survived a one hour exposure to 0.1% bendiocarb are compared with RNA from control females (C -blue) which were exposed to control papers for one hour. Each black arrow represents an individual array (20) with the direction indicating which dye the sample is labelled with (Cy3 or Cy5).

3.4.4 RNA extraction and quality assessment

RNA was extracted from batches of five test or control females using a modified 'phenol-chloroform like' extraction method. Mosquitoes were removed from emersion in RNAlater® (Ambion) (-20°C) using fine forceps and placed onto filter paper to remove excess preservative solution. Five mosquitoes were then placed into 1.5ml nuclease-free tubes and homogenised in 200µl of TRI Reagent (Ambion) using RNAse free plastic pestles. Tubes were then filled to 1ml with TRI Reagent and incubated for 5 minutes at room temperature (RT). After centrifugation (12,000RPM, 4°C, 10 minutes) the supernatant was transferred to a fresh 1.5ml tube to which 100µl of bromocholoropropane (BCP, Sigma-Aldrich) was added as an alternative to chloroform. Tubes were shaken vigorously for 15 seconds and incubated for 10

minutes at RT before centrifugation to separate phases (12,000RPM, 4°C, 15 minutes). The upper aqueous phase, containing RNA, was carefully transferred to a clean 1.5ml tube to which 500µl of isopropanol (Sigma-Aldrich) was added and vortexed before incubation at RT for ten minutes. RNA was pelleted via centrifugation (12,000RPM, 4°C, 10 minutes) and the supernatant removed before the pellet was washed with 75% ethanol. RNA pellets were allowed to air dry for 3-5 minutes before being dissolved in 30µl of nuclease free water (Sigma-Aldrich).

To remove any DNA contamination, RNA was DNAse treated using the Turbo DNA-*free*™ kit (Ambion) according to manufacturer's instructions.

RNA concentration was then measured using a NanoDrop™ spectrophotometer (Thermo Scientific) and RNA quality assessed using the Agilent 2100 Bioanalyzer and RNA 600 Nano Kit (Agilent) according to manufacturer's guidelines. See Figure 3.3 below for example traces from the Bioanalyzer and Appendix 11 for quantitative and qualitative information on RNA pools.

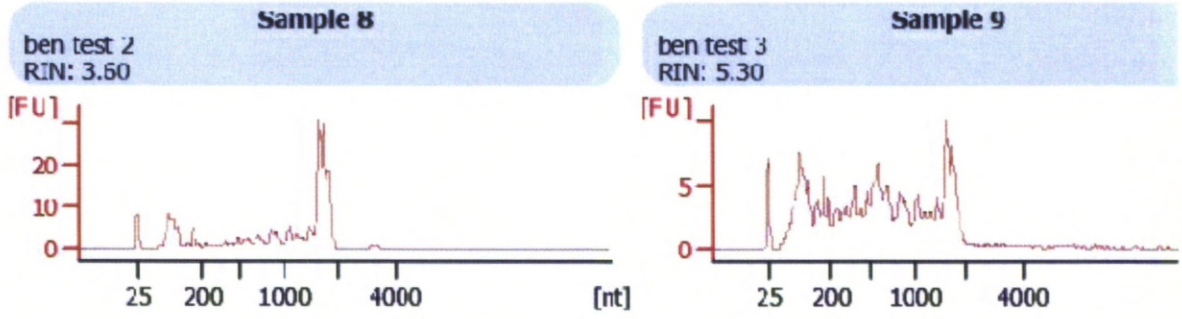

Figure 3.3. Example Bioanalyzer traces from Agilent 2100 Bioanalyzer showing the quality of *An. gambiae* RNA extracted from pools of five females mosquitoes using a modified phenol-chloroform-like extraction method and Turbo DNAse treatment (Ambion).

RNA from females surviving exposure to 0.1% bendiocarb for one hour is shown (test pools 2 and 3). Pool 2 is an example of good quality RNA with 3 peaks at ~ 2000nt representing 18s ribosomal RNA and 28s after it is cleaved, with a relatively flat profile between ~100 and 2000nt. Test 3 pool is an example of degraded RNA characterised by a high and jagged profile between 100-2000nt, this sample was not used for microarray experiments.

3.4.5 RNA amplification and labelling for Microarray

RNA pools of the selected genetic background which were qualitatively determined to be of good quality were selected for microarray labelling (Appendix 11). The concentration of each pool was re-measured using the NanoDrop™ (Thermo Scientific) on the day of labelling and 2000µg in a maximum volume of 8.3µl taken from each pool for labelling.

Samples were labelled with both cyanine dyes (Cy3 and Cy5) using the Quick Amp Labelling Two-Color Kit (Agilent) and RNA Spike-In Kit (Agilent) according to manufacturers protocol version 5.7 (Agilent, 2008). A schematic of the labelling/amplification process is shown in Figure 3.4. All sample incubation steps were performed in a thermal cycler (Dyad, Biorad/MJ Research). Amplified cRNA samples were purified using a column kit (RNeasy Mini Kit, Qiagen) and eluted in 30µl of RNase-free water (Invitrogen).

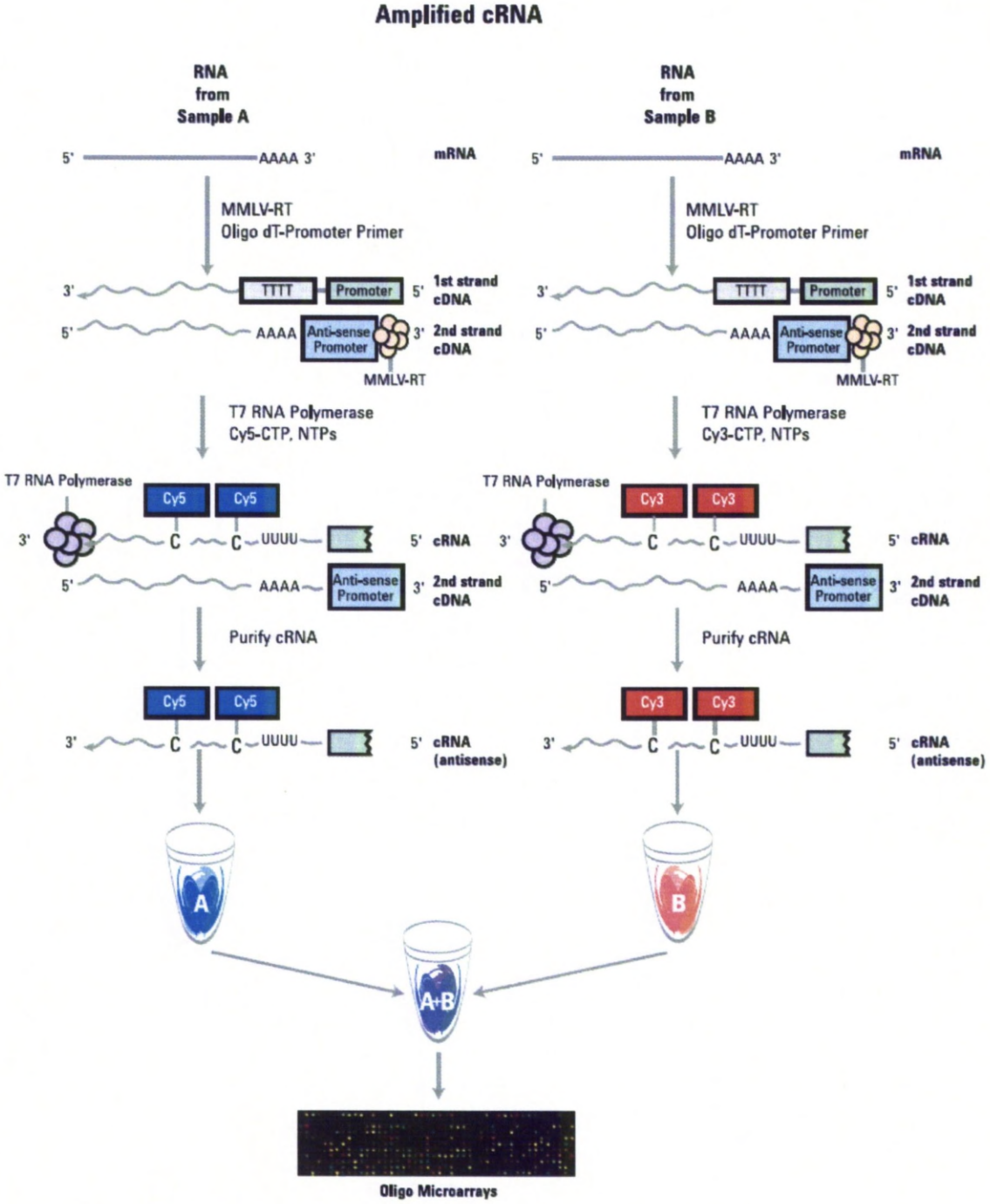

Figure 3.4. Schematic of the RNA amplification and labelling process using the Agilent Quick Amp labelling kit.

Figure taken from the Two-Colour Microarray-Based Gene Expression Analysis (Quick Amp Labelling) Protocol version 5.7 (Agilent, 2008). Messenger RNA is isolated from two samples via its poly(A) tail and converted to double stranded complementary DNA by MMLV-RT (Moloney Murine Leukaemia Virus Reverse Transcriptase). T7 RNA polymerase then creates Cy3 and Cy5 labelled complementary RNA (antisense) (cRNA) which is purified and combined before hybridisation on the oligo-microarray.

Amplification and dye incorporation of samples was assessed using the *Microarray Measurement* setting of the NanoDrop™ (Thermo Scientific). After blanking with water (Invitrogen), 1µl of each labelled sample was measured under the RNA-40 setting and cRNA concentration (ng/µl), absorbance ratio (260nm/280nm) and Cy dye concentration (pmol/µl) were recorded. Yield (µg of cRNA) and specific activity (pmol Cy per µg of cRNA) were determined using the following equations:

Yield:

Concentration cRNA (ng/µl)*30µl (elution volume)/1000 = µg of cRNA

Specific activity:

Conc. of Cy3 or Cy5 (pmol/µl)/conc. cRNA (ng/µl)*1000 = pmol Cy per µg cRNA

Yield > 825ng with a specific activity of >8.0pmol/µg of cRNA is required to proceed to the microarray hybridisation step

An aliquot of each cRNA sample for hybridisation was also analysed for quality on the Agilent 2100 Bioanalyzer using the RNA 600 Nano Kit (Agilent) according to manufacturer's guidelines. Good quality labelled RNA was expected to give a wide peak between 100-2000 nucleotides (Figure 3.5).

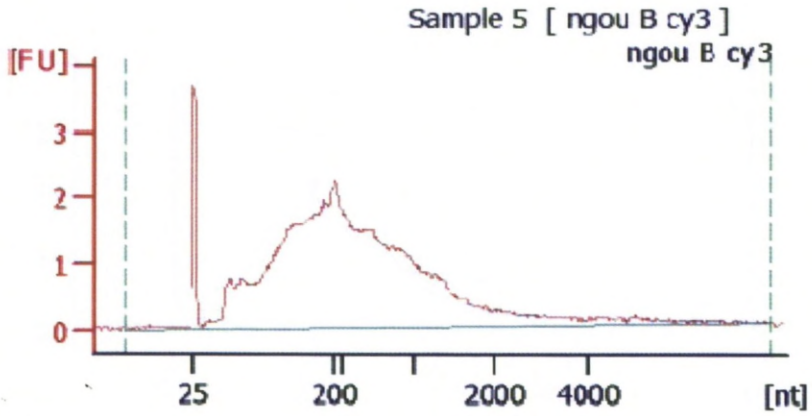

Figure 3.5 Example Bioanalyzer trace from a Cy3 labelled, amplified cRNA sample.

A wide peak between ~100-2000 nucleotides is expected for good quality labelled RNA.

3.4.6 Microarray hybridisation

Pre-hybridisation, 835ng of each labelled cRNA to be compared on a single array were combined and fragmented for 30 minutes at 60°C using reagents from the Gene Expression Hybridization kit (Agilent) following the manufacturer's protocols. To stop the reaction 55µl of 2x GEx Hybridization Buffer HI-RPM (Agilent) was added to the samples and mixed via pipetting. Each tube was spun down briefly and kept on ice until hybridisation.

After placing a clean 4x44K gasket slide into a SureHyb chamber base (Agilent Microarray Hybridization Chamber Kit), 100µl of each hybridisation sample was placed into the centre of each gasket using a "drag and dispense" pipette action. A 4x44K array was then slowly lowered onto the gaskets and the chamber fully assembled. The array gasket sandwich was then placed into a hybridization oven (Agilent) and rotated at 10RPM for 17 hours at 65°C.

After hybridisation, arrays were washed in GE wash buffers 1 and 2 (Agilent) according to manufacturer's instructions. An additional treatment with acetonitrile (Sigma) followed by a Stabilization and Drying Solution (Agilent) was performed to protect against ozone-induced degradation of cyanine dyes and to reduce background noise on arrays.

Arrays were then stored in light protective boxes (Corning) until scanned.

3.4.7 Microarray scanning and feature extraction

Arrays were scanned using G2505C/G2539A Series Microarray Scanner utilising the *Default* profile and settings according to the Agilent Microarray Scanner System User Manual (v 7.0) (Appendix 12).

High (100% photo-multiplier tube (PMT)) and low (10% PMT) extended dynamic range (XDR) scan images were combined and extracted using Feature Extraction (FE) software GE2_10.5_Dec08 (Agilent) and the custom array grid template (022094_D_F_20081124.XML). Default parameters and *FULL* Text output were selected and the resultant QC (Quality Control) reports and result files exported for further analysis.

QC reports were consulted to give an indication of array quality. A QC score of 11/11 indicates the eleven main array parameters were passed; an example QC report is provided in Appendix 14. QC parameters include signals from spike-in controls, spatial distribution of outliers and signals from non-control spots. A score between 8 and 11 is deemed successful (Agilent pers. comm.). A full list of QC parameters can be found in the Agilent Feature Extraction Software Reference Guide v 10.5. In addition ".tif" scan images were visualised in the FE software under the 'LOG scale' view to check for artefacts on the array which may have affected spot signal intensities (Figure 3.9).

3.4.8 Microarray analysis

Microarray normalisation and analysis was performed using R console 2.10.1 employing free software packages available through the CRAN repository (http://cran.r-project.org/) and Bioconductor project (http://www.bioconductor.org).

The complete R script used for analysis can be found in Appendix 13.

3.4.8.1 Microarray normalisation

Normalisation of raw spot intensities was performed using the Limma 3.2.3 software package (Smyth, 2005). A target file was created containing the file names of each array and information including Cy sample labelling, slide number and date the array was performed. After reading in array target files, spots were weighted so that spots with signal less than or equal to background signals were excluded from analysis.

Probes were then ordered according to ID so that duplicates were adjacent and a 'SpotFile' created providing information on which probe group each spot belonged to i.e. Detox probes, control probes (Table 3.4).

Background intensities were not removed from raw spot intensities (method="none") as background normalisation of Agilent arrays may increase the variability in subsequent MA plots (red/green intensity ratio ('M') versus average intensity ('A')) and reduce correlation in dye swap results (Zahurak *et al.*, 2007). In addition high background fluorescence is not normally an issue with this array platform which was confirmed by visual inspection of array scans. An offset of 50 was added to all spot intensities to shrink log-ratios towards zero at low intensities which avoids 'fanning' in this intensity region (Smyth, 2009).

Within array normalisation was then performed using the LOESS (method="loess") function employing mean spot intensity values of probes not expected to be differentially expressed. The LOESS function is derived from LOWESS which employs a locally weighted least squares estimate of a regression fit (Cleveland, 1979). This is a moving window approach whereby a linear regression line is fitted to an area (window) of data and the process repeated for the next window of data creating a smooth regression line. Visualisation of intensity trends in the data via MA plots and box-plots of dye signal distribution was employed pre and post normalisation to establish whether further manipulations were required.

Between array normalisation was not necessary as box-plot distributions for each array were similar after within array normalisation. Between array normalisation may also increase variability and mean square error (Zahurak *et al.*, 2007).

3.4.8.2 ANOVA analysis

For statistical analysis of normalised intensities, a Java™ (Oracle Corporation) based graphical user interface (GUI) for an R package, MAANOVA (MicroArray Analysis Of Variance), was employed. This package fits either a mixed or fixed effect ANOVA model to data and calculates F statistics and p-values by permuting or consulting p-value tables. The software was developed by the Churchill group and is freely available at http://churchill.jax.org/software/jmaanova.shtml.

Kerr *et al.* (2000) first proposed the ANOVA model for microarray analysis as a method of calculating relative expression values which are not biased by 'ancillary' sources of variation. The ANOVA method estimates the relative expression of each gene within a sample and employs the weighted average expression of a gene in all samples to calculate relative expression within a given sample (Cui and Churchill, 2003). We employed a mixed effect model which allowed factors in the experiment to be defined as random or non-random. Random factors include gene and array specific effects, while condition (*i.e.* test or control) is a non-random effect. The ANOVA model treats these factors as sources of variance and expects correlation between measurements matched for a specific factor such as dye. By accounting for all sources of variation and correlations between matched factors we can be confident that differences seen are due to the non-random condition (test or control) rather than random factors (Cui and Churchill, 2003).

The data files from Limma normalisation were manipulated to meet the MAANOVA criteria (Wu *et al.*, 2009) and control spots removed as they were no longer required for analysis. An experimental 'design file' was created containing a row for every sample in the experiment with details of the array hybridised, dye label, sample ID and condition (test/control). Subsequently the amended data file was read into the J/MAANOVA GUI and log transformed before a mixed effect model was fitted. The mixed effect model included four factors (Array, Dye, Sample, Group) with Array and Sample factors designated as random.

A universal F test was then used to compare Group (condition – test/control). A minimum of six replicates per Group is recommended to calculate p-values via permutation (Cui and Churchill, 2003). As each Group in this experiment contains 5 replicates, tabulated p-values were employed after confirmation that residual plots were normally distributed. Results tables were created from both unadjusted and False Discovery Rate (FDR) (Benjamini and Hochberg, 1995; Storey and Tibshirani, 2003) corrected tabulated p-values. FDR corrects for multiple comparisons and controls the expected proportion of false positives (type 1 errors).

Volcano plots displaying log fold change for each gene and relative p-values were created within the J/MAANOVA GUI.

3.5 Results

3.5.1 RNA extraction and quality assessment

Ten RNA pools each consisting of either 5 female *An. gambiae s.s.* which survived 1 hour exposure to 0.1% bendiocarb or control papers were selected on the basis of yield (Table 3.1) and RNA quality determined via Bioanalyzer analysis (Figure 3.6, Figure 3.7). Yields varied between pools after RNA extraction with a mean concentration of 496.8ng/µl (111.6 SD) in 30µl recorded. Post DNAse treatment the average concentration was lower due to a slight increase in volume (30µl to 37.5 µl) required for enzyme kinetics. Mean concentration was 380.3ng/µl (116.6 SD); though yields were comparable before and after DNAse treatment.

Bioanalyzer traces (Figure 3.6 and Figure 3.7) revealed RNA of acceptable quality, however levels of degradation varied between samples. For instance control samples 1 and 2 and test samples 1 and 2 displayed the best quality traces with very little 'noise' between 200 and 2000 nucleotides (nt). While control samples 3 and 4 and test sample 7 showed a more jagged and raised profile between 200-2000nt indicative of low level RNA degradation. However levels of degradation in these samples were trivial when compared to other pools extracted and material collected from the field limited so samples were deemed acceptable for use.

Table 3.1. NanoDrop™ spectrophotometer (Thermo Scientific) readings from ten pools of RNA (5 control and 5 test) for microarray analysis using a 4x44K *An. gambiae s.s.* whole genome array.

Control pools consist of phenol-chloroform extracted RNA from five female *An. gambiae s.s* mosquitoes collected in Ghana 2008 which have been exposed to bendiocarb control papers for one hour. Test pools consist of RNA extracted from five female *An. gambiae s.s.* from the same region, which survived one hour exposure to 0.1% bendiocarb. NanoDrop™ readings were taken after RNA extraction and repeated post DNase treatment with Turbo DNAse (Ambion).

		tot. RNA NanoDrop™			NanoDrop™ post DNAse		
GROUP	POOL	260:280	230:260	ng/ µl	260:280	230:260	ng/ µl
CONTROL	1	1.88	0.58	726.6	1.94	0.48	603.2
	2	1.94	0.80	680.4	1.92	0.45	591.8
	3	2.00	1.08	397.1	1.99	1.01	278.8
	4	1.97	0.92	459.6	1.96	0.88	347.0
	5	1.92	1.05	452.1	1.97	0.93	324.8
TEST	1	1.86	1.34	469.0	1.93	1.12	353.6
	2	1.98	1.31	444.5	1.93	0.85	318.2
	5	2.00	1.17	419.9	2.00	0.78	306.2
	7	1.97	1.35	466.3	1.99	1.25	336.8
	8	1.96	0.63	452.8	1.93	0.46	342.8

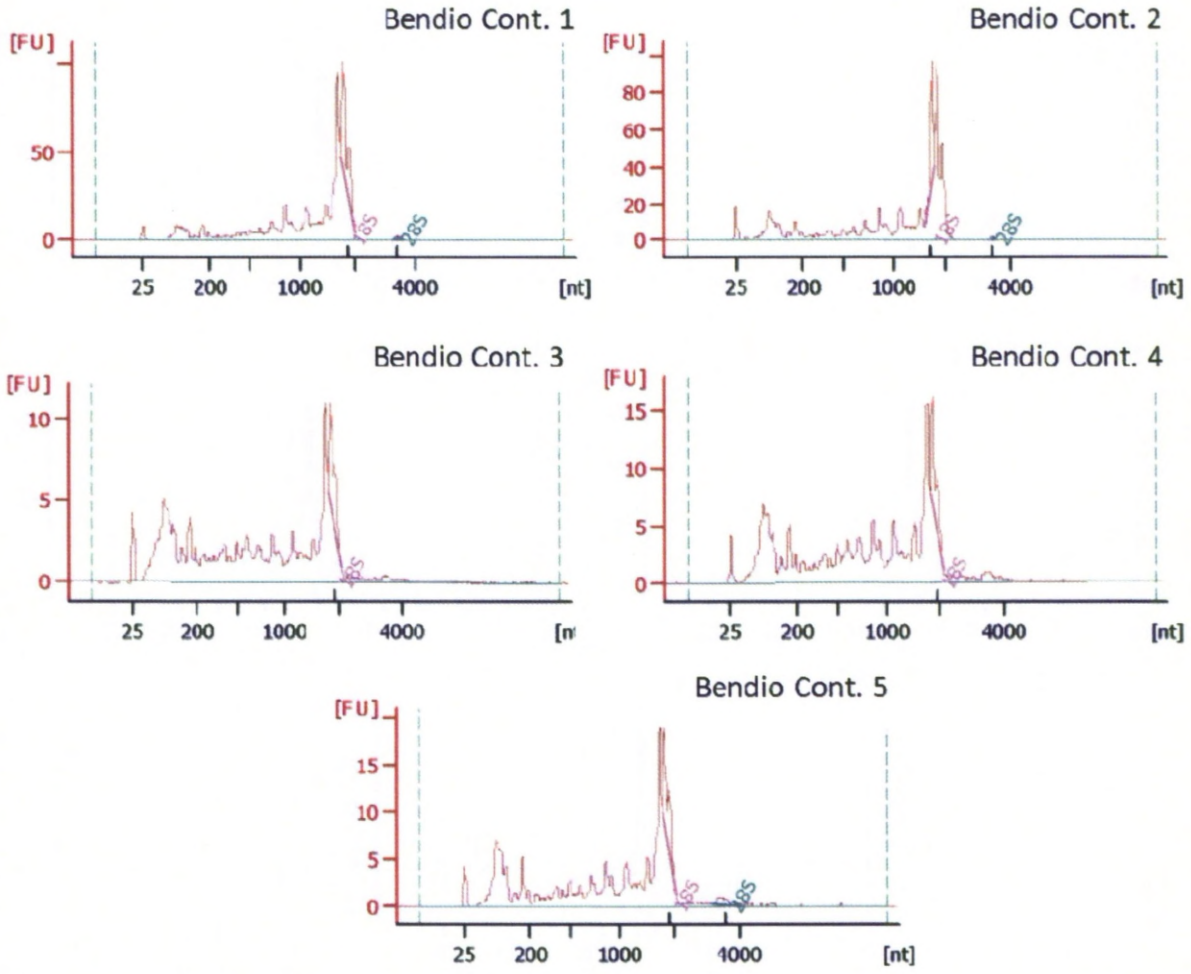

Figure 3.6. Bioanalyzer traces showing RNA quality from five control pools of RNA selected for microarray analysis.

Control pools consist of 5 female *An. gambiae s.s.* mosquitoes collected from field sites in Ghana and exposed to bendiocarb WHO control papers for 1 hour. RNA was extracted using a modified phenol-chloroform-like method and subsequently DNAse treated (Ambion). 18 and 28S ribosomal RNA is seen as three peaks at ~ 2000 nt, the profile between 200-2000nt is indicative of RNA quality, with a jagged profile indicating potential RNA degradation.

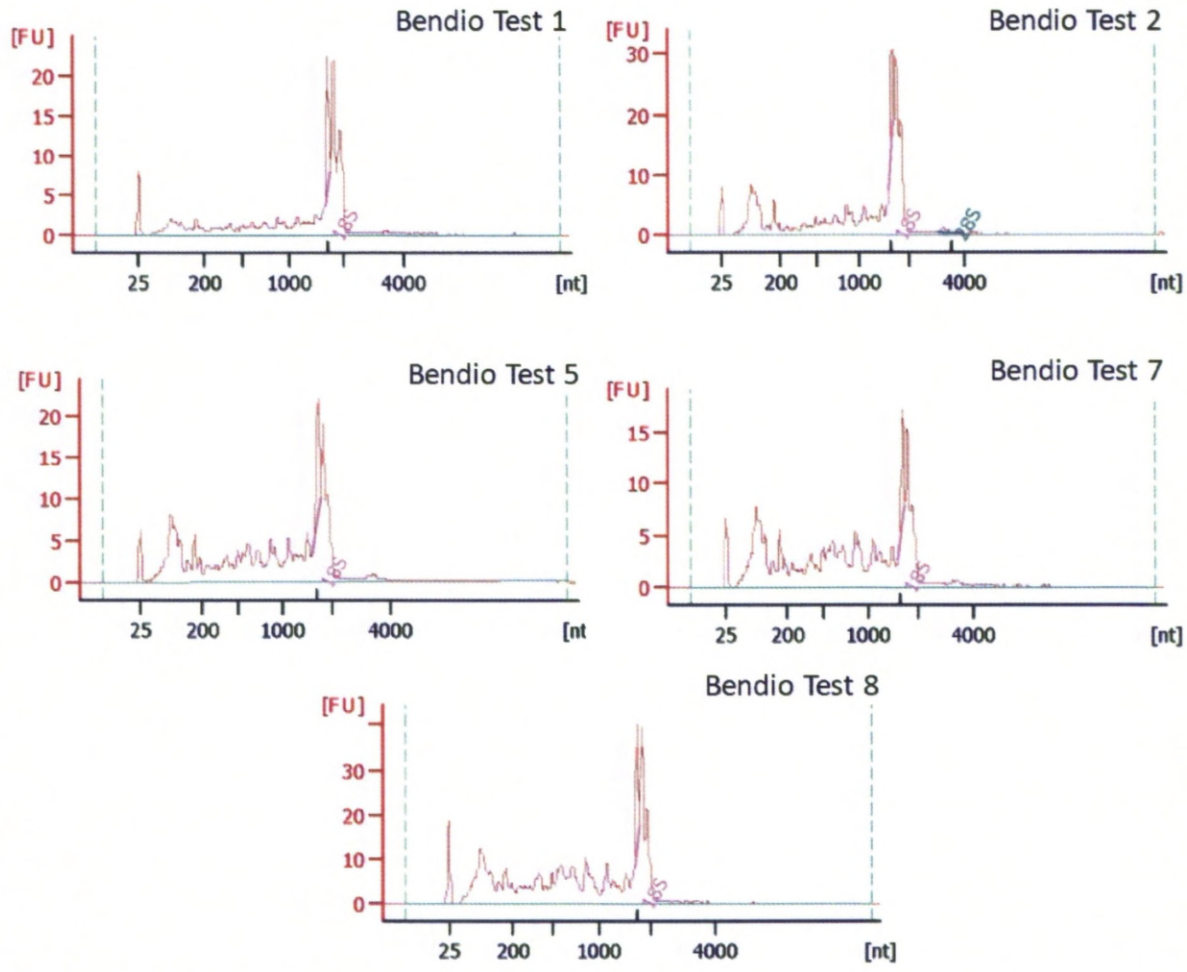

Figure 3.7. Bioanalyzer traces showing RNA quality from five test pools of RNA selected for microarray analysis.

Test pools consist of 5 female *Anopheles gambiae s.s.* mosquitoes collected from field sites in Ghana and exposed to 0.1% bendiocarb papers (WHO) for 1 hour. RNA was extracted using a modified phenol-chloroform-like method and subsequently DNAse treated (Ambion). 18 and 28S ribosomal RNA is seen as three peaks at ~ 2000 nt, the profile between 200-2000nt is indicative of RNA quality, with a jagged profile indicating potential RNA degradation.

3.5.2 RNA labelling

All ten RNA pools (5 control, 5 test) were successfully labelled with both Cy3 and Cy5 dyes using the Quick Amp Labelling Kit (Agilent). NanoDrop™ readings taken after labelling (Table 3.2 and Table 3.3) were used to calculate yield and specific activity. All labelled samples had a yield greater than 825ng and specific activity of 8 or more apart from Cy3 labelled control pool 1 (Table 3.2) with an activity of 7.6; however as this value was close to 8 re-labelling was not warranted.

Table 3.2. NanoDrop™ readings from Cy3 labelled RNA pools for microarray analysis.

Readings are shown for ten RNA pools each consisting of 5 female *An. gambiae s.s.* exposed for one hour to either bendiocarb control papers (CONTROL) or 0.1% bendiocarb papers (TEST). RNA was labelled with Cy3 cyanine dye using the Quick Amp Labelling Kit (Agilent). Calculated yield in µg ((ng/µl *30[elution volume])/1000) and specific activity pmol Cy3/µg (([pmol/µl]/[ng/µl])*1000) are also shown.

		Cy 3 LABELLED				
GROUP	**POOL**	**260:280**	**pmol**	**ng/µl**	**Yield (µg cRNA)**	**Specific activity (pmol Cy3/µg cRNA)**
CONTROL	1	2.27	6.2	814.8	24.44	7.61
	2	2.12	8.5	864.9	25.95	9.83
	3	2.13	6.4	690.9	20.73	9.26
	4	2.23	5.5	628.7	18.86	8.75
	5	1.87	14.5	822.7	24.68	17.62
TEST	1	2.17	10.9	738.3	22.15	14.76
	2	1.98	7.6	705.3	21.16	10.78
	5	2.22	5.5	624.8	18.74	8.80
	7	2.34	9	648	19.44	13.89
	8	2.32	5.4	530.3	15.91	10.18

Table 3.3. NanoDrop™ readings from Cy5 labelled RNA pools for microarray analysis.

Readings are shown for ten RNA pools each consisting of 5 female *An. gambiae s.s.* exposed for one hour to either bendiocarb control papers (CONTROL) or 0.1% bendiocarb papers (TEST). RNA was labelled with Cy5 cyanine dye using the Quick Amp Labelling Kit (Agilent). Calculated yield in µg ((ng/µl *30[elution volume])/1000) and specific activity pmol Cy5/µg (([pmol/µl]/[ng/µl])*1000) are also shown.

			Cy 5 LABELLED			
GROUP	**POOL**	**260:280**	**pmol**	**ng/µl**	**Yield (µg cRNA)**	**Specific activity (pmol Cy3/µg cRNA)**
CONTROL	1	2.26	10.1	671.9	20.16	15.03
	2	2.02	13.4	756.3	22.69	17.72
	3	2.11	11.4	698.5	20.96	16.32
	4	2.18	12.9	832.3	24.97	15.50
	5	2.22	11.3	735.6	22.07	15.36
TEST	1	2.26	8.8	709.5	21.29	12.40
	2	1.89	13.9	790.2	23.71	17.59
	5	2.13	11.8	775.2	23.26	15.22
	7	2.18	10.4	665.2	19.96	15.63
	8	2.28	5	566	16.98	8.83

Bioanalyzer analysis post labelling confirmed the quality of cRNA with large peaks between 100-2000 nucleotides recorded in all samples (example traces Figure 3.8). All pools were pursued for microarray hybridisation.

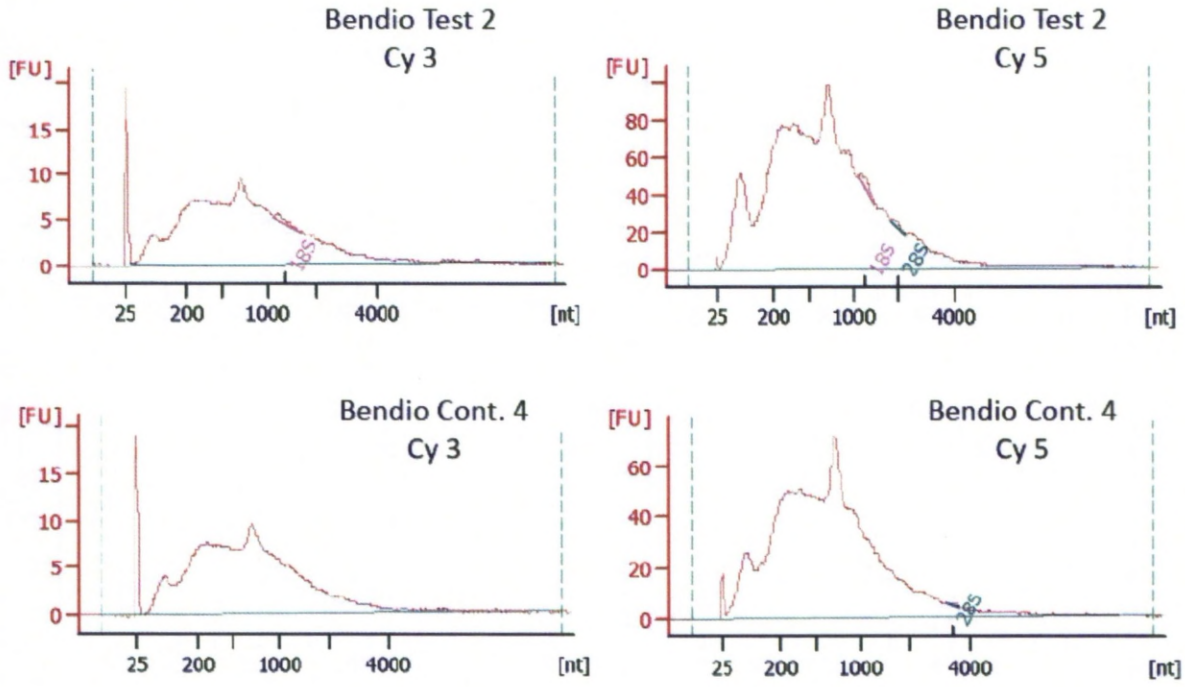

Figure 3.8. Example Bioanalyzer traces from Cy3 and Cy5 labelled cRNA pools.

Traces from RNA test pool 2 and control pool 4 are shown for both Cy3 and Cy5 labelled cRNA. Large peaks between 200 and 2000 nucleotides are typical of amplified labelled RNA. Note that the peaks from Cy5 labelled samples have a higher fluorescence value (y-axis [FU]) due to the excitation of this cyanine dye under Bioanalyzer conditions.

3.5.3 Microarray hybridisation, scanning and feature extraction

All 20 microarrays were successfully hybridised producing good quality scan images with low background signal (example Figure 3.9). After feature extraction, QC reports for each array produced maximum scores of 11.

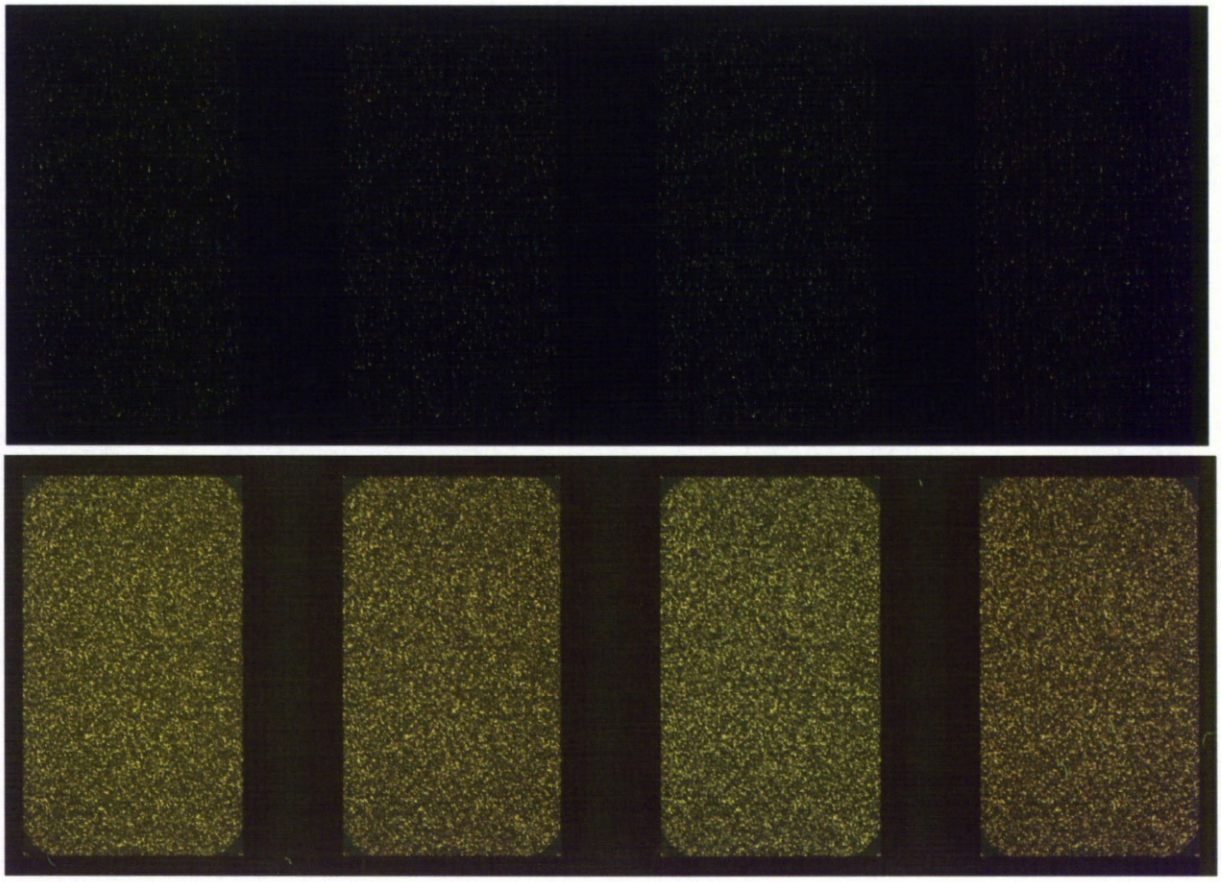

Figure 3.9. Example scan images from a custom Agilent 4x44K microarray (AGAMIMMUNODETOX) in both normal (top) and Log view (bottom).

Little background noise is visible on the four array slide in either normal or Log view, indicating successful hybridisation and washing.

3.5.4 Microarray normalisation

Array out-put files from Agilent feature extraction were read into Limma using a weight function which removed spot signals below background level. By default, Limma reads the mean foreground and median background signal for each spot on the array. A spot type file was created which colour coded each probe group on the array for identification in graphical plots (Table 3.4).

Table 3.4. Spot type colour code for probe groups on a custom Agilent 4x44K microarray (AGAMIMMUNODETOX).

Spot type	Control type	Probe name	Colour
detoxtarget	0	AG_DETOX_P*	brown
immunitytarget	0	AG_IDB_P*	blue
WGtarget	0	A_90_P*	black
CV	0	CV_A_90_P*	yellow
Brightcorner	1	*BrightCorner	orange
Darkcorner	1	DarkCorner	grey
positivecontrol	1	*	green
negativecontrol	-1	*	red

A series of graphical data plots were created to visualise the intensity trends before normalising microarray spot intensities. These plots included MA plots of red/green intensity ratio (M) versus the average intensity (A) for each spot, indicating potential dye bias, and density plots of Cy3 and Cy5 dye signals to assess overall dye distributions. Normalisation involved an offset of 50 being added to each signal and LOESS within array normalisation. The MA and density plots were re-created post-normalisation to ascertain whether data had been successfully normalised. Box-plots of dye ratio in each array were produced to establish whether between array normalisation was required. Examples of pre and post normalisation plots are shown in Figure 3.11and Figure 3.11 indicating that the normalisation process was successful; MA plots were less skewed and distributed around zero (Figure 3.10) while Cy3 and Cy5 densities were similar post normalisation (Figure 3.11). The box-plot of post normalised dye ratios indicated the spread of 'M' values (width of each box) in each of the 20 arrays was similar, so between array normalisation was not required (Figure 3.12).

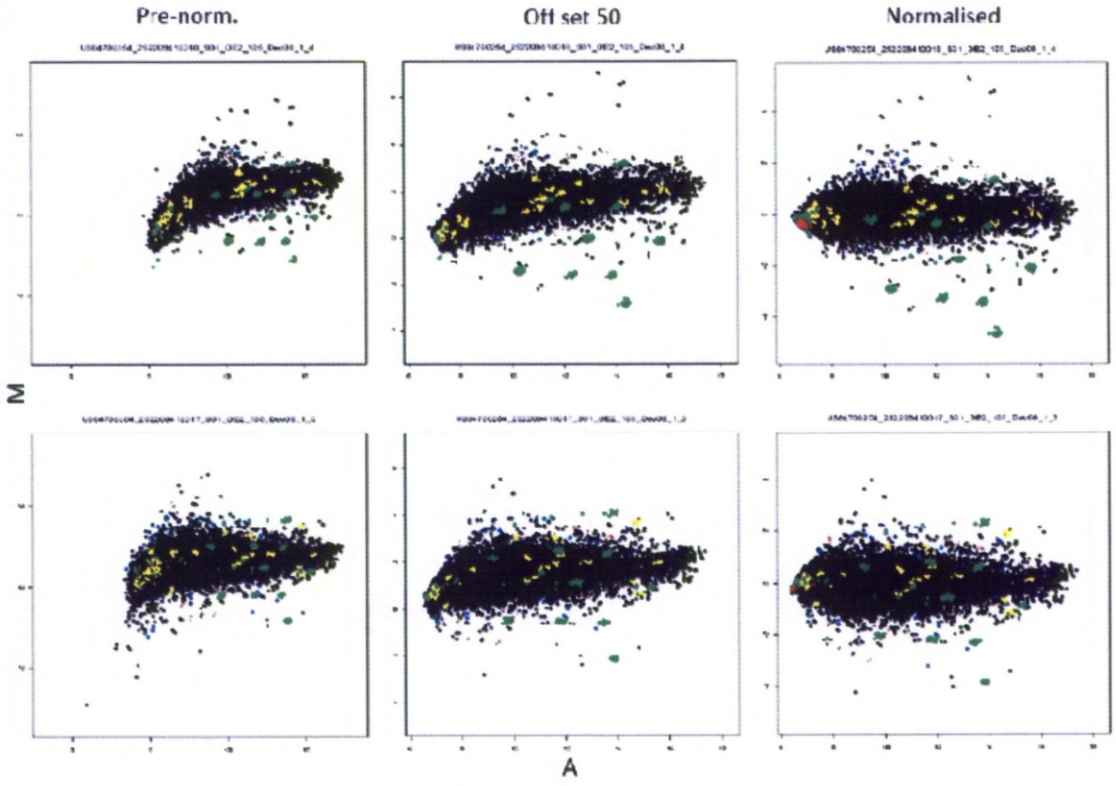

Figure 3.10. Examples of MA plots pre-normalisation of microarray data (left), after offsetting of 50 (middle) and post LOESS within array normalisation (right).

MA plots depict Cy5/Cy3 intensity ratio (M) versus the average intensity (A) for each spot. For spot colour key refer to Table 3.4.

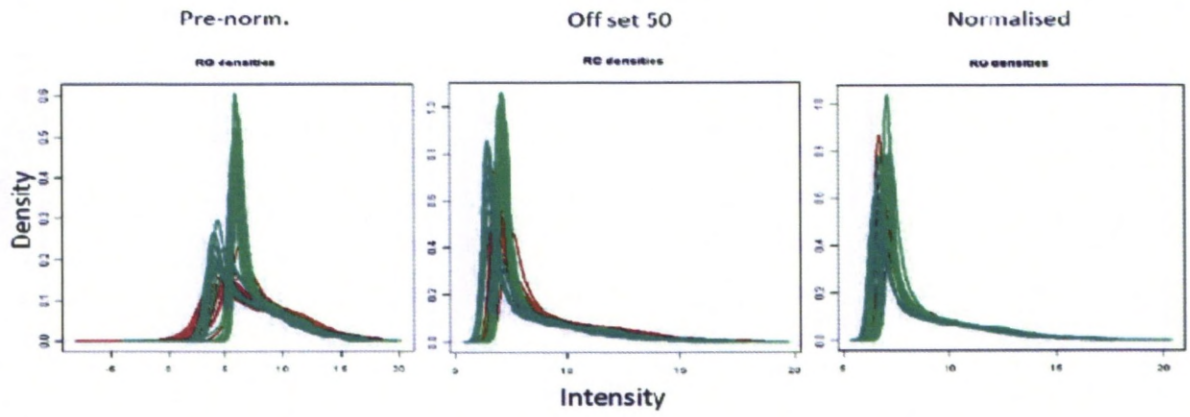

Figure 3.11. Examples of Cy3 Cy5 dye density plots from pre normalised microarray data (left), after offsetting of 50 (middle) and post LOESS within array normalisation (right).

Post normalisation the two dye signals align.

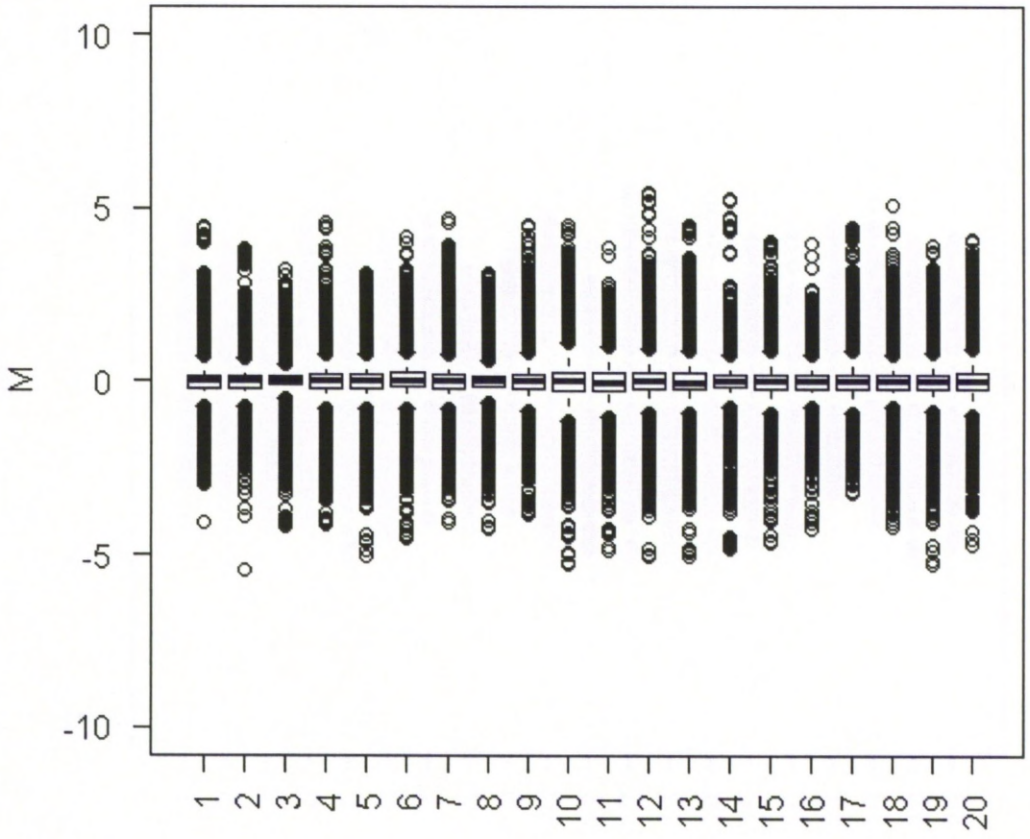

Figure 3.12. Box-plot of 'M' value distribution for each array (1-20 x-axis) in a microarray experiment comparing female *An. gambiae s.s.* which had survived one hours exposure to 0.1% bendiocarb with non-exposed controls post normalisation (offset 50 followed by LOESS within array normalisation).

M= red/green intensity ratio. The spread of M values for each array are equivalent with the width of each box similar, the absence of obvious outlier arrays meant between array normalisation was not required.

3.5.5 Microarray analysis of variance

After F-testing for statistical significance, tabulated p-values (& FDR adjusted values) were calculated for the $\log_2$ expression ratio of each array probe. A volcano plot (a scatter graph showing $\log_2$ fold change versus p-value) was created using the J/MAANOVA GUI and annotated to show a selection of genes over and under-expressed in the test group with regards to un-exposed mosquitoes (control group) (Figure 3.13)

Two strategies were used to identify candidates for additional validation.

The first was **p value ranking of all probes followed by selection of the most significant p-values (top 1%)**. A similar approach was taken by Muller *et al.* (2008b) when comparing resistant mosquitoes to non-exposed controls due to the constraints this design has on fold-change.

The second is **selection of probes with ≥ two-fold change in both test and control groups followed by p-value ranking.** This second approach was taken to identify other potentially over-expressed candidates which may not have the most significant p-values due to inherent biological 'noise' and the comparison between resistant mosquitoes and controls of mixed phenotype.

1. P-value ranking

The top 1% of ranked p-values comprised of 435 probes (from a total of 43,446 gene probes). The first fifty of these probes are listed in Table 3.5 with the full table in Appendix 15. The majority of probes in the top 50 were either ESTs (24/50) or novel genes which are yet to be defined in Ensembl (26/50). Sequences for expressed sequence tags were retrieved from GenBank (http://www.ncbi.nlm.nih.gov/genbank/) and BLAST searched in VectorBase (http://www.vectorbase.org/Tools/BLAST/) to find the closest transcript match. For novel genes, Gene Ontology terms supplied by VectorBase were employed to provide some indication of potential protein function.

Three probes with the highest p-values irrespective of test/control expression ratio are labelled on the volcano plot (two probes for AGAP008627 and probe for transcript AGAP005784 (Figure 3.13). Transcript AGAP05784 is marginally over-expressed in the test group with an expression ratio of 1.19 and 'peptidylglycine monoxygenase activity' listed under GO terms. While transcript AGAP008627 has GO term 'N-acetyltransferase activity' and the two probes for this transcript show mean expression in the control group 1.84 times that in the test. In common with the control group two probes for a transcript with putative 'N-acetyltransferase activity' (AGAP010142) are over-expressed in the test group with mean fold change of 1.55 [0.0033 SD] and p-values within the top 13. Other potential genes of interest in the top 50 ranked p-value list include putative membrane transporters showing 1.23-1.32 over expression in the test group

From the full list of 435 probes with p-values in the top 1% a number of additional candidates were identified (Appendix 15). A group of three cytochrome P450 genes (*CYP325C1*, *C2* and *C3*) all had probes which were over-expressed in the test group ranging from 1.56-2.27 fold. While two probes for a putative P450, AGAP002195, had a mean expression ratio of 1.89 [0.0127 SD] in the bendiocarb test group.

Putative transmembrane proteins show elevated expression in both the test (2.46 fold, CX819818*)* and control (1.85 fold, AGAP010471) groups. With potential signal peptides and transferases over-expressed in the control group (2.84 and 2.95 fold respectively). Two odorant binding protein showed elevated expression in the test group with two probes for *OBP19* having an expression ratio of 1.48 and *OBP45* showing marginal over expression (1.09 fold). In addition, a putative odorant receptor (AGAP004974-RA) probe was 1.25 fold over-expressed in the test group. While probes for a glutathione-S-transferase gene *GSTS1_1* displayed ~ 2 fold elevated expression in the test group along with probe for carboxylesterase gene *COEJHE2E* (3.23 fold over-expressed in the bendiocarb test group).

P-value ranking was employed rather than a cut-off for significance due to poor p-values post FDR correction. Inherent variability in field specimens in combination with the design which compared resistant mosquitoes to those of mixed phenotype is likely to have attributed to variation in the data set. This issue will be discussed further (see Discussion).

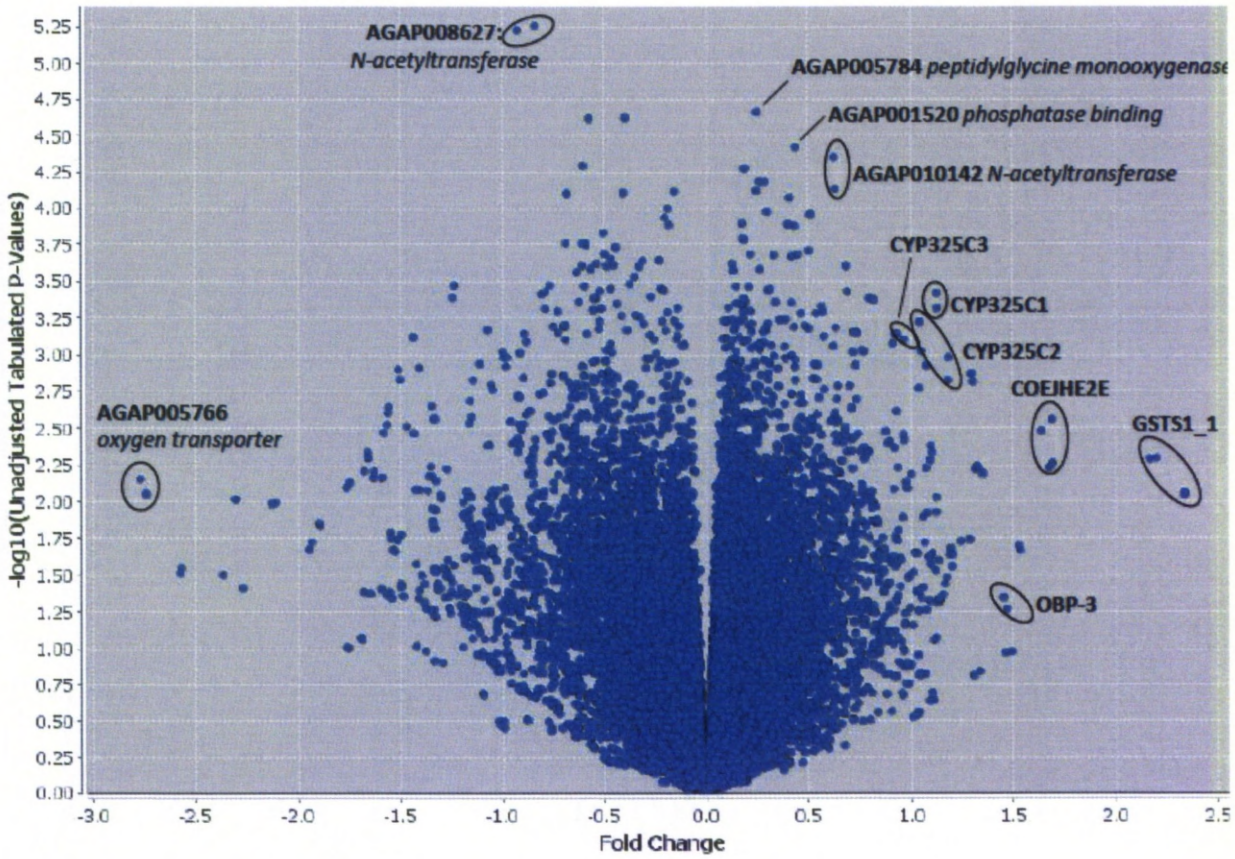

Figure 3.13. Volcano plot from microarray analysis of female *An. gambiae s.s.* surviving 1 hours exposure to 0.1% bendiocarb and unexposed females from the same area of Ghana, West Africa using a custom 4x44K *An. gambiae s.s.* whole genome microarray.

Differential expression is show as log2 fold change with positive values (right hand side) representing the resistant (surviving) mosquitoes, unadjusted tabulated p-values are shown post ANOVA analysis (F-testing)

Table 3.5. Bendiocarb microarray top-table showing top 50 p-values from all probe results ranked according to unadjusted p-value (Pval).

False discovery rate adjusted p-values (adjPval) log fold change (logFC) and fold change (FC) also shown. logFC highlighted in grey are genes over-expressed in the bendiocarb resistant *An. gambiae* s.s. samples, log FC in white are over-expressed in the non-exposed *An. gambiae s.s.* controls. Probes for genes with similar properties or from the same gene families are colour highlighted and descriptions in inverted commas are derived from Gene Ontology terms.

GeneName	Probe ID	Accession no.	Description	logFC	Pval	adjPval	FC
AgaP_AGAP008627	A_90_P094201.	AGAP008627-PA	'N-acetyltransferase activity'	-0.83	7.18E-06	0.114497	1.78
AgaP_AGAP008627	A_90_P094201.	AGAP008627-PA	'N-acetyltransferase activity'	-0.92	9.34E-06	0.114497	1.89
AgaP_AGAP005784	A_90_P030913.	AGAP005784-PA	'peptidylglycine monooxygenase activity'	0.25	2.67E-05	0.167306	1.19
AgaP_AGAP007610	A_90_P089073.	AGAP007610-PA	'signal peptide'	-0.58	2.73E-05	0.167306	1.49
BM603228	A_90_P038148.	BM603228	EST 'nucleotide binding' AGAP001645-RA	-0.57	4.01E-05	0.170065	1.48
AgaP_AGAP003902	A_90_P014972.	AGAP003902-PA	'binding'	-0.39	4.50E-05	0.170065	1.31
AgaP_AGAP001520	A_90_P050625.	AGAP001520-PA	novel	0.44	5.69E-05	0.170065	1.36
AgaP_AGAP005784	A_90_P030913.	AGAP005784-PA	'peptidylglycine monooxygenase activity'	0.25	5.99E-05	0.170065	1.19
AgaP_AGAP010142	A_90_P067856.	AGAP010142-PA	'N-acetyltransferase activity'	0.63	7.08E-05	0.170065	1.54
AgaP_AGAP004938	A_90_P121598.	AGAP004938-RA	metal ion binding	-0.60	7.94E-05	0.170065	1.52
AgaP_AGAP009594	A_90_P079266.	AGAP009594-PA	Leucine-rich repeat-containing protein 50	0.29	8.49E-05	0.170065	1.22
AgaP_AGAP010142	A_90_P067856.	AGAP010142-PA	'N-acetyltransferase activity'	0.63	1.16E-04	0.170065	1.55
BX030375	A_90_P002700.	BX030375	EST	-0.40	1.21E-04	0.170065	1.32
AgaP_AGAP005937	A_90_P018241.	XM_556607	'cellular process'	0.19	1.21E-04	0.170065	1.14
CR535150	A_90_P021113.	CR535150	EST AGAP008388-RA	0.26	1.24E-04	0.170065	1.20
CR535796	A_90_P050960.	CR535796	EST 'NAD or NADH binding'	-0.68	1.24E-04	0.170065	1.60
AgaP_AGAP004255	A_90_P035721.	XM_313174	novel	0.41	1.25E-04	0.170065	1.33
AgaP_AGAP003736	A_90_P059845.	XM_310268	'transmembrane'	0.30	1.25E-04	0.170065	1.23
BX602661	A_90_P118910.	BX602661	EST 'actin binding'	-0.61	1.38E-04	0.170321	1.53

BX621435	A_90_P024771.	BX621435	EST NAP1	0.51	1.44E-04	0.170321	1.42
AL932434	A_90_P042161.	AL932434	phosphoprotein phosphatase activity' AGAP003017	-0.27	1.92E-04	0.170321	1.20
AgaP_AGAP001520	A_90_P050625.	XM_321604	novel	0.43	1.95E-04	0.170321	1.35
BX009069	A_90_P111060.	BX009069	EST 'DNA binding' AGAP004990	-0.18	1.97E-04	0.170321	1.13
AgaP_AGAP007221	A_90_P078231.	XM_308564	'nucleotide binding'	-0.15	2.18E-04	0.170321	1.11
AgaP_AGAP006518	A_90_P025600.	XM_316550	'protein transport'	-0.50	2.19E-04	0.170321	1.42
AL930868	A_90_P122313.	AL930868	EST NAP1 'membrane transport' AGAP005701	0.40	2.20E-04	0.170321	1.32
AJ284743	A_90_P115041.	AJ284743	EST 'signal peptide' AGAP000385	0.18	2.27E-04	0.170321	1.14
BX009069	A_90_P111060.	BX009069	EST 'DNA binding' AGAP004990	-0.20	2.32E-04	0.170321	1.15
BX006423	A_90_P083807.	BX006423	EST ' isomerase activity' AGAP010217	-0.31	2.33E-04	0.170321	1.24
BM617347	A_90_P123690.	BM617347	EST - AGAP006017	-0.18	2.51E-04	0.170321	1.14
BX621435	A_90_P024771.	BX621435	EST NAP1	0.50	2.59E-04	0.170321	1.42
CR526388	A_90_P118516.	AGAP009737	EST 'Elongation factor G, mitochondrial Precursor '	0.21	2.60E-04	0.170321	1.16
AgaP_AGAP005937	A_90_P018241.	XM_556607	'cellular process'	0.18	2.61E-04	0.170321	1.13
AgaP_AGAP007941	A_90_P069487.	XM_317536	'amino transferase'	-0.60	2.63E-04	0.170321	1.52
CR535796	A_90_P050960.	CR535796	EST 'NAD or NADH binding'	-0.68	2.64E-04	0.170321	1.61
BM603228	A_90_P038148.	BM603228	EST 'nucleotide binding' AGAP001645	-0.59	2.65E-04	0.170321	1.50
AgaP_AGAP007941	A_90_P069487.	XM_317536	'amino transferase'	-0.58	2.71E-04	0.170321	1.50
BX625088	A_90_P120896.	BX625088	EST	0.22	2.71E-04	0.170321	1.17
AgaP_AGAP006359	A_90_P033083.	XM_316380	novel	-0.44	2.78E-04	0.170321	1.36
no_name	AG_IDB_P0334.	IDB_LRR09594	LRR leucine rich repeat	0.33	2.80E-04	0.170321	1.26
BX058738	A_90_P042909.	BX058738	EST ' electron carrier activity/cytochrome c1'	0.45	3.11E-04	0.170321	1.36
AgaP_AGAP006518	A_90_P025600.	XM_316550	protein transport'	-0.49	3.11E-04	0.170321	1.41
BX058738	A_90_P042909.	BX058738	EST ' electron carrier activity/cytochrome c1'	0.42	3.35E-04	0.170321	1.34
AgaP_AGAP007610	A_90_P089073.	AGAP007610-PA	'signal peptide'	-0.46	3.36E-04	0.170321	1.38
BX050182	A_90_P098240.	BX050182	'ribosomal'	0.23	3.42E-04	0.170321	1.17

BM628651	A_90_P112361.	BM628651	EST - AGAP006017	0.68	3.44E-04	0.170321	1.61
AgaP_AGAP000325	A_90_P108110.	XM_310792	'lysine-tRNA ligase activity'	-0.54	3.47E-04	0.170321	1.45
AgaP_AGAP002548	A_90_P069050.	XM_312391	'transmembrane/signal peptide'	0.27	3.49E-04	0.170321	1.20
AJ284743	A_90_P115041.	AJ284743	EST 'signal peptide' AGAP000385	0.19	3.60E-04	0.170321	1.14
AgaP_AGAP005535	A_90_P045709	XM_556176	novel	-0.45	3.68E-04	0.170321	1.37

2. Fold change cut-off

Results 'top-tables' were created showing all genes approximately 2-fold over-expressed in either the test or control groups (≥0.95 $\log_2$), ranked according to p-value along with accession numbers and gene descriptions/Gene Ontology. The two fold cut-off is highly arbitrary but is convention in many microarray studies (Vontas *et al.*, 2007; Muller *et al.*, 2007; Strode *et al.*, 2006; Huang *et al.*, 2001) and produced a manageable fold change derived data set to pursue in addition to the p-value ranked results set. However if resources had permitted a 'self-self' hybridisation could have been performed to calculate a fold-change cut-off according to the method of Yang *et al.* (2002).

Genes over-expressed in the Test group

A summary table containing the first 50 probes greater than 0.95 $\log_2$ over expressed in the test group is presented below (Table 3.6) and the full table in Appendix 16. Probes displaying greater than 0.95 $\log_2$ expression ratios in favour of the test group will be discussed in the context of unadjusted p-values.

The cytochrome P450 gene family (*CYP325C1-3*) identified via p-value ranking also met the 2 fold cut-off so were top of the p-value ranked table of probes displaying ≥0.95 $\log_2$ fold-change in the bendiocarb resistant test group. Encouragingly, duplicate probes for the same gene displayed very similar fold changes, which is an indication of consistency across the arrays. Although not all probes for *CYP325C* genes displayed more than two fold over-expression. The full set of probe results for the *CYP325C* family is shown in Appendix 19.

CYP325C1 probe AG_DETOX_P457 gave the highest fold change (2.18 mean, 0.0025 standard deviation for two duplicate probes) and lowest p-values (p=0.000472-0.00624). The other two probes and their duplicates for this gene did show a similar up-regulation in the test group however the fold-changes were not equivalent with a mean of 1.58 (0.0257 SD.) and p-values between p=0.002123-0.004572. Two probes for *CYP325C2* gave greater than two fold over-expression (2.07-2.27 fold) with p-values ranging from p= 0.000759 -0.001961. Two additional duplicated probe for this gene were also expressed at a higher level in the test group but were just outside of the 2 fold cut-off (1.78 mean fold, 0.1012 SD, p= 0.001259-0.004358). The *CYP325C3* gene provided the least consistent results across unique probes. Only one probe (AG_DETOX_P461) reached the >0.95 $\log_2$ fold defined cut off (1.95 fold change, p= 0.001048) with the duplicate probe just below (0.93 $\log_2$ FC). The remaining probes designed against *CYP325C3* gave a mean fold change of 1.24 (0.2219, SD.).

All probes sets designed against the glutathione-S-transferase gene identified via p-value ranking, *GSTS1_1*, met the expression ratio cut-off. Whole genome designed probe BX050330 provided the lowest p-value (0.001596 - 0.002042) with a mean fold

change of 2.08 (0.0368 standard deviation). The remaining probes for *GSTS1_1* provided fold changes between 5.07-2.01 with a mean fold change of 3.20 (1.3951 SD.) p=0.006264-0.027569.

Similarly all probes (3 duplicate sets) for carboxylesterase gene *COEJHE2E,* also identified via p-values, passed the cut-off criteria. The mean fold change for the 6 individual probes was 3.21 with a standard deviation of 0.0447 reflecting concordance between probes. Unadjusted p-values were between p=0.003548-0.007368.

Two cuticular proteins were also amongst those up-regulated in the test group. Probes for *CPLC4* and *CPLC17* were ~ 2 fold over expressed with the mean expression of C4 and C17, 1.95 and 2.05 respectively. P-values ranged between 0.004589-0.026909.

In addition to the odorant binding proteins identified through p-value ranking (*OBP19* and *OBP45*) *OBP3* was found to be over-expressed in the test group with a higher fold change compared to the previous OBPs identified (2.75 fold change [0.0214 SD], p= 0.051484-0.061408); however p-values were not in the top 1%.

Genes over-expressed in the control group

A table of genes greater than two fold over-expressed in the control groups (or under-expressed in the test group) was also created and can be found in Appendix 17.

The majority of genes which appeared to be down regulated in the test group or alternatively more expressed in the control group are not well characterised (Appendix 17). However probes for two cytochrome P450s *CYP6M1* and *CYP9J3* and three GSTs, *GSTO1, GSTD1_5, GSTD1* along with carboxylesterase *COEA2G* showed greater than 0.95 $\log_2$ expression ratio in favour of control RNA pools. In addition a number of genes potentially linked to metabolic processes were expressed at a higher level in the control group including various peroxidises and caspases.

Table 3.6. Bendiocarb microarray top-table showing gene probes with greater than ~ 2 fold (>0.95 log_2 fold) over expression in the bendiocarb resistant group of *An. gambiae s.s.* compared with the unexposed controls from the same area of Ghana sorted according to unadjusted p-values (Pval).

False discovery rate adjusted p-values (adjPval) log fold change (logFC) and fold change (FC) are also shown. Probes for genes with similar properties or from the same gene families are colour highlighted and gene descriptions in inverted commas are derived from Gene Ontology terms.

Gene name	Probe ID	Accession no.	Description	logFC	Pval	adjPval	FC
CYP325C1	AG_DETOX_P457	AGAP002207-RA	Cytochrome P450 monooxygenase;	1.12	4.72E-04	0.170321	2.17
CYP325C1	AG_DETOX_P457	AGAP002207-RA	Cytochrome P450 monooxygenase;	1.13	6.24E-04	0.173082	2.19
CYP325C2	AG_DETOX_P460	AGAP002205-RA	Cytochrome P450 monooxygenase;	1.04	7.59E-04	0.173706	2.06
CYP325C3	AG_DETOX_P461	AGAP009696-RA	Cytochrome P450 monooxygenase;	0.96	1.05E-03	0.177791	1.95
CYP325C2	AG_DETOX_P460	AGAP002205-RA	Cytochrome P450 monooxygenase;	1.05	1.25E-03	0.18496	2.07
CYP325C2	AG_DETOX_P459	AGAP002205-RA	Cytochrome P450 monooxygenase;	1.18	1.39E-03	0.18496	2.27
BX050330	A_90_P082266.	AGAP010404	EST 'Glutathione S-transferase GSTS1_1'	1.08	1.60E-03	0.188937	2.11
CX819818	A_90_P086003.	AGAP010701-RA	EST 'Membrane prot.'	1.30	1.81E-03	0.188937	2.46
CYP325C2	AG_DETOX_P459	AGAP002205-RA	Cytochrome P450 monooxygenase;	1.18	1.96E-03	0.188937	2.27
CX819818	A_90_P086003.	AGAP010701-RA	EST 'Membrane prot.'	1.30	2.02E-03	0.188937	2.46
BX050330	A_90_P082266.	AGAP010404	EST 'Glutathione S-transferase GSTS1_1'	1.04	2.04E-03	0.188937	2.06
COEJHE2E	AG_DETOX_P400	AGAP005834-RA	Carboxylesterase;	1.69	3.55E-03	0.193737	3.23
COEJHE2E	AG_DETOX_P400	AGAP005834-RA	Carboxylesterase;	1.64	4.20E-03	0.193737	3.12
CPLC4	A_90_P125986.	AGAP003334-RA	Cuticular Protein 4	0.96	4.59E-03	0.193737	1.95
CPLC4	A_90_P125986.	AGAP003334-RA	Cuticular Protein 4	0.96	4.95E-03	0.193737	1.95
AgaP_AGAP000356	A_90_P033026.	XM_310756	AGAP000356-PA extracellular prot	1.10	5.36E-03	0.193737	2.14
AgaP_AGAP000356	A_90_P033026.	XM_310756	AGAP000356-PA extracellular prot	1.10	5.90E-03	0.193737	2.14
GSTS1_1	AG_DETOX_P704	AGAP010404-RA	Glutathione S-transferase;	2.20	6.26E-03	0.193737	4.59

GSTS1_1	AG_DETOX_P704	AGAP010404-RA	Glutathione S-transferase;	2.17	6.32E-03	0.193737	4.50
HPX3	AG_IDB_P0449.	AGAP003714-RA	'response to oxidative stress'	1.09	6.70E-03	0.193737	2.13
COEJHE2E	AG_DETOX_P064	AGAP005834-RA	Carboxylesterase;	1.69	6.87E-03	0.193737	3.23
COEJHE2E	AG_DETOX_P399	AGAP005834-RA	Carboxylesterase;	1.69	6.97E-03	0.193737	3.23
COEJHE2E	AG_DETOX_P399	AGAP005834-RA	Carboxylesterase;	1.68	7.22E-03	0.193737	3.20
AgaP_AGAP012394	A_90_P014238.	XM_320164	'oxidoreductase activity'	1.32	7.22E-03	0.193737	2.50
HPX3	AG_IDB_P1022.	AGAP003714-RA	'response to oxidative stress'	1.00	7.26E-03	0.193737	2.00
AgaP_AGAP005848	A_90_P020836.	XM_315874	'vasopressin receptor activity/binding'	1.33	7.32E-03	0.193737	2.51
COEJHE2E	AG_DETOX_P064	AGAP005834-RA	Carboxylesterase;	1.68	7.37E-03	0.193737	3.20
HPX3	AG_IDB_P0449.	AGAP003714-RA	'response to oxidative stress'	1.07	7.54E-03	0.193737	2.10
AgaP_AGAP012394	A_90_P014238.	XM_320164	'oxidoreductase activity'	1.31	7.61E-03	0.193737	2.48
AgaP_AGAP005848	A_90_P020836.	XM_315874	'vasopressin receptor activity/binding'	1.35	7.94E-03	0.193737	2.55
PX14	AG_DETOX_P238	AGAP003714-RA	Peroxidase;	0.98	9.34E-03	0.193737	1.97
PX14	AG_DETOX_P238	AGAP003714-RA	Peroxidase;	0.98	1.02E-02	0.193737	1.97
GSTS1_1	AG_DETOX_P216	AGAP010404-RA	Glutathione S-transferase;	2.34	1.06E-02	0.193737	5.06
GSTS1_1	AG_DETOX_P216	AGAP010404-RA	Glutathione S-transferase;	2.34	1.07E-02	0.193737	5.06
AgaP_AGAP012395	A_90_P001781.	XM_320163	AGAP012395-PA	1.12	1.18E-02	0.193737	2.17
AgaP_AGAP012395	A_90_P001781.	XM_320163	AGAP012395-PA	1.12	1.46E-02	0.193737	2.17
AgaP_AGAP005009	A_90_P025801.	XM_315115	AGAP005009-PA	1.08	1.49E-02	0.193737	2.11
AgaP_AGAP005009	A_90_P025801.	XM_315115	AGAP005009-PA	1.06	1.57E-02	0.195002	2.08
BX008814	A_90_P091076.	BX008814	EST unknown	0.96	1.59E-02	0.195002	1.95
BX467419	A_90_P038422.	BX467419	EST 'sphingomyelin phosphodiesterase activity'	1.29	2.18E-02	0.195136	2.45
BX467419	A_90_P038422.	BX467419	EST 'sphingomyelin phosphodiesterase activity'	1.24	2.23E-02	0.195136	2.36
CPLC17	A_90_P056541.	AGAP006149-RA	Cuticular Protein 17	1.06	2.31E-02	0.195136	2.08

AgaP_AGAP008487	A_90_P062948.	XM_316957	'sphingomyelin phosphodiesterase activity'	1.53	2.39E-02	0.195136	2.89
AgaP_AGAP010404	A_90_P043882.	AGAP010404-RA	Glutathione S-transferase (GST class-sigma)	1.19	2.52E-02	0.195136	2.28
AgaP_AGAP008487	A_90_P062948.	XM_316957	'sphingomyelin phosphodiesterase activity'	1.53	2.57E-02	0.195136	2.89
CPLC17	A_90_P056541.	AGAP006149-RA	Cuticular Protein 17	1.01	2.69E-02	0.195136	2.01
AgaP_AGAP004156	A_90_P053637.	XM_313040	AGAP004156-PA	1.10	2.71E-02	0.195136	2.14
AgaP_AGAP005717	A_90_P022831.	AGAP005717-RA	Lysozyme c-6.	1.03	2.73E-02	0.195136	2.04
LRR	AG_IDB_P0367.	IDB_LRR11930	leucine rich repeat	1.20	2.74E-02	0.195136	2.30
AgaP_AGAP010404	A_90_P043882.	AGAP010404-RA	Glutathione S-transferase (GST class-sigma)	1.19	2.76E-02	0.195136	2.28

3.6 Discussion

3.6.1 Genes most significantly up-regulated

The top 50 differentially expressed probes after p-value ranking highlighted genes with links to N-acetyltransferase activity, membrane transport and DNA binding (Table 3.5). However all 50 genes were either defined as novel genes in Ensembl or were ESTs which on BLAST search were derived from novel transcripts, making it potentially difficult to infer links to phenotype and to pursue genes proteomically. In addition the fold changes recorded for these genes were low, with a maximum expression ratio of 1.89 in favour of the control group and the mean fold change of 1.35 [0.1809 SD]. Fold change restriction in microarray experiments comparing resistant individuals to a group of mixed phenotype has been highlighted previously (Muller *et al.*, 2008b). When considering all 435 genes with p-values in the top 1%, candidates in the test group with defined properties and higher expression ratios were identified including P450 genes *CYP325C1-3*, odorant binding proteins 19 and 45, *GSTS1_1* and carboxylesterase *COEJHE2E*.

3.6.1.1 N-acetyltransferase activity

Two genes with N-acetyltransferase activity listed under their GO terms were amongst the very top p-values for all probes. These were AGAP010142 over-expressed 1.55 fold in the test group and AGAP008627 over-expressed 1.84 fold in the controls. N-acetyltransferases are enzymes which catalyse the transfer of acetyl groups from acetyl-CoA to arylamine compounds and in humans are involved in the metabolism of drugs (Sim *et al.*, 2008) such as anti tubercular compound isoniazid (Evans *et al.*, 1960), as well as having a role in carcinogen metabolism (Upton *et al.*, 2001).

The putative N-acetyltransferase over-expressed in the test group is orthologous to dopamine-N-acetyltransferase (DAT) from both *Drosophila melanogaster* (E value 7.5 e^{-41}) and *Culex quinquefasciatus* (E-value 1.3 e^{-07}), sharing between 50-60% sequence ID. DATs use Coenzyme A to acetylate dopamine to N-acetlydopamine which is a 'classic' sclerotization compound in insects (Hintermann *et al.*, 1996; Karlson and Sekeris, 1962). Sclerotization involves cross-linking of proteins and chitin to harden the cuticle particularly during eclosion and moulting (Hopkins and Kramer, 1992). Enhanced sclerotization through elevated DAT levels may result in a cuticle which is more resistant to insecticide penetration; though this is speculative. However, cuticular mechanisms of resistance are implicated in a other insect species (Wood *et al.*, 2010; Pedrini *et al.*, 2009; Puinean *et al.*, 2010) with cuticle thickening potentially involved in some cases (Pedrini *et al.*, 2009; Wood *et al.*, 2010).

The gene elevated in the control group with N-acetyltransferase gene ontology does not possess any characterised gene orthologues. It is therefore more difficult to speculate on a role for this protein.

3.6.2 Genes greater than two-fold expressed

This second approach for candidate gene identification was employed to provide a broad picture of differential expression between the two groups in light of poor p-values. A dual approach enabled identification of the most significantly expressed genes and those with the highest differential expression. A number of genes and protein families were identified by both approaches.

3.6.2.1 Genes greater than two-fold expressed in the test group

Genes of potential interest identified in the test group by this approach represented enzyme families with previous links to insecticide resistance (P450, GSTs, esterases) as well as a number of novel gene sets.

Cuticular genes *CPLC4* and *CPLC17* were up-regulated ~ 2 fold in the test group with respect to non-exposed controls which may suggest a role for the cuticle in bendiocarb resistance (see *Candidate genes* Discussion section). While three probes for a putative peroxidase gene, AGAP003714, indicated greater expression in the test group (~2 fold). Peroxidases catalyse the conversion of reactive oxygen species such as hydrogen peroxide to less dangerous compounds. Elevated expression of these enzymes may be linked to an oxidative stress response induced by insecticide exposure. Enhanced peroxidase expression has been previously reported in a DDT resistant strain of *An. gambiae* (David *et al.*, 2005).

Two trypsin precursor genes also appeared to be expressed at a higher level in the bendiocarb test group with a mean expression ratio of 2.10 recorded for trypsin precursors 6 and 7. Trypsins are serine proteases normally associated with digestion. In mosquitoes trypsins have been linked to plasmodium infection (Shahabuddin *et al.*, 1996) and blood meal digestion in *Anopheles* (Marinotti *et al.*, 2006; Billingsley and Hecker, 1991; Muller *et al.*, 1993). Trypsins 6 and 7 are among a number of trypsins which exhibited constitutive expression in unfed *An. gambiae* females via northern blot and quantitative PCR analysis (Muller *et al.*, 1993). Potential links to insecticide resistance have been previously inferred with two trypsin genes reportedly over-expressed in a deltamethrin resistant strain of in *Culex pipiens pallens* compared to a susceptible strain (Gong *et al.*, 2005). Subsequent co-expression of these two trypsins in a mosquito cell line resulted in significantly improved viability on exposure to deltamethrin when compared to cells not expressing the proteases. However it may be difficult to infer a role *in vivo* from results derived from cell lines, with techniques such as gene knock-down in adult insects a more convincing approach to validate a role in resistance (Bautista *et al.*, 2009; Lycett *et al.*, 2006). Proteases are also over expressed in DDT resistant strains of *Drosophila* (Pedra *et al.*, 2004) and elevated protease activity is associated with DDT resistance in house fly *Musca domestica* (Ahmed *et al.*, 1998). It is hypothesised that elevated proteolytic activity may have a role in meeting energy demands during stress response on exposure to insecticides. Proteases may act to increase the supply of amino-acid precursors for synthesis of

detoxifying enzymes or perform proteolytic activation/modification of enzymes involved in resistance (Pedra *et al.*, 2004; Gong *et al.*, 2005).

Genes identified by both p-value ranking and fold change cut-off approaches were those with the highest and most significant over expression. In the test group these included Glutathione-S-transferase *GSTS1_1* and P450 genes *CYP325C1-3*. However some inconsistencies in probe signals resulted in mean fold changes of below 2 for these P450 genes (mean fold change: 1.78 - C1, 1.97- C2 and 1.38 - C3). A single probe for carboxylesterase gene *COEJHE2E* had a p-value in the top 1% with a fold change of 3.23 in favour of the test group. While all six probes designed for this esterase gene displayed greater than 2 fold expression ratios. Representatives of the odorant binding protein family (OBP3, 19, 45) were also identified by both approaches along with a candidate odorant receptor gene (see *Candidate genes* Discussion section).

3.6.2.2 Genes greater than two-fold expressed in the control group

Various novel genes as well as members of the 'detox' enzyme families were represented in probes over-expressed >0.95 $\log_2$ fold in the control group. Probes for two cytochrome P450s, *CYP6M1* and *CYP9J3*, showed elevated expression of 3.50 and 2.14 fold respectively. *CYP6M1* is a member of the family most commonly associated with insecticide resistance, the CYP6 family (Hemingway *et al.*, 2004); however, this gene has not been linked to resistance to date unlike CYP6 member *CYP6M2* (Muller *et al.*, 2007; Djouaka *et al.*, 2008). In common with *CYP6M1* a link between P450 *CYP9J3* and insecticide resistance has not been reported. These two enzymes may have a role in general xenobiotic detoxification.

Four GSTs also exhibited elevated expression in the control group and three of these were from the Delta class; *GSTD1*, *D1_5* and *D10* which had expression ratios between 2.08-2.44. The delta class of GST along with the epsilon class are two insect specific GST classes (Ding *et al.*, 2003). Gene *GSTD1* contains five exons which are alternatively spliced to produce four variants in *An. gambiae* including *GSTD1_5* (Ranson *et al.*, 1998). The probe showing elevated expression for *GSTD1* originated from the Agilent designed probe sets and on BLAST search aligned with splice variant D1_5 so is likely to confirm the results from the *GSTD1_5* probes. *GSTD1* was previously shown to be elevated in a DDT resistant strain of *Drosophila* via western blot, with recombinant *GSTD1* protein exhibiting DDTase activity (Tang and Tu, 1994). Recombinant *GSTD1_5* and *D1_6* from a resistant strain of *An. gambiae* also exhibited some DDTase activity (Ranson *et al.*, 1997); however, this only accounted for 6% of DDT metabolism in this strain with an epsilon class GST, *GSTe2*, subsequently found to be responsible for the majority of DDTase activity (Ranson *et al.*, 2001). Delta class GSTs may have a more general role in defence against oxidative stress potentially induced by insecticide exposure. A delta class GST over-expressed in a pyrethroid resistant strain of plant hopper *Nilaparvata lungens* showed high peroxidase activity when expressed as a recombinant protein (Vontas *et al.*, 2001; Vontas *et al.*, 2002).

The authors suggest a role for this GST in protecting against pyrethroid induced lipid peroxidation.

An Omega class GST, *GSTO1*, was over-expressed 1.96 fold in the control group with expression of this class of GST also likely to be linked to an antioxidant response. *GSTO1* is classed among a number of 'anti-oxidant' genes in the honey bee (Corona and Robinson, 2006) while recombinant *GSTO1-1* from *Anopheles cracens* showed thiol transferase activity suggestive of a role in oxidative stress response (Wongtrakul *et al.*, 2010), with a similar role reported for human Omega GSTs (Board *et al.*, 2000).

A number of hypotheses could explain the elevated expression of these genes in the control group with respect to the test. Firstly it must be emphasised that this control group has not undergone selection so is likely to contain a mixture of resistance phenotypes. Therefore genes elevated in the control group are not likely to be genes linked specifically to susceptibility or a lack of resistance rather representative of expression in the general population. The test group has also been effectively selected for its ability to survive bendiocarb exposure so individuals in the test group are likely to share a similar expression profile which enhances insecticide detoxification when exposed to bendiocarb. The elevated expression of genes conferring resistance may be costly to the insect, so resources may need to be diverted within the mosquito to meet these energy demands. This may involve constitutive or induced diversion of transcription and protein synthesis away from general house-keeping genes towards specific proteins which aid survival when exposed to lethal compounds. For instance general detoxifying enzymes may be replaced in favour of specific insecticide metabolising proteins or more efficient stress-responders. This could explain the 'detox' gene expression profiles of the test and control groups, whereby the control groups appear to over-express general detoxifiers whereas the test group may have diverted expression to specific enzymes potentially capable of detoxifying bendiocarb. The elevated trypsin profile of the test group, as discussed, may have a role in meeting amino-acid demand for expression of detoxifying enzymes. This reasoning is, however, speculative and it remains to be seen whether genes over-expressed in the test group do indeed have a role in bendiocarb resistance.

An alternative hypothesis for differential expression is allelic variation between the test and control groups. Resistant individuals may carry mutated genes encoding proteins with potentially enhanced metabolic activity against insecticides. If these mutations are coincident with a lack of microarray probe hybridisation, spurious expression ratios in favour of the control group could arise. To test this hypothesis extensive sequencing of genes in the test and control pools would be required.

3.6.3 Candidate genes

3.6.3.1 Odorant binding proteins - OBPs 3, 19, 45

Odorant binding proteins, a peptide family with representatives elevated in the test group, are not members of any 'detox' gene families putatively linked to resistance to date. The identification of these genes as potential candidates supports the use of a genome-wide screen rather than a candidate gene approach.

Insect OBPs are classically involved in detection of semio-chemical signals released from potential hosts and/or mates which enter the insect antennae via diffusion through pores (Zhou, 2010). Once inside the insect, compounds are transported by OBPs and other proteins to the olfactory receptors across the sensillum lymph (Zhou *et al.*, 2010). OBPs are thought to change conformity when bound to a chemical compound (Zhou *et al.*, 2010) and can also display broad ligand specificity (Pelosi *et al.*, 2006). Recombinant *OBP22* from *Ae. aegypti* was able to interact with a number of organic compounds, while expression of *OBP22* was detected in the antennae/proboscis as well as male reproductive apparatus, spermatheca in females and tracheal spiracles, via immunofluorescence (Li *et al.*, 2008). The combination of broad ligand specificity and expression profile led authors to hypothesise multiple functions for this protein. While in *An. gambiae* the tissue specific expression of 32 putative OBP was analysed using semi-quantitative reverse transcription PCR, with OBP expression recorded in the head, legs and body of both sexes (Li *et al.*, 2005). The expression of *OBP3* was reported to be restricted to head tissue with higher expression in females; the authors postulated a role in olfaction from the *OBP3* tissue profile. However PCR based quantification from dissected animals may be prone to cross-tissue contamination and may be limited to dissectible tissues, while a fully quantitative approach would have been more accurate. As microarray experiments here employed whole mosquito extracts the tissue specific expression profile of *OBP3* cannot be inferred.

Three OBPs in total were over-expressed in the test group with *OBP3* showing the highest level of expression at a ratio of 2.75 while OBP19 and OBP45 probes results were more significant according to p-values though expression ratios lower (1.09-1.48). Interestingly a potential odorant receptor was also up-regulated 1.25 fold with p-values within the top 1% of all probes. It is feasible that these OBPs may have broad substrate specificity potentially interacting with insecticides entering the mosquito on exposure. Indeed non-pheromone binding has been reported in OBPs of silk worm *Bombyx mori* (Hooper *et al.*, 2009) and *Ae. aegypti* (Li *et al.*, 2008). OBPs potentially interacting with insecticidal volatiles may then transport these compounds to odorant receptors instigating signal transduction, with potential behavioural and/or detoxification responses ensuing.

3.6.3.2 Cuticular genes – *CPLC4* and *17*

Both *CPLC4* and *CPLC17* are members of a family of 18 low sequence complexity cuticular proteins annotated in *An. gambiae s.s.* (He *et al.*, 2007). Over-expression of

two cuticular protein pre-cursors of the same CPLC family was also reported from microarray analysis of pyrethroid resistant *An. gambiae s.s* (Djouaka *et al.*, 2008; Awolola *et al.*, 2009). While gene *CPLC8* was reportedly over-expressed ~ 5 fold in a pyrethroid resistant strain of *An. stephensi* when compared to a susceptible colony via microarray. The authors speculated a role in cuticle thickening and reduced insecticide penetration (Vontas *et al.*, 2007).

Reduced cuticular penetration has previously been associated with pyrethroid resistance in *Musca domestica* (Plapp and Hoyer, 1968; DeVries and Georghiou, 1981), deltamethrin resistance in the cotton bollworm *Helicoverpa armigera* (Ahmad *et al.*, 2006) and S-Fenvalerate resistance in the diamondback moth *Plutella xylostella* (Noppun *et al.*, 1989). While cuticle thickening, with the potential to slow insecticide absorption increasing the efficacy of detoxification, has been linked to insecticide resistance in Chagas vector *Triatoma infestans* and malaria vector *An. funestus*. Pedrini *et al.* (2009) reported a significant increase in cuticle thickness in a pyrethroid resistant strain of *T. Infestans* compared with a susceptible strain using scanning electron microscopy (SEM), with cuticle from resistant specimens almost double the width of those from susceptible insects ($p<0.0001$). Although as the two separate laboratory strains were isolated from different regions of Argentina there are potential strain related and geographical confounders. While in *An. funestus* cuticle thickening was reportedly associated with permethrin resistance with 9.5-10% increase in mean cuticle thickness, measured again by SEM, in mosquitoes more resistant to permethrin compared to less resistant individuals of the same strain (Wood *et al.*, 2010). However sample sizes in this study were small and differences in cuticle only significant to $p<0.05$. In addition both studies employed laboratory mosquitoes so the potential relationship between cuticle thickness and resistance in field populations must be investigated.

The possible role of cuticular proteins in preventing insecticide penetration remains unclear, with factors such as size differences in the insects compared or environmental exposures potentially confounding results. However a recent microarray study identified over-expression of several cuticular proteins in a neonicotinoid resistant strain of aphid *Myzus persicae,* with subsequent *in vivo* penetration assays with radiolabelled insecticide providing evidence of reduced penetration in the resistant strain (Puinean *et al.*, 2010). The cuticular genes identified in this aphid strain may provide a starting point for techniques such as RNAi, linking specific cuticular gene expression to the reduced penetration phenotype. Gene knock-down methodologies could be applied to the cuticular proteins identified in this microarray study to establish whether any change in resistance phenotype results from reducing cuticular protein expression. However this would require insects displaying the bendiocarb resistant phenotype through either development of a stable resistant colony of insects from the field or performing gene knock-down experiments on wild caught insects in the field, which may be logistically challenging. Further support for a cuticular role in the bendiocarb resistance reported here is the previously discussed up-regulation of a dopamine-N-acetyltransferase ortholog with potential links to cuticle sclerotization. The

sclerotization process involves cross-linking of cuticle proteins, with both *CPLC4* and *CPLC17* implicated in multimer formation via protein cross-linking in *An. gambiae* from protein electrophoresis and Mass Spectrometric analysis (He *et al.*, 2007).

3.6.3.3 Glutathione-S-transferase - *GSTS1_1*

Glutathione-S-transferases are a large family of detoxifying enzymes, consisting of six classes in insects, which have been linked to resistance to all major classes of insecticides (reviewed by Enayati *et al.* (2005)). Elevated expression of sigma class GSTs has been reported in pyrethroid resistant strains of *An. gambiae* (Muller *et al.*, 2007; Strode *et al.*, 2006) and *An. arabiensis* (Muller *et al.*, 2008a), with some authors predicting a role in protection against oxidative stress for this class of GST in insects (Singh *et al.*, 2001)(Dowd *et al.* in prep). Elevated expression of *GSTS1_1* in the bendiocarb test group may be linked to direct insecticide metabolism or alternatively have a role in protection against oxidative stresses induced by bendiocarb exposure. Recombinant *GSTS1_1* could be assayed for activity with bendiocarb to establish whether a direct metabolism pathway is feasible. Other insect GSTs have been implicated in direct metabolism of insecticides such as DDT (Ranson *et al.*, 2001; Tang and Tu, 1994) and various organophosphates (Wei *et al.*, 2001; Huang *et al.*, 1998). While *GSTS1_1* mediated detoxification of reactive oxygen species can be assessed using model substrates such as hydroperoxides (Simmons *et al.*, 1989) and 4-hydroxynonenal, a product of lipid peroxidation (Bruns *et al.*, 1999). Vontas *et al.* (2002) employed cumene and t-butyl hydroperoxides to detect peroxidase activity in recombinant GST (*nlgst1-1*) from the rice brown plant hopper *Nilaparata lugens*. This GST was implicated in pyrethroid resistance via detoxification of pyrethroid induced lipid peroxidation products (Vontas *et al.*, 2001).

3.6.3.4 Cytochrome P450s – *CYP325C1-3*

Three *CYP325C* genes were over-expressed in the test group with microarray probes for these genes both over 2 fold expressed and within the top 1% of ranked p-values. However the high sequence similarity of these P450s (see Chapter 4) means there is potential for cross probe hybridisation, which is a potential pitfall of microarray profiling of closely related genes. Cytochrome P450 monooxygenases are associated with resistance to a number of insecticidal compounds including pyrethroids (Muller *et al.*, 2008b; Yang *et al.*, 2006; Hardstone *et al.*, 2007), organophosphates (Sabourault *et al.*, 2001), neonicotinoids (Puinean *et al.*) and organochlorine DDT (Chiu *et al.*, 2008; Daborn *et al.*, 2002). To date P450s have not been definitively linked to carbamate resistance, however, these enzymes display broad substrate specificity and catalytic diversity, with over 100 P450s gene reported in *An. gambiae s.s.* (Ranson *et al.*, 2002).

Cytochrome P450s metabolise toxins via an oxidative catalytic cycle which requires electron donation from redox partners such as NADPH-cytochrome P450 reductase and cytochrome b5 (Feyereisen, 1999). Dual recombinant protein systems expressing

P450s along with the required redox partners have been successful in validating P450 involvement in insecticide metabolism (Muller *et al.*, 2008b; Chiu *et al.*, 2008).

A similar approach could be taken with this family of enzymes to establish whether any of the three proteins are capable of bendiocarb metabolism.

3.6.3.5 Carboxylesterase – *COEJHE2E*

Esterases are commonly associated with organophosphate and carbamate resistance in strains of aphid species *Myzus persicae*, plant hopper *Nilaparvata lugens* and mosquitoes of the *Culex* genus via gene amplification, elevated expression and/or allelic variation (reviewed by Hemingway (2000) & Oakeshott *et al.* (2005)). Amplification of carboxylesterase gene *E4* or paralogue gene *FE4* in aphid species *Myzus persicae* has been linked to organophosphate, pyrethroid and carbamate resistance (Field *et al.*, 1988; Field and Devonshire, 1998). Devonshire and Moore (1982) demonstrated that esterase *E4* extracted from *M. persicae* homogenate was capable of hydrolysing organophosphates, pyrethroids and carbamates using an assay based on insecticidal inhibition of the esterase and subsequent re-activation when the enzyme successfully hydrolysed the compound. The slower re-activation of *E4* when incubated with carbamate insecticides suggested the primary mode of esterase mediated carbamate resistance would be via sequestration of the insecticide followed by slow metabolic turn over. In *Anopheles* species esterases have been indirectly linked to pyrethroid resistance via elevated expression (Vulule *et al.*, 1999; Matowo *et al.*, 2010).

Elevated expression of esterase *COEJHE2E* in the bendiocarb test group may have resulted from alterations in regulatory regions controlling transcription. Over expression of a number of esterase genes have been linked to resistance in *Culex pipiens* (Raymond *et al.*, 1998). While organophosphate resistance in *Culex quinquefasciatus* is linked to amplification and over-expression of two esterases (*Estα21, Estβ21*) (Hemingway *et al.*, 1998). Analysis of the intergenic spacer between these two head-to-head orientated genes identified potential regulatory elements involved in elevated expression (Hawkes and hemingway, 2002).

An alternative hypothesis for elevated *COEJHE2E* transcript levels is gene duplication/amplification which has been reported in a number of insect species including *M. persicae*, *Cx pipiens*, *Cx tritaeniorhynchus* and *Nilaparvata lugens* (Field and Devonshire, 1998; Raymond *et al.*, 1998; Karunaratne *et al.*, 1998; Vontas *et al.*, 2000). Amplification of the *COEJHE2E* gene within the genome of bendiocarb resistant field samples could be investigated through various molecular methods. For instance, Southern blotting which involves enzymatic digestion of gDNA followed by gel separation and probing with specific gene fragments; this technique was employed by Karunaratne *et al.* when investigating esterase amplification in *Cx tritaeniorhynchus* (1998). Alternatively esterase copy number ratio can be calculated using Real-Time quantitative PCR employing a gene of known copy number as a standard (Vontas *et*

al., 2000), or duplication directly identified by sequencing of BAC clones (Wondji *et al.*, 2009).

The potential involvement of *COEJHE2E* in bendiocarb resistance may be through direct hydrolysis of the insecticide or via sequestration, the latter being the primary mechanism hypothesised for *M. persicae* esterase *E4* and carbamate compounds (Devonshire and Moores, 1982).

3.6.4 Effect of biological variation on the consistency of gene expression

Overall a number of interesting candidates were identified via microarray comparison of bendiocarb selected and unselected female *An. gambiae s.s.* from the Urban Accra region of Ghana. However it is evident that variation in the dataset affected levels of fold change and significance.

A test versus control design is inherently "noisy" as individuals selected for a degree of resistance are compared to a set of unselected controls which are likely to contain a mixture of resistance phenotypes. This experimental design can also constrain the maximum fold changes calculated via microarray as addressed by Muller *et al.* (2008b). However the use of material which did not survive exposure, i.e. susceptible females, was not an option due to RNA degradation *post mortem*. Biological and phenotypic variation in combination with potentially small sample sizes are the probable reasons behind the poor p-values generated for microarray expression ratios (Morey *et al.*, 2006).

However, field-collected and phenotyped material is obviously important when trying to identify genes associated with operationally relevant insecticide resistance. In addition a number of authors have questioned the ability of lab colonies to act as proxy for wild populations of insects (Macdonald and Long, 2004; Riehle *et al.*, 2006). Biological variation is unavoidable when studying wild caught material; although variation can be reduced by increased replication, matching of samples for known genetic variables and standardising experimental conditions. Indeed samples were matched according to molecular form (S) as well as the presence of *ace*-1 mutation *G119S* in an attempt to reduce variability. Field collected mosquitoes were also exposed to standardized rearing conditions within the field insectaries until 3-5 days old to reduce potential expression variation associated with environmental differences. Unfortunately, material collected with the chosen genetic background (S form *G119S* heterozygotes) of adequate quality on extraction of RNA was limited, so increasing sample size was not an option. An alternative strategy would have been to add a reference pool to the interwoven loop design, for instance a fully susceptible strain from the same geographical area, which would hopefully aid identification of genes consistently over-expressed in the 'resistant' group.

3.6.5 Intragenic, inter-probe variation in fold-change

Overall duplicate probes provided consistent estimates of differential expression illustrated by the close proximity of duplicates on the volcano plot (Figure 3.13) and

consistent fold changes and p-values between duplicates in result tables (Table 3.5 and Table 3.6). This suggests that hybridization was uniform across the array. However there was inconsistency in expression ratios reported for unique probes designed for the same gene. For instance in the *CYP325C1-3* group all three gene probe sets displayed fluctuating fold changes. For variant C3 fold-change varied between 1.07-1.95 depending on probe design, while for C1, 1.60-2.18 fold change was recorded and for C2 expression ratios varied between 1.69-2.27. Probes designed for gene *GSTS1_1* also produced varying expression ratios (2.1-5.1 fold). These differences may be a result of poor probe quality and/or allelic variation. The probe design algorithms employed in eArray use various parameters including T_m and GC content to select the 'best' probe for a gene which is indicated by a 'BC' score (for design parameters consult Appendix 10). For detox genes, such as *GSTS1_1* and *CYP325C1-3*, three unique probes were designed per gene and it likely that there is variation in probe sensitivity and specificity which may affect hybridisation.

Allelic variation may have caused inconsistencies in probe results. It is estimated that the *An. gambiae s.s.* genome has a SNP density of $1.6X10^{-3}$ (approximately 1 SNP every 620 bp) (Holt *et al.*, 2002) although it is noted that the majority of these polymorphisms are intronic. While Wilding *et al.* (2009) reported one SNP every 34bp in a SNP discovery study focusing on gene families previously linked to resistance (insecticide targets and detox genes). The discrepancy between SNP frequencies recorded by Holt and Wilding may be explained by high levels of polymorphism in 'detox' genes which supports substrate plasticity. In addition, the sequencing in Holts study employed a highly in-bred strain (PEST) compared to geographically diverse samples in Wilding *et al.*'s investigation. Genes highlighted as producing inconsistencies in probe results were members of the 'detox' gene group representing P450s and GSTs (*CYP325C1-3* and *GSTS1_1*) so could potentially harbour higher SNP densities compared to non-detox genes. Indeed a number of SNPs were found in genes *GSTS1_1*, *CYP325C1-3* and candidate *OBP3* compared with Ensembl sequences (Chapter 4). *CYP325C* genes sequenced from field derived cDNA showed divergence from Ensembl published sequences (Chapter 4), with 2-3 differential SNPs found to be coincident with microarray probe positions. However the *GSTS1_1* and *OBP3* SNPs were not within the 60mer microarray probes, which may suggest that probe quality according Agilent scoring is the potential cause for discrepancy.

It is likely that sequence differences between probe and target resulted in some loss of signal, which is an unavoidable problem when using microarrays in highly variable field populations. Recently developed RNA sequencing technologies may over-come inherent issues of allelic variation in microarray analysis as transcripts are sequenced rather than detected via sequence specific probes. Next generation sequencing also has the potential to resolve expression profiles of closely related genes, as microarray probes may bind to conserved regions within gene families such as the *CYP325C1-3* genes which share ~ 96% sequence identity (Chapter 4) or to shared exons of alternatively spliced genes such as *GSTS1_1* and *S1_2* (Ding *et al.*, 2003). Although

probe designs were checked against the *An. gambiae s.s.* genome to try and ensure gene specificity.

In addition, NGS may provide greater sensitivity for the detection of differentially expressed genes especially those with low overall expression levels. Sultan *et al.* (2008) reporting 25% more genes detected using RNA-sequencing compared to microarray on the same sample set; although correlation between results from both platforms was high (Pearson coefficient = 0.88). On the other hand Feng *et al.* (2010) found that Agilent 4x44K microarrays were able to identify more transcripts (17, 362) compared with Illumina based NGS Digital Gene Expression tag profiling (DGE) (15,938) with a 13,221 transcript overlap. An absence of restriction site *NlaIII* (used to create short cDNA sequencing fragments) in nearly 2000 transcripts meant these were not detected by DGE; however the potential for microarray false positives was also noted. The authors found the technical repeatability of microarray data was marginally higher than DGE though the increased sensitivity of NGS to low copy number transcripts was a major advantage. It is important to consider the advantages and disadvantages of both technologies before planning an investigation of gene expression (as reviewed by Roh (2010)). For instance NGS provides an open system requiring no previous knowledge of the background genome and the depth of coverage and SNP information provision is a notable advantage over microarray technology. However NGS is also a much lower through-put method compared with microarray and is also currently much more costly. Moreover, the analysis of NGS at present is highly complex, requiring sequence read annotation and mapping before expression differences can be calculated. Microarrays are therefore recommended for more routine studies of gene expression differences across many samples (Roh *et al.*, 2010), as in this study.

3.6.6 Potential for gene induction

A number of studies have attempted to assess the effects of insecticide exposed on the induction of gene expression in insects (Willoughby *et al.*, 2006) (Vontas *et al.*, 2005; David *et al.*, 2010). Immediate up-regulation of genes involved in detoxification (Poupardin *et al.*, 2008; David *et al.*, 2010) and cuticle development (David *et al.*, 2010) have been recorded in *Ae. aegypti* post sub-lethal insecticide exposure. While gene induction at a series of times points post exposure was explored using microarrays in *An. gambiae s.s.* (Vontas *et al.*, 2000). Permethrin exposure induced expression of 75 EST from a 20,000 EST array, including genes implicated in detoxification, redox and stress response. This induction of expression peaked at ~6 hours post insecticide exposure with minimal induction reported at 2 and 24 hours after exposure. A similar 6 hour induction peak and a return to baseline expression levels 24hours post exposure, was recorded in a separate study in which wild caught *An. gambiae s.s.* from Ghana were exposed to sub-lethal doses of pyrethroids and compared to non-exposed insects using a candidate gene array (The Detox Chip, (David *et al.*, 2005)) (E. Warr, unpublished data.).

Mosquitoes were therefore allowed to recover for 24 hours after insecticide exposure in this study in an attempt to rule out gene induction, as at this point, base-line gene expression should have been resumed (Vontas *et al.*, 2000)(E. Warr, unpublished data). The importance of gene induction relative to constitutive expression in insecticide resistance remains unclear. Furthermore, there is a particular lack of research into induction of genes at extended time periods post insecticide exposure (>24 hours). The recovery period could have been extended in this study to try and rule out gene induction in exposed animals influencing differential expression on comparison with non-exposed controls. However, as there is currently no published data on induction at times periods greater than 24 hours, there is no 'ideal' recovery time in which induction ceases. Results from previous studies therefore informed the decision on recovery time.

To eliminate the potential influence of gene induction post insecticide exposure, alternative experimental designs could have been employed. For example, male and female mosquitoes surviving insecticide exposure (test group) could have been mated to produce an F1 generation in which both parents were insecticide 'resistant'. These non-exposed F1 mosquitoes could have then been compared to non-exposed controls taken from the sampled population and differential expression assessed.

3.6.7 Conclusion

This experiment was the first whole genome array performed on field collected *An. gambiae* and was intended to be an exploratory study of expression associated with bendiocarb resistance in highly diverse field populations. The experimental design comparing tested mosquitoes to un-exposed controls in addition to the biological complexity of wild caught samples meant the results set was potentially noisy and rigorous p-value correcting unfeasible. However a number of potentially interesting genes were identified which will be validated via alternative methods of expression analysis as well as recombinant protein experiments. In addition, issues highlighted in this study have enabled informed decision making and experimental design for a subsequent microarray experiment (Chapter 5).

Chapter 4: Validation of candidate genes from a microarray experiment investigating bendiocarb resistance in *An. gambiae s.s.* from Ghana, West Africa

4.1 Abstract

Whole genome microarray profiling of bendiocarb selected *An. gambiae s.s.* from Ghana, West Africa identified a number of differentially expressed genes when compared to unexposed controls from the area. In an attempt to validate microarray candidates, a selection of genes including enzymes from families previously linked to insecticide metabolism as well as a number of novel protein candidates, were pursued using a combination of multiplex quantitative PCR, for confirmation of expression ratios, and recombinant protein assay, to identify potential interactions with insecticide bendiocarb. Quantitative PCR provided good correlation with microarray derived fold change for a number of genes in the test and control pools employed in the array experiment. Recombinant protein assay of candidate *GSTS1_1* suggested that this glutathione-S-transferase was not interacting directly with bendiocarb. However, fluorescent displacement assays with protein *OBP3* (odorant binding protein) implicated a potential protein-insecticide interaction.

4.2 Introduction

Microarray experiments, especially those on a whole genome scale, inevitably produce large numbers of candidate genes. Whilst statistical techniques can act as a primary filter, the subsequent validation of these genes is the bottleneck in the process.

An initial first step, undertaken in the majority of microarray experiments, is to validate the expression levels recorded via microarray using an alternative quantitative approach. Real-Time quantitative PCR, involving the detection of accumulated gene specific PCR products through incorporation of a fluorescent dye, is frequently employed to validate a selection of microarray candidates (Morey *et al.*, 2006; Abercrombie *et al.*, 2008; Dong *et al.*, 2006; Erickson *et al.*, 2009; Beckham *et al.*, 2010). Alternatively a multiplex quantitative PCR approach, such as GeXP (Beckman-Coulter), may be taken to establish expression ratios for a number of genes concomitantly (Felix *et al.*, 2010; Muller *et al.*, 2008b; Ossovskaya *et al.*, 2010). Although, it must also be noted that a number of factors can affect correlations between microarray and quantitative PCR determined expression ratios, including differences in chemistry (Freedman *et al.*, 1999) and normalisation procedures (Morey *et al.*, 2006).

Once confident in expression results, candidate genes should be validated biologically to assess influence over the phenotypic trait of interest. A common *in vivo* approach for candidate validation is RNA interference (RNAi), which involves the introduction of gene specific double stranded RNA and subsequent breakdown of corresponding

mRNA through the RNA Induced Silencing Complex (RISC) (Hannon, 2002). This 'loss of function' technique has been employed to validate genes identified in the immune response to *Plasmodium* infection in *An. gambiae*, highlighting seven genes which influenced resistance to parasite infection (Dong *et al.*, 2006). RNAi also successfully identified a microarray candidate associated with colon cancer which severely reduced tumour growth in a cancer cell-line (Williams *et al.*, 2003).

Gene knock-down studies in insects may require stable insect colonies displaying a phenotype of interest. With regards to insecticide resistance in wild mosquito populations, this may involve colonising field material, which is laborious and problematic. Alternatively, RNAi may be performed in the field, though this requires suitable laboratory facilities and availability of regents and equipment.

A better approach to biological validation may be 'gain of function' assays, in which over-expression of a candidate gene *in vivo* produces a greater phenotype. For instance, GAL4/UAS (Brand and Perrimon, 1993) driven transgenic over-expression of gene *CYP6G1*, a microarray candidate associated with insecticide resistance in *Drosophila* (Le Goff *et al.*, 2003), was necessary and sufficient to confer DDT resistance in a susceptible strain (Daborn *et al.*, 2002). Gain of function studies in insects may also circumvent colonisation of insects with a phenotype of interest, requiring only an insect strain of similar genetic background in which to express genes; however, transgenic techniques involved in gene over-expression are potentially more complex compared with RNAi.

Recombinant expression of microarray candidates is an *in vitro* proteomic approach which can be used to explore a genes involvement in a phenotype. A number of metabolic enzymes identified as candidates in insecticide resistance studies have been expressed using recombinant systems and assayed with insecticidal compounds (Muller *et al.*, 2008b; Stevenson *et al.*, 2011; Chiu *et al.*, 2008). Direct *in vitro* metabolism of pyrethroid insecticides via P450s *CYP6P3* (Muller *et al.*, 2008b) and *CYP6M2* (Stevenson *et al.*, 2011), was detected using HPLC analysis, implicating these enzymes in *An. gambiae s.s.* pyrethroid resistance. While another *An. gambiae s.s.* P450, *CYP6Z1*, indentified in a microarray study (David *et al.*, 2005) , was subsequently found to metabolise DDT when expressed in an *E. coli* system (Chiu *et al.*, 2008). *In vitro* techniques are beneficial in the absence of live insect material. Moreover, protein assays and potential crystallisation of recombinant proteins provide greater detail of the reaction kinetics and allow the influence of allelic variation to be rapidly assessed.

4.3 Methods

4.3.1 GeXP multiplex quantitative PCR

To confirm expression ratios obtained from the microarray comparison of bendiocarb resistant and control mosquitoes, a multiplex quantitative PCR approach using the GeXP system (Beckman-Coulter) was employed. Twenty one genes which displayed a range of expression ratios were selected for analysis.

4.3.1.1 GeXP multiplex design and validation

The genes selected for validation from initial bendiocarb microarray results are listed in Table 4.1. These genes were selected to provide a range of fold changes and included 8 genes which had probes >0.95 $\log_2$ over-expressed in the bendiocarb resistant group. In addition a splice variant (*GSTS1_2*) which is closely related to putative candidate gene *GSTS1_1* was included to investigate whether differential expression may have resulted from non-specific hybridisation. *GSTS1_1* shares 75% identity with gene *GSTS1_2*, therefore primers was designed to divergent regions of these splice variants in an attempt to rule out cross hybridisation.

The Genomelab GeXP eXpress Profiler software was used for primer design based on cDNA sequences obtained from Ensembl (http://metazoa.ensembl.org/index.html). Multiplex primer pairs were designed for all 21 genes along with a standard house-keeping gene (ribosomal S7) and a spike in reference gene (KAN^r) based on Primer3 software algorithms (Rozen and Skaletsky, 2000) according to the GenomeLab™ User's Guidelines (A29142-AB, 2007).

All primer designs contained a universal tag in addition to gene specific regions and were designed to produce fragments between 100-400bp in length with matching primer melting temperatures (T_m ~60°C). The assay relies on each gene producing a PCR fragment of unique size which is fluorescently labelled via universal sequences and separated by capillary electrophoresis (Figure 4.1). The GeXP assay is able to detect up to 30 genes within one sample; however expression of specific genes within a sample cannot be accurately measured due to differing PCR efficiencies. Therefore a comparative sample is employed to provide expression ratios for each gene.

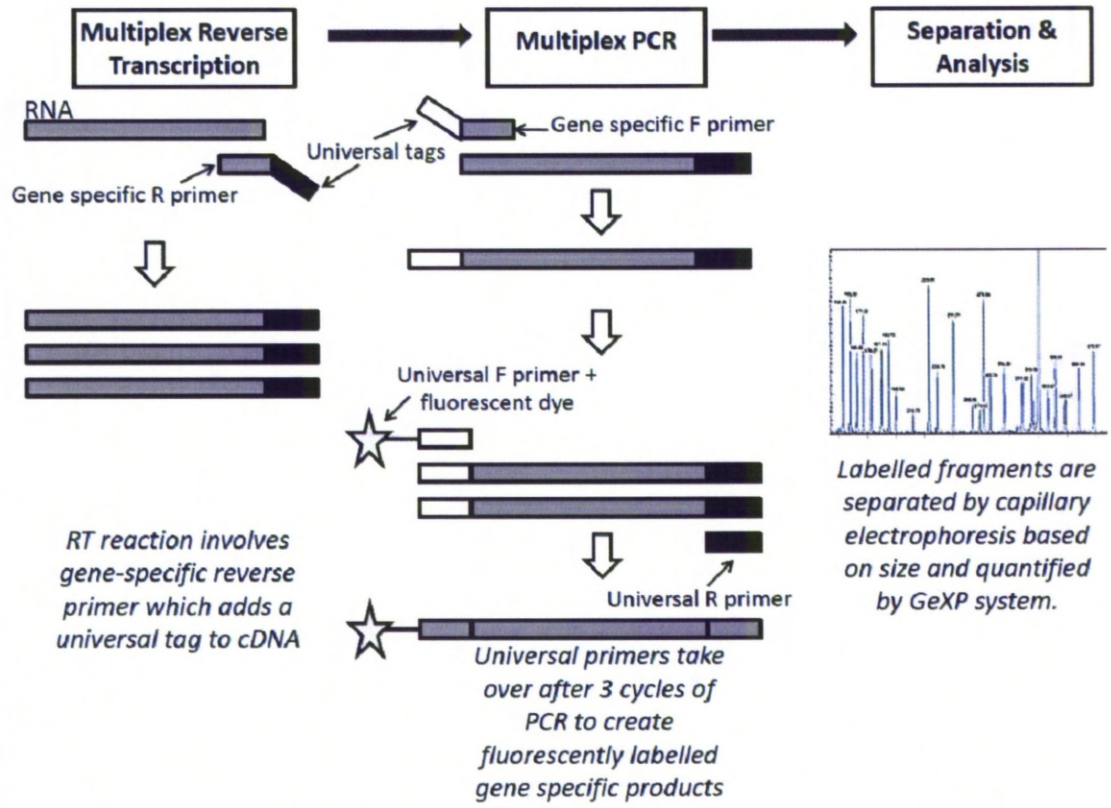

Figure 4.1. Schematic of gene expression analysis using the Beckman Coulter GeXP system.

Full primer sequences were exported for BLAST (Altschul *et al.*, 1990) analysis against the *An. gambiae s.s.* genome (Ensembl P3) before synthesis. If cross hybridisation was predicted, gene primers were removed from the multiplex and new primers designed to a more specific region of the gene where possible. This was necessary for the *CYP325C1-3* genes as there was high similarirty between the three cDNA sequences deposited in VectorBase (http://agambiae.vectorbase.org/). It was not possible to design locus specific primers for *CYP325C2* due to high identity with both *CYP325C1* (96%) and *CYP325C3* (94.5%) (Figure 4.2). The final set of primers (Appendix 20) were tested individually by employing each forward primer singularly with a multiplex of all 23 reverse primers (including S7 and KAN^r) following the GeXP Chemistry Protocol (A29143AC 2009) and checking for the predicted fragment size.

4.3.1.2 GeXP sample preparation and analysis

The ten total RNA pools used in the microarray experiment (5 test, 5 control) were employed for GeXP analysis. RNA was diluted to a concentration of 20ng/µl using nuclease-free water (Sigma) with 5µl used per reaction. Reverse transcription (RT) and PCR amplification were performed using the GenomeLab™ GeXP Start Kit (Beckman-Coulter) according to manufacturer's guidelines with all incubations carried out in a Bio-Rad/MJ Dyad thermal cycler. The RT reaction contained 100ng of each

RNA pool in a 20µl reaction with 1X RT reaction buffer, 1unit/µl reverse-transcriptase, 5µl of KAN^r RNA control and 50nM reverse primer plex. Template-free and reverse transcriptase free controls were also performed. Incubation steps were 48°C for 1 minute, 37°C for five minutes, 42°C for one hour and 95°C for five minutes. PCRs were performed post RT in a total volume of 20µl with 9.3µl of each RT reaction as template. Each PCR contained 1X PCR buffer, 5mM $MgCl_2$ (ABgene), 0.7µl of Thermo-Start DNA Polymerase (ABgene) and 20nM forward primer plex. PCR cycling conditions were 95°C for 10 minutes followed by 35 cycles of 94°C for 30 seconds, 55°C for 30 seconds, 70°C for 1 minute and a final hold at 4°C.

Post PCR, 1µl of undiluted product was added to 38.5µl of sample loading buffer and 0.5µl of DNA size standard (GeXP Start Kit) and analysed on the GenomeLab™ GeXP Genetic Analysis System following the User's guide. Once separated, fragments were examined in the Fragment Analysis module of the GeXP software and exclusion parameters set to filter out unwanted peaks. Subsequently, analysis was performed using the eXpress Analysis software after peak binning to adjust for fragments 1-2 base pairs outside of their expected size range. Three technical replicates for each of the 10 RNA pools were performed. Gene expression was normalised against the house-keeping gene ribosomal S7, and mean normalised expression and fold difference between test and control were calculated for each gene.

4.3.2 Recombinant protein validation

4.3.2.1 *GSTS1_1* – bendiocarb inhibition assays

Both *GSTS1_1* and splice variant *GSTS1_2* have previously been cloned and expressed at LSTM, from the insecticide susceptible Kisumu strain of *An. gambiae s.s.* by A. Dowd *et al.*

Two protein aliquots were provided courtesy of A. Dowd. *GSTS1_1* aliquot 017 (batch 99) contained protein at 0.072mg/ml and GS*TS1_2* aliquot 018 (batch 100) has a concentration of1.4mg/ml.

Bendiocarb inhibition assay

This assay was based on a protocol devised by Dowd *et al.* (in prep).

Each *GSTS1* splice variant was pre-incubated with insecticide bendiocarb before measuring the activity with known GST substrate 1-choro-2,4-dinitrobenzene (CDNB) (Habig *et al.*, 1974). CDNB is commonly used to measure GST enzyme activity via a conjugation reaction:

$$\text{GSH + CDNB} \xrightarrow{\textit{GST}} \text{GS-DNB Conjugate + HCl} \qquad (3)$$

The product of this reaction, GS-DNB, absorbs at 340nm which can be read on a spectrophotometer. The rate of conjugate production is directly proportional to GST activity in the reaction.

The activity of *GSTS1* splice variants pre-incubated with bendiocarb was compared with protein not exposed to the insecticide to ascertain whether the compound was inhibiting enzyme activity. If inhibition of CDNB activity occurs it is inferred that the insecticide is interacting with the recombinant protein.

Ten micro-grams of each protein were incubated in triplicate with 200µM bendiocarb for five minutes on a shaking incubator (700RPM) at 25°C. Reactions were performed in 0.1M potassium phosphate buffer (pH 6.5) with 2.5mM glutathione (GSH) and 10% methanol (Sigma), to increase bendiocarb solubility, at a total volume of 1ml. Post incubation, CDNB was introduced to a final concentration of 1mM and briefly mixed before being immediately read at 340nm on a spectrophotometer (Cary 4000) at 25°C for 1 minute. The same reactions were performed with each of the two enzyme variants minus bendiocarb for un-exposed comparative readings. The changes in absorbance over one minute were plotted, minus a blank containing no enzyme, to produce a gradient reading for each reaction. The slope for each protein was then used to calculate the CDNB specific activity (µmol/mg/min) for each enzyme using the following equation:

((gradient – blank) * (1/9.6[a]) * (assay vol. / enzyme vol.)) / (conc. Enzyme mg/ml) (4)

Where [a] is the CDNB molar extinction coefficient; a measure of CDNB light absorbance at 340nm.

The mean specific activity for both *GSTS1_1* and *GSTS1_2* ± bendiocarb was calculated along with 95% Confidence Intervals, and mean specific activities used to calculate percentage inhibition of GST activity.

4.3.2.2 *OBP3* bendiocarb interaction

A fluorescent displacement assay employing recombinant *OBP3* from *An. gambiae s.s.* to assess potential interaction with bendiocarb, was performed by collaborator J-J. Zhou at the Biological Chemistry Division, Rothamsted Research, Harpenden. This research group are currently working on a number of odorant binding proteins in vectors of disease as well as crop pests to investigate olfactory responses during host-seeking and mating (Zhou *et al.*, 2010). For a protocol outline see Appendix 18.

4.3.2.3 *CYP325C1-3* recombinant protein expression

Cyp325C gene sequences

All three *CYP325C* genes (C1-3) had representative probes which were over-expressed in the bendiocarb test group compared to the non-exposed controls. Overall, signals from *CYP325C2* probes were the most consistent with a mean expression ratio of 1.97 [0.2275 SD.]. Database sequences for all three genes were aligned and pair-wise DNA distance calculated using MEGA 4.0.2 (http://www.megasoftware.net/) which revealed high sequence identity (94.6-96%, Figure 4.2). Consequently there was a strong likelihood that probes for one gene may have cross-hybridised to another.

Sequences were retrieved from two databases as the automatic annotation of gene *CYP325C3* in Ensembl was questionable; the deposited cDNA sequence contained no defined 3' or 5' untranslated regions (UTRs) or an ATG start codon. Moreover the cDNA sequence also contained suspected indels which were not present in the manually curated *CYP325C3* sequence deposited in the insect P450 database (http://p450.sophia.inra.fr/index_link.html). Therefore the *CYP325C3* sequence from the P450 website was employed for pair-wise comparison and primer design.

Although similar in sequence, these genes do not share chromosomal location with *CYP325C3* reportedly situated on chromosome 3R while C1 and C2 are in close proximity to each other (~8.5Kb apart) on 2R. This may suggest a gene duplication and translocation event creating *CYP325C3* on chromosome 3R. Such an event may also explain the high levels of sequence similarity between genes.

In a separate study, a SNP association mapping approach identified a large cluster of resistance associated SNPs on chromosome 3R in permethrin resistant *An. gambiae s.s.* from Ghana (D. Neafsey, Broad institute). Within the 52 genes listed for this region was *CYP325C3*. This independent implication of a *CYP325C* gene in insecticide resistance further supported the decision to pursue protein expression and insecticide interaction studies with this P450 gene family.

CYP325C1-3 cloning and sequence analysis

Primers were designed to conserved 3' and 5' regions of the *CYP325C1-3* gene sequences (CYP325C-generic Fb: ATGTGGTGGTTTTGGTTGG, CYP325-generic-R: ACTGCACCTACATCCTCTTCTC). These primers were likely to capture all three *CYP325C* P450s and were successful in amplifying full length product (~1.5Kb) from a generic cDNA pool constructed from non-selected female mosquitoes from the field site sampled for microarray investigation. This fragment was subsequently cloned into the pJET 1.2/blunt holding vector (Fermentas) and plasmids prepared via mini-prep. Briefly 1µl of PCR product was ligated with 50ng of pJET 1.2/blunt cloning vector and 5 units of T4 DNA ligase in 1x Reaction Buffer (Fermentas) for 30 minutes at room temperature. Five micro-litres of each ligation reaction were transformed into 50µl of chemically competent *E.coli* cells (JM109 AllianceBio) via heat-shock. Cells were added to the ligation mixture in round bottomed 14ml tubes on ice and after incubation for 20 minutes, tubes were place in a water bath at 42°C for 45 seconds before placing back on ice for 2 minutes.

Cells were recovered in 950µl of SOC medium (0.5% yeast extract, 2% tryptone, 10mM NaCl, 2.5mM KCl, 10mM $MgCL_2$, 10mM $MgSO_4$, 20mM Glucose) in a 37°C shaking incubator (150 RPM) for one hour before plating 50 and 100µl on LB-amp-agar plates (1% tryptone, 0.5% yeast extract, 200mM NaCl, 1.5% agar, 100µg/ml ampicillin). Plate colonies were screened after overnight incubation at 37°C by mixing a single colony with 5µl of Luria-Bertani (LB) medium (1% tryptone, 0.5% yeast extract, 200mM NaCl) and using 1µl in a PCR reaction with external plasmid primers (Fermentas). The remaining 4 µl from positive clones was used to inoculate 5ml of LB containing 100 µg/ml ampicillin prior to overnight culture in a 37°C shaking incubator at 150RPM.

Cells were pelleted after taking 100µl for glycerol stocks (50:50, cells to 80% glycerol) and plasmid DNA extracted using the QIAprep Spin Miniprep Kit (Qiagen) according to manufacturer's instructions eluting in 50µl of elution buffer (10mM Tris-Cl, pH 8.5).

On sequencing, approximately three *CYP325C* gene variants were revealed (Figure 4.3); although sequence identity between clones was high (97.9-100%, Figure 4.2). Pair-wise percentage identity was calculated for these sequences along with database deposited *CYP325C* sequences using MEGA 4.0.2 which revealed that sequenced clones more closely resembled the *CYP325C2* database sequence (98.4-99.4% ID, Figure 4.2). This was supported by the construction of a DNA distance tree using cloned genes and Ensembl/P450 database sequences; a neighbour-joining algorithm was employed based on pair-wise DNA distance using all codon positions with 2000 repetitions for bootstrap calculations (Figure 4.3).

	Database sequences			**Sequenced clones**				
	C3	C2	C1	A4	A3	A2-3	A2-2	A2-1	A2
C3									
C2	94.6								
C1	95.1	96.0							
A4	95.0	98.4	95.9						
A3	94.3	98.9	95.9	97.9					
A2-3	95.2	98.5	96.0	100.0	98.1				
A2-2	94.4	98.8	95.6	98.6	98.3	98.7			
A2-1	94.7	99.1	95.9	98.7	98.5	98.9	99.7		
A2	95.0	98.4	96.0	99.6	97.9	99.7	98.5	98.7	

Percentage identity

Figure 4.2. Pair-wise DNA distance of three database deposited *An. gambiae s.s. CYP325C* genes and sequences obtained from field collected *An. gambiae s.s.* material from Ghana, West Africa.

CYP325C1 (C1) and *CYP325C2* (C2) sequences were obtained from Ensembl (P3) (http://metazoa.ensembl.org/index.html) while *CYP325C3* (C3) was obtained from the Insect P450 database (http://p450.sophia.inra.fr). A2-A4 represent clones obtained through amplification from Ghanaian cDNA using primers designed to conserved 5' and 3' regions in all three database deposited *CYP325C* genes. Pair-wise percentage identity was calculated using MEGA 4.0.2 (http://www.megasoftware.net/).

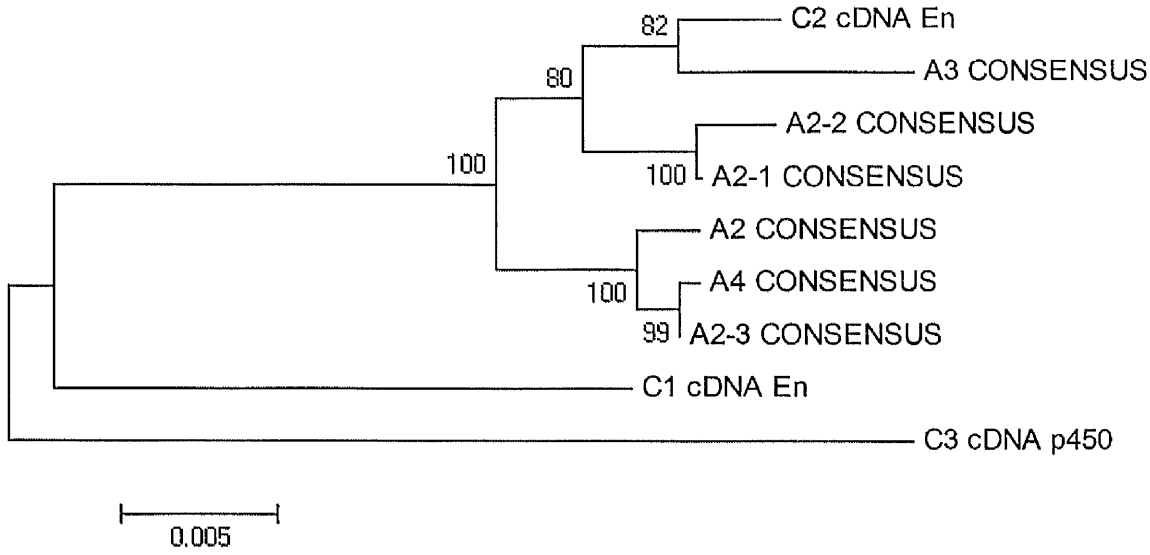

Figure 4.3. DNA distance tree for *An. gambiae s.s CYP325C* genes and database deposited sequences.

This neighbour-joining dendrogram was produced using the MEGA 4.0.2 (http://www.megasoftware.net/) software from pair-wise DNA distance using all codon positions and 2000 repetitions for bootstrap values. *CYP325C1* (C1) and C2 sequences were taken from Ensembl (P3) (http://metazoa.ensembl.org/index.html) while *CYP325C3* (C3) was taken from the insect P450 database (http://p450.sophia.inra.fr/). A2, A2_1, A2_2, A2_3, A3 and A4 are all cloned and sequenced genes amplified from a generic Ghanaian cDNA pool using primers designed to amplify all three genes.

P450 expression system (ompA+2 leader sequence)

Using the sequenced *CYP325C* genes as template, recombinant protein expression was pursued following protocols designed for the expression of human cytochrome P450 monooxygenases in *E.coli* (Pritchard *et al.*, 2005). The system involves modification of the N-terminal coding region of the P450 of interest through fusion with an endogenous *E.coli* signal peptide (*ompA* leader) to allow functional expression in the bacterial system. The 21 amino acid leader sequence along with two spacer amino acids (alanine-proline), essential for subsequent *ompA* cleavage, were introduced to the P450 via fusion PCR (Figure 4.4). Two restriction enzyme recognition sites are introduced to the full length *ompA*-P450 fragment, *NdeI* and *XbaI* in Figure 4.4, to enable ligation into the expression vector which also contains an ampicillin resistance gene. *NdeI* is employed as the 5' restriction site in the forward primer as it incorporates the initiation codon (ATG) and is the most 5' site in the expression vector (pCW). The 3' site can be selected from a number of recognition sequences in the vectors multiple cloning site. This system is advantageous in that it does not require changes to the coding region of the P450.

Microsomal cytochrome P450 enzymes also require co-expression with redox partner cytochrome P450 reductase (CPR) and/or the presence of cytochrome b_5 to function

(Feyereisen, 1999). This involves co-transformation of *E.coli* with the P450 encoding plasmid and a plasmid containing a redox partner. This second plasmid also confers resistance to a different antibiotic *i.e.* chloramphenicol to enable screening for both transformations. Once co-transformed and induced, cells are lysed and spheroplasts, containing the P450 complex, are pelleted.

To confirm P450 activity analytical procedures are performed including spectral detection of a peak at 450nm, hence P'450', while CPR activity is confirmed via cytochrome c reductase activity.

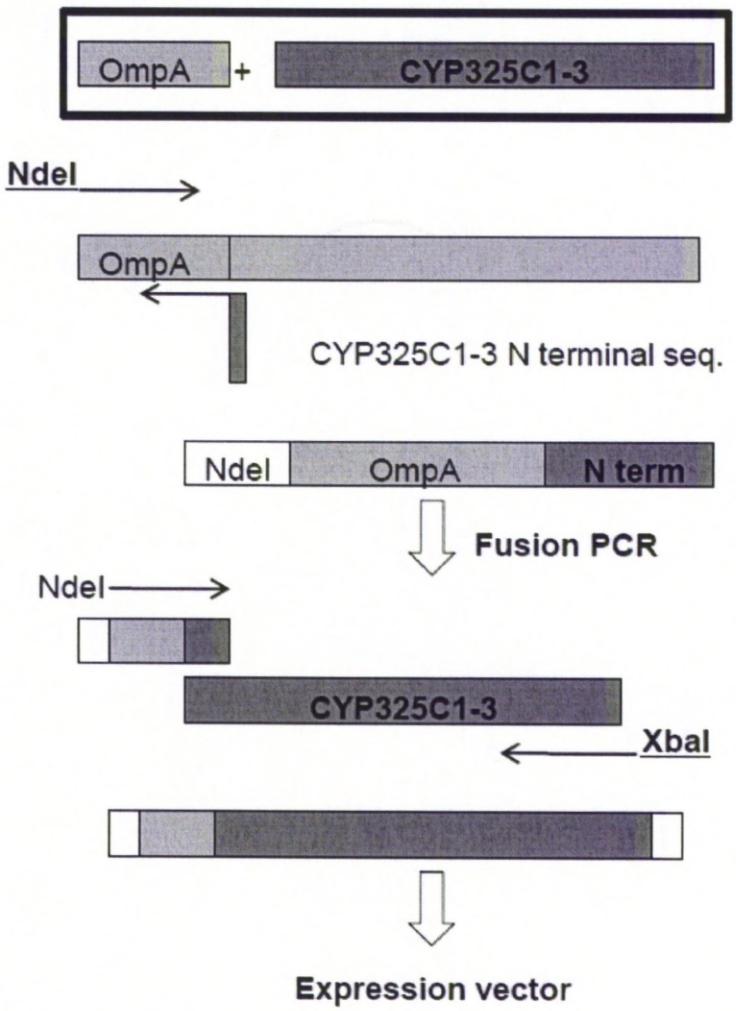

Figure 4.4. *OmpA* leader expression system.

OmpA specific forward and reverse primers are used to amplify the *ompA* fragment from a template containing the full *ompA* sequence. The forward primer incorporates an *NdeI* restriction enzyme site while the reverse primer contains a section of sequence from the N-terminus of the P450 to be expressed (*CYP325C1-3*). Fusion PCR is performed using this small *ompA* fragment along with the full length sequence of the P450 to be expressed. The forward *ompA-NdeI* primer is employed along with a P450 specific reverse primer designed to the C-terminus and incorporating an *XbaI* restriction site. The two restriction sites (*NdeI* and *XbaI*) are then used to ligate the *ompA*-P450 construct into a pre-cut expression vector.

Fusion PCR primer design

An *ompA* specific forward primer incorporating an *NdeI* restriction site, plus 6 'junk' bases to aid restriction enzyme docking, was designed to amplify a 99 bp linker fragment with two separate linker primers (PCR 1). The first linker primer (Natural or nat) contained the initial 21 shared bases of *CYP325C1-3* (ATGTGGTGGTTTTGGTTGGTG), and the final 21 bases of the *ompA* leader fragment. The second (rich) primer was modified with the second codon changed to an alanine from tryptophan as this base at position two is the most conserved pattern in highly expressed bacterial proteins (Tats *et al.*, 2006). Codon usage was also manipulated to increase AT richness as this aids DNA melting and subsequent transcription (Imai *et al.*, 1993). A total of 3 changes were made to the 'rich' primer compared with the 'natural' primer. After PCR 1, this linker fragment was employed in the fusion PCR (2) (Figure 4.5) with the *ompA*+2 forward primer, P450 template and a P450 specific reverse primer for the 3' end of the gene. This primer was designed with two possible restriction sites, *XbaI* and *EcoRI*.

Primers for linker fragment (product: 99bp)

NdeI *ompA* 5'
ompA+2 F – ggaggt**catatg**aaaaagacagctatcgcg

CYP325C1-3, 5' ***Pro|Ala*** *ompA* 3'
Linker (nat) –caccaaccaaaaccaccacat**cggagc**ggcctgcgctacggtagcgaa

AT-rich ***Ala***
Linker (rich) –cac***taa***ccaaaacca***cgc***cat**cggagc**ggcctgcgctacggtagcgaa

Primers for Cyp325C+*ompA* (product: 1,578bp)

ompA+2 F – ggaggt**catatg**aaaaagacagctatcgcg

XbaI /EcoRI CYP325C1-3, 3'
Cyp325C1-3 R+RE – ***tctaga/gaattc***ctacatcctcttctctacctgtac

Figure 4.5. *OmpA*+2 P450 cloning primers for *CYP325C* genes from *An. gambiae s.s.*.

Primers in PCR 1 are used to produce the *ompA* linker fragment while PCR 2 primers are employed in fusion PCR creating the *ompA*-P450 construct.

PCR conditions for PCR1 and 2

PCR 1 employed a plasmid template containing the *ompA* leader sequence (provided by M. Paine). Fifty microlitre reactions were performed using 1µl of plasmid template, 1 unit of taq (KAPA), 200µM dNTPs, 1X PCR buffer (KAPA), 0.4µM of the *ompA*+2 forward primer and 0.4µM of either the natural or rich primer. Reactions were incubated at 95°C for 2 minutes followed by 35 cycles of 95°C for 30 seconds, 60°C for 30 seconds, 72°C for 6 seconds followed by a final extension of 72°C for 2 minutes.

Only the 'rich' forward primer PCR resulted in strong products when run on a 1.5% agarose gel (0.5µg/ml ethidium bromide). This 99bp fragment was PCR purified using a column based kit (QIAquick - Qiagen) according to manufacturer's instructions and eluted in 50µl of nuclease free water (Sigma).

For PCR 2, three clones previously amplified using *CYP325C1-3* conserved primers (A3, A2, A2_1) were selected as template. These clones represented the three major sequence variants amplified from cDNA originating from Ghanaian *An. gambiae s.s.* (Figure 4.3). Fifty microlitre reactions contained 2 units of hi-fidelity taq (HiFi – KAPA), 1X HiFi buffer, 0.3mM dNTPs, 0.3mM *ompA2*+F primer, 0.3mM *CYP325C1-3* R+RE reverse primer with 50ng each of purified 99bp linker and plasmid *CYP325C1-3* template. Reactions were incubated at 95°C for 2 minutes followed by 30 cycles of 98°C for 20 seconds, 45°C for 15 seconds, 72°C for 1.5 minutes then a final extension at 72°C for 2 minutes. A 5µl aliquot of each PCR reaction was visualised on a 1.5% agarose gel (0.5µg/ml ethidium bromide). The remaining 45µl was then run on a 1% agarose gel without ethidium bromide and stained with 0.002% methylene blue (w/v, Sigma) solution for 1-2 hours until bands were visible. This staining was used to avoid damage to DNA cause by exposure to UV light and to increase transformation efficiency (Huang *et al.*, 2010). Bands were excised from the gel and DNA extracted using a column based kit according to instructions (QIAquick Gel Extraction Kit – Qiagen) and eluted in 50µl nuclease free water (Sigma). DNA concentration was then measured on the NanoDrop™ spectrophotometer (Thermo Scientific).

NdeI and XbaI restriction digest

An aliquot of cloning vector pCW containing a 2.5kb insert was obtained (courtesy of M. Paine) and 1µl used to transform 50µl of chemically competent *E.coli* cells (JM109) via heat-shock at 42°C. Cells were plated and positive colonies selected for 5ml overnight culture. Three overnight cultures were mini-prepped to acquire plasmid DNA using a column based kit (QIAprep Spin Miniprep – Qiagen) which was subsequently digested with *NdeI* and *XbaI* to remove the insert ready for *CYP325C* ligation. Three micro-grams of pCW vector were digested in a 50µl reaction with 4 units of *NdeI* (Promega) and 1 unit *XbaI* (Promega) in 1x buffer (D – Promega) and 1X BSA (Promega). The reaction was incubated for 3 hours at 37°C before being run on a 0.8% agarose gel (0.5µg/ml ethidium bromide) and briefly exposed to UV before excising the double cut vector and purifying using a column based kit (QIAquick Gel Extraction Kit – Qiagen).

The PCR amplified *CYP325C1-3* gene fragments were also digested with *NdeI* and *XbaI* to produce complementary sticky ends for ligation into the pCW vector. Approximately 3µg of each PCR amplified gene was digested in a 50µl reaction as described previously. Digests were purified using a PCR purification column (QIAquick – Qiagen) to remove small DNA fragments

Ligation and transformation

A three to one ligation ratio was calculated following Manufacturers' recommendations (Promega) for the 1.5Kb insert and 100ng of the 5Kb cut vector using the following equation:

$$\text{ng of insert} = \frac{\text{ng vector X Kb insert}}{\text{Kb vector}} \text{ X 1:3}$$

Twenty micro-litre ligations were set up using 90ng of digested insert, 100ng of digested plasmid with 1 unit of T4 ligase (Promega) in 1X ligase buffer and incubated at 4°C overnight. Transformation via heat-shock was performed using 5µl of each ligation mixture and 50µl of chemically competent *E.coli* (DH5α). Post cell recovery and plating overnight, colonies were screened using P450 specific primers.

After this screen all colonies were negative for CYP325C inserts due to an internal *NdeI* restriction site in the CYP353C sequences which was missed during initial screens (see Results and Discussion). As *NdeI* is required to be the 5' restriction site for this method of cloning it was decided to continue with these primers and constructs. Sub-cloning of the ompA-P450 fragments into a holding vector then restriction digest with a serial dilution of *NdeI* to optimise the full length product was pursued.

Sub-cloning into holding vector

OmpA-CYP325C fragments were re-amplified using the PCR 2 protocol previously outline. Two micro-litres were run on a 1.5% agarose gel (0.5µg/ml ethidium bromide) to check fragments, before PCR products were cleaned via a column (QIAquick – Qiagen) and eluted in 30µl of nuclease free water (Sigma). *CYP325C1-3* PCR amplified fragments were then cloned into pJET 1.2/blunt holding vector (Fermentas) following previously outline protocols. Colonies were screened using *ompA*+2F and *Cyp325C1-3* R+RE primers and positive colonies grown in 5ml cultures overnight before mini-prepping (QIAprep Spin Miniprep – Qiagen). Plasmids were sequenced to confirm the correct genes had been cloned.

This work has been taken forward in collaboration with C. Wilding and is summarised below:

A serial dilution of restriction enzyme *NdeI* was employed with previously outlined digest conditions to optimise the full length P450 product and avoid internal *NdeI* digestion of the *CYP325C1-3* genes in the pJET holding vector. Aliquots of each product were visualised on an agarose gel before the optimal conditions were identified and the digest repeated. Products from the repeat digest were run out on a 0.8%

agarose gel and stained with methylene blue before full length *ompA*+2-*CYP325C1-3* products were excised and DNA extracted using a column based kit (QIAquick Gel Extraction Kit – Qiagen). Ligation into pCW vector was then repeated as previously outlined and plasmids sequenced to confirm that genes were in the correct expression vector.

Chemically competent *E.coli* cells (JM109) were then co-transformed with the pCW vector containing the *CYP325C1-3* genes along with a plasmid containing redox partner Cytochrome P450 Reductase (CPR) (Courtesy of M. Paine). Colonies were screened via P450 and CPR specific PCR to confirm both plasmids were present before positive colonies were used to inoculate 200ml of Terrific Broth (Fisher Scientific) containing selective antibiotics (ampicillin (P450-pCW) and chloramphenicol (CPR)). These cultures were then incubated at 37°C, 150RPM, and once optimal OD had been reached (~0.7-0.8 OD), were induced with IPTG, and heme precursor aminolevulinic acid (ALA) added before incubating at a reduced temperature (20°C) for 1-2 days. Aliquots of the induced culture were then removed and checked for P450 peaks via spectrophotometric analysis following the carbon monoxide spectra protocol of Prichard *et al.* (2005).

In an attempt to increase P450 expression levels, gene synthesis is being pursued which will optimise the codon usage for this *E.coli* expression system (Gustafsson *et al.*, 2004). Once optimised, work will continue on the expression and purification of *CYP325C* proteins, which if successful will be followed by insecticide metabolism assays.

4.4 Results

4.4.1 GeXP Expression validation

4.4.1.1 Primer testing

Eleven out of 21 GeXP primer pairs along with S7 and KAN^r produced a single product of predicted size. PCR product was absent for 6 genes and 4 primer pairs produced either multiple peaks or peaks of an unexpected size. A summary of primer optimisation results is shown in Table 4.1.

Table 4.1. Summary of GeXP results for 21 genes using a multiplex of reverse primers and individual forward primers.

Primers were designed for genes identified from a microarray study. Predicted product sizes from Ensembl sequence data are shown with actual fragment sizes recorded using the GeXP system (Beckman-Coulter).

GENE NAME	Predicted product size	Product	Fragment sizes
AD28229	151	Yes	153.77
COEJHE2E	299	Yes	299.61
CYP15B1	144	Yes	145.45
CYP9J5	332	Yes	332.6
DH31	186	Yes	184.33
GSTO1	235	Yes	232.32
GSTs1_1	172	Yes	172.16
GSTS1_2	137	Yes	138.5
LRR15	215	Yes	215
tryp-4 prec	207	Yes	207.24
tryp-7 prec	242	Yes	243.5
CX819818	158	Additional	158.94 & 222.49
PX14	193	Additional	191.39 & 221.14
CTL9	200	Incorrect size	232.33
CYP325C1	165	Incorrect size	224.16
COEJHE1E	249	No	-
CYP325C3	179	No	-
LRR17	280	No	-
OBP3	270	No	-
PGRPS2	313	No	-
tryp-6 prec	306	No	-

As the majority of primer pairs gave products of expected size (15/23 including S7 and KAN^r) and the GeXP analysis software allows unwanted peaks to be removed from analysis it was decided to continue with this multiplex design.

4.4.1.2 Final multiplex

After full analysis of GeXP fragments using the eXpress Analysis programme, ribosomal S7 normalised gene expression levels were calculated for each gene in all 10 RNA pools (5 test, 5 control). Normalised gene expression for each RNA pool from three technical replicates (15 measurements per gene) were combined to produce a mean expression value per gene per treatment (test or control) (Figure 4.7) and an expression ratio per gene (test/control) (Appendix 21).

When $\log_2$ ratios for GeXP measured expression were plotted against those calculated from microarray data the Pearson Correlation Coefficient was 0.7867 and significant

when assessed by t-test($p<0.001$) (http://faculty.vassar.edu/lowry/ch4apx.html) (Figure 4.6). Eleven of the 13 genes showed good correlation and therefore concordance with previously determined microarray fold changes; however two genes, *CYP15B1* and *CX819818* appeared to be outliers (Figure 4.6).

GeXP data replicates were generally consistent indicated by standard deviation measurements (Appendix 21) and standard error bars (Figure 4.7). Although less precision in expression measurement were recorded for genes *CYP15B1* and trypsins 4 and 7. The variation in results for *CYP15B1* may have contributed the lack of correlation with microarray data.

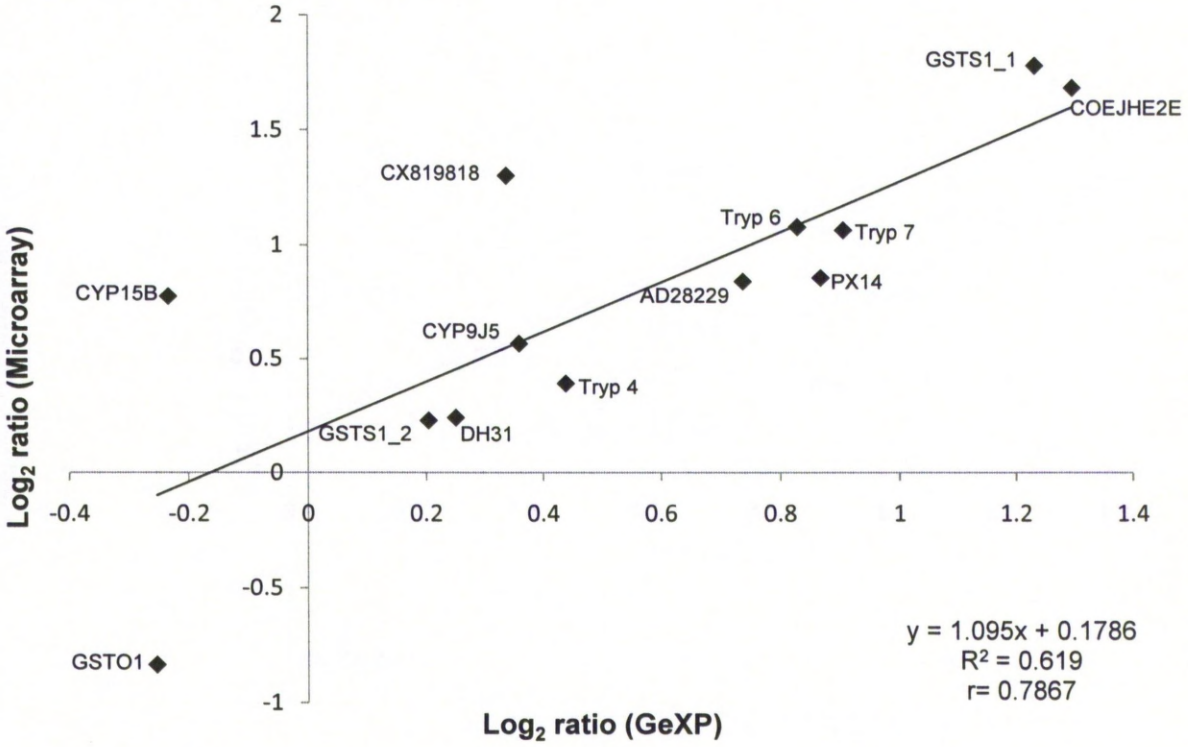

Figure 4.6. Correlation between log_2 ratios measured for 13 genes by GeXP gene expression analysis and whole genome microarray.

The expression of 13 genes from a microarray experiment comparing bendiocarb resistant and non-exposed *An. gambiae s.s.* was validated using the GeXP multiplex quantitative PCR system (Beckman-Coulter). Three replicate expression measurements from 5 resistant and 5 control RNA pools were made for each of the 13 genes via GeXP and mean ribosomal S7 normalised expression used to calculate the resistant/control ratio. Microarray expression ratios were calculated from an interwoven loop experimental design comparing the same ten RNA pools using an ANOVA based model.

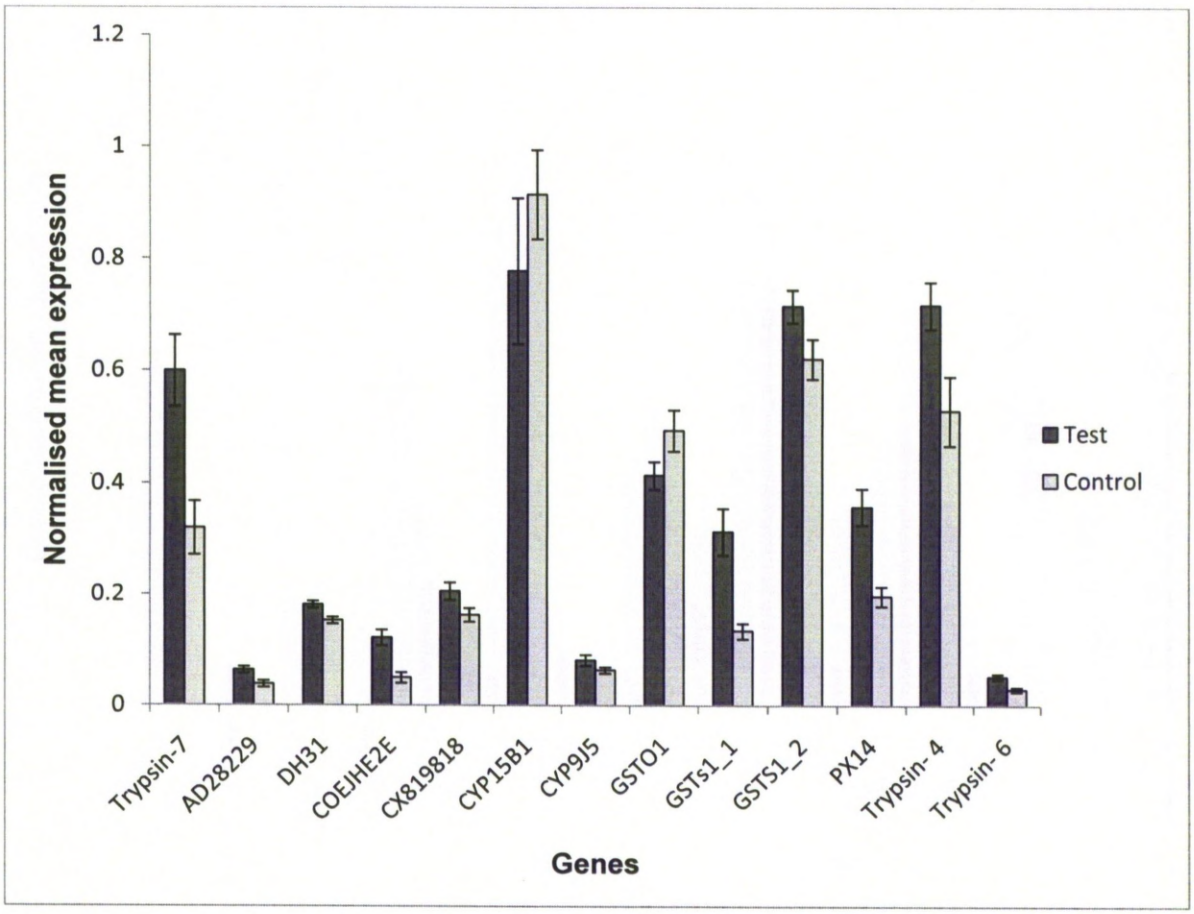

Figure 4.7. GeXP gene expression results for 13 genes selected from a microarray study comparing bendiocarb resistant *An. gambiae s.s* (Test) and non-exposed Controls.

Mean gene expression and standard errors from 5 test and 5 control RNA pools (3 technical replicates) are shown normalised against housekeeping gene ribosomal S7. See Appendix 21 for details.

4.4.2 Recombinant protein validation

4.4.2.1 *GSTS1_1* assay

Mean specific activity of recombinant enzyme *GSTS1_1* with substrate CDNB was higher (0.62 µmol/mg/min) than that of *GSTS1_1* pre-incubated with insecticide bendiocarb (0.53 µmol/mg/min); however this 13.5% activity inhibition was not significant when assessed by t-test (t=0.6655, p=0.5421) (Table 4.2). No apparent inhibition was recorded with GSTS1 splice variant *GSTS1_2*; no significant difference in activity was recorded ± bendiocarb (t=0.8413, p=0.4475).

Some inconsistency was witnessed in the third experimental replicate especially in the case of *GSTS1_1*, where CDNB activity was higher post bendiocarb incubation in opposition to the two previous replicates.

Table 4.2. Bendiocarb GST-CDNB activity inhibition results for *An. gambiae s.s. GSTS1_1* and *GSTS1_2*.

Recombinant enzyme was incubated ± bendiocarb (200µM), and CDNB activity slope recorded over 1 minute for each protein (3 replicates). Specific activity (µmol/mg/min) was calculated from each slope (specific activity = ((gradient – blank) * (1/9.6[a]) * (assay vol. / enzyme vol.)) / (conc. Enzyme mg/ml)) and mean specific activity used to determine percentage inhibition. CDNB Activity ± bendiocarb (200µM) was assessed using a t-test for each splice variant and the p-values for significance supplied.

	Rep. 1 slope	Specific activity µmol/mg/min	Rep. 2 slope	Specific activity	Rep. 3 slope	Specific activity	Specific activity			Specific activity mean [95% CI] µmol/mg/min	% inhibition	t-test p-val
							Mean	SD	S.E.M			
GSTS1_1	0.0849	0.7591	0.0785	0.6979	0.0498	0.3990	0.6187	1.93E-01	1.11E-01	0.6205±0.4788	13.54	0.54
GSTS1_1 + bendio	0.0746	0.6526	0.0575	0.4792	0.0569	0.4729	0.5349	1.02E-01	5.89E-02	0.5349±0.2534		
GSTS1_2	0.1367	1.3047	0.1246	1.1786	0.1207	1.1380	1.2071	8.69E-02	5.02E-02	1.2071±0.2160	-4.63	0.45
GSTS1_2 + bendio	0.1249	1.1817	0.1340	1.2766	0.1392	1.3307	1.2630	7.54E-02	4.35E-02	1.2630±0.1874		

4.4.2.2 *OBP3* displacement assay (Assays conducted by J-J. Zhou, Rothamsted Research)

Results from this initial assay suggested bendiocarb was responsible for a degree of competitive fluorophore displacement from *An. gambiae s.s.* recombinant odorant-binding protein 3. Fluorescence was reduced by over 30% at bendiocarb concentrations greater than 12µM compared to 1-NPN bound to *OBP3* alone (Table 4.3), and a steady reduction in fluorescence was recorded as bendiocarb concentration was increased (Figure 4.8). According to results from a number of displacement assays, bendiocarb had the 5th best affinity to *OBP3* from 26 compounds measured (J-J. Zhou, pers. comm.).

Table 4.3. Bendiocarb induced displacement of fluorophore N-phenyl-1-naphthylamine (1-NPN) from *An. gambiae s.s.* recombinant odorant binding protein three (*OBP3*).

Bendiocarb was introduced to recombinant *OBP3* bound to ligand 1-NPN at increasing concentrations and the change in fluorescence, on 1-NPN displacement, recorded. Percentage reduction in fluorescence was calculated with respect to *OBP3* bound to 1-NPN in the absence of bendiocarb. All assays were performed by J-J. Zhou, Rothamsted Research.

Fluorescence intensity reading (407nm)	[Bendiocarb] µM	1-NPN fluorescence (%)	Reduction (%)
113.80	0	100.00	0.00
99.11	2	87.09	12.91
94.03	4	82.63	17.37
89.57	6	78.71	21.29
85.89	8	75.47	24.53
81.74	10	71.83	28.17
78.53	12	69.01	30.99
76.92	14	67.59	32.41
72.79	16	63.96	36.04

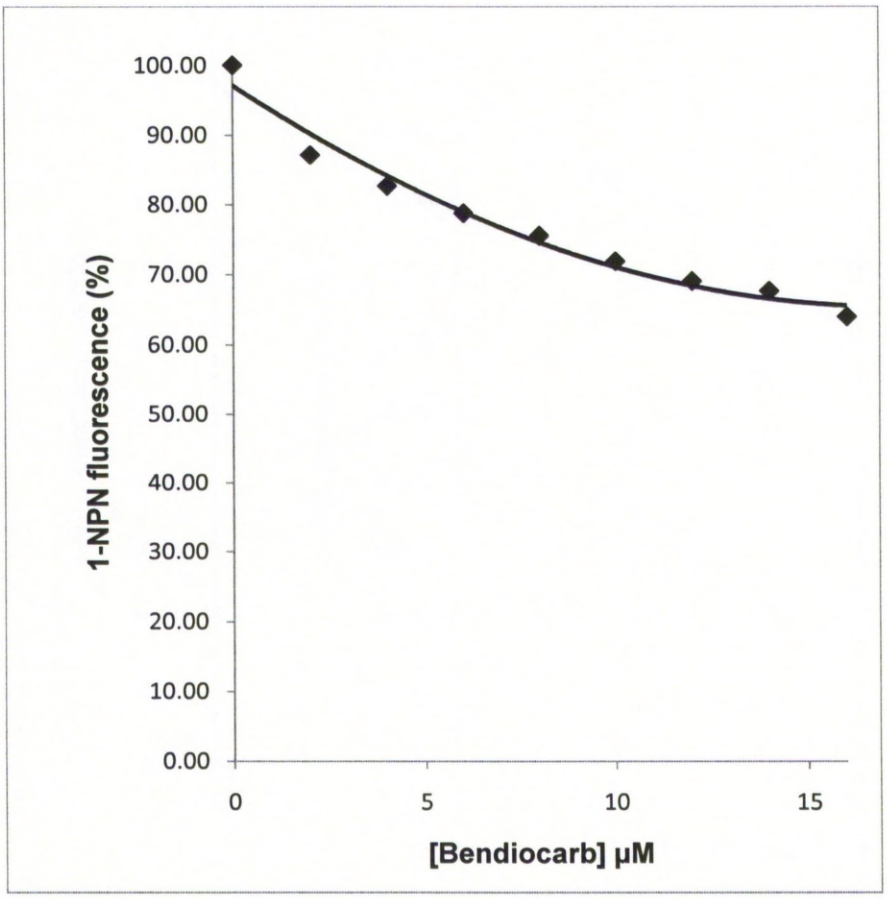

Figure 4.8. Competitive displacement of fluorophore N-phenyl-1-naphthylamine (1-NPN) from *An. gambiae s.s.* recombinant odorant binding protein three (*OBP3*) after incubation with carbamate insecticide bendiocarb at 2µM concentration increments up to 16 µM.

The change in fluorescence on 1-NPN displacement was recorded at each bendiocarb concentration and percentage reduction calculated with respect to *OBP3* bound to 1-NPN alone. All assays were performed by J-J. Zhou, Rothamsted Research.

4.4.2.3 CYP325C sequencing and cloning

Sequencing of *CYP325C1-3* genes amplified from field derived cDNA identified three *CYP325C* variants. These genes, amplified with primers designed to conserved 3' and 5' regions of all three P450s, more closely resembled the *CYP325C2* sequence from the Ensembl database (version P3) (Figure 4.3).

Using the *ompA*+2 P450 expression system (Pritchard *et al.*, 2005), fusion PCR successfully introduced the 21 amino-acid *ompA* leader sequence plus 5' and 3' restriction sites to clones representing the three sequenced *CYP325C* variants. However an internal *NdeI* restriction site meant that subsequent cloning and ligation into the pCW expression vector was unsuccessful.

The *CYP325C* constructs were cloned into a holding vector and using a serial dilution of *NdeI,* a partial digest successfully excised the three CYP325C sequences which were then cloned into the pCW expression vector. P450 expression plasmids were co-

transformed with cytochrome P450 reductase into competent *E.coli* cells (JM109 AllianceBio). Post IPTG induction, P450 expression levels were poor indicated by low P450 peaks on spectrophotometric analysis. Optimisation of protein expression is being pursued.

4.5 Discussion

4.5.1 Expression validation

Overall, microarray and GeXP expression ratios were in concordance with a significant Pearson correlation coefficient ($R=0.787$, $p<0.001$) (Figure 4.6). Genes which showed good correlation included Trypsins 4, 6 and 7, *COEJHE2E*, *GSTS1_2*, *CYP9J5* and *DH31*. Of the 13 genes for which microarray and GeXP fold changes were compared two potential outliers, *CYP15B1* and *CX819818,* were identified which affected overall correlation; without these outliers $R= 0.9783$. The GeXP results for outlier *CYP15B1* were highly variable with some sample replicates failing to provide signal. This variation was reflected in large standard deviations (Appendix 21) and error bars (Figure 4.7) and inevitably affected concordance between microarray and GeXP data.

Microarray and GeXP fold changes were compared for a subset (13) of the 21 genes for which GeXP primers were originally designed as 2 genes produced GeXP peaks of incorrect size while primers for 6 genes failed to produce a PCR product (Table 4.1). A similar absence of PCR product was witnessed by Muller *et al.* (2008b) for two genes when attempting to validate microarray data via GeXP. Failure of GeXP product amplification may be a result of poor primer design, divergence between primer and sample sequences or even an absence of transcript in biological samples. While discrepancies between observed and expected GeXP peak size may be explained by gene length polymorphisms between samples and database sequences; though this hypothesis must be explored through sequencing.

As primers for GeXP were designed from database sequences rather than sequence data from field sample the level of polymorphism, whether single nucleotides (SNPs) or length differences, between field samples and database sequences is unknown. Polymorphisms in regions of primer binding may affect amplification and therefore expression results. Indeed this may have been an issue with gene *OBP3* as GeXP primers designed to the Ensembl sequence failed to provide a peak. On subsequent sequencing of this gene from Ghanaian field samples the GeXP reverse primer was found to be situated in an area of high sequence polymorphism. However extensive sequencing of all genes in all samples was not feasible.

The multiplex GeXP system has been previously employed for microarray validation (Muller *et al.*, 2008b; Felix *et al.*, 2010) with varying success. Felix *et al.* (2010) measured expression of a subset of 20 *An. gambiae* genes identified via microarray as differentially expressed in two tissue types and found good concordance between the two measurements (Pearson correlation coefficient, $R =0.85- 0.884$). While correlation between microarray and GeXP expression levels of 14 genes was poor in an earlier study ($R=0.129$) which was thought to be associated with low levels of fold change

(Muller *et al.*, 2008b). This technique has the advantage of measuring the expression levels of up to 30 genes in a single assay with automatic primer design in the Genomelab GeXP eXpress software circumventing laborious primer construction. Another notably benefit of this system is the small amounts of RNA required for each assay (100ng/reaction to measure up to 30 genes compared to ~33ng/reaction for single locus qPCR), which is especially important when material is limited.

Exact concordance between microarray and qPCR validation techniques is unlikely due to inherent differences in the reverse transcription and amplification chemistries (Freedman *et al.*, 1999) as well as the normalisation process, which for microarray is based on a global normalisation while qPCR data is routinely normalises against reference genes (Morey *et al.*, 2006). Differences in the primer/probe positions used for the microarray and qPCR can also affect agreement between methods (Etienne *et al.*, 2004). However good concordance was recorded for eleven genes using GeXP quantitative PCR.

For three microarray candidate genes for which GeXP was either unsuccessful (*OBP3, CYP325C1-3*) or failed to produce expression ratios which matched microarray data (*GSTS1_1*), a "gold standard" approach of Real-Time qPCR was taken (Appendix 22). This method of validation is employed extensively in microarray experiments (Burney *et al.*, 2007; Abercrombie *et al.*, 2008; Dong *et al.*, 2006; Vontas *et al.*, 2005) and can be used for relative or absolute gene expression analysis. Absolute transcript quantification using plasmid standard curves was pursued which gives this method a potential advantage over GeXP in accuracy. While sequence data from field samples was employed, where possible, for primer design to ensure primers were placed in areas of conserved sequence. However plasmid preparation and sequencing before primer design is laborious and was therefore only feasible for a small number of genes.

Expression ratios calculated from the Real-Time qPCR analysis were not in support of microarray data (Results, Appendix 22). A number of confounding effects potentially impacted qPCR results including: suboptimal amplification efficiency due to restricted primer positioning; high sequence identity between closely related genes hindering gene specific primer design; and unknown levels of polymorphism in field samples due to poor sequence reads (Discussion, Appendix 22).

Ultimately, biological variation has the greatest impact on expression data (Morey *et al.*, 2006) and cannot be controlled for; although 'noise' can be reduced by adequate replication. As discussed in the previous chapter biological variation and experimental design effected p-values calculated for microarray probes and may have also influenced the qPCR validation results reported here. However this investigation was designed to be exploratory with an aim to identify genes with putative involvement in phenotype. Subsequent proteomic analysis and techniques such as RNA interference will be required to confidently associate genes with insecticide resistance.

4.5.2 Recombinant protein validation

4.5.2.1 *GSTS1_1*

The Sigma class glutathione-S–transferase, *GSTS1,* in *An. gambiae* is a cytosolic GSTs which is alternatively spliced creating two variants, *GSTS1_1* and *GSTS1_2* (Ding *et al.*, 2003). Alternative splicing of GSTs in *An. gambiae* is a common mechanism of increasing heterogeneity and therefore substrate plasticity (Ranson *et al.*, 1998). Overall the *GSTS1_1* and *S1_2* variants share 75% sequence identity, however higher identity is recorded in the N terminal region (93%) due to a common first exon. *GSTS1_1* was found to be over-expressed in the bendiocarb resistant *An. gambiae s.s.* field samples from Ghana, West Africa when compared to non-exposed controls with a mean expression ratio of 3.43 recorded via microarray. The C terminus or 3' region of each gene was the region employed for microarray probe design so cross-hybridisation of probes was unlikely, while qPCR primers were selected to be outside the common exon.

GSTS1_1-CDNB activity was inhibited by ~14% when the enzyme was pre-incubated with 200µM bendiocarb for 5 minutes prior to assay (Table 4.2) with no apparent inhibition recorded for splice variant *GSTS1_2.* However confidence intervals for specific activity were overlapping which suggests differences in GST activity ± bendiocarb were not significant. Moreover, both enzymes have previously displayed much higher levels of CDNB inhibition with other insecticide classes (Dowd *et al.* in prep.). Dowd *et al* characterised the activity of recombinant *An. gambiae GSTS1_1* and *S1_2* with various GST substrates and insecticides. Activity inhibition levels of ~50% were recorded in both variants after incubation with 100µM permethrin, while *GSTS1_1* also showed ~40% reduction in CDNB activity when pre-incubated with DDT and two additional pyrethroid insecticides.

GSTS1_2 was previously reported to be over-expressed in a pyrethroid (permethrin) resistant colony of *An. gambiae s.s.* also originating from Ghana, West Africa (Muller *et al.*, 2007). In a microarray comparison *GSTS1_2* was 2.5 fold over-expressed in the Ghanaian (Odumasy) colony compared with a susceptible strain (Kisumu); though as Kisumu originated from East Africa, geographical differences may have confounded results. However further evidence of a potential link to resistance came from a subsequent microarray experiment from the same group where they reported 3.5 fold over-expression of *GSTS1_2* in a population of *An. arabiensis* after a deltamethrin spraying campaign in Cameroon (Muller *et al.*, 2008a). While both variant *GSTS1_1* and *S1_2* have been demonstrated as up-regulated in adult stages of a permethrin resistant colony of *An. gambiae* compared to larval stages via microarray analysis (Strode *et al.*, 2006).

The *GSTS1_1*-CDNB inhibition levels reported by Dowd *et al.* with insecticides of various classes in combination with previous links to pyrethroid resistance suggest that any bendiocarb interaction is likely to be non-specific. Indeed the inhibition witnessed post incubation with at least 8 different compounds advocates a more generalist role for *GSTS1_1 (Dowd et al. in prep,).* Grant and Matsumura (1989) predicted Sigma

GSTs in *Ae. aegypti* to have a "general" role in resistance to a number of different insecticides after elevated Sigma class GST activity was recorded in DDT and pyrethroid resistant strains. *GSTS1_1* could also be involved in binding and trafficking xenobiotics such as insecticides in the cytosol of insects. Other cytosolic GSTs are reportedly involved in binding non-substrate ligands (Zucker *et al.*, 1995; Vargo and Colman, 2001; Oakley *et al.*, 1999) thus contributing to intracellular transport, sequestration and disposition of compounds (Hayes *et al.*, 2005).

Some authors suggest a role for Sigma class GSTs in protection against oxidative stress in insects (Dowd *et al.* in prep (Singh *et al.*, 2001; Ranson *et al.*, 2005)). Singh *et al.* reported on a *Drosophila* sigma class GST (*DmGSTS1-1*) displaying high glutathione conjugating activity with a product of lipid peroxidation (4-hydroxynonenal). While expression of Sigma class GSTs in the house-fly *Musca domestica* was found to be concentrated in the indirect flight muscles (Franciosa and Berge, 1995), which may be necessary to protect this highly metabolic tissue from oxidative stress (Ranson *et al.*, 2005).

Oxidative stress protection may be integral to insecticide resistance through circumventing insecticide induced stresses. Vontas *et al.* (2001) reported elevation of GSTs with high levels of peroxidise activity in a pyrethroid resistant strain of planthopper, *Nilaparvata lugens*. The authors hypothesised a role for these enzymes in protecting the insect from pyrethroid induced oxidative stress and lipid peroxidation, as when GSTs were inhibited *in vivo,* increased susceptibility to pyrethroid induced lipid peroxidation was witnessed. Subsequently a specific GST was isolated (*nlgst1-1*) which was over-expressed in the resistant strain and was shown to have high levels of peroxidation activity in recombinant protein assays (Vontas *et al.*, 2002). Similar peroxidase activity has also been demonstrated in recombinant *An. gambiae GSTS1_1* through activity with 4-hydroxynonenal (4-HNE) (Dowd *et al.* in prep). A role in protecting the insect from pyrethroid induced lipid peroxidation may also explain the elevated expression of *GSTS1* variants in pyrethroid resistant strains of *An. gambiae* (Muller *et al.*, 2007; Muller *et al.*, 2008a; Strode *et al.*, 2006). Carbamate insecticides are also likely to induce oxidative stresses within the mosquito, with mammalian model systems providing evidence of this action (Kamboj *et al.*, 2006; Gupta *et al.*, 2007).

4.5.2.2 *OBP3*

OBP3 is a classical odorant-binding protein characterised by 6 conserved cysteine residues (Xu *et al.*, 2003) and is one of approximately 66 identified in *An. gambiae s.s.* to date (Zhou *et al.*, 2010). Currently a group at Rothamsted Research institute are investigating the roles of various *An. gambiae* OBPs in host seeking and mate attraction.

Carbamate insecticide bendiocarb was found to competitively displace *fluorophore* 1-NPN from recombinant A*n. gambiae OBP3* with more than 30% displacement recorded at bendiocarb concentrations >12µM. This preliminary experiment may suggest interaction between *OBP3* and bendiocarb and a potential role in trafficking insecticides *in vivo*.

OBPs are mainly implicated in the transport of semio-chemicals across the sensillium of the antennae to odorant receptors (Zhou *et al.*, 2010). However OBP binding studies to-date suggest broad ligand specificity (Pelosi *et al.*, 2006). Multiple functions have been predicted for *Ae. aegypti* OBPs through evidence of interaction with numerous organic compounds (Li *et al.*, 2008), while the silk worm (*Bombyx mori*) OBP, *PBP1,* was recently reported to bind non-pheromone compounds with a higher affinity than that for pheromones (Hooper *et al.*, 2009). A role for OBPs outside of semio-chemical detection and transport is supported, in part, by the *in vivo* expression of these proteins in areas other than the sensory appendages of the insect head. *Ae. aegypti OBP22* was reportedly expressed in the male and female reproductive appendages and in the respiratory spiracles via immuno fluorescence (Li *et al.*, 2008), while *Drosophila* OBPs were expressed in gustatory sensilla supporting a role in taste (Galindo and Smith, 2001). Li *et al.* (2005) recorded expression of *An. gambiae* OBPs in the legs and body as well as the head through semi-quantitative reverse transcription PCR; however they reported *OBP3* expression to be restricted to the head.

If indeed *OBP3* expression in *An. gambiae* is restricted to the mosquito head a role outside of classical semio-chemical olfaction is still possible. *OBP3* could be involved in trafficking insecticide volatiles to sensory receptors inducing signalling cascades and expression of genes involved in behavioural and or metabolic resistance. This is a speculative hypothesis; however the elevated expression of a putative OBP receptor in the bendiocarb resistant samples may support involvement of an insecticide trafficking and signalling pathway. As tissue specific profiling was not performed in this study and Li *et al.* (2005) relied on semi-quantitative PCR for expression profiling, it will be important to characterise the *in vivo* expression of *OBP3* in bendiocarb resistant field samples. Techniques such as absolute qPCR of cDNA derived from tissue dissection or *in situ* hybridisation could be employed to characterise the tissue specific profile.

Another consideration is the ability of an *in vitro* assay with recombinant protein to act as proxy for conditions *in vivo*. It is likely that concentrations of bendiocarb and *OBP3* used in the displacement assay are potentially unrepresentative of physiological levels in the insect and further experiments are required to support a role *in vivo*. For instance an immunofluorescence approach (Lycett *et al.*, 2006) employing an *OBP3* specific antibody and a fluorescently tagged bendiocarb compound may indicate co-localisation and binding (Huang *et al.*, 2005). While potential *OBP3*-bendiocarb binding could be explored further through co-crystallisation (Sandler *et al.*, 2000) or application of surface plasmon resonance biosensors. This technique involves ligand immobilisation on to a metal surface and the introduction of a protein solution, with interaction measured by changes in optical reflectivity of light shone onto the surface (Szabo *et al.*, 1995).

Sequence polymorphism may be important in potential OBP-ligand interaction if non-synonymous SNPs occur in the ligand binding site. High levels of polymorphism were recorded in *OBP3* sequenced from Ghanaian field samples with six polymorphic nucleotides identified when compared with the sequence employed in recombinant protein expression. Four of these SNPs were non-synonymous and the consequences

of the resultant amino acid changes unknown. The *OBP3* alleles identified from field material should therefore be investigated to establish whether amino acid alterations alter protein structure and function. Computational modelling could be applied to establish whether amino-acid changes occur within the substrate binding region; however this technique benefits from a predetermined crystal structure. To date only the crystal structure for *OBP-1* from *An. gambiae* has been resolved (Wogulis *et al.*, 2006), and as this gene shares only ~30% identity with *OBP3*, its structure may be an unsuitable scaffold. Alternatively, recombinant proteins from variant alleles could be expressed and fluorescent displacement assays repeated to establish whether the *OBP3*-bendiocarb interaction is affected.

Finally two additional OBPs were also found to be over-expressed in the bendiocarb resistant group, *OBP 19* and 45 with expression ratios of 1.48 and 1.09 respectively. The possible role for these two proteins in resistance has yet to be explored and fluorescent displacement assays with recombinant protein would be a first step in establishing a role in bendiocarb resistance.

The potential link between odorant binding proteins and insecticide resistance is an interesting and novel area of research. The use a genome-wide expression screening exposed the possible OBP association, which would have been overlooked using an array limited to detoxification associated candidates.

4.5.2.3 *CYP325C1-3*

Over 100 cytochrome P450s have been reported in *An. gambiae* (Ranson *et al.*, 2002) with a number of these genes, most commonly from the CYP6 family, associated with insecticide resistance (Muller *et al.*, 2008b; Nikou *et al.*, 2003; David *et al.*, 2005; Djouaka *et al.*, 2008).

The expression of three *CYP325C* genes, C1-3, appeared to be up-regulated in carbamate resistant field samples. Comparison of database sequences suggested that these genes were closely related sharing ~95% sequence identity (Figure 4.2). On isolation and sequencing of *CYP325C* genes from field samples, the transcripts uncovered bore most resemblance to variant *CYP325C2*. Support for involvement of this gene group in insecticide resistance was also provided by a SNP association study in which an association spike, identified in a pyrethroid resistant Ghanaian population, contained gene *CYP325C3*.

The CYP325 gene family has been previously linked to pyrethroid resistance in a number of gene expression studies. A homologue of *An. gambiae* P450, *CYP325C1*, was reportedly 1.85 fold over expressed in a permethrin resistant strain of *An. stephensi* via microarray analysis (Vontas *et al.*, 2007). While in *An. gambiae*, expression of gene *CYP325A3* was elevated (1.72 fold) in a laboratory colony with reduced susceptibility to permethrin (David *et al.*, 2005), and P450, *CYP325D2*, reportedly 5.1 fold over expressed in permethrin resistant field isolates from Benin, West Africa (Djouaka *et al.*, 2008). More recently all three *CYP325C1-3* genes were found to be up-regulated (>10 fold) in *An. arabiensis* larvae exposed to polluted breeding sites in Bobo Dioulassou, Burkina Faso (Jones *et al.* unpublished data).

Expression levels were compared to larvae from unpolluted sites using an 8x15K whole genome *An. gambiae* microarray designed in house (Mitchell and Wilding, 2009). This may suggest a role for xenobiotics in breeding sites to induce expression of this P450 family. Cross induction of detoxification enzymes via exposure to xenobiotics has been recorded elsewhere (David *et al.*, 2010; Poupardin *et al.*, 2008; Djouaka *et al.*, 2007), and as discussed (Chapter 2), the breeding sites sampled in this study were often visibly polluted. However, control samples were also selected from similar areas and are therefore likely to have been exposed to similar pollutants.

As yet, cytochrome P450s have not been directly linked to bendiocarb resistance, though this does not rule out a role in metabolism of this carbamate. However as resistance to other compounds (DDT, pyrethroids) has been recorded in the area ((Klinkenberg *et al.*, 2008), Chapter 2, Appendix 8) in combination with previous links to pyrethroid resistance, the possibility that *CYP325C* expression may be associated with resistance to other insecticides must be considered.

To explore the potential role of this P450 family in resistance to bendiocarb and other insecticides the expression of *CYP325C1-3* genes identified via sequencing of field isolates was pursued using an ompA+2 *E.coli* expression system. Ultimately recombinant protein would be assayed with insecticidal compounds to test for metabolism. The P450 expression construct along with redox partner cytochrome P450 reductase were successfully co-transformed into an *E.coli* cell line; however protein expression levels upon induction were low. Optimisation of the codon usage for this *E.coli* expression system will hopefully provide improved levels of protein expression. Alternatively these proteins could be expressed using a baculovirus insect cell line system which is more tailored for the expression of eukaryote and specifically insect proteins (Shotkoski *et al.*, 1996; McCarroll and King, 1997).

A final consideration is the potential for allelic variation within the *CYP325C* genes to alter metabolic activity. Allelic variation within detoxifying enzymes has been linked to insecticide resistance in a number of previous studies (see Discussion, Chapter 6). *CYP325C* genes sequenced in this study were isolated from a generic cDNA pool produced from unselected mosquitoes collected from the field. In future work these genes should be isolated from the test and control pools employed in the microarray study to establish whether segregating SNPs are present. Moreover, if non-synonymous SNPs are present, allelic variants should be expressed to establish whether amino acid changes affect the metabolic activity of the recombinant P450 protein.

4.5.3 Conclusions

In this study a number of promising candidates were identified via microarray analysis with GeXP expression profiling providing validation for many of these genes. This data provides a starting point for functionally validating candidates with this process initiated for a number of proteins identified.

Previous microarray expression profiling of insecticide resistant insects has provided insight into key proteins involved in the resistance phenotype. Muller *et al.* (2008b)

identified a P450 enzyme, *CYP6P3*, through microarray analysis of a pyrethroid resistant *An. gambiae s.s.* population which was subsequently found to metabolise pyrethroids. While another *An. gambiae s.s.* P450, *CYP6Z1*, identified as over-expressed in a DDT resistant strain via microarray analysis (David *et al.*, 2005) was later found to metabolise DDT in recombinant protein assays (Chiu *et al.*, 2008). Although good examples of candidate identification and validation exist there is still a deficiency in functional characterisation of genes compared to the abundance of expression data, which needs to be addressed.

In bendiocarb resistant field samples it was possible to identify potential roles for recombinant protein candidates despite the presence of the target site mutation *G119S* associated with carbamate and organophosphate resistance. Target site insensitivity alone does not fully explain the resistance phenotype as mechanisms of insecticide entry and transport within the insect are still poorly characterised. Odorant binding proteins and a putative OBP receptor were up-regulated in the bendiocarb resistant group with preliminary evidence for an interaction between recombinant *OBP3* and bendiocarb. OBPs may have a role in insecticide movement and receptor signalling *in vivo,* while cuticular proteins identified from microarray results may be involved in insecticide entry, possibly slowing movement across the cuticle.

Moreover, the breakdown and excretion of xenobiotics compounds is still necessary even in the presence of an insensitive target site. Elevated expression of enzymes from gene families associated with xenobiotic breakdown such as *GSTS1_1*, *COEJHE2E* and *CYP325C1-3* may highlight potential routes of insecticide removal or compensatory mechanisms for stresses induced by insecticide exposure. It is conceivable that this Ghanaian field population possesses a combination of metabolic and target site resistance to bendiocarb, a phenomenon which has been reported elsewhere (Corbel *et al.*, 2007; Perera *et al.*, 2008).

4.5.4 Future perspectives

A number of potential metabolic candidates highlighted in the microarray study are currently being pursued at an off-site protein production facility (OPPF- Oxford) in both *E.coli* and insect cell-line systems. If successful these recombinant proteins will be assayed with bendiocarb and other insecticides to establish whether these enzymes are implicated in metabolic resistance. Enzymatic breakdown of insecticidal compounds can be measured using techniques such as high-performance liquid chromatography (HPLC)(Muller *et al.*, 2008b) and mass spectrometry (MS) (Smith *et al.*, 2006; Olsson *et al.*, 2004) which provide a profile of metabolites.

Studying the 'chemical finger-prints' left behind by the metabolism of insecticides may provide insight into the *in vivo* metabolic processes behind resistance as well as providing candidate metabolites which may undergo subsequent break down. By comparing a resistant versus a susceptible strain post-insecticide exposure, techniques such as nuclear magnetic resonance (NMR), MS and HPLC can be employed to compare metabolite profiles. Such techniques are employed extensively in the study of plant resistance to crop pests (Prashant *et al.*, 2010; Mirnezhad *et al.*, 2009; Widarto *et*

al., 2006). Metabolomic techniques have also been applied to insects with a recent study exploring an insect-bacterial symbiotic relationship in aphids via NMR (Wang *et al.*, 2010b), while MS systems are in development for metabolome analysis in *Drosophila* (Kamleh *et al.*, 2009).

It is important to provide evidence of a role for promising resistance candidates *in vivo*. Techniques such as transgenic over-expression or expression knock down through RNA interference (RNAi) can be employed to validate candidates within the mosquito. Transgenic mosquitoes which over-express a gene of interest can be bioassayed for potential changes to insecticide resistance phenotype. Gene expression may be introduced via transgenesis with viruses such as *An. gambiae* densovirus (Ren *et al.*, 2008) or through germ-line transformation with expression constructs employing inducible promoters (Shin *et al.*, 2003). For this technique a susceptible transgenic mosquito colony would be required in which to over-express genes of interest to establish links to insecticide resistance.

RNAi involves introducing double stranded RNA, complementary to a gene of interest, into the insect. The complementary RNA is processed *in vivo* by the enzyme dicer creating small interfering RNAs which can bind to an RNA-inducing silencing complex and silence complementary transcripts (Hannon, 2002). RNA can be introduced into the mosquito via microinjection (Blandin *et al.*, 2002) or feeding (Zhang *et al.*, 2010) and gene knockdowns of between 50-90% have been reported in mosquitoes (Zhang *et al.*, 2010; Boisson *et al.*, 2006; Lycett *et al.*, 2006). Once a gene has been knocked down the potential effects on resistance phenotype can be accessed via insecticide bioassays. However this technique requires resistant mosquitoes known to over-express the allele of interest. This would normally involve colonisation and insecticide selection of a field isolated population which can be problematic and time consuming.

However single gene knock-down or over-expression may not reveal the cause of resistance as the often quantitative nature of this trait means a number of proteins are likely to be involved in producing the phenotype. Furthermore, there are likely to be alternative routes of metabolism if one pathway is interrupted.

Identifying key indicators of resistance will enable the development of tools to monitor resistance in the field as well as providing candidates for genetic manipulation to potentially increase susceptibility and prolong the active life-span of current insecticidal compounds. The use of whole genome profiling will hopefully provide insight into the complete resistance process from routes of insecticide entry, to transport and pathways involved in metabolism and excretion. In addition, by exposing the potential routes of metabolic break-down, tailored design of new compounds potentially resistant to metabolism or incorporation of enzyme inhibitors rendering insect proteins incapable of breaking down insecticides are possibilities.

Chapter 5: Microarray analysis of DDT resistant *An. gambiae s.s.* from Ghana, West Africa

5.1 Abstract

The expression profile of a highly DDT resistant population of *An. gambiae s.s.* from Ghana, West Africa was investigated using a 15K whole genome microarray. Due to a lack of susceptible material from the field site, the resistant Ghanaian population was compared with two West African colony controls displaying no or little DDT resistance. A number of genes were consistently and significantly over-expressed in the DDT resistant samples with respect to both colony strains. These included a number of candidates which have previously been linked to insecticide resistance including *CYP6M2* and *CYP6Z3*. However a number of 'novel-candidates' were identified including a cuticular gene, *CPR9*, an arrestin, *ARR2*, and a putative membrane transporter. The identification of potentially novel resistance mechanisms highlights the importance of whole transcriptome analysis when performing genotype-phenotype association studies.

5.2 Introduction

In light of increasing evidence of pyrethroid resistance in Ghana, West Africa (Coetzee *et al.*, 2006; Klinkenberg, 2006; Muller *et al.*, 2008b) it was decided to phenotype wild populations of *An. gambiae s.s.* for resistance to alternative insecticides which may replace pyrethroids for mosquito control. One of these compounds, DDT, has seen resurgence in use for indoor residual spraying (IRS) throughout Sub-Saharan Africa, bolstered by WHO advocacy for its employment (WHO, 2006c). In 2007 over 10 African countries were using DDT routinely in IRS regimes (WHO, 2007a).

However in 2008, extremely high levels of DDT resistance were recorded in Southern Ghana, with an LT_{30} of approximately 6 hours when mosquitoes were exposed to 4% DDT under standard WHO bioassay conditions (Chapter 2). The presence of DDT resistance in Ghana had also been alluded to in previous work (Coetzee *et al.*, 2006); though reporting has been inconsistent (Kristan *et al.*, 2003).

5.3 Aims

The aim of this study was to profile the transcriptome of the highly DDT resistant field population of Ghanaian *An. gambiaes.s.* using a custom designed whole genome microarray. Resistant mosquitoes were defined as 3-5 day old females which survived a 6 hour exposure to 4% DDT under WHO bioassay conditions. These mosquitoes were compared to two West African strains of *An. gambiae s.s.* which displayed no or little resistance to DDT. Laboratory colonies were used in this experiment, as opposed to the sympatric control methodology employed in chapter 3, due to the high levels of DDT resistance which prevented identification of DDT susceptible individuals.

Expression level validation of a promising candidate gene was performed using reverse-transcription Real-Time quantitative PCR.

5.4 Methods

5.4.1 Microarray design - 8x15K 'AGAM_15K'

An Agilent 8x15K microarray was designed based on the 4x44K *An. gambiae s.s.* whole genome array used in chapter 3. The new design provided whole genome coverage but simultaneously reduced costs and increased through-put, providing eight arrays on a single slide. The entire coding transcriptome from the latest edition of the *An. gambiae s.s.* genome (Ensembl AgamP3.5, 2009) was employed in probe design, with additional probe coverage for the detoxification gene families previously implicated in insecticide resistance.

The design process was again performed using the Agilent on-line design package eArray (https://earray.chem.agilent.com/earray/), and probe design settings outlined previously were employed (Appendix 10).

The 8x15K platform had 15,208 available spots which were allocated as follows:

Whole genome probes

A single database for whole genome probe design was employed, the Ensembl AgamP3.5 annotation (September 2009), which contained 12,604 *An. gambiae s.s.* genes. A file containing all known transcripts was downloaded from VectorBase (http://agambiae.vectorbase.org/Help/AgamP3.5) and all non-coding RNAs including rRNA, snRNA and miRNAs were removed. The resultant 14,103 coding RNAs where submitted to eArray in FASTA format and a single probe designed for each transcript. This yielded 14,071 probes due to sequence duplication in the original downloaded file.

Detoxification gene probes

The 281 detox gene sequences, 230 genes from David *et al.* (2005) plus additional sequences subsequently added to this candidate microarray design, were submitted to eArray in FASTA format and three unique probes designed per gene. Three unique probes rather than duplicate probes were employed to provide greater coverage of each gene, potentially compensating for unknown polymorphisms in target sequences.

'Coefficient of Variance' (CV) probes

The first 25 CV genes from the Agilent custom *An. gambiae* microarray were selected and each replicated 10 times. These probes were used for quality control.

5.4.2 Sample pooling and experimental design

Characterisation of DDT resistance phenotype in Ghanaian *An. gambiae s.s.*, and subsequent genotype assessment, uncovered high levels of DDT resistance in both molecular forms (Chapter 2). The M molecular form provided the majority of phenotyped samples (63%), and so was selected for the microarray study. The predominant *kdr* target site genotype in M form samples was *L1014F* heterozygous, and so phenotyped individuals with this genetic background were pursued. By standardising molecular form and the presence/absence of target site mutations we hoped to control for potential confounding (Muller *et al.*, 2008b)."

Data on all M form individuals heterozygous for the *L1014F* mutation were tabulated. Females that survived DDT exposure were selected as 'test' animals. These females were drawn from collection area one (Chapter 2; Figure 2.1).

The extremely high levels of DDT resistance recorded in the field were likely to have compounded issues surrounding underestimation of fold-change when comparing insecticide selected and non-selected animals (Muller *et al.*, 2008b). Therefore material with greater DDT susceptibility was used as a comparator. Currently, to our knowledge, no DDT susceptible colonies originating from Ghana are available which in part reflects the high levels of resistance recorded in this study. A compromise was achieved by selecting two M-form West African colonies which displayed no or low level DDT resistance. Firstly the N'gousso colony originating from Cameroon, which is a fully DDT susceptible *An. gambiae s.s.* strain possessing no known *kdr* mutations (C. Witzig pers. comm.), colonised in the N'gousso district of Yaounde in 2006 (kindly provided by A. Christophe). Secondly a colony from Akron, Benin (provided by MR4, http://www.mr4.org/Reagents/tabid/303/Default.aspx?Cid=222) with low level DDT resistance (LT_{50} ~ 1 hour) and *kdr* West at a lower frequency compared to the Ghanaian population (frequency (R): Ghanaian M form controls 0.58, Akron M form colony 0.29). The use of *two* West African control colonies will hopefully adjust for potential confounding effects such as geographical origin, as candidate genes will be selected on the basis that they are differentially expressed in the DDT resistant samples with respect to both colonies.

Colony samples underwent the same treatment as DDT control samples in the field (Chapter 2). Adult females were selected at 3-5 days old and were exposed to DDT control tubes (no insecticide) for 6 hours before being held for 24 hours. Mosquitoes were then immersed in RNAlater (Ambion) and refrigerated overnight (4°C) before being transferred to the freezer (-20°C) until extractions were performed.

All RNA extractions were performed on pools of 5 female mosquitoes grouped by test or control status. All RNA was extracted using TRI Reagent (Ambion) and DNAse treated as previously described (Chapter 3). RNA extractions of DDT selected field samples was performed concomitantly to bendiocarb phenotyped extractions in which 5 females per pool were also employed. However due to the between pool variability observed in the bendiocarb microarray experiment it was decided to increase pool size. Two pools of RNA were combined based on similar concentrations and Bioanalyzer profiles to create single RNA pools comprising of ten female mosquitoes. RNAs were

combined in 1.5ml nuclease free tubes producing a total volume of 60µl. Samples were then evaporated in a SpeedVac (Thermo Scientific) at 35°C until volumes were halved (30µl), which took approximately 15 minutes. The colony material was treated in an identical manner. Combined pools were measured on the Bioanalyzer for quality and re-measured on the NanoDrop™ spectrophotometer (Thermo Scientific) for concentration readings.

An interwoven loop design employing three RNA pools from each experimental group was selected for microarray analysis (Figure 5.1).

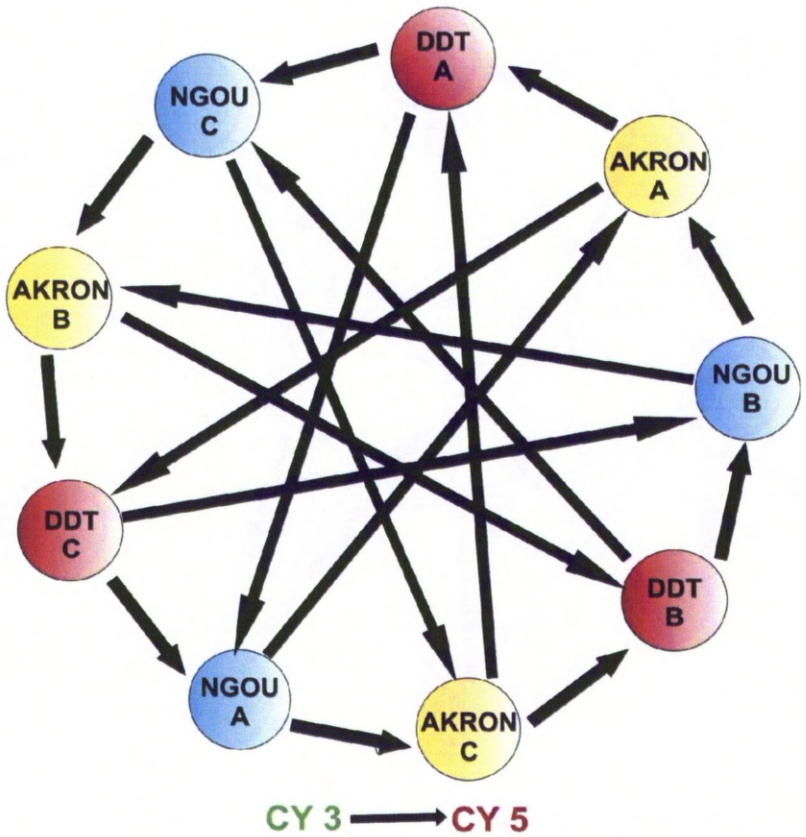

Figure 5.1. Interwoven microarray experimental loop design for a comparison between DDT resistant field collected *An. gambiae s.s.* M forms from Ghana, West Africa and two *An. gambiae s.s.* M form laboratory colonies originating from West Africa.

The N'gousso (NGOU) colony originates from Cameroon and is fully susceptible to DDT while the Akron colony was colonised from Benin and displays low level DDT resistance. The arrows indicate the direction of Cy dye labelling.

5.4.3 RNA amplification and labelling

The pools selected for microarray analysis were re-measured on the NanoDrop™ for concentration on the day of labelling and 100ng in a total volume of 1.5µl selected for microarray labelling. The reduced amount of starting material compared to the bendiocarb experiment (Chapter 3) is due to the use of a *Low Input* Quick Amp

Labelling Kit (Agilent). This kit replaced the original Quick Amp labelling kit (Agilent) and is designed for an RNA sample input range between 25-200ng. Each RNA pool was labelled separately with Cy3 and Cy5 dyes according to manufacturer's protocol version 6 (G4140-90050 Agilent).

After labelling, NanoDrop™ measurements were made from which yield and specific activities were calculated as previously outlined (Chapter 3). All samples passed Agilent recommendations of yield > 825ng and specific activity>6.0 pmol of Cy per µg cRNA. Sample quality was also assessed using the Bioanalyzer.

5.4.4 Microarray hybridisation, scanning and feature extraction

For hybridisation, 300ng of each labelled sample to be compared on a single array, were combined and fragmented according to guidelines. Subsequently 40µl of each hybridisation sample was added to a gasket for the 8x15K platform and the array lowered on top. Array hybridisation, washing, scanning and feature extraction were performed as previously outlined.

5.4.5 Microarray analysis

Microarray normalisation was performed in the R console (2.10.1) using Limma 3.2.3 software and the Java™ based GUI for package MAANOVA employed for analysis of normalised signal intensities as previously described. As the experiment contained three experimental groups, a pair-wise t-test approach was employed to compare groups. Each of the three pair-wise tests was defined separately within the contrast matrix of the J/MAANOVA window and results tables and corresponding volcano plots produced. Within the volcano plots, genes which were over-expressed in the DDT test group in one comparison were highlighted and were identifiable in the second comparison. By this method, genes consistently differentially expressed in the DDT group irrespective of comparison were identified.

5.4.6 Real-Time quantitative PCR validation

A promising gene candidate showing elevated expression in the DDT group with respect to both laboratory colonies was validated through a Real-Time quantitative PCR approach as described in Appendix 22.

5.5 Results

5.5.1 RNA extraction and quality assessment

Individual RNA pools extracted from 5 mosquitoes were assessed quantitatively and qualitatively (Appendix 23). Concentrations post-DNAse treatment ranged between 345.9-680.4ng/µl with N'gousso samples producing more RNA than the other two groups, which may reflect size differences. Subsequently two pools matched by concentration and RNA Bioanalyzer profile were combined to create three pools for each experimental group consisting of RNA from ten females (Appendix 23). Final

pool concentrations were between and 563.1 – 1227.8 ng/µl, with a mean concentration of 892.0 ng/µl [217.2 SD]. Combined pools were re-measured on the NanoDrop and assessed on the Bioanalyzer to check quality (Figure 5.2). Quality was acceptable for all pools and no degradation was endured during the SpeedVac concentration process, which was determined by comparing pre and post combination RNA traces. The DDT resistant Ghanaian samples showed marginally more degradation compared to the laboratory samples (Figure 5.2), which is to be expected considering insecticide exposure, transit from the field and duration in storage.

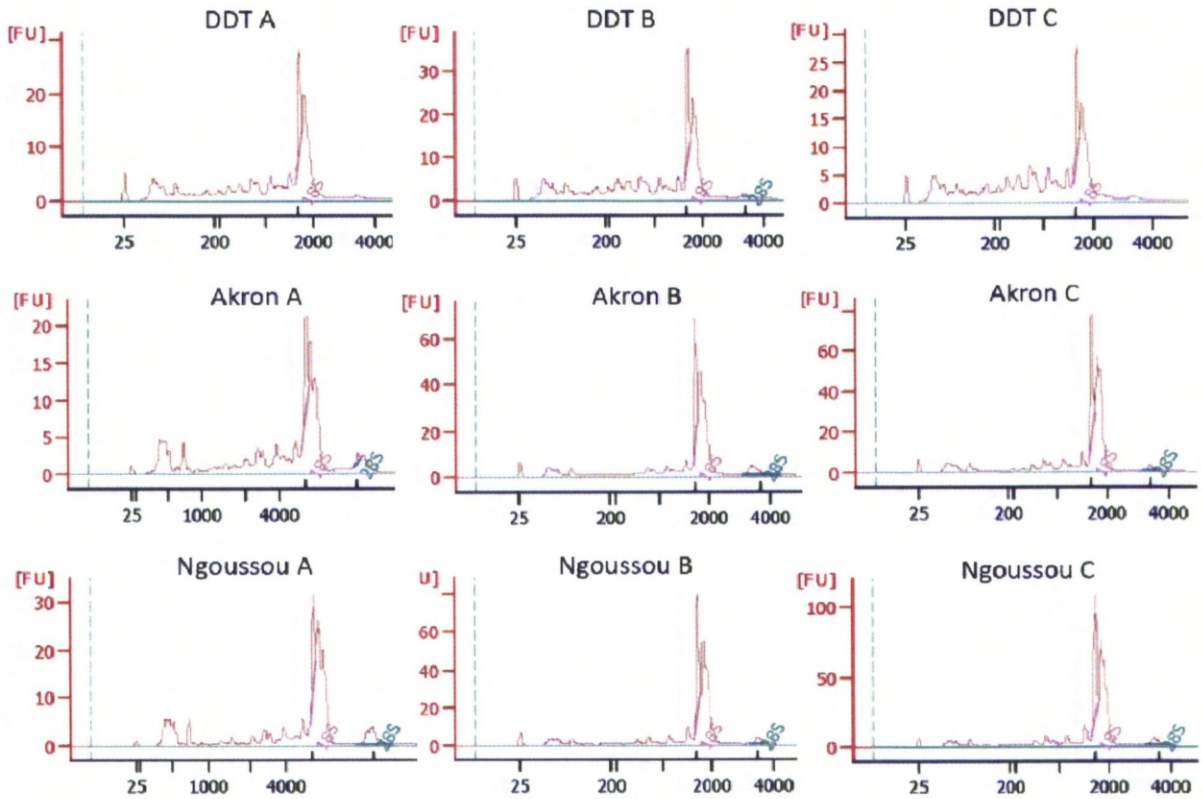

Figure 5.2. Bioanalyzer traces from combined RNA pools extracted from DDT resistant Ghanaian samples (DDT) and West African colony samples (Akron and N'gousso).

Note the different x-axis scale in N'gousso A and Akron A is attributed to inadequate ladder denaturation; however the visual profiles are as expected.

5.5.2 RNA amplification and labelling

Successful labelling and amplification was achieved for all 9 RNA pools indicated by yields > 825ng and specific activities > 6.0 pmol Cy/µg cRNA (Table 5.1). Yields were between 6.16 – 10.24µg (7.45 µg mean [1.04 SD]) with specific activity ranging from 17.58-33.97 which was way above threshold levels. Bioanalyzer analysis revealed an expected distribution of cRNA between 100-2000 nucleotides in all samples. Therefore all pools were employed in microarray hybridisation.

Table 5.1. NanoDrop™ spectrophotometer (Thermo Scientific) readings from Cy3 and Cy5 labelled RNA pools for microarray analysis.

Nine RNA pools were analysed in total consisting each of 10 female *An. gambiae s.s.* exposed for 6 hours to either DDT control papers (AKRON & N'GOUSSO) or 4% DDT papers (TEST). RNA was labelled with cyanine dye using the Low Input Quick Amp Labelling Kit (Agilent). Calculated yield in µg ((ng/µl *30[elution volume])/1000) and specific activity in pmol Cy/µg (([pmol/µl]/[ng/µl])*1000) are also shown.

GROUP	POOL	Cy	260:280	pmol	ng/µl	Yield	Specific activity
TEST	A	3	2.24	4.9	256.7	7.70	19.09
		5	2.27	5.9	243.3	7.30	24.25
	B	3	2.26	4.7	212.5	6.38	22.12
		5	2.30	5.2	226.6	6.80	22.95
	C	3	2.25	5.0	224.4	6.73	22.28
		5	2.28	4.7	205.4	6.16	22.88
AKRON	A	3	2.24	9.4	276.7	8.30	33.97
		5	2.31	5.3	254.9	7.65	20.79
	B	3	2.19	6.9	282.5	8.48	24.42
		5	2.22	3.5	195.0	5.85	17.95
	C	3	2.23	6.5	341.4	10.24	19.04
		5	2.27	7.4	276.0	8.28	26.81
N'GOUSSO	A	3	2.23	8.6	243.2	7.30	35.36
		5	2.27	6.5	268.0	8.04	24.25
	B	3	2.20	4.7	236.8	7.10	19.85
		5	2.24	4.4	236.0	7.08	18.64
	C	3	2.22	5.8	269.4	8.08	21.53
		5	2.25	3.9	221.9	6.66	17.58

5.5.3 Microarray hybridisation, scanning and feature extraction

All 18 arrays were successfully hybridised with scans of good quality and low background. Post feature extraction the QC reports for each array produced scores between 10 and 11, which were well above the threshold (QC 8-9) suggested by Agilent.

5.5.4 Microarray analysis

Normalisation was performed as previously described with an off-set of 50 and LOESS within array normalisation. MA and Cy density plots were produced at each stage to check normalisation procedures were sufficient. Box-plots of dye-ratio in each array showed similar distributions (Figure 5.3) so an additional between array normalisation step was not required.

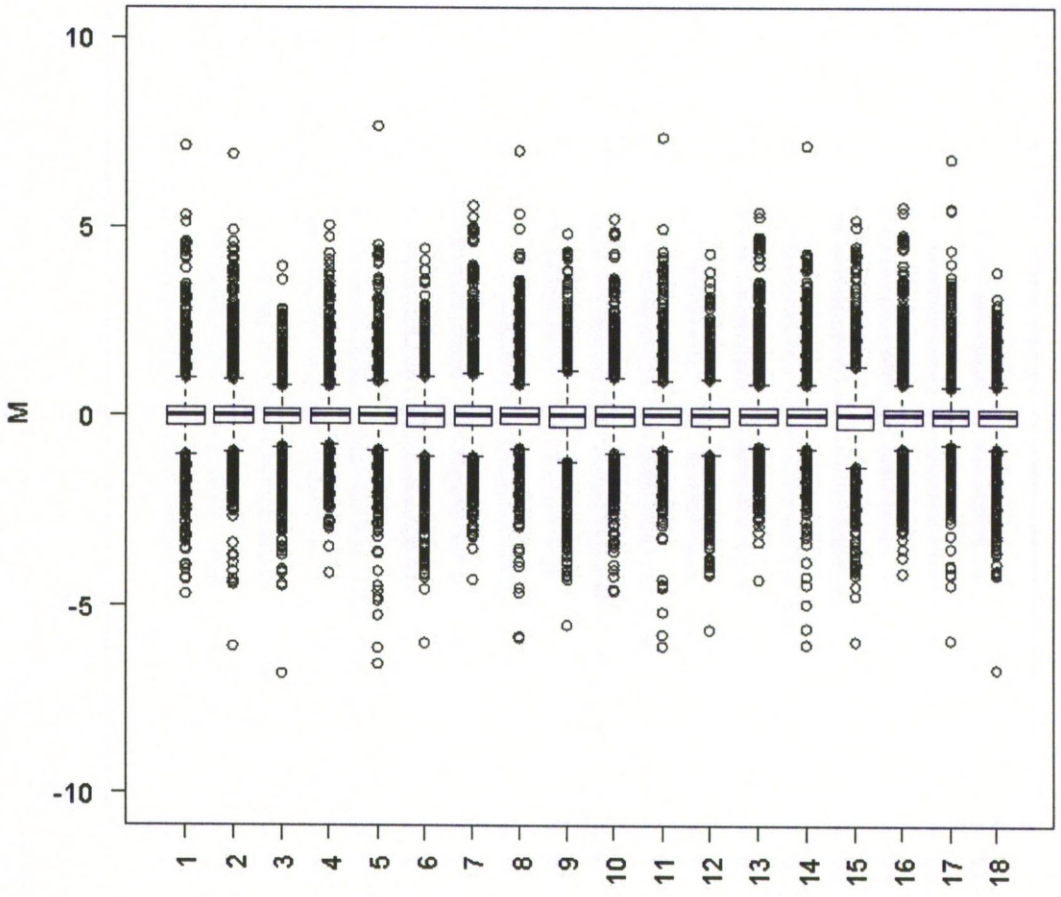

Figure 5.3. Box-plot of 'M' value distribution for each of 18 microarrays employed in an experiment comparing a DDT resistant Ghanaian *An. gambiae s.s.* population with two West African control strains.

The x-axis shows each array (1-18) after normalisation which involved off-setting intensity values by 50 and LOESS within array normalisation. The y axis indicates M values where M= red/green dye intensity ratio. As the distribution of signals were similar in each array (box width and position similar) no between array normalisation was required.

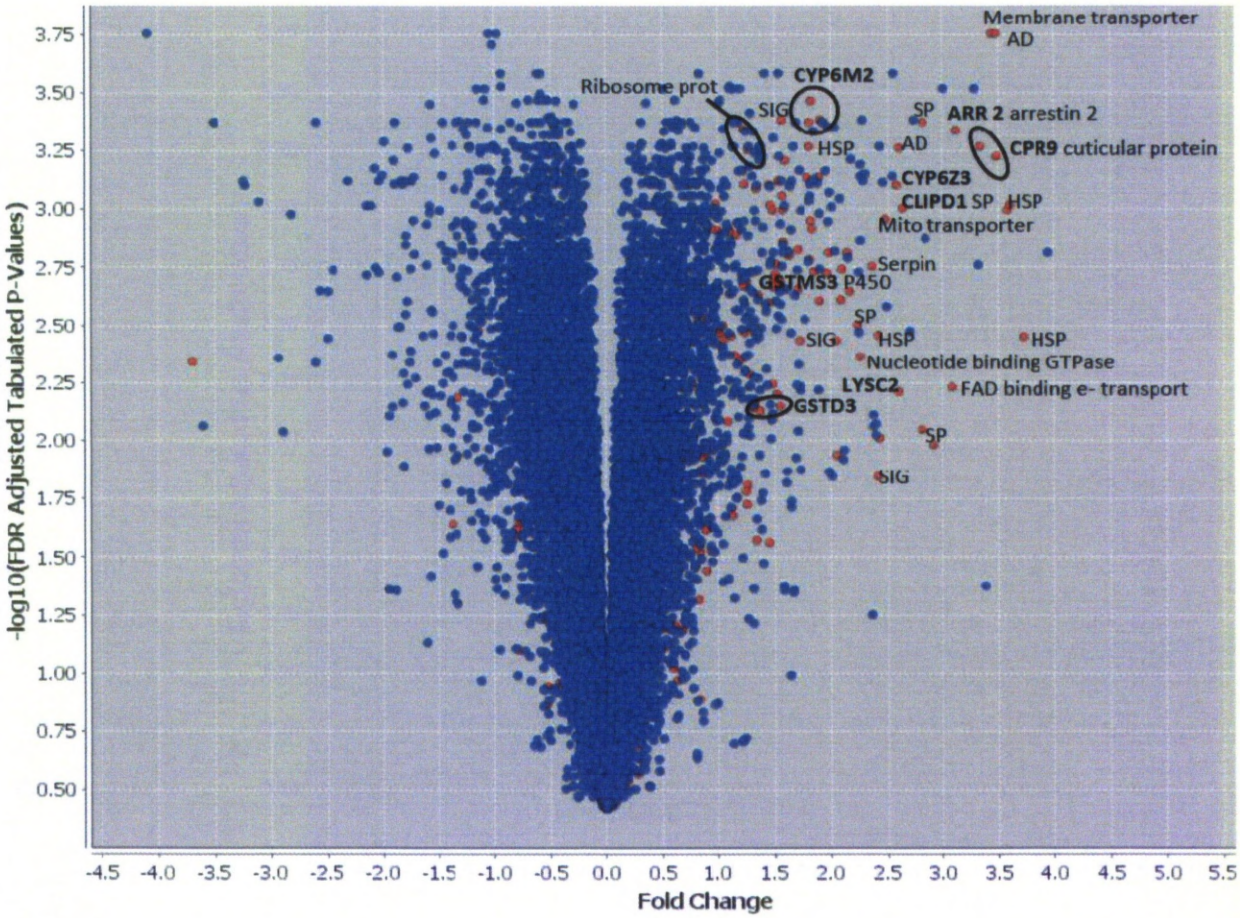

Figure 5.4. Volcano plot showing genes consistently over-expressed in DDT resistant field collected *An. gambiae s.s.* from Ghana (positive values) compared with two West African control colonies.

Comparison between DDT resistant field samples and the N'gousso Cameroonian control is shown with probes highlighted in red also over expressed in the DDT group compared with the Beninese Akron control (>0.95 Log_2 fold over-expressed in the DDT group compared to Akron, adjusted p-value p <0.05 (~1.30 $-log_{10}$ False Discovery Rate adjusted p-value). SP – Serine Peptidase, HSP – Heat Shock Protein, AD – Alcohol Dehydrogenase, SIG – Signal peptide, **bold text** = named gene in Ensembl

5.5.5 Genes over-expressed in DDT resistant Ghanaian samples

A volcano plot displaying log_2 fold ratios and $-log_{10}$ FDR adjusted p-values was created within the J/MAANOVA GUI comparing the DDT resistant Ghanaian group with the Beninese Akron colony. Within this plot probes which were >0.95 Log_2 fold (~2 fold) over-expressed in the DDT resistance group with an adjusted p-value <0.05 (~1.30 – log_{10} FDR adjusted p-value) were selected and defined as a probe group. A volcano plot comparison was then made between the DDT Ghanaian group and Cameroonian (N'gousso) colony and the DDT versus Akron defined probe group highlighted (Figure 5.4).

From these combined volcano plots, gene probes over-expressed ≥ 2 fold (adjusted p<0.05) in the DDT group with respect to both Akron and N'gousso colonies were

identified. Probes for genes including *CYP6M2*, *CYP6Z3*, cuticular protein *CPR9*, arrestin 2 (*ARR2*) and various heat shock proteins, transporter proteins, ribosomal proteins and proteases were all over-expressed in the DDT group (Figure 5.4).

A subset of genes displaying the most significant and consistent over-expression in the DDT Ghanaian group were selected and tabulated (Table 5.2). Probes for these gene had expression ratios >0.95$\log_2$ with a p values < 0.001 (~3.00 –log10 FDR adjusted p) in both comparisons.

The most highly and consistently over-expressed probe in this subset was for gene AGAP010326-RA, which was expressed 10.78 fold higher in the DDT group compared to N'gousso (FDR p=0.000212) and 9.04 higher compared to Akron (FDR p=0.000249). This gene is currently listed as novel in Ensembl and has 'membrane transporter' listed under its Gene Ontology (GO) terms.

For cytochrome P450 *CYP6M2*, four probes were consistently over-expressed in the DDT group in both the Akron and N'gousso comparison; though the whole genome probe CUST_11496 was identical in sequence to a separately designed DETOX probe (DETOX_439_P14). DDT group expression was 5.68-6.68 fold higher compared with Akron (p<0.0005) and 3.51-3.83 higher compared to N'gousso (p<0.001). The lower expression ratio compared to N'gousso was explained by *CYP6M2* showing marginally higher expression in the N'gousso colony compared with Akron (1.66 mean fold change [0.05475 SD], p< 0.01).

In addition, novel gene AGAP003691-RA, with GO term 'serine type endopeptidase' was approximately 7 fold over expressed in the DDT group compared to both colonies (p<0.001) while a novel transcript AGAP0055499-RA with potential 'alcohol dehydrogenase' activity was between 6-8 fold more highly expressed in the DDT group (p<0.001).

Table 5.2. Microarray probes which were significantly (FDR corrected p<0.001) greater that 0.95 Log_2 over-expressed in the Ghanaian DDT resistant group of *An. gambiae s.s.* compared to both Akron and N'gousso laboratory colonies.

FC – fold change, FDR p – False Discovery Rate adjusted p-value post ANOVA analysis. Description indicates gene name if defined in Ensembl or Gene Ontology terms where genes are listed as 'novel' in Ensembl. * indicates two identical probes for gene *CYP6M2*.

Description	Gene Name	Probe ID	DDT vs AK		DDT vs Ngou	
			FC	FDR p	FC	FDR p
Transmembrane signal peptide	AGAP001717-RA	CUST_7558_PI4	2.27	9.43E-04	2.95	5.37E-04
Ribosomal protein constituent	AGAP002364-RA	CUST_4143_PI4	2.50	5.65E-04	2.34	6.41E-04
Ribosomal protein constituent (alternative transcript)	AGAP002364-RB	CUST_4142_PI4	2.58	8.91E-04	2.41	8.77E-04
Serine type endopeptidase	AGAP003691-RA	CUST_5955_PI4	7.13	5.65E-04	7.04	5.72E-04
Alcohol dehydrogenase family	AGAP005499-RA	CUST_933_PI42	7.64	7.41E-04	6.06	8.22E-04
Membrane transporter	AGAP010326-RA	CUST_8992_PI4	9.04	2.49E-04	10.78	2.12E-04
Metalloendopeptidase activity	AGAP013468-RA	CUST_5671_PI4	2.04	3.87E-04	2.35	3.07E-04
CYP6M2	AGAP008212-RA	CUST_11496_PI*	5.88	2.49E-04	3.55	4.54E-04
CYP6M2	CYP6M2	DETOX_439_PI4*	6.68	2.49E-04	3.83	5.72E-04
CYP6M2	CYP6M2	DETOX_441_PI4	6.19	2.49E-04	3.77	5.70E-04

Another CYP6 P450, *CYP6Z3*, was consistently over-expressed in the DDT group in both comparisons (Table 5.6); however as with gene *CYP6M2*, the whole genome probe (CUST_11501_PI) was identical to a DETOX probe (DETOX_490_P14) which was reflected in consistent fold-changes and p-values. *CYP6Z3* was over-expressed 3.24-5.97 fold in the DDT group compared to N'gousso ($p<0.01$) according to all four probes with an overall mean fold change of 4.48 [1.1240 SD]. While, with respect to Akron, the DDT group over-expressed *CYP6Z3* 2.34-3.80 fold ($p<0.05$) when consulting three of the four gene probes. However DETOX probe 492_ P14 produced a fold change marginally in favour of the Akron colony (-1.24 fold change) though the p-value was not significant ($p=0.2060$). This probe also gave the lowest fold change and least significant p-value in the N'gousso comparison (3.24 fold change $p= 0.0081$).

Two transcripts of cuticular protein gene, *CPR9*, were up-regulated in the DDT group with a probe for *CPR9* transcript AGAP002726-RB, 11.14 fold over-expressed compared with N'gousso ($p<0.001$) and 4.5 fold compared with Akron ($p<0.005$). The probe for *CPR9* transcript AGAP002726-RA, was 10.06 fold more highly expressed in the DDT group compared to N'gousso ($p<0.001$) and 4.21 with respect to Akron ($p<0.005$).

Two Glutathione S transferase genes showed elevated expression in the DDT resistant group. Firstly microsomal GST, *GSTMS3*, was over-expressed 3.26 fold [0.1026 SD] compared with the Akron colony ($p<0.005$) and 2.88 fold [0.0898 SD] compared to the N'gousso colony ($p<0.005$) when all four probes were averaged. Secondly a delta class GST, *GSTD3*, showed 3.50 fold [0.3563 SD] over expression compared to Akron ($p<0.01$) and 2.62 fold [0.2101 SD] compared to N'gousso ($p<0.05$). However p-values were less significant compared with *GSTMS3* results.

A number of novel genes not yet named in the Ensembl database were consistently over-expressed in the DDT resistant group compared to controls. These included putative heat shock proteins (HSP), ribosomal constituents, serine proteases, signal peptides and alcohol dehydrogenases. Five HSPs were identified as over-expressed in both the DDT/Akron and DDT/N'gousso comparisons (

Table 5.3), including two genes, AGAP007159 and AGAP007158 which were over-expressed between 11.92-21.06 fold ($P<0.005$).

Table 5.3. Putative heat shock proteins up-regulated in a DDT resistant group of Ghanaian *An. gambiae* s.s. with respect to two West African laboratory colonies (Akron and N'gousso).

Microarray fold change (FC) calculated from an 8x15K whole genome array experiment using an ANOVA based analysis are shown along with false discover rate (FDR) corrected p-values (Benjamini and Hochberg, 1995)

Gene Name	Probe ID	DDT vs Ak		DDT vs Ngou	
		FC	FDR p	FC	FDR p
AGAP007161-RA	CUST_2696_PI4	2.72	1.37E-03	3.51	7.42E-04
AGAP004583-RA	CUST_7222_PI4	3.90	1.15E-02	5.35	4.94E-03
AGAP004581-RA	CUST_7220_PI4	7.53	1.29E-02	13.16	5.01E-03
AGAP007158-RA	CUST_2690_PI4	21.06	1.23E-03	11.92	1.51E-03
AGAP007159-RA	CUST_2691_PI4	21.03	1.15E-03	12.06	1.38E-03

As highlighted in Figure 5.4 a whole genome probe for an Arrestin gene, *ARR2*, was significantly up-regulated in the DDT group in both comparisons. Compared to the N'gousso colony the DDT group over-expressed *ARR2* 8.68 fold ($p<0.001$) while compared with Akron the fold difference was lower (~2 fold) though still significant ($p<0.05$).

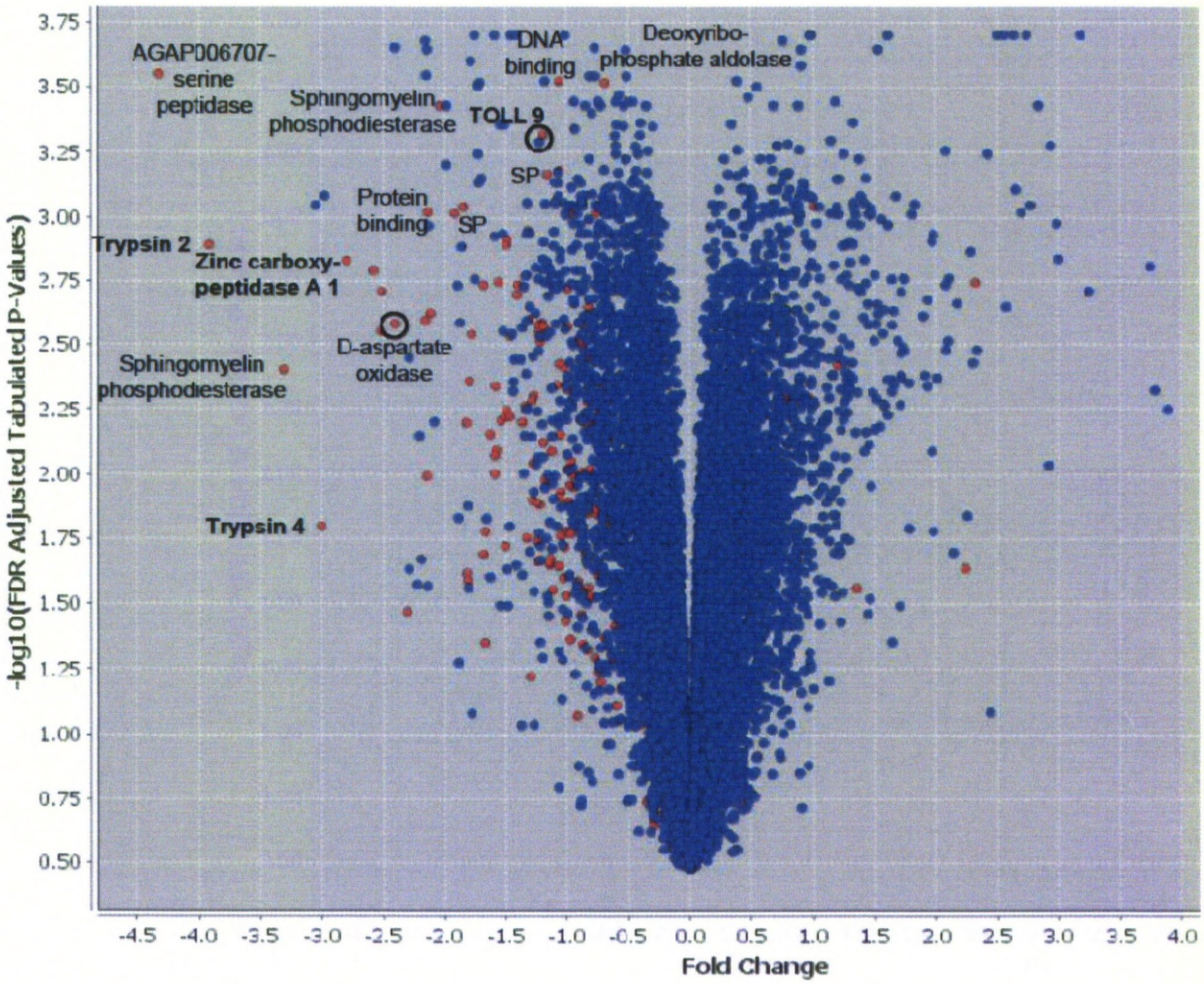

Figure 5.5. Volcano plot showing genes consistently over-expressed in West African control colonies (negative values) compared with field collected DDT resistant *An. gambiae s.s.* from Ghana.

Comparison between DDT resistant field samples and the Akron Beninese control is shown with probes highlighted in red also over expressed in the N'gousso control group compared with the DDT resistant samples (>0.95 Log_2 fold over-expressed in the N'gousso group compared to DDT resistant, adjusted p-value p <0.05 (~1.30 $-log_{10}$ False Discovery Rate adjusted p-value). SP – Serine Peptidase, **bold text** = named gene in Ensembl.

5.5.6 Genes consistently over-expressed in colony samples compared to DDT resistant field samples

A number of gene probes were consistently over-expressed in the colony material (Akron and N'gousso) compared to the DDT resistant field collected samples (Figure 5.5). A subset of the most significant ($P<0.001$) genes which show >0.95 $\log_2$ expression ratios in favour of colony material are shown in Table 5.4.

A transcript of gene AGAP006707 was highly over-expressed in colony material; 20.10 fold compared to Akron and 11.33 compared to N'gousso. This gene is currently defined as novel in Ensembl but has serine peptidase activity listed under GO terms. Other genes which showed greater expression levels in colony samples included *TOLL 9,* a toll-like receptor, and three addition novel genes with putative DNA binding, sphinomyelin phosphodiesterase and peptidase activities. Two trypsin genes (trypsin 2 and trypsin 4) were also more highly expressed in the colony material.

5.5.7 Genes over-expressed in each colony

In the Beninese Akron colony, probes designed for gene *ace-1* encoding acetylcholineesterase were consistently over expressed compared to both DDT resistant Ghanaian samples (3.96-5.30 fold $p<0.001$) and N'gousso Cameroonian colony samples (3.93-5.28 fold $p<0.001$). In addition, proteins with putative functions in ATPase/nucleotide binding, signalling and tRNA synthesis were also up regulated (Appendix 24).

In the N'gousso Cameroonian colony two novel peptides showed consistently high levels of up-regulation compared to the Akron colony and Ghanaian field samples. The probe for transcript AGAP011475-RA was over-expressed 17.10 fold compared with the Ghana group ($p<0.001$) and 16.01 compared with Akron ($p<0.0001$); this transcript was not from an Ensembl defined gene however GO terms suggested metalloendopeptidase activity. Secondly a transcript putatively linked to IgE was up-regulated 11.32 and 14.02 fold compared to Ghanaian and Akron groups respectively ($p<0.001$). Ten additional probes were over-expressed in the N'gousso group at ratios >0.95 Log2 and $p<0.001$ (Appendix 25), these included *CASPS4* a caspase gene and genes with putative involvement in DNA/actin binding, pseudouridine synthesis, splicosome and ribosome complexes and protein signalling.

Table 5.4. Microarray probes which were significantly (FDR corrected p<0.001) greater that 0.95 Log_2 over-expressed in colony mosquitoes (Akron & N'gousso) compared to DDT resistant field samples.

Description provides gene names or where genes are novel GO term information from the Ensembl database (http://metazoa.ensembl.org/Anopheles_gambiae/Info/Index). FC – fold change.

Description	Gene Name	Probe ID	Cf. AK FC	AK FDR p	Cf. Ngou FC	Ngou FDR p
serine peptidase	AGAP006707-RA	CUST_2216_PI4	20.10	3.52E-04	11.33	5.72E-04
DNA binding	AGAP000281-RA	CUST_8629_PI4	2.09	4.30E-04	2.22	3.92E-04
sphingomyelin phosphodiesterase	AGAP004054-RA	CUST_6425_PI4	4.12	5.65E-04	3.71	5.98E-04
serine-type endopeptidase activity	AGAP004700-RA	CUST_32_PI422	2.23	9.28E-04	2.98	4.75E-04
TOLL9 Toll-like Receptor	AGAP006974-RA	CUST_2566_PI4	2.29	7.41E-04	2.34	6.18E-04
'binding'	AGAP007683-RA	CUST_3245_PI4	2.09	9.28E-04	2.03	8.46E-04

5.5.8 Candidate gene validation

5.5.8.1 *CYP6M2* Real-Time quantitative PCR

The P450 gene *CYP6M2* was pursued for expression validation using the Real-Time quantitative PCR approach described elsewhere (Appendix 22). Sequence data for this P450 gene from Ghanaian field collected mosquitoes was consulted during primer design to ensure conserved regions were selected (primer sequences Appendix 26). However, it was not possible to design intron spanning primers due to high levels of polymorphism characteristic of cytochrome P450 genes (Wilding *et al.*, 2009).

A plasmid containing the coding region of *An. gambiae s.s. CYP6M2* was provided courtesy of M. Paine and the ribosomal S7 plasmid (Appendix 22) was again used for normalization.

For the *CYP6M2* plasmid standard curve a linear relationship was recorded between concentration and C_t value with R^2= 0.9998 (Appendix 27). PCR amplification efficiency was 104%, within the recommended range (90-105%). Mean normalised *CYP6M2* copy number in the Ghanaian DDT resistant RNA pools was 1.72E-01 [0.0050 SD] which was higher than both Akron (7.30E-02 [0.0074 SD]) and N'gousso (1.13E-01 [0.0057 SD]).

Higher *CYP6M2* normalised expression was recorded in the Ghanaian field samples compared to both Akron and N'gousso, with 2.36 fold greater expression compared to Akron and 1.54 fold compared to the Cameroonian N'gousso colony (Table 5.5).

However, the expression ratios recorded via microarray were over double those calculated from Real-Time qPCR results (Table 5.5, column 8).

Table 5.5. Real-time quantitative PCR results for gene *CYP6M2* for sample pools employed in a microarray experiment comparing DDT resistant Ghanaian *An. gambiae s.s.* field samples with two West African laboratory colonies.

DDT A-C = DDT resistant Ghanaian cDNA pools, AKRON A-C = Akron colony controls, N'GOUSSO A-C = N'gousso colony controls. Absolute expression levels are calculated from a plasmid standard curve (Appendix 26) and normalised against ribosomal S7 copy number. A group mean was taken of normalised *CYP6M2* copy numbers and compared against other groups to calculate test/control ratios; these are listed along with mean ratios calculated from microarray probes (microarray FC).

Target	C(t) Mean	C(t) Std. Dev	Copy no. 1pg= 1.20E+05 copies	S7 copy no.	*CYP6M2* copy no./S7 copy no.	Group Mean	Std. Dev
DDT A	23.70	0.2051	7.27E+04	4.37E+05	1.66E-01		
DDT B	22.87	0.0340	8.68E+04	4.95E+05	1.75E-01		
DDT C	23.17	0.0288	8.14E+04	4.66E+05	1.75E-01	1.72E-01	0.0050
AKRON A	26.15	0.1896	4.28E+04	5.26E+05	8.14E-02		
AKRON B	26.98	0.1344	3.58E+04	5.12E+05	6.98E-02		
AKRON C	26.60	0.0853	3.88E+04	5.74E+05	6.77E-02	7.30E-02	0.0074
N'GOUSSO A	24.01	0.3354	6.79E+04	5.82E+05	1.17E-01		
N'GOUSSO B	24.21	0.1265	6.50E+04	5.67E+05	1.15E-01		
N'GOUSSO C	24.25	0.0052	6.44E+04	6.07E+05	1.06E-01	1.13E-01	0.0057

	Test/ Control	FC	microarray FC [SD]
	DDT/AKRON	2.36	6.11 [0.44]
	DDT/N'GOUSSO	1.53	3.67 [0.16]
	N'GOUSSO/AKRON	1.54	1.66 [0.05]

5.6 Discussion

5.6.1 *CYP6M2*

Members of the cytochrome P450 family have been repeatedly linked to resistance to insecticides such as DDT (Chiu *et al.*, 2008; Daborn *et al.*, 2002) and pyrethroids (Muller *et al.*, 2008b; Yang *et al.*, 2006), with the CYP6 family implicated in the majority of cases. The DDT resistant M-form Ghanaian group displayed elevated expression of *CYP6M2* compared to both control groups in the microarray experiment.

Real-Time qPCR also supported microarray findings with normalised expression ratios in favour of the DDT resistant field samples in both colony comparisons; however expression ratios were not equivalent between methods (Table 5.5) with greater fold-change recorded via microarray.

A number of factors can effect correlation between microarray and qPCR data, and have been addressed elsewhere (Chapter 4, Appendix 22). P450 genes are also known to be highly polymorphic (Wilding *et al.*, 2009) and SNP differences may cause loss of signal in both microarray and qPCR experiments. Due to time constraints sequencing of this gene in groups compared was not possible. However sequence data from Ghanaian *An. gambiae s.s.* collected ~ 40km from the field site (courtesy of P. Muller) were consulted for the design of Real-Time qPCR primers and aligned with Ensembl and East African colony sequences allowing conserved regions to be selected.

P450 gene families often contain genes with high sequence similarity; *CYP6M2* shares ~72% identity with *CYP6M4*. Cross signal from closely related transcripts may confound expression results. The forward Real-Time qPCR primer for *CYP6M2* shared 21/22 nucleotides with *CYP6M4* while the reverse primer shared 12/21; however due to sequence polymorphisms there was little option for moving these primers. Spurious *CYP6M4* amplification may have masked *CYP6M2* signal, with M4 expressed at similar levels across groups. Though this is speculative at present and extensive sequencing of qPCR amplified product is required to confirm this hypothesis.

A number of studies have reported over-expression of *CYP6M2* in resistant *An. gambiae s.s.* populations. It was over-expressed in a permethrin resistant S form colony originating from Odumasy in Ghana when compared to a susceptible lab colony using a candidate gene microarray approach (Muller *et al.*, 2007). The use of an insecticide selected laboratory colony and a susceptible comparison colony originating from East Africa may have confounded these results. However a later microarray study (Muller *et al.*, 2008b) reported *CYP6M2* over-expression in pyrethroid selected S form females from field sites in the same region of Ghana. These mosquitoes had survived permethrin exposure at a predetermined LT_{50} and were compared to non-exposed controls from the same area.

Elevated *CYP6M2* levels were reported in other West African *An. gambiae* populations, again linked to a pyrethroid resistant phenotype. Djouaka *et al.* (2008) reported on

three WHO defined permethrin resistant M forms populations from Benin and Nigeria where 2.4 to 2.7 fold *CYP6M2* over-expression was recorded via microarray compared with a susceptible Nigerian population (FDR $p<0.005$). The authors also suggest that recombinant *CYP6M2* demonstrated metabolic activity with pyrethroids in preliminary tests *in vitro* ("*Stevenson, personal commun.*").

Finally in a study unlinked to insecticide resistance, elevated *CYP6M2* was reported in an M form colony of *An. gambiae* originating from Cameroon on infection with *Plasmodium berghei* (Felix *et al.*, 2010). In this microarray study *CYP6M2* was over-expressed 4.23 fold in the mid-gut and 2.73 fold in the fat body isolated from *Plasmodium* infected insects compared to non-infected controls on day one of infection. The authors hypothesised a broad role for *CYP6M2* in detoxification in light of previous association with resistance.

Elevated expression of *CYP6M2* has been reported in both M and S molecular forms of *An. gambiae s.s.* and in both males and females with the only link between these studies being the West African origin of mosquitoes employed. Although it is unlikely that over-expression of this P450 is an artefact of West African populations as elevated expression has been reported when comparing mosquitoes with different resistance phenotypes originating from the same area (Djouaka *et al.*, 2008; Muller *et al.*, 2008b). In three studies a link to permethrin resistance is speculated as elevated expression is reported in insects surviving permethrin exposure, with one study citing preliminary evidence of permethrin, deltamethrin and cypermethrin metabolism by recombinant *CYP6M2* (Djouaka *et al.*, 2008). It is possible that elevated *CYP6M2* expression in the DDT resistant Ghanaian population reported here may be linked to pyrethroid resistance as preliminary bioassays suggested both permethrin and deltamethrin (WHO defined) resistance in two areas of the Accra region sampled (Appendix 8). However it is also possible that the other West African populations exhibiting *CYP6M2* elevated expression are also resistant to DDT, with *CYP6M2* expression related to resistant to this organochlorine rather than pyrethroid compounds. DDT resistance has been previously reported in Ghana (Coetzee *et al.*, 2006), Benin (Yadouleton *et al.*, 2010) and Nigeria (Awolola *et al.*, 2005). Alternatively *CYP6M2* may have a role in metabolism of both DDT and pyrethroids or perhaps performs a more generalist function in xenobiotic detoxification; a hypothesis supported by the work of Felix *et al.* (2010) in which up-regulation was associated with parasite infection.

In order to establish the role of *CYP6M2* in these detoxification processes, recombinant protein must be screened with an array of compounds to elucidate potential metabolism. Sequence differences between mosquito strains and resistance phenotypes should also be considered as a role for allelic variation in altering P450 metabolic activity is supported by previous work (Amichot *et al.*, 2004; Mao *et al.*, 2007).

5.6.2 *CYP6Z3*

CYP6Z3 was over-expressed in the DDT resistant field samples compared to both colony comparisons. A mean expression ratio of 4.48 ($p<0.01$) was recorded compared with N'gousso when all four probes were consulted, and compared to the Akron colony, a 2.85 ($p<0.05$) ratio was calculated from three probes.

CYP6Z3 shares high sequence identity with closely related P450s *CYP6Z2* (93.8% ID) and *CYP6Z1* (71.4%ID) which, as speculated for *CYP6M2* and *M4*, is likely to have confounded results. *CYP6Z3* microarray probes shared high sequence identity with *CYP6Z2* (Table 5.6, column 2), with DETOX probe, _492_PI4, showing greatest similarity (87%ID to *CYP6Z2*). This probe also produced the lowest fold changes and least significant p-values in both comparisons (Table 5.6). When *CYP6Z2* microarray results were consulted (Appendix 28), this gene was found to be more highly expressed in the Akron colony compared to N'gousso. While two out of the four Z2 probes showed elevated expression in the Akron colony with respect to the DDT group. The cross signal from *CYP6Z2* in the Akron colony is likely to have affected the results for *CYP6Z3* probe DETOX_492_P14; this probe was therefore removed from overall fold change calculations.

Table 5.6. *CYP6Z3* microarray probes results from an 8x15K whole genome array experiment comparing three groups of *An. gambiae* s.s mosquitoes.

A DDT resistant Ghanaian field strain (DDT) was compared to two West African laboratory colonies (AKRON (Ak) and N'GOUSSO (Ngo)) which were also compared with each other in a looped microarray experimental design. Percentage identity (% ID) for microarray probes to closely related P450 genes, *CYP6Z2* (Z2) and *CYP6Z1* (Z1) are shown along with microarray fold change (FC) and false discovery rate (FDR) corrected p-values (Benjamini and Hochberg, 1995)

CYP6Z3 probes	% ID to Z2	% ID to Z1	DDT vs Ak		DDT vs Ngou		Ak v Ngo	
			FC	FDR p	FC	FDR p	FC	FDR p
CUST_11501_PI*	70	47	2.34	1.48E-02	4.29	2.65E-03	1.83	4.06E-02
DETOX_491_PI4	80	47	3.80	3.13E-03	5.97	1.18E-03	1.57	5.96E-02
DETOX_490_PI4*	70	47	2.39	1.25E-02	4.42	2.28E-03	1.85	3.58E-02
DETOX_492_PI4	87	57	-1.24	2.06E-01	3.24	8.09E-03	4.02	6.81E-03

Elevated expression of *CYP6Z3* has been previously reported in a permethrin resistant Ghanaian colony of *An. gambiae s.s.* (Muller *et al.*, 2007). Three fold over-expression

was recorded when males of resistant Odumasy strain were compared to males of the susceptible Kisumu strain ($p=1.54x10^{-07}$). However significant over-expression ($p<0.001$) was not reported in females. *CYP6Z3* expression levels are also reportedly very low in adult mosquitoes, being undetectable via Real-Time qPCR (Nikou *et al.*, 2003). The authors suggest differential expression in males may reflect 'noise' when comparing genes of low level expression as a female-male comparison showed no differential expression. However from microarray results presented here, the raw signal intensities for *CYP6Z3* were in the upper 50% of ranked intensities and were equivalent to those recorded for P450 *CYP6M2*. This suggests expression is easily detectable via microarray in adults, and *CYP6Z3* over-expression in the DDT resistant group is not a spurious result.

CYP6Z3 is one of three *CYP6Z* genes in *An. gambiae* situated in a cluster of P450s on chromosome arm 3R. This P450 cluster is within the boundaries of a major quantitative trait locus associated with pyrethroid resistance (Ranson *et al.*, 2004). Closely related P450 *CYP6Z2* (~94%ID) is situated within this cluster and was also reportedly over-expressed in the study by Muller *et al.* (2007), with 4.2 fold over-expression in permethrin resistant females ($p=8.6x10^{-07}$) and 8 fold in males ($2.27x10^{-11}$). However as speculated in this study, potential cross hybridisation between probes for *CYP6Z2* and *Z3* may have affected results.

Subsequently recombinant *CYP6Z2* was found to bind, though not metabolise two pyrethroid compounds (permethrin and cypermethrin) (McLaughlin *et al.*, 2008) Although the lack of metabolism may have been attributed to an absence of cytochrome b5, which is necessary to drive some P450 metabolism (Schenkman and Jansson, 2003) such as *CYP6D1* pyrethroid breakdown in the *Musca domestica* (Zhang and Scott, 1996). *CYP6Z2* did however display broad substrate specificity with a preference for plant toxins, which in combination with peak expression during larval stages (Nikou *et al.*, 2003) lead the authors to hypothesise an important role in chemo-protection during aquatic stages of the mosquito (McLaughlin *et al.*, 2008).

Elevated expression of *CYP6Z1* (which shares 69.8% ID with Z3) was reported in a permethrin resistant colony compared to a susceptible control via microarray and Real-Time qPCR (David *et al.*, 2005; Nikou *et al.*, 2003) and was also up-regulated in the DDT resistant ZAN/U strain of *An. gambiae* (David *et al.*, 2005). A later study found recombinant *An. gambiae CYP6Z1* capable of metabolising DDT (Chiu *et al.*, 2008). However recombinant *CYP6Z2* was unable to break-down this insecticide, which was attributed to a smaller binding site and the presence of three residues which potentially blocked DDT binding. *CYP6Z1* was also over-expressed in the DDT resistant Ghanaian group compared to controls in this microarray study, however expression ratios and p-values were lower than for *CYP6Z3*. Compared with Akron *CYP6Z1* mean over expression was 1.84 fold ([0.0220 SD] $p<0.01$) and 1.55 fold ([0.0107SD] $p<0.05$) compared with N'gousso. This may support a role for *CYP6Z1* in conferring high levels of DDT resistance in the Ghanaian samples via metabolism.

The ability of *CYP6Z3* to metabolise DDT or pyrethroids has not been tested to date. *CYP6Z3* does share high sequence identity with *CYP6Z2* and contains 2 of the 3

residues highlighted by Chiu *et al.* (2008) which potentially blocked DDT binding. However the authors emphasize that only a small number of amino-acids in the substrate binding site define substrate preference, with amino acid differences between *CYP6Z3* and *Z2* in this region. For a definitive answer the metabolic profile of recombinant *CYP6Z3* must be investigated.

5.6.3 *ARR2* arrestin

A probe for gene *ARR2*, encoding Arrestin 2, was over-expressed in the DDT resistant Ghanaian samples. Arrestins are a major class of soluble proteins which together with G protein–coupled receptor kinases (GRKs) stop or arrest intercellular signalling via desensitisation and modulation of G-protein-coupled receptors (GPCRs) (Krupnick and Benovic, 1998). Arrestins may also have a role in the internalisation and recycling of GPCRs post agonist exposure which re-sensitises receptors following stimulation (Goodman *et al.*, 1996; Laporte *et al.*, 2000). The arrestin proteins are grouped into those having visual and non-visual roles, in which proteins associated with vision are restricted to photoreceptor cells in retinal tissues modulating rhodopsin signalling (Krupnick and Benovic, 1998). In *Drosophila* photoreceptor response, amplitude and termination was found to be associated with arrestin concentration (Belušič *et al.*, 2010). While non-visual arrestins interact with a broad spectrum of GPCRs including those involved in olfactory signalling.

However insect 'visual arrestin' genes have been found to be expressed in the antennae with involvement in olfactory signalling regulation. This suggests some insect arrestins do not conform to visual/non-visual classification and are in fact broadly utilised (Merrill *et al.*, 2002). Merrill *et al.* (2003) investigated the role of three arrestins in *An. gambiae* including *ARR2* which shows similarity to *Drosophila* gene *DMARR2* associated with olfactory as well as visual signalling (Belušič *et al.*, 2010). *ARR 2* was found to be expressed in the head, antennae and body carcasses of both adult males and females, however expression was notably higher in male antennae and in female body carcasses. Enhanced levels of *ARR2* transcript detected in male antennae lead authors to postulate a preferential role in male olfactory behaviours such as pheromone detection (Merrill *et al.*, 2003). Although the expression of *ARR2* in the body of female mosquitoes may be linked to olfactory and gustatory neurones on mosquito tarsi or alternatively non-sensory GPCR signalling (Merrill *et al.*, 2003).

Supported by tissue specific profiling and links to olfaction, a speculative role for *ARR2* in DDT resistant mosquitoes may be an olfactory/sensory response to DDT volatiles detected in mosquito tarsi or antennae. The bendiocarb microarray and subsequent proteomic assays suggested a putative role for odorant binding protein 3 (*OBP3*) in carbamate resistance through interaction with bendiocarb (Chapter 4). OBPs are known to be involved in the activation of odorant receptors of which ~ 79 have been identified as GPCRs (Hill *et al.*, 2002). Although no OBPs were specifically identified as >0.95 $\log_2$ over-expressed in the DDT group, three OBPs (*OBP4*, 20 and a 'novel' OBP) were between 1.4-1.6 fold over-expressed compared to both laboratory colonies ($p<0.05$). It is also likely that a number of OBPs have yet to be annotated as odorant binding proteins in Ensembl, and so expression ratios could not be identified from

results. There is potential for arrestins to be involved in modulating odorant GPCR response on stimulation with insecticidal compounds, which may reach receptors via OBP mediated transport. In addition a number of novel receptors and signalling proteins were identified as 'top-hits' in the DDT resistant group (Table 5.2) which may also be implicated in chemosensory signalling pathways.

5.6.4 *CPR9* cuticular protein

In common with the bendiocarb resistance microarray experiment (Chapter 3) the cuticular protein group was represented in the list of genes putatively over-expressed in the DDT resistant samples. Cuticular protein *CPR 9* was 11.14 over-expressed with respect to N'gousso ($p<0.001$) and 4.50 fold compared to Akron ($p<0.01$). This gene is part of a major family of structural proteins known as RR after Rebers and Riddiford (1988) who identified a consensus sequence shared by these proteins. Three forms of this consensus have been identified (RR-1, RR-2 and RR-3) with *CPR9* in the RR-1 group (He *et al.*, 2007). The RR-1 genes have been associated with proteins in the flexible part of the cuticle rather than more rigid sections (Willis *et al.*, 2005). Cornman *et al.* (2008) investigated the RR gene family in *An. gambiae* and compared the rate of non-synonymous nucleotide mutations with synonymous mutations (Ka/Ks) as a measure of natural selection. They found that the RR-1 class had a higher evolutionary rate compared with RR-2 with more genes displaying Ka/Ks>1 indicative of positive selection (higher rate of amino acid substation then expected under neutrality). While RR-1 gene *CPR-9* displayed one of the highest mean pair-wise Ka/Ks values recorded.

The potential involvement of cuticular genes in the resistance phenotype has been discussed previously (Chapter 3 & 4) with evidence of a direct involvement still illusive. However signs of positive selection within a cuticular gene up regulated in a DDT resistant population may support an advantageous role for *CPR9* in the insecticide resistance phenotype. Although potential confounders of cuticular gene expression such as mosquito size, geographical origins and environmental exposures must also be considered. A technique such as RNA interference could be employed to sequentially knock-down cuticular proteins putatively involved in resistance to establish whether insecticide susceptibility is increased.

5.6.5 Novel membrane transporter

A putative membrane transporter was the most significantly and highly over-expressed gene in the DDT resistant group (Table 5.2). The probe for transcript AGAP010326-RA was between 9.04-10.78 fold over expressed with p-values < 0.0005. In the Ensembl database GO terms listed for this transcript included 'transporter activity' and 'integral to the membrane'. Interestingly this gene was also one of the 'top hits' in a separate microarray experiment also comparing DDT resistant mosquitoes with non-exposed controls (C. Jones pers. comm.). The same 8x15K microarray was used in this later study to compare *An. arabiensis* which had survived 1 hours exposure to 4% DDT with

non-exposed controls from the same area; 3.2 fold over-expression of AGAP010326-RA was recorded.

From consultation of related proteins in Ensembl, two aquaporin protein orthologues were found in *Aedes aegypti* (44% ID) and *Pediculus humanis* (35% ID). Aquaporins are membrane proteins involved in the movement of water and other solutes across biological membranes, and in insects are often associated with the excretory/osmoregulatory Malpighian tubules system (Kaufmann *et al.*, 2005; Spring *et al.*, 2009). It is feasible that up-regulation of this putative aquaporin in the DDT resistant samples is related to efflux/excretion of toxins derived from DDT itself or endogenous products from a stress response, such as reactive oxygen species, induced by insecticide exposure. The Malpighian tubule system is involved in the removal of toxins from the haemolymph, which is believed to be the route of DDT transport *in vivo* (Winter *et al.*, 1975). However DDT is highly hydrophobic meaning a water based excretion mechanism is unlikely to be the major route of efflux unless protection is provided by some form of transporter protein. Such transporters are speculated to have a role in the movement of DDT in the aqueous haemolymph of the cockroach with haemolymph lipoproteins able to bind DDT *in vitro* (Winter *et al.*, 1975).

The potential for this putative transporter to have a role outside of water movement should also be considered as the highest percentage identity to an aquaporin was still below 50%. To investigate the role of this protein further it will be important to determine where this protein is expressed *in vivo* which could involve techniques such as *in situ* hybridisation employing complementary RNA or immuno-staining of tissue preparations with specific antibody. While expression systems such as *Xenopus* (Warmke *et al.*, 1997) could be employed in which membrane bound protein can be expressed and transport of solutes across the membrane assessed. A *Xenopus* system was successfully employed by Duchesne *et al.* (2003) to study solute transport by an aquaporin from *Ae. aegypti*.

5.6.6 Glutathione-S-transferases

Glutathione-S-transferases are detoxifying enzymes which have been associated with resistance to insecticides including DDT (Lumjuan *et al.*, 2005; Ranson *et al.*, 2001), pyrethroids (Vontas *et al.*, 2001) and organophosphates (Huang *et al.*, 1998; Wei *et al.*, 2001). In insects there are two main GST groups classified according to location within the cell. Firstly, cytosolic GSTs which are soluble proteins with at least six classes in insects, and secondly, the membrane bound microsomal GSTs comprising of a single class in insects, which in *An. gambiae* has three members (Ranson *et al.*, 2005).

A microsomal GST, *GSTMS3* (named GSTMIC3 in Ensembl) was approximately 3 fold over-expressed in the DDT resistant group ($p<0.005$). Microsomal GSTs are believed to catalyse a similar range of reactions compared to cytosolic GSTs (Jakobsson *et al.*, 1997; Toba and Aigaki, 2000); however to date have not been implicated in insecticide resistance.

The second GST which was over-expressed in the DDT resistant mosquitoes was a delta class cytosolic protein, *GSTD3*, which was between 2.62-3.50 fold over-

expressed compared to colony insects ($p<0.05$). The delta class of GSTs is the largest class in insects and is one of two classes, along with Epsilon GSTs, which are specific to insects (Enayati *et al.*, 2005). In *An. gambiae* there are 12 Delta class GSTs which are situated in two closely linked clusters on chromosome 2R, divisions 18B and 19D (Ding *et al.*, 2003). These clusters are thought to be a result of segmental duplication (Ding *et al.*, 2003; Holt *et al.*, 2002). *GSTD3* is situated in the division 19 cluster along with *GSTD4, D6, D10, D11* and *D12*, and shares highest sequence identity with *GSTD4* (69.2% ID) and *GSTD10* (62.8% ID). However neither D4 nor D10 GSTs were significantly up-regulated in the DDT resistant group which suggests cross hybridisation of microarray probes was not an issue.

The delta class GSTs have expanded independently in different insect families, which may suggest a role in adaptation to specific ecological niches via detoxification of environmental xenobiotics (Ranson *et al.*, 2005). Insecticides may be considered environmental xenobionts with delta class GSTs previously linked to pyrethroid (Vontas *et al.*, 2001), organophosphate (Wang *et al.*, 1991) and DDT (Ranson *et al.*, 1997; Tang and Tu, 1994) resistance. *GSTD1* expression in drosophila was shown to be elevated in a DDT resistant strain and following recombinant protein expression, *GSTD1* exhibited DDTase activity (Tang and Tu, 1994). Two Delta class GSTs from a DDT resistant strain of *An. gambiae*, *GSTD1* splice variants *GSTD1_5* and D1_6, also exhibited limited DDTase activity in a recombinant protein system (Ranson *et al.*, 1997). However this activity only accounted for 6% of DDT metabolism in this strain with an epsilon class GST, *GSTE2*, subsequently found to be responsible for the majority of DDTase activity (Ranson *et al.*, 2001).

In the absence of elevated *GSTE2* expression, it is feasible that *GSTD3* contributes to the high levels of DDT resistance recorded in this Ghanaian field population. Alternatively this Delta class GST may have a more general role in defence against oxidative stress potentially induced by insecticide exposure. For example, a delta class GST, over-expressed in a pyrethroid resistant strain of plant hopper *Nilaparvata lungens*, showed high peroxidise activity when expressed as a recombinant protein (Vontas *et al.*, 2001; Vontas *et al.*, 2002). Leading the authors to suggest a role for this GST in protecting against pyrethroid induced lipid peroxidation.

5.6.7 Heat shock proteins

Heat shock proteins (HSPs) are molecular chaperones which enable cells to cope with stresses induced by the denaturation of other peptides; thus HSP expression can be associated with resistance to stress (for an extensive review of HSP see Feder and Hofmann (1999)). HSP expression in mosquitoes has been linked to the immune response on challenge with bacteria (Oduol *et al.*, 2000), *Plasmodium* (Vlachou *et al.*, 2005) and O'nyong-nyong virus in *An. gambiae* (Sim *et al.*, 2005), as well as dehydration tolerance in *Ae. aegypti* (Benoit *et al.*, 2010). An indirect link between HSP expression and insecticide resistance in mosquitoes was made by Patil *et al.* (1996) where a 40-50% increase in propoxur (carbamate) tolerance was witnessed in *An. stephensi* and *Ae. Aegypti* after exposure to 'HSP inducing' sub-lethal temperatures when compared to insects unexposed to elevated temperatures. While

in a reciprocal experiment propoxur exposure was seen to increase heat tolerance. However no direct measurement of HSP expression was made and the data reported lacked error measurements. Yoshimi *et al.* (2002) reported induction of *HSP70* in the aquatic midge *Chironomus yoshimatsui* on exposure to insecticides ethofenprox and fenitrothion compared to non-exposed controls, while midges collected from field sites post insecticide spraying showed higher *HSP70* expression than those collected three months after cessation of spraying.

DDT exposure in the Ghanaian group may have induced expression of HSPs with a number of putative HSPs over-expressed in the DDT resistant mosquitoes compared with the controls. Five putative HSPs were identified as over-expressed between 2.72 and 21.06 fold in the DDT group compared to the lab colonies (p-values between $p<0.05$-$p<0.001$) (

Table 5.3). With two potentially closely related genes AGAP007158 and AGAP007159 showing consistently high expression ratios (11.92-21.06 fold) in both comparisons ($p<0.005$). These two genes are situated within ~1.5Kb of each other on chromosome 2L, however as these genes share high sequence identity (~98%) cross probe hybridisation was likely. HSPs could also be constitutively expressed at higher levels in the DDT resistant group which may be an adaptive advantage to insecticide induced stresses. However the environmental exposures and stresses in the field compared to the laboratory cannot be ruled out as a cause of elevated HSP expression in these field collected mosquitoes.

There is potential for a link between either induced or constitutive expression of HSPs and exposure to insecticides which is partially supported by preliminary findings (Patil *et al.*, 1996; Yoshimi *et al.*, 2002). However further investigation into the possible role of HSPs in the resistance phenotype is required and the current poor annotation of putative HSP genes in Ensembl must be addressed to facilitate this process.

5.6.8 Genes over-expressed in colony material

In total, six genes were consistently ≥2 fold over-expressed in the colony control material compared with the DDT resistant field samples with p-values <0.001 (Table 5.4). Of these six only one gene, *TOLL9*, is named in Ensembl with the rest listed as novel genes with corresponding GO terms supplied. *TOLL9* is a toll-like receptor gene which was ~ 2.3 fold over-expressed in the colony material. Toll-like receptors in insects are associated with the innate immune response to bacterial and fungal challenge (Bilak *et al.*, 2003; Hoffmann and Reichhart, 2002; Michel *et al.*, 2001). Luna *et al.* (2002) characterised four of ~11 Toll related genes in *An. gambiae* (Christophides *et al.*, 2002); *AgTOLL*, *AgTOLL6*, *AgTREX*, and *AgTOLL9* that are orthologous to *Drosophila TOLL* genes. *TOLL9* was found to be expressed during all developmental stages with elevated levels reported in the gut of adult mosquitoes. On larval bacterial challenge (*E.coli*) *AgTOLL9* was also found to be 'weakly' up-regulated. The authors speculate a roll for *AgTOLL9* in anti-*Plasmodium* responses due to high levels of expression in the adult gut and the gene location within a locus potentially associated with *Plasmodium* melanisation (Crews-Oyen *et al.*, 1993).

The potential for *Plasmodium* infection to induce expression is not conceivable in this case as females selected for experimentation were not blood fed or exposed to *Plasmodium* infected animals. However as these two lab colonies were raised under the same conditions in-house it is possible that mosquitoes were exposed to bacterial/fungal challenge during laboratory rearing which may have triggered expression of this *TOLL* gene. Elevated expression in the gut reported by Luna *et al.* (2002) may suggest a trigger present in an ingested substance such as food or water. Although no obvious bacterial/fungal contamination of rearing environment was witnessed and water, larval food and sucrose sources were changed regularly.

A putative sphinomyelin phosphodiesterase or sphinomyelinase (SMase) was also up-regulated ~ 4 fold in the laboratory control colonies. SMase is an enzyme responsible for the breakdown of sphingomyelin forming ceramide and phosphocholine (Kitatani *et al.*, 2008). Ceramide is proposed to be a co-ordinator of cellular stress responses (Hannun, 1996) with many inducers of stress causing ceramide accumulation which can take place via activation of SMase (Andrieu-Abadie *et al.*, 2001). It is possible that laboratory specific stresses may have triggered SMase expression in colony material with these stresses also potentially triggering *TOLL9* over-expression.

As highlighted, the current lack of annotation for a number of up-regulated genes has prevented a more thorough discussion of the potential roles in phenotype. As annotation of the *An. gambiae s.s.* genome improves these candidate genes can be revisited.

5.6.8.1 *Ace-1* and *OBP3* expression in the Akron colony

The Akron colony was originally provided by MR4 (Malaria Research and Reference Reagent Resource Center http://www.mr4.org/Home/tabid/93/Default.aspx) where the colony had been selected for carbamate resistance. The target of carbamate insecticides as well as organophosphate (OP) compounds is acetylcholinesterase encoded by gene *ace-1*. A mutation in this gene at position 119 (*G119S*) is associated with resistance to both carbamates and OPs in mosquitoes (Weill *et al.*, 2004) though incurs potential fitness costs (Djogbenou *et al.*, 2010). It is believed that in some insects carrying the *G119S* mutation, the *ace-1* gene is duplicated so that a functional copy of the gene is also present (Djogbenou *et al.*, 2009). See Chapter 2 for further discussion of *ace-1* and resistance to insecticides.

Over-expression of *ace-1* was recorded in the Akron colony compared to both N'gousso and Ghanaian field samples with a mean fold change of 4.55 [0.6132 SD] ($p<0.001$). This colony is known to carry the *G119S* mutation though has not been under bendiocarb selection since arriving at LSTM 12 months prior to the experiment. This elevated expression may reflect a constitutive over-expression of the *ace-1* enzyme as a mechanism to over-come the lethal effects of carbamate insecticide during selection at MR4.

Alternatively, higher levels of the *ace-1* transcript may have resulted from gene duplication in the Akron colony and both the mutant and wild-type genes being expressed concomitantly. To date the *ace-1* duplication has been studied via indirect

methods based on departure from Hardy-Weinberg phenotypic frequency equilibrium (HWE) (Lenormand *et al.*, 1998). Using this technique the duplication has been detected in *An. gambiae s.s* .populations from Burkina Faso (Djogbenou *et al.*, 2008b) and Côte d'Ivoire (Djogbenou *et al.*, 2009) and is suspected to be present in Benin (Djogbenou *et al.*, 2010), where the Akron colony originates. Ghanaian field samples from this study were tested for potential *G119S* divergence from HWE; however no departure was detected (Chapter 2). To test this hypothesis, departure from HWE could be investigated in the Akron colony; however ideally empirically evidence of a duplicated *ace-1* gene should be sought via Southern blotting or extensive sequencing of individual mosquitoes.

Of note with regards to carbamate resistance and the Akron colony is the up-regulation of *OBP3* within this strain; 2.91 fold over expression was recorded versus the DDT group (adjusted $p<0.01$) and ~1.5 fold elevated expression compared with N'gousso, though p-value was not significant (adjusted $p=0.1234$). *OBP3* was up-regulated in the bendiocarb selected Ghanaian samples investigated in Chapters 3 and 4 with a role in insecticide transport speculated. Elevated *OBP3* expression in the carbamate selected Akron colony may support the involvement of this odorant binding protein in bendiocarb resistance.

5.7 Conclusions

Overall, p-values calculated using the MAANOVA approach were much more significant compared with the bendiocarb microarray study (Chapter 3). A global F test comparing all three experimental groups resulted in 1339 probes (from 14914) producing p-values <0.001 after multiple testing correction. A number of factors may have contributed to these improved p-values. Firstly, the number of biological samples per RNA pool was double that of the bendiocarb experiment, with RNA extracts combined to create pools representative of 10 mosquitoes. Inherent biological variance within field samples was potentially a major contributor to the 'noisy' microarray data generated in the bendiocarb experiment, and the smaller biological sample size (n=5) may have compounded this variation.

Secondly, due to the high levels of DDT resistance recorded in the field, the experimental design employed two comparative laboratory populations with no or little resistance to DDT. If the previous design of test versus unexposed field control (Chapter 3) had been pursued, fold-changes would have been vastly underestimated (Muller *et al.*, 2008b). Non-exposed field controls represent a mixture of resistance phenotypes, and in areas where populations exhibit high levels of insecticide resistance, a large proportion of the control samples will be resistant. This adds to the variability in the data-set and means true genotype-phenotype associations may be more difficult to identify.

The pursued experimental design enabled a highly resistant population to be characterised through comparison to a fully susceptible control, with secondary confirmation of gene 'hits' in an independent comparison. This provided more confidence in results and also contributed to improved p-values. Sourcing control

populations from two different West African countries potentially addressed geographical confounding, while the treatment of material in an identical manner to field collected samples reduced variation associated rearing, age, sample storage and RNA extraction.

Furthermore, the use of a fully interwoven loop design and ANOVA based analysis, helped to reduce the variance associated with ancillary factors (Kerr, 2000). This allowed overall selection of the most significant hits associated with experimental group (test/control) from all three mosquito populations.

However, ideally, colony material would not be compared with insects collected from the field due to the potential effects of colonisation on gene expression. Indeed, a number of genes were consistently over-expressed in both lab colonies compared with the DDT samples (Table 5.4, Figure 5.5); though conversely these genes may be down-regulated in the resistant field population. However, suitable susceptible material from the field site was not available and this compromise comparison provided an option for characterising resistance in this region of Ghana. Moreover, a number of resistance associated genes have been identified through the use of colonised material (Muller *et al.*, 2007; Marcombe *et al.*, 2009; David *et al.*, 2005). *CYP6M2* was up-regulated in a pyrethroid resistant *An. gambiae s.s.* colony from Ghana, when compared to an *East* African susceptible strain (Muller *et al.*, 2007); recombinant expression and analysis of this P450 has recently provided evidence of pyrethroid metabolise *in vitro* (M. Paine, pers. comm.). Gene *CYP6Z1,* over-expressed in a DDT resistant colony from Tanzania when compared to a Kenyan susceptible strain (David *et al.*, 2005), has also been shown to metabolise DDT *in vitro* (Chiu *et al.*, 2008).

5.8 Future perspectives

Proteomic analysis of candidate *CYP6M2* is currently being pursued as this protein has been previously expressed in-house from *An. gambiae s.s.* and assays developed to quantify metabolism with insecticidal compounds (Stevenson/Paine *et al.* in prep). However it will be essential to establish whether the expressed protein differs in amino acid sequence to the wild caught DDT resistant Ghanaian samples. For other P450 candidates (*CYP6Z3, Z1*) the OmpA+2 leader system (Pritchard *et al.*, 2005), employed in Chapter 4, could be employed for expression. The DDT detoxification activity of *CYP6Z1* cloned from DDT resistant Ghanaian field samples could be compared with that of previously characterised *An. gambiae CYP6Z1* derived from a pyrethroid resistant laboratory colony (Chiu *et al.*, 2008).

Microarray experiments on a whole genome scale inevitably produce a number of potentially interesting avenues to pursue. Though the constraints of time and funding dictate that the most 'promising' candidates are often pursued first which inevitably are those from gene families with previous links to the phenotype of interest. However novel mechanisms and pathways of resistance should not be over-looked and whole transcriptome profiling has the advantage of highlighting these potentially novel routes.

A number of genes from protein families not previously linked to insecticide metabolism were identified as up-regulated in DDT resistant field samples including arrestin *ARR2*, cuticular protein *CPR9*, a putative membrane transporter/aquaporin, heat-shock proteins and signalling peptides. Speculative involvement of these proteins in chemosensory networks, cuticle modifications and efflux systems have been discussed; however further investigation of the roles of these genes *in vivo* is required which may take the form of tissue specific expression profiling, DDT interaction prediction via protein modelling, or gene knock-downs to establish any loss or gain of phenotype. These investigations will be aided by improved genome annotation.

Chapter 6: Allelic variation in *An. gambiae s.s.* Glutathione-S-Transferase E2 is associated with increased detoxification of DDT.

6.1 Abstract

Metabolic insecticide resistance in mosquitoes is most commonly associated with elevated expression of detoxifying enzymes. Glutathione-S-transferase E2 is an enzyme linked to DDT resistance in *An. gambiae s.s.* and *Ae. aegypti* through over-expression (Prapanthadara *et al.*, 1993; Lumjuan *et al.*, 2005), with recombinant *GSTe2* able to metabolise DDT *in vitro* (Ranson *et al.*, 2001). On sequencing of *GSTe2* from a DDT resistant and susceptible strain of *An. gambiae s.s.*, non-synonymous polymorphisms which segregated with phenotype were identified. These amino-acid substitutions were found to be proximal to the DDT binding domain when the crystal structure of *GSTe2* (Wang *et al.*, 2008) was consulted. Recombinant protein expression and DDTase analysis revealed that the *GSTe2* allele from the resistant strain exhibited a greater maximum enzyme velocity, while proteins from the susceptible strain lost activity at higher DDT concentrations characteristic of substrate inhibition. This supports a role for *GSTe2* allelic variation as well as elevated expression in the DDT resistance phenotype.

6.2 Introduction

As a prelude to investigating the role of *cis* and *trans* regulation in the expression of genes linked to insecticide resistance, a number of genes were selected for expression level determination and polymorphism discovery via sequencing. The approach taken was based on the work of Wittkopp *et al.* (2004; 2008), in which genes exhibiting consistent differential expression between strains of interest were selected, and the relative abundance of strain specific transcripts in F1 hybrids analysed via fixed polymorphisms. Various metabolic genes linked to insecticide resistance in *An. gambiae s.s.* were initially chosen including cytochrome P450s and glutathione-S-transferases. Among these genes was an epsilon class glutathione-S-transferase, *GSTe2*, which displayed elevated expression in a DDT resistant strain of *An. gambiae s.s* originating from Zanzibar (ZAN/U) (Prapanthadara *et al.*, 1993). Recombinant *GSTe2* from the ZAN/U strain was subsequently found to metabolise DDT *in vitro* (Ranson *et al.*, 2001). Similarly, in *Ae. aegypti*, the ortholog *AaGSTe2* was linked to DDT resistance via elevated expression, with recombinant proteins also displaying DDTase activity (Lumjuan *et al.*, 2005).

In previous studies, *GSTe2* was found to be between 5-8-fold over-expressed in the ZAN/U strain compared to a susceptible East African colony (Kisumu) (Ranson *et al.*, 2001; Ding *et al.*, 2003), with *cis* regulatory elements in the promoter region implicated in elevated expression (Ding *et al.*, 2005). Ding *et al.* (2005) computationally predicted

the 352bp promoter region for *GSTe2*, which was confirmed by cloning this region into a reporter vector. The promoter activity from the ZAN/U strain was compared with that of the susceptible Kisumu strain and exhibited 2.8-fold higher activity, with a number of potential regulatory factors identified. A two base pair adenosine indel had most effect on activity; when two adenosine residues where introduced into the ZAN/U promoter, approximately a 7-fold reduction in activity was recorded. However the reciprocal experiment, whereby two adenosines were removed from the Kisumu promoter, failed to increase expression of luciferase in the reporter vector system. In addition, when the *GSTe2* promoter region was investigated by sequencing of DDT-phenotyped field and colony mosquitoes for this study, however, no correlation between resistance phenotype and the presence of the 2bp adenosine indel or other putative regulatory elements was found (Appendix 29).

Initial expression level comparison between in-house ZAN/U and Kisumu colonies at LSTM was performed via reverse-transcriptase quantitative PCR employing the GeXP system (Beckman-Coulter) as outlined in chapter 4. Ribosomal S7 normalised *GSTe2* expression from ten RNA pools derived from the ZAN/U colony was 2.3 fold greater than that of the susceptible Kisumu colony (Figure 6.4).

However, expression levels in the ZAN/U colony did not match the 5-8-fold elevated levels previously reported (Ranson *et al.*, 2001; Ding *et al.*, 2003). In addition, *GSTe2* expression levels appeared to be variable in the ZAN/U strain, indicated by large error bars in Figure 6.4. It should be noted that two different methods were compared, with previous work employing absolute quantification via Real-Time qPCR, which may be more accurate compared to the relative expression measured via GeXP.

The Wittkopp approach for investigating *cis/trans* regulation requires consistent differential gene expression between strains of at least 2 fold (2004). The inconsistencies in *GSTe2* expression within and between studies meant this approach was abandoned. However during sequencing of *GSTe2* in the ZAN/U and Kisumu colonies a number of differential non-synonymous single nucleotide polymorphisms (SNPs) were identified between strains. The identification of these SNPs in combination with the continued presence of DDT resistance in the ZAN/U colony (~ LT_{50} of 40-50 minutes with 4% DDT), lead to a hypothesis that allelic variation in *GSTe2* may play a role in the resistance phenotype.

6.3 Methods

6.3.1 GeXP analysis of *GSTe2* gene expression between the ZAN/U and Kisumu strain of *An. gambiae s.s*

RNA extraction

Mosquito pupae from the DDT resistant ZAN/U and susceptible Kisumu strain were separated according to sex and harvested 3 days post emergence. RNA was extracted from pools of ten female *An. gambiae s.s* from each strain using the PicoPure™ RNA Isolation Kit (Arcturus). Briefly, mosquitoes were placed into a nuclease free 1.5ml tubes and homogenised in 150µl of extraction buffer using RNAse

free pestles. The homogenate was then incubated at 42°C for 30 minutes before centrifuging for 2 minutes at 14,000RCF to pellet cell debris. The supernatant was transferred to a new 1.5ml tube to which 150µl of 70% ethanol was added before transferring the whole volume to a pre-conditioned purification column. To bind the RNA, the column was centrifuged for 2 minutes at 0.1 RCF followed by a 30 second spin at 16,000RCF to remove flow through. The column was then washed with 100µl of Wash Buffer 1 before centrifuging for 1 minute at 8,000RCF. An on column DNAse treatment step was then performed using an RNAse-Free DNase Set (Qiagen) according to manufacturer's instructions, before washing the column twice with 100µl of Wash Buffer 2 followed by a 2 minute centrifugation at 16,000RCF. The column was then transferred to a new 1.5ml tube and RNA eluted in 30µl of Elution Buffer. RNA concentration was measured on the NanoDrop® (Thermo Scientific) before samples were labelled and stored at -80°C.

GeXP *GSTe2* quantitative PCR

Primers for gene *GSTe2* and housekeeping gene ribosomal S7 were selected from a multiplex previously designed using the Genomelab GeXP eXpress Profiler software (Courtesy of P. Muller).

In total, ten pools of RNA, each containing extract from ten female mosquitoes from each strain (Kisumu and ZAN/U), were employed in GeXP *GSTe2* expression quantification. RNA was diluted to 20ng/µl using nuclease-free water (Sigma) with 5µl used per reaction. Reverse transcription (RT) and PCR amplification were performed using the GenomeLab™ GeXP Start Kit (Beckman-Coulter) according to manufacturer's guidelines with all incubations carried out in a Bio-Rad/MJ Dyad thermal cycler.

PCR amplified fluorescently labelled *GSTe2* and ribosomal S7 fragments were separated by capillary electrophoresis and analysed using the GenomeLab™ GeXP Genetic Analysis System according to the User's guide. For each of the ten biological replicates per strain, the peak areas for the fragments corresponding to *GSTe2* (297.8 bp) and S7 (164.8bp) were employed to calculate normalised *GSTe2* expression (*GSTe2* peak area/ S7 peak area). Mean normalised *GSTe2* expression was then calculated from the ten pools for each strain, and the expression ratio determined by dividing the mean normalised *GSTe2* expression for the ZAN/U strain by that of the Kisumu strain.

6.3.2 Sequencing *GSTe2* coding region from the ZAN/U and Kisumu strains of *An. gambiae s.s.*

The *GSTe2* gene was sequenced from DNA extracted from the Kisumu and ZAN/U strains of *An. gambiae s.s.* and segregating polymorphisms identified. Both of these laboratory colonies are of the S molecular form and originate from East Africa. The DDT resistant ZAN/U strain was originally colonised from Zanzibar, Tanzania in 1982 and displays DDT resistance in the absence of target site mutation *kdr.* ZAN/U adults have been maintained under DDT selection at LSTM and display an LT_{50} of

approximately 40-50 minutes with 4% DDT. The Kisumu strain is fully susceptible to DDT and originates from Kisumu in Western Kenya.

DNA was extracted from ten individual female mosquitoes of 3-5 days old from both the Kisumu and ZAN/U strain. Total DNA was purified from single insects using the DNeasy® Blood and Tissue spin column kit (Qiagen) according to manufacturer's guidelines with a final elution volume of 200µL. The concentration of DNA extracts was measured on the NanoDrop™ spectrophotometer (Thermo Scientific) before being stored at -20°C until required.

All twenty DNA extracts were confirmed as *An. gambiae sensu stricto* of the S molecular form using a PCR-RFLP approach (Fanello *et al.*, 2002).

Gene *GSTe2* (accession AGAP009194) is situated on chromosome 3R at position 28,597,686-28,598,594 (AgamP3.5). The gene is comprised of three exons separated by two introns and from start to stop codon the sequence is 843bp. Primers were designed to amplify a 680bp fragment from gene *GSTe2*; the forward primer positioned in exon 1 at 28,598,483-28,598,461 (CCATTCATCTGCTAACGGGTGAC), and the reverse primer in exon 3 at 28,597,837-28,597,817 (GTTCGCTTCCTCGTAGTAGGG).

PCR amplification was performed using HotStarTaq® polymerase and buffers (Qiagen) and products were purified using the QIAquick® Spin kit (Qiagen) before being sent for sequencing (Macrogen). Each reaction contained 1µl of DNA solution, 0.2mM of each primer, 1.5mM $MgCl_2$, 0.2mM dNTP mix, 1 X HotStar PCR buffer and 0.8 units of HotStarTaq polymerase at a total volume of 25µl. Cycling conditions were 95°C for 15 minutes, 30 cycles of 95°C for 30 seconds, 48°C for 30 seconds, 72°C for one minute, with a final extension of 10 minutes at 72°C.

6.3.2.1 Cloning of *GSTe2* for sequencing from Kisumu mosquitoes

GSTe2 was isolated from eight of the Kisumu individuals using a cloning approach. The *GSTe2* 680bp fragment was amplified from the 8 selected DNA extracts following previous protocols and cloned using the pGEM®-T Easy Vector (Promega). After checking for product on a 1.5% agarose gel, 1-2 µl of each PCR product, depending on band intensity, were using in a ligation reaction with pGEM®-T Easy Vector (Promega) following manufacturer's instructions. After over-night ligation at 4°C, 2µl of each reaction were transferred to 14ml round bottomed tubes on ice before 50µl of chemically competent *E.coli* cells (JM109 AllianceBio) were introduced followed by 20 minute incubation on ice. Cells were then heat-shocked by placing tubes in a water bath at 42°C for 45 seconds before returning tubes to ice for 2 minutes.

Cells were recovered in 950µl of SOC medium (0.5% yeast extract, 2% tryptone, 10mM NaCl, 2.5mM KCl, 10mM $MgCL_2$, 10mM $MgSO_4$, 20mM Glucose) in a 37°C shaking incubator (150 RPM) for one hour before plating. LB-amp-agar plates (1% tryptone, 0.5% yeast extract, 200mM NaCl, 1.5% agar, 100µg/ml ampicillin) were spread with 100µl of 100mM IPTG (Isopropyl β-D-1-thiogalactopyranoside) and 20µl of 50mg/ml X-gal (bromo-chloro-indolyl-galactopyranoside) solution using a flamed spreader and allowed to dry before 50 and 100µl of each transformed culture were spread onto

plates. Post over-night incubation at 37°C, 4 white colonies (positives) and one blue (negative) colony were selected from each transformation using a sterile pipette tip and were spotted onto gridded fresh plates (LB-amp-X-gal -IPTG agar). The remainder of the colony was introduced into a PCR mix for screening with plasmid specific primers (M13). Each 15µl reaction contained 0.3mM of each M13 primer, 2mM $MgCl_2$, 0.2mM dNTPs, 0.4 units of taq polymerase (Bioline) and 1X PCR buffer (Bioline). Cycling conditions were 95°C for 5 minutes, 30 cycles of 94°C for 30 seconds, 55°C for 30 seconds, 72°C for 30 seconds, with a final extension of 5 minutes at 72°C.

Five microlitres of each PCR reaction were visualised on a 1.5% agarose gel containing 0.5µg/ml ethidium bromide and a selection (12) of clones were chosen for sequencing. Selected clones were identified on the gridded plates and used to inoculate 5ml of Luria-Bertani (LB) medium (1% tryptone, 0.5% yeast extract, 200mM NaCl) containing 100µg/ml ampicillin in 14ml round bottomed tubes. Cultures were grown overnight at 37°C in a shaking incubator (150RPM). Subsequently cells were pelleted from 1.5ml of each over-night culture via centrifugation and glycerol stocks prepared from the remaining culture (50:50, cells to 80% glycerol). Plasmid DNA was extracted from cell pellets using a column based mini-prep kit (QIAprep Spin Miniprep Kit, Qiagen) according to manufacturer's instructions eluting in 50 µl of elution buffer (10mM Tris-Cl, pH 8.5) before aliquots were sent for sequencing (Macrogen).

6.3.3 Sequence alignment and polymorphism identification

All sequences were aligned alongside the full *GSTe2* genomic sequence obtained from VectorBase (http://www.vectorbase.org/) using CodonCode Aligner software (CodonCode Corporation) and polymorphisms highlighted. Non-synonymous single nucleotide polymorphisms (SNPs) were identified by translating the *GSTe2* exons into corresponding amino-acids.

6.3.4 Computational modelling of *GSTe2* and SNPs identified

Recombinant protein expression of the *GSTe2* variants identified was pursued alongside computational modelling of the mutations within the *GSTe2* protein structure.

The crystal structure of *GSTe2* from the susceptible Kisumu strain of *An. gambiae s.s.* was resolved by Wang *et al.* (2008); however they were unable to co-crystallise the enzyme with a DDT substrate. The authors pursued computational docking of DDT into the putative binding site, a hydrophobic pocket adjacent to the glutathione (GSH) binding site. From modelling the DDT-*GSTe2* interaction a number of key residues with potential importance in DDT binding were identified.

The *GSTe2* crystal structure with computationally docked DDT was provided as a Protein Database (PDB) file courtesy of L. Chen (Laboratory for Structural Biology, University of Alabama). The positions of the *GSTe2* variants identified in this study were then interrogated by protein modeller Daniel Rigden and protein crystallographer Olga Mayans (School of Biological Sciences, Liverpool University). The amino-acid changes identified in the ZAN/U and Kisumu sequence data were over-laid onto the

model and the potential importance in DDT binding inferred by consulting the residues highlighted by Wang *et al.* (2008). A structural model of *GSTe2* highlighting the potentially important variant residues identified in the Kisumu-ZAN/U comparison was produced using the open-source PyMOL (www.pymol.org) software based on the *An. gambiae GSTe2* crystal structure obtained from the Protein Database (PDBcode: 2IMI).

6.3.5 *GSTe2* recombinant protein expression

Expression of the three variant *GSTe2* enzymes identified was pursued using an *E.coli* system. RNA was extracted from three batches of ten 1 day old female mosquitoes from either the Kisumu or ZAN/U strain using the PicoPure™ RNA Isolation Kit (Arcturus) as previously outline. RNA concentration was measured using the NanoDrop™ spectrophotometer and approximately 2µg were employed from each pool for complementary DNA synthesis. First strand cDNA synthesis was performed using SuperScript™ III Reverse Transcriptase (Invitrogen) as outlined in Appendix 22. The cDNA sequence of gene *GSTe2* was amplified from all six cDNA pools (3 Kisumu, 3 ZAN/U) using primers situated in the 5' and 3' untranslated regions (UTR) of the cDNA sequence (forward– CGCTGCGAAAATGTCCAACC, reverse-TACCTTTTTAAGCCTAGCATTC), which produced a 683bp fragment.

PCR reactions contained 1µl of cDNA solution, 0.4µM of each primer, 2mM $MgCl_2$, 0.2mM dNTP mix, 1X Kapa PCR buffer A and 1 unit of KapaTaq (Kapa Biosystems) in a total volume of 25µl. Cycling conditions were 94°C for 5 minutes, followed by 35 cycles of 94 °C for 30 seconds, 55 °C for 30 seconds, 72 °C for 45 seconds, with a final extension of ten minutes at 72 °C. The amplified *GSTe2* fragment from each cDNA pool was then cloned into a holding vector (pGEM®-T Easy Vector (Promega)) using 1µl of PCR product in each ligation as previously outlined. Colonies were screened with plasmid specific primers (M13) before clones were selected from each strain (Kisumu/ZAN/U) for sequencing. Selected clones were used to inoculate a 5ml over-night culture from which plasmid DNA was extracted using a column based mini-prep kit (QIAprep Spin Miniprep Kit, Qiagen) as outlined. An aliquot of each plasmid was then sent for sequencing (Macrogen) with the remainder stored at -20°C.

6.3.5.1 *GSTe2* coding region amplification incorporating restriction sites

After sequence alignment, clones were selected which represented the *GSTe2* allelic variants previously identified through genomic sequencing (Table 6.2). New primers were designed to amplify the *GSTe2* coding sequence from the start to stop codon incorporating restrictions enzyme sites to enable cloning into the pET-15b (Novagen) expression vector (Figure 6.1). The forward primer included the *NdeI* restriction site which contains the ATG start codon, while the reverse primer contained the restriction site for enzyme *BamHI* which incorporated a TAG stop codon (Table 6.1). Each primer also included three thymine bases to aid restriction enzyme docking.

Table 6.1. Primers designed to amplify the coding region of *GSTe2* incorporating the *NdeI* and *BamHI* restriction enzyme sites for subsequent cloning into expression vector pET-15b (Novagen).

GSTe2 cDNA_RE_F	***NdeI*** 5'-TTT***CATATG***TCCAACCTTGTAC-3'
GSTe3 cDNA_RE_R	***BamHI*** 5'-TTT***GGATCC***TAAGCCTTAGCATTC-3'

The *GSTe2* coding region including restriction sites was amplified in 50µl reactions from 1µl of plasmid template containing the cDNA sequence from selected clones (Kisumu 1B, 2B and ZAN/U 1C (Table 6.2)). Each reaction contained 0.4µM of each primer, 2mM $MgCl_2$, 0.2mM dNTP mix, 1X Kapa PCR buffer A and 1 unit of KapaTaq (Kapa Biosystems). Reactions were incubated for 5 minutes at 94°C, followed by 35 cycles of 94 °C for 30 seconds, 55 °C for 30 seconds, 72 °C for 45 seconds, with a final extension of ten minutes at 72 °C. Five microlitres of product was loaded onto a 1.5% agarose gel containing 0.5µg/ml ethidium bromide and visualised on a gel imaging system. The remaining 45µl were purified using a column based kit (QIAquick PCR purification kit, Qiagen) and eluted in 30 µl of elution buffer (10mM Tris-Cl, pH 8.5) before storage at -20°C.

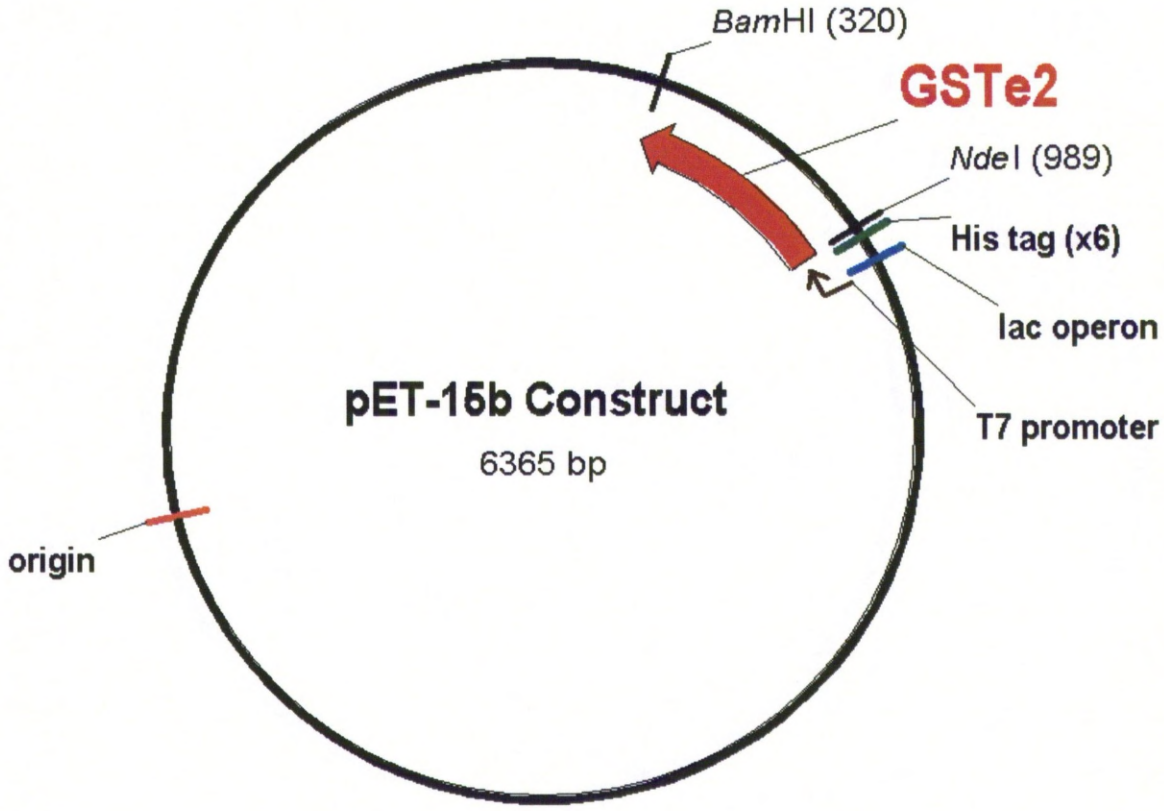

Figure 6.1. The *GSTe2* protein expression construct.

GSTe2 coding sequence was cloned into the pET-15b (Novagen) expression vector via incorporation of *NdeI* and *BamHI* restriction sites. The expression vector also adds a 6X histidine tag to expressed proteins and expression is under the control of a T7 promoter.

6.3.5.2 Restriction enzyme digest and isolation

Amplified and restriction tagged *GSTe2* coding fragments were cloned into a holding vector (pGEM®-T Easy Vector (Promega)) to aid visual confirmation of the restriction digest on agarose gels and for variant sequence validation.

Sequential restriction digests were performed on a pET-15b construct containing a previously cloned gene and all three *GSTe2* variants in the pGEM® holding vector. Four micrograms of each plasmid went into a 50µl digest containing 2µl (20 units) of *NdeI* (Promega), 1X buffer D (Promega) and 1X BSA (Promega). The reaction was incubated at 37°C for one hour before checking 5 µl of the digest on a 1.5% agarose gel (0.5µg/ml ethidium bromide) to ensure total digestion.

The *NdeI* cut plasmids were then purified using the Merlin DNA Purification System (Lyer, 1993, http://www.bio.net/bionet/mm/methods/1994-April/013440.html) which employs a silica based DNA purification resin and Wizard® Minicolumns (Promega) driven by syringe pressure. Briefly, 600µl of resin solution IV (1.5% w/v Diatomaceous Earth (Sigma) in 7M Guanidine hydrochloride buffer) was added to the 40 µl *NdeI* digest and mixed thoroughly before being passed through the wizard column via syringe action. The bedded resin was washed twice with wash buffer V (200mM NaCL,

20mM Tris.HCL, pH 7.5, 5mM EDTA, 50% ethanol) before the column was centrifuged at full speed for 3 minutes to remove residual ethanol. DNA was then eluted into a clean 1.5ml tube by adding 40µl of warmed water (75°C) and spinning the column for 30 seconds at full speed.

All 40µl of each cut plasmid was then digested with *BamHI* in a 50 µl reaction containing 20 units of *BamHI* (Promega), 1X buffer E (Promega) and 1X BSA (Promega). Reactions were incubated for one hour at 37°C before purification via the Merlin DNA Purification System as described. Purified double digested pET-15b plasmids were then treated with Shrimp Alkaline Phosphatise (SAP) to prevent re-ligation of digested fragments. One microlitre (1 unit) of SAP (Promega) was added to the 40µl digest in 1X SAP buffer (Promega) in a total volume of 50µl. The reaction was incubated at 37°C for 30 minutes before heat inactivation (65°C, 15 minutes).

Restriction cut pET-15b plasmid and *GSTe2* coding region fragments were isolated via gel electrophoresis and band excision. Initially 3 µl of each digest was run on a 1.5% agarose gel (0.5µg/ml ethidium bromide) to visually inspect digestion. The remainder of each digest was then run on a 1% gel (0.5µg/ml ethidium bromide) for ~ 1hour before being visualised at a low UV intensity through a series of glass plates to reduced UV exposure and related DNA damage. Correct bands were excised and DNA extracted using a Merlin based gel extraction system. Gels pieces were incubated with 600µl of Solubilisation and Binding Buffer (Buffer QG, Qiagen) containing 1.5% w/v Diatomaceous Earth (Sigma) and incubated at 50°C until gel was fully dissolved. The mixture was then passed through a Wizard® Minicolumn, washed and eluted as previously described before DNA concentration was measured on the NanoDrop™ spectrophotometer.

6.3.5.3 Ligation & transformation

A three to one ligation ratio was calculated for the 665bp insert and 50ng of the 5.7Kb cut pET-15b vector using the equation outline in Chapter 4. Twenty nanograms of each *GSTe2* variant were ligated with 50ng of pET-15b in a 10µl reaction containing 0.5µl T4 ligase (200 units) (NEB) and 1X ligase buffer (NEB). Reactions were incubated at 16°C overnight.

Five microlitres of each ligation reaction was used to transform 50µl of chemically competent *E.coli* cells (DH5α (Invitrogen)). The ligation mixture was added to cell aliquots on ice and gently mixed before incubating on ice for 30 minutes. Cells were then heat-shocked for 20 seconds at 42°C then placed on ice for 2 minutes. Cells were then recovered in 950µl of SOC medium and incubated at 37°C for 1 hour at 225 RPM before plating 100µl and 200µl on LB-amp agar plates and incubating over-night at 37°C.

Colonies were screened by mixing a scraping with 5µl of LB medium and using 1µl in a PCR reaction with a gene specific primer (GSTe2 cDNA_RE_F), and an external plasmid primers (T7_R). Each 20 µl reaction contained 0.5µM of each primer, 2mM $MgCl_2$, 0.2mM dNTP mix, 1 X buffer (Kapa Biosystems) and 1 unit of KapaTaq (Kapa Biosystems). Reactions were incubated at 94°C for 5 minutes followed by 30 cycles at

94 °C for 30 seconds, 55 °C for 30 seconds, 72 °C for 45 seconds and a final extension at 72 °C for ten minutes. Bands were visualised on a 1% agarose gel (0.5µg/ml ethidium bromide).

The remaining 4µl from positive clones was used to inoculate 5ml of LB containing 100µg/ml ampicillin prior to overnight culture in a 37°C shaking incubator at 150RPM. Cells were pelleted after taking 100µl for glycerol stocks (50:50, cells to 80% glycerol) and plasmid DNA extracted using the QIAprep Spin Miniprep Kit (Qiagen) according to manufacturer's instructions eluting in 50µl of elution buffer (10mM Tris-Cl, pH 8.5). Aliquots of each plasmid were sent for sequencing (Sheffield University, UK) to confirm the correct variants were cloned into the pET-15b vector.

6.3.5.4 BL21 cell transformation and induction

On confirmation that the correct *GSTe2* variants were cloned into the pET-15b vector, 1µl of each variant plasmid was used to transform 50µl of BL21(DE3) cells (NEB), a chemically competent *E.coli* protein expression cell strain. The plasmid was added to cell aliquots on ice and gently mixed before incubating on ice for 30 minutes. Cells were then heat-shocked for 45 seconds at 42°C before being placed back on ice for a further 2 minutes. Subsequently cells were recovered in 950µl of SOC medium and incubated at 37°C for 1 hour at 225 RPM before plating 100µl and 200µl on LB-amp agar plates and incubating over-night at 37°C.

Colonies were then screened with a gene specific (GSTe2 cDNA_RE_F) and external plasmid primers (T7_R) as previously described. A 5ml over-night LB-Amp culture was set up from a selection of positive clones from each *GSTe2* variant from which an LB-amp-agar streak plate and glycerol stocks (50:50, cells to 80% glycerol) were produced.

6.3.5.5 *GSTe2* protein induction and purification

A scraping from streak plates representing each variant was used to inoculate 200ml of LB-Amp medium (100µg/ml ampicillin) in a 500ml baffle flask, and the culture incubated at 37 °C, 150RPM. The optical density (OD) was read at regular intervals using spectrophotometric readings from 300µl of culture, blanked with LB, at a wavelength of 595 nm. When OD reached 0.8, protein expression was induced by addition of IPTG to a final concentration of 1mM. Induced cultures were incubated at a reduced temperature (30°C) over-night at 150RPM.

Cells were then harvested in 250ml plastic flasks (Sorvall) via centrifugation at 10,000g for 10 minutes at 4 °C. The cell pellet was then re-suspended in 20ml of TSE buffer (50mM trisma base, 1mM EDTA, 150mM NaCl, pH 7.4) before being transferred to a 50ml plastic tube (Falcon) and placed at -80°C. After two rounds of freeze-thawing, 14µl of 14M 2-mercaptoethanol (final concentration 10mM), 25µl MgCl2 (final concentration 1.25mM) and 250U Benzonase® were added to re-suspended cells before sonicating 4 times for 30 seconds using a Sonics Materials Vibracell sonicator.

To remove cell debris, lysed cells were then centrifuged at 10,000g for 20 minutes at 4°C and the supernatant filtered through a 0.2µm filter (Sartorius Stedim Biotech) into a clean 50ml tube (Falcon) before being placed on ice. The pET-15b expression vector adds an N-terminal 6X Histidine tag (2.2kDa) (Figure 6.1) to recombinant proteins which was exploited for *GSTe2* protein purification using metal affinity chromatography.

A 1ml Ni-NTA Agarose column (Qiagen) was poured from 50% slurry into a 25ml plastic column and equilibrated with 10 column volumes (10ml) of Wash buffer (50 mM sodium phosphate, 200mM NaCl, 20mM imidazole, pH 8). Before applying the crude cell lysate to the column, 0.35g NaCl (0.3M final concentration) and 400µl of 1M imidazole (20mM final concentration) were added to the lysate and mixed by inverting. The lysate was then applied to the Ni-NTA column and flow-through containing un-bound proteins collected (Flow). Bound protein was washed with 10 column volumes of Wash buffer and flow through collected before eluting the bound protein in 10ml of Elution buffer (50 mM sodium phosphate, 0.3M NaCl, 250 mM imidazole, pH 8) collected in 0.5ml fractions.

The protein concentration of eluted fractions was assessed using the Bio-Rad Protein Assay reagent (Bio-Rad), based on the Bradford method (1976). Five microlitres of each protein fraction were added to 250µl of Bio-Rad Protein reagent in a microtitre plate and incubated for one minute at room temperature before reading absorbance on a spectrophotometer at 595nm. The five fractions (2.5ml) containing the highest concentrations of protein were pooled then added to a PD-10 Desalting Column (GE healthcare) which had been pre-equilibrated with 25ml of 2X Storage buffer (100mM sodium phosphate, 20mM Dithiothreitol (DTT), pH 7.4). Protein was eluted from the PD-10 column by the addition of a further 3.5ml of 2X Storage buffer and aliquoted into 0.5ml volumes to which an equal volume of 80% glycerol was added before storage at -80°C.

Protein concentration was determined using a commercial Protein Quantification Kit (Fluka – Sigma-Aldrich) based on the Bradford protein assay (1976) (Appendix 30). The purification process was also visualised using Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). Aliquots of crude cell lysate, the Flow fraction (unbound protein) and purified eluted protein were visualised using the NuPAGE® SDS-PAGE system (Invitrogen) (Appendix 31). Finally GST activity was confirmed for each purified recombinant variant using the GST substrate 1-chloro-2, 4-dinitrobenzene (CDNB) in a standard colorimetric activity assay (Appendix 32).

6.3.6 DDTase assay

The DDT dehydrochlorinase activity of all three *GSTe2* variants (K1B, K2B and Z1C) was assessed using an enzymatic assay and High Performance Liquid Chromatography method based on that of Prapanthadara *et al.* (1993). *GSTe2* catalyses the dehydrochlorination of DDT in the presence of glutathione (GSH) to produce metabolite 1,1-dichloro-2,2-bis(p-chlorophenyl) ethylene (DDE) (Figure 6.2). Both DDT and DDE can be separated and identified on a HPLC column based on retention time in the presence of a mobile solvent, and quantified using peak area and standard curves.

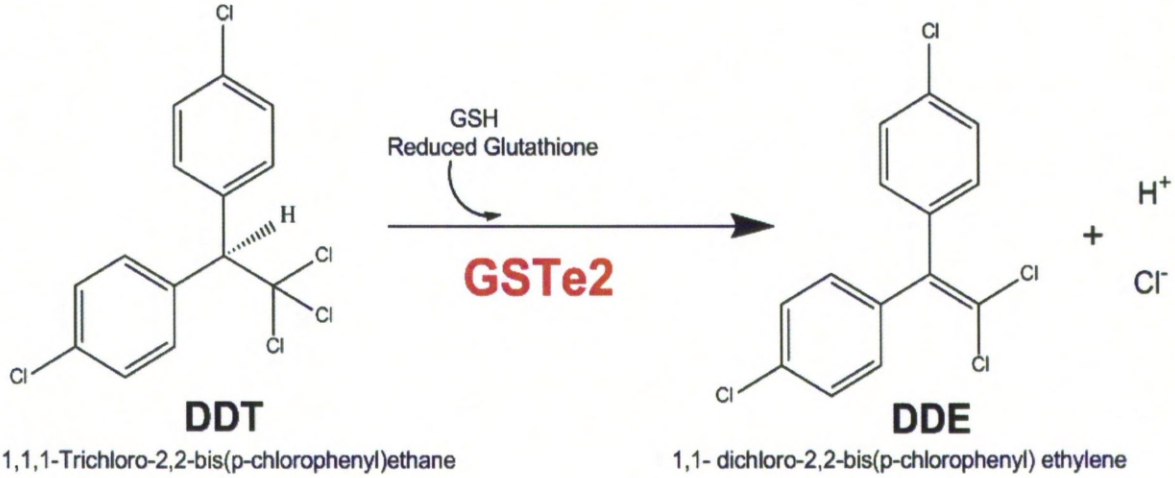

Figure 6.2. DDT dehydrochlorination catalysed by *GSTe2*.

GSTe2 mediated dehydrochlorination of DDT to form DDE in the presence of reduced glutathione.

6.3.6.1 DDT, DDE and spike-in Dicofol standard curve

A doubling dilution series from 200 - 12.5µg/ml was produced for DDT, DDE and spike-in standard Dicofol. Approximately 5mg of each compound in powdered form (ChemService) was accurately measured before being dissolved in 1ml of acetonitrile (Sigma-Aldrich). This stock solution was used to prepare the doubling dilution series using 50% acetonitrile and 50% 0.1M potassium phosphate buffer (pH 6.5). A blank containing only buffer and acetonitrile was also prepared for each dilution series.

Reverse-phase HPLC was employed using a silica based stationary phase and 90%:10% methanol:water mobile phase which separated molecules according to their polarity. DDT, closely related compound Dicofol, and DDE are detected via UV absorbance with a preferred wavelength of 232nm. Dicofol is the most polar of the three compounds and therefore elutes first during HPLC analysis at approximately 9 minutes while DDE is less polar than DDT so elutes last. DDE also contains a double c-c bond which means it has a higher UV absorbance (H. Ismail pers. comm.).

To produce standard curves, 150µl of each chemical dilution was transferred to glass HPLC vials (Chromacol) and fitted with a penetrable lid. The mobile solvent phase (90% methanol:10% water) was pumped through the HPLC system (Ultimate 3000) at a rate of 1ml/minute and 20µl of each sample injected. Data acquisition was set at 18 minutes as DDE elutes at approximately 14 minutes with DDT eluting at ~12 minutes, and the UV_2 channel (232nm wavelength) selected.

Compound concentration (µg/ml) was then plotted against the HPLC peak area at UV wavelength 232nm to produce a standard curve with the intercept fixed at zero (Appendix 33). The equation of this curve was employed to assess DDT, DDE and Dicofol concentration in future assays.

6.3.6.2 Dilution series (Michaelis-Menten enzyme kinetics)

To compare enzyme activity between variants, the enzyme kinetics according to substrate concentration was investigated. A doubling dilution series of DDT from 200-3.125 µg/ml was employed using optimised reaction parameters (Appendix 34). Each variant *GSTe2* protein was assayed at each DDT concentration and a series of three technical replicates performed.

The DDE area from the HPLC trace of each reaction was normalised against the Dicofol spike-in area (DDE area/Dicofol area), and the normalised area used to calculate micrograms of DDE produced per ml reaction using the standard curve equation (Appendix 33). The concentration of DDE in nanomoles was then determined using the molecular weight of DDE (318.03g/mol) and used to calculate the enzyme rate, *nmol DDE/mg GSTe2 protein/min*, for 60µg (0.06mg) of enzyme.

From substrate (DDT) concentration and product (DDE) formation a Michaelis-Menten plot (Figure 6.3) was produced to compare the kinetics of each *GSTe2* allele. The plot represents the rise in metabolic rate for a given enzyme as substrate concentration is increased until the enzyme reaches saturation point and rate begins to plateau. The enzyme reaction can be defined as follows (Lehninger, 1982):

$$E+S \leftrightarrow ES \rightarrow PE$$

The first part of the equation represents reversible binding of the enzyme (E) and substrate (S) to form the ES complex. While second part shows the enzyme-substrate complex breaking down to form product (P) and free enzyme, which is the slower, rate limiting step. The rate of the reaction is maximum when almost all enzyme is present as ES complex with high substrate concentration driving the formation of ES in the top equation until all free enzyme is occupied/saturated.

The Michaelis-Menten plot was employed to calculate the maximum enzyme rate (V_{max}) for each protein when all active sites are saturated and the Michaelis-Menten Constant (K_M) which is the substrate concentration for an enzyme at half its maximum velocity (Figure 6.3). This constant is used to define the relationship between substrate and enzyme velocity.

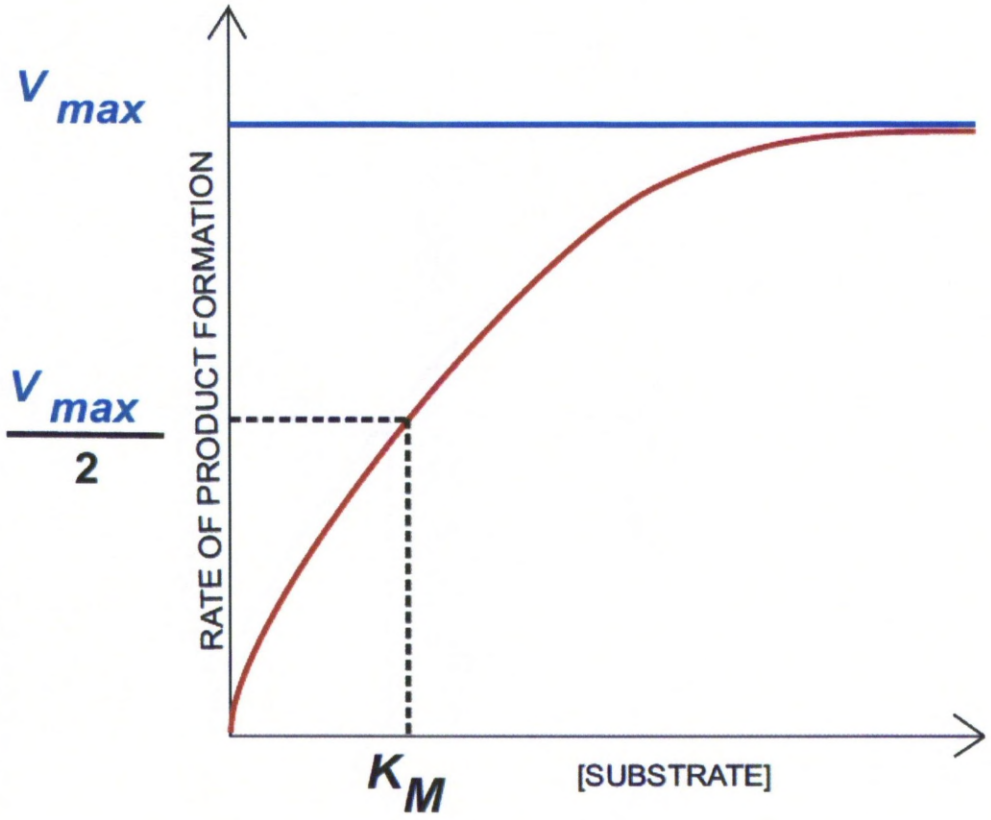

Figure 6.3. Michaelis-Menten plot for an enzyme showing the relationship between substrate concentration and metabolic rate.

The V_{max} is the maximum rate for an enzyme, while the K_M is the substrate concentration at which an enzyme reaches half its maximum velocity.

6.4 Results

6.4.1 GeXP analysis of *GSTe2* gene expression between the ZAN/U and Kisumu strain of *An. gambiae s.s*

Mean normalised expression of gene *GSTe2* in the ZAN/U samples was 2.3 fold greater than that of the susceptible Kisumu colony. Although expression of *GSTe2* was more variable within the ZAN/U samples, indicated by the error bars in Figure 6.4.

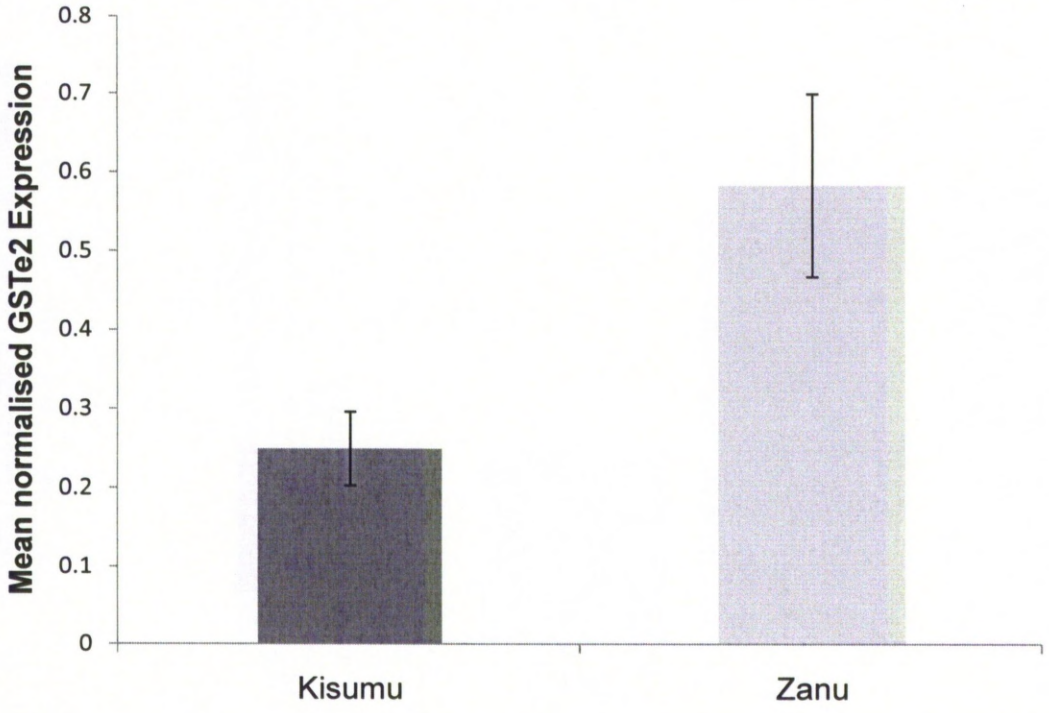

Figure 6.4. Mean normalised expression of *GSTe2* in female *An. gambiae s.s.* of the DDT resistant ZAN/U strain and susceptible Kisumu strain.

Expression of *GSTe2* and ribosomal S7 were assessed from ten RNA pools comprised of ten 3 day old female mosquitoes using the GeXP quantitative PCR system (Beckman-Coulter). The ZAN/U colony showed 2.34 fold greater expression of *GSTe2* compared with the Kisumu colony. *GSTe2* expression was normalised against housekeeping gene ribosomal S7. Standard error of the normalised mean expression is also indicated.

6.4.2 Sequencing *GSTe2* coding region from the ZAN/U and Kisumu strain of *An. gambiae s.s.*

A total of three strain segregating polymorphisms were uncovered in the sequenced section of *GSTe2* from ten individual mosquitoes of the Kisumu and ZAN/U strain (Figure 6.5). Firstly an intronic SNP at position 202 where an A>T change from the database deposited sequence was present in all Kisumu samples compared with ZAN/U. Secondly within exon two an A>G change was witnessed at position 285 in all Kisumu individuals; however this change was synonymous. Lastly a non-synonymous SNP was identified in exon two at position 415 where a T>C change was recorded in all ZAN/U individuals. The presence of a thymine at position two of this codon caused an isoleucine to threonine change in the encoded amino acid (I114T).

Further partially segregating polymorphisms included a 12bp insertion at intronic position 163 which was present in approximately half of Kisumu individuals but absent in all ZAN/U mosquitoes analysed. A second partially segregating non-synonymous SNP at 524C>G caused a phenylalanine to leucine shift (F120L) in approximately half the Kisumu individuals sequenced, while all ZAN/U individuals had a phenylalanine at this position.

It was also noted that the encoded amino-acid at segregating position 114 in ZAN/U was different to that previously recorded for this strain by previous authors (Ranson *et al.*, 2001; Ding *et al.*, 2003; Ortelli *et al.*, 2003). The ZAN/U sequence reported in these previous investigations possessed an asparagine residue at this position rather than a threonine.

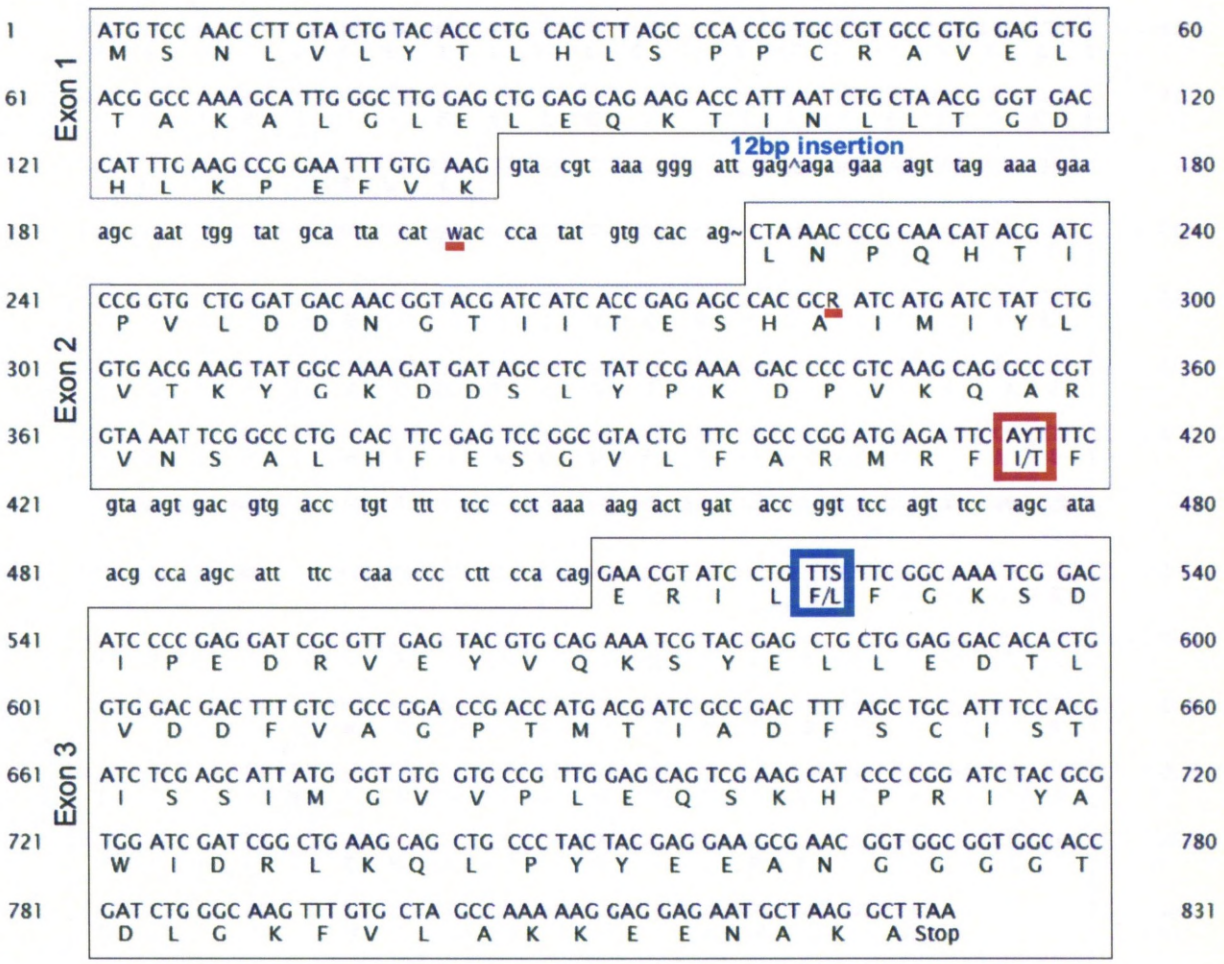

Figure 6.5. *GSTe2* genomic DNA sequence showing segregating (red) and partially segregating (blue) polymorphisms between two strains of *An. gambiae s.s.* (Kisumu and ZAN/U).

Universal code for base ambiguity applied, W= A or T, R= A or G, Y= C or T, S= C or G.

6.4.3 Computational modelling of *GSTe2* and SNPS identified

Wang *et al.* (2008) compared the crystal structure *An. gambiae s.s. GSTe2* with that of delta class GST, *GSTD1-6*, which displays much lower metabolic activity with DDT (Ranson *et al.*, 1997). From this comparison the authors were able to define structural evidence for the high DDTase activity of enzyme *GSTe2* and highlight a number of residues which were potentially important in the predicted DDT-*GSTe2* interaction. A hydrophobic pocket in close proximity to the GSH binding site was proposed as the site of DDT binding. Predicted to be of particular importance was the inclination of the

upper section of the H4 helix (depicted in turquoise in Figure 6.6) which brought residues 112, 116 and 120 closer to the GSH binding site. These residues also helped to form a pocket 'cap' for the putative DDT binding site which would potentially increase hydrophobicity and therefore affinity for the hydrophobic DDT molecule.

The variable mutation found at position 120, *F120L*, in the Kisumu strain therefore had potential to effect the formation of this putative pocket cap as the aromatic, hydrophobic residue phenylalanine (F) was replaced by a smaller aliphatic, hydrophobic leucine (L) residue. The F120 residue is predicted to make hydrophobic contact with one of the aromatic rings of the DDT molecule (Figure 6.6). A Leucine residue at this position, being smaller, would not reach as far meaning the potential DDT interaction would be weaker (Dan Rigden, pers. comm.).

The fixed difference at position 114 (*I114T*), also situated in close proximity to the putative DDT binding pocket (Figure 6.6), was predicted to have even greater effect on the GST-DDT interaction. A destabilizing polar group (a hydroxyl) was introduced in a hydrophobic core region of the protein by the Isoleucine to Threonine change in ZAN/U which had potentially marked effects on protein configuration (Olga Mayans, pers. comm.).

These observations provided good support for pursuing recombinant protein expression of the *GSTe2* variants identified.

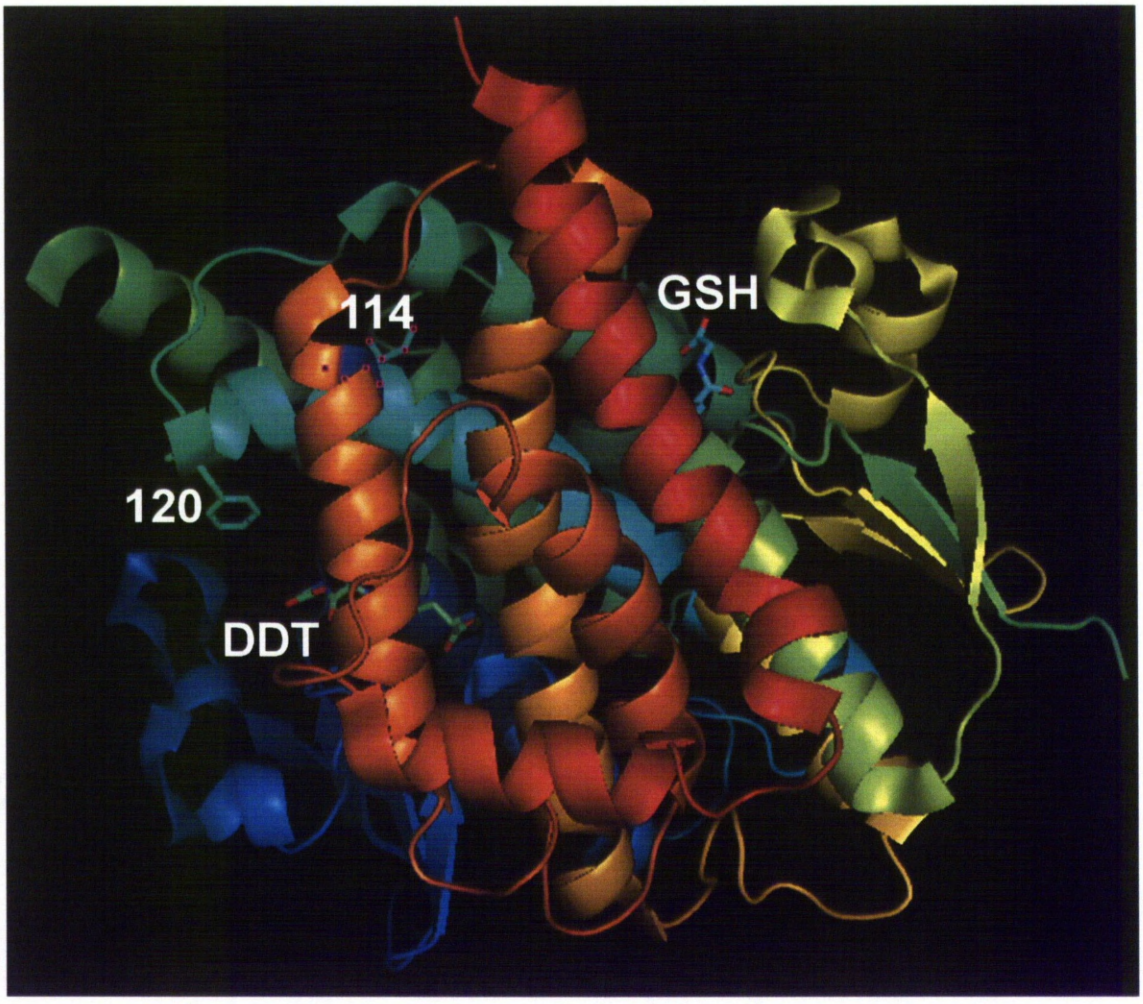

Figure 6.6. *GSTe2* crystal structure showing the position of two polymorphisms (114, 120) identified in strains of *An. gambiae s.s.*.

The *GSTe2*-GSH crystal structure with computationally docked DDT was provided by L. Chen (Laboratory for Structural Biology, University of Alabama). This structure was manipulated using the PyMOL (www.pymol.org) software to locate the positions of variant residues identified in a comparison of the DDT resistant ZAN/U and susceptible Kisumu strains.

6.4.4 Recombinant protein expression

Expression of three variant proteins polymorphic for mutations proximal to the putative DDT binding domain, at positions 114 and 120 (Table 6.2), was pursued using an *E.coli* expression system.

Table 6.2. *GSTe2* allelic variants from the *An. gambiae s.s.* Kisumu and ZAN/U strain pursued for recombinant protein expression.

The position of variant amino-acids proximal to the putative DDT binding site are shown.

Cloned variant	Amino-acid position	
	114	120
Kisumu 1B	Isoleucine	Leucine
Kisumu 2B	Isoleucine	Phenylalanine
ZAN/U 1C	Threonine	Phenylalanine

Each protein variant was successfully expressed and purified (Figure 6.7) with comparable final protein concentrations in the range of 2-4mg/ml (Table 6.3). All three variants also exhibited activity with model substrate CDNB which confirmed the expressed proteins were glutathione-S-transferases (Table 6.3).

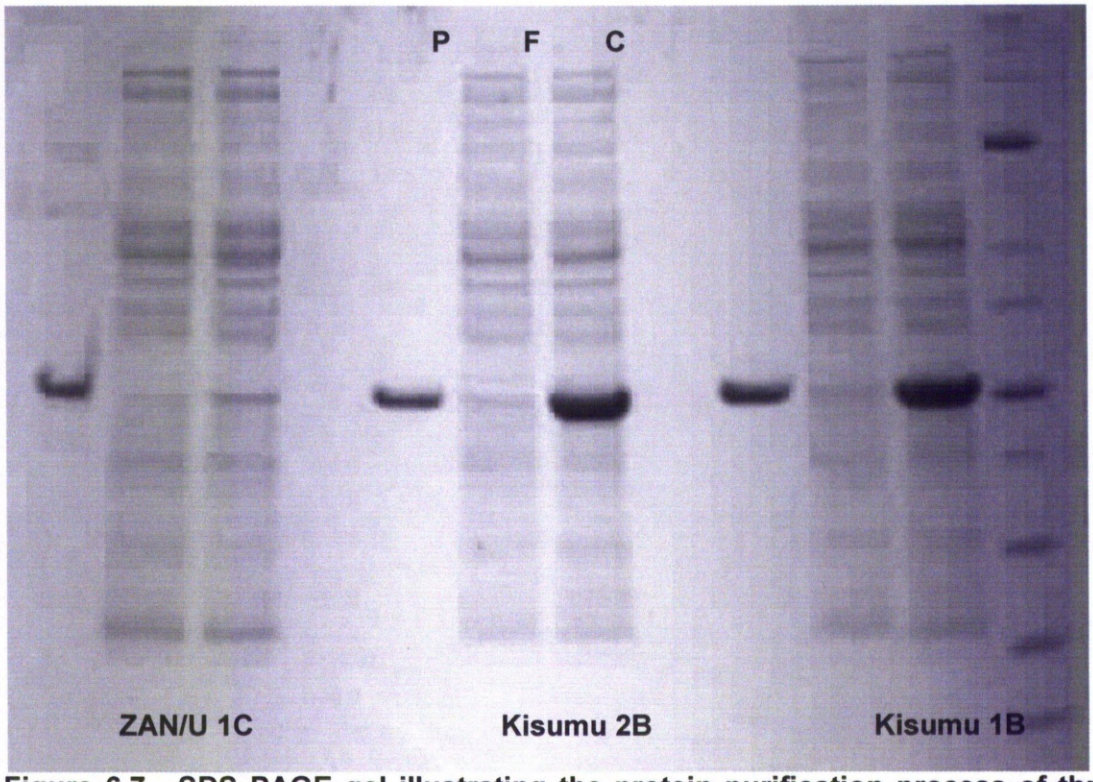

Figure 6.7. SDS PAGE gel illustrating the protein purification process of three recombinant variants of *An. gambiae s.s. GSTe2.*

P= purified *GSTe2* (2.47µg each), F= Flow through, containing unbound proteins (5µl), C= crude cell lysate (5µl) after induction. Refer to Table 6.2 for information on protein variants. NOTE. A second induction and purification of the ZAN/U 1C variant was pursued to achieve protein expression more comparable to the Kisumu variants.

Table 6.3. Recombinant *GSTe2* protein concentration and specific activity with substrate CDNB.

For details of variable residues see Table 6.2. Protein concentrations were determined using a commercial assay (Fluka – Sigma-Aldrich) based on Bradford assay chemistry (1976). CDNB activity was determined by colorimetric assay and spectrophotometric readings (Appendix 32).

	Kisumu 1B	Kisumu 2B	ZAN/U 1C
Protein conc. (mg/ml)	2.7	2.3	4.6
CDNB activity (µmoles/ml/min)	42.70	49.06	32.66
Specific activity (µmoles/mg)	15.85	21.33	7.10

6.4.5 DDTase assay

Initial DDTase assays were performed to determine optimal conditions for kinetic analysis of each variant *GSTe2* enzyme with a substrate (DDT) dilution series (Appendix 34). Standard curves for DDT, DDE and spike-in Dicofol were produced (Appendix 33) to determine chemical concentration from HPLC peak areas (Figure 6.8).

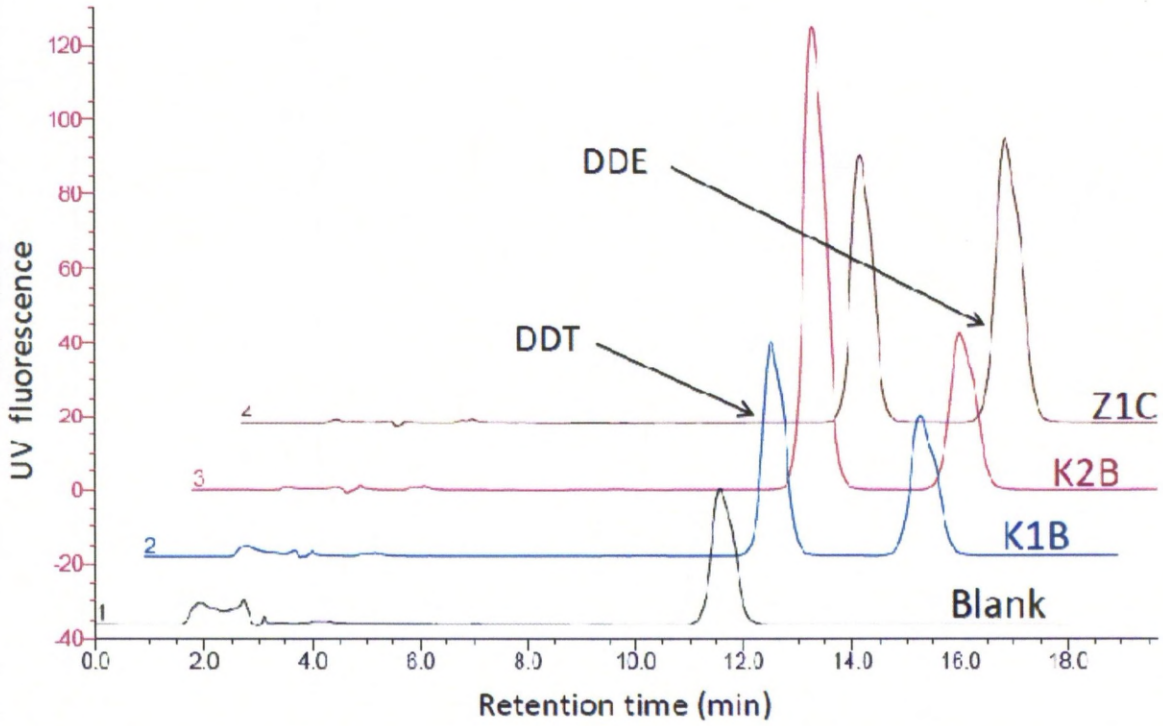

Figure 6.8. Example High Performance Liquid Chromatography (HPLC) traces from preliminary DDTase assays comparing three variant *GSTe2* recombinant enzymes.

K1B refers to Kisumu 1B variant, K2B = Kisumu 2B, Z1C = ZAN/U 1C, for details of variant amino acids refer to Table 6.2. The Blank contains DDT, GSH and reaction buffer in absence of *GSTe2* protein. *NOTE. These initial traces do not contain a Dicofol spike in peak.* For assay protocol see Appendix 34.

Three replicates per *GSTe2* variant enzyme per DDT dilution were performed and mean enzyme rates (nmol DDE/mg DDT/min) and specific activities (nmol DDE/µg *GSTe2*) calculated (Table 6.4.) At lower concentrations, 8.82-35.26µM/ml, all three variant enzymes displayed comparable activity (Table 6.4, Figure 6.9). However at DDT concentrations of 70.52µM/ml (fourth dilution in Figure 6.9) and above the ZAN/U *GSTe2* protein displayed a significantly higher mean enzyme rate compared to both Kisumu variants. This difference was most pronounced at the highest DDT concentration, 564.19µM/ml, where the mean rate for the ZAN/U enzyme was almost three times that of the two Kisumu proteins.

Table 6.4. Mean activity of three variant *GSTe2* enzymes with substrate DDT.

A dilution series of insecticide DDT was employed and three replicate DDTase assays performed with each enzyme at each DDT concentration from which the mean enzyme rate and standard error (STD Error) were calculated. Specific activity per µg of recombinant *GSTe2* enzyme was calculated using the total amount of enzyme in a single assay (60µg).

		Kisumu 1B			**Kisumu 2B**			**ZAN/U 1C**		
DDT µg/ml	DDT µM/ml	Mean rate nmol/mg/min	STD Error	Specific Activity nmol DDE/ µg *GSTe2*	Mean rate nmol/mg/min	STD Error	Specific Activity nmol DDE/ µg *GSTe2*	Mean rate nmol/mg/min	STD Error	Specific Activity nmol DDE/ µg *GSTe2*
3.125	8.82	1.55	0.06	0.09	1.70	0.15	0.10	1.56	0.21	0.09
6.25	17.63	3.75	0.02	0.23	4.29	0.09	0.26	4.17	0.21	0.25
12.5	35.26	7.89	1.07	0.47	7.28	0.14	0.44	8.60	0.53	0.52
25	70.52	8.78	0.89	0.53	9.87	0.41	0.59	12.36	1.46	0.74
50	141.05	9.73	1.03	0.58	12.05	1.29	0.72	15.23	0.84	0.91
100	282.10	8.04	1.01	0.48	10.78	1.08	0.65	18.64	1.26	1.12
200	564.19	7.09	0.66	0.43	7.73	1.86	0.46	20.36	2.24	1.22

Kisumu variant 1B exhibited lower activity compared to 2B between DDT concentrations of 70.52 and 564.19µM/ml; however this difference was only significant at 282.10µM/ml with the rates comparable at the highest concentration. The DDTase activity of the two Kisumu proteins appeared to peak at a DDT concentration of 141.05µM/ml after which both proteins began to lose activity. The mean enzyme rate for both Kisumu *GSTe2* alleles at the top DDT concentration (564.19 µM/ml) dropped to a level similar to that recorded at a substrate concentration of 35.26µM/ml (Table 6.4, Figure 6.9). The ZAN/U recombinant *GSTe2* on the other hand, showed a continued rise in enzyme rate with activity yet to reach an obvious plateau (saturation point) even at the highest DDT concentration, though the rate appeared to have slowed.

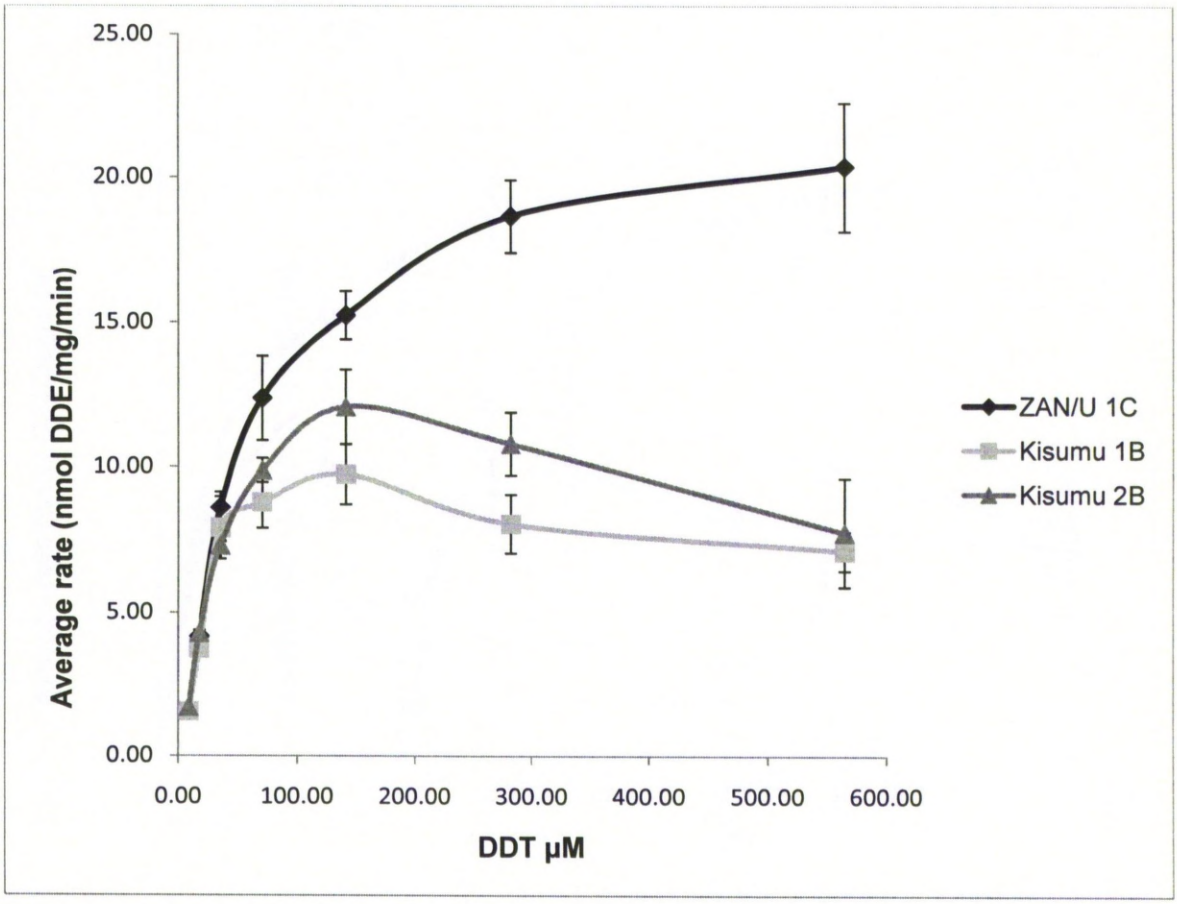

Figure 6.9. Comparison of *GSTe2* catalysed DDT metabolism for three variant recombinant proteins over a DDT dilution series.

Three allelic variants of enzyme *GSTe2* from *An. gambiae s.s.* (Table 6.2) are compared over a range of DDT concentrations and the mean production of DDE plotted from three replicate assays.

From the enzyme activity plots, kinetic parameters V_{max} (maximum enzyme rate) and K_M (substrate concentration at half maximum velocity) for each enzyme were calculated (Figure 6.3, Table 6.5). The maximum enzyme rate for the ZAN/U allele

was more than double that of the two Kisumu *GSTe2* alleles with the K_M more than three times higher in the DDT resistant strain.

Table 6.5. Enzyme kinetic parameters of three *GSTe2* alleles with substrate DDT.

Three variant *GSTe2* proteins (Table 6.2) were expressed from a DDT resistant (ZAN/U) and susceptible (Kisumu) strain of *An. gambiae s.s.* and assayed with substrate DDT over a range of concentrations. The maximum enzyme rate (V_{max}) and substrate concentration at half the maximum rate (K_M) were calculated for each protein from a Michaelis-Menten plot (Figure 6.9).

	Kisumu 1 B	Kisumu 2 B	ZAN/U 1C
K_M	16.7	21.3	66.4
V_{max}	9.2	11.1	22.9

6.5 Discussion

When the expression of gene *GSTe2* in the DDT resistant ZAN/U colony was compared with that of the susceptible Kisumu strain, expression level differences previously reported (Ding *et al.*, 2003; Ranson *et al.*, 2001) were not reproduced. The lower fold change recorded in combination with the identification of segregating non-synonymous polymorphisms within *GSTe2* suggested the potential involvement of allelic variation in the resistance phenotype.

A number of polymorphisms were identified within the *GSTe2* cDNA sequence which conferred amino-acid changes within the encoded protein. When these variant residues were mapped onto the crystal structure of the protein two mutant positions were found to be in close proximity to the putative DDT binding domain. Variation at position 120 was predicted to have implications for the formation of a hydrophobic pocket cap, while I114T may have marked effects on the conformation of the DDT binding pocket (Olga Mayans, Dan Rigden pers. comm.). Mutated/polymorphic GSTs are associated with a number of biological consequences in humans, potentially increasing the susceptibility to diseases such as cancer (Hayes and Strange, 2000; McIlwain *et al.*, 2006). While in mosquitoes Wongtrakul *et al.* (2003) demonstrated via site directed mutagenesis that two amino-acid substitutions were adequate to confer marked differences in activity parameters V_{max} and K_M for an *Anopheles dirus* GST, *GSTD3-3*, with five different substrates compared to wild type protein. The mutations within *GSTD3-3* were situated *outside* of the active site but produced cumulative effects, with the authors concluding that small changes in tertiary structure play an important role in modulating enzyme activity.

Moreover, it was discovered that the ZAN/U colony originally studied possessed an asparagine residue at position 114 (Ranson *et al.*, 2001; Ding *et al.*, 2003; Ortelli *et al.*, 2003) compared to a threonine in the colony employed here. A change from N>T at position 114 was predicted to place a hydrophobic residue within the hydrophobic DDT

binding pocket. An asparagine at 114, however, would have placed a non-hydrophobic amino-acid at this hydrophobic position and may have resulted in unpredictable rearrangement of this protein region with knock on effects for the positioning of residues within the binding pocket and cap (Dan Rigden pers. comm.). Without a crystal model of the original ZAN/U allele we cannot fully predict the outcome of N114. However in combination with reduced *GSTe2* expression levels and segregating allelic variation it is tempting to suggest that mutations found in the current ZAN/U strain confer a *GSTe2* enzyme with higher DDTase activity, which requires less protein expression to provide the same resistance phenotype. To test this directly the 'original' *GSTe2* sequence should be expressed and the DDTase activity compared to that of the allele expressed in this study.

Elevated expression of detoxifying enzymes is commonly linked to insecticide resistance in mosquitoes (David *et al.*, 2005; Djouaka *et al.*, 2008; Lumjuan *et al.*, 2005; Amenya *et al.*, 2008; Muller *et al.*, 2008b; Ranson *et al.*, 2001) as well as other insect species (Wang *et al.*, 1991; Le Goff *et al.*, 2003; Vontas *et al.*, 2002). However there is also arguably an important role for enzyme mutation conferring or enhancing the resistance phenotype.

One of the earliest examples of a causal link between amino-acid substitution and insecticide resistance involves mutated ali-esterases associated with hydrolysis of aliphatic esters and organophosphate resistance (Townsend and Busvine, 1969; Oppenoorth and van, 1960; Newcomb *et al.*, 1997; Campbell *et al.*, 1998). In the sheep blowfly, *Lucilia cuprina,* Newcomb *et al.* (1997) highlighted a *G137D* substitution within carboxylesterase gene *E3* which conferred broad-spectrum organophosphate (OP) hydrolase activity towards dimethyl OPs, while simultaneously reducing normal ali-esterase activity. The mutation, predicted to be situated within the active site, was found to correlate with resistance across 15 strains, while recombinant expression and enzyme assay confirmed the involvement of the *G137D* mutation . Subsequently a second mutation was identified within the same gene, *Y251L*, which conferred broad-spectrum 'malathion' resistance to diethyl OPs in the blowfly (Campbell *et al.*, 1998). The *G137D* mutation has since been found to confer OP resistance in the housefly *Musca domestica* (Claudianos *et al.*, 1999).

The activity of the ZAN/U allele with synthetic GST substrate CDNB was lower than that of the two susceptible alleles. This could be akin to the mutated ali-esterase example whereby the function of the wild-type protein is sacrificed for insecticide metabolising properties. However this is speculative at present and the function of the ZAN/U *GSTe2* protein with substrates likely to be metabolised *in vivo* must be examined.

In metabolic assays with substrate DDT, the ZAN/U *GSTe2* allele exhibited greater maximum enzyme velocity (V_{max}) compared to variants isolated from the susceptible Kisumu strain. Mao *et al.* (2007) recorded a higher V_{max} in a mutated P450 enzyme which was associated with a higher rate of plant toxin metabolism in the Parsnip web worm, *Depressaria pastinacella*, compared to the wild-type enzyme. The difference in activity was attributed to a single amino-acid change, *V95A*, which affected the interaction with P450-reductase.

The ZAN/U *GSTe2* protein also conformed to Michaelis-Menten enzyme kinetics producing a rectangular hyperbola saturation curve in which the enzyme rate rose rapidly with substrate concentration before reaching a plateau (Figure 6.3). The Kisumu alleles, on the other hand, exhibited non-Michaelis-Menten kinetics. After initially rising with substrate concentration, enzyme rate for both Kisumu proteins peaked before activity declined at higher substrate concentrations (Figure 6.9); this profile is typical of enzymes exhibiting substrate inhibition (Figure 6.10). Lin *et al.* (2001) examined cytochrome P450 substrate inhibition in 13 recombinant human P450 enzymes and found 6 displayed characteristic inhibition curves (Figure 6.10). They proposed a kinetic model for the inhibition assuming that these enzymes carried two substrate binding sites; a catalytic and an inhibitory site. A similar hypothesis was suggested by Vincent *et al.* (2005) when examining glucosamine-6-phosphate deaminase substrate inhibition with a second bound glucosamine 6-phosphate molecule potentially obscuring the catalytic binding site.

It is possible that the putative substrate inhibition witnessed in the Kisumu alleles was caused by a secondary DDT binding. A DDT molecule may have bound to a non-catalytic site obscuring the catalytic binding pocket or potentially interfered with GSH binding at the G site. An alternative explanation may be the higher affinity for DDT in the Kisumu alleles suggested by the lower K_M recorded. A higher DDT affinity suggests the enzyme-substrate complex (ES) forms more readily and the enzyme becomes saturated at a lower concentration. However the product formation (ES $\rightarrow$ PE) and release of free enzyme may occur slowly, with the Kisumu alleles less likely to release bound DDT or product DDE. In contrast the ZAN/U allele, with lower DDT affinity (K_M), easily releases DDT and potentially DDE and is therefore able to convert more substrate to product at higher substrate concentrations, without becoming saturated. The exact cause of this inhibition is as yet unknown and requires further investigation of the binding kinetics and potential for a secondary DDT docking site. Isothermal titration calorimetry would provide quantitative information on the binding affinity of *GSTe2* alleles for substrate DDT. This technique relies on detection of changes in the temperature derivative of Gibbs energy caused by bonding rearrangements when a protein moves from a free to substrate-bound state (Ladbury and Chowdhry, 1996), and is commonly used to explore enzyme-substrate interactions (Olsen, 2006; Feng *et al.*, 2004). The potential for a secondary inhibitory DDT interaction could be explored via further computational modelling or co-crystallisation of DDT with the *GSTe2* variants.

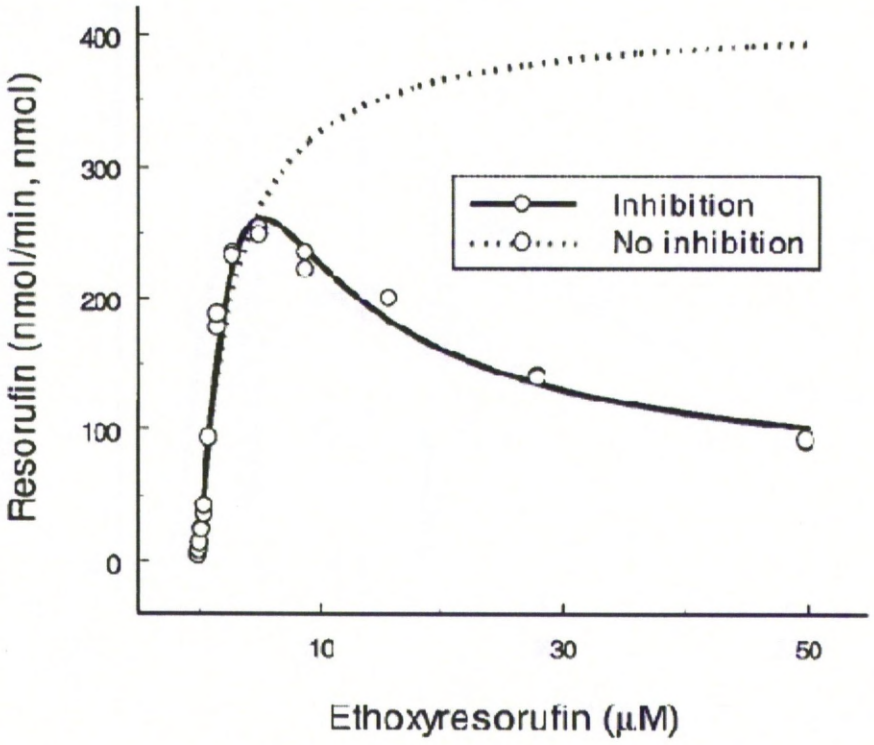

Figure 6.10. Substrate inhibition of *CYP1A2*-catalysed O-deethlation of ethoxyresorufin.

Figure taken from Lin et al. (2001).

Allelic variation with *GSTe2* appears to have profound effect on catalytic activity with substrate DDT *in vitro*. However to be relevant to resistance in the field we must consider the implications of these results *in vivo*. The physiological concentrations of both *GSTe2* and DDT on insecticide exposure are fundamental factors which will influence the enzyme kinetics *in vivo* and must be investigated to infer biological consequences of results presented here. The current DDT spraying regime recommended by WHO delivers 100-200μg DDT/cm^2 (WHO, 2006b), which is within the range of the top DDT concentration employed in assays. However the concentration of DDT found *in vivo* is likely to be several orders of magnitude lower as the insect cuticle acts as a major barrier to the uptake of lipophilic insecticides. Hence the concentration of DDT used in the assays was greater than those present in a single mosquito. The use of labelled insecticide which can be traced within the mosquito on exposure would provide indication on the physiological DDT levels and localisation within the insect. Each assay also contained 60μg of recombinant *GSTe2* protein. Accurate western blot analysis may provide a measure of total *GSTe2* content in a mosquito; however localisation of the expressed protein within the insect is fundamental as this may increase enzyme concentration in the presence of DDT. Immune-staining of GSTe2 *in vivo* in combination with a labelled insecticide may provide information on the potential co-localisation of enzyme and insecticide on exposure.

The role of allelic variation in conferring DDT resistance has been investigated in the cytochrome P450 enzyme family. A *Drosophila melanogaster* DDT resistance strain was found to possess a P450 enzyme, *CYP6A2*, containing three amino-acid changes compared to the wild-type peptide (Amichot *et al.*, 2004). Mutant enzymes possessing one, two or all three altered residues were expressed and DDTase activity compared to wild-type protein. The triple mutant was found to have enhanced DDT metabolising activity, however all mutant enzymes were less stable than the wild-type protein. The *An. gambiae* P450s *CYP6Z1* and Z2 genes were among five up-regulated in the DDT resistant ZAN/U strain (David *et al.*, 2005). When computationally modelled with DDT very different metabolic profiles were predicted despite high sequence identity (~ 70% amino acid ID). *CYP6Z1* was predicted to possess a larger substrate binding site compared to Z2, which could accommodate the DDT molecule (Chiu *et al.*, 2008). This prediction was confirmed via recombinant expression and DDTase assay in which only *CYP6Z1* was able to metabolise this insecticide. This result may suggest a role for *CYP6Z1* in conferring DDT resistance in the ZAN/U colony investigated here, however this needs to be investigated via expression and functional level analysis.

It is clear from previous studies and results presented here that allelic variation within detoxifying enzymes may have a role in conferring or enhancing insecticide resistance. Thorough investigation of metabolic resistance should not rely solely on detection of elevated enzyme levels through biochemical assays (WHO, 1998a) or microarray studies (David *et al.*, 2005; Muller *et al.*, 2007). Systematic analysis of enzymes implicated for potential allelic variants, and where possible modelling and protein expression should also be performed. The advent of next generation sequencing for expression analysis and a number of large scale SNP discovery endeavours (Wilding *et al.*, 2009; Neafsey *et al.*, 2010) will hopefully provide more information on polymorphisms which segregate with phenotype.

6.5.1 Future perspectives

To establish whether the mutations highlighted here are having the predicted effect on DDT binding these variants should be pursued using protein crystallography. Wang *et al.* (2008), as discussed, resolved the crystal structure for a Kisumu *GSTe2* allele which matched clone 2B at sites 114 and 120. However they were unable to co-crystallise with DDT, and so resorted to computational modelling of this interaction. To understand how this protein and substrate interact, co-crystallisation or soaking of crystallised proteins with a DDT solution should be performed. Although GSH binding did not affect the conformation of the protein crystallized by Wang *et al.* (2008), it will be important to crystallise with this co-factor as the effect of glutathione binding on conformation of variant proteins identified in this study is unknown.

To confirm an *in vivo* role in resistance, over-expression of the variant alleles could be pursued in a transgenic system such as the UAS/GAL4 (Brand and Perrimon, 1993) or piggyBac system (Grossman *et al.*, 2001). The UAS/GAL4 system has been successfully employed in *Drosophila* to evaluate the potential for eight cytochrome P450 enzymes to conferring insecticide resistance (Daborn *et al.*, 2007).

While *An. gambiae* proteins have also been expressed and investigated using this system in *Drosophila* (Walker *et al.*, 2008).

Finally, here we have focused on the implications of non-synonymous SNPs within the protein *GSTe2*. However there is increasing evidence that synonymous SNPs are not inconsequential as previously thought and can in fact alter mRNA structure and stability as well as protein folding and activity (Hunt *et al.*, 2009). A synonymous SNP in the *Multidrug Resistance 1* gene product, P-glycoprotein, altered drug and inhibitor interacts without an alteration in amino acid sequence; it was hypothesised that the resultant rare codon influenced translation rates and thus protein folding, altering the active sites (Kimichi-Sarfaty *et al.*, 2007). While Duan *et al.* (2003) noted that mRNA stability and translation of the human dopamine receptor D2 was increased in the presence of a specific synonymous SNP; however the presence of a second synonymous SNP, which had no influence alone, voided the effects of the first SNP illustrating an additive effect. The consequences of synonymous SNPs, though potentially harder to quantify, should also be considered when investigating allelic variations in the future.

Chapter 7: Conclusions

7.1 Whole genome expression profiling of insecticide resistant populations

Transcriptional profiling of insecticide resistant *An. gambiae s.s.* from the field provided a number of interesting gene candidates potentially implicated in the resistance phenotype. This work is unique in that it is the first time whole genome microarray analysis has been applied to the study of insecticide resistance in wild caught mosquito populations. Previously, microarray approaches for studying insecticide resistance have relied upon small scale candidate gene arrays, often employing laboratory colonised mosquitoes. The candidate approach was validated, in part, by this study as a number of detoxification genes previously associated with resistance were implicated. However in addition, a number of novel candidate genes and protein families were implicated in both DDT and bendiocarb resistance. These genes would have been missed if a 'detox only' candidate gene approach had been taken.

7.1.1 Microarray candidates

7.1.1.1 Novel candidates

Cuticular genes were up-regulated in both DDT and bendiocarb resistant populations with *CPLC4* and *CPLC17* up-regulated in bendiocarb resistant samples and *CPR9* over expressed in DDT selected mosquitoes. Previous studies have suggested a role for the cuticle in resistance by reducing the ability of insecticides to reach their target-site (Puinean *et al.*, 2010; Wood *et al.*, 2010; Plapp and Hoyer, 1968). The identification of cuticular genes in this study supports the involvement of cuticle modification in the resistance phenotype and this neglected area of resistance research should be addressed in future work. The experiments of Puinean *et al.* (2010) illustrate how microarray identification and subsequent use of radiolabelled insecticides can implicate cuticular genes *in vivo*.

Chemosensory networks were potentially implicated in the bendiocarb resistant samples with up regulation of three odorant binding proteins (OBPs) as well as a putative OBP receptor. One of these genes, *OBP3*, was subsequently found to interact with bendiocarb in recombinant protein assays. In addition, *OBP3* was also up-regulated in a carbamate resistant laboratory colony employed as a control in the DDT microarray study, which may further support involvement of this protein in bendiocarb resistance. Furthermore, three OBPs were up-regulated ~ 1.5 fold in the DDT resistant group with a speculated role for arrestin gene *ARR2* in modulating OBP receptor response. Again little is known about the detection and movement of insecticide volatiles *in vivo*. These proteins may be involved in signalling pathways which instigate behavioural and/or metabolic responses on insecticide exposure. Alternatively 'OBPs' may have a role in trafficking compounds such as insecticides to specific organs *in vivo* where metabolic breakdown takes place. The term odorant binding protein is perhaps a misnomer as proteins from this class have been shown to

interact with non-odorant substrates (Hooper *et al.*, 2009; Pelosi *et al.*, 2006; Li *et al.*, 2008) as well exhibit expression outside of the sensory appendages (Li *et al.*, 2008; Galindo and Smith, 2001), which may suggest involvement in general transport of molecules *in vivo*.

The current poor annotation of a number of *An. gambiae s.s.* genes in Ensembl hindered the identification and pursuit of microarray candidates; although issues with annotation are likely to be a problem in a number of animal systems including *Ae. aegypti*, where fewer contigs have been successfully mapped (D. Neafsey, Broad institute, pers. comm.). The identification of currently un-annotated genes in separate resistance studies provides good evidence for involvement in phenotype and may support functional investigation; *e.g.* the putative membrane transporter also identified as up-regulated in DDT resistance mosquitoes from a separate study (C. Jones, pers. comm.). Ultimately, improved annotation will facilitate the functional pursuit of genes implicated in the resistance phenotype.

7.1.1.2 Cytochrome P450s

Cytochrome P450s have been frequently associated with insecticide resistance (Daborn *et al.*, 2002; Djouaka *et al.*, 2008; Le Goff *et al.*, 2003; Muller *et al.*, 2007; Muller *et al.*, 2008b; Nikou *et al.*, 2003). Members of this large protein family were represented in genes over-expressed in field collected *An. gambiae s.s.* phenotyped for both bendiocarb and DDT resistance. However high sequence similarity between closely related genes, often arranged in chromosomal clusters, may have confounded expression results in both microarray experiments. For instance the *CYP325C1-3* family up-regulated in bendiocarb resistant samples share 95-96% identity, suggesting cross probe hybridisation was likely in array experiments. This similarity also prevented single locus primer design for subsequent qPCR. A similar situation was suspected for probes designed against DDT microarray candidate *CYP6Z3* and closely related gene *CYP6Z2* (94% ID). Sequence similarity is likely to be a blanket issue in large gene families, although probe design was directed towards areas of divergence in an attempt to minimise cross-hybridisation. Next generation sequencing approaches for gene expression analysis may hopefully circumvent these issues; however reads would need to bridge regions of divergence within areas of high identity to correctly assign sequences to each closely related gene, which may be more difficult when short reads are generated.

Gene duplication is likely to have generated many of these closely related P450s and provides broad substrate specificity for these proteins *in vivo* (Feyereisen, 1999). Moreover, it is likely that closely related proteins may have similar detoxification profiles. However, some empirical evidence *in vitro* suggests that only a few amino acid differences can affect the metabolic profile of closely related proteins (Mao *et al.*, 2007; Amichot *et al.*, 2004; Wongtrakul *et al.*, 2003). Recombinant *CYP6Z1* for instance was able to metabolise DDT while *CYP6Z2* which shares 70% amino acid identity was unable to break down this insecticide (Chiu *et al.*, 2008).

In mosquitoes, elevated expression of P450s has been most commonly linked to pyrethroid resistance (Muller *et al.*, 2007; Wondji *et al.*, 2009; Komagata *et al.*, 2010;

Nikou *et al.*, 2003; Muller *et al.*, 2008b; Djouaka *et al.*, 2008). This is likely to be due, in part, to a large proportion of studies focusing on pyrethroid resistance, as this is the only class permitted for ITN use, as well as being employed extensively in IRS. P450 up-regulation in pyrethroid resistant insects may also be linked to underlying resistance to other compounds; multiple insecticide resistance has been previously reported in other mosquito populations (Corbel *et al.*, 2007; Dabire *et al.*, 2008; Perera *et al.*, 2008), and is present in the Ghanaian population sampled in this study (Chapter 2). Furthermore, a single P450 may be involved in resistance to more than one compound, a phenomenon documented in *Drosophila* where the P450 *CYP6G1* is associated with resistance to DDT, neonicotinoids, and organophosphate malathion (Daborn *et al.*, 2002; Joußen *et al.*, 2008; Le Goff *et al.*, 2003). Similarly, cytochrome P450s have also been shown to display broad substrate range with non-insecticidal xenotoxins (Li *et al.*, 2004; McLaughlin *et al.*, 2008). In *An. gambiae CYP6Z1* was found to be up-regulated in both permethrin resistant (RSP) and DDT resistant (ZAN/U) colonies via microarray (David *et al.*, 2005), while *CYP6M2* was up regulated in the DDT resistant field samples profiled here as well as in pyrethroid resistant mosquitoes (Muller *et al.*, 2007; Djouaka *et al.*, 2008). As yet, the ability of these P450 proteins to metabolise both DDT and pyrethroids has not been assessed.

The evidence (Daborn *et al.*, 2002; Joußen *et al.*, 2008; Li *et al.*, 2004; McLaughlin *et al.*, 2008) suggests that P450 enzymes can interact with numerous substrates *in vivo* with the reported high SNP density within this gene family (Wilding *et al.*, 2009) supportive of substrate plasticity. These enzymes play a key role in detoxification of xenobiotics which may be naturally derived such as plant toxins or synthetic such as insecticides introduced into the mosquito habitat. The answer to the problem of resistance may not be to knock out or inhibit single P450 genes as with over 100 P450s in *An. gambiae* (Ranson *et al.*, 2002) it is likely another broad spectrum P450 or closely related enzyme can take the place of an inactivated protein. However the identification of resistance linked P450 enzymes is important for screening new insecticidal compounds so potential routes of metabolic breakdown can be identified. In addition it is essential to identify markers of resistance which can be screened for in field populations to provide indication of resistance development and spread.

7.1.2 Field versus colony comparison

Characterising insecticide resistance associated gene expression in field material is preferable over the use of artificially selected laboratory colonies. Insects selected for resistance in the lab are likely to have undergone population bottlenecks, inbreeding, and exposure to conditions unrepresentative of the wild. Muller *et al.*(2007) reported ~5 fold over expression of P450, *CYP6M2*, in mosquitoes colonised from Ghana and selected for pyrethroid resistance in the laboratory, when compared to a susceptible strain from East Africa. However when wild caught mosquitoes from the same region of Ghana were selected using a pyrethroid LT_{50}, and compared to unexposed controls from the same area, *CYP6M2* fold difference was just 1.29, with a putative fold change of 1.64, when adjusted for the presence of resistant individuals in the controls (2008b). Although *CYP6M2* was still significantly over-expressed in the later study ($p= 6.28x10^{-07}$). Larger fold differences may be expected when comparing an insecticide selected

colony to a fully susceptible lab strain, which is supported, in part, by the greater fold changes witnessed in the DDT microarray experiment compared to the bendiocarb results. This highlights the importance of considering expression ratios of less than 2 fold, which show statistical significance, when comparing field material in a test-control scenario; p-value ranking rather than an arbitrary fold change cut off was employed by Muller *et al.* (2008b) for candidate selection, and was also utilized in this study.

Of perhaps more concern was the inability of Muller's study on colonised, pyrethroid selected mosquitoes (2007) to identify a potentially key resistance associated gene, *CYP6P3*. This gene was subsequently identified as over-expressed in field collected mosquitoes from the same area, and was able to metabolise pyrethroids in recombinant protein assays (Muller *et al.*, 2008b). It is possible that elevated expression of this P450 was lost during the colonisation process; although an increase in *CYP6P3* expression levels in the Ghanaian population during the intermittent years between studies must also be considered.

Characterisation of resistance in wild-caught insects is potentially more informative for vector control in the field. However, the use of colony control material in this study enabled characterisation a highly (DDT) resistant field population, in absence of a susceptible sympatric population and where un-exposed controls from the area were an unsuitable comparison. Though not ideal, every effort was made to reduce the variation in the experiment associated with genomic background, geographical origin, sample treatment, storage and extraction. Moreover, the use of two independent control colonies provided more confidence in results, as candidate genes in the resistant field samples were selected on the basis of being significantly over-expressed compared to both controls.

Subsequent candidate follow-up will confirm whether genes identified in this field versus colony experiment are indeed involved in the resistance phenotype. However the identification of two gene candidates: firstly a P450 already linked to insecticide resistance (*CYP6M2*) (Muller *et al.*, 2007; Djouaka *et al.*, 2008)(M. Paine, pers. comm.); and secondly, a putative membrane transporter independently identified as over-expressed in DDT resistant field material, suggests this design was effective in identifying resistance associated expression.

7.1.3 Pathways & regulation

The use of whole genome arrays, as opposed to previous candidate gene approaches, has the potential to provide information on the networks and pathways involved in the resistance phenotype. Examining signatures of co-regulation within a data set can highlight networks of genes which are interacting to produce a final gene product or phenotype. However this type of analysis requires well annotated genomes, knowledge of pathways, as well as large experimental data sets to provide the power to identify pathways from thousands of genes. By combining results from experiments characterising expression related to the same resistance phenotype, a meta-analysis could be performed to identify common pathways.

Understanding networks and pathways involved in resistance will provide new targets for insecticide synergists as well as for control via genetic manipulation. Targeting endpoint products such as a single P450 or GST is unlikely to be a successful control option as alterative enzymes are likely to take over the role. However through network analysis, hub proteins which interact with and are essential for the function of a number of proteins maybe targeted. Alternatively, if identifiable, regulatory proteins such as transcription factors which drive the expression of key genes may be inhibited.

Cytochrome P450 reductase (CPR) is an example of a hub protein which interacts with numerous cytochrome P450 enzymes providing electrons which are essential for protein function (Paine *et al.*, 2005). In mosquitoes a single CPR provides electrons for over 100 P450 proteins, therefore by inactivating this gene P450 based resistance could be eliminated from the population. Indeed Lycett *et al.* (2006) reported increased susceptibility to pyrethroids when the CPR gene was knocked down by between 50-90% in various tissues of *An. gambiae*. The potential for such hub proteins to control activity of numerous peptides linked to a trait of interest may warrant a shift of focus from the transcriptome and proteome to the 'interactome' (Lehner *et al.*, 2006; He and Zhang, 2006).

7.1.4 Tissue specific expression

Expression profiling of insecticide resistant material to date, has mainly involved whole insect extracts (Muller *et al.*, 2008b; Strode *et al.*, 2006; Djouaka *et al.*, 2008; David *et al.*, 2005; Pedra *et al.*, 2004; Komagata *et al.*, 2010). This can partly be attributed to the small amounts of RNA produced from a single insect and the difficulties involved in dissecting specific tissues to provide sufficient material without degrading RNA, especially under field conditions. However the localised expression of genes in specific tissues is potentially integral to the resistance phenotype and by uncovering these patterns we may be able to piece together the movement and breakdown of insecticides *in vivo*. The focus of metabolism can then be targeted by gene knockdown or tissue specific transgenic expression of metabolic inhibitors. While new insecticides could be designed which avoid these 'hot spots' of metabolism or synergists employed which accumulate in these tissues. Moreover, when the whole organism is subjected to expression profiling, highly expressed genes potentially unlinked to resistance may mask localised gene expression (Wang *et al.*, 2004). The improvement of microarray labelling techniques, which can now require as little as 20ng of starting material, in combination with better RNA preservatives will make tissue specific expression profiling a more feasible option.

Tissue specific profiling in *Anopheles* has been applied to the study transcription post bloodmeal; Marinotti *et al.* (2006) employed a *Plasmodium/Anopheles* genome array to evaluate accumulation of specific transcripts in the fat body, midgut and ovaries of female mosquitoes. While Wang *et al.* (2004) performed tissue specific microarray analysis of the Malpighian tubules in *Drosophila* and found the transcriptome differed significantly from that obtained using whole flies; 307 genes were present only in the tubule arrays and enzyme families putatively linked to insecticide resistance were enriched over 25 fold in the tubules including GSTs and P450s. Preliminary data from

An. gambiae also support expression of metabolic enzymes predominantly in the Malpighian tubules with *GSTe2* and *CYP6M2* over-expressed in this tissue (P. Muller personal communication).

7.1.5 Secondary metabolism

Another consideration yet to be addressed in published metabolism studies is the potential for proteins to break down the insecticide metabolites produced by other enzymes. This area of metabolic resistance is far more complex and requires understanding of the metabolic profiles produced *in vivo* by detoxifying enzymes. *In vitro* techniques such as HPLC may provide information on insecticide metabolites formed by recombinant enzymes and these metabolites, once identified, can be assayed with other proteins. Such an approach has been performed at LSTM on metabolites of P450 mediated pyrethroid break down (B. Stevenson pers. comm.). The potential role for enzymes in secondary metabolism of insecticides suggests that an absence of direct compound break down in recombinant assays should not rule out involvement in the detoxification process.

7.2 Allelic variation

Allelic variation within genes putatively linked to insecticide resistance was a topic highlighted in both the microarray and recombinant *GSTe2* studies. The higher SNP density within enzyme families commonly linked to resistance compared to elsewhere in the genome (Wilding *et al.*, 2009) has potential to enhance substrate plasticity. This allelic variation also potentially confounded expression analysis using probes designed to a single allele from the genome database (Ensembl), as the levels of polymorphism in field material was unknown. SNP density within candidate genes on sequencing also restricted the positioning of primers for subsequent qPCR.

As highlighted in previous studies (Amichot *et al.*, 2004; Newcomb *et al.*, 1997; Mao *et al.*, 2007) these polymorphisms have the potential to affect the metabolic activity of enzymes. This was found to be the case when allelic variation within gene *GSTe2* was explored using a recombinant protein system and *in vitro* assay with insecticide DDT (Chapter 6).

Previous studies of insecticide resistance have relied heavily on analysis of elevated enzymes linked to resistance through biochemical assays or microarray expression profiling. However the potential for allelic variation within genes linked to resistance to affect phenotype has been somewhat overlooked. This is possibly a result of the time and cost involved in identifying candidate genes then performing extensive sequencing of individual samples to establish whether SNPs correlate with phenotype.

However the extensive SNP discovery efforts in *An. gambiae* and subsequent SNP array development (Weetman *et al.*, 2010; Neafsey *et al.*, 2010) has enabled mapping of genomic differentiation between populations. Moreover, the advent of next-generation sequencing will provide a technique in which gene expression and allelic variation can be concomitantly measured within the same samples. This will hopefully

provide much needed information on variation within resistance linked genes which can then be explored further.

References.

Abbott, W. S. (1925) A method of computing the effectiveness of an insecticide. Vol. 18, pp. 265 - 267.

Abercrombie, J. M., Halfhill, M. D., Ranjan, P., Rao, M. R., Saxton, A. M., Yuan, J. S. and Stewart, C. N., Jr. (2008) Transcriptional responses of Arabidopsis thaliana plants to As (V) stress. BMC Plant Biology, 8, 87.

Achee, N. L., Sardelis, M. R., Dusfour, I., Chauhan, K. R. and Grieco, J. P. (2009) Characterization of spatial repellent, contact irritant, and toxicant chemical actions of standard vector control compounds. Journal of the American Mosquito Control Association, 25, 156-167.

Achonduh, O. A., Gbewonyo, W. S. K., Boakye, D. A. and Wilson, M. D. (2008) Susceptibility Status of Anopheles gambiae s.l. (Diptera:Culicidae) from Cabbage Growing Areas Associated with Pyrethroid and Organophosphate Use in Accra, Ghana. West African Journal of Applied Ecology, 12.

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

Agilent (2008) Two-Color Microarray-Based Gene Expression Analysis (Quick Amp Labelling). Agilent Technologies, Santa Clara, California, USA.

Ahmad, M., Denholm, I. and 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.

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

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. (1990) Basic local alignment search tool. Journal of Molecular Biology, 215, 403-410.

Amenya, D. A., Naguran, R., Lo, T. C., Ranson, H., Spillings, B. L., Wood, O. R., Brooke, B. D., Coetzee, M. and 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.

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

Andrieu-Abadie, N., Gouazé, V., Salvayre, R. and Levade, T. (2001) Ceramide in apoptosis signaling: relationship with oxidative stress. Free Radical Biology and Medicine, 31, 717-728.

Ansede, J. H. and Thakker, D. R. (2004) High-throughput screening for stability and inhibitory activity of compounds toward cytochrome P450-mediated metabolism. Journal of Pharmaceutical Sciences, 93, 239-255.

Anto, F., Asoala, V., Anyorigiya, T., Oduro, A., Adjuik, M., Owusu-Agyei, S., Dery, D., Bimi, L. and Hodgson, A. (2009) Insecticide resistance profiles for malaria vectors in the Kassena-Nankana district of Ghana. Malaria Journal, 8, 81.

Appawu, M., Owusu-Agyei, S., Dadzie, S., Asoala, V., Anto, F., Koram, K., Rogers, W., Nkrumah, F., Hoffman, S. L. and Fryauff, D. J. (2004) Malaria transmission dynamics at a site in northern Ghana proposed for testing malaria vaccines. Tropical Medicine & International Health, 9, 164 - 170.

Appawu, M. A., Baffoe-Wilmot, A., Afari, E. A., Nkrumah, F. K. and Petrarca, V. (1994) Species composition and inversion polymorphism of the Anopheles gambiae complex in some sites of Ghana, West Africa. Acta Tropica, 56, 15 - 23.

Asidi, A., N' Guessan, R., Koffi, A., Curtis, C., Hougard, J.-M., Chandre, F., Corbel, V., Darriet, F., Zaim, M. and Rowland, M. (2005) Experimental hut evaluation of bednets treated with an organophosphate (chlorpyrifos-methyl) or a pyrethroid

(lambdacyhalothrin) alone and in combination against insecticide-resistant Anopheles gambiae and Culex quinquefasciatus mosquitoes. Malaria Journal, 4, 25.
Awolola, T. S., Oduola, O. A., Strode, C., Koekemoer, L. L., Brooke, B. and Ranson, H. (2009) Evidence of multiple pyrethroid resistance mechanisms in the malaria vector Anopheles gambiae sensu stricto from Nigeria. Medical and Veterinary Entomology, 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. and Coetzee, M. (2005) Distribution of the molecular forms of Anopheles gambiae and pyrethroid knock down resistance gene in Nigeria. Acta Tropica, 95, 204 - 209.
Baginsky, S., Hennig, L., Zimmermann, P. and Gruissem, W. (2010) Gene Expression Analysis, Proteomics, and Network Discovery. Plant Physiology, 152, 402-410.
Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. A., Selker, E. U., Cresko, W. A. and Johnson, E. A. (2008) Rapid SNP Discovery and Genetic Mapping Using Sequenced RAD Markers. PLoS One, 3, e3376.
Bass, C., Nikou, D., Donnelly, M., Williamson, M., Ranson, H., Ball, A., Vontas, J. and Field, L. (2007) Detection of knockdown resistance (kdr) mutations in Anopheles gambiae: a comparison of two new high-throughput assays with existing methods. Malaria Journal, 6, 111.
Bass, C., Nikou, D., Vontas, J., Williamson, M. S. and Field, L. M. (2010) Development of high-throughput real-time PCR assays for the identification of insensitive acetylcholinesterase (ace-1R) in Anopheles gambiae. Pesticide Biochemistry and Physiology, 96, 80-85.
Bautista, M. A. M., Miyata, T., Miura, K. and Tanaka, T. (2009) RNA interference-mediated knockdown of a cytochrome P450, CYP6BG1, from the diamondback moth, Plutella xylostella, reduces larval resistance to permethrin. Insect Biochemistry and Molecular Biology, 39, 38-46.
Baylin, S. B. (2005) DNA methylation and gene silencing in cancer. Nature Clinical Practise Oncology, 2.
Beales, P. F. and Gilles, H. M. (2002) Rationale and techniques of malaria control. In Essential Malariology(Eds, Warrell, D. A. and Gilles, H. M.) Arnold, London, pp. 107.
Becker, N., Petric, D., Zgomba, M., Boase, C., Dahl, C., Lane, J. and Achim, K. (2003) Chemical Control. In *Mosquitoes and their control*Kluwer Academic/Plenum Publishers, London, pp. 387-393.
Beckham, J. T., Wilmink, G. J., Opalenik, S. R., Mackanos, M. A., Abraham, A. A., Takahashi, K., Contag, C. H., Takahashi, T. and Jansen, E. D. (2010) Microarray analysis of cellular thermotolerance. Lasers in Surgery and Medicine, 42, 752-765.
Belušič, G., Pirih, P. and Stavenga, D. (2010) Photoreceptor responses of fruitflies with normal and reduced arrestin content studied by simultaneous measurements of visual pigment fluorescence and ERG. Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology, 196, 23-35.
Benjamini, Y. and Hochberg, Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society, Series B (Methodological), 57, 289-300.
Benoit, J. B., Lopez-Martinez, G., Phillips, Z. P., Partrick, K. R. and Denilinger, D. L. (2010) Heat shock proteins contribute to mosquito dehydration tolerance. Journal of Insect Physiology, 56, 151-156.
Bertone, P. and Snyder, M. (2005) Prospects and Challenges in Proteomics. Plant Physiology, 138, 560-562.
Biedler, J. and Tu, Z. (2010) Evolutionary analysis of the kinesin light chain genes in the yellow fever mosquito Aedes aegypti: gene duplication as a source for novel early zygotic genes. BMC Evolutionary Biology, 10, 206.
Bilak, H., Tauszig-Delamasure, S. and Imler, J. L. (2003) Toll and Toll-like receptors in Drosophila. Biochemical Society Transactions, 31.

Billingsley, P. F. and Hecker, H. (1991) Blood digestion in the mosquito, Anopheles stephensi Liston (Diptera: Culicidae): activity and distribution of trypsin, aminopeptidase, and alpha-glucosidase in the midgut. Journal of Medical Entomology, 28, 865-871.

Blandin, S., Moita, L. F., Kocher, T., Wilm, M., Kafatos, F. C. and Levashina, E. A. (2002) Reverse genetics in the mosquito Anopheles gambiae: targeted disruption of the Defensin gene. EMBO Rep, 3, 852-856.

Blanford, S., Chan, B. H. K., Jenkins, N., Sim, D., Turner, R. J., Read, A. F. and Thomas, M. B. (2005) Fungal Pathogen Reduces Potential for Malaria Transmission. Science, 308, 1638-1641.

Board, P. G., Coggan, M., Chelvanayagam, G., Easteal, S., Jermiin, L. S., Schulte, G. K., Danley, D. E., Hoth, L. R., Griffor, M. C., Kamath, A. V., Rosner, M. H., Chrunyk, B. A., Perregaux, D. E., Gabel, C. A., Geoghegan, K. F. and Pandit, J. (2000) Identification, characterization, and crystal structure of the Omega class glutathione transferases. Journal of Biological Chemistry, 275, 24798-24806.

Boisson, B., Jacques, J. C., Choumet, V., Martin, E., Xu, J., Vernick, K. and Bourgouin, C. (2006) Gene silencing in mosquito salivary glands by RNAi. FEBS Letters, 580, 1988-1992.

Bourguet, D., Guillemaud, T., Chevillon, C. and Raymond, M. (2004) Fitness costs of insecticide resistance in natural breeding sites of the mosquito Culex pipiens. Evolution, 58, 128 - 135.

Bourguet, D., Raymond, M., Bisset, J., Pasteur, N. and Arpagaus, M. (1996) Duplication of the Ace.1 locus in Culex pipiens mosquitoes from the Caribbean. Biochemical Genetics, 34, 351-362.

Bozdech, Z., Zhu, J., Joachimiak, M., Cohen, F., Pulliam, B. and DeRisi, J. (2003) Expression profiling of the schizont and trophozoite stages of Plasmodium falciparum with a long-oligonucleotide microarray. Genome Biology, 4, R9.

Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Analytical Biochemistry, 72, 248-254.

Brand, A. H. and Perrimon, N. (1993) Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. Development, 118, 401-415.

Brooke, B. D. (2008) kdr: can a single mutation produce an entire insecticide resistance phenotype? Transactions of the Royal Society of Tropical Medicine and Hygiene, 102, 524-525.

Brooke, B. D., Hunt, R. H. and Coetzee, M. (2000) Resistance to dieldrin + fipronil assorts with chromosome inversion 2La in the malaria vector Anopheles gambiae. Medical and Veterinary Entomology, 14, 190-194.

Brooke, B. D., Hunt, R. H., Matambo, T. S., Koekemoer, L. L., Van Wyk, P. and Coetzee, M. (2006) Dieldrin resistance in the malaria vector *Anopheles gambiae* in Ghana. Medical and Veterinary Entomology, 20, 294-299.

Brooke, B. D., Kloke, G., Hunt, R. H., Koekemoer, L. L., Temu, E. A., Taylor, M. E., Small, G., Hemingway, J. and Coetzee, M. (2001) Bioassay and biochemical analyses of insecticide resistance in southern African *Anopheles funestus* (Diptera: Culicidae). Bulletin of Entomological Research, 91, 265-272.

Bruns, C. M., Hubatsch, I., Ridderström, M., Mannervik, B. and Tainer, J. A. (1999) Human glutathione transferase A4-4 crystal structures and mutagenesis reveal the basis of high catalytic efficiency with toxic lipid peroxidation products. Journal of Molecular Biology, 288, 427-439.

Burney, R. O., Talbi, S., Hamilton, A. E., Vo, K. C., Nyegaard, M., Nezhat, C. R., Lessey, B. A. and Giudice, L. C. (2007) Gene Expression Analysis of Endometrium Reveals Progesterone Resistance and Candidate Susceptibility Genes in Women with Endometriosis. Endocrinology, 148, 3814-3826.

Busvine, J. R. (1951) Mechanism of resistance to insecticides in houseflies. Nature, 168, 193-195.

Campbell, P. M., Newcomb, R. D., Russel, R. J. and Oakeshott, J. G. (1998) Two different amino acid substitutions in the ali-esterase, E3, confer alternative

types of organophosphorus insecticide resistance in the sheep blowfly, Lucilia cuprina. Insect Biochemistry and Molecular Biology, 28, 139 - 150.
Caputo, B., Nwakanma, D., Jawara, M., Adiamoh, M., Dia, I., Konate, L., Petrarca, V., Conway, D. and della Torre, A. (2008) Anopheles gambiae complex along The Gambia river, with particular reference to the molecular forms of An. gambiae s.s. Malaria Journal, 7, 182.
Carino, F. A., Koener, J. F., Plapp, F. W. and Feyereisen, R. (1994) Constitutive overexpression of the cytochrome P450 gene CYP6A1 in a house fly strain with metabolic resistance to insecticides. Insect Biochemistry and Molecular Biology, 24, 411-418.
Carson, R. L. (1963) Silent Spring, Hamish Hamilton, London.
Carter, R. and Mendis, K. N. (2002) Evolutionary and Historical Aspects of the Burden of Malaria. Clinical Microbiology Reviews, 15, 564-594.
Catania, F., Kauer, M. O., Daborn, P. J., Yen, J. L., Ffrench-Constant, R. H. and SchlÖTterer, C. (2004) World-wide survey of an Accord insertion and its association with DDT resistance in Drosophila melanogaster. Molecular Ecology, 13, 2491-2504.
Catteruccia, F., Crisanti, A. and Wimmer, E. A. (2009) Transgenic technologies to induce sterility. Malaria Journal, 16.
CDC (2010) Health Information for Travelers to Ghana.
Chaerkady, R. and Pandey, A. (2008) Applications of Proteomics to Lab Diagnosis. Pathology: Mechanisms of Disease, 3, 485-498.
Chandre, F., Darrier, F., Manga, L., Akogbeto, M., Faye, O., Mouchet, J. and Guillet, P. (1999) Status of pyrethroid resistance in Anopheles gambiae sensu lato. Bulletin of the World Health Organisation, 77, 230 - 234.
Chandre, F., Darriet, F., Duchon, S., Finot, L., Manguin, S., Carnevale, P. and Guillet, P. (2000) Modifications of pyrethroid effects associated with kdr mutation in Anopheles gambiae. Medical and Veterinary Entomology, 14, 81-88.
Chinery, W. A. (1984) Effects of ecological changes on the malaria vectors Anopheles funestus and the Anopheles gambiae complex of mosquitoes in Accra, Ghana. American Journal of Tropical Medicine and Hygiene, 87, 75 - 81.
Chiu, T. L., Wen, Z., Rupasinghe, S. G. and 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.
Christophides, G. K., Zdobnov, E., Barillas-Mury, C., Birney, E., Blandin, S., Blass, C., Brey, P. T., Collins, F. H., Danielli, A., Dimopoulos, G., Hetru, C., Hoa, N. T., Hoffmann, J. A., Kanzok, S. M., Letunic, I., Levashina, E. A., Loukeris, T. G., Lycett, G., Meister, S., Michel, K., Moita, L. F., Muller, H.-M., Osta, M. A., Paskewitz, S. M., Reichhart, J.-M., Rzhetsky, A., Troxler, L., Vernick, K. D., Vlachou, D., Volz, J., von Mering, C., Xu, J., Zheng, L., Bork, P. and Kafatos, F. C. (2002) Immunity-Related Genes and Gene Families in Anopheles gambiae. Science, 298, 159-165.
Clark, A. G. and Shamaan, N. A. (1984) Evidence that DDT-dehydrochlorinase from the house fly is a glutathione S-transferase. Pesticide Biochemistry and Physiology, 22, 249-261.
Claudianos, C., Russell, R. J. and Oakeshott, J. G. (1999) The same amino acid substitution in orthologous esterases confers organophosphate resistance on the house fly and a blowfly. Insect Biochemistry and Molecular Biology, 29, 675-686.
Cleveland, M. S. (1979) Robust Locally Weighted Regression and Smoothing Scatterplots. Journal of the American Statistical Associtation, 74, 829-836.
Coetzee, M., van Wyk, P., Booman, M., Koekemoer, L. L. and Hunt, R. H. (2006) Insecticide resistance in malaria vector mosquitoes in a gold mining town in Ghana and implications for malaria control. Bulletin de la Société de pathologie exotique, 99, 400 - 403.

Cohen, J. and Dupas, P. (2010) Free Distribution or Cost-Sharing? Evidence from a Randomized Malaria Prevention Experiment*. Quarterly Journal of Economics, 125, 1-45.
Coleman, M., Casimiro, S., Hemingway, J. and Sharp, B. (2008) Operational impact of DDT reintroduction for malaria control on Anopheles arabiensis in Mozambique. Journal of Medical Entomology, 45, 885-890.
Cook, P. E. and Sinkins, S. P. (2010) Transcriptional profiling of Anopheles gambiae mosquitoes for adult age estimation. Insect Molecular Biology, 19, 745-751.
Corbel, V., Chabi, J., Dabire, R., Etang, J., Nwane, P., Pigeon, O., Akogbeto, M. and Hougard, J.-M. (2010) Field efficacy of a new mosaic long-lasting mosquito net (PermaNet(R) 3.0) against pyrethroid-resistant malaria vectors: a multi centre study in Western and Central Africa. Malaria Journal, 9, 113.
Corbel, V., Chandre, F., Brengues, C., Akogbeto, M., Lardeux, F., Hougard, J. and Guillet, P. (2004) Dosage-dependent effects of permethrin-treated nets on the behaviour of Anopheles gambiae and the selection of pyrethroid resistance. Malaria Journal, 3, 22.
Corbel, V., Darriet, F., Chandre, F. and Hougard, J. M. (2002) Insecticide mixtures for mosquito net impregnation against malaria vectors. Parasite, 9, 255 - 259.
Corbel, V., N'Guessan, R., Brengues, C., Chandre, F., Djogbenou, L., Martin, T., Akogbeto, M., Hougard, J. M. and Rowland, M. (2007) Multiple insecticide resistance mechanisms in Anopheles gambiae and Culex quinquefasciatus from Benin, West Africa. Acta Tropica, 101, 207-216.
Corbett, J. R., Wright, K. and Baillie, A. C. (1984) The biochemical mode of action of pesticides, Academic Press, London.
Cornman, R. S., Togawa, T., Dunn, W. A., He, N., Emmons, A. and Willis, J. (2008) Annotation and analysis of a large cuticular protein family with the R&R Consensus in Anopheles gambiae. BMC Genomics, 9, 22.
Corona, M. and Robinson, G. E. (2006) Genes of the antioxidant system of the honey bee: annotation and phylogeny. Insect Molecular Biology, 15, 687-701.
Craig, D. W. and Stephan, D. A. (2005) Applications of whole-genome high-density SNP genotyping. Expert Review of Molecular Diagnostics, 5, 159-170.
Craig, M. H., Kleinschmidt, I., Le Sueur, D. and Sharp, B. L. (2004) Exploring 30 years of malaria case data in KwaZulu-Natal, South Africa: Part II. The impact of non-climatic factors. Tropical Medicine & International Health, 9, 1258-1266.
Crews-Oyen, A. E., Kumar, V. and Collins, F. H. (1993) Association of two esterase genes, a chromosomal inversion, and susceptibility to Plasmodium cynomolgi in the African malaria vector Anopheles gambiae. The American Journal of Tropical Medicine and Hygiene, 49, 341-347.
Cui, F., Weill, M., Berthomieu, A., Raymond, M. and Qiao, C.-L. (2007) Characterization of novel esterases in insecticide-resistant mosquitoes. Insect Biochemistry and Molecular Biology, 37, 1131-1137.
Cui, X. and Churchill, G. A. (2003) Statistical tests for differential expression in cDNA microarray experiments. Genome Biology, 4, 210.
Curtis, C., Maxwell, C., Lemnge, M., Kilama, W. L., Steketee, R. W., Hawley, W. A., Bergevin, Y., Campbell, C. C., Sachs, J., Teklehaimanot, A., Ochola, S., Guyatt, H. and Snow, R. W. (2003) Scaling-up coverage with insecticide-treated nets against malaria in Africa: who should pay? The Lancet Infectious Diseases, 3, 304-307.
Curtis, C. F., Maxwell, C. A., Finch, R. and Njunwa, K. J. (1998) Comparison of use of a pyrethroid for house spraying or bednet treatment against Tanzanian malaria vectors. Tropical Medicine & International Health, 3.
Curtis, C. F., Maxwell, C. A., Magesa, S. M., Rwegoshora, R. T. and Wilkes, T. J. (2006) Insecticide-treated bed-nets for malaria mosquito control. Journal of the American Mosquito Control Association, 22, 501-506.
Curtis, C. F., Myamba, J. and Wilkes, T. J. (1996) Comparison of different insecticides and fabrics for anti-mosquito bednets and curtains. Medical and Veterinary Entomology, 10, 1 - 11.

Czeher, C., Labbo, R., Arzika, I. and Duchemin, J.-B. (2008) Evidence of increasing Leu-Phe knockdown resistance mutation in Anopheles gambiae from Niger following a nationwide long-lasting insecticide-treated nets implementation. Malaria Journal, 7, 189.

D'Alessandro, U., Olaleye, B., Langerock, P., Aikins, M. K., Thomson, M. C., Cham, M. K., Greenwood, B. M., McGuire, W., Bennett, S. and Cham, B. A. (1995) Mortality and morbidity from malaria in Gambian children after introduction of an impregnated bednet programme. The Lancet, 345, 479-483.

Dabire, K., Diabate, A., Djogbenou, L., Ouari, A., N'Guessan, R., Ouedraogo, J.-B., Hougard, J.-M., Chandre, F. and Baldet, T. (2008) Dynamics of multiple insecticide resistance in the malaria vector Anopheles gambiae in a rice growing area in South-Western Burkina Faso. Malaria Journal, 7, 188.

Dabire, K. R., Diabate, A., Namontougou, M., Djogbenou, L., Kengne, P., Simard, F., Bass, C. and Baldet, T. (2009) Distribution of insensitive acetylcholinesterase (ace-1R) in Anopheles gambiae s.l. populations from Burkina Faso (West Africa). Tropical Medicine & International Health, 14, 396 - 403.

Daborn, Boundy, Yen, Pittendrigh and ffrench, C. (2001) DDT resistance in Drosophila correlates with over-expression and confers cross-resistance to the neonicotinoid imidacloprid. Molecular Genetics and Genomics, 266, 556-563.

Daborn, P. J., Lumb, C., Boey, A., Wong, W., ffrench-Constant, R. H. and Batterham, P. (2007) Evaluating the insecticide resistance potential of eight Drosophila melanogaster cytochrome P450 genes by transgenic over-expression. Insect Biochemistry and Molecular Biology, 37, 512-519.

Daborn, P. J., Yen, J. L., Bogwitz, M., LeGoff, G., Feil, E., Jeffers, S., Tijet, N., Perry, T., Heckel, D. and Batterham, P. (2002) A single P450 allele associated with insecticide resistance in global populations of *Drosophila*. Science, 297, 2253-2256.

Darriet, F., N' guessan, R., Koffi, A. A., Konan, L., Doannio, J. M. C., Chandre, F. and Carnevale, P. (2000) Impact of the resistance to pyrethroids on the efficacy of impregnated bednets used as a means of prevention against malaria: results of the evaluation carried out with deltamethrin SC in experimental huts. Bulletin de la Societe de Pathologie Exotique, 93, 131 - 134.

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

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

Davies, T. G. E., Andrias, O. O. R., Linda, M. F., Wallace, B. A. and Martin, S. W. (2008) Knockdown resistance to DDT and pyrethroids: from target-site mutations to molecular modelling. Pest Management Science, 64, 1126-1130.

de Souza, D., Kelly-Hope, L., Lawson, B., Wilson, M. and Boakye, D. (2010) Environmental Factors Associated with the Distribution of *Anopheles gambiae s.s* in Ghana; an Important Vector of Lymphatic Filariasis and Malaria. PLoS ONE, 5, e9927.

della Torre, A., Tu, Z. and Petrarca, V. (2005) On the distribution and genetic differentiation of Anopheles gambiae s.s. molecular forms. Insect Biochemistry and Molecular Biology, 35, 755-769.

Devine, G. and Killeen, G. (2010) The potential of a new larviciding method for the control of malaria vectors. Malaria Journal, 9, 142.

Devine, G. J., Zamora Perea, E., Killeen, G. F., Stancil, J. D., Clark, S. J. and Morrison, A. C. (2009) Using adult mosquitoes to transfer insecticides to Aedes aegypti larval habitats. Proceedings of the National Academy of Sciences of the United States of America, 106, 11530 - 11534.

Devonshire, A. L. and Field, L. M. (1991) Gene Amplification and Insecticide Resistance. Annual Review of Entomology, 36, 1-21.

Devonshire, A. L., Field, L. M., Foster, S. P., Moores, G. D., Willamson, M. S. and Blackman, R. L. (1999) The evolution of insecticide resistance in the peach-potato aphid, *Myzus persicae*. In Insecticide resistance: from mechanisms to management. Based on a Discussion Meeting held at the Royal Society, London, in April 1998. (Eds, Denholm, I., Pickett, J. A. and Devonshire, A. L.) CABI Publishing, Oxon.

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

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

Diabate, A., Baldet, T., Chandre, C., Dabire, K. R., Kengne, P., Guiguemde, T. R., Simard, F., Guillet, P., Hemingway, J. and Hougard, J. M. (2003) KDR mutation, a genetic marker to assess events of introgression between the molecular M and S forms of Anopheles gambiae (Diptera: Culicidae) in the tropical savannah area of West Africa. Journal of Medical Entomology, 40, 195 - 198.

Diabate, A., Baldet, T., Chandre, F., Akogbeto, M., Darriet, F., Brengues, C., Guiguemde, T. R., Guillet, P., Hemingway, J. and Hougard, J. M. (2002) The role of agricultural use of insecticides in resistance to pyrethroids in Anopheles gambiae s.l. in Burkina Faso. The American Journal of Tropical Medicine and Hygiene, 67, 617 - 622.

Diabate, A., Dao, A., Yaro, A. S., Adamou, A., Gonzalez, R., Manoukis, N. C., Traora, S. F., Gwadz, R. W. and Lehmann, T. (2009) Spatial swarm segregation and reproductive isolation between the molecular forms of Anopheles gambiae. Proceedings of the Royal Society B, 276, 4215-4222.

Dimopoulos, G., Casavant, T. L., Chang, S., Scheetz, T., Roberts, C., Donohue, M., Schultz, J., Benes, V., Bork, P., Ansorge, W., Soares, M. B. and Kafatos, F. C. (2000) Anopheles gambiae pilot gene discovery project: Identification of mosquito innate immunity genes from expressed sequence tags generated from immune-competent cell lines. Proceedings of the National Academy of Sciences of the United States of America, 97, 6619-6624.

Dimopoulos, G., Christophides, G. K., Meister, S., Schultz, J., White, K. P., Barillas-Mury, C. and Kafatos, F. C. (2002) Genome expression analysis of Anopheles gambiae: Responses to injury, bacterial challenge, and malaria infection. Proceedings of the National Academy of Sciences of the United States of America, 99, 8814-8819.

Ding, Y., Hawkes, N., Meredith, J., Eggleston, P., Hemingway, J. and Ranson, H. (2005) Characterization of the promoters of Epsilon glutathione transferases in the mosquito Anopheles gambiae and their response to oxidative stress. Biochemical Journal, 387, 879-888.

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

Djenontin, A., Chabi, J., Baldet, T., Irish, S., Pennetier, C., Hougard, J.-M., Corbel, V., Akogbeto, M. and Chandre, F. (2009) Managing insecticide resistance in malaria vectors by combining carbamate-treated plastic wall sheeting and pyrethroid-treated bed nets. Malaria Journal, 8, 233.

Djogbenou, L., Chandre, F., Berthomieu, A., Dabire, R., Koffi, A., Alout, H. and Weill, M. (2008a) Evidence of Introgression of the ace-1R Mutation and of the ace-1R Duplication in West African Anopheles gambiae s.s. PLoS One, 3.

Djogbenou, L., Dabire, R., Diabate, A., Kengne, P., Akogbeto, M., Hougard, J. M. and Chandre, F. (2008b) Identification and Geographic Distribution of the ACE-1R

Mutation in the Malaria Vector Anopheles gambiae in South-Western Burkina Faso, West Africa. The American Journal of Tropical Medicine and Hygiene, 78, 298-302.
Djogbenou, L., Labbe, P., Chandre, F., Pasteur, N. and Weill, M. (2009) Ace-1 duplication in Anopheles gambiae: a challenge for malaria control. Malaria Journal, 8, 70.
Djogbenou, L., Noel, V. and Agnew, P. (2010) Costs of insensitive acetylcholinesterase insecticide resistance for the malaria vector Anopheles gambiae homozygous for the G119S mutation. Malaria Journal, 9, 12.
Djouaka, R., Bakare, A., Bankole, H., Doannio, J., Coulibaly, O., Kossou, H., Tamo, M., Basene, H., Popoola, O. K. and Akogbeto, M. (2007) Does the spillage of petroleum products in Anopheles breeding sites have an impact on the pyrethroid resistance? Malaria Journal, 6, 159.
Djouaka, R. F., Bakare, A. A., Coulibaly, O. N., Akogbeto, M. C., Ranson, H., Hemingway, J. and Strode, C. (2008) Expression of the cytochrome P450s, CYP6P3 and CYP6M2 are significantly elevated in multiple pyrethroid resistant populations of Anopheles gambiae s.s. from Southern Benin and Nigeria. BMC Genomics, 9, 538.
Dong, Y., Aguilar, R., Xi, Z., Warr, E., Mongin, E. and Dimopoulos, G. (2006) Anopheles gambiae immune response to human and rodent plasmodium parasite species. PLoS Pathogens, 2, e52.
Donnelly, M. J., Corbel, V., Weetman, D., Wilding, C. S., Williamson, M. S. and Black Iv, W. C. (2009) Does kdr genotype predict insecticide-resistance phenotype in mosquitoes? Trends in Parasitology, 25.
Du, W., Awolola, T. S., Howell, P., Koekemoer, L. L., Brooke, B. D., Benedict, M. Q., Coetzee, M. and Zheng, L. (2005) Independent mutations in the Rdl locus confer dieldrin resistance to Anopheles gambiae and An. arabiensis. Insect Molecular Biology, 14, 179-183.
Duan, J., Wainwright, M. S., Comeron, J. M., Saitou, N., Sanders, A. R., Gelernter, J. and Gejman, P. V. (2003) Synonymous mutations in the human dopamine receptor D2 (DRD2) affect mRNA stability and synthesis of the receptor. Human Molecular Genetics, 12, 205-216.
Duchesne, L., Hubert, J.-F., Verbavatz, J.-M., Thomas, D. and Pietrantonio, P. V. (2003) Mosquito (Aedes aegypti) aquaporin, present in tracheolar cells, transports water, not glycerol, and forms orthogonal arrays in Xenopus oocyte membranes. European Journal of Biochemistry, 270, 422-429.
Duggan, D. J., Bittner, M., Chen, Y., Meltzer, P. and Trent, J. M. (1999) Expression profiling using cDNA microarrays. Nature Genetics, 21.
Elissa, N., Mouchet, J., Riviere, F., Meunier, J. Y. and Yao, K. (1993) Resistance of Anopheles gambiae s.s to pyrethroids in Cote d'Ivoire. Annales des Sociétés Belges de Médecine Tropicale, 73, 291 - 294.
Elliott, R. and Ramakrishna, V. (1956) Insecticide resistance in Anopheles gambiae Giles. Nature, 177, 532-533.
Emerson, J. J., Cardoso-Moreira, M., Borevitz, J. O. and Long, M. (2008) Natural Selection Shapes Genome-Wide Patterns of Copy-Number Polymorphism in Drosophila melanogaster. Science, 320, 1629-1631.
Enayati, A. A., Ranson, H. and Hemingway, J. (2005) Insect glutathione transferases and insecticide resistance. Insect Molecular Biology, 14, 3-8.
Erickson, S. M., Xi, Z., Mayhew, G. F., Ramirez, J. L., Aliota, M. T., Christensen, B. M. and Dimopoulos, G. (2009) Mosquito Infection Responses to Developing Filarial Worms. PLoS Neglected Tropical Diseases, 3, e529.
Etang, J., Fondjo, E., Chandre, F., Morlais, I. and Brengues, C. (2006) First report of knockdown mutations in the malaria vector Anopheles gambiae from Cameroon. American Journal of Tropical Medicine and Hygiene, 74, 795-797.
Etienne, W., Meyer, M. H., Peppers, J. and Meyer, R. A. J. (2004) Comparison of mRNA gene expression by RT-PCR and DNA microarray. Biotechniques, 36, 618-620.

Evans, D. A., Manley, K. A. and Mc, K. V. (1960) Genetic control of isoniazid metabolism in man. British Medical Journal, 2, 485-491.
Fanello, C., Santolamazza, F. and della Torre, A. (2002) Simultaneous identification of species and molecular forms of *Anopheles gambiae* complex by PCR-RFLP. Medical and Veterinary Entomology, 16, 461-464.
Farenhorst, M. and Knols, B. G. J. (2007) Fungal entomopathogens for the control of adult mosquitoes: a look at the issues. Proceedings of the Netherlands Entomological Society Meeting, 18, 51-59.
Feder, M. E. and Hofmann, G. E. (1999) Heat-Shock proteins, molecular chaperones, and the stress response: Evolutionary and Ecological Physiology. Annual Review of Physiology, 61, 243-282.
Felix, R. C., Muller, P., Ribero, V., Ranson, H. and Silveira, H. (2010) Plasmodium infection alters Anopheles gambiae detoxification gene expression. BMC Genomics, 11.
Feng, L., Liu, H., Liu, Y., Lu, Z., Guo, G., Guo, S., Zheng, H., Gao, Y., Cheng, S., Wang, J., Zhang, K. and Zhang, Y. (2010) Power of Deep Sequencing and Agilent Microarray for Gene Expression Profiling Study. Molecular Biotechnology, 45, 101-110.
Feng, M., Patel, D., Dervan, J. J., Ceska, T., Suck, D., Haq, I. and Sayers, J. R. (2004) Roles of divalent metal ions in flap endonuclease-substrate interactions. Nature Structural & Molecular Biology, 11, 450-456.
Feyereisen, R. (1999) Insect P450 Enzymes. Annual Review of Entomology, 44, 507-533.
ffrench-Constant, R. H., Rocheleau, T. A., Steichen, J. C. and Chalmers, A. E. (1993) A point mutation in a Drosophila GABA receptor confers insecticide resistance. Nature, 363, 449-451.
Ffrench, C. and Richard, H. (2007) Which came first : Insecticides or resistance? Trends in Genetics, 23, 1-4.
Field, L. M. and Devonshire, A. L. (1998) Evidence that the E4 and FE4 esterase genes responsible for insecticide resistance in the aphid Myzus persicae (Sulzer) are part of a gene family. Biochemical Journal, 330, 169-173.
Field, L. M., Devonshire, A. L., Ffrench-Constant, R. H. and Forde, B. G. (1989) Changes in DNA methylation are associated with loss of insecticide resistance in the peach-potato aphid Myzus persicae (Sulz.). Federation of European Biochemical Societies, 243, 323-327.
Field, L. M., Devonshire, A. L. and Forde, B. G. (1988) Molecular evidence that insecticide resistance in peach-potato aphids (Myzus persicae Sulz.) results from amplification of an esterase gene. Biochemical Journal, 251, 309-312.
Fillinger, U. and Lindsay, S. W. (2006) Suppression of exposure to malaria vectors by an order of magnitude using microbial larvicides in rural Kenya. Tropical Medicine & International Health, 11, 1629-1642.
Fillinger, U., Ndenga, B., Githeko, A. and Lindsay, S. W. (2009) Integrated malaria vector control with microbial larvicides and insecticide-treated nets in western Kenya: a controlled trial. Bulletin of the World Health Organisation, 87, 655 - 665.
Fournier, D., Bride, J. M., Poirie, M., Berge, J. B. and Plapp, F. W., Jr. (1992) Insect glutathione S-transferases. Biochemical characteristics of the major forms from houseflies susceptible and resistant to insecticides. The Journal of Biological Chemistry, 267, 1840-1845.
Franciosa, H. and Berge, J. B. (1995) Glutathione S-transferases in housefly (Musca domestica): Location of GST-1 and GST-2 families. Insect Biochemistry and Molecular Biology, 25, 311-317.
Fransen, K., Visschedijk, M. C., van Sommeren, S., Fu, J. Y., Franke, L., Festen, E. A., Stokkers, P. C., van Bodegraven, A. A., Crusius, J. B., Hommes, D. W., Zanen, P., de Jong, D. J., Wijmenga, C., van Diemen, C. C. and Weersma, R. K. (2010) Analysis of SNPs with an effect on gene expression identifies UBE2L3

and BCL3 as potential new risk genes for Crohn's disease. Human Molecular Genetics, 19, 3482-3488.
Freedman, W. M., Walker, S. J. and Vrana, K. E. (1999) Quantitative RT-PCR: pitfalls and potential. Biotechniques, 26, 112-122, 124-115.
Fu, G., Lees, R. S., Nimmo, D., Aw, D., Jin, L., Gray, P., Berendonk, T. U., White-Cooper, H., Scaife, S., Kim Phuc, H., Marinotti, O., Jasinskiene, N., James, A. A. and Alphey, L. (2010) Female-specific flightless phenotype for mosquito control. Proceedings of the National Academy of Sciences, 107, 4550-4554.
Galindo, K. and Smith, D. P. (2001) A Large Family of Divergent Drosophila Odorant-Binding Proteins Expressed in Gustatory and Olfactory Sensilla. Genetics, 159, 1059-1072.
Gassmann, A. J., Carrière, Y. and Tabashnik, B. E. (2009) Fitness Costs of Insect Resistance to Bacillus thuringiensis. Annual Review of Entomology, 54, 147-163.
Geng, Y.-J., Gao, S.-T., Huang, D.-N., Zhao, Y.-R., Liu, J.-p., Li, X.-H. and Zhang, R.-L. (2009) Differentially expressed genes between female and male adult *Anopheles anthropophagus*. Parasitology Research, 105, 843-851.
GhanaStatisticalService (2010).
Gibbs, J. R. and Singleton, A. (2006) Application of Genome-Wide Single Nucleotide Polymorphism Typing: Simple Association and Beyond. PLoS Genetics, 2, e150.
Gilad, Y., Rifkin, S. A. and Pritchard, J. K. (2008) Revealing the architecture of gene regulation: the promise of eQTL studies. Trends in Genetics, 24, 408-415.
Gilles, H. M. (2002) Historical outline. In Essential Malariology(Eds, Warrel, D. A. and Gilles, H. M.) Arnold, London, pp. 1-8.
Gillies, M. T. and Coetzee, M. (1987) Publication of the South African Institute for Medical Research, Johannesburg.
Gong, M., Shen, B., Gu, Y., Tian, H., Ma, L., Li, X., Yang, M., Hu, Y., Sun, Y., Hu, X., Li, J. and Zhu, C. (2005) Serine proteinase over-expression in relation to deltamethrin resistance in Culex pipiens pallens. Archives of Biochemistry and Biophysics, 438, 53-62.
Goodman, O. B., Jr., Krupnick, J. G., Santini, F., Gurevich, V. V., Penn, R. B., Gagnon, A. W., Keen, J. H. and Benovic, J. L. (1996) Beta-arrestin acts as a clathrin adaptor in endocytosis of the beta2-adrenergic receptor. Nature. 1996 Oct 3;383(6599):447-50.
Grant, D. F. and Matsumura, F. (1989) Glutathione S-Transferase 1 and 2 in Susceptible and Insecticide Resistant *Aedes aegypti*. Pesticide Biochemistry and Physiology, 33, 132-143.
Grieco, J. P., Achee, N. L., Chareonviriyaphap, T., Suwonkerd, W., Chauhan, K., Sardelis, M. R. and Roberts, D. R. (2007) A New Classification System for the Actions of IRS Chemicals Traditionally Used For Malaria Control. PLoS ONE, 2, e716.
Grossman, G. L., Rafferty, C. S., Clayton, J. R., Stevens, T. K., Mukabayire, O. and Benedict, M. Q. (2001) Germline transformation of the malaria vector, Anopheles gambiae, with the piggyBac transposable element. Insect Molecular Biology, 10, 597-604.
Guillet, P., N'Guessan, R., Darriet, F., Traore-Lamizana, M., Chandre, F. and Carnevale, P. (2001) Combined pyrethroid and carbamate 'two-in-one' treated mosquito nets: field efficacy against pyrethroid-resistant Anopheles gambiae and Culex quinquefasciatus. Medical and Veterinary Entomology, 15, 105-112.
Gupta, R. C., Milatovic, S., Dettbarn, W.-D., Aschner, M. and Milatovic, D. (2007) Neuronal oxidative injury and dendritic damage induced by carbofuran: Protection by memantine. Toxicology and Applied Pharmacology, 219, 97-105.
Gustafsson, C., Govindarajan, S. and Minshull, J. (2004) Codon bias and heterologous protein expression. Trends in Biotechnology, 22, 346-353.
Habig, W. H., Pabst, M. J. and Jakoby, W. B. (1974) Glutathione S-Transferases. The Journal of Biological Chemistry, 249, 7130-7139.

Hannon, G. J. (2002) RNA interference. Nature, 418, 244-251.
Hannun, Y. A. (1996) Functions of Ceramide in Coordinating Cellular Responses to Stress. Science, 274, 1855-1859.
Hardstone, M. C., Leichter, C., Harrington, L. C., Kasai, S., Tomita, T. and Scott, J. G. (2007) Cytochrome P450 monooxygenase-mediated permethrin resistance confers limited and larval specific cross-resistance in the southern house mosquito, Culex pipiens quinquefasciatus. Pesticide Biochemistry and Physiology, 89, 175-184.
Hargreaves, K., Hunt, R. H., Brooke, B. D., Mthembu, J., Weeto, M. M., Awolola, T. S. and Coetzee, M. (2003) Anopheles arabiensis and An. quadriannulatus resistance to DDT in South Africa. Medical and Veterinary Entomology, 17, 417-422.
Hargreaves, K., Koekemoer, L. L., Brooke, B., Hunt, R. H., Mthembu, J. and Coetzee, M. (2000) Anopheles funestus resistant to pyrethroid insecticides in South Africa. Medical and Veterinary Entomology, 14, 181 - 189.
Hartley, C. J., Newcomb, R. D., Russell, R. J., Yong, C. G., Stevens, J. R., Yeates, D. K., La Salle, J. and Oakeshott, J. G. (2006) Amplification of DNA from preserved specimens shows blowflies were preadapted for the rapid evolution of insecticide resistance. Proceedings of the National Academy of Sciences, 103, 8757-8762.
Hawkes, N. and hemingway, J. (2002) Analysis of the promoters for the B-esterase genes associated with insecticide resistance in the mosquito *Culex quinquefasciatus*. Biochimica et Biophysica Acta (BBA), 1574, 51-62.
Hawley, W. A., Phillips-Howard, P. A., ter Kuile, F. O., Terlouw, D. J., Vulule, J. M., Ombok, M., Nahlen, B. L., Gimnig, J. E., Kariuki, S. K., Kolzac, M. S. and Hightower, A. W. (2003) Community-wide effects of permethrin-treated bed nets on child mortality and malaria morbidity in western Kenya. The American Journal of Tropical Medicine and Hygiene, 68, 121 - 127.
Hayes, J. D., Flanagan, J. U. and Jowsey, I. R. (2005) Glutathione Transferases. Annual Review of Pharmacology and Toxicology, 45, 51-88.
Hayes, J. D. and Strange, R. C. (2000) Glutathione S-Transferase Polymorphisms and Their Biological Consequences. Pharmacology, 61, 154-166.
He, N., Botelho, J. M. C., McNall, R. J., Belozerov, V., Augustine Dunn, W., Mize, T., Orlando, R. and Willis, J. H. (2007) Proteomic analysis of cast cuticles from Anopheles gambiae by tandem mass spectrometry. Insect Biochemistry and Molecular Biology, 37, 135-146.
He, X. and Zhang, J. (2006) Why Do Hubs Tend to Be Essential in Protein Networks? PLoS Genetics, 2, e88.
Helvig, C., Koener, J. F., Unnithan, G. C. and Feyereisen, R. (2004) CYP15A1, the cytochrome P450 that catalyzes epoxidation of methyl farnesoate to juvenile hormone III in cockroach corpora allate. Proceedings of the National Academy of Sciences of the United States of America, 101, 4024-4029.
Hemingway, J. (2000) The molecular basis of two contrasting metabolic mechanisms of insecticide resistance. Insect Biochemistry and Molecular Biology, 30, 1009-1015.
Hemingway, J., Beaty, B. J., Rowland, M., Scott, T. W. and Sharp, B. L. (2006) The Innovative Vector Control Consortium: improved control of mosquito-borne diseases. Trends in Parasitology, 22, 308-312.
Hemingway, J., Hawkes, N., Prapanthadara, L. A., Indrananda Jayawardenal, K. G. and Ranson, H. (1998) The role of gene splicing, gene amplification and regulation in mosquito insecticide resistance. Philosophical Transactions of The Royal Society London. Series B, Biological Sciences, 353, 1695-1699.
Hemingway, J., Hawkes, N. J., McCarroll, L. and Ranson, H. (2004) The molecular basis of insecticide resistance in mosquitoes. Insect Biochemistry and Molecular Biology, 34, 653-665.
Henry, M. C., Assi, S. B., Rogier, C., Dossou-Yovo, J., Chandre, F., Guillet, P. and Carnevale, P. (2005) Protective efficacy of lambda-cyhalothrin treated nets in

Anopheles gambiae pyrethroid resistance areas of Cote d'Ivoire. The American Journal of Tropical Medicine and Hygiene, 73, 859 - 864.
Hick, C. A., Field, L. M. and Devonshire, A. L. (1996) Changes in the methylation of amplified esterase DNA during loss and reselection of insecticide resistance in peach-potato aphids, Myzus persicae. Insect Biochemistry and Molecular Biology, 26, 41-47.
Hill, C. A., Fox, A. N., Pitts, R. J., Kent, L. B., Tan, P. L., Chrystal, M. A., Cravchik, A., Collins, F. H., Robertson, H. M. and Zwiebel, L. J. (2002) G protein-coupled receptors in Anopheles gambiae. Science. , 298, 176-178.
Hintermann, E., Grieder, N. C., Amherd, R., Brodbeck, D. and Meyer, U. A. (1996) Cloning of an arylalkylamine N-acetyltransferase (aaNAT1) from Drosophila melanogaster expressed in the nervous system and the gut. Proceedings of the National Academy of Sciences of the United States of America, 93, 12315-12320.
Hittinger, C. T., Johnston, M., Tossberg, J. T. and Rokas, A. (2010) Leveraging skewed transcript abundance by RNA-Seq to increase the genomic depth of the tree of life. Proceedings of the National Academy of Sciences, 107, 1476-1481.
Hoffmann, J. A. and Reichhart, J. M. (2002) Drosophila innate immunity: an evolutionary perspective. Nature Immunology, 3, 121-126.
Holt, R. A., Subramanian, G. M., Halpern, A., Sutton, G. G., Charlab, R., Nusskern, D. R., Wincker, P., Clark, A. G., Ribeiro, J. M., Wides, R., Salzberg, S. L., Loftus, B., Yandell, M., Majoros, W. H., Rusch, D. B., Lai, Z., Kraft, C. L., Abril, J. F., Anthouard, V., Arensburger, P., Atkinson, P. W., Baden, H., de Berardinis, V., Baldwin, D., Benes, V., Biedler, J., Blass, C., Bolanos, R., Boscus, D., Barnstead, M., Cai, S., Center, A., Chaturverdi, K., Christophides, G. K., Chrystal, M. A., Clamp, M., Cravchik, A., Curwen, V., Dana, A., Delcher, A., Dew, I., Evans, C. A., Flanigan, M., Grundschober-Freimoser, A., Friedli, L., Gu, Z., Guan, P., Guigo, R., Hillenmeyer, M. E., Hladun, S. L., Hogan, J. R., Hong, Y. S., Hoover, J., Jaillon, O., Ke, Z., Kodira, C., Kokoza, E., Koutsos, A., Letunic, I., Levitsky, A., Liang, Y., Lin, J. J., Lobo, N. F., Lopez, J. R., Malek, J. A., McIntosh, T. C., Meister, S., Miller, J., Mobarry, C., Mongin, E., Murphy, S. D., O'Brochta, D. A., Pfannkoch, C., Qi, R., Regier, M. A., Remington, K., Shao, H., Sharakhova, M. V., Sitter, C. D., Shetty, J., Smith, T. J., Strong, R., Sun, J., Thomasova, D., Ton, L. Q., Topalis, P., Tu, Z., Unger, M. F., Walenz, B., Wang, A., Wang, J., Wang, M., Wang, X., Woodford, K. J., Wortman, J. R., Wu, M., Yao, A., Zdobnov, E. M., Zhang, H., Zhao, Q., et al. (2002) The genome sequence of the malaria mosquito Anopheles gambiae. Science, 298, 129 - 149.
Hooper, A. M., Dufour, S., He, X., Muck, A., Zhou, J.-J., Almeida, R., Field, L. M., Svatos, A. and Pickett, J. A. (2009) High-throughput ESI-MS analysis of binding between the Bombyx mori pheromone-binding protein BmorPBP1, its pheromone components and some analogues. Chemical Communications, 5725-5727.
Hopkins, T. L. and Kramer, K. J. (1992) Insect Cuticle Sclerotization. Annual Review of Entomology, 37, 273-302.
Hougard, J. M., Corbel, V., N'Guessan, R., Darriet, F., Chandre, F., Akogbeto, M., Baldet, T., Guillet, P., Carnevale, P. and Traore-Lamizana, M. (2003) Efficacy of mosquito nets treated with insecticide mixtures or mosaics against insecticide resistant Anopheles gambiae and Culex quinquefasciatus (Diptera: Culicidae) in Cote d'Ivoire. Bulletin of Entomological Research, 93, 491 - 498.
Huang, H.-S., Hu, N.-T., Yao, Y.-E., Wu, C.-Y., Chiang, S.-W. and 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, H., Stok, J. E., Stoutamire, D. W., Gee, S. J. and Hammock, B. D. (2005) Development of Optically Pure Pyrethroid-Like Fluorescent Substrates for Carboxylesterases. Chemical Research in Toxicology, 18, 516-527.

Huang, W. L., Li, C. Z., Chen, Z. R., He, W., Zhou, Y., Zhou, Z. G., Liu, S. W. and Zhou, C. (2010) [Effects of UV-induced DNA damage on vector ligation and transformation into bacterial cells]. Nan Fang Yi Ke Da Xue Xue Bao, 30, 111-113.

Huang, Y., Prasad, M., Lemon, W. J., Hampel, H., Wright, F. A., Kornacker, K., LiVolsi, V., Frankel, W., Kloos, R. T., Eng, C., Pellegata, N. S. and de la Chapelle, A. (2001) Gene expression in papillary thyroid carcinoma reveals highly consistent profiles. Proceedings of the National Academy of Sciences, 98, 15044-15049.

Hudson, M. E. (2008) Sequencing breakthroughs for genomic ecology and evolutionary biology. Molecular Ecology Resources, 8, 3-17.

Huestis, D. L. and Marshall, J. L. (2009) From gene expression to phenotype in insects: non-microarray approaches for transcriptome analysis. BioScience, 59.

Hunt, R., Sauna, Z. E., Ambudkar, S. V., Gottesman, M. M. and Kimchi-Sarfaty, C. (2009) Silent (synonymous) SNPs: should we care about them? Methods in Molecular Biology, 578, 23-39.

Imai, T., Globerman, H., Gertner, J. M., Kagawa, N. and Waterman, M. R. (1993) Expression and purification of functional human 17 alpha-hydroxylase/17,20-lyase (P450c17) in Escherichia coli. The Journal of Biochemical Chemistry, 268, 19681-19689.

IRAC (2009) Insecticide resistance: causes and actions. (Ed, Center, I. R. A. C. I. i. c. w. t. S. R. I. P. M.).

Jakobsson, P.-J., Mancini, J. A., Riendeau, D. and Ford-Hutchinson, A. W. (1997) Identification and Characterization of a Novel Microsomal Enzyme with Glutathione-dependent Transferase and Peroxidase Activities. Journal of Biological Chemistry, 272, 22934-22939.

Jones, D. G. (1998) Piperonyl butoxide: the insecticide synergist, Academic Press, London.

Joußen, N., Heckel, D. G., Haas, M., Schuphan, I. and Schmidt, B. (2008) Metabolism of imidacloprid and DDT by P450 CYP6G1 expressed in cell cultures of Nicotiana tabacum suggests detoxification of these insecticides in Cyp6g1-overexpressing strains of Drosophila melanogaster, leading to resistance. Pest Management Science, 64, 65-73.

Justice, R. W., Dimitratos, S., Walter, M. F., Woods, D. F. and Biessmann, H. (2003) Sexual dimorphic expression of putative antennal carrier protein genes in the malaria vector Anopheles gambiae. Insect Molecular Biology, 12, 581-594.

Kamboj, A., Kiran, R. and Sandhir, R. (2006) Carbofuran-induced neurochemical and neurobehavioral alterations in rats: attenuation by <i>N</i>-acetylcysteine. Experimental Brain Research, 170, 567-575.

Kamleh, M. A., Hobani, Y., Dow, J. A. T., Zheng, L. and Watson, D. G. (2009) Towards a platform for the metabonomic profiling of different strains of Drosophila melanogaster using liquid chromatography–Fourier transform mass spectrometry. FEBS Journal, 276, 6798-6809.

Karlson, P. and Sekeris, C. E. (1962) N-Acetyl-dopamine as Sclerotizing Agent of the Insect Cuticle. Nature, 195, 183-184.

Karunaratne, S., vaughan, A., Paton, M. G. and Hemingway, J. (1998) Amplifcation of a serine esterase gene is involved in insecticide resistance in Sri Lankan Culex tritaeniorhynchus. Insect Molecular Biology, 7, 307-315.

Kaufmann, N., Mathai, J. C., Hill, W. G., Dow, J. A. T., Zeidel, M. L. and Brodsky, J. L. (2005) Developmental expression and biophysical characterization of a Drosophila melanogaster aquaporin. American Journal of Physiology - Cell Physiology, 289, C397-C407.

Keating, J., Macintyre, K., Mbogo, C., Githure, J. and Beier, J. (2004) Characterization of potential larval habitats for Anopheles mosquitoes in relation to urban land-use in Malindi, Kenya. International Journal of Health Geographics, 3, 9.

Kerr, M. K., Churchill, G.A., (2001) Experimental design for gene expression microarrays. Biostatistics, 2, 183-202.

Kerr, M. K., Mitchell, M., Churchill, G.A., (2000) Analysis of variance for gene expression microarray data. Journal of Computational Biology, 7, 819-837.
Khan, J., Saal, L. H., Bittner, M. L., Chen, Y., Trent, J. M. and Meltzer, P. S. (1999) Expression profiling in cancer using cDNA microarrays. Electrophoresis, 20, 223-229.
Kimichi-Sarfaty, C., Mi Oh, J., Kim, I.-W., Sauna, Z. E., Calcagno, A. M., Ambudkar, S. V. and Gottesman, M. M. (2007) A "Silent" Polymorphism in the MDR1 Gene Changes Substrate Specificity. Science, 315, 525-528.
Kitatani, K., Idkowiak-Baldys, J. and Hannun, Y. A. (2008) The sphingolipid salvage pathway in ceramide metabolism and signaling. Cellular Signalling, 20, 1010-1018.
Kleinschmidt, I., Schwabe, C., Benavente, L., Torrez, M., Ridl, F. C., Segura, J. L., Ehmer, P. and Nchama, G. N. (2009a) Marked Increase in Child Survival after Four Years of Intensive Malaria Control. The American Journal of Tropical Medicine and Hygiene, 80, 882-888.
Kleinschmidt, I., Schwabe, C., Shiva, M., Segura, J. L., Sima, V., Mabunda, S. J. A. and Coleman, M. (2009b) Combining Indoor Residual Spraying and Insecticide-Treated Net Interventions. The American Journal of Tropical Medicine and Hygiene, 81, 519-524.
Klinkenberg, E. (2006) Impact of irrigated urban agriculture on malaria transmission in two cities in Ghana. In Vector group, Liverpool School of Tropical Medicine, Vol. PhD University of Liverpool, Liverpool.
Klinkenberg, E., McCall, P. J., Wilson, M., Amerasinghe, F. and Donnelly, M. (2008) Impact of urban agriculture on malaria vectors in Accra, Ghana. Malaria Journal, 7, 151.
Knols, B. G. J., Bukhari, T. and Farenhorst, M. (2010) Entomopathogenic fungi as the next-generation control agents against malaria mosquitoes. Future Microbiology, 5, 339-341.
Kokoza, V., Ahmed, A., Woon Shin, S., Okafor, N., Zou, Z. and Raikhel, A. S. (2010) Blocking of Plasmodium transmission by cooperative action of Cecropin A and Defensin A in transgenic Aedes aegypti mosquitoes. Proceedings of the National Academy of Sciences, 107, 8111-8116.
Kolaczinski, J. and Curtis, C. (2001) Laboratory evaluation of fipronil, a phenylpyrazole insecticide, against adult Anopheles (Diptera: Culicidae) and investigation of its possible cross-resistance with dieldrin in Anopheles stephensi. Pest Management Science, 57, 41-45.
Komagata, O., Kasai, S. and Tomita, T. (2010) Overexpression of cytochrome P450 genes in pyrethroid-resistant Culex quinquefasciatus. Insect Biochemistry and Molecular Biology, 40, 146-152.
Kononen, J., Bubendorf, L., Kallioniemi, A., Barlund, M., Schrami, P., Leigthon, S., Torhorst, J., Milhatsch, M. J., Sauter, G. and Kallioniemi, O. P. (1998) Tissue microarrays for high-throughput molecular profiling of tumor specimens. Nature Medicine, 4, 844-847.
Korenromp, E. L., Miller, J., Cibulskis, R. E., Kabir Cham, M., Alnwick, D. and Dye, C. (2003) Monitoring mosquito net coverage for malaria control in Africa: possession vs. use by children under 5 years. Tropical Medicine & International Health, 8, 693-703.
Kouznetsov, R. L. (1977) Malaria control by application of indoor spraying of residual insecticides in tropical Africa and its impact on community health. Tropical Doctor, 7, 81-91.
Kristan, M., Fleischmann, H., della Torre, A., Stich, A. and Curtis, C. F. (2003) Pyrethroid resistance/susceptibility and differential urban/rural distribution of Anopheles arabiensis and An. gambiae s.s. malaria vectors in Nigeria and Ghana. Medical and Veterinary Entomology, 17, 326 - 332.
Kristensen, M., Hansen, K. K. and Jensen, K.-M. V. (2005) Cross-Resistance Between Dieldrin and Fipronil in German Cockroach (Dictyoptera: Blattellidae). Journal of Economic Entomology, 98, 1305-1310.

Krupnick, J. G. and Benovic, J. L. (1998) The role of the receptor kinases and arrestins in G protein coupled receptor regulation. Annual Review of Pharmacology and Toxicology, 38, 289-319.
Kumar, S., Christophides, G. K., Cantera, R., Charles, B., Han, Y. S., Meister, S., Dimopoulos, G., Kafatos, F. C. and Barillas-Mury, C. (2003) The role of reactive oxygen species on Plasmodium melanotic encapsulation in Anopheles gambiae. Proceedings of the National Academy of Sciences of the United States of America, 100, 14139-14144.
Ladbury, J. E. and Chowdhry, B. Z. (1996) Sensing the heat: the application of isothermal titration calorimetry to thermodynamic studies of biomolecular interactions. Chemistry & Biology, 3, 791-801.
lal, A., Sui, I. M. and Riggins, G. J. (1999) Serial analysis of gene expression: probing transcriptomes for molecular targets. Current Opinion in Molecular Therapeutics, 1.
Laporte, S. A., Oakley, R. H., Holt, J. A., Barak, L. S. and Caron, M. G. (2000) The interaction of beta-arrestin with the AP-2 adaptor is required for the clustering of beta 2-adrenergic receptor into clathrin-coated pits. Journal of Biological Chemistry, 275, 23120-23126.
Lawniczak, M. K. N., Emrich, S. J., Holloway, A. K., Regier, A. P., Olson, M., White, B., Redmond, S., Fulton, L., Appelbaum, E., Godfrey, J., Farmer, C., Chinwalla, A., Yang, S. P., Minx, P., Nelson, J., Kyung, K., Walenz, B. P., Garcia-Hernandez, E., Aguiar, M., Viswanathan, L. D., Rogers, Y. H., Strausberg, R. L., Saski, C. A., Lawson, D., Collins, F. H., Kafatos, F. C., Christophides, G. K., Clifton, S. W., Kirkness, E. F. and Besansky, N. J. (2010) Widespread Divergence Between Incipient Anopheles gambiae Species Revealed by Whole Genome Sequences. Science, 330, 512-514.
Le Goff, G., Boundy, S., Daborn, P. J., Yen, J. L., Sofer, L., Lind, R., Sabourault, C., Madi-Ravazzi, L. and ffrench-Constant, R. H. (2003) Microarray analysis of cytochrome P450 mediated insecticide resistance in Drosophila. Insect Biochemistry and Molecular Biology, 33, 701-708.
Le Goff, G., Hamon, A., Bergé, J.-B. and Amichot, M. (2005) Resistance to fipronil in Drosophila simulans: influence of two point mutations in the RDL GABA receptor subunit. Journal of Neurochemistry, 92, 1295-1305.
Lehner, B., Crombie, C., Tischler, J., Fortunato, A. and Fraser, A. G. (2006) Systematic mapping of genetic interactions in Caenorhabditis elegans identifies common modifiers of diverse signaling pathways. Nature Genetics, 38, 896-903.
Lehninger, A. L. (1982) Enzymes. In Principles of Biochemistry(Eds, Anderson, S. and Fox, J.) Worth Publishers, Inc., New York.
Lengeler, C. (2004) Insecticide-treated bed nets and curtains for preventing malaria. The Cochrane Collaboration, CD000363.
Lenormand, T., Bourguet, D., Guillemaud, T. and Raymond, M. (1999) Tracking the evolution of insecticide resistance in the mosquito Culex pipiens. Nature, 400, 861-864.
Lenormand, T., Guillemaud, T., Bourguet, D. and Raymond, M. (1998) Appearance and Sweep of a Gene Duplication: Adaptive Response and Potential for New Functions in the Mosquito Culex pipiens. Evolution, 52, 1705-1712.
Li, S., Picimbon, J.-F., Ji, S., Kan, Y., Chuanling, Q., Zhou, J.-J. and Pelosi, P. (2008) Multiple functions of an odorant-binding protein in the mosquito Aedes aegypti. Biochemical and Biophysical Research Communications, 372, 464-468.
Li, X., Baudry, J., Berenbaum, M. R. and Schuler, M. A. (2004) Structural and functional divergence of insect CYP6B proteins: From specialist to generalist cytochrome P450. Proceedings of the National Academy of Sciences, 101, 2939-2944.
Li, Z. X., Pickett, J. A., Field, L. M. and Zhou, J. J. (2005) Identification and expression of odorant-binding proteins of the malaria-carrying mosquitoes Anopheles gambiae and Anopheles arabiensis. Archives of Insect Biochemistry and Physiology, 58, 175-189.

Lin, Y., Lu, P., Tang, C., Mei, Q., Sandig, G., Rodrigues, A. D., Rushmore, T. H. and Shou, M. (2001) Substrate Inhibition Kinetics for Cytochrome P450-Catalyzed Reactions. Drug Metabolism and Disposition, 29, 368-374.
Lines, J. D. (1988) Do agricultural insecticides select for insecticide resistance in mosquitoes - a look at the evidence. Parasitology Today, 4, S17 - S20.
Lines, J. D., Myamba, J. and Curtis, C. F. (1987) Experimental hut trials of permethrin-impregnated mosquito nets and eave curtains against malaria vectors in Tanzania. Medical and Veterinary Entomology, 1, 37 - 51.
Liu, N. N., Liu, H. Q., Zhu, F. and Zhang, L. (2007) Differential expression of genes in pyrethroid resistant and susceptible mosquitoes, Culex quinquefasciatus (S.). Gene, 394, 61 - 68.
Lockwood, J. A., Byford, R. L., Story, R. N., Sparks, T. C. and Quisenberry, S. S. (1985) Behavioral resistance to the pyrethroids in the Horn Fly, Haematobia irritans (Diptera: Muscidae). Environmental Entomology, 14, 873-880.
Lumjuan, N., McCarroll, L., Prapanthadara, L.-a., Hemingway, J. and Ranson, H. (2005) Elevated activity of an Epsilon class glutathione transferase confers DDT resistance in the dengue vector, Aedes aegypti. Insect Biochemistry and Molecular Biology, 35, 861-871.
Luna, C., Wang, X., Huang, Y., Zhang, J. and Zheng, L. (2002) Characterization of four Toll related genes during development and immune responses in Anopheles gambiae. Insect Biochemistry and Molecular Biology, 32, 1171-1179.
Lycett, G. J., McLaughlin, L. A., Ranson, H., Hemingway, J., Kafatos, F. C., Loukeris, T. G. and Paine, M. J. I. (2006) Anopheles gambiae P450 reductase is highly expressed in oenocytes and in vivo knockdown increases permethrin susceptibility. Insect Molecular Biology, 15, 321-327.
Lynd, A., Ranson, H., McCall, P. J., Randle, N. P., Black, W. C., Walker, E. D. and Donnelly, M. J. (2005) A simplified high-throughput method for pyrethroid knock-down resistance (kdr) detection in Anopheles gambiae. Malaria Journal, 4, 16.
Lynd, A., Weetman, D., Barbosa, S., Egyir-Yawson, A., Mitchell, S., Pinto, J., Hastings, I. and Donnelly, M. J. (2010) Field, genetic and modelling approaches show strong positive selection acting upon an insecticide resistance mutation in Anopheles gambiae s.s. Molecular Biology & Evolution, 27, 1117-1125.
Mabaso, M. L. H., Sharp, B. and Lengeler, C. (2004) Historical review of malarial control in southern African with emphasis on the use of indoor residual house-spraying. Tropical Medicine & International Health, 9, 846-856.
Macdonald, S. J. and Long, A. D. (2004) A Potential Regulatory Polymorphism Upstream of hairy Is Not Associated With Bristle Number Variation in Wild-Caught Drosophila. Genetics, 167, 2127-2131.
Macintyre, K., Keating, J., Okbaldt, Y. B., Zerom, M., Sosler, S., Ghebremeskel, T. and Eisele, T. P. (2006) Rolling out insecticide treated nets in Eritrea: examining the determinants of possession and use in malarious zones during the rainy season. Tropical Medicine & International Health, 11, 824-833.
Maharaj, R., Mthembu, D. J. and Sharp, B. L. (2005) Impact of DDT re-introduction on malaria transmission in KwaZulu-Natal. South African Medical Journal, 95, 871-874.
Malcom, C. A. (1988) Current status of pyrethroid resistance on anophelines. Parasitology Today, 4, S12 - S14.
Mao, W., Rupasinghe, S. G., Zangerl, A. R., Berenbaum, M. R. and Schuler, M. A. (2007) Allelic Variation in the Depressaria pastinacella CYP6AB3 Protein Enhances Metabolism of Plant Allelochemicals by Altering a Proximal Surface Residue and Potential Interactions with Cytochrome P450 Reductase. Journal of Biological Chemistry, 282, 10544-10552.
Marcombe, S., Poupardin, R., Darriet, F., Reynaud, S., Bonnet, J., Strode, C., Brengues, C., Yebakima, A., Ranson, H. and Corbel, V. (2009) Exploring the molecular basis of insecticide resistance in the dengue vector Aedes aegypti: a case study in Martinique Island (French West Indies). BMC Genomics, 10, 494.

Mardis, E. R. (2008) Next-Generation DNA Sequencing Methods. Annual review of Genomics and Human Genetics, 9, 387-402.
Marinotti, O., Calvo, E., Nguyen, Q. K., Dissanayake, S., Ribeiro, J. M. C. and James, A. A. (2006) Genome-wide analysis of gene expression in adult Anopheles gambiae. Insect Molecular Biology, 15, 1-12.
Martinez-Torres, D. F., Chandre, F., Williamson, M. S., Darriet, F., Berge, J. B., Devonshire, A. L., Guillet, P., Pasteur, N. and Pauron, D. (1998) Molecular characterization of pyrethroid knockdown resistance (kdr) in the major malaria vector Anopheles gambiae s.s. Insect Molecular Biology, 7, 179 - 184.
Matambo, T. S., Abdalla, H., Brooke, B. D., Koekemoer, L. L., Mnzava, A., Hunt, R. H. and Coetzee, M. (2007) Insecticide resistance in the malarial mosquito Anopheles arabiensis and association with the kdr mutation. Medical and Veterinary Entomology, 21, 97 - 102.
Mathenge, E. M., Gimnig, J. E., Kolczak, M., Ombok, M., Irungu, L. W. and Hawley, W. A. (2001) Effect of permethrin-impregnated nets on exiting behavior, blood feeding success, and time of feeding of malaria mosquitoes (Diptera: Culicidae) in western Kenya. Journal of Medical Entomology, 38, 531-536.
Matowo, J., Kulkarni, M., Mosha, F., Oxborough, R., Kitau, J., Tenu, F. and Rowland, M. (2010) Biochemical basis of permethrin resistance in Anopheles arabiensis from Lower Moshi, north-eastern Tanzania. Malaria Journal, 9, 193.
McCarroll, L. and King, L. A. (1997) Stable insect cell cultures for recombinant protein production. Current Opinion in Biotechnology, 8, 590-594.
McCart, C., Buckling, A. and ffrench-Constant, R. H. (2005) DDT resistance in flies carries no cost. Current Biology, 15, R587-R589.
McIlwain, C. C., Townsend, D. M. and Tew, K. D. (2006) Glutathione S-transferase polymorphisms: cancer incidence and therapy. Oncogene, 25, 1639-1648.
McLaughlin, L. A., Niazi, U., Bibby, J., David, J. P., Vontas, J., Hemingway, J., Ranson, H., Sutcliffe, M. J. and Paine, M. J. I. (2008) Characterization of inhibitors and substrates of *Anopheles gambiae* CYP6Z2. Insect Molecular Biology, 17, 125-135.
Merrill, C. E., Pitts, R. J. and Zwiebel, L. J. (2003) Molecular characterization of arrestin family members in the malaria vector mosquito, Anopheles gambiae. Insect Molecular Biology, 12, 641-650.
Merrill, C. E., Riesgo-Escovar, J., Pitts, R. J., Kafatos, F. C., Carlson, J. R. and Zwiebel, L. J. (2002) Visual arrestins in olfactory pathways of Drosophila and the malaria vector mosquito Anopheles gambiae. Proceedings of the National Academy of Sciences of the United States of America, 99, 1633-1638.
Metropolis, H., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H. and Teller, E. (1953) Equations of state calculations by fast computing machines. Journal of Chemical Properties, 21, 1087.
Michel, T., Reichhart, J.-M., Hoffmann, J. A. and Royet, J. (2001) Drosophila Toll is activated by Gram-positive bacteria through a circulating peptidoglycan recognition protein. Nature, 414, 756-759.
Miller, T. A., Kennedy, J. M. and Collins, C. (1979) CNS Insensitivity to Pyrethroids in the Resistant kdr Strain of House Flies. . Pesticide Biochemistry and Physiology, 12, 224-230.
Mireji, P. O., Keating, J., Hassanali, A., Mbogo, C. M., Nyambaka, H., Kahindi, S. and Beier, J. C. (2008) Heavy metals in mosquito larval habitats in urban Kisumu and Malindi, Kenya, and their impact. Ecotoxicology and Environmental Safety, 70, 147-153.
Mirnezhad, M., Romero-González, R. R., Leiss, K. A., Choi, Y. H., Verpoorte, R. and Klinkhamer, P. G. L. (2009) Metabolomic analysis of host plant resistance to thrips in wild and cultivated tomatoes. Phytochemical Analysis, 21, 110-117.
Mongkalangoon, P., Grieco, J. P., Achee, N. L., Suwonkerd, W. and Chareonviriyaphap, T. (2009) Irritability and repellency of synthetic pyrethroids on an Aedes aegypti population from Thailand. Journal of Vector Ecology, 34, 217-224.

Morey, J. S., Ryan, J. C. and Van Dolah, F. (2006) Microarray validation: factors influencing correlation between oligonucleotide microarrays and real-time PCR. Biological Procedures Online, 8, 175-193.
Muller, H. M., Crampton, J. M., della Torre, A., Sinden, R. and Crisanti, A. (1993) Members of a trypsin gene family in Anopheles gambiae are induced in the gut by blood meal. EMBO Journal, 12, 2891-2900.
Muller, P., Chouaibou, M., Pignatelli, P., Etang, J., Walker, E. D., Donnelly, M. J., Simard, F. and 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.
Muller, P., Donnelly, M. and Ranson, H. (2007) Transcription profiling of a recently colonised pyrethroid resistant Anopheles gambiae strain from Ghana. BMC Genomics, 8, 36.
Muller, P., Warr, E., Stevenson, B. J., Pignatelli, P. M., Morgan, J. C., Steven, A., Yawson, A. E., Mitchell, S. N., Ranson, H. and Hemingway, J. (2008b) Field-caught permethrin-resistant Anopheles gambiae overexpress CYP6P3, a P450 that metabolises pyrethroids. PLOS Genetics, 4.
Mutero, A., Pralavorio, M., Bride, J. and Fournier, D. (1994) Resistance-Associated Point Mutations in Insecticide-Insensitive Acetylcholinesterase. Proceedings of the National Academy of Sciences, 91, 5922-5926.
Mwangi, T. W., Ross, A., Marsh, K. and Snow, R. W. (2003) The effects of untreated bednets on malaria infection and morbidity on the Kenyan coast. Transactions of the Royal Society of Tropical Medicine and Hygiene, 97, 369 - 372.
N'Guessan, R., Asidi, A., Boko, P., Odjo, A., Akogbeto, M., Pigeon, O. and Rowland, M. (2010) An experimental hut evaluation of PermaNet® 3.0, a deltamethrin-piperonyl butoxide combination net, against pyrethroid-resistant Anopheles gambiae and Culex quinquefasciatus mosquitoes in southern Benin. Transactions of the Royal Society of Tropical Medicine and Hygiene, 104, 758-765.
N'Guessan, R., Corbel, V., Akogbeto, M. and 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, 199-206.
N'Guessan, R., Darriet, F., Guillet, P., Carnevale, P., Traore-Lamizana, M., Corbel, V., Koffi, A. A. and Chandre, F. (2003) Resistance to carbosulfan in Anopheles gambiae from Ivory Coast, based on reduced sensitivity of acetylcholinesterase. Medical and Veterinary Entomology, 17, 19-25.
Nabarro, D. (1999) Roll Back Malaria. Parassitologia, 41, 501-504.
Najera, J. A. and Zaim, M. (2001) Insecticides for indoor residual spraying, World Health Organistion, WHOPES.
Neafsey, D. E., Lawniczak, M. K. N., Park, D. J., Redmond, S. N., Coulibaly, M. B., Traoré, S. F., Sagnon, N., Costantini, C., Johnson, C., Wiegand, R. C., Collins, F. H., Lander, E. S., Wirth, D. F., Kafatos, F. C., Besansky, N. J., Christophides, G. K. and Muskavitch, M. A. T. (2010) SNP Genotyping Defines Complex Gene-Flow Boundaries Among African Malaria Vector Mosquitoes. Science, 330, 514-517.
NetMark (2010).
Newcomb, R. D., Campbell, P. M., Ollis, D. L., Cheah, E., Russell, R. J. and 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.
Newcomb, R. G. (1998) Two-sided confidence intervals for the single proportion: comparison of seven methods. Statistics in Medicine, 17, 857-872.
Nikou, D., Ranson, H. and Hemingway, J. (2003) An adult-specific CYP6 P450 gene is overexpressed in a pyrethroid-resistant strain of the malaria vector, Anopheles gambiae. Gene, 318, 91-102.

Niwa, R., Matsuda, T., Yoshiyama, T., Namiki, T., Mita, K., Fujimoto, Y. and Kataoka, H. (2004) CYP306A1, a cytochrome P450 enzyme, is essential for ecdysteroid biosynthesis in the prothoracic glands of Bombyx and Drosophila. Journal of Biological Chemistry, 279, 35942-35949.
Noor, A. M., Amin, A. A., Akhwale, W. S. and Snow, R. W. (2007) Increasing coverage and decreasing inequity in insecticide-treated bed net use among rural Kenyan children. PLoS Medicine, 4.
Noppun, V., Saito, T. and Miyata, T. (1989) Cuticular penetration of S-fenvalerate in fenvalerate-resistant and susceptible strains of the diamondback moth, Plutella xylostella (L.). Pesticide Biochemistry and Physiology, 33, 83-87.
Ntow, W. J., Drechsel, P., Botwe, B. O., Kelderman, P. and Gijzen, H. J. (2008) The Impact of Agricultural Runoff on the Quality of Two Streams in Vegetable Farm Areas in Ghana. Journal of Environmental Quality, 37, 696-703.
Ntow, W. J., Gijzen, H. J., Kelderman, P. and Drechsel, P. (2006) Farmer perceptions and pesticide use practices in vegetable production in Ghana. Pest Management Science, 62, 356-365.
Oakeshott, J. G., Devonshire, A. L., Claudianos, C., Sutherland, T. D., Horne, I., Campbell, P. M., Ollis, D. L. and Russell, R. J. (2005) Comparing the organophosphorus and carbamate insecticide resistance mutations in cholin- and carboxyl-esterases. Chemico-Biological Interactions, 157-158, 269-275.
Oakley, A. J., Lo Bello, M., Nuccetelli, M., Mazzetti, A. P. and Parker, M. W. (1999) The ligandin (non-substrate) binding site of human pi class glutathione transferase is located in the electrophile binding site (H-site). Journal of Molecular Biology, 291, 913-926.
Oduol, F., Xu, J., Niaré, O., Natarajan, R. and Vernick, K. D. (2000) Genes identified by an expression screen of the vector mosquito Anopheles gambiae display differential molecular immune response to malaria parasites and bacteria. Proceedings of the National Academy of Sciences of the United States of America, 97, 11397-11402.
Olsen, S. N. (2006) Applications of isothermal titration calorimetry to measure enzyme kinetics and activity in complex solutions. Thermochimica Acta, 448, 12-18.
Olsson, A. O., Baker, S. E., Nguyen, J. V., Romanoff, L. C., Udunka, S. O., Walker, R. D., Flemmen, K. L. and Barr, D. B. (2004) A Liquid Chromatography−Tandem Mass Spectrometry Multiresidue Method for Quantification of Specific Metabolites of Organophosphorus Pesticides, Synthetic Pyrethroids, Selected Herbicides, and DEET in Human Urine. Analytical Chemistry, 76, 2453-2461.
Ono, M., Swanson, J. J., M. Field, L., Devonshire, A. L. and D. Siegfried, B. (1999) Amplification and methylation of an esterase gene associated with insecticide-resistance in greenbugs, Schizaphis graminum (Rondani) (Homoptera: Aphididae). Insect Biochemistry and Molecular Biology, 29, 1065-1073.
Oppenoorth, F. J. and van, A. (1960) Allelic genes in the housefly producing modified enzymes that cause organophosphate resistance. Science, 132, 298-299.
Oppenoorth, F. J., van der Pas, L. J. T. and Houx, N. W. H. (1979) Glutathione S-transferase and hydrolytic activity in a tetrachlorvinphos-resistant strain of housefly and their influence on resistance. Pesticide Biochemistry and Physiology, 11, 176-188.
Ortelli, F., Rossiter, L. C., Vontas, J., Ranson, H. and Hemingway, J. (2003) Heterologous expression of four glutathione transferase genes genetically linked to a major insecticide-resistance locus from the malaria vector Anopheles gambiae. Biochemical Journal, 373, 957-963.
Ossovskaya, V., Koo, I. C., Kaldjian, E. P., Alvares, C. and Sherman, B. M. (2010) Upregulation of Poly (ADP-Ribose) Polymerase-1 (PARP1) in Triple-Negative Breast Cancer and Other Primary Human Tumor Types. Genes & Cancer, 1, 812-821.
OXITEC (2010) Open field trial demonstrates effectiveness of RIDL system for suppressing a target wild mosquito population. Oxitec Limited, Oxfordshire, England.

Paine, M. J., Scrutton, N. S., Munro, A. W., Guitierrez, A., Roberts, G. C. K. and Wolf, C. R. (2005) Electron transfer partners of cytochrom P450. In Cytochrome P450: Structure, mechanism, and biochemistry(Ed, Ortiz de Montellano, P. R.) Kluwer Academic/Plenum Publishers, New York, pp. 115-148.
Patil, N. S., Lole, K. S. and Deobagkar, D. N. (1996) Adaptive larval thermotolerance and induced cross-tolerance to propoxur insecticide in mosquitoes Anopheles stephensi and Aedes aegypti. Medical and Veterinary Entomology, 10, 277-282.
Patterson, L. H. and Murray, G. I. (2002) Tumour Cytochrome P450 and Drug Activation. Current Pharmaceutical Design, 8.
Pedra, J. H. F., McIntyre, L. M., Scharf, M. E. and Pittendrigh, B. R. (2004) Genome-wide transcription profile of field- and laboratory-selected dichlorodiphenyltrichloroethane (DDT)-resistant Drosophila. Proceedings of the National Academy of Sciences, 101, 7034-7039.
Pedrini, N., Mijailovsky, S. J., Girotti, J. R., Stariolo, R., Cardozo, R. M., Gentile, A. and Juarez, M. P. (2009) Control of pyrethroid-resistant Chagas disease vectors with entomopathogenic fungi. PLoS Neglected Tropical Diseases, 3, e434.
Peffly, R. L. (1959) Insecticide resistance in anophelines in eastern Saudi Arabia. Bull World Health Organ, 20, 757-776.
Pelosi, P., Zhou, J., Ban, L. and Calvello, M. (2006) Soluble proteins in insect chemical communication. Cellular and Molecular Life Sciences, 63, 1658-1676.
Perera, M., Hemingway, J. and Karunaratne, S. H. P. P. (2008) Multiple insecticide resistance mechanisms involving metabolic changes and insensitive target sites selected in anopheline vectors of malaria in Sri Lanka. Malaria Journal, 7, 168.
Phuc, H., Andreasen, M., Burton, R., Vass, C., Epton, M., Pape, G., Fu, G., Condon, K., Scaife, S., Donnelly, C., Coleman, P., White-Cooper, H. and Alphey, L. (2007) Late-acting dominant lethal genetic systems and mosquito control. BMC Biology, 5, 11.
Pinto, J., Lynd, A., Elissa, N., Donnelly, M. J., Costa, C., Gentile, G., Caccone, A. and do Rosario, V. E. (2006) Co-occurrence of East and West African kdr mutations suggests high levels of resistance to pyrethroid insecticides in Anopheles gambiae from Libreville, Gabon. Medical and Veterinary Entomology, 20, 27 - 32.
Pinto, J. o., Lynd, A., Vicente, J. L., Santolamazza, F., Randle, N. P., Gentile, G., Moreno, M., Simard, F. d. r., Charlwood, J. D., do RosÃ¡rio, V. l. E., Caccone, A., della Torre, A. and Donnelly, M. J. (2007) Multiple Origins of Knockdown Resistance Mutations in the Afrotropical Mosquito Vector Anopheles gambiae. PLoS ONE, 2, e1243.
Plapp, F. W. and Hoyer, R. F. (1968) Insecticide resistance in the house fly: Decreased rate of absorption as the mechanism of action of a gene that acts as an intensifier of resistance. Journal of Economic Entomology, 61.
Pluess, B., Tanser, F. C., Lengeler, C. and Sharp, B. L. (2010) Indoor residual spraying for preventing malaria. Cochrane Database of Systematic Reviews, 14.
PMI (2010) President's Malaria Initiative: Malaria operational plan - year three (FY2010) GHANA.
Poupardin, R., Reynaud, S., Strode, C., Ranson, H., Vontas, J. and David, J.-P. (2008) Cross-induction of detoxification genes by environmental xenobiotics and insecticides in the mosquito Aedes aegypti: Impact on larval tolerance to chemical insecticides. Insect Biochemistry and Molecular Biology, 38, 540-551.
Prapanthadara, L. A., Hemingway, J. and Ketterman, A. J. (1993) Partial Purification and Characterization of Glutathione S-Transferases Involved in DDT Resistance from the Mosquito Anopheles gambiae. Pesticide Biochemistry and Physiology, 47, 119-133.
Prashant, M., Pandey, A., Tiwari, M., Chandrashekar, K., Sidhu, O. P., Asif, M. H., Chakrabarty, D., Singh, P. K., Trivedi, P. K., Nath, P. and Tuli, R. (2010) Modulation of Transcriptome and Metabolome of Tobacco by Arabidopsis Transcription Factor, AtMYB12, Leads to Insect Resistance. Plant Physiology, 152, 2258-2268.

Pritchard, M. P., McLaughlin, L. and Friedberg, T. (2005) Cytochrome P450 Protocols: Establishment of Functional Human Cytochrome P450 Monooxygenase System in Escherichia coli. In Methods in Molecular Biology Vol. 320 (Eds, Phillips, I. R. and Shephard, E. A.) Humana Press Inc., NJ, pp. 19-29.

Protopopoff, N., Verhaeghen, K., Van Bortel, W., Roelants, P., Marcotty, T., Baza, D., D'Alessandro, U. and Coosemans, M. (2008) A significant increase in kdr in Anopheles gambiae is associated with an intensive vector control intervention in Burundi highlands. Tropical Medicine & International Health, 13, 1479 - 1487.

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

Rajatileka, S., Black, W. C., Saavedra-Rodriguez, K., Trongtokit, Y., Apiwathnasorn, C., McCall, P. J. and Ranson, H. (2008) Development and application of a simple colorimetric assay reveals widespread distribution of sodium channel mutations in Thai populations of Aedes aegypti. Acta Tropica, 108, 54 - 57.

Ramakrishna, V. and Elliott, R. (1959) Insecticide resistance in Anopheles gambiae in Sokoto Province. Transactions of the Royal Society of Tropical Medicine and Hygiene, 53, 102-109.

Ramirez, J. L., Garver, L. S. and Dimopoulos, G. (2009) Challenges and approaches for mosquito targeted malaria control. Current Molecular Medicine, 9, 116-130.

Ranson, H., Abdallah, H., Badolo, A., Guelbeogo, W., Kerah-Hinzoumbe, C., Yangalbe-Kalnone, E., Sagnon, N. F., Simard, F. and Coetzee, M. (2009) Insecticide resistance in Anopheles gambiae: data from the first year of a multi-country study highlight the extent of the problem. Malaria Journal, 8, 299.

Ranson, H., Claudianos, C., Ortelli, F., Abgrall, C., Hemingway, J., Sharakhova, M. V., Unger, M. F., Collins, F. H. and Feyereisen, R. (2002) Evolution of Supergene Families Associated with Insecticide Resistance. Science, 298, 179-181.

Ranson, H., Collins, F. and Hemingway, J. (1998) The role of alternative mRNA splicing in generating heterogeneity within the Anopheles gambiae class I glutathione S-transferase family. Proceedings of the National Academy of Sciences of the United States of America, 95, 14284-14289.

Ranson, H., Hemingway, J., Helmut, S. and Lester, P. (2005) Mosquito Glutathione Transferases. In Methods in Enzymology, Vol. Volume 401 Academic Press, pp. 226-241.

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

Ranson, H., N'Guessan, R., Lines, J., Moiroux, N., Nkuni, Z. and Corbel, V. (2010) Pyrethroid resistance in African anopheline mosquitoes: what are the implications for malaria control? Trends in Parasitology, 13, 13.

Ranson, H., Paton, M. G., Jensen, B., McCarroll, L., Vaughan, A., Hogan, J. R., Hemingway, J. and Collins, F. H. (2004) Genetic mapping of genes conferring permethrin resistance in the malaria vector, *Anopheles gambiae*. Insect Molecular Biology, 13, 379-386.

Ranson, H., Prapanthadara, L. a. and Hemingway, J. (1997) Cloning and characterization of two glutathione S-transferases from a DDT-resistant strain of Anopheles gambiae. Biochemical Journal, 324, 97-102.

Ranson, H., Rossiter, L., Ortelli, F., Jensen, B., Wang, X., Roth, C. W., Collins, F. H. and Hemingway, J. (2001) Identification of a novel class of insect glutathione S-transferases involved in resistance to DDT in the malaria vector Anopheles gambiae. Biochemical Journal, 359, 295-304.

Rashatwar, S. and Matsumura, F. (1985) Reduced calcium sensitivity of the sodium channel and the Na+/Ca2+ exchange system in the Kdr-type, DDT and pyrethroid resistant German cockroach, Blattella germanica. Comparative Biochemistry and Physiology C, 81, 97-103.

Raymond, M., Chevillion, C., Guillemaud, T., Lenormand, T. and Pasteur, N. (1998) An overview of the evolution of overproducted esterases in the mosquito *Culex pipiens*. Philosophical Transactions of The Royal Society London. Series B, Biological Sciences, 353, 1707-1711.

Rebers, J. E. and Riddiford, L. M. (1988) Structure and expression of a Manduca sexta larval cuticle gene homologous to Drosophila cuticle genes. Journal of Molecular Biology, 203, 411-423.

Reimer, L., Fondjo, E., Patchoka, S., Diallo, B., Lee, Y., Ng, A., Ndjemai, H. M., Atangana, J., Traore, S. F., Lanzaro, G. and Cornel, A. J. (2008) Relationship Between kdr Mutation and Resistance to Pyrethroid and DDT Insecticides in Natural Populations of Anopheles gambiae. Journal of Medical Entomology, 45, 260-266.

Reimer, L., Tripet, F., Slotman, M., Spielman, A., Fonjo, E. and Lanzaro, G. C. (2005) An unusual distribution of the kdr gene among populations of Anopheles gambiae on the island of Bioko, Equatorial Guinea. Insect Molecular Biology, 14, 683 - 688.

Ren, X., Hoiczyk, E. and Rasgon, J. L. (2008) Viral Paratransgenesis in the Malaria Vector *Anopheles gambiae*. PLoS Pathogens, 4, e1000135.

Richmond, C. S., Glasner, J. D., Mau, R., Jin, H. and Blattner, F. R. (1999) Genome-wide expression profiling in Escherichia coli K-12`. Nucleic Acid Research, 27, 3821-3835.

Riehle, M. M., Markianos, K., Niare, O., Xu, J., Li, J., Toure, A. M., Podiougou, B., Oduol, F., Diawara, S., Diallo, M., Coulibaly, B., Ouatara, A., Kruglyak, L., Traore, S. F. and Vernick, K. D. (2006) Natural Malaria Infection in Anopheles gambiae Is Regulated by a Single Genomic Control Region. Science, 312, 577-579.

Rodriguez, A. D., Penilla, R. P., Rodriguez, M. H. and Hemingway, J. (2006) Insecticide resistance management in a multi-resistant malaria vector scenario: A Mexican trial shows sustainability. Vol. 18 Public Health Journal.

Roh, S. W., Abell, G. C. J., Kim, K.-H., Nam, Y.-D. and Bae, J.-W. (2010) Comparing microarrays and next-generation sequencing technologies for microbial ecology research. Trends in Biotechnology, 28, 291-299.

Ross, R. (1910) The prevention of malaria, Murray., London.

Rozen, S. and Skaletsky, H. (2000) Primer3 on the WWW for general users and for biologist programmers. Methods in Molecular Biology, 132, 365 - 386.

RTI (2008) Indoor Residual Spraying (IRS) for Malaria Control Indefinite Quantity Contract (IQC) Task Order 1 Quarterly Report.

Ruden, D. M., Chen, L., Possidente, D., Possidente, B., Rasouli, P., Wang, L., Lu, X., Garfinkel, M. D., Hirsch, H. V. B. and Page, G. P. (2009) Genetical toxicogenomics in Drosophila identifies master-modulatory loci that are regulated by developmental exposure to lead. NeuroToxicology, 30, 898-914.

Sabourault, C., Guzov, V. M., Koener, J. F., Claudianos, C., Plapp, F. W. and Feyereisen, R. (2001) Overproduction of a P450 that metabolizes diazinon is linked to a loss-of-function in the chromosome 2 ali-esterase (MdαE7) gene in resistant house flies. Insect Molecular Biology, 10, 609-618.

Salinas, A. E. and Wong, M. G. (1999) Glutathione S-Transferases - A Review. Current Medicinal Chemistry, 6, 279-309.

Sandler, B. H., Nikonova, L., Leal, W. S. and Clardy, J. (2000) Sexual attraction in the silkworm moth: structure of the pheromone-binding-protein-bombykol complex. Chemistry & Biology, 7, 143-151.

Santolamazza, F., Calzetta, M., Etang, J., Barrese, E., Dia, I., Caccone, A., Donnelly, M., Petrarca, V., Simard, F., Pinto, J. and della Torre, A. (2008) Distribution of

knock-down resistance mutations in Anopheles gambiae molecular forms in west and west-central Africa. Malaria Journal, 7, 74.
Sawicki, R. M. (1978) Unusual response of DDT-resistant houseflies to carbinol analogues of DDT. Nature, 275.
Schena, M., Shalon, D., Davis, R. W. and Brown, P. O. (1995) Quantitative monitoring of gene expression pattern with complementary DNA microarrays. Science, 270, 467-470.
Schenkman, J. B. and Jansson, I. (2003) The many roles of cytochrome b5. Pharmacology & Therapeutics, 97, 139-152.
Scholte, E.-J., Ng'habi, K., Kihonda, J., Takken, W., Paaijmans, K., Abdulla, S., Killeen, G. F. and Knols, B. G. J. (2005) An Entomopathogenic Fungus for Control of Adult African Malaria Mosquitoes. Science, 308, 1641-1642.
Schuster, S. C. (2008) Next-generation sequencing transforms today's biology. Nature Methods, 5, 16-18.
Scott, J., Brogdon, W. and Collins, F. (1993) Identification of single specimens of the Anopheles gambiae complex by PCR. The American Journal of Tropical Medicine and Hygiene, 49, 520 - 529.
Service, M. (2004) Medical entomology for students, Cambridge University Press, Cambridge.
Service, M. W. (1993) Mosquito Ecology. London , Chapman & Hall.
Shahabuddin, M., Lemos, F. J., Kaslow, D. C. and Jacobs-Lorena, M. (1996) Antibody-mediated inhibition of Aedes aegypti midgut trypsins blocks sporogonic development of Plasmodium gallinaceum. Infection and Immunity, 64, 739-743.
Sharma, S. N., Shukle, R. P., Raghavendra, K. and Subbarao, S. K. (2005) Impact of DDT spraying on malaria transmission in Bareilly District, Uttar Pradesh, India. Journal of Vector Borne Disease, 42, 54-60.
Sharp, B., Ridl, F., Govender, D., Kuklinski, J. and Kleinschmidt, I. (2007) Malaria vector control by indoor residual insecticide spraying on the tropical island of Bioko, Equatorial Guinea. Malaria Journal, 6, 52.
Shen, B., Dong, H.-Q., Tian, H.-S., Ma, L., Li, X.-L., Wu, G.-L. and Zhu, C.-L. (2003) Cytochrome P450 genes expressed in the deltamethrin-susceptible and -resistant strains of Culex pipiens pallens. Pesticide Biochemistry and Physiology, 75, 19-26.
Shin, S. W., Kokoza, V. A. and Raikhel, A. S. (2003) Transgenesis and reverse genetics of mosquito innate immunity. Journal of Experimental Biology, 206, 3835-3843.
Shotkoski, F., Zhang, H.-G., Jackson, M. B. and ffrench-Constant, R. H. (1996) Stable expression of insect GABA receptors in insect cell lines promoters for efficient expression of Drosophila and mosquito Rdl GABA receptors in stably transformed mosquito cell lines. FEBS Letters, 380, 257-262.
Silverman, J. and Ross, M. H. (1994) Behavioral resistance of field-collected German cockroaches (Blattodea: Blattellidae) to baits containing glucose. Environmental Entomology, 23, 425-430.
Sim, C., Hong, Y. S., Vanlandingham, D. L., Harker, B. W., Christophides, G. K., Kafatos, F. C., Higgs, S. and Collins, F. H. (2005) Modulation of Anopheles gambiae gene expression in response to o'nyong-nyong virus infection. Insect Molecular Biology, 14, 475-481.
Sim, E., Lack, N., Wang, C.-J., Long, H., Westwood, I., Fullam, E. and Kawamura, A. (2008) Arylamine N-acetyltransferases: Structural and functional implications of polymorphisms. Toxicology, 254, 170-183.
Simmons, T. W., Jamall, I. S. and Lockshin, R. A. (1989) Selenium-independent glutathione peroxidase activity associated with glutathione S-transferase from the housefly, Musca domestica. Comparative Biochemistry and Physiology - Part B: Biochemistry 94, 323-327.
Singh, S. P., Coronella, J. A., Benes, H., Cochrane, B. J. and Zimniak, P. (2001) Catalytic function of Drosophila melanogaster glutathione S transferase

DmGSTS1-1 (GST-2) in conjugation of lipid peroxidation end products. European Journal of Biochemistry, 268, 2912-2923.
Skalsky, R., Vanlandingham, D., Scholle, F., Higgs, S. and Cullen, B. (2010) Identification of microRNAs expressed in two mosquito vectors, Aedes albopictus and Culex quinquefasciatus. BMC Genomics, 11, 119.
Smith, C. A., Want, E., O'Maille, G., Abagyan, R. and Siuzdak, G. (2006) XCMS: Processing mass spectometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. Analytical Chemistry, 78, 779-787.
Smyth, G. (2005) limma: Linear Models for Microarray Data. In Bioinformatics and Computational Biology Solutions using R and Bioconductor(Eds, Gentleman, R., Carey, V., Dudoit, S., Irizarry, R. and Huber, W.) New York: Springer Verlag, pp. 397-420.
Smyth, G. (2009) Correct Intensities for Background. In Linear Models for Microarray Data, Vol. 2010 (Ed, Smyth, G.) R Documentation.
Snow, R. W. and Marsh, K. (2010) Malaria in Africa: progress and prospects in the decade since the Abuja Declaration. The Lancet, 376, 137-139.
Snow, R. W., Rowen, K. M., Lindsay, S. M. and Greenwood, B. M. (1988) A trial permethrin treated bed nets prevention of malaria in Gambia children. 82, 838 - 842.
Soderlund, D. M. and Bloomquist, J. R. (1989) Neurotoxic Actions of Pyrethroid Insecticides. Annual Review of Entomology, 34, 77-96.
Soderlund, D. M. and Knipple, D. C. (2003) The molecular biology of knockdown resistance to pyrethroid insecticides. Insect Biochemistry and Molecular Biology, 33, 563-577.
Song, X. J., Huang, W., Shi, M., Zhu, M. Z. and Lin, H. X. (2007) A QTL for rice grain width and weight encodes a previously unknown RING-type E3 ubiquitin ligase. Nature Genetics, 39, 623-630.
Sparks, T. C., Lockwood, J. A., Byford, R. L., Graves, J. B. and Leonard, B. R. (1989) The role of behavior in insecticide resistance. Pesticide Science, 26, 383-399.
Spring, J. H., Robichaux, S. R. and Hamlin, J. A. (2009) The role of aquaporins in excretion in insects. The Journal of Experimental Biology, 212, 358-362.
Stevenson, B. J., Bibby, J., Pignatelli, P., Muangnoicharoen, S., O'Neill, P. M., Lian, L.-Y., Müller, P., Nikou, D., Steven, A., Hemingway, J., Sutcliffe, M. J. and Paine, M. J. I. (2011) Cytochrome P450 6M2 from the malaria vector Anopheles gambiae metabolizes pyrethroids: Sequential metabolism of deltamethrin revealed. Insect Biochemistry and Molecular Biology, In Press, Corrected Proof.
Storey, J. D. and Tibshirani, R. (2003) Statistical significance for genomewide studies. Proceedings of the National Academy of Sciences, 100, 9440-9445.
Strode, C., Steen, K., Ortelli, F. and Ranson, H. (2006) Differential expression of the detoxification genes in the different life stages of the malaria vector *Anopheles gambiae*. Insect Molecular Biology, 15, 523-530.
Strode, C., Wondji, C. S., David, J. P., Hawkes, N. J., Lumjuan, N., Nelson, D. R., Drane, D. R., Karunaratne, S., Hemingway, J. and Black, W. C. (2008) Genomic analysis of detoxification genes in the mosquito Aedes aegypti. Insect Biochemistry and Molecular Biology, 38, 113 - 123.
Stump, A. D., Atieli, F. K., Vulule, J. M. and Besansky, N. J. (2004) Dynamics of the pyrethroid knockdown resistance allele in Western Kenyan populations of Anopheles gambiae in response to insecticide treated bed net trials The American Journal of Tropical Medicine and Hygiene, 70, 591-596.
Sultan, M., Schulz, M. H., Richard, H., Magen, A., Klingenhoff, A., Scherf, M., Seifert, M., Borodina, T., Soldatov, A., Parkhomchuk, D., Schmidt, D., O'Keeffe, S., Haas, S., Vingron, M., Lehrach, H. and Yaspo, M.-L. (2008) A Global View of Gene Activity and Alternative Splicing by Deep Sequencing of the Human Transcriptome. Science, 321, 956-960.
Szabo, A., Stolz, L. and Granzow, R. (1995) Surface plasmon resonance and its use in biomolecular interaction analysis (BIA). Current Opinion in Structural Biology, 5, 699-705.

Tang, A. H. and Tu, C. P. (1994) Biochemical characterization of Drosophila glutathione S-transferases D1 and D21. Journal of Biological Chemistry, 269, 27876-27884.
Tats, A., Remm, M. and Tenson, T. (2006) Highly expressed proteins have an increased frequency of alanine in the second amino acid position. BMC Genomics, 7.
Ter Kuile, F. O., Terlouw, D. J., Phillips-Howard, P. A., Hawley, W. A., Friedman, J. F., Kariuki, S. K., Shi, Y. P., Kolczak, M. S., Lal, A. A., Vulule, J. M. and Nahlen, B. L. (2003a) Reduction of malaira during pregnancy by permethrin-treated bednets in an area of intense perennial malaria transmission in Westerb Kenya The American Journal of Tropical Medicine and Hygiene, 68, 50-60.
Ter Kuile, F. O., Terlouw, D. J., Phillips-Howard, P. A., Hawley, W. A., Friedman, J. F., Kolczak, M. S., Kariuki, S. K., Shi, Y. P., Kwena, A. M., Vulule, J. M. and Nahlen, B. L. (2003b) Impact of permethrin-treated bed nets on malaria and all-cause morbidity in young children in an area of intense perennial malaria transmission in western Kenya: cross-sectional survey. The American Journal of Tropical Medicine and Hygiene, 68, 100 - 107.
The Complex Trait Consortium (2003) The nature and identification of quantitative trait loci: a community's view. Nature Reviews Genetics, 4, 911-916.
Toba, G. and Aigaki, T. (2000) Disruption of the Microsomal glutathione S-transferase-like gene reduces life span of Drosophila melanogaster. Gene, 253, 179-187.
Tomancak, P., Beaton, A., Weiszmann, R., Kwan, E., Shu, S., Lewis, S. E., Richards, S., Ashburner, M., Hartenstein, V., Celniker, S. E. and Rubin, G. M. (2002) Systematic determination of patterns of gene expression during Drosophile embryogenesis. Genome Biology, 3.
Tomlin, C. D. S. (1997) The Pesticide Manual, British Crop Protection Council, Farnham, UK.
Toure, Y. T., Petrarca, V., Traore, S. F., Coulibaly, A., Maiga, H. M., Sankare, O., Sow, M., Di Deco, M. A. and Coluzzi, M. (1994) Ecological genetic studies in the chromosomal form Mopti of Anopheles gambiae s.str. in Mali, west Africa. Genetica, 94, 213 - 223.
Toure, Y. T., Petrarca, V., Traore, S. F., Coulibaly, A., Maiga, H. M., Sankare, O., Sow, M., Di Deco, M. A. and Coluzzi, M. (1998) The distribution and inversion polymorphism of chromosomally recognized taxa of the Anopheles gambiae complex in Mali, West Africa. Parassitologia, 40, 477 - 511.
Townsend, M. G. and Busvine, J. R. (1969) The mechanism of malathion-resistance in the blowfly Chrysomya putoria. Entomologia Experimentalis et Applicata, 12, 243-267.
Turker, M. S. (2002) Gene silencing in mammalian cells and the spread of DNA methylation. Oncogene, 21, 5388-5393.
Turner, T. L., Hahn, M. W. and Nuzhdin, S. V. (2005) Genomic islands of speciation in Anopheles gambiae. PLOS Biology, 3, e285.
United-Nations (2009) World Population Prospects (2008 revision). (Ed, Secretariat, P. D. o. t. D. o. E. a. S. A. o. t. U. N.) United Nations, New York. .
Upton, A., Johnson, N., Sandy, J. and Sim, E. (2001) Arylamine N-acetyltransferases - of mice, men and microorganisms. Trends in Pharmacological Sciences, 22, 140-146.
USAID (2007) Saving the lives of mother and children in Africa - first annual report. In The President's Malaria Initiative
van den Berg, H. and Takken, W. (2010) Integrated vector management (IVM): from concept to practice. Malaria Journal, 9, I12.
Vargo, M. A. and Colman, R. F. (2001) Affinity Labeling of Rat Glutathione S-Transferase Isozyme 1-1 by 17β-Iodoacetoxy-estradiol-3-sulfate. Journal of Biological Chemistry, 276, 2031-2036.
Vaughan, A., Hawkes, N. and Hemingway, J. (1997) Co-amplification explains linkage disequilibrium of two mosquito esterase genes in insecticide-resistant Culex quinquefasciatus. Biochemical Journal, 325, 359-365.

Verhaeghen, K., Van Bortel, W., Roelants, P., Backeljau, T. and Coosemans, M. (2006) 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.

Vijverberg, H. P. M., Van Der Zalm, J. M. and Van Den Bercken, J. (1982) Similar mode of action of pyrethroids and DDT on sodium channel gating in myelinated nerves. Nature, 295, 601-603.

Vincent, F., Davies, G. J. and Brannigan, J. A. (2005) Structure and Kinetics of a Monomeric Glucosamine 6-Phosphate Deaminase. Journal of Biological Chemistry, 280, 19649-19655.

Vinciotti, V., Khanin, R., D'Alimonte, D., Liu, X., Cattini, N., Hotchkiss, G., Bucca, G., de Jesus, O., Rasaiyaah, J., Smith, C. P., Kellam, P. and Wit, E. (2005) An experimental evaluation of a loop versus a reference design for two-channel microarrays. Bioinformatics, 21, 492-501.

Vlachou, D., Schlegelmilch, T., Christophides, G. K. and Kafatos, F. C. (2005) Functional genomic analysis of midgut epithelial responses in Anopheles during Plasmodium invasion. Current Biology, 15, 1185 - 1195.

Vontas, J., Bass, C., Koutsos, A. C., David, J. P., Kafatos, F. C., Louis, C., Hemmingway, J., Christophides, G. K. and Ranson, H. (2005) Gene expression in insecticide resistant and susceptible Anopheles gambiae strains constitutively or after insecticide exposure. Insect Molecular Biology, 14, 509 - 521.

Vontas, J., David, J. P., Nikou, D., Hemingway, J., Christophides, G. K., Louis, C. and Ranson, H. (2007) Transcriptional analysis of insecticide resistance in Anopheles stephensi using cross-species microarray hybridization. Insect Molecular Biology, 16, 315-324.

Vontas, J., Ranson, H. and Alphey, L. (2010) Transcriptomics and disease vector control. BMC Genomics, 8, 52.

Vontas, J. G., Small, G. J. and Hemingway, J. (2000) Comparison of esterase gene amplification, gene expression and esterase activity in insecticide susceptible and resistant strains of the brown planthopper, Nilaparvata lugens (Stål). Insect Molecular Biology, 9, 655-660.

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

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

Vulule, J. M., Beach, R. F., Atieli, F. K., McAllister, J. C., Brogdon, W. G., Roberts, J. M., Mwangi, R. W. and Hawley, W. A. (1999) Elevated oxidase and esterase levels associated with permethrin tolerance in Anopheles gambiae from Kenyan villages using permethrin-impregnated nets. Medical and Veterinary Entomology, 13, 239-244.

Vuylsteke, M., Peleman, J. D. and van Eijk, M. J. T. (2007) AFLP-based transcript profiling (cDNA-AFLP) for genome-wide expression analysis. Nature Protocols, 2, 1399-1413.

Walker, W. B., Smith, E. M., Jan, T. and Zwiebel, L. J. (2008) A functional role for Anopheles gambiae Arrestin1 in olfactory signal transduction. Journal of Insect Physiology, 54, 680-690.

Wang, J., Kean, L., Yang, J., Allan, A., Davies, S., Herzyk, P. and Dow, J. (2004) Function-informed transcriptome analysis of Drosophila renal tubule. Genome Biology, 5, R69.

Wang, J. Y., McCommas, S. and Syvanen, M. (1991) Molecular cloning of a glutathione S-transferase overproduced in an insecticide-resistant strain of the housefly (*Musca domestica*). Molecular & General Genetic, 227, 260-266.

Wang, M.-H., Marinotti, O., James, A. A., Walker, E., Githure, J. and Yan, G. (2010a) Genome-Wide Patterns of Gene Expression during Aging in the African Malaria Vector <italic>Anopheles gambiae</italic>. PLoS One, 5, e13359.
Wang, Y., Carolan, J. C., Hao, F., Nicholson, J. K., Wilkinson, T. L. and Douglas, A. E. (2010b) integrated metabonomic-proteomic analysis of an insect-bacterial symbiotic system. Journal of Proteome Research, 9, 1257-1267.
Wang, Y., Qiu, L., Ranson, H., Lumjuan, N., Hemingway, J., Setzer, W. N., Meehan, E. J. and Chen, L. (2008) Structure of an insect epsilon class glutathione S-transferase from the malaria vector Anopheles gambiae provides an explanation for the high DDT-detoxifying activity. Journal of Structural Biology, 164, 228-235.
Warmke, J. W., Reenan, R. A. G., Wang, P., Qian, S., Arena, J. P., Wang, J., Wunderler, D., Liu, K., Kaczorowski, G. J., Ploeg, L. H. T. V. d., Ganetzky, B. and Cohen, C. J. (1997) Functional Expression of Drosophila para Sodium Channels. The Journal of General Physiology, 110, 119-133.
Weetman, D., Wilding, C. S., Steen, K., Morgan, J. C., Simard, F. and Donnelly, M. J. (2010) Association Mapping of Insecticide Resistance in Wild *Anopheles gambiae* Populations: Major Variants Identified in a Low-Linkage Disequilbrium Genome. PLoS ONE, 5, e13140.
Wei, S. H., Clark, A. G. and Syvanen, M. (2001) Identification and cloning of a key insecticide-metabolizing glutathione S-transferase (MdGST-6A) from a hyper insecticide-resistant strain of the housefly Musca domestica. Insect Biochemistry and Molecular Biology, 31, 1145-1153.
Weill, M., Chandre, F., Brengues, C., Manguin, S., Akogbeto, M., Pasteur, N., Guillet, P. and Raymond, M. (2000) The kdr mutation occurs in the Mopti form of Anopheles gambiaes.s. through introgression. Insect Molecular Biology, 9, 451-455.
Weill, M., Lutfalla, G., Mogensen, K., Chandre, F., Berthomieu, A., Berticat, C., Pasteur, N., Philips, A., Fort, P. and 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. and 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.
White, B. J., Cheng, C., Simard, F., Costantini, C. and Besansky, N. J. (2010) Genetic association of physically unlinked islands of genomic divergence in incipient species of Anopheles gambiae. Molecular Ecology, 19, 925-939.
WHO (1957) Insecticides. 7th report of the expert committee on insecticides. . In WHO Tech. Rep. Ser. 125(Ed, WHO) Geneva.
WHO (1998a) *Test procedures for insecticide resistance monitoring in malaria vectors, bio-efficacy and persistence of insecticides on treated surfaces.* WHO/CDS/CPC/MAL/98.12. World Health Organisation. , Geneva.
WHO (1998b) *Techniques to detect insecticide resistance mechanisms (field and laboratory manual*). WHO/CDS/CPC/MAL/98.6. World Health Organisation, Geneva.
WHO (2000) The African Summit on Roll Back Malaria. Vol. (WHO/CDS/RBM/2000.17) Abuja, Nigeria.
WHO (2001) Supplies for monitoring insecticide resistance in disease vectors. Vol. WHO/CDS/CPE/PVC/20012 2001.
WHO (2002) Mortality and burden of disease estimates for WHO member states in 2002
WHO (2005a) Strategic orientation paper on prevention and control of malaria - for national and international programme officers involved in malaria control at country level. WHO/HTM/MAL/2005.1105. In ROLL BACK MALARIA DEPARTMENT OPERATIONS SUPPORT AND CAPACITY DEVELOPMENT(Ed, Aregawi, M.) WHO.
WHO (2005b) WHO Country Office Ghana Annual report.

WHO (2006a) Indoor residual spraying: Use of indoor residual spraying for scaling up global malaria control and elimination. In Global Malara Programme.
WHO (2006b) Pesticides and their application for the control of vectors and pests of public health importance. WHO/CDS/WHOPES/GCDPP/2006. In WHO Pesticide evaluation scheme (WHOPES)World Health Organisation, Department of Control of Neglected Tropical Diseases. .
WHO (2006c) WHO gives indoor use of DDT a clean bill of health for controlling malaria. WHO.
WHO (2007a) Implementation of Indoor Residual Spraying of Insecticides for Malaria Control in the WHO African Region Report. (Ed, Africa, W. H. O. f.) Vector Biology and Control Unit Division of Healthy Environments and Sustainable Development.
WHO (2007b) Insecticide-treated mosquito nets: a WHO position statement. In Global Malaria Programme.
WHO (2007c) The Use of DDT in Malaria Vector Control: WHO position statement. In Global Malaria Programme.
Widarto, H., Van Der Meijden, E., Lefeber, A., Erkelens, C., Kim, H., Choi, Y. and Verpoorte, R. (2006) Metabolomic Differentiation of Brassica rapa Following Herbivory by Different Insect Instars using Two-Dimensional Nuclear Magnetic Resonance Spectroscopy. Journal of Chemical Ecology, 32, 2417-2428.
Wilding, C., Weetman, D., Steen, K. and Donnelly, M. (2009) High, clustered, nucleotide diversity in the genome of Anopheles gambiae revealed through pooled-template sequencing: implications for high-throughput genotyping protocols. BMC Genomics, 10, 320.
Willamson, M. S., Denholm, I., Bell, C. A. and Devonshire, A. L. (1993) Knockdown resistance (kdr) to DDT and pyrethroid insecticides maps to a sodium channel gene locus in the housefly (Musca domestica). Molecular & General Genetics, 240, 17-22.
Willamson, M. S., Martinez-Torres, D., Hick, C. A. and Devonshire, A. L. (1996) Identification of mutations in the housefly para-type sodium channel gene associated with knockdown resistance (kdr) to pyrethroid insecticides. Molecular & General Genetic, 252, 51-60.
Williams, N. S., Gaynor, R. B., Scoggin, S., Verma, U., Gokaslan, T., Simmang, C., Fleming, J., Tavana, D., Frenkel, E. and Becerra, C. (2003) Identification and Validation of Genes Involved in the Pathogenesis of Colorectal Cancer Using cDNA Microarrays and RNA Interference. Clinical Cancer Research, 9, 931-946.
Willis, J. H., Iconomidou, V. A., Smith, R. F. and Hamodrakas, S. J. (2005) Cuticular proteins. In Comprehensive Molecular Insect Science, Vol. 4 (Eds, Gilbert, L. I., Iatrou, K. and Gill, S. S.) Elsevier, Oxford, UK, pp. 79-110.
Willoughby, L., Chung, H., Lumb, C., Robin, C., Batterham, P. and Daborn, P. J. (2006) A comparison of Drosophila melanogaster detoxification gene induction responses for six insecticides, caffeine and phenobarbital. Insect Biochemistry and Molecular Biology, 36, 934-942.
Winter, C. E., Giannotti, O. and Holzhacker, E. L. (1975) DDT-lipoprotein complex in the American cockroach hemolymph: A possible way of insecticide transport. Pesticide Biochemistry and Physiology, 5, 155-162.
Wittkopp, P. J., Haerum, B. K. and Clark, A. G. (2004) Evolutionary changes in *cis* and *trans* gene regulation. Nature, 430, 85-88.
Wittkopp, P. J., Haerum, B. K. and Clark, A. G. (2008) Independent effects of cis- and trans-regulatory variation on gene expression in Drosophila melanogaster. Genetics, 178, 1831-1835.
Wogulis, M., Morgan, T., Ishida, Y., Leal, W. S. and Wilson, D. K. (2006) The crystal structure of an odorant binding protein from Anopheles gambiae: Evidence for a common ligand release mechanism. Biochemical and Biophysical Research Communications, 339, 157-164.

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

Wondji, C. S., Morgan, J., Coetzee, M., Hunt, R. H., Steen, K., Black IV, W. C., Hemingway, J. and Ranson, H. (2007) Mapping a Quantitative Trait Locus (QTL) conferring pyrethroid resistance in the African malaria vector *Anopheles funestus*. BMC Genomics, 8.

Wongtrakul, J., Pongjaroenkit, S., Leelapat, P., Nachaiwieng, W., Prapanthadara, L. A. and Ketterman, A. J. (2010) Expression and characterization of three new glutathione transferases, an epsilon (AcGSTE2-2), omega (AcGSTO1-1), and theta (AcGSTT1-1) from Anopheles cracens (Diptera: Culicidae), a major Thai malaria vector. Journal of Medical Entomology, 47, 162-171.

Wongtrakul, J., Udomsinprasert, R. and Ketterman, A. J. (2003) Non-active site residues Cys69 and Asp150 affected the enzymatic properties of glutathione S-transferase AdGSTD3-3. Insect Biochemistry and Molecular Biology, 33, 971-979.

Wood, O. R., Hanrahan, S., Coetzee, M., Koekemoer, L. L. and Brooke, B. D. (2010) Cuticle thickening associated with pyrethroid resistance in the major malaria vector Anopheles funestus. Parasites & Vectors, 3, 67.

Wren, J. D., Kulkarni, A., Joslin, J., Butow, R. A. and Garner, H. R. (2002) Cross-hybridization on PCR-spotted microarrays. Engineering in Medicine an Biology Magazine, IEEE, 21, 71.

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

Wu, H., Yang, H. and Churchill, G. (2009) Package 'maanova'. In Tools for analyzing MicroArray experiments.CRAN, http://churchill.jax.org/.

Xu, P. X., Zwiebel, L. J. and Smith, D. P. (2003) Identification of a distinct family of genes encoding atypical odorant-binding proteins in the malaria vector mosquito, *Anopheles gambiae*. Insect Molecular Biology, 12, 549-560.

Xu, X., Dong, Y., Abraham, E. G., Kocan, A., Srinivasan, P., Ghosh, A. L., sinden, R. E., Ribero, J. M. C., Jacobs-Lorena, M., Kafatos, F. C. and Dimopoulos, G. (2005) Transcriptome analysis of Anopheles stephensi–Plasmodium berghei interactions. Molecular & Biochemical Parasitology, 142, 76-87.

Yadouleton, A., Asidi, A., Djouaka, R., Braima, J., Agossou, C. and Akogbeto, M. (2009) Development of vegetable farming: a cause of the emergence of insecticide resistance in populations of Anopheles gambiae in urban areas of Benin. Malaria Journal, 8, 103.

Yadouleton, A., Padonou, G., Asidi, A., Moiroux, N., Bio-Banganna, S., Corbel, V., N'Guessan, R., Gbenou, D., Yacoubou, I., Gazard, K. and Akogbeto, M. (2010) Insecticide resistance status in Anopheles gambiae in southern Benin. Malaria Journal, 9, 83.

Yang, I., Chen, E., Hasseman, J., Liang, W., Frank, B., Wang, S., Sharov, V., Saeed, A., White, J., Li, J., Lee, N., Yeatman, T. and Quackenbush, J. (2002) Within the fold: assessing differential expression measures and reproducibility in microarray assays. Genome Biology, 3.

Yang, Y., Chen, S., Wu, S., Yue, L. and Wu, Y. (2006) Constitutive overexpression of multiple cytorchrome P450 genes associated with pyrethroid resistance in Helicoverpa armigera. Journal of Economic Entomology, 99, 1784-1789.

Yawson, A. E., McCall, P. J., Wilson, M. D. and Donnelly, M. J. (2004) Species abundance and insecticide resistance of Anopheles gambiae in selected areas of Ghana and Burkina Faso. Medical and Veterinary Entomology, 18, 372 - 377.

Yawson, A. E., Weetman, D., Wilson, M. D. and Donnelly, M. J. (2007) Ecological Zones Rather Than Molecular Forms Predict Genetic Differentiation in the Malaria Vector Anopheles gambiae s.s. in Ghana. Genetics, 175, 751 - 761.

Yoshida, S., Shimada, Y., Kondoh, D., Kouzuma, Y., Ghosh, A. K., Jacobs-Lorena, M. and Sinden, R. E. (2007) Hemolytic C-type lectin CEL-III from sea cucumber expressed in transgenic mosquitoes impairs malaria parasite development. PLoS Pathogens, 3.

Yoshimi, T., Minowa, K., Karouna-Renier, N. K., Watanabe, C., Sugaya, Y. and Miura, T. (2002) Activation of a stress-induced gene by insecticides in the midge, Chironomus yoshimatsui. Journal of Biochemical and Molecular Toxicology, 16, 10-17.

Yukich, J., Tediosi, F. and Lengeler, C. (2007) Operations, costs and costeffectiveness of five insecticide-treated net programs (Eritrea, Malawi, Tanzania, Togo, Senegal) and two indoor residual spraying programs (Kwa-Zulu-Natal, Mozambique). Swiss Tropical Institute, Basel, Switzerland., pp. 130.

Zahurak, M., Parmigiani, G., Yu, W., Scharpf, R., Berman, D., Schaeffer, E., Shabbeer, S. and Cope, L. (2007) Pre-processing Agilent microarray data. BMC Bioinformatics, 8, 142.

Zaim, M., Aitio, A. and Nakashima, N. (2000) Safety of pyrethroid treated mosquito nets. Medical and Veterinary Entomology, 14, 1-5.

Zaim, M. and Jambulingam, P. (2007) *Global Insecticide Use for Vector-Borne Disease Control.* WHO/CDS/NTD/WHOPES/GCDPP/2007.2, WHO.

Zhang, M. and Scott, J. G. (1996) Cytochrome b5 Is Essential for Cytochrome P450 6D1-Mediated Cypermethrin Resistance in LPR House Flies. Pesticide Biochemistry and Physiology, 55, 150-156.

Zhang, X., Zhang, J. and Zhu, K. Y. (2010) Chitosan/double-stranded RNA nanoparticle-mediated RNA interference to silence chitin synthase genes through larval feeding in the African malaria mosquito (Anopheles gambiae). Insect Molecular Biology.

Zhou, J.-J., Field, L. M. and He, X. L. (2010) Insect Odorant-Binding Proteins: Do They Offer an Alternative Pest Control Strategy? Outlooks on Pest Management, 21, 31-34.

Zhou, J. J. (2010) Odorant-binding proteins in insects. Vitamins and Hormones, 83, 241-272.

Zucker, S. D., Goessling, W., Ransil, B. J. and Gollan, J. L. (1995) Influence of glutathione S-transferase B (ligandin) on the intermembrane transfer of bilirubin. Implications for the intracellular transport of nonsubstrate ligands in hepatocytes. The Journal of Clinical Investigation, 96, 1927-1935.

Field, Genetic, and Modeling Approaches Show Strong Positive Selection Acting upon an Insecticide Resistance Mutation in *Anopheles gambiae s.s.*

Amy Lynd,[1] David Weetman,[1] Susana Barbosa,[2] Alexander Egyir Yawson,[3] Sara Mitchell,[1] Joao Pinto,[4] Ian Hastings,[2] and Martin J. Donnelly*[,1]

[1]Vector Group, Liverpool Tropical School of Medicine, Liverpool, United Kingdom

[2]Molecular and Biochemical Parasitology Group, Liverpool Tropical School of Medicine, Liverpool, United Kingdom

[3]Biotechnology and Nuclear Agriculture Research Institute, Ghana Atomic Energy Commission, Accra, Ghana

[4]Centro de Malária e outras Doenças Tropicais, Instituto de Higiene e Medicina Tropical, Universidade Nova de Lisbon, Lisbon, Portugal

***Corresponding author:** E-mail: m.j.donnelly@liv.ac.uk.

Associate editor: John H McDonald

Research article

Abstract

Alleles subject to strong, recent positive selection will be swept toward fixation together with contiguous sections of the genome. Whether the genomic signatures of such selection will be readily detectable in outbred wild populations is unclear. In this study, we employ haplotype diversity analysis to examine evidence for selective sweeps around knockdown resistance (*kdr*) mutations associated with resistance to dichlorodiphenyltrichloroethane and pyrethroid insecticides in the mosquito *Anopheles gambiae*. Both *kdr* mutations have significantly lower haplotype diversity than the wild-type (nonresistant) allele, with *kdr L1014F* showing the most pronounced footprint of selection. We complement these data with a time series of collections showing that the *L1014F* allele has increased in frequency from 0.05 to 0.54 in 5 years, consistent with a maximum likelihood–fitted selection coefficient of 0.16 and a dominance coefficient of 0.25. Our data show that strong, recent positive selective events, such as those caused by insecticide resistance, can be identified in wild insect populations.

Key words: *Anopheles gambiae*, malaria, insecticide resistance, selection.

Introduction

The sequencing of the *Anopheles gambiae* genome has opened up the possibility for genome-wide single nucleotide polymorphism (SNP)–based association mapping studies that have been successful in identifying positively selected loci in the human genome (Sabeti et al. 2002, 2007; Bersaglieri et al. 2004). The resolution of the association mapping approach is defined by the probability that recombination will have broken down the association between markers and a trait-associated functional polymorphism. Data from extensive resequencing of (primarily) detoxification genes in samples from wild populations of *A. gambiae* revealed a very high frequency of segregating sites (Wilding et al. 2009), consistent with high rates of recombination (Begun and Aquadro 1992; Begun et al. 2007) and/or a long history of outbreeding. In isofemale lab strains of *Drosophila spp.*, it has been possible to observe selective sweeps around insecticide resistance–associated loci (Schlenke and Begun 2004; Aminetzach et al. 2005), but how long these signatures persist in wild populations is unknown. In this paper, we use linkage disequilibrium (LD)–based haplotype diversity analysis (Sabeti et al. 2006) to investigate the pattern of molecular genetic variation associated with insecticide resistance mutations at the pyrethroid and dichlorodiphenyltrichloroethane (DDT) knockdown resistance locus, *kdr*, in the African malaria mosquito *A. gambiae s.s.* Furthermore, as a corollary of this indirect genetic approach we demonstrate, using a series of temporal collections, a dramatic increase in *kdr* frequency in a population of *A. gambiae s.s.* over a period of approximately 72 generations. Data from these temporal collections are used to estimate the selection and dominance coefficients operating on *kdr* in the field to illustrate the potential levels of selection necessary to produce the patterns of LD we observe.

Insecticide-treated bed nets are the principal method for preventing malaria in sub-Saharan Africa. Currently, pyrethroids are the only class of insecticides licensed for use on nets, and there is concern that resistance will compromise control programs. To date the most commonly recorded resistance mechanism is termed "knockdown resistance" and results from single–base pair mutations in the voltage-gated sodium channel. The sodium channel gene, located within division 20C near the centromere of chromosome 2L, codes for a protein that is the target site of pyrethroid insecticides. Two alternative single–base pair mutations have been found in *A. gambiae*, and these *kdr* mutations can cause target-site insensitivity to pyrethroids as well as cross-resistance to DDT. The substitutions cause amino acid changes at codon 1014 within the transmembrane structure of segment 6 in domain II of the voltage-gated sodium channel (numbering according to the

Table 1. Origin and *kdr* Genotype of Specimens Used in the Study.

Population	Year Collected	Total *N*	Form	Number of Each *kdr* Genotype					
				L1014S/ L1014S	*L1014F/ L1014F*	*L1014F/ L1014S*	*L1014S/ wt*	*L1014F/ wt*	Wt/wt
Asembo Bay, Kenya, 00°10′S, 34°22′E	2005[1]	48	S	11	—	—	17	—	20
Dienga, Gabon, 01°52′S, 12°40′E	1999–2000[2]	30	S	—	—	—	4	2	24
Bakoumba, Gabon, 01°49′S, 13°01′E	1999–2000[2]	42	S	—	5	8	5	7	17
Libreville, Gabon, 00°22′N, 09°26′E	1999–2000[2]	73	S	34	8	31	—	—	—
Okyereko and Accra area, Ghana, 05°24.9′N, 00°36.6′W, 05°38′N, 00°15′E	2002[3]	35	S	—	33	—	—	2	—
Okyereko, Ghana, 05°24.9′N, 00°36.6′W	2002[3]	30	M	—	—	—	—	2	28

NOTE.—The population name and total numbers of each DNA sample utilized. Molecular form is indicated, and the numbers of each *kdr* genotype are shown. Additional information on the collection sites may be obtained from the publications where the specimens are originally described: [1]Müller et al. (2008), [2]Pinto et al. (2006), and [3]Yawson et al. (2004); wt, wild type.

housefly *para* sequence, GenBank X96668). The *L1014F* mutation, a leucine to phenylalanine change, was first observed in West Africa (Martinez-Torres et al. 1998), and the same substitution has been observed in a diverse array of insects (Davies et al. 2007a). A second substitution, *L1014S*, was observed more recently in East African *A. gambiae* (Ranson et al. 2000) and involves the adjacent base of the same codon, resulting in a leucine to serine change.

There are two incipient species within the nominal taxon *A. gambiae s.s.* that are characterized by mutations on the X chromosome and are termed M and S form. The distribution of the *kdr* mutation is not uniform either within or between forms, although in general *kdr* alleles have been found at much higher frequencies in *A. gambiae s.s.* S-form samples compared with M-form samples (reviewed in Santolamazza et al. 2008). The reasons for the differences in distribution remain unclear because little is known about the origins of the *kdr* mutations and the selection pressures acting upon them in wild populations. In a sample from Benin, the *L1014F* was found in tight LD with two upstream intronic polymorphisms in both M- and S-form individuals. The two upstream polymorphisms associated with the *L1014F* variant were not found in wild-type M-form individuals but were common in wild-type S-form individuals, suggestive of an introgression event from S-form to M-form populations (Weill et al. 2000). This linkage between *kdr* and the intronic polymorphisms was not seen in M-form individuals from Bioko Island and was thought to indicate de novo mutation (Reimer et al. 2008). More recently, a study of S-form specimens from 15 countries suggested that the *L1014F* and *L1014S* mutations have both arisen independently on at least two separate occasions (Pinto et al. 2007).

Samples were obtained from three regions in sub-Saharan Africa; Kenya (East Africa) *A. gambiae*, S molecular form, *kdr L1014S* allele present; Ghana (West Africa) both M and S molecular form, *kdr L1014F* allele present; Gabon (Central Africa) S molecular form, both *L1014S* and *L1014F* *kdr* alleles present.

These population samples allow us to address a number of questions.

1. Available evidence suggests that the *L1014S* mutation has high penetrance for a DDT-resistant phenotype but lower penetrance for a pyrethroid-resistant phenotype than the *L1014F* mutation (Ranson et al. 2000). DDT was banned in Kenya in 1990, and we can investigate the signature of positive selection associated with weaker selection or recombination and relaxed selection.
2. The populations from central Africa are some of the few locations where both *L1014F* and *L1014S* alleles are observed sympatrically (Santolamazza et al. 2008). Indeed, in an earlier study, a significant, albeit marginal, *L1014F/ L1014S* heterozygote excess was observed in samples from Libreville, Gabon (Pinto et al. 2006). By comparing patterns of LD around the three alleles, we investigate whether the unusually high frequency of the *L1014S* allele in these populations (63%; Pinto et al. 2006) is a result of a recent selective sweep.
3. In many S-form populations in West Africa, including our collections from Ghana, the *L1014F* allele is close to fixation. In the absence of wild-type alleles, we are unable to control for local variation in recombination rates (Sabeti et al. 2007), and it is therefore impossible to ascribe patterns of LD to a positive selection event. Recently developed approaches such as cross-population extended haplotype homozygosity (EHH) have been developed to allow interpopulation comparisons in instances where alleles proceed to near fixation in some populations (Sabeti et al. 2007), but in our system resistance alleles may have multiple origins, presenting a confounding variable (Pinto et al. 2007). However, the presence of sympatric M-form populations in southern Ghana (Yawson et al. 2004, 2007) allows us to both document the increase in frequency of the same *L1014F* haplotype, following an introgression event, over a period of 5 years and estimate the selection and dominance coefficients associated with the signatures of positive selection.

Materials and Methods

Sample Sites, DNA Extraction, and Species Identification

Adult female *A. gambiae s.s.* mosquitoes used in this study were obtained from aspirator and pyrethroid knockdown collections from the field in various geographic locations (table 1). DNA was extracted from single female *A. gambiae* using either a modified Livak method or a phenol–chloroform method (Livak 1984; Ballinger-Crabtree et al. 1992). Species identification polymerase

chain reaction (PCR) was carried out on *A. gambiae s.l.* according to the protocol (Scott et al. 1993). Reactions were then digested with *Cfo*I restriction enzyme for 24 h at 37 °C in order to type *A. gambiae s.s* mosquitoes to M and S form (Fanello et al. 2002), and products visualized under UV light after electrophoresis on a 2% agarose Tris/borate/EDTA (TBE) gel with ethidium bromide. *Kdr* genotypes were determined by allele-specific PCR, heated oligonucleotide ligation assay (Lynd et al. 2005), or Taqman assay (Bass et al. 2007) depending upon year of collection.

Sodium Channel SNP Identification

The voltage-gated sodium channel gene is nearly 74 kbp in length and is composed of 35 exons including two duplicate exons (Davies et al. 2007a). Ten regions of the sodium channel were amplified by PCR for direct sequencing. Where possible, primers were designed to bind within exons to produce amplicons that spanned an intron with a maximum size of 1.5 kbp. Exons (numbering as Davies et al. 2007a) 1–2, 3, 4, 7–9, 13–14, 15–17, 20c, 23–24, 28–30, and 32–33 were selected as targets for sequencing. Primer and amplification details are provided (supplementary table 1, Supplementary Material online). Sequencing for SNP detection was carried out on up to 12 individuals of known *kdr* genotype from Ghana, São Tomé, Gabon, Angola, Mozambique, Malawi, and Kenya, from a susceptible laboratory strain (KISUMU), and from a permethrin tolerant resistant laboratory strain (reduced susceptibility to permethrin), both originating from Kenya. PCR products were cleaned using a Mini Elute PCR Purification kit (Qiagen) and then sequenced in both directions. Sequences were aligned using Bioedit software version 7.0.5.2 (Hall 1999) and then manually annotated for polymorphisms and ambiguities.

In addition, seven M-form individuals from Accra, Ghana, homozygous for the *L1014F* allele were bidirectionally sequenced across PCR amplicons 13–14, 15–17, and 21 to determine the associated haplotype of the *kdr* allele in this population.

SNP Screening

SNPs discovered through resequencing were screened in the large-scale SNP detection study using the SNPStart Primer Extension Kit on the Beckman CEQ 8000 Genetic Analysis System. Details of SNPs both included and excluded from the SNP screening are given in supplementary table 2, Supplementary Material online. Multiplex PCR was carried out to amplify the regions of DNA containing SNPs of interest, including a region of exon 20 and the preceding intron to allow high-throughput detection of the *kdr* mutation and three other well-characterized SNPs (Weill et al. 2000; Diabate et al. 2004; Pinto et al. 2006) (primers and reaction conditions detailed in supplementary table 3, Supplementary Material online). Products were visualized on a 2% TBE agarose gel. Successfully multiplexed samples were prepared for subsequent SNP extension by *Exo*I/shrimp alkaline phosphatase (SAP) enzymatic digestion. Interrogation primers were then designed for each individual SNP chosen for investigation according to the manufacturers' recommendations (supplementary table 4, Supplementary Material online). Single-base extension to the 3′ end of the interrogation primer by a dye terminator molecule, corresponding to the nucleotide found at the SNP location, was carried out using a GenomeLab SNPStart Primer Extension Kit (Beckman Coulter, Amersham, UK). The SAP-digested product was then scored on the Beckman CEQ 8000 Genetic Analysis System.

Data Analysis

As reviewed exhaustively by Sabeti et al. (2006), there are numerous statistical tests of positive selection which differ in their ability to detect selection events on different timescales. For the present SNP data set, it is not possible to use the suite of sequence-based tests that compare synonymous/nonsynonymous differences or detect an excess of rare alleles. We are therefore fortunate that on the timescales in which the emergence, and selection, of insecticide resistance is likely to occur, estimates of interpopulation divergence (e.g., based on *F* statistics) and screens of LD around selected versus wild-type alleles are likely to be the two most powerful analytical approaches. With the sample sizes available in our study, single-marker analyses based on *F*-statistic estimates would perform better as indicators of selection when markers can be typed at a more coarse scale, with consequently enhanced signal:noise ratio. However, with sample size constraints the signal would be difficult to localize. By contrast, long-range haplotype analyses, such as EHH (Sabeti et al. 2002) analysis, perform very well at a fine physical scale in identifying narrow candidate regions (Sabeti et al. 2006).

EHH analysis was carried out to assess the patterns of LD associated with wild type and the two *kdr* alleles. EHH can be defined as the probability that two random chosen chromosomes carrying the core (e.g., the wild-type or *kdr* allele) haplotype of interest are identical by descent. This approach first identifies core haplotypes surrounding the locus of interest and then examines the decay in LD from these core haplotypes to the surrounding loci. The resulting EHH can be used as evidence of recent positive selection at a locus in haplotypes that have high frequency and high EHH (Sabeti et al. 2002). EHH analysis requires haplotype information that cannot be empirically determined from the genotype data gathered by the methods used in this study. Therefore, haplotypes were inferred using PHASE software version 2.1.1 using default parameters (Stephens et al. 2001; Stephens and Scheet 2005). PHASE utilizes a Bayesian coalescent–based approach to determine phase and allows for varying rates of recombination at each SNP interval. The method is based on the idea that an unresolved haplotype is more likely to be the same or be similar to a previous haplotype. This approach was found to outperform other methods available for autosomal human data sets (Stephens et al. 2001; Stephens and Scheet 2005). Data were analyzed together rather than as separate subpopulations because 1) previous studies found this to be more accurate and 2) haplotype determination methods of this nature are relatively insensitive to departures

from Hardy–Weinberg equilibrium so are fairly robust to population substructuring. This approach is also more conservative than determining haplotypes for individual populations because the latter is liable to lead to an underestimation in differences in haplotype frequencies (Stephens and Scheet 2005). Phase reconstruction was executed ten times upon the total data set, and differences in counts of best haplotypes were noted.

The estimated haplotypes obtained from PHASE were used as input for EHH analysis implemented by SWEEP version 2.1.1 (Sabeti et al. 2002). Core haplotypes were selected manually to include only the two adjacent *kdr*-causing loci. Significance of EHH values is usually assigned through comparison to an empirically generated null distribution from other regions of the genome. However, given that we had already identified the causal mutations of interest, we were able to make a comparison of patterns of LD around wild-type and resistant cores. The primary advantage of this approach is that it is not subject to the genome-wide variations in recombination rate which can affect the null distribution approach in species lacking detailed recombination maps. Significant differences in EHH values were determined in two ways: 1) Within country samples, at individual SNP positions with nonoverlapping 95% confidence intervals (CIs). These CIs were calculated at each SNP position using a bootstrapping procedure, carried out in SAS version 9 software. Resampling was carried out 1,000 times. 2) Across all SNPs within and among country samples, the diversity of the different *kdr* allele–bearing haplotypes was compared using sign tests, implemented by SPSS 14. Where exact sign test probabilities could not be calculated, a Monte Carlo procedure with 10,000 permutations was performed. The sequential Bonferroni procedure was applied to determine statistical significance following correction for multiple testing (Holm 1979). Although our data—EHH values at each SNP position—are not independent, it is this nonindependence caused by LD that will cause departure from the null hypothesis of equality of median EHH values. Therefore, the null hypothesis remains that there is no difference in median EHH between *kdr* and wild-type alleles. Bifurcation plots were also created using the SWEEP software. In a bifurcation plot, the core haplotype is represented as a black circle. Each SNP, moving out from the core both upstream and downstream of the *kdr* locus, is a potential site for a bifurcation that would result from the presence of two segregating alleles. Therefore, the diagram provides a means of displaying the breakdown in LD at increasing distance from the core haplotypes. The radius of the circle at each node is proportional to the number of individuals with that haplotype.

Calculation of Selection and Dominance Coefficients

The spread of the *L1014F* allele was modeled using the standard recursive population genetic formula:

$$p' = \frac{p^2(1+s) + p(1-p)(1+hs)}{\bar{W}}, \quad (1)$$

where p is the frequency of the *L1014F* allele, p' is the frequency in the next generation, s is the selective coefficient of the resistance mutation, h is the dominance coefficient (1 = complete dominance, 0 = complete recessivity), and W is the normalizing factor (Maynard-Smith 1998).

Tracking allele frequencies over time requires three input parameters: initial allele frequency at time zero, s, and h. Estimates of all three unknown parameters were obtained by maximum likelihood assuming a binomial distribution of observed allele frequencies around the predicted frequency. The analysis was performed in R (http://www.r-project.org) using maximum likelihood functions and optimizing routines. The generation time was set at the standard of one generation per calendar month (Lehmann et al. 1998).

Results

SNP Discovery and Screening

Ten genomic regions of a combined length of ≈6.5 kb of DNA, spanning a region of ≈73 kb of the voltage-gated sodium channel, were amplified and sequenced in *A. gambiae s.s.* individuals from seven countries across sub-Saharan Africa. A total of 62 potential SNPs were found, of which 14 were exonic (supplementary table 1, Supplementary Material online). Six intronic indels were observed, usually in poly-A or tandem AT repeats (supplementary table 1, Supplementary Material online). On average, there was one SNP every 106 bp, which represents a low SNP frequency for *A. gambiae*, but similar to other genes in the same genomic locality (chromosome 2L division 20; Wilding et al. 2009). Thirty-two SNPs, including the two *kdr* mutations, were selected for screening in 258 individuals. In S-form individuals, the SNP adjacent to the core in the upstream (centromeric) direction was excluded from further analysis as it was found to be monomorphic. Details of the populations and associated *kdr* genotypes are given in table 1. The genotypic data were resolved into haplotypes with ten runs of the analysis. In only one instance, did the replicate runs resolve a novel estimated haplotype, which in a subsequent comparative analysis was found to exert no qualitative effect on the results. Therefore, all analyses reported here are based upon the haplotypes resolved in the vast majority of the phasing runs.

EHH analysis was carried out to assess the patterns of LD associated with the wild-type and the two *kdr* alleles. The intronic SNPs that have been used to identify the origin of the *kdr* mutations were the proximate SNPs in the centromeric direction (Weill et al. 2000; Pinto et al. 2007). LD decay was examined between these core haplotypes and the remaining 29 or 30 SNP loci (for S or M forms, respectively).

Only two core-alleles were present in the western Kenyan sample: wild type and *L1014S*. In the downstream telomeric direction, EHH decays at a similar rate for both wild type and *L1014S*, but there was a marked contrast between alleles in the centromeric direction, with entirely nonoverlapping confidence limits from just a few kilobases away from the core (figs. 1 and 2*A*). In the Gabonese collection, the difference between resistance-associated alleles

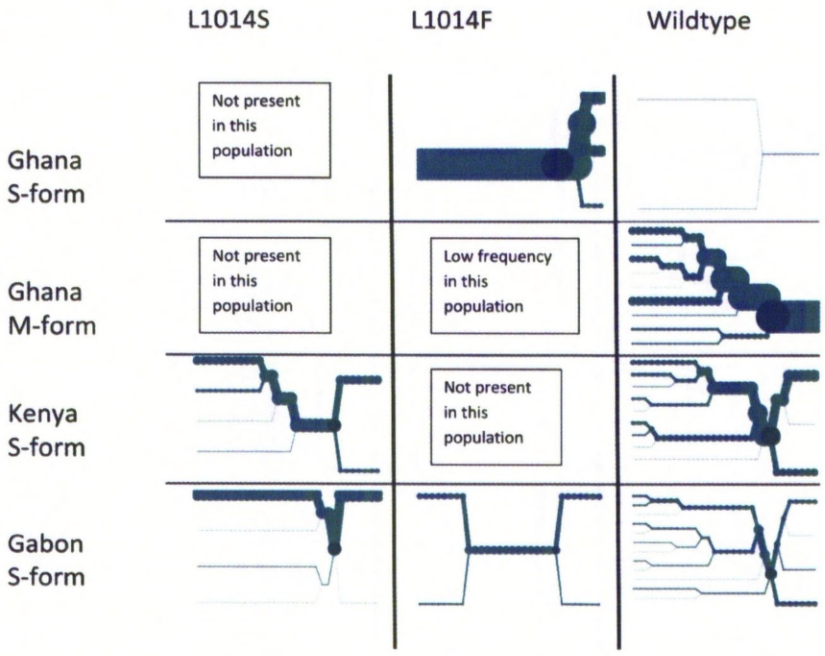

FIG. 1. Bifurcation plots showing patterns of recombination in the centromeric (5′ toward the left) and telomeric (3′ toward the right) directions. The core is marked by the dark circle, and each of the 29/30 SNPs is represented by a node and a recombination event is represented by a bifurcation. The diameter of the circle at each SNP node is proportional to the numbers of individuals with the same long-range haplotype at that position. No bifurcation plot is shown for the *L1014F* core in Ghanaian M-form populations as only a single haplotype was observed (see Results).

and wild type was even more marked with significantly lower EHH in the wild type in both centromeric and telomeric directions less than 5 kb from the core (fig. 2*B*). Indeed, both the *L1014F* and the *L1014S* resistance mutations showed little haplotype bifurcation in the Gabon samples over the length of the sodium channel (fig. 1), suggesting a relatively recent origin for both these mutations accompanied by a strong selective sweep. The patterns of LD are most marked around the resistant *L1014F* haplotype in Ghanaian S-form samples in which the *L1014F kdr* allele was at very high frequency (figs. 1 and 2*C*), as would be expected given the near fixation of this allele in southern Ghana in the S molecular form (mean frequency = 0.96; 95% CI 0.95–0.97) (Yawson et al. 2004). The presence of only two wild-type haplotypes in the sample prevent any meaningful comparison of LD decay, but it should be noted that there was complete LD over the entire 64-kb length of the sodium channel in the centromeric direction. The wild-type allele, observed in the Ghanaian M-form populations (figs. 1 and 2*C*), showed marked LD, only in the telomeric direction, between exons 20 and 32, the opposite directional asymmetry to the *L1014F* mutation in Ghana S-form populations. Although simulation studies have shown that LD decay may be asymmetric even when rates of mutation and recombination are constant (Kim and Stephan 2002), it is possible that the LD observed in these samples may reflect the presence of one or more hitherto overlooked selectively advantageous mutants, although we cannot rule out recombination with unsampled haplotypes (supplementary table 5, Supplementary Material online). Davies et al. (2007b) have summarized that there are a number of additional nonsynonymous changes observed in a variety of taxa, and detailed association mapping studies are presently underway to investigate this phenomenon. Comparing overall levels of EHH for the whole 72.6-kb regions typed, it is interesting to note that median EHH values are statistically indistinguishable for the same allele typed in different populations (table 2) and that a clear hierarchy of evidence for selective sweeps emerged. Median EHH levels were highest for the *L1014F* resistance mutation, followed by those for the *L1014S* mutations, with the lowest for the wild-type allele (table 2). The only exception to this pattern was within the Gabonese sample, the only one in which both resistance alleles were present, where median EHH was equal for the two resistance alleles. Nevertheless, despite the possibilities of different origins of the same allele, and local variation in recombination rates, EHH levels across the genomic region investigated suggest some degree of commonality in selection across populations for each allele, although the actual rate of change in LD with distance can be quite complex and dependent on direction from the core (figs. 1 and 2).

We investigated temporal change in the frequency of the *L1014F* allele and associated haplotype in sympatric populations of M-form individuals in a subset of the populations previously described by Yawson et al. (2004). Using the data reported in Yawson et al. (2004), we estimated the

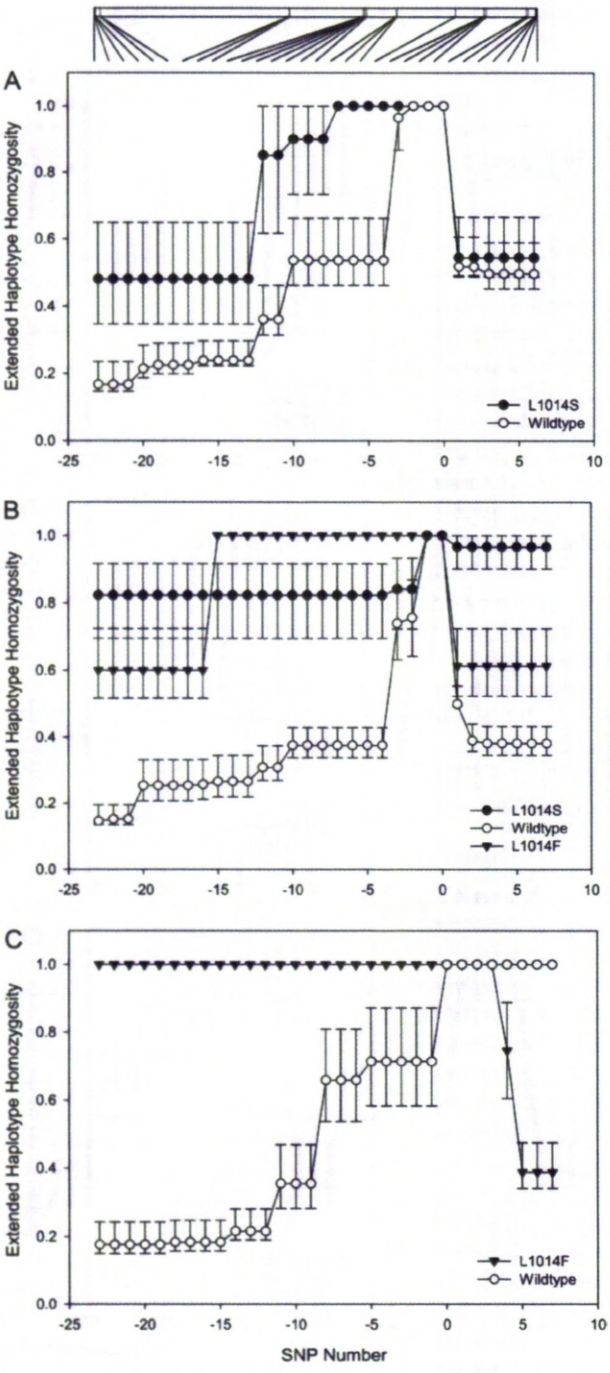

FIG. 2. EHH analysis showing LD decay with increasing distance from the core (marked as the origin on the *x* axis). The 95% CIs were estimated by bootstrapping (see Materials and Methods). The *x* axis is ordinal, negative numbers are in the centromeric direction and positive numbers in the telomeric direction. The scale bar at the top of the figure is 72.6 kb in length and shows the physical distance between the SNPs. (*A*) Kenya data for *L1014S* and wild-type alleles; (*B*) Gabon data for *L1014S*, *L1014F*, and wild-type alleles, and (*C*) Ghana data for *L1014F* (S form) and wild type (M form).

L1014F allele frequency in M-form populations from around Accra, southern Ghana (≈30 km diameter collection area), during 2002 ($freq_{L1014F}$ = 0.03; 95% CI 0.01–0.05). Additional screening in 2007 and 2008 from the same greater Accra regions revealed that within 5 years, this frequency had reached $freq_{L1014F}$ = 0.54 (95% CI 0.49–0.60; fig. 3). The data from years 2007 and 2008 are reported here for the first time. Phasing of the SNP genotypes of two M-form individuals with a wild-type/*L1014F* genotype showed that the *L1014F*-associated haplotype was identical to that found in the S form. This was confirmed by sequences obtained from seven M-form individuals collected from Accra, Ghana in 2008, which were homozygous for the *L1014F* allele (supplementary table 6, Supplementary Material online). Therefore, the *L1014F* allele, which has increased in frequency in M-form populations, is the same that has been putatively swept toward fixation in sympatric S-form populations. Introgression of *kdr* alleles between forms has been documented previously (Weill et al. 2000) and is unsurprising given that in southern Ghana there is a low but temporally stable level of interform matings (Yawson et al. 2004, 2007).

Using a maximum likelihood estimation procedure with random starting values for selection coefficient (s), dominance (h), and initial allele frequency (p_0), the parameter estimates converged to s = 0.163 (standard deviation [SD] = 0.052), h = 0.249 (SD = 0.142), and initial frequency p_0 = 0.025 (SD = 0.008) (fig. 3).

Discussion

These data show that there is marked LD around *kdr* mutations, loci exhibiting high penetrance, and, for *L1014F* at least, subject to strong recent positive selection. Despite similar median EHH levels, there were differences in the patterns of LD associated with the *L1014S* mutation in Kenya and Gabon. In Kenyan samples, the rate of dissipation of LD around the *L1014S* core was quite rapid suggesting that the mutation has not been subject to as recent or as strong a selective sweep as the same mutation in Gabon (or indeed as the *L1014F* mutation in Ghana). This is as predicted if the serine resistance allele was primarily selected by the use of DDT in the latter part of the 20th century rather than by the more recent use of pyrethroids in agriculture and insecticide control programs. In *Culex* mosquitoes, the equivalent *L1014S* mutation gives low levels of *kdr* to pyrethroids compared with the *L1014F* mutation but confers high levels of DDT resistance (Martinez-Torres et al. 1998; Ranson et al. 2000). Stump et al. (2004) investigated the change in allele frequency of the *L1014S* allele before and after the commencement of a large-scale ITN project in Asembo Bay, Western Kenya, the site of our collections (Stump et al. 2004). The frequency of the *L1014S* allele in the region approximately 10 years before bed net introduction was approximately 0.04 (95% CI 0.02–0.08). In 2002, 15 years after this initial survey and 5 years after the introduction of nets, the frequency of the *L1014S* allele had increased, nonsignificantly to only 0.075 (95% CI 0.05–0.12). This suggests that there is little selective advantage for this

Table 2. Comparison of Median EHH Levels between Alleles at the *kdr* Loci.

	Kenya *L1014S* (S form)	Kenya Wild Type (S form)	Gabon *L1014S* (S form)	Gabon Wild Type (S form)	Gabon *L1014F* (S form)	Ghana Wild Type (M form)
Kenya wild type (S form)	0.0001					
Gabon *L1014S* (S form)	**0.26 NS**	**0.0001**				
Gabon wild type (S form)	0.0001	**0.026NS**	0.0001			
Gabon *L1014F* (S form)	**0.0001**	**0.0001**	**1.00NS**	**0.0001**		
Ghana wild type (M form)	0.005	**1.00NS**	0.005	**0.86NS**	0.005	
Ghana *L1014F* (S form)	**0.0005**	**0.0001**	**0.0001**	**0.0001**	**0.04NS**	**0.0003**

NOTE.—Probabilities from sign tests are shown. The values followed by NS were not significant after sequential Bonferroni corrections. Values that are underlined indicate that the EHH values were significantly higher for the sample given in the column heading; values that are in bold indicate that the EHH values were significantly higher for the sample given in the row heading.

mutation in the present environment, although it should be noted that in a neighboring district in Uganda, a recent study reported that the *L1014S* mutation was at a frequency of 0.85 (95% CI 0.83–0.87) (Ramphul et al. 2009). An alternative explanation would be that in Uganda there is an epistatic interaction between *L1014S* and some, as yet unidentified locus, which may affect the selection, and indeed dominance coefficients, and thereby result in a higher *L1014S* frequency.

The high frequency and marked LD associated with *L1014S* in Gabon may be a result of the co-occurrence in genotypes, though not haplotypes (supplementary table 5, Supplementary Material online), with *L1014F*. A recent study from Cameroon showed that although *L1014F/L1014S* heterozygotes were significantly less resistant to permethrin than *L1014F* homozygotes, *L1014F/L1014S* heterozygotes were significantly more resistant to all insecticides tested than *L1014F/L1014*-wild type heterozygotes (Reimer et al. 2008). Repetitive mutation at the *1014* locus could, at least in part, be responsible for the patterns of LD around the *kdr* locus in the Gabonese data. Indeed, there is evidence for repeated mutations of *kdr* alleles across the species range of *A. gambiae* (Pinto et al. 2007). However, we argue that on the recent timescales on which *kdr* has arisen and spread it is more parsimonious to assume that recombination is the dominant influence on patterns of LD rather than high rates of repetitive mutations.

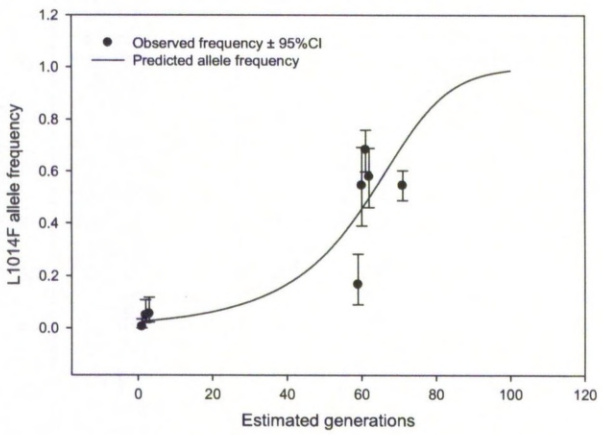

FIG. 3. Observed and predicted changes in *L1014F* allele frequency in the *Anopheles gambiae* M-form populations from southern Ghana. Observed data obtained from surveys conducted in 2002, 2006, and 2007. First collection point (Generation 1) was June 2002. Data from 2002, first three data points, are taken from Yawson et al. (2004); all other data are novel. One generation per month is assumed following Lehmann et al. (1998). The 95% CIs for each observed data point were calculated according to Newcombe (1998). Expected data generated from simultaneous maximum likelihood estimates of initial frequency and selection and dominance coefficients (see Materials and Methods).

Although *kdr* is the best-documented resistance mechanism in *A. gambiae*, there are many other resistance-associated loci. Microarray and recombinant protein expression work has shown that resistant mosquitoes over express a small number of enzymes that catalyze insecticide degradation (Ortelli et al. 2003; Müller et al. 2007; Chiu et al. 2008; Müller et al. 2008). LD-based screens could be a powerful way of identifying regions of the genome carrying the scars of recent selection that regulate such overexpression. However, whether association mapping approaches will effectively identify genes subject to much older and comparatively weaker selection is currently unclear. The bounded estimate of the selection coefficient reported here is at the upper limit of estimates generated to date and of a similar magnitude to estimates generated for resistance alleles in the mosquito *Culex pipiens* (Labbe et al. 2009). In human populations, mutations associated with resistance to malaria infection such as G6PD and sickle cell trait have coefficients of selection of 0.02–0.05 (Tishkoff and Williams 2002) and 0.05–0.18 (Li 1975), respectively. In the third actor in the malaria transmission cycle of *Plasmodium falciparum*, a selection coefficient of 0.1 has been obtained for the locus *dhfr* that confers resistance to the chemotherapeutic agent, pyrimethamine (Nair et al. 2003).

Together with strong and recent positive selection, the major determinant of LD around selected loci will be the rate of recombination. Indications of dramatic variation in the recombination rate across the *A. gambiae* genome have already been reported (Pombi et al. 2006; Black et al. 2008), and it is possible that, being close to the centromere of chromosome 2L, the sodium channel locus is in an area of reduced recombination. However, our Kenyan data are consistent with rates of recombination sufficient to reduce the region hitchhiked with a selectively advantageous locus in a relatively short period of time. Indeed, detection of the signatures of selection for loci with low selection coefficients will be more logistically challenging in *A. gambiae* than humans because of much lower background levels of LD (Weetman D, Wilding CS, Steen K, Donnelly MJ,

unpublished data). We attempted to amplify microsatellites from around the sodium channel to fully define the extent of the swept region as has been done for drug resistance loci in *P. falciparum* (Wootton et al. 2002; Nair et al. 2003). However, the sodium channel is situated in a region with an abundance of repetitive sequences and it was not possible to identify unique locus-specific microsatellite primer pairs.

Given the apparently high selection pressure on the *L1014F* mutation, it is curious that there are no studies, with adequate sample size, that have observed either of the *kdr* alleles at fixation (Santolamazza et al. 2008). One explanation would be that of overdominance; however, insecticide bioassays studies suggest that this is unlikely to be the case (Chandre et al. 2000; Reimer et al. 2008), and our estimate of the dominance coefficient shows the *kdr L1014F* allele to be partially recessive. Therefore, it is likely that there is some fitness cost to the *L1014F* allele and that this could be attributable to heterogeneity in exposure to pyrethroids in the environment or a consequence of an Hill–Robertson effect where selection at a *kdr* locus can interfere with the selection at nearby beneficial mutations (Hill and Robertson 1966).

The data presented herein show that it is possible to detect genomic signatures of strong positive selection in pest species with large effective population size and generally low levels of LD. We suggest that such approaches are likely to extremely powerful in many nonmodel taxa subject to similar selective events.

Supplementary Material

Supplementary tables 1–6 are available at *Molecular Biology and Evolution* online (http://www.mbe.oxfordjournals.org/).

Acknowledgments

We thank Nelson Cuamba, Eveline Klinkenberg, and Pie Muller for providing wild caught specimens and the associate editor for constructive comments. This work was supported by the National Institutes for Health (U01 AI58542 to Dr Edward Walker, Michigan State University), the Innovative Vector Control Consortium, and the Royal Society.

References

Aminetzach YT, Macpherson JM, Petrov DA. 2005. Pesticide resistance via transposition-mediated adaptive gene truncation in *Drosophila*. *Science* 309:764–767.

Ballinger-Crabtree ME, Black WC IV, Miller BR. 1992. Use of genetic polymorphisms detected by the random-amplified polymorphic DNA polymerase chain reaction (RAPD-PCR) for differentiation and identification of *Aedes aegypti* subspecies and populations. *Am J Trop Med Hyg.* 47:893–901.

Bass C, Nikou D, Donnelly MJ, Williamson MS, Ranson H, Ball A, Vontas J, Field LM. 2007. Detection of knockdown resistance (*kdr*) mutations in *Anopheles gambiae*: a comparison of two new high-throughput assays with existing methods. *Malar J.* 6e:111.

Begun DJ, Aquadro C. 1992. Levels of naturally occurring DNA polymorphism correlate with recombination rates in *D. melanogaster*. *Nature* 356:519–520.

Begun DJ, Holloway AK, Stevens K, et al. (13 co-authors). 2007. Population genomics: whole-genome analysis of polymorphism and divergence in *Drosophila simulans*. *PLoS Biol.* 5:e310.

Bersaglieri T, Sabeti PC, Patterson N, Vanderploeg T, Schaffner SF, Drake JA, Rhodes M, Reich DE, Hirschhorn JN. 2004. Genetic signatures of strong recent positive selection at the lactase gene. *Am J Hum Genet.* 74:1111–1120.

Black WC IV, Gorrochetegui-Escalante N, Randle NP, Donnelly MJ. 2008. The yin and yang of linkage disequilibrium: mapping of genes and nucleotides conferring insecticide resistance in insect disease vectors. *Transgenesis and the Management of Vector-Borne Disease.* 627:71–83.

Chandre F, Darriet F, Duchon S, Finot L, Manguin S, Carnevale P, Guillet P. 2000. Modifications of pyrethroid effects associated with *kdr* mutation in *Anopheles gambiae*. *Med Vet Entomol.* 14:81–88.

Chiu T, Wen Z, Rupasinghe S, Schuler M. 2008. Comparative molecular modeling of *Anopheles gambiae* CYP6Z1, a mosquito P450 capable of metabolizing DDT. *Proc Natl Acad Sci U S A.* 105:8855–8860.

Davies TGE, Field LM, Usherwood PNR, Williamson MS. 2007a. A comparative study of voltage-gated sodium channels in the Insecta: implications for pyrethroid resistance in Anopheline and other Neopteran species. *Insect Mol Biol.* 16:361–375.

Davies TGE, Field LM, Usherwood PNR, Williamson MS. 2007b. DDT, pyrethrins, pyrethroids and insect sodium channels. *IUBMB Life.* 59:151–162.

Diabate A, Brengues C, Baldet T, et al. (11 co-authors). 2004. The spread of the Leu-Phe *kdr* mutation through *Anopheles gambiae* complex in Burkina Faso: genetic introgression and *de novo* phenomena. *Trop Med Int Health.* 9:1267–1273.

Fanello C, Santolamazza F, della Torre A. 2002. Simultaneous identification of species and molecular forms of the *Anopheles gambiae* complex by PCR-RFLP. *Med Vet Entomol.* 16:461–464.

Hall TA. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser.* 41:95–98.

Hill WG, Robertson A. 1966. The effect of linkage on limits to artificial selection. *Genet Res.* 8:269–294.

Holm S. 1979. A simple sequentially rejective multiple test procedure. *Scand J Stat.* 6:65–70.

Kim Y, Stephan W. 2002. Detecting a local signature of genetic hitchhiking along a recombining chromosome. *Genetics* 160: 765–777.

Labbe P, Sidos N, Raymond M, Lenormand T. 2009. Resistance gene replacement in the mosquito *Culex pipiens*: fitness estimation from long-term cline series. *Genetics* 182:303–312.

Lehmann T, Hawley WA, Grebert H, Collins FH. 1998. The effective population size of *Anopheles gambiae* in Kenya: implications for population structure. *Mol Biol Evol.* 15:264–276.

Li W. 1975. The first arrival time and mean age of a deleterious mutant gene in a finite population. *Am J Hum Genet.* 27: 274–286.

Livak KJ. 1984. Organization and mapping of a sequence on the *Drosophila melanogaster* X and Y chromosomes that is transcribed during spermatogenesis. *Genetics* 107:611–634.

Lynd A, Ranson H, McCall PJ, Randle NP, Black WC, Walker ED, Donnelly MJ. 2005. A simplified high-throughput method for pyrethroid knock-down resistance (*kdr*) detection in *Anopheles gambiae*. *Malar J.* 4:16.

Martinez-Torres D, Chandre F, Williamson MS, Darriet F, Berge JB, Devonshire AL, Guillet P, Pasteur N, Pauron D. 1998. Molecular

characterization of pyrethroid knockdown resistance (*kdr*) in the major malaria vector *Anopheles gambiae s.s. Insect Mol Biol.* 7:179–184.

Maynard-Smith J. 1998. Evolutionary genetics. Oxford University Press, Oxford.

Müller P, Chouaibou M, Pignatelli P, Etang J, Walker ED, Donnelly MJ, Simard F, Ranson HW. 2008. Pyrethroid tolerance is associated with elevated expression of antioxidants and agricultural practice in *Anopheles arabiensis* sampled from an area of cotton fields in Northern Cameroon. *Mol Ecol.* 17: 1145–1155.

Müller P, Donnelly MJ, Ranson H. 2007. Transcription profiling of a recently colonised pyrethroid resistant *Anopheles gambiae*-strain from Ghana. *BMC Genomics.* 8:e36.

Müller P, Warr E, Stevenson BJ, et al. (12 co-authors). 2008. Field-caught permethrin-resistant *Anopheles gambiae* overexpress CYP6P3, a P450 that metabolises pyrethroids. *PLoS Genet.* 4:e1000286.

Nair S, Williams JT, Brockman A, et al. (12 co-authors). 2003. A selective sweep driven by pyrimethamine treatment in southeast asian malaria parasites. *Mol Biol Evol.* 20:1526–1536.

Newcombe RG. 1998. Two-sided confidence intervals for the single proportion: comparison of seven methods. *Stat Med.* 17: 857–872.

Ortelli F, Rossiter LC, 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.

Pinto J, Lynd A, Elissa N, Donnelly MJ, Costa C, Gentile G, Caccone A, Do Rosario VEI. 2006. Co-occurrence of East and West African *kdr* mutations suggests high levels of resistance to pyrethroid insecticides in *Anopheles gambiae* from Libreville, Gabon. *Med Vet Entomol.* 20:27–32.

Pinto J, Lynd A, Vicente JL, et al. (13 co-authors). 2007. Multiple origins of knockdown resistance mutations in the Afrotropical mosquito vector *Anopheles gambiae. PLoS ONE.* 2:e1243.

Pombi M, Stump AD, Della Torre A, Besansky NJ. 2006. Variation in recombination rate across the X chromosome of *Anopheles gambiae. Am J Trop Med Hyg.* 75:901–903.

Ramphul U, Boase T, Bass C, Okedi LM, Donnelly MJ, Müller P. Forthcoming. 2009. Insecticide resistance and its association with target-site mutations in natural populations of *Anopheles gambiae*from eastern Uganda. *Trans R Soc Trop Med Hyg.* 103: 1121–1126.

Ranson H, Jensen B, Vulule JM, Wang X, Hemingway J, Collins FHI. 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 Mol Biol.* 9: 491–497.

Reimer L, Fondjo E, Patchoke S, et al. (11 co-authors). 2008. Relationship between *kdr* mutation and resistance to pyrethroid and DDT insecticides in natural populations of *Anopheles gambiae. J Med Entomol.* 45:260–266.

Sabeti PC, Reich DE, Higgins JM, et al. (17 co-authors). 2002. Detecting recent positive selection in the human genome from haplotype structure. *Nature* 419:832–837.

Sabeti PC, Schaffner SF, Fry B, Lohmueller J, Varilly P, Shamovsky O, Palma A, Mikkelsen TS, Altshuler D, Lander ESW. 2006. Positive natural selection in the human lineage. *Science* 312:1614–1620.

Sabeti PC, Varilly P, Fry B, et al. (12 co-authors). 2007. Genome-wide detection and characterization of positive selection in human populations. *Nature* 449:913–918.

Santolamazza F, Calzetta M, Etang J, et al. (11 co-authors). 2008. Distribution of knock-down resistance mutations in *Anopheles gambiae*molecular forms in west and west-central Africa. *Malar J.* 7:e74.

Schlenke TA, Begun DJ. 2004. Strong selective sweep associated with a transposon insertion in *Drosophila simulans. Proc Natl Acad Sci U S A.* 101:1626–1631.

Scott JA, Brogdon WG, Collins FH. 1993. Identification of single specimens of the *Anopheles gambiae* complex by the polymerase chain reaction. *Am J Trop Med Hyg.* 49:520–529.

Stephens M, Scheet P. 2005. Accounting for decay of linkage disequilibrium in haplotype inference and missing-data imputation. *Am J Hum Genet.* 76:449–462.

Stephens M, Smith NJ, Donnelly P. 2001. A new statistical method for haplotype reconstruction from population data. *Am J Hum Genet.* 68:978–989.

Stump AD, Atieli FK, Vulule JM, Besansky NJ. 2004. Dynamics of the pyrethroid knockdown resistance allele in western Kenyan populations of *Anopheles gambiae* in response to insecticide-treated bed net trials. *Am J Trop Med Hyg.* 70:591–596.

Tishkoff SA, Williams SM. 2002. Genetic analysis of African populations: human evolution and complex disease. *Nat Rev Genet.* 3:611–621.

Weill M, Chandre F, Brengues C, Manguin S, Akogbeto M, Pasteur N, Guillet P, Raymond M. 2000. The *kdr* mutation occurs in the Mopti form of *Anopheles gambiae s.s.* through introgression. *Insect Mol Biol.* 9:451–455.

Wilding CS, Weetman D, Steen K, Donnelly MJ. 2009. High, clustered, nucleotide diversity in the genome of *Anopheles gambiae*revealed by SNP discovery through pooled-template sequencing: implications for high-throughput genotyping protocols. *BMC Genomics.* 10:e320.

Wootton JC, Feng XR, Ferdig MT, Cooper RA, Mu JB, Baruch DI, Magill AJ, Su XZI. 2002. Genetic diversity and chloroquine selective sweeps in *Plasmodium falciparum. Nature* 418: 320–323.

Yawson AE, McCall PJ, Wilson MD, Donnelly MJ. 2004. Species abundance and insecticide resistance of *Anopheles gambiae* in selected areas of Ghana and Burkina Faso. *Med Vet Entomol.* 18:372–377.

Yawson AE, Weetman D, Wilson MD, Donnelly MJ. 2007. Ecological zones rather than molecular forms predict genetic differentiation in the malaria vector *Anopheles gambiae s.s.* in Ghana. *Genetics* 175:751–761.