INSECTICIDE RESISTANCE AND BIONOMICS IN LABORATORY REARED AND FIELD CAUGHT *Anopheles funestus* GILES (DIPTERA: CULICIDAE)

Belinda Lea Spillings

A thesis submitted to the Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree of Doctor of Philosophy

Johannesburg, 2012.

DECLARATION

I, Belinda Lea Spillings declare that this thesis is my own work. It is being submitted for the degree of Doctor of Philosophy in the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination at this or any other University.

……

28th day of MAY, 2012

PUBLICATIONS AND PRESENTATIONS ARISING FROM THIS STUDY

PUBLICATIONS

Spillings BL, Coetzee M, Koekemoer LL and Brooke BD. 2008. The Effect of a Single Blood Meal on the Phenotypic Expression of Insecticide Resistance in the Major Malaria Vector *Anopheles funestus*. *Malaria Journal*. **7**: 226

Amenya DA, Naguran R, Lo TC, Ranson H, **Spillings BL**, Wood OR, Brooke BD, Coetzee M and Koekemoer LL. 2008. Overexpression of a Cytochrome P450 (*CYP6P9*) in a Major African Malaria Vector, *Anopheles funestus*, Resistant to Pyrethroids. *Insect Molecular Biology* **17**: 19–25

Spillings BL, Brooke BD, Koekemoer LL, Chiphwanya J, Coetzee M and Hunt RH. 2009. A New Species Concealed by *Anopheles funestus* Giles, a Major Malaria Vector in Africa. *American Journal of Tropical Medicine and Hygiene* **81**: 510 – 515

Christian RN, Matambo TS, **Spillings BL**, Brooke BD, Coetzee M and Koekemoer LL. 2011. Age Related Pyrethroid Resistance is not a Function of P450 Gene Expression in the Major African Malaria Vector, *Anopheles funestus* (Diptera: Culicidae). *Genetics and Molecular Research* **10**: 3220 – 3229

ORAL PRESENTATIONS

Spillings BL, Koekemoer LL, Brooke BD, Kikankie C, Coetzee M and Hunt RH. 2008. Malaria Vector Control: Dealing with Non-Vector Species. Faculty of Health Sciences Research Day and Faculty Postgrad Expo. Johannesburg, South Africa.

Spillings BL, Koekemoer LL and Brooke BD. 2008. The Effect of a Blood Meal on Insecticide Resistance in *Anopheles funestus*. XXIII International Congress of Entomology. Durban, South Africa.

Hunt, RH, Spillings BL, Koekemoer LL, Brooke BD, Kikankie C and Coetzee M. 2008. Malaria Vector Control: Dealing with Non-Vector Species. XXIII International Congress of Entomology. Durban, South Africa.

Spillings BL, Koekemoer LL, Brooke BD and Coetzee M. 2009. Molecular Investigations of the Effect of a Blood Meal on the Insecticide Resistance in *Anopheles funestus*. 5th MIM Pan-African Malaria Conference. Nairobi, Kenya.

Spillings BL, Brooke BD, Coetzee M and Koekemoer LL. 2011. The Effect of Blood-feeding on Gene Expression in Insecticide Resistant *Anopheles funestus*. XVII Congress of the Entomological Society of Southern Africa. Bloemfontein, South Africa.

POSTER PRESENTATIONS

Spillings BL, Koekemoer LL and Brooke BD. 2008. The Effect of a Single Blood Meal on Insecticide Resistance in *Anopheles funestus.* Faculty of Health Sciences Research Day and Faculty Postgrad Expo. Johannesburg, South Africa.

Spillings BL, Koekemoer LL, Brooke BD, Coetzee M and Hunt RH. 2008. Malaria Vector Control: Dealing with Non-Vector Species. NICD Academic Day. Johannesburg, South Africa.

Spillings BL, Koekemoer LL and Brooke BD. 2008. The Effect of a Single Blood Meal on Insecticide Resistance in *Anopheles funestus.* NICD Academic Day. Johannesburg, South Africa.

Spillings BL, Koekemoer LL, Brooke BD, Coetzee M and Hunt RH. 2009. A New Species of the *Anopheles funestus* Group from Malawi. 5th MIM Pan-African Malaria Conference. Nairobi, Kenya.

Spillings BL, Koekemoer LL, Brooke BD, Coetzee M and Hunt RH. 2009. Malaria Vector Control: Dealing with Non-Vector Species. Molecular and Cell Biology Department Symposium. Johannesburg, South Africa.

ABSTRACT

Malaria is transmitted by the mature, blood feeding portion of mosquito vector populations. Malaria vector control programs based on indoor residual spraying (IRS) of insecticides are designed to target resting adult *Anopheles* mosquitoes before or after they have blood fed.

When a female mosquito acquires a blood meal, she could also ingest harmful xenobiotics that are present in the blood. During the resting period after feeding, many processes are initiated in order to assist in the digestion and assimilation of the blood. Ultimately, this enables the mosquito to absorb those amino acids needed for the biosynthesis of yolk proteins, which are essential for subsequent egg maturation. Since the regulation of xenobiotic (including insecticides) detoxification enzyme systems is likely to be altered in response to the ingestion of blood, this study aimed to investigate the effect of a blood meal on insecticide tolerance in insecticide resistant and susceptible southern African strains of the major malaria vector *Anopheles funestus*.

Through the use of CDC bottle bioassays it was demonstrated that blood fed *An. funestus* carrying a pyrethroid resistant phenotype are even more tolerant of pyrethroid intoxication than their unfed counterparts. Using another major malaria vector, *An.* gambiae, microarray analysis revealed that a general increase in delta class glutathione-s-transferase (GST) expression occured in response to a blood meal. One gene, *GSTD3*, was over-expressed in both blood fed *An. gambiae* and *An. funestus*. Although this gene could not be validated with real time quantitative PCR, it serves as a viable target for future investigations.

Since the pyrethroid resistant phenotype of southern African *An. funestus* has been linked to the over-expression of the duplicate copy gene *CYP6P9*, the expression levels of both copies

of this gene were investigated. *CYP6P9* and its copy, *CYP6P13*, showed a small but significant increase in expression in response to a blood meal. The increased expression of these major effect genes in response to blood feeding may be responsible for the increase in insecticide tolerance seen in the bottle bioassays.

In an effort to repeat these experiments on wild caught *An. funestus*, field material was collected from Karonga in northern Malawi. Specimens were morphologically identified as members of the *An. funestus* group. However, attempts to molecularly identify them to species level failed. Through the use of *ITS2* and *D3* sequence analysis, cytogenetics and cross mating studies it was possible to conclude that these wild caught specimens were a new species. They have been provisionally named *An. funestus*-like.

ACKNOWLEDGEMENTS

I would like to extend my gratitude to my supervisors, Dr Basil Brooke, Prof Lizette Koekemoer and Prof Maureen Coetzee for their advice, encouragement and guidance throughout the development of this thesis. Special thanks go to Dr Brooke for initiating this study and for his exceptional ability to critique the results and outputs that have been generated herein. Prof Koekemoer is thanked for her patience and guidance through the molecular aspects of this research project. It is a pleasure to thank Professors Maureen Coetzee and Richard Hunt for their encouragement and enthusiasm, and for sharing their vast academic experience. I also thank Prof Hunt for teaching me the basics of field work, which I have come to enjoy immensely.

My gratitude is also extended to the following people and institutions that have facilitated different aspects of this work. Paladin Resources Ltd is thanked for facilitating the field collections and kindly providing accommodation in Karonga, northern Malawi; Mr Matthee for assisting with field trip logistics; Dr Wirtz for *P. falciparum* positive controls and monoclonal antibody *P. falciparum* 2A10. Mr Vezenegho is thanked for performing the *P. falciparum* sporozoite ELISA on the Malawian specimens; Mr Kikanie for his assistance with the field collections; Mr Lo for laboratory assistance; Dr Brooke and Prof Hunt are thanked for performing the cross-mating studies and cytogenetics. Prof Ranson of the Liverpool School of Tropical Medicine is thanked for kindly supplying the *An. gambiae* 'detox chip'.

Thank you to all the staff and students of the Vector Control Reference Unit (NHLS) and the Malaria Entomology Research Unit (University of the Witwatersrand) for help and guidance along the way.

Lastly, I wish to thank all the co-authors who have contributed to the publications that have arisen from this work. (Author contributions have been listed with each publication.)

This project was funded through the following grants: South African Medical Research Council grant to Dr Basil Brooke; the South African Malaria Initiative; National Health Laboratory Service Research Trust grant to Prof Lizette Koekemoer; National Research Foundation grant to Prof Lizette Koekemoer; and the South African Research Chair Initiative of the Department of Science and Technology and the National Research Foundation awarded to Prof Maureen Coetzee.

ETHICS APPROVAL

University
of the Witwatersrand,
Johannesburg

Human Research Ethics Committee (Medical)
(formerly Committee for Research on Human Subjects (Medical)

Secretariat: Research Office, Room 9H10003, 10th floor, Senate House • Telephone: +27 11 717 1234 • Fax: +27 11 339 8708
Private Bag 3, Wits 2050, South Africa

Ref: W-CJ-101214-1

15/12/2010

TO WHOM IT MAY CONCERN:

Waiver: This certifies that the following research does not require clearance from the Human Research Ethics Committee (Medical).

Investigator: **Ms Belinda Lea Spillings (Student No 0201444H).**

Project title: **Insecticide resistance and bionomics in laboratory reared and field caught Anopheles funestus Giles (Diptera: Culicidae).**

Reason: This is a laboratory study using mosquitos. No humans are involved.

UNIVERSITY OF THE WITWATERSRAND
2010 -12- 14
JOHANNESBURG

Professor Peter Cleaton-Jones
Chair: Human Research Ethics Committee (Medical)

copy: Anisa Keshav, Research Office, Senate House, Wits

TABLE OF CONTENTS

CHAPTER 3 – DETOXIFICATION GENE TRANSCRIPTION ANALYSIS OF THE EFFECT OF A BLOOD MEAL ON FEMALE MOSQUITOES

CHAPTER 4 – *ANOPHELES FUNESTUS-LIKE*: THE DISCOVERY OF A NEW MEMBER SPECIES OF THE *ANOPHELES FUNESTUS* GROUP

LIST OF FIGURES

LIST OF TABLES

NOMENCLATURE /ABBREVIATIONS

ACT - Artemisinin combined therapy

BLAST - Basic Local Alignment Search Tools

CDC – Centers for Disease Control and Prevention

DDT - Dichloro-diphenyl-trichloroethane

CO1 – Cytochrome oxidase, subunit 1

Cq – Quantification cycle

CYP6P9 – Cytochrome P450, gene family 6, sub family P, gene 9

CYP6P13 – Cytochrome P450, gene family 6, sub family P, gene 13

DNA – Deoxyribonucleic Acid

E8 – Elimination Eight

EST - Expressed sequence tag

FUN18S – 18s rDNA gene, specifically of *An. funestus*

GapDH – Glyceraldehyde-3-phosphate dehydrogenase

GST – Glutathione-S-Transferase

GSTD3 – Glutathion-S-Transferase Delta Class, 3

IRS – Indoor Residual Spraying

IVM – Integrated Vector Management

HKG – Housekeeping gene (also known as a reference gene)

LD50 - Lethal dose required to kill 50% of insecticide exposed individuals

LSDI - Lubombo Spatial Development Initiative

mRNA - Messenger RNA

MA Plot – A graph representing the distribution of the red and green intensity ratios (M) by the average intensity (A)

MIAME – Minimum Information for About a Microarray Experiment

MIQE - Minimum Information for Publication of Quantitative Real-Time PCR Experiments

ND5 – NADH dehydrogenase, subunit 5

PCR - Polymerase chain reaction

PMT - photomultiplier

qPCR – Quantitative Real-time PCR

QTL - Quantitative trait locus

RFLP – Restriction fragment length polymorphism

RG – Reference Gene (also known as a housekeeping gene – HKG)

RIN – RNA integrity number

RNA – Ribonucleic acid

RPL8 – Ribosomal protein, large subunit 8

RPL19 – Ribosomal protein, large subunit 19

rRNA – Ribosomal RNA

RPS7 – Ribosomal protein, small subunit 7

RPS26 – Ribosomal protein, small subunit 26

SMRS – Specific Mate Recognition System

WHO – World Health Organization

CHAPTER 1 - GENERAL INTRODUCTION

1.1 INTRODUCTION

Malaria is one of the leading vector borne diseases worldwide and is responsible for the suffering of many individuals and family units in Sub-Saharan Africa. The World Health Organization (WHO) estimates that worldwide, there were over 225 million cases of malaria during 2009, with African countries carrying the bulk (78%) of the disease burden (http://www.who.int/malaria/world_malaria_report_2010). Currently, malaria is the leading cause of death in children under the age of five years. The impact of malaria places an increased burden on health care systems and results in decreased levels of productivity within the home and work environments (Chima *et al.*, 2003). The effects of this disease place a great strain on the economies of third world countries and malaria has, in the past, been implicated as one of the contributing factors to reduced development in African countries.

Malaria is an ancient disease that has been alluded to in Egyptian papyri as early as 1500 B.C. Many great scientists and doctors, including Hippocrates and Laveran, have worked to elucidate the aetiology of the disease (Garnham, 1966). The link between malaria and mosquitoes was suggested by Dr. Beauperthuy, one of the early malaria researchers, in the mid 1800's. He suggested that the mosquito absorbed a "decay toxin" from its marsh-like environment and further proposed that this "toxin" was passed from the mosquito to humans thereby causing malaria (Boyce, 1910). In 1897 in India, Major Ronald Ross suggested that mosquitoes belonging to the sub-family Anophelinae were the vectors of avian malaria and that these mosquitoes were the intermediate host for the malaria parasite (Boyce, 1910). At the same time, Italian researchers Bignami and Grassi showed that

Plasmodium falciparum was transmitted to humans through the infective bite of an anopheles mosquito. Grassi and his colleagues were the first to show the full life cycle of *P. falciparum* and *P. vivax* through the mosquito (Garnham, 1966). This elucidation of the *Plasmodium*-anopheline link has allowed for the development of malaria control efforts to limit the effect of the disease.

Many current malaria control programmes continue to focus on the timely diagnosis and cost effective treatment of malaria. The parasitic protist responsible for this disease belongs to the Genus *Plasmodium*, with the dominant African species being *P. falciparum*. Many African countries still rely on chloroquine based treatment, however artemisinin combined therapy (ACT) is becoming more widely available especially in areas that have seen the emergence of chloroquine resistant *Plasmodium*. In the face of emerging drug resistance there is an ever increasing need for an integrated approach that embraces the importance of adequate vector control (Oaks *et al.*, 1991; Miller and Greenwood, 2002).

Vector surveillance and vector control are essential for assisting in the reduction of the incidence of malaria. Vector control aims to interrupt the malaria cycle prior to parasite transmission. Many early vector control programmes relied heavily on the use of indoor residual spraying (IRS) to target the indoor resting mosquitoes and larviciding to target the aquatic stages. In an African setting larviciding is not always feasible due to the extensive distribution of larval habitats which require frequent larvicide application due to the low persistence of the insecticides (Coosemans and Carnevale, 1995). Modern vector control programmes generally follow an integrated vector management system (IVM) as promoted by WHO-AFRO. An IVM can be loosely defined as ‘a process of evidence-based decision-making procedures aimed at planning, implementing, monitoring and evaluating targeted,

cost-effective and sustainable combinations of vector control measures' (Manga *et al.*, 2004). Since IVMs aim to target vector mosquitoes at multiple points of their life cycle they generally combine the use of IRS with a combination of any of the following tools: bednets, biological control, environmental management, improved personal protection and community education. Unfortunately, there is no universal approach to malaria vector control. The epidemiology of the disease varies between as well as within countries and is dependant on the vector mosquito and *Plasmodium* species present.

1.1.1 Malaria Vector Control in South Africa

South Africa has a good track record in terms of the successful implementation of malaria vector control. Malaria used to be prevalent from the northeastern border with Mozambique and Zimbabwe to as far south as Port St Johns on the east coast and as far inland as Pretoria (le Sueur *et al*., 1993). It is primarily the implementation of successful provincial IRS programmes that has seen the distribution of malaria shrink towards the northern and eastern border regions of South Africa.

South Africa experienced a devastating malaria outbreak in KwaZulu-Natal during the 1931 malaria season. This prompted the first trial applications of pyrethrum insecticides in an attempt to control adult indoor resting mosquitoes (Coetzee and Hunt, 1998). In 1932, stimulated by the success of this trial, the widespread use of pyrethrum in house-spraying programmes began.

DDT was used extensively during World War II to protect soldiers from disease carrying insects such as lice, mites and fleas. Following the successes of DDT during the war and the fact that South Africa had started to produce their own DDT stocks, there was a shift to

replace pyrethrums with DDT in the IRS campaigns in South Africa. This shift towards DDT resulted in the malaria mosquitoes, *An. funestus* and *An. gambiae,* largely being eliminated from this region (De Meillon, 1986; Sharp and le Sueur, 1996; Coetzee and Hunt, 1998).

During 1996, due to pressure to discontinue the use of DDT, South Africa began using the pyrethroid deltamethrin for IRS (Govere *et al*., 2002; Coetzee and Fontenille, 2004). Despite continued IRS, malaria incidence rose dramatically from ±10 000 cases/annum in 1995 to ± 60 000 case/annum by the end of 2000 (Figure 1.1). Hargreaves *et al.* (2000) showed that *An. funestus* had returned to the KwaZulu-Natal region and that this population was resistant to pyrethroids. A return to the use of DDT in 2000 resulted in a steady drop in malaria cases, with a reported decrease of 91% of malaria incidence in the KwaZulu-Natal region by 2002 (Department of Health, unpublished data; Maharaj *et al.*, 2005).

In 2000, the Lubombo Spatial Development Initiative (LSDI), a joint programme between the governments of Swaziland, Mozambique and South Africa, was launched. The main goal of the LSDI was to develop the communities within the border regions between the member countries in order to stimulate the region to become a globally competitive economic zone (Sharp *et al.*, 2007). Malaria control was considered a prerequisite to achieve this goal. Intensive IRS and antimalarial campaigns were initiated, treatment policies were reviewed and parasite prevalence surveys routinely carried out. During the initial seven year period of the LSDI, malaria cases within the targeted areas dropped significantly and the LSDI has contributed to a significant reduction in imported malaria cases within South Africa (Sharp *et al*., 2007).

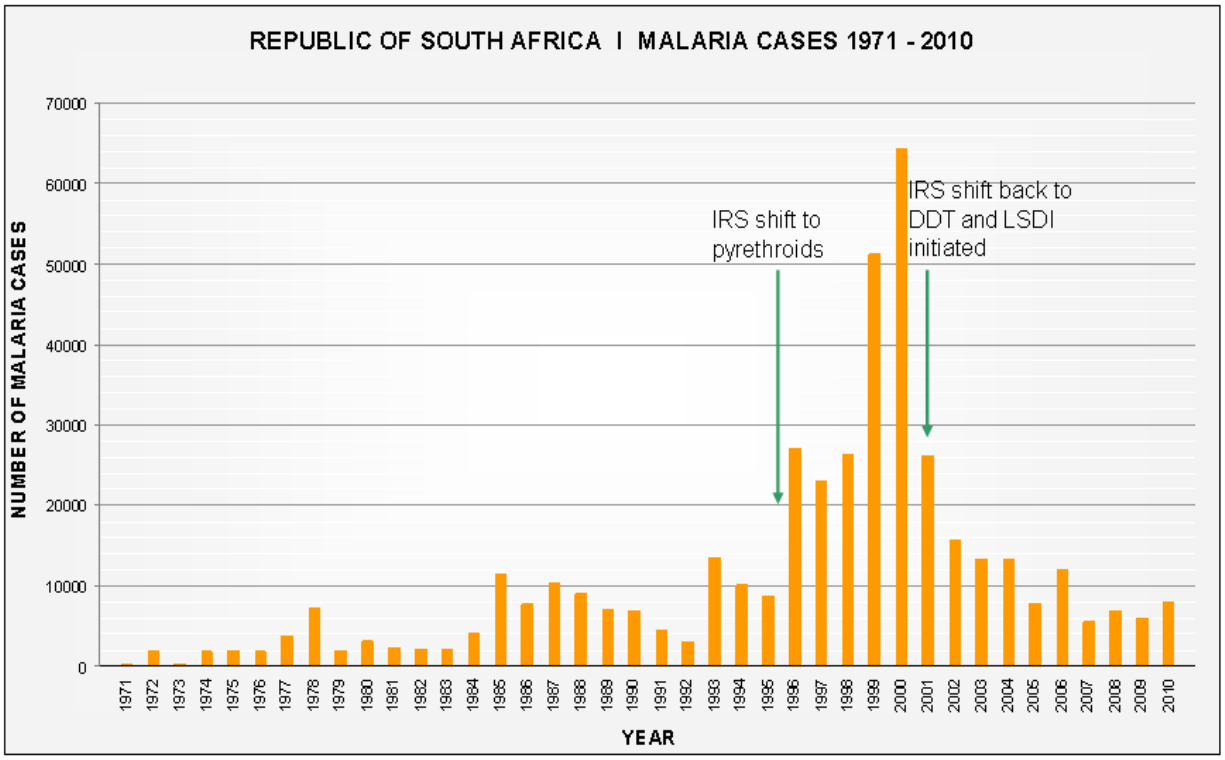


Figure 1.1. The number of malaria cases in South Africa per year, from 1971 to 2010 (unpublished data, Department of Health, South Africa. Available at www.fidssa.co.za/B_FIDSSA2011Presen.asp under the presentation titled: Progress towards Malaria Elimination in South Africa).

On the back of a highly successful LSDI programme, the concept of the Elimination Eight (E8) was born (www.malariaeliminationgroup.org). In March 2009, the Windhoek Resolution was drafted and this was presented at the SADC Ministers of Health Meeting, where the E8 was formally adopted. This is a regional and cross-border collaborative effort aimed at eliminating malaria in the four southernmost African countries, namely Botswana, Namibia, South Africa and Swaziland. These are considered to be the most likely to reach elimination by 2015. In order for elimination to be successful, support from the secondary bordering countries (Angola, Mozambique, Zambia and Zimbabwe) is essential. These secondary countries will need to scale up vector control activities in their southern most regions in order to create malaria free zones north of the elimination countries' borders (www.malariaeliminationgroup.org).

Since South Africa is now striving towards malaria elimination, an understanding of basic anopheline biology and ecology, as well as the processes involved in the development of insecticide resistance in vector populations, is becoming ever more important. Although DDT is currently being used in combination with carbamates and pyrethroids as part of South Africa's IRS campaign, multiple insecticide resistances have already started to appear in South African anophelines (Hargreaves *et al*., 2003; Mouatcho *et al.* 2009). The discovery of insecticide resistance in affected *An. funestus* populations has prompted concern of the possibility of development of additional resistance phenotypes and has highlighted the need to improve our understanding of resistance mechanisms and the implementation of novel methods of control.

1.1.2 Southern African Malaria Vectors

The three major malaria vectors in southern Africa belong to the *Anopheles gambiae* complex and the *Anopheles funestus* group (Gillies and De Meillon, 1968). Species from both groups often occur in sympatry and hence accurate identification of each species present is required for effective vector control.

The *An. gambiae* complex is comprised of seven morphologically similar species which can only be differentiated at the molecular or chromosomal level. Within this group *An. gambiae* Giles and *An. arabiensis* Patton are the major malaria vectors, whereas the remaining members (*An. merus* Dönitz, *An. melas* Theobald and *An. bwambae* White) are incidental or localized vectors (Gillies and De Meillon, 1968; Gillies and Coetzee, 1987). *Anopheles quadriannulatus* Theobald sp A and B (Hunt *et al*., 1998) have never been implicated as malaria vectors. *Anopheles gambiae s.l.* larvae are opportunistic in terms of

larval habitat and have been found in a wide range of temporary water bodies, including water filled pots, tyre tracks and hoof/foot prints.

The southern African malaria vector *An. funestus* Giles belongs to the *An. funestus* group. This group comprises nine members which are difficult to distinguish at the adult stage. However, identification can to some extent be more easily made at the egg and larval stages (Gillies and De Meillon, 1968; Gillies and Coetzee, 1987). *Anopheles funestus*, *An. rivulorum* Leeson, *An. leesoni* Evans, *An. parensis* Gillies and *An. vaneedeni* Gillies and Coetzee are commonly found in southern Africa. Of these, *An. funestus* is the only major malaria vector. *Anopheles funestus* is extremely anthropophilic and preferentially feeds on humans, generally during the second half of the night (Gillies and De Meillon, 1968). Blood fed female *An. funestus* have been shown to rest indoors until the second day of the gonotrophic cycle, only exiting the shelter when they are half gravid (Gillies and De Meillon, 1968). The preferred larval habitats of *An. funestus* are clean, permanent, vegetated water bodies such as marshes, ponds and swamps. The presence of this species generally coincides with seasonal fluctuations in rainfall, appearing in the middle of rainy seasons and becoming more abundant towards the start of drier periods (Gillies and De Meillon, 1968). It is suggested that changes in the level of the water table as well as climatic temperatures, which are less pronounced during the warm rainy seasons, exert an effect on the abundance of *An. funestus* (Gillies and De Meillon, 1968). Effective control of *An. funestus* can be achieved by the implementation of IRS because this species is highly anthropophilic.

Field methods used to distinguish the members of the *An. gambiae* complex and *An. funestus* group mostly rely on the use of the morphological keys (Gillies and Coetzee,

1987). *Anopheles funestus* group mosquitoes can be distinguished from *An. gambiae* complex members by the absence of pale spots on the legs combined with the absence of a pale spot in the third main dark area of vein one of the wings. Mosquitoes belonging to the *An. gambiae* complex and *An. funestus* group can only be identified to species level using molecular techniques, including restriction fragment length polymorphism (RFLP) (Garros *et al*., 2004), multiplex PCR (Scott *et al*., 1993; Koekemoer *et al*., 2002) and real-time hydrolysis probe analysis (Bass *et al*, 2008).

The accurate identification of field caught material is imperative during malaria vector control efforts since this will directly affect the costs involved as well as considerations of which insecticide to use. The ability to distinguish between morphologically similar vector and non-vector species, especially in areas of sympatry, is essential to the success of any vector control programme. Furthermore, the ability to identify which species carry insecticide resistance phenotypes allows for vector control programmes to be effective and cost efficient.

1.1.3 Mechanisms of Insecticide Resistance

The development of insecticide resistance has been demonstrated across most insect orders, with an ever increasing prevalence of multiple resistances (Hemingway and Ranson, 2000). The WHO defines insecticide resistance as the "development of an ability in a strain of an organism to tolerate doses of a toxicant which would prove lethal to the majority of individuals in a normal (susceptible) population of the species" (Zlotkin, 1999). Insecticide resistance predominantly occurs through target site insensitivity or enzymatic sequestration and detoxification (Hemingway *et al*., 2004).

Target site insensitivity generally occurs due to one or more point mutations in an insecticide target site. Altered target sites which translate into insecticide resistance include: the sodium channel gene which is associated with knock down resistance to pyrethroids and DDT, otherwise known as *Kdr* resistance (Martinez-Torres *et al*., 1997; Ranson *et al*., 2000); the acetylcholinesterase gene (Weill *et al*., 2002; N'Guessan *et al*., 2003) which confers resistance to organophosphates and carbamates and the GABA receptors which confer resistance to the cyclodienes (Hemingway and Ranson, 2000). Knock-down resistance has been identified in a vast number of insect species, including *An. gambiae* (Martinez-Torres *et al*., 1998), *Culex quinquefasciatus* (Xu *et al*., 2006) and *Musca domestica* (Williamson *et al*., 1996).

Enzymatic sequestration and detoxification of insecticides occurs when one or more detoxifying enzymes exhibit an increased level of activity. Detoxifying enzymes generally belong to one of three broad classes, namely the monooxygenases, glutathione S-transferases (GSTs) and non-specific esterases. The monooxygenase class encompasses the cytochrome P450 groups which have been implicated in pyrethroid resistance in *C. quinquefasciatus* (Xu *et al*., 2005), *C. pipiens pipiens* (McAbee *et al*., 2003), *An. gambiae* (Nikou *et al*., 2003; Djouaka *et al*., 2008), *An. arabiensis* (Müller *et al*., 2008; Munhenga *et al*., 2011) and *An. funestus* (Brooke *et al*., 2001; Wondji *et al*., 2007; Amenya *et al*., 2008).

The development of pyrethroid resistance in southern African *An. funestus* has been linked to increased levels of P450 monooxygenase activity (Brooke *et al*., 2001; Amenya *et al*., 2008; Cuamba *et al*., 2010). The P450 monooxygenases are a superfamily of enzymes that are involved in the detoxification of xenobiotics and participate in the breakdown of

endogenous metabolic compounds (Scott, 1999). P450 monooxygenases generally metabolize insecticides by incorporating an oxygen atom into the insecticide molecule, which leads to inactivation and detoxification.

The availability of the *An. gambiae* whole genome sequence has made it possible to identify, through Basic Local Alignment Search Tools (BLAST), consensus regions or sequences that encode P450-like genes. Using the conserved P450 heme binding region, Ranson *et al*. (2002) were able to identify 111 putative P450 genes in the *An. gambiae* genome. A number of P450 studies have suggested that increased levels of insecticide tolerance may be linked to increased levels of CYP4, CYP6, CYP9 and/or CYP12 expression (Scott, 1999). Since the P450 family is large it has been difficult to link a resistance phenotype to any one particular CYP enzyme. However, pyrethroid resistance in *An. gambiae* has been linked to the overexpression of CYP6Z1 (Nikou *et al*., 2003), CYP6Z2 and CYP6M2 (Müller *et al*., 2007) genes, which all occur within a cluster of P450 genes on chromosome 3R in *An. gambiae* adults (Nikou *et al*., 2003). Two major and one minor quantitative trait locus (QTL) have been identified in *An. gambiae* (Ranson *et al*., 2004). More recent research into these resistance "hotspots" in *An. funestus* has revealed that a QTL, which accounts for approximately 60% of resistance to pyrethroids in a southern African population of *An. funestus*, exists on chromosome 2R (Wondji *et al*., 2007). This collection of genes is primarily made up of a number of CYP6 P450 genes, thereby confirming the importance of CYP6s in insecticide resistance and is supported by evidence showing that CYP6P9, localized within this QTL, is overexpressed in pyrethroid resistant *An. funestus* (Wondji *et al*., 2007; Amenya *et al*., 2008).

1.1.4 Factors Influencing Insecticide Resistance

Life stage studies in *An. gambiae* and *An. funestus* have shown that gene expression linked to the detoxification of insecticides changes through the different stages of the mosquito life cycle (Strode *et al*., 2006; Christian *et al*., 2011). These changes in gene expression could have an effect on the relative levels of insecticide resistance at each life stage. Hunt *et al*. (2005) have shown that age plays a role in reduced survival of adult *An. funestus* exposed to pyrethroids.

One of the major biological process that may influence gene expression and hence insecticide resistance is blood feeding. During and after a female mosquito has taken a blood meal a suite of digestive processes are initiated in order to utilize the blood nutrients. Sanders *et al*. (2003) analysed gene expression within the midgut of *Aedes aegypti* following a murine blood meal. The study showed that the expressions of approximately 330 genes are altered after a blood meal and that the genes spanned a broad spectrum of processes including: nutrient uptake and metabolism; peritrophic matrix formation and stress responses. Of particular interest was the increased expression of two P450s whose putative function is detoxification (Sanders *et al*., 2003). The upregulation of cytochrome P450s in response to an avian blood meal has also been identified in *C. pipiens* (Baldridge and Feyereisen, 1986). The increase in P450 activity could be associated with the detoxification of xenobiotics present in the blood meal, however no literature on the direct effect of blood feeding on insecticide resistance could be found.

1.1.5 Laboratory Tools for Investigating Insecticide Resistance

Insecticide resistance can be determined in a population by carrying out WHO susceptibility assays (WHO, 1998). These bioassays score the percentage mortality after a sample group has been exposed to a particular insecticide, at a particular dosage, for a set time period. The data gained from these experiments indicate the presence or absence of resistance but do not provide information on the mechanism/s involved. The combination of bioassays with the use of synergists (compounds that abrogate the effect of insecticide resistance mechanisms) assists in determining which metabolic system could be responsible for the resistance phenotype, but is only one step in the process. Target site insensitivity may be indicated when cross-resistance is observed in bioassays. Biochemical analyses can be used to quantify metabolic enzyme systems in association with resistance as well as to determine the presence of target site insensitivity in altered acetycholinesterase systems. These results can also be confirmed by molecular tools such as PCR which can identify currently described mutations (Martinez-Torres *et al*., 1998; Ranson *et al*., 2000; Mutero *et al*., 1994; ffrench-Constant *et al*., 1993).

The majority of the resistance mechanisms that have been assessed to date are metabolic in nature and hence may involve a number of genes. Research into these systems has in the past been slow and the speed of resistance gene identification has been limited. Current technologies such as microarrays allow for studies into multiple gene expression simultaneously. The development of the *An. gambiae* detox chip (David *et al*., 2005) has allowed researchers to target their investigations directly at the suite of detoxification genes that are most likely to play a role in the development of insecticide resistance. Vontas *et al*. (2005) used this chip to show that a number of genes exhibit altered expression patterns after adult mosquitoes are exposed to insecticides. This altered gene

expression included the increased expression of a cytochrome P450. Further, the *An. gambiae* detox chip has been used to study detoxification gene expression in other species such as *An. arabiensis* (Müller *et al*., 2008) and *An. funestus* (Christian *et al*., 2011b).

1.1.6 Species Concepts

The debate about how new species arise, otherwise referred to as the process of speciation, is often confused with the need to define what a species is (Hey, 2001). Since many theories of speciation rely on a particular idea of what a species is, and the idea of a species depends on the concept of speciation, each becomes somewhat dependent upon the other. Perhaps a good place to start is to try and define the word 'species'. To the lay person, a species can be defined as "a group of living organisms consisting of similar individuals capable of exchanging genes or interbreeding" (Oxford Dictionary: www.oxforddictionaries.com). This is a biological definition dependent on the ability of organisms to reproduce. A second definition relates more to taxonomy: "The species is the principal natural taxonomic unit, ranking below a genus and denoted by a Latin binomial" (Oxford Dictionary: www.oxforddictionaries.com).

1.1.6.1 Taxonomic Concepts

Taxonomic systems are old concepts that aimed to rank organisms in a simple and practical manner. These systems assumed that species remain unchanged, without any alteration through time. Aristotle was one of the first scientists to attempt to classify species according to a '*scala naturae*' in his book *Historia animalium*. He used gross morphology as well as a scale which was dependant on the organism's degree of perfection or potentiality (Bowler, 1992). The organisms that fell into the higher degrees were those

most similar to humans, i.e. those that gave birth to warm, live young. The organisms that fell into the lower degrees of the scale reproduce by laying cold, dry eggs.

Aristotle's system was simplistic and he limited it to the animal kingdom. Linnaeus (1758) refined the taxonomical concept by separating all organisms into three broad Kingdoms: *Regnum Animale*, *Regnum Vegetabile* and *Regnum Lapideum* (animal, plant and mineral Kingdoms). He then divided each Kingdom into Classes composed of several Orders. The end point of his classification was a binomial given to each organism or species. Although Linnaeus based many of his classifications on structural similarities between organisms, he still did not leave any room for the idea of evolution. He states in his publication *Systema Naturae* that 'species are as many as were created in the beginning by the Infinite" (Linnaeus, 1758).

Linnaeus' binomial system is still used today and he is considered the father of modern taxonomy. The first part of the binomial is the genus name to which the organism belongs, followed by the 'identifier' term. This identifier was generally chosen to describe the species itself according to a particular character.

1.1.6.2 Biological Concepts

Today, it is commonly accepted that classing organisms according to their gross morphology and 'degrees of perfection' is entirely subjective and that the process of speciation involves genetic, biological and ecological components. Biological concepts of speciation can be broadly separated into two groups. Relational (or isolation) concepts are concerned with the physical location and separation of populations in order for speciation to occur. Genetical concepts (including the recognition concept) mostly assume that

ecological reproductive isolation has already occurred and are more focused on the genetic mechanisms that have led to species divergence.

Dobzhansky and Mayr proposed theories of speciation that have been referred to as isolation concepts (Paterson, 1978). Their concepts are concerned with how the isolation of populations leads to divergent speciation. Hence, they are relational concepts. Dobzansky (1937) suggested that isolating 'mechanisms' are responsible for delimiting the species gene pool. He defines isolating mechanisms as 'all the mechanisms hindering or preventing the interbreeding of racial complexes or species' (Dobzhansky, 1937). These isolating mechanisms were generally related to the geographical positioning of the populations as a method for limiting their interaction. During this period of separation (allopatry), and given enough time, each population would randomly acquire unique mutations. If sufficient functional variation accumulated in each genome to prevent successful interbreeding of populations and sympatry was re-established, speciation can be considered to have taken place. This is an allopatric mode of speciation (Figure 1.2).

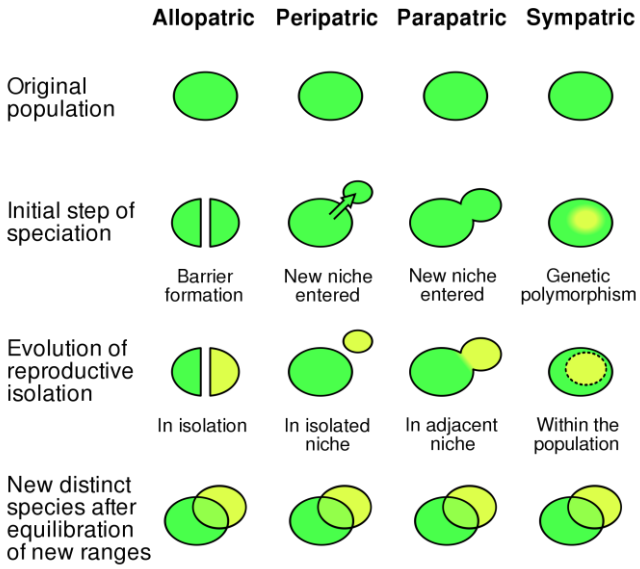


Figure 1.2. Modes of speciation. (Ilmari Karonen, http://en.wikipedia.or/wiki/File:Speciation_ modes .svg)

Mayr (1969) defined a species as "groups of interbreeding natural populations that are reproductively isolated from other such groups". He strongly supported the concepts of allopatric and peripatric speciation, but remained unconvinced of the plausibility of sympatric speciation. Further, he suggested that reproductive isolating mechanisms could be due to pre-mating or post-mating isolation (Mayr, 1963). Pre-mating isolation includes seasonal and habitat isolation where potential mates cannot meet due to different preferences or adaptations to habitat and climate; ethological isolation where potential mates can meet but not mate; and mechanical isolation where potential mates meet and attempt to copulate but fail. Post-mating isolation includes gamete inviability and or mortality whereby copulation takes place however the ovum cannot be fertilized; zygote mortality where the ovum is fertilized but the resulting zygote dies; hybrid inviability where the resulting zygote survives but has reduced viability; and hybrid sterility where the zygote is fully viable but cannot reproduce to form the next generation due to partial or complete sterility.

The idea of hybrid inviability and sterility favoured the re-birth and development of the concept of speciation through reinforcement. The Wallace Effect, first suggested by Alfred Wallace in his book Darwinism (1889), lays the foundations for theories of speciation through reinforcement. These theories suggest that if two populations of different species were to come into contact with one another and mate successfully, their hybrid offspring would be less fit and therefore less competitive relative to pure-bred progeny. Hence, natural selection would select against hybrid survival (negative heterosis). The reproductive gap between the two species groups would thus be reinforced.

Paterson (1978) argued that the reinforcement theory could not be valid. He reasoned that if two populations came into contact and random mating occurred to produce less fit

hybrids, the smaller of the two populations would eventually become extinct. Further, he highlighted inconsistencies in experiments designed to prove reinforcement theories. In these experiments researchers manipulated the population numbers such that each population was maintained at equal levels at each successive generation (Koopman, 1950; Knight *et al*., 1956; Paterniani, 1969; Crossley, 1974). This manipulation prevented the 'weaker' of the populations from dying out, thereby giving false evidence in favour of the reinforcement theory.

Since hybrid sterility and/or decreased hybrid fitness is the foundation of the reinforcement theory, the issues of defining species in terms of hybrid sterility needed to be investigated (Paterson, 1988). Historically, early Christian scientists viewed hybrid sterility as a Divine measure designed to ensure the purity of breeds (Lyell, 1832). The example of a sterile mule, being the hybrid progeny of a horse and a donkey, has often been used as evidence of a Divine mechanism to keep species populations pure and genetic lines uncontaminated. This concept of hybrid sterility breaks down when one considers polyploidy in angiosperms. The triploid offspring generated from a diploid and tetraploid parental pair of the same angiosperm are more or less sterile. In terms of hybrid sterility this would suggest that the parental types are different species. However, we know that this is not the case and the sterility is actually due to meiotic differences in the gametes. For this reason, hybrid sterility in the reinforcement concept is not a suitable model for speciation.

In an effort to propose a clear and logical process of speciation, Paterson (1985) put forward his recognition concept. He defines a species as the 'most inclusive population of individual biparental organisms which share a common fertilization system'. In a natural, unforced environment members of a species will evolve a common specific mate

recognition system (SMRS). The system includes all the co-adapted signals and receptors required for two potential mates to recognize one another, copulate and successfully produce offspring. The SMRS includes olfactory, auditory, visual, tactile and chemo-sensory stimuli which could form a series of signals between two potential mates. The SMRS in a biparental system would have to evolve such that both the male and female parts of the system develop in parallel. Although the recognition concept applies best to mobile organisms, it can also be applied to sessile organisms such as plants, mussels and oysters (Paterson, 1985). In plants, the SMRS may evolve to include pollen vectors which are specific organisms that transport pollen from one plant to another, as well as surface receptors on the stigma that recognizes pollen from like species.

The co-adaption of the male and female SMRS is required for individuals of a species to retain compatibility. It is the adaptation of the components of the SMRS, for example to new habitats or environmental strain, which causes alterations in the SMRS between individuals. With this in mind, the end point of speciation in terms of the recognition concept is the point at which the new species' SMRS is no longer recognized by the original population and mating cannot occur successfully (Paterson, 1980). Speciation is thus viewed as an incidental effect of adaptation, not a direct consequence.

There are at least 24 different concepts that have been proposed to solve the species problem (Hey, 2001), some more philopsophical and others more practical. The recognition concept aims to explain speciation through the most basic biological need or process, i.e. reproduction. It is not a relational concept that tries to explain speciation in one population in relation to another population. It highlights individual interactions within a population, with the common goal to produce progeny. For this reason the recognition

concept is favoured here over other species concepts because it deals with the issues of speciation in the most logical, simple and biologically relevant manner.

1.2 RESEARCH AIMS AND HYPOTHESIS

When a female mosquito takes a blood meal, altered gene expression occurs in order to accommodate and utilize the nutrients. It is hypothesized that the suite of genes responsible for the detoxification of xenobiotics in the blood meal may influence the subsequent level of susceptibility to insecticide following insecticide exposure. Since P450 monooxygenase mediated pyrethroid resistance has been identified in the major South African malaria vector, *An. funestus*, it is imperative to understand the external factors that influence P450 regulation.

Hence, in female *An. funestus* mosquitoes carrying a pyrethroid resistance phenotype, the process of acquiring and digesting a blood meal may affect their insecticide resistance status. Potentially, an altered insecticide resistance profile may impact on the current control strategies employed by malaria vector control programmes. The aim of this study was to determine whether there is a correlation between blood feeding and altered susceptibility to pyrethroids in *An. funestus*, and if so, to quantify the phenotypic and detoxification gene expression effects.

1.3 OBJECTIVES

The aims of this study can broadly be separated into four main objectives:

1. Determine, using bioassays, the effect of a blood meal on pyrethroid susceptibility in laboratory reared pyrethroid resistant and susceptible *An. funestus* females (Chapter 2).
2. Analyse changes in detoxification gene expression using microarrays in laboratory reared pyrethroid resistant and susceptible *An. funestus* and *An. gambiae* females in response to blood feeding (Chapter 3).
3. Verify any changes observed on selected genes using quantitative real-time PCR (qPCR), based on the results obtained from microarray analysis (Chapter 3).
4. Characterise wild populations of *An. funestus* (Chapter 4).

CHAPTER 2 –THE EFFECT OF A SINGLE BLOOD MEAL ON THE PHENOTYPIC EXPRESSION OF INSECTICIDE RESISTANCE IN *ANOPHELES FUNESTUS*

2.1 INTRODUCTION

The primary malaria vector during the 1999/2000 malaria outbreak in South Africa was *An. funestus* and control of this mosquito currently relies on effective indoor residual spraying (IRS). Although South Africa's IRS campaign has adopted a mosaic spray approach using DDT (dichloro-diphenyl-trichloroethane), carbamates and pyrethroids (Maharaj *et al*., 2005), the occurrence of pyrethroid resistance in the *An. funestus* population (Hargreaves *et al*., 2000; Brooke *et al*., 2001) is a major concern for vector control. In addition, the development of insecticide resistance in *An. arabiensis* populations in South Africa (Hargreaves *et al*., 2003; Mouatcho *et al*., 2009) induces additional cause for concern, and an understanding of the mode, expression and inheritance of insecticide resistance mechanisms has become increasingly important.

Insecticide resistance in insect populations is predominantly based on improved enzymatic sequestration and detoxification as well as by the alteration of insecticide target sites leading to insensitivity to insecticide (Hemingway *et al*., 2004). Improved enzymatic detoxification has been linked to three broad classes of enzymes, namely monooxygenases, glutathione-S-transferases (GSTs) and non-specific esterases. Pyrethroid resistance in *Culex quinquefasciatus* (Xu *et al*., 2005), *Culex pipiens pipiens* (McAbee *et al.*, 2003), *An. gambiae* (Nikou *et al*., 2003) and *An. funestus* (Brooke *et al*., 2001; Wondji *et al*., 2007; Amenya *et al*., 2008; Wondji *et al*., 2009) has been linked to the increased activity of

cytochrome P450s, members of the monooxygenase class of detoxification enzymes. Further, the P450 monooxygenases have been implicated in the detoxification of xenobiotics (including drugs, pesticides and plant toxins) as well as endogenous metabolic products in insects (Scott, 1999).

Since many major biological processes affect gene expression it is possible that insecticide detoxification gene expression may be stimulated by processes other than insecticide exposure. The upregulation of cytochrome P450s in response to a blood meal has been demonstrated in *C. pipiens* (Baldridge and Feyereisen, 1989) and *Aedes aegypti* (Sanders *et al*., 2003). It is hypothesized that the detoxification of xenobiotics and toxic blood components in the *An. funestus* midgut may inadvertently result in an increased ability to tolerate insecticide intoxication.

2.2 RATIONALE

Vector control relies on the use of insecticide chemicals to significantly reduce the number of malaria vectors by targeting that portion of the female population that takes blood meals and subsequently rests indoors. It has been suggested that the intake of a blood meal may assist female mosquitoes to tolerate higher doses of insecticide through vigour tolerance. It is hypothesized that in addition to vigour tolerance, during the process of blood digestion, detoxification mechanisms required for the neutralizing of harmful components in the blood meal may also confer an increased ability to tolerate insecticide intoxication through increased enzyme regulation.

2.3 AIMS AND OBJECTIVES

The aim of this study was to determine whether a change in pyrethroid insecticide tolerance occurs when *An. funestus* mosquitoes have taken a blood meal. This was carried out through the following objectives:

- To determine the percentage mortalities, for pyrethroid resistant and susceptible colonies of *An. funestus*, following exposure to varying dosages of permethrin for both unfed and blood fed cohorts
- To determine the lethal dose required to kill 50% of permethrin exposed individuals (LD50) in each of the unfed study groups
- To determine if a change in insecticide tolerance occurs for the resistant and susceptible colonies of *An. funestus* when exposed to the LD50 dose of permethrin, post blood feeding

2.4 MATERIALS AND METHODS

2.4.1 Mosquito colonies. *Anopheles funestus* laboratory colonies have been established and are maintained at the Vector Control Reference Unit of the National Institute for Communicable Diseases, NHLS (Johannesburg, South Africa). All colonies are maintained under standard insectary conditions (Hunt *et al*., 2005). The two *An. funestus* colonies used were: FUMOZ-RH, which originates from southern Mozambique and has been intensively selected for pyrethroid resistance (Hunt *et al*., 2005). FUMOZ-R (as in previously published work) and FUMOZ-RH refer to the same colony. The second colony, FANG, originates from Angola and is susceptible to pyrethroids.

2.4.2 Insecticide dose-response experiments. The process of blood meal digestion may activate detoxification systems required to detoxify xenobiotics present in the blood.

Hence, it was decided that insecticide susceptibility should be investigated at different stages during the blood digestion process. Susceptibility to permethrin was investigated during the early stage of blood digestion at four hours post blood feeding based on the assumption that those genes involved in the digestion process would have been transcribed by that time. Susceptibility to permethrin was also investigated during the later stage of the digestion process at 18 hours post blood feeding to allow for the possibility that different genes may have been upregulated by that time.

Three to four day old female cohorts from each colony were collected. Each cohort was divided into two groups, one was fed on a 10% sucrose solution and the other was blood-fed. Blood meals were offered in a darkened room with an ambient temperature of 25°C. Only females that took blood were subsequently tested for susceptibility to permethrin. Following blood-feeding, a 10% sucrose solution was made available to all the females for either 4 hours or 18 hours prior to permethrin exposure.

Dose-mortality responses comparing blood-fed versus unfed samples from the permethrin resistant and susceptible *An. funestus* colonies were assayed according to the CDC bottle bioassay method (Brogdon and McAllister, 1998). Glass bottles (250 ml volume) were coated with the following range of permethrin concentrations (µg of permethrin/250 ml bottle): 0.1 µg, 1 µg, 10 µg, 25 µg, 50 µg, 100 µg, 250 µg, 500 µg and 1000 µg. Appropriate amounts of permethrin (Sigma) were dissolved in 1 ml acetone as a carrier. Each bottle was used a maximum of three times before being discarded.

Approximately 20 to 25 females were used per bottle and the insecticide exposures lasted one hour. Following exposure to permethrin, all the females were transferred to

polystyrene cups with access to a 10% sucrose solution. Percentage mortality was recorded 24 hours post-exposure for each permethrin concentration. For each *An. funestus* colony, eight to twelve cohorts were used. Each cohort was used as a single replicate such that one cohort could be used for the full dose range of exposures, in parallel for both the blood fed and unfed groups. The mean percentage mortality was calculated at each insecticide dose and the dose response graphs created. The data for each replicate of each cohort was log transformed to allow for the calculations of the 50% lethal dose (LD50) value, using regression analysis. The mean LD50 and standard deviation could then be calculated for blood-fed and unfed, resistant and susceptible mosquitoes. An example of the regression analysis can be found in Appendix A.

2.4.3 Dose specific responses following a blood meal. The WHO defines the discriminating dosage of insecticide to be used in resistance assays as twice the amount of insecticide required to kill 100% of an insecticide susceptible sample of the same species (WHO/CDS/MAL/98.12). The susceptible FANG colony showed 100% mortality at approximately 50 µg/ 250 ml bottle. It was thus decided that investigations at 100 µg/250 ml bottle would be appropriate for dose specific assays against FUMOZ-RH. Insecticide dosages of 2 µg and 5 µg/250 ml bottle were chosen for dose specific assays against the susceptible FANG colony based on results from the dose-response experiments where the range induced approximately 50% mortality.

Three to four day old female cohorts from each colony were removed and divided into two groups: one for blood-feeding and one to be fed on a 10% sucrose solution. Blood meals were offered four hours prior to the one hour permethrin exposures. Twenty to twenty five females were exposed per bottle through 9 to 11 replicates. Final mortality was recorded

24 hours post exposure and comparisons between blood-fed and unfed groups for each *An. funestus* colony were based on 2 sample *t*-tests and one-way ANOVA.

2.5 RESULTS

2.5.1 Lethal dose response graphs. Dose response graphs were generated for the insecticide susceptible FANG (Figure 2.1) and permethrin resistant FUMOZ-RH (Figure 2.2) colonies. No significant difference in susceptibility to permethrin between the unfed and blood-fed groups for both FUMOZ-RH and FANG ($p > 0.05$) was evident across the full dosage range, regardless of the lapse of time between blood-feeding and permethrin exposure. However, the FUMOZ-RH colony showed consistently higher levels of permethrin tolerance in the blood-fed group as compared to the unfed.

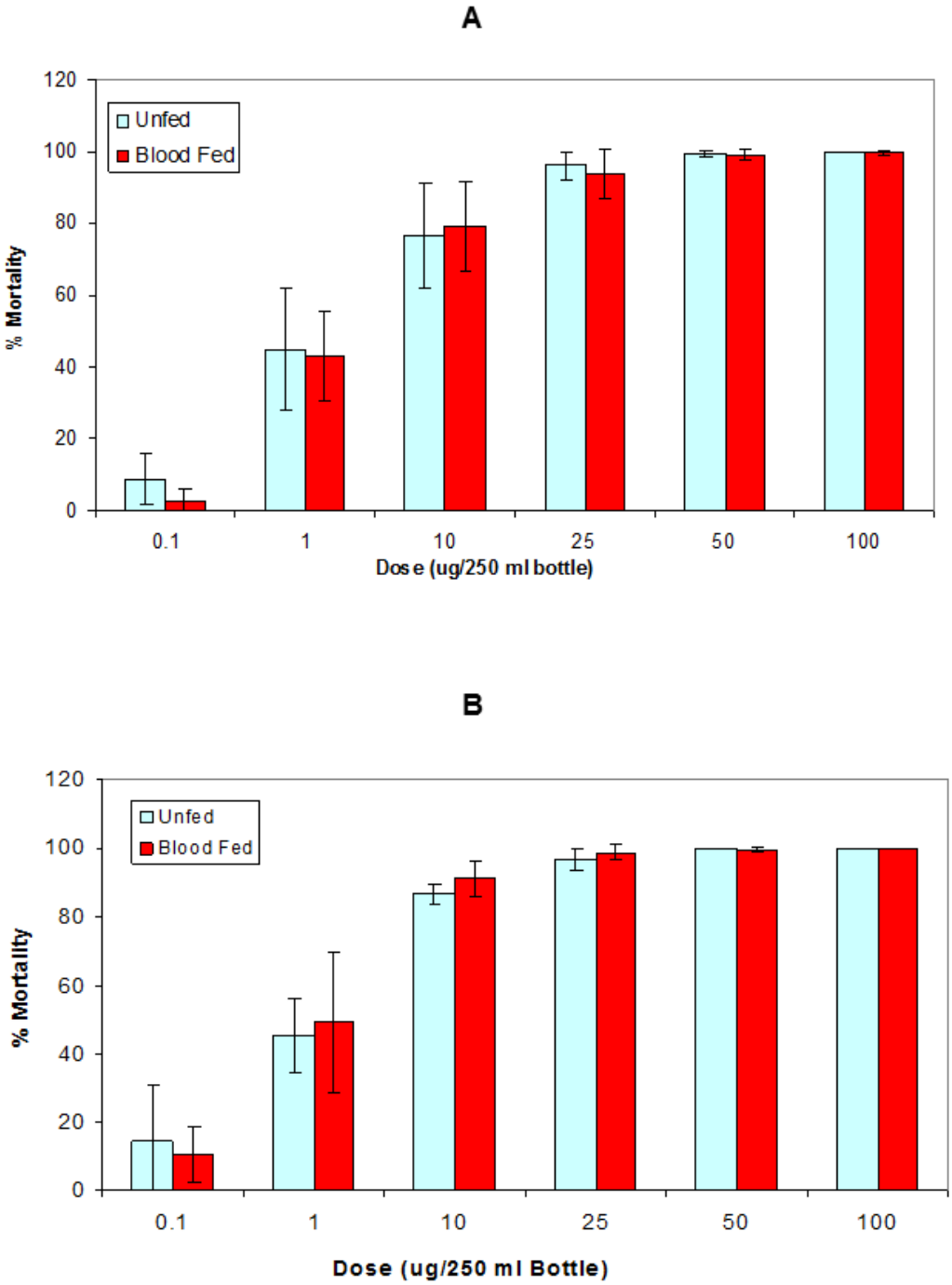


Figure 2.1. Percentage mortalities 24 hours post exposure for the pyrethroid susceptible *An. funestus* colony (FANG), in response to permethrin exposures, with either (A) blood-feeding 4 hours prior to permethrin exposure or (B) blood-feeding 18 hours prior to permethrin exposure.

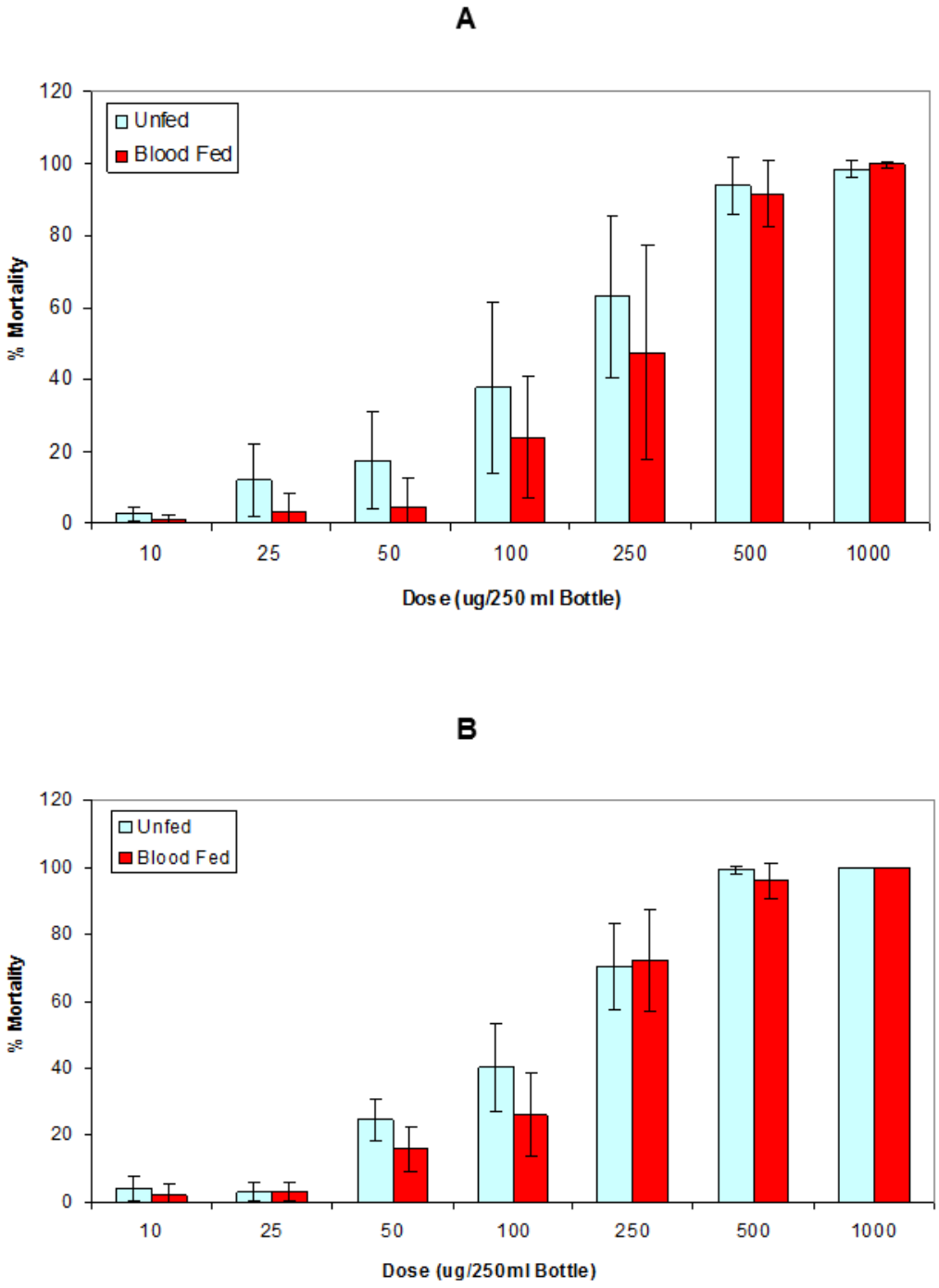


Figure 2.2. Percentage mortalities 24 hours post exposure for the pyrethroid resistant *An. funestus* colony (FUMOZ-RH), in response to permethrin exposures with either (A) blood-feeding 4 hours prior to permethrin exposure or (B) blood-feeding 18 hours prior to permethrin exposure.

Figure 2.3 shows the dose of permethrin required to produce 50% mortality in each of the colonies, for each of the treatment times (permethrin exposure at either 4 hours or 18 hours post blood-feeding). FUMOZ-RH showed significantly higher levels of insecticide tolerance as compared to the susceptible FANG colony ($p<0.05$). The permethrin dose required to kill 50% of the resistant FUMOZ-RH samples was approximately 70 to 80 times greater than that for the susceptible FANG colony.

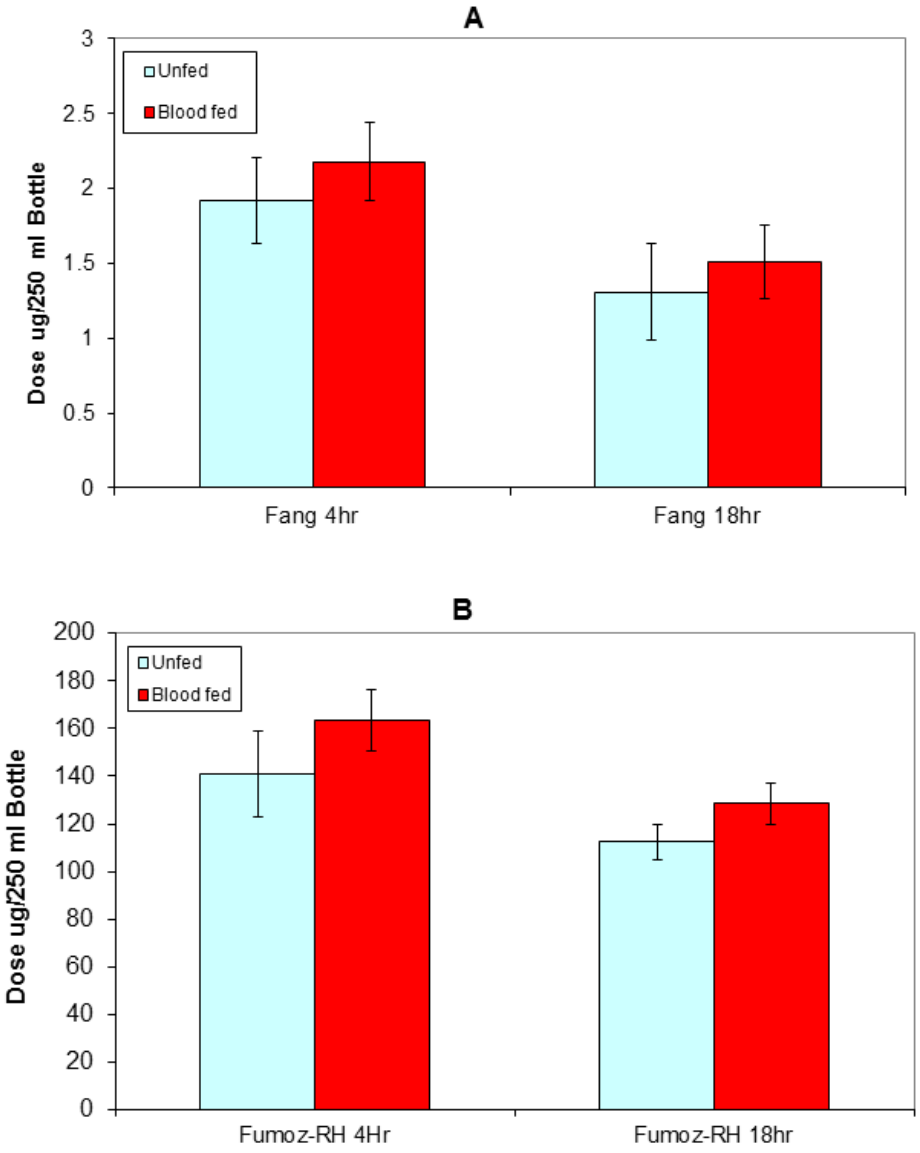


Figure 2.3. Comparison of dosages required to produce 50% mortality, 24 hours post permethrin exposure, (A) in the susceptible *An. funestus* colony (FANG) and (B) the resistant *An. funestus* colony (FUMOZ-RH). Permethrin exposures were carried out at either 4 hours or 18 hours post blood-feeding.

2.5.2 Dose specific responses following a blood meal. The susceptible FANG colony showed no significant difference in response to permethrin exposure between the unfed and blood-fed groups, for both of the insecticide dosages tested ($p > 0.05$, Figure 2.4). The difference in response to permethrin exposure between blood-fed and unfed cohorts from the FUMOZ-RH colony was highly significant ($p < 0.001$) with the blood-fed cohorts showing a mean percentage mortality approximately five times lower than that of the unfed cohorts (Table 2.1).

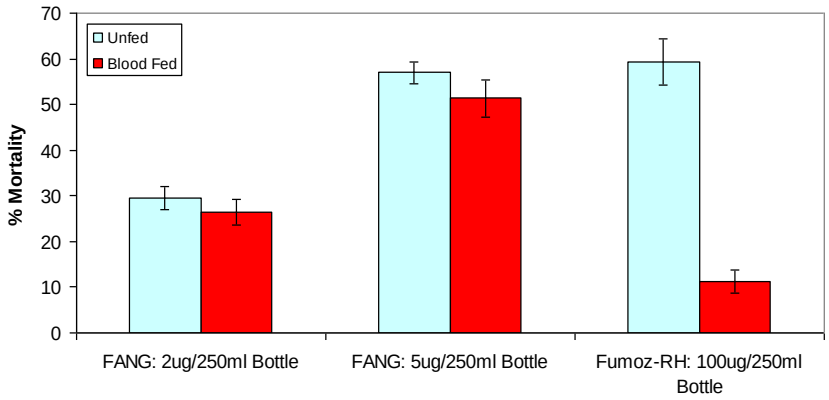


Figure 2.4. Comparison of percentage mortalities 24 hours post permethrin exposure at chosen discriminating dosages for the susceptible (FANG) and resistant (FUMOZ-RH) *An. funestus* colonies. Permethrin exposures were carried out 4 hours post blood- feeding at 2 and 5 µg/250 ml bottle for FANG and 100 µg/250 ml bottle for FUMOZ-RH.

Table 2.1. Mean percentage mortalities at the discriminating dosages for the susceptible *An. funestus* colony (FANG) and the resistant *An. funestus* colony (FUMOZ-RH).

	Dose/250 ml Bottle	Mean % Mortality	SE	n	p
FANG Unfed	2µg	29.45%	2.57%	213	> 0.05
FANG Blood-fed	2µg	26.31%	2.80%	216	
FANG Unfed	5µg	57.00%	2.42%	261	> 0.05
FANG Blood-fed	5µg	51.33%	4.14%	246	
FUMOZ-RH Unfed	100µg	59.21%	5.01%	244	< 0.001
FUMOZ-RH Blood-fed	100µg	11.37%	2.54%	245	

All exposures on blood-fed individuals were carried out 4 hours post blood-feeding. "SE" = standard error; "n" = sample size; "p" = significance of difference between the unfed and blood-fed groups following 2 sample t-tests.

2.6 DISCUSSION

The development of insecticide resistance in southern African *An. funestus,* and its dramatic effect on malaria transmission in South Africa, has highlighted the need to investigate this phenotype and its controlling factors. Pyrethroid resistance in southern African *An. funestus* has been linked to elevated levels of monooxygenase cytochrome P450 activity as the primary mode of resistance (Brooke *et al*., 2001; Wondji *et al.*, 2007; Amenya *et al*., 2008; Wondji *et al*., 2009). It has subsequently been demonstrated that the resistance phenotype is inherited as a single, autosomal, incompletely dominant genetic factor (Okoye *et al*., 2008) and that there is no compromise in reproductive and physiological fitness associated with resistance (Okoye *et al*., 2007), leading to the prediction that pyrethroid resistance can be expected to spread readily within and between *An. funestus* populations in affected areas. If insecticide application is to remain effective, then this scenario must ultimately consider the response to insecticide exposure of older, blood-feeding females, which form that proportion of the population actively transmitting malaria.

The application of an adapted CDC bottle bioassay method (Brogdon and McAllister, 1998) allowed for the quantification and comparison of the levels of insecticide tolerance in both insecticide resistant and susceptible *An. funestus* colonies, in response to the effect of blood-feeding. The results presented in this chapter indicate that the permethrin resistant colony (FUMOZ-RH), which has been intensively selected for pyrethroid resistance, has a 70- to 80-fold increase in insecticide tolerance as compared to the insecticide susceptible colony (FANG). Although the insecticide dose response graphs did not highlight any significant differences in insecticide tolerance between any of the blood-fed and unfed cohorts, the blood-fed resistant FUMOZ-RH colony consistently required

higher dosages than its unfed counterpart in order to produce the same level of mortality. The lack of statistically significant differences between blood-fed and unfed cohorts may be an artifact of wide variation in response to insecticide exposure between batches of mosquitoes over successive generations.

The direct comparison of percentage mortality following exposure to discriminating dosages of permethrin showed that a blood meal did not significantly alter the degree of insecticide tolerance in the fully insecticide susceptible colony of *An. funestus*. This result suggests that vigour tolerance through increased body mass (and subsequent increased dilution of internalized insecticide) does not offer a significant measure of insecticide resistance. However, similar comparisons between blood-fed and unfed, insecticide resistant females from the FUMOZ-RH colony showed a significant increase in insecticide tolerance in association with a single blood meal. This result suggests that the presence of a blood meal combined with an already effective insecticide detoxification mechanism significantly enhances the expression of the resistance phenotype.

Given that IRS campaigns aim to target the biting portion of a vector population that rests indoors and that insecticide resistance phenotypes within *An. funestus* populations are becoming more prevalent, the data presented here warrant further consideration. The results presented here suggest that the presence of a blood meal and/or the process of its digestion activate a series of insecticide detoxification pathways which "prime" the mosquito for contact with insecticide, in all likelihood through the increased expression of P450 genes hypothetically associated with blood meal digestion and insecticide detoxification.

2.7 CONCLUSION

The fully insecticide susceptible *An. funestus* colony did not show any significant alteration in susceptibility to insecticide following a blood meal suggesting that vigour tolerance through increased body mass does not play a significant role in tolerance to insecticide intoxication. The decrease in insecticide susceptibility in the pyrethroid resistant colony of *An. funestus* following a blood meal suggests that insecticide detoxification mechanisms involved in insecticide resistance may further be stimulated by the presence of a blood meal prior to insecticide exposure, thereby leading to enhanced expression of the resistance phenotype. This finding may be significant in terms of the criteria that are used to evaluate resistance phenotypes determined by WHO bioassay (WHO/CDS/MAL/98.12) in field populations, because blood-fed female mosquitoes may show enhanced expression of the resistance phenotype, possibly allowing for earlier detection of insecticide resistance.

2.8 Spillings BL, Coetzee M, Koekemoer LL and Brooke BD. 2008. The effect of a single blood meal on the phenotypic expression of insecticide resistance in the major malaria vector *Anopheles funestus*. *Malaria Journal* **7**: 226

Contribution to publication:

Belinda Spillings carried out the laboratory work and data analysis, wrote the first and subsequent drafts of the manuscript.

2.9 Amenya DA, Naguran R, Lo T-C M, Ranson H, Spillings BL, Wood OR, Brooke BD, Coetzee M and Koekemoer LL. 2008. Overexpression of a cytochrome P450 (CYP6P9) in a major African malaria vector, *Anopheles funestus*, resistant to pyrethroids. *Insect Molecular Biology* **17**: 19-25

Contribution to publication:

Belinda Spillings carried out the adult insecticide dose response experiments and the data analysis. She also contributed towards comments on the manuscript.

CHAPTER 3 – DETOXIFICATION GENE TRANSCRIPTION ANALYSIS OF THE EFFECT OF A BLOOD MEAL ON FEMALE MOSQUITOES

3.1 INTRODUCTION

In the past, gene transcription studies relied on a "one-gene, one-experiment" setup (Muyal *et al.*, 2008) and generating a global overview of gene transcription in response to a treatment or event was time consuming and costly. The advent of microarray technology brought about a shift from the tedious traditional molecular methods to the high-throughput screening of selected subsets of genes and even whole genomes simultaneously.

The concept of DNA microarray technology was developed at Stanford University Medical Centre, USA, with the first account of the use of a cDNA microarray for expression profiling published by Schena *et al.* (1995). A microarray can be described as a "two-dimensional arrangement of specific biological probes (e.g., DNA or protein molecules) deposited in an addressable fashion on a glass slide or other substance (e.g. polymer coated glass, plastics, nitrocellulose)" (Barbulovic-Nad *et al*., 2006).

During gene transcription studies, microarrays can be used to analyse the differential expression of groups of genes within two comparable biological sources at a particular point in time. In any living cell, at any given point in time, genes are induced in response to any number of factors, including internal circadian rhythms and external factors or stimuli. The first stage of information transfer described in the central dogma of molecular biology is transcription, which in many eukaryotes results in pre-mRNAs that are often further processed to form mature mRNAs (Klug and Cummings, 1997). Each messenger

RNA (mRNA) that results from a transcriptional event is the product of a specific gene and has the potential to be translated into functional protein. Isolation and identification of the mRNAs present at a particular point in an experiment or treatment gives a snap-shot view of gene expression within the sample and in essence provides information on the functional state of each gene in relation to the treatment concerned (Schena, 2002).

Isolated mRNA targets can be labeled with a Cyanine-5 (Target sample) or Cyanine-3 (Reference sample) fluorescent dye. The two mRNA samples are then combined prior to hybridisation onto a single microarray. Hybridisation allows complimentary targets and probes to bind, producing double stranded molecules anchored onto the array. Post-hybridisation washing serves to remove unbound targets so that only the bound targets are detected by the laser scanner. Two colour microarrays enable researchers to perform direct comparisons between two samples due to the competitive binding of each sample (Schulze and Downward, 2001), thus simultaneous detection of up and down regulation of gene targets within a single array is achievable.

The Anopheles gambiae Detoxification chip

The first published cDNA microarray experiment used 45 probes specific to 14 complete sequences and 31 expressed sequence tags (ESTs) of *Arabidopsis thaliana* (Schena *et al*., 1995). More recently developed arrays have expanded the number to be in excess of four million probes within a single array (e.g. Roche NimbleGen, www.nimblegen.com/products/expression/eukarya/index.html). These types of arrays are used to generate an unbiased view of gene expression across the entire genome of the organism. Knowledge of the full genome sequence is required to produce these tiling arrays.

Tiling arrays targeted at anopheline species are currently unavailable. However, Affymetrix® has developed their GeneChip® Plasmodium/Anopheles Genome Array, which incorporates more than 4,700 *P. falciparum* and 16,000 *An. gambiae* transcripts onto one array (www.affymetrix.com). This single colour system is suitable for investigating host-parasite based relationships thereby giving a global picture of gene expression responses resulting from *Plasmodium* infection. Two smaller, highly specialized arrays have been developed to investigate immunity and oxidative stress (Dimopoulos *et al*., 2002) and metabolic based insecticide resistance (David *et al*., 2005) in *An. gambiae.* The latter of these mini-microarrays is called the detox chip and it contains probes that represent approximately 230 putative detoxification genes, each printed four times within the array. It has been used to profile insecticide resistance phenotypes for *An. gambiae* (Müller P *et al*, 2007; Djouaka *et al*., 2008) as well as the closely related *An. arabiensis* (Müller *et al.*, 2008) and the unrelated *An. funestus* (Christian *et al*., 2011b).

Standardisation of microarray experiments and downstream quantification

Many different aspects of a microarray experiment can affect the quality, reproducibility and interpretation of the results. Basic guidelines have been established to assist researchers in standardising their experiments to an international level of acceptability. These MIAME (Minimum Information About a Microarray Experiment) guidelines include: descriptions of RNA extraction and amplification protocols; reporting of efficiencies and screening techniques to ensure suitable quality of biological samples; scanning parameters; annotations; data processing etc. (Brazma *et al*., 2001). In order to publish microarray findings, many journals now require that microarray experiments follow the MIAME guidelines and that data (raw and analysed) be submitted to a

microarray repository. These repositories are open-access and retrospective studies are able to utilize the data therein.

Since the detox chip is being utilized as an investigative tool in this study, all genes that are found to be upregulated will need to be validated using quantitative real time PCR (qPCR). qPCR allows one to analyse the transcript abundance after each cycle of amplification thereby resulting in real time data acquisition. Relative expression values can be determined for a target gene or gene of interest by comparing the change in transcript abundance of the gene in one biological sample versus another (possibly treated) biological sample. In order to ensure that the observed changes in target gene expression are due to a biological effect and not differences in quantity of starting cDNA, gene expression levels of a reference gene (RG) or housekeeping gene (HKG) must be monitored in parallel to all the target genes. The RG is a gene that should exhibit stable expression levels between all samples, including the untreated and treated cohorts, in order to validate direct comparisons of target gene expression between samples (Derveaux *et al*., 2010). As with microarray experiments, a set of guidelines has been published to assist researchers in ensuring that their qPCR experiments conform to a minimum set of requirements (Bustin *et al*., 2009). Issues such as RNA extraction methodology; RNA quantification; cDNA synthesis; qPCR chemistries; RG suitability as well as data analyses are covered in the MIQE (Minimum Information for Publication of Quantitative Real-Time PCR Experiments) guidelines (Bustin *et al*., 2009; Bustin, 2010). Real-time PCR reviews covering these topics are frequently published (see Baker, 2011).

3.2 RATIONALE

In Chapter 2 it was shown that blood feeding in *An. funestus*, in samples where pyrethroid resistance already occurred, led to increased levels of pyrethroid resistance. This effect was not apparent in *An. funestus* females that were fully susceptible to insecticides. These findings provided the impetus to investigate the effect of a blood meal on detoxification gene transcription. Previous gene transcription studies that have investigated the effect of blood feeding have utilised an *An. gambiae* strain (Pink Eye Standard) that is susceptible to insecticides (Davidson, 1956; Holt *et al*., 2002 in supplementary information). Since a response to the blood meal was only seen in *An. funestus* mosquitoes carrying a resistance phenotype, the transcription experiment warranted repeating with an insecticide resistant *An. gambiae* strain as well as *An. funestus.*

3.3 AIMS AND OBJECTIVES

The aim of this study was to investigate the use of the *An. gambiae* detox chip to detect potential gene targets that are upregulated in response to a blood meal in *An. funestus* and *An. gambiae*. Gene expression was investigated three hours post blood meal in order to generate data that was comparable to currently published research, such as Marinotti *et al*. (2006). Although the detox chip was designed for profiling expression within *An. gambiae*, it has been successfully applied to other closely related species. All potential gene targets were then analysed for relative expression using quantitative real time PCR. Specific objectives were:

- To investigate the effect of a blood meal on gene expression in *An. funestus* FUMOZ-RH and *An. gambiae* GAH using the *An. gambiae* detox chip
- To identify genes that have increased levels of expression in both the *An. funestus* and *An. gambiae* microarrays
- To validate, through the use of real-time qPCR, the levels of gene expression, for the above genes, in response to the presence of a blood meal

3.4 MATERIALS AND METHODS

3.4.1 Mosquito colonies and sample preparation. The *An. funestus* adults used in this study originated from the FUMOZ-RH laboratory colony. This colony originates from southern Mozambique and carries high levels of P450-mediated pyrethroid resistance (Brooke *et al.*, 2001). The *An. gambiae* S form adults were drawn from the GAH laboratory colony which originates from the Ahafo region in Ghana. The GAH colony exhibits extensive insecticide resistance (bendiocarb, DDT, dieldrin, permethrin, deltamethrin). Most importantly, the pyrethroid resistance in GAH is mediated by P450 and esterase metabolism in conjunction with *kdr* (Kaiser *et al.*, 2010).

Both of these colonies are maintained under standard insectary conditions (25 ± 2°C, 80% RH) with a 12 hour day/night cycle which includes a 30 minute dusk/dawn transition. Both male and female adults were placed in 5L cages on the day of emergence, with access to 10% sucrose. At three days of age, the males were removed and the cohort of females split into two groups. Both groups of females were then placed in artificial dusk cycles for 30 minutes, after which they were placed in complete darkness for a minimum of 30 minutes. One of the cages was offered a human blood meal. After the blood meal, both groups of females were allowed to rest for three hours.

Three separate biological replicates were used for this experiment. It was ensured that the adults used for each replicate were from different cohorts collected on different days.

3.4.2 RNA preparation and cDNA synthesis. After the three hour rest period, blood fed females and their unfed counterparts were sacrificed for RNA extractions. Approximately 10-15 females per group were placed in 1.5 ml microcentrifuge tubes, on ice, for immediate RNA extraction. Total RNA extractions were carried out using the PicoPure^{TM} RNA isolation kit (Arcturus) following manufacturer's guidelines. In order to limit genomic DNA contamination, a DNase treatment (RNase-Free DNase Set, Qiagen) was included in the RNA isolation procedure. The total RNA quantity and quality was assessed by spectrophotometry using the RNA-40 setting on the Nanodrop machine (Nanodrop Technologies, UK). Samples with less than 200 ng.μl^{-1} of RNA were excluded from the following steps.

Approximately five micrograms of total RNA was amplified using the RiboAmp™ amplification kit (Arcturus) according to manufacturer's instructions. The quantity and quality of product was confirmed using the Nanodrop machine (minimum cutoff values: 1000 ng.µl^{-1} and a 260/280 ratio of at least 1.7). The resulting aRNA was then reverse transcribed to cDNA using Superscript III reverse transcriptase (Invitrogen); random hexamers (Invitrogen) and fluorescently tagged Cy3- and Cy5-dUTPs (Amersham Biosciences). The Lucidea Universal Score Card (Amersham Biosciences) was used as a RNA spike-in control in all cDNA target preparations. A 1M NaOH, 20 mM EDTA solution was used to stop the reaction and degrade the original aRNA template, after which the targets were combined and cleaned using the illustra™ CyScribe™ GFX™ Purification Kit (GE Healthcare). After purification, dye incorporation was assessed using the Nanodrop machine and the microarray (33-factor) settings. A minimum cDNA yield of 15 ng.µl^{-1}; minimum dye incorporation of 0.1 pmol.µl^{-1} for each Cy dye and a 260/280 ratio of less than two, were the cutoff values used to control the efficiency of labeling and purification. Finally, 5 µg of poly dA oligo (Sigma) was added to each target set, combined and then vacuum dried using an Eppendorf Concentrator 5301 (Eppendorf, Hamburg) set at 45°C for 30 minutes. Dried pellets were resuspended in 15.5 µl of long hybridisation buffer (Corning Inc.) and denatured at 95°C for 5 minutes immediately before loading onto the array. A detailed, step by step listing of the protocol can be found in Appendix B.

3.4.3 Microarray preparation and hybridisation. The Pronto! Universal Microarray Kit (Corning Inc) was used to prepare all the microarray slides used in this study. The kit is a three step process, divided into: Pre-Soak and Pre-Hybridisation Protocol, Hybridisation Protocol and Post-Hybridisation Wash Protocol. In order to decrease the stringency of the array binding for the *An. funestus* experiments, all wash solutions were prepared and used

at 3x the recommended concentration (Christian *et al*., 2011b). The same solutions were prepared according to manufacturer's guidelines for the *An. gambiae* experiments.

Hybridisation was carried out at 38°C and 42°C for the *An. funestus* and the *An. gambiae* experiments respectively (Christian *et al*., 2011b; David *et al*, 2005). Hybridisation was carried out for 16 hours in hybridisation chambers that had approximately 40 µl of 3x SSC buffer distributed around the inner edges. Hybrislips (Grace Biolabs) were used to cover each array, to prevent evaporation of the added targets. Post-Hybridisation washes were carried out at 38°C for half the specified times (Christian *et al*., 2011b). Wash steps requiring ambient temperature were carried out as close to 25°C as possible. The full protocol can be found in Appendix B.

Experimental Design

This study used a direct comparison design with included dye swap. Briefly, for each replicate used in the hybridisations, the targets were labeled alternately with Cy3 or Cy5, in a dye swap manner thereby resulting in 2 hybridisations for each replicate. The use of three biological replicates, each with a dye swap, assisted in the control of intra-experimental variation. A final technical repeat, with dye swap, was carried out using one of the existing biological replicates. This group served to control for inter-experimental variation. In total, 11 arrays were hybridised and analysed for *An. gambiae* GAH and 8 arrays for *An. funestus* FUMOZ-RH.

Microarray Scanning

The microarrays were scanned using a GenePix 4000B scanner (Axon Instruments, Molecular Devices, USA). The photomultplier (PMT) range was adjusted to fall within

500 – 750V for each channel of detection thereby resulting in an overall intensity ratio as close to 1 as possible. The resulting TIFF files were analysed using the GenePix Pro 6.0 software (Axon Instruments). Preprocessing according to the following cutoff values was used to determine spot quality. If a spot met either of the following criteria, it was excluded from analysis: a signal to noise ratio of less than 3 and or spot intensity values of greater than 65,000. The raw intensity values were then imported into the Limma 2.4 software running in the R 2.9.0 package of Bioconductor for analysis (Gentleman *et al.*, 2004).

Microarray Data Analysis

Bioconductor is an open source bioinformatics software project that makes available a number of tools for the analysis of high throughput data. The project and the programs used in this microarray data analysis are available at www.bioconductor.org (Gentleman *et al.*, 2004). R is a statistical programming language and an interface in which one can utilize a range of statistical analysis packages. The full program code used for the *An. gambiae* GAH analysis can be found in Appendix C. Briefly, before any data transformations or normalizations were carried out, MA plots and image density plots for the Cy dyes were generated for each array. An MA plot displays the distribution of the data with regards to the red or green intensity ratio (M) plotted against the average intensity (A) for each spot. Background corrections were carried out for each spot using the 'normexp' method with an offset of 50, which ensures that the resulting corrected intensities are positive values (Ritchie *et al.*, 2007). Normalization within arrays was carried out using the global 'loess' method, without any additional offsets (Smyth and Speed, 2003). Boxplots were generated to assess the level of background distribution for each Cy dye. These boxplots could then be used to determine whether a particular array was an 'outlier' and should be removed from future analyses. Normalization between arrays was then carried

out using aquantile normalization which ensures that the average intensities (A) for each array exhibit the same empirical distribution whilst leaving the log ratios (M) unchanged (Smyth and Speed, 2003). The 'duplicateCorrelation' function was used to take into consideration that each individual spot is repeated four times within each array (Smyth *et al.*, 2005), after which differential expression was assessed using the linear model approach implemented in the 'limfit' and 'eBayes' functions (Smyth, 2004). The empirical Bayes statistic results in the calculation of the log fold change, *p*-value and *b*-statistic (*b*-stat) value for each gene on the array.

The top five up and down regulated genes were compared between the *An. gambiae* arrays and the *An. funestus* arrays. Only genes with $p < 0.001$ and a *b*-stat value >2.95 were considered.

(Note: To calculate the %probability that the gene shows true differential expression, one can use the *b*-stat value in the following function:

$$\% \text{ probability} = \frac{\exp^{(b\text{-stat value})}}{(1+ \exp^{(b\text{-stat value})}]} \times 100$$

In general, *b*-stat values of >2.95 correlate to a >95% probability that the data represents a true event or difference. A *b*-stat value >4.5 suggests a probability of >99% that the gene is differentially expressed.

3.4.4 Quantitative Real Time PCR. Quantitative PCR was carried out using the RNA samples generated for the microarray experiments and the standard curves produced below. The Bio-Rad CFX96™ Real-Time PCR Detection System (Bio-Rad Laboratories Inc.) was

used to perform the real-time qPCR experiments. The target genes of interest for qPCR analysis include *CYP*6P9; *CYP6P13*; *CYP6Gen* and *GSTD3*.

CYP6P9, CYP6P13 and *CYP6Gen* were chosen for analysis in real time as they have already been implicated in the pyrethroid resistance phenotype of FUMOZ-RH (Amenya *et al.,* 2008; Matambo *et al*., 2010). *GSTD3* was the only potential gene upregulated in both the *An. funestus* and the *An. gambiae* microarray studies. However, *GSTD3* has not been investigated in *An. funestus* and does not seem to have an orthologue in the *An. funestus* EST database (Serazin *et al.*, 2009; Gregory *et al*., 2011). Since the *An. funestus* genome has not been sequenced the percentage similarity between many of the genes of *An. funestus* and its closest relative *An. gambiae* remains unknown. Primers thus needed to be designed for all three potential gene targets (see below). The CYP6Gen primer pair is described by Amenya *et al*. (2008). However, it has subsequently been found that the primer pair may target *CYP6* genes indiscriminately and have hence been named CYP6Gen in this study.

The 25 µl reactions were set-up using: 12.5 µl IQ SYBR Green Supermix (Bio-Rad Laboratories Inc.); 1 µl Forward primer (10µM); 1 µl Reverse primer (10µM); 1 µl cDNA and 9.5 µl nuclease-free water. The thermal cycling conditions are described in table 3.1 below.

Table 3.1 Thermal cycling conditions for qPCR experiments

Initial denaturation	95°C for 3 minutes
Amplification – cycle 36-40 times	95°C for 10 seconds 55/59.1°C* for 15 seconds 72°C for 15 seconds single fluorescence detection at 510-530 nm
Melt curve detection	Ramp from 60°C to 95°C: Increment of 0.5°C for 5 seconds, single fluorescence detection at 510-530 nm

* 55°C for *CYP6P13* and *CYP6Gen* amplification; 59.1°C for *CYP6P9* amplification. Annealing temperature for the reference gene was run at the same temperature as the target.

Target Primer design for Quantitative Real Time PCR

CYP6P9, CYP6P13 and CYP6Gen Primers

Since *CYP6P9* (EU450763) and *CYP6P13* (EF152577) share 93.7% nucleotide similarity, alignments of these two genes were carried out to identify regions of dissimilarity. Primers for the *CYP6P9* gene were designed manually to fall with either their first, last, or preferably both nucleotides falling on a region of dissimilarity. All possible primer pairs were then entered into the freely available Netprimer program (Premier Biosoft International: www.premierbiosoft.com) to assess the possible self-primer dimers, cross-primer dimers and hairpin formations. All primer pairs were confirmed for sequence specificity using the Megablast algorithm on the NCBI website (http://blast.ncbi.nlm.nih.gov). The two best primer pairs (with respect to specificity and dimer formation) were chosen and each pair was then assessed for their performance in real-time qPCR, see table 3.2 below. The primers targeted to the *CYP6P13* region were designed by Riann Christian and are detailed in the table 3.2 below (Christian *et al*. 2011a). The CYP6Gen primer pair was originally described by Amenya *et al.* (2008).

GSTD3 Primers

Since the nucleotide sequence for the potential *GSTD3* gene in *An. funestus* is unknown, alignments of *An. gambiae* sequences (XM 313667; AF513638; BX046867) with the GSTD3 detox chip probe were carried out for primer design. Two primer pairs were designed and analysed for suitability in the Netprimer program as above. Sequence specificity was confirmed using Megablast as above.

The first set of primers (GSTD3 For and Rev) was designed to amply a large region of the *GSTD3* target. The primers were tested on five different *An. gambiae* colonies and the resulting amplicons were sequenced (Inqaba Biotechnologies, Pretoria, SA) to ensure specificity to the *GSTD3* gene. The primers were then applied to *An. funestus* genomic DNA extracted from three FUMOZ-RH individuals. In an effort to reduce the non-specifics obtained in this PCR, attempts were made to optimize both salt concentration and annealing temperature. The amplicons that resulted from the 45°C annealing reaction were ligated into a pGem®-T Easy Vector (Promega Corporation) and then transformed into *E. cloni*® 10G Chemically Competent Cells (Lucigen® Corporation). Transformants were screened for insert using the *LacZ* insertional inactivation system. White transformants containing potential inserts were screened using the SP6/T7 PCR screening method described in the manufacturer's guidelines. Five representative clones for each insert length obtained were sequenced (Inqaba Biotechnologies). Sequences were aligned using the DNAStar Lasergene 7 package (DNAStar Inc., Madison, WI). The resulting contigs were blasted for sequence identification using Megablast as above and manually aligned to the GSTD3 probe from the *An. gambiae* detox chip.

All of the contigs obtained from the sequencing were translated in all six reading frames using SeqBuilder in the DNAStar Lasergene 7 package (DNAStar Inc., Madison, WI). These amino acid sequences were then aligned to the translated GSTD3 probe sequence. The best fit translated sequence was then aligned to multiple *An. gambiae* and *Ae. aegypti* delta-class GSTs.

The second primer pair (GD3 For and Rev) was designed to give a small amplicon which could be used for real-time analysis of the *GSTD3* gene.

Table 3.2 Target genes for quantitative real time amplification

Primer	Sequence	Primer Tm	Annealing Temp	Amplicon length	Product Tm	Reference
CYP6P9 For1	5' TGC ATT CGG GAT TGA GTG TA 3'	58.3°C	55.0°C	209bp	81°C	This study
CYP6P9 Rev1	5' ATT CCA CCG TTT CCT TAA CA 3'	56.3°C				
CYP6P9 For2	5' AGA TGT GAT TGG CAC CTG T 3'	58.0°C	55.0°C	232bp	82°C	This study
CYP6P9 Rev2	5' TCG ATA TTC CAC CGT TTC CT 3'	58.3°C				
CYP6P13 For	5' CTG GAT CTC CTA ATT ATG ATG AAG TTT TTC 3'	61.9°C	59.0°C	132bp	81°C	Christian *et al*., 2011a
CYP6P13 Rev	5' GTT CAC CGT CTC GCG GAC T 3'	64.5°C				
CYP6PGen For	5' GAG GAA GTG AAG AAG CGA CAT C 3'	62.7°C	55.0°C	141bp	84.5°C	Amenya *et al*., 2008
CYP6PGen Rev	5' TGA CGG TGA GAA GCG GAA C 3'	62.3°C				
GD3 For	5' GAG CAC TTC CTT ACC GAA CG 3'	62.5°C	45 – 55.0°C	103bp*	~	This study
GD3 Rev	5' CGT ACT TCA GCC AGT TCA GTG3'	62.6°C				
GSTD3 For	5' TCG CCG TAG TCA GTT CAG ATG 3'	62.6°C	45 – 55.0°C	537bp*	~	This study
GSTD3 Rev	5' CGG TTC GAG ATC GTA CTT CAG 3'	62.6°C				

* Predicted amplicon size based on *An. gambiae* cDNA sequence.

Reference Primer Design for Quantitative Real Time PCR

Reference genes (RG) or housekeeping genes (HKG) are used to "calibrate" target gene expression levels. These RGs should have stable expression between the untreated and treated samples. Generally, more than one RG is required to perform an accurate assessment of the expression levels of the target gene of interest (Vandersompele *et al*., 2002; Bustin *et al*., 2009; Chervoneva *et al*., 2010; Derveaux *et al*., 2010). In order to choose the most suitable candidate RGs, a number of potential genes are analysed for suitability.

The RGs screened for suitability in this study are listed in table 3.3 below. A box and whisker plot was drawn to show the distribution of the quantification cycle (Cq) values for each HKG. The following freely available Microsoft Excel Visual Basic macros were used to assist in the assessment of HKG suitability: Bestkeeper (Pfaffl *et al.*, 2004); geNorm (Vandesompele *et al.*, 2002) and Normfinder (Andersen *et al*., 2004).

Bestkeeper uses raw Cq values to rank candidate HKGs based on the standard deviation of the Cqs and performs repeated pairwise correlation analyses between the HKGs. Ultimately, the assessment results in a 'Power of HKG' value and *p*-value for each gene, for which the smallest possible value is preferred.

geNorm calculates an *M*-value for each gene by assessing the average pairwise variation of the relative starting quantities entered by the user. The least stable gene, having the highest *M*-value, is then eliminated from the assessment and the remaining genes are re-analysed. This

step-wise elimination of the least suitable genes eventually results in the two most favourable candidates. Acceptability of a reference gene depends on whether the M-value is below 0.5 for homogenous sample sets or less than 1.0 for heterogenous sample sets (Taylor *et al*., 2010).

NormFinder uses linear expression quantities to determine the variation within and between the test groups (eg. blood fed and unfed samples) and uses these values to derive a stability value for ranking the candidate RGs. The program will then determine the most suitable RG for a single reference gene experiment as well as the best pair of RGs in multiple reference gene experiments.

Amplicons derived from the test runs were ligated, cloned and transformed, as above. Clones were sequenced by Inqaba Biotech (Pretoria, South Africa) to ensure that the correct reference genes were being targeted.

Standard Curve Preparation for Quantitative Real Time PCR

Total RNA was extracted from 10 – 15 female, 3-day old, unfed, FUMOZ-RH mosquitoes using the Trizol™ RNA extraction protocol (Invitrogen). Pestles were cleaned with RNaseZap (Sigma), rinsed with DEPC-treated water and then autoclaved for 20 minutes. The final RNA pellet was resuspended in nuclease-free water (Ambion) and subjected to DNase treatment (Qiagen) to remove any DNA carryover. In order to inactivate the DNase enzyme, the RNA samples were briefly denatured at 72˚C for 10 minutes. The resulting RNA was quantified

using a Nanodrop spectrophotometer and quality was assessed by formaldehyde gel analysis and Agilent 2100BioAnalyzer (detailed protocols below).

Table 3.3 Table of candidate reference genes. Note that the sequence for the RPS7 primer pairs originates from Amenya *et al*., 2008. All other primer pairs were designed in this study.

Primer	Sequence	Primer Tm	Genbank Accession No/EST Database No	Amplicon length
RPS7 For	5' TTA CTG CTG TGT ACG ATG CC 3'	60.4°C		
RPS7 Rev	5' GAT GGT GGT CTG CTG GTT C 3'	62.3°C	EF450776.1	134bp
RPL8 For	5' CAT CAG CAC ATT GGT AAG GC 3'	60.4°C	CD664267.1/	
RPL8 Rev	5' GTT TTC GCT TCC CGT TTT TC 3'	58.4°C	AF-NORA-contig_176	305bp
RPL19 For	5' GAA ACA CCA ACT CCC GAC A 3'	60.2°C		
RPL19 Rev	5' TCA ACA GGC GAC GCA ACA C 3'	62.3°C	DQ910355.1	223bp
RPS26 For	5' GAT AAG GCA ATC AAG AAG TTC G 3'	59.0°C	CD577850.1/	
RPS26 Rev	5' TAC GGA CAA CCT TCG AGT GG 3'	62.5°C	AF-NORA-contig_270	160bp
CO1 For	5' TAG GAG CCC CTG ATA TAG CTT TC 3'	62.8°C		
Co1 Rev	5' ACT GTT CAT CCT GTT CCT GCT C 3'	62.7°C	AY423059.1	123bp
ND5 For	5' TAG AAT TTT ATT AGG GTG GGA TGG 3'	59.4°C		
ND5 Rev	5' ATC TCC AAT TCG ATT TGA TAA TGC 3'	57.7°C	AY727744.1	122bp
Fun 18S For	5' GTG TAC TTG GGC GTT ACT CTG TG 3'	64.6°C		
Fun 18s Rev	5' CTT TGA GCA CTC TAA TTT GTT CAA G 3'	59.7°C	AF417780.1	116bp
GapDH For	5' GAC TGC CAC TCG TCC ATC 3'	62.2°C		
GapDH Rev	5' CCT TGG TCT GCA TGT ACT TG 3'	60.4°C	EZ966147.1	139bp

cDNA was synthesized using the High Capacity RNA-to-cDNA Kit (Applied Biosystems). An acceptable yield of cDNA was >2000 $ng.\mu l^{-1}$ with a 260/280 ratio of >1.8. cDNA generated in this manner was used to create the standard curves used in the real-time analyses below. A two-fold serial dilution of the cDNA was generated and used as the standard curve template in the real time qPCR reactions. Only experiments that had a standard curve with an efficiency of 95-102% and an R^2 >0.97 were used for analysis. A minimum of 7 out of the 11 standards were required to create the final standard curve using the Bio-Rad CFX Manager Software Version 1.5 (Bio-Rad Laboratories Inc).

Assessment of RNA Quality

The quality of the extracted RNA was assessed using both formaldehyde gel electrophoresis and the BioAnalyzer micro-fluidics platform developed by Agilent Technologies Inc.

Formaldehyde Gel Electrophoresis

Before preparing the gel, all equipment including gel tank, casting tray, combs, 200ml erlenmeyer flask etc was soaked in 0.1M NaOH, for a minimum of 1 hour. All the equipment was rinsed with DEPC-treated water and excess water shaken off. A 1.3% agarose gel was prepared by dissolving 0.65g agarose in 44ml DEPC-treated water in an erlenmeyer flask. The solution was boiled well to ensure that the agarose had completely dissolved. Whilst allowing the molten agarose to cool to 55-60°C, the opening of the flask was covered with foil to prevent further evaporation. Whilst the agarose was cooling, 5 ml 10x MOPS running buffer

(0.4M MOPS pH 7.0; 0.1M Na-acetate; 10 mM EDTA) was mixed with 1.5 ml 37% formaldehyde in a 50 ml falcon tube. In a fume hood, the cooled agarose was combined with the MOPS-formaldehyde mix and 3 µl ethidum bromide (0.5 $mg.ml^{-1}$). The formaldehyde agarose mix was then poured into a casting tray with the appropriate combs in place. The gel was allowed to set for at least half an hour. Before using the gel, it was pre-electrophoresed for at least 20 minutes at 5V/cm using 1x MOPS running buffer in the gel tank.

Each RNA sample was prepared in a 0.5 ml PCR tube by mixing: 2.5 µl 10x MOPS running buffer; 4.4 µl 37% formaldehyde solution; 12.5 µl deionised formamide; up to 5.6 µl RNA sample and nuclease-free water to a final total volume of 25 µl. The solution was vortexed and briefly spun down. Heat denaturation of the mixture was carried out in a thermal cycler at 55°C for 10 minutes then 70°C for 5 minutes. Immediately after denaturation the samples were placed on ice, until they were loaded on the gel. Prior to loading the samples, the 25 µl denatured sample was mixed with 5 µl Orange-G loading buffer (0.35% w/v Orange-G; 30% w/v Ficoll; 1 mM EDTA) and 0.2 µl ethidium bromide (0.5 $mg.ml^{-1}$). The samples were electrophoresed for no more than 2 hours at 5V/cm, with an RNA High Range molecular weight marker (Fermentas) included.

Agilent 2100 Bioanalyzer

The RNA samples and RNA 6000 Nano chips were prepared and run according to manufacturer's instructions. Briefly, the chips were loaded with prepared gel-dye mix using the syringe and priming stations provided. A marker solution was added to all sample and

molecular weight marker wells. Heat denatured RNA ladder was added to the molecular weight marker well and 1 µl of RNA sample to each of the sample wells (up to 12). The chip was briefly vortexed and then inserted into the Agilent Bioanalyzer machine. The Bioanalyzer machine, through the use of microcapillaries and laser detection, analyses the RNA present in the samples. The software generates a digital electropherogram which is then analysed for the presence and absence of degradation products. An RNA Integrity Number (RIN) is calculated whereby a RIN value of 1 is indicative of completely degraded RNA and a value of 10 is fully intact RNA. The RIN value is calculated through the use of an algorithm that utilises a number of different input factors, including: the presence of the 28S and 18S peaks; the 28S:18S peak ratio; peak area; as well as the presence of pre-, inter- and post-region peaks.

Quantitative Real Time PCR Data Analysis

The resulting qPCR data in the form of quantification cycle (Cq) values was analysed according to the ΔCq relative quantification method (Pfaffl, 2001). Four biological samples were amplified in triplicate for each run. Unfed mosquitoes represented the untreated samples and blood fed mosquitoes represented the treated samples. The mean Cq value for each was used in the following equation to determine the fold over-expression or relative expression ratio.

$$\text{Relative Expression Ratio} = \frac{(\text{Efficiency of target})^{\Delta Cq \text{ Target (treated – untreated)}}}{(\text{Efficiency of HKG})^{\Delta Cq \text{ HKG (treated – untreated)}}}$$

3.5 RESULTS

Assessment of RNA Quality

As part of the MIQE and MIAME guidelines, RNA integrity needs to be assessed to ensure that the results of any expression experiment are trustworthy. Figures 3.1 and 3.2 below show total RNA extracted from FUMOZ-RH mosquitoes for use in standard curve preparation. All RNA samples used to generate the microarray data and real time analyses were run on either formaldehyde gels or Agilent RNA 6000 nano chips. The presence of a strong 18S RNA band and a weaker 28S band was observed for all RNA extracts. A RIN (RNA Integrity Number) value of > 6.0 was achieved for all the samples used in this study.

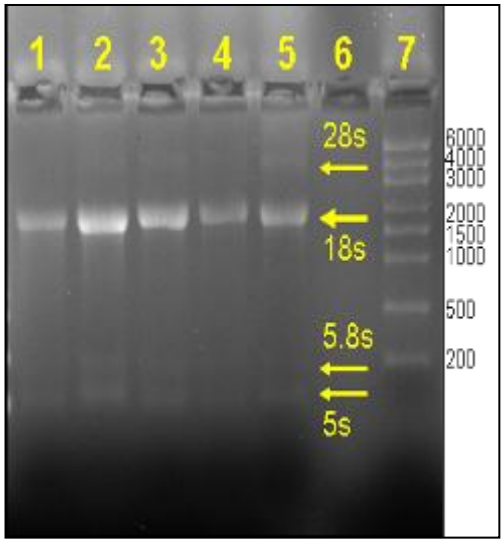


Figure 3.1 Formaldehyde gel electrophoresis of RNA used to prepare standard curves for real-time quantitative PCR. Lanes 1-5: FUMOZ-RH Total RNA. Lane 6: Empty. Lane 7: RNA Molecular Weight marker. (1.3% agarose gel, Fermentas High Range RNA Ladder SM0423).

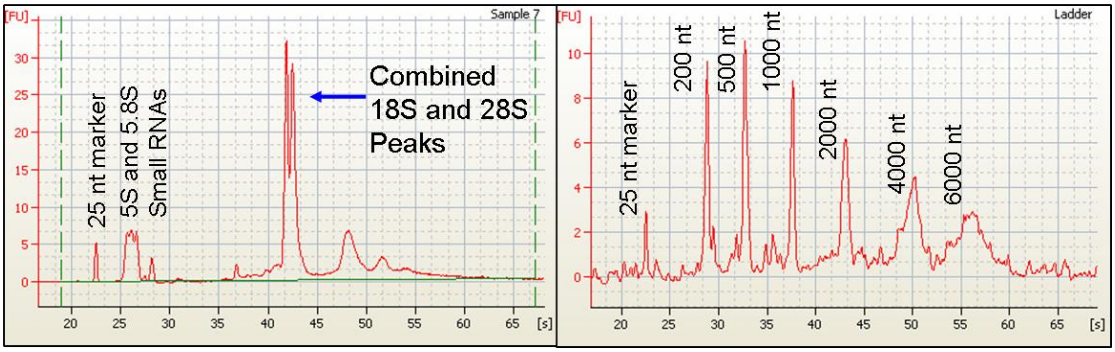


Figure 3.2 Agilent Bioanalyzer run of FUMOZ-RH RNA used to create cDNA for the standard curves used in the real-time quantitative PCR experiments. A: Typical FUMOZ-RH RNA profile. B: RNA ladder run in parallel to the RNA samples.

Microarray analysis of the effect of a blood meal on detoxification gene expression in blood fed versus unfed An. gambiae

Normalization of microarray data is necessary to ensure that differences in gene expression are due to true gene effects and not artifacts introduced by technological differences such as dye-biases and print-tip efficiencies. Figure 3.3 depicts the variation in background fluorescence intensities for both the red and green channels after internal normalization for the *An. gambiae* GAH strain on each of the detox chips. These plots allow the slides to be analysed for extreme outliers in terms of fluorescence signals. A slide that has too many extreme outliers and a median that is greatly removed from the average can then be excluded from further analysis. The plots below show that although the overall background fluorescence for the green channel is more dispersed than that of the red channel, the slides included below can be analysed together.

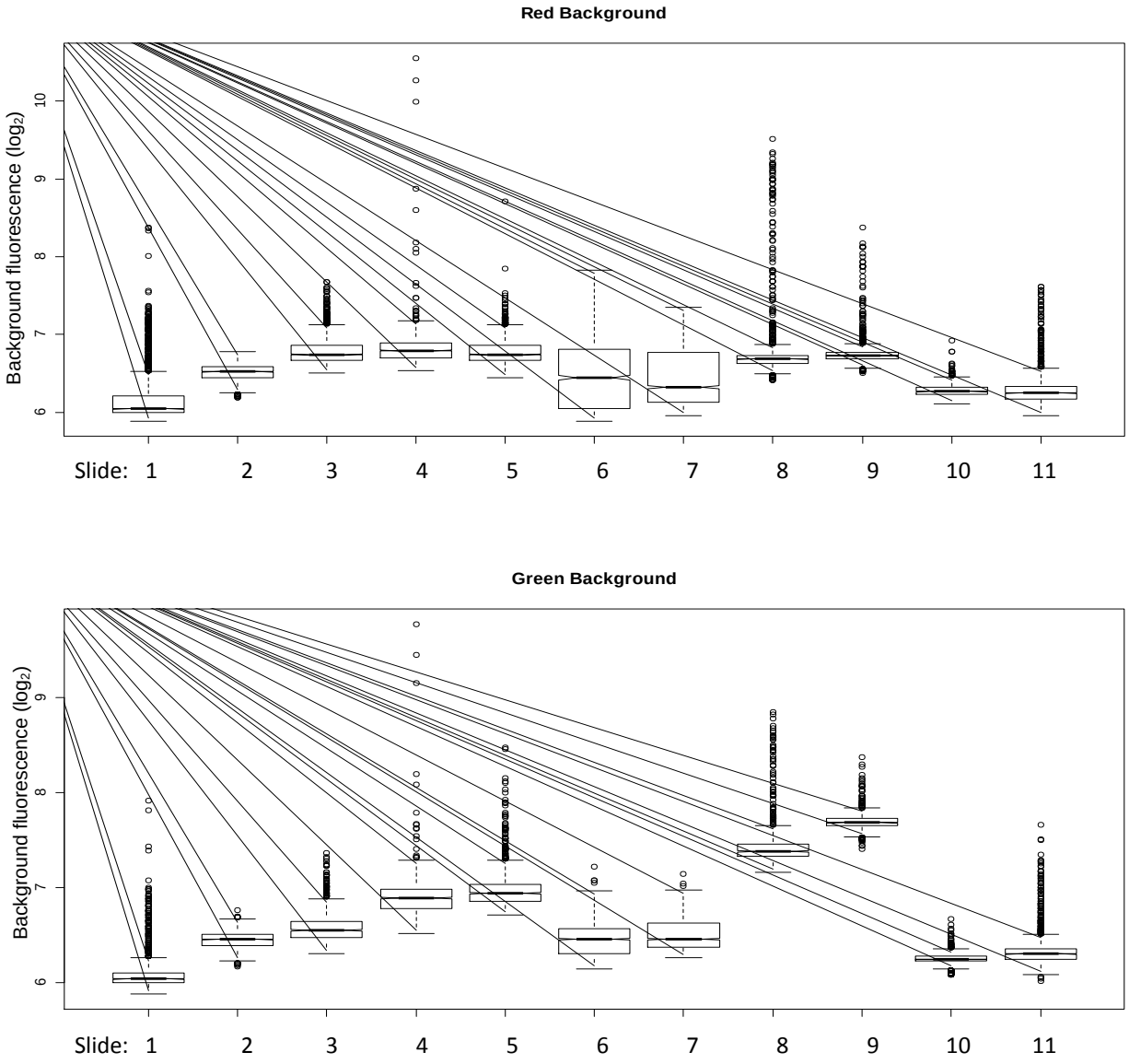


Figure 3.3 Box and whisker plots depicting the $\log_2$ background fluorescence after internal normalization for the *An. gambiae* GAH microarrays.

After normalizing the fluorescence between the slides, the differential expression of genes in response to a blood meal could be calculated using Bayes statistics. A volcano plot was generated (Figure 3.4A) highlighting the top 10 up and down regulated genes, after which the

results were filtered according to *b*-stat values to reveal those that had the highest probability of reflecting a true change in expression levels (Figure 3.4B). The top five down regulated genes belong to the monooxygenase P450 groups Cyp6 and Cyp9 and displayed log fold changes between 0.2 to 0.3 fold change in expression (Table 3.4). *CYP9J5* was excluded from the analysis due to its *b*-statistic of 2.21, falling below the cutoff of 2.95, even though the *p*-value <0.001 suggests significance. Four of the most up-regulated genes belong to the Delta-class glutathione-S-transferases, each displaying at least a 2 fold increase in expression (Table 3.4). However, *GSTD1-3* had a *b*-statistic value of 1.96 which suggests only an 88% probably of being a true change in differential expression and was excluded from the analysis since it fell below the *b*-statistic cut-off. The fifth most up-regulated gene was found to be a glutaredoxin gene, *GRX1*. This glutaredoxin gene was only slightly upregulated with a final fold change of 1.76.

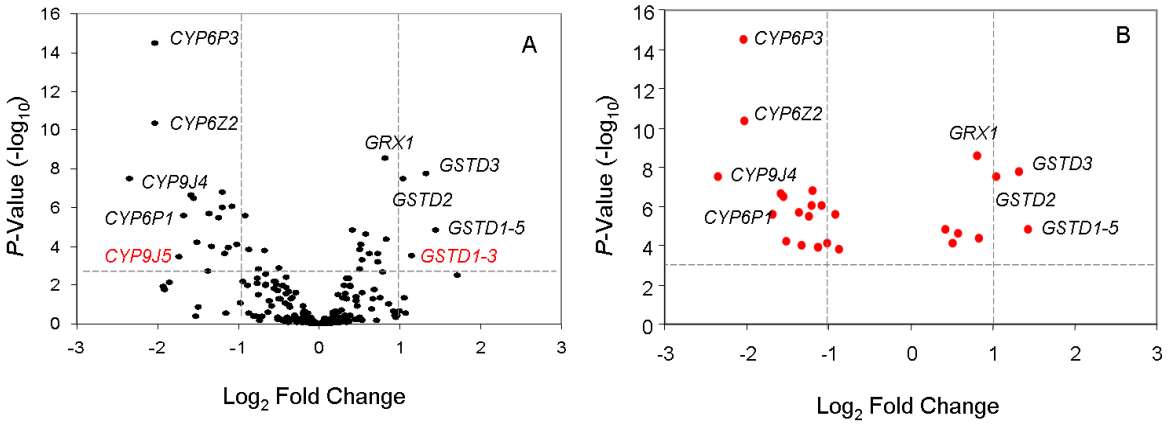


Figure 3.4 Volcano plots for *An. gambiae* GAH on the detox chip (Cons corr: 0.69). A depicts all the genes from the analysis, with top 5 up- and down-regulated genes identified. B depicts the top 4 up- and down-regulated genes after filtering according to *b*-statistic values. Note: *CYP9J5* and *GSTD1-3* are filtered out due to their low *b*-statistic value. Up-regulated genes possess a $\log_2$ fold change greater than 1. Down regulated genes possess a $\log_2$ fold change less than 1.

Table 3.4 Top differentially expressed genes in *An. gambiae* GAH, in response to the presence of a blood meal. Genes highlighted in red fell just below the *b*-statistic cut off of 2.95 even though the P-value was significant.

Gene	Vector Base Gene ID	Function	Location	Fold Change	*P*-value	*b*-statistic value
Up-Regulated Genes						
GSTD1-5	AGAP004164	Glutathione S-Transferase	2R	2.71	1.55E-05	5.67
GSTD3	AGAP004382	Glutathione S-Transferase	2R	2.50	1.83E-08	14.26
GSTD1-3	AGAP004164	Glutathione S-Transferase	2R	2.21	3.06E-04	1.96
GSTD2	AGAP004165	Glutathione S-Transferase	2R	2.07	3.30E-08	13.02
GRX1	AGAP011107	Glutaredoxin	3L	1.76	2.83E-09	15.90
Down-Regulated Genes						
CYP6P1	AGAP002868	Cytochrome P450	2R	0.31	2.78E-06	8.08
CYP9J5	AGAP012296	Cytochrome P450	3L	0.30	3.42E-04	2.37
CYP6Z2	AGAP008218	Cytochrome P450	3R	0.25	4.57E-11	20.28
CYP6P3	AGAP002865	Cytochrome P450	2R	0.24	3.28E-15	30.24
CYP9J4	AGAP012292	Cytochrome P450	3L	0.20	3.30E-08	13.25

Microarray analysis of the effect of a blood meal on detoxification gene expression in blood fed versus unfed An. funestus

Figure 3.5 shows the distribution of the background fluorescence for the hybridisation of blood fed and unfed *An. funestus* onto the *An. gambiae* detox chip. Similarly to Figure 3.3 above, these plots allow the slides to be analysed for extreme outliers in terms of fluorescence signals. Again, the overall background fluorescence for the green channel is more dispersed than that of the red channel and in this experiment fewer slides could be included and analysed together due to greater levels of slide variation.

After normalization the differential expression of genes in response to the presence of a blood meal could be calculated using Bayes statistics. The resulting volcano plot (Figure 3.6) depicts the five most up-regulated genes in response to a blood meal. Due to the large amount of variation between slides, the consensus correlation is low and a lower fold change ($\log_2$=0.5) has been used as a cut-off for significance. There are no significantly down regulated genes as can be seen by the absence of genes falling within the upper left hand quadrant of the volcano plot.

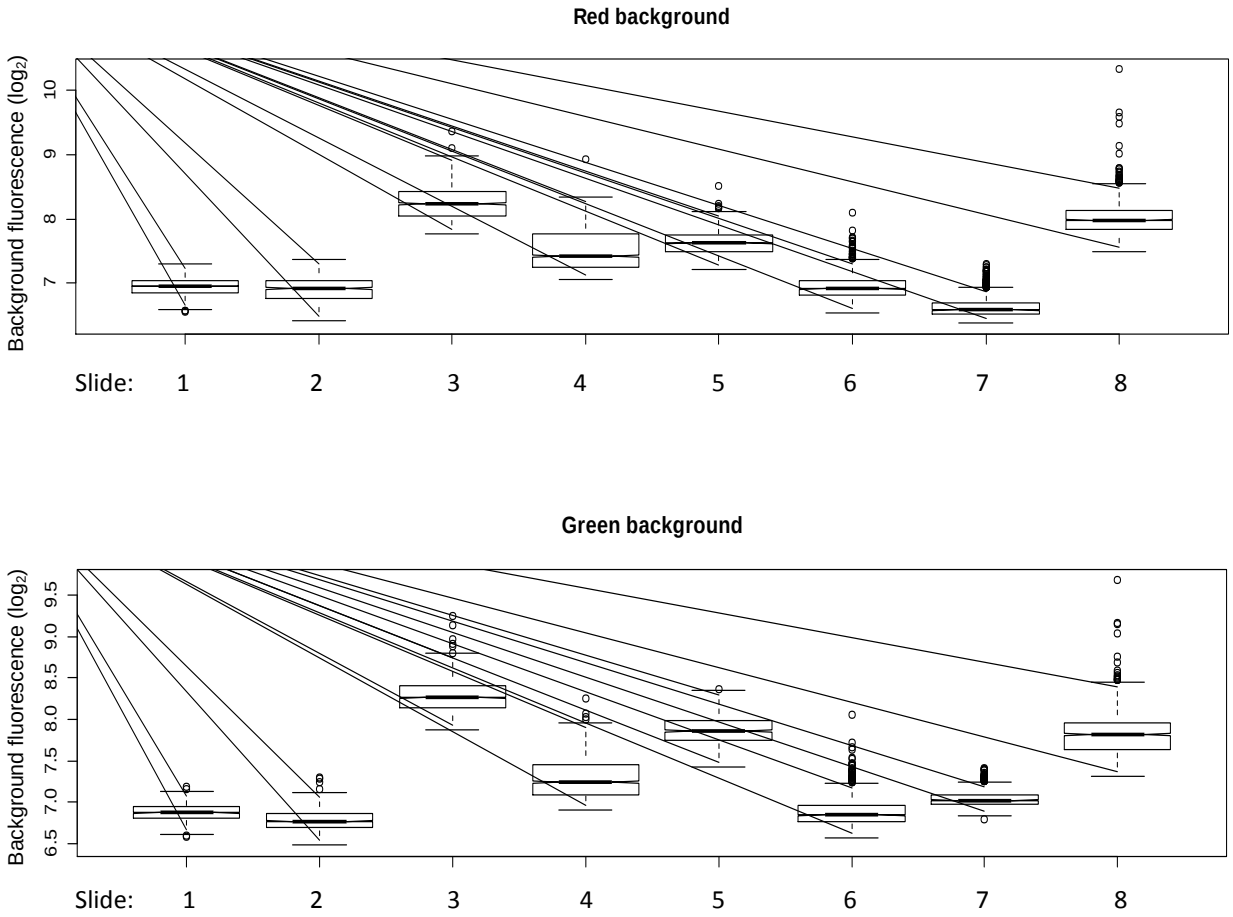


Figure 3.5 Box and whisker plots depicting the $\log_2$ background fluorescence after internal normalization for the *An. funestus* FUMOZ-RH microarrays.

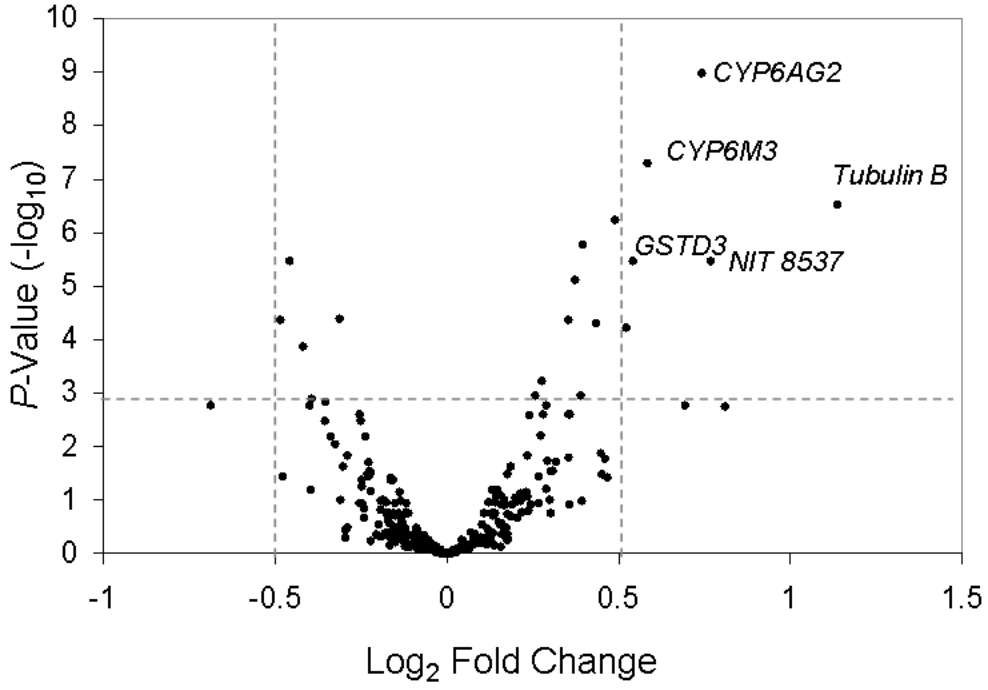


Figure 3.6 Volcano plot for *Anopheles funestus* FUMOZ-RH on the detox chip (Cons corr: 0.53). All the genes from the analysis are depicted, with the top 5 up-regulated genes identified. Note: none of the genes fell within in the significantly down-regulated quadrant. Genes considered to be up-regulated possess a $\log_2$ fold change greater than 0.5.

Table 3.5 lists the top five up-regulated genes. Tubulin B is the only gene that displays a significant increase in expression, with a doubling of expression levels in response to the blood meal. Two cytochrome P450 genes, a glutathione-S-transferase gene and a nitrilase are all marginally up-regulated with significant *p*-values and *b*-statistics >2.95. *GSTD3* is the only gene that displays up-regulation in both the *An. gambiae* GAH and *An. funestus* FUMOZ-RH microarray experiments.

Table 3.5 Top differentially expressed genes in *An. funestus* FUMOZ-RH, in response to the presence of a blood meal. Vector Base Gene ID or *genBank ID, Function and Location for the *An. gambiae* gene probes found to be up-regulated.

Gene	Vector Base Gene ID	Function	Location	Fold Change	*P*-value	*b*-statistic value
Up-Regulated Genes						
Tubulin B	AGAP010510	Tubulin - Structural	3L	2.20	3.04E-07	6.52
NIT 8537	AGAP003515	Nitrilase	2R	1.70	3.47E-06	5.46
CYP6AG2	*GB: AY745224	Cytochrome P450	2R	1.67	1.08E-09	8.97
CYP6M3	AGAP008213	Cytochrome P450	3R	1.50	5.19E-08	7.28
GSTD3	AGAP004382	Glutathione S-Transferase	2R	1.46	3.47E-06	5.46

Reference Gene Assessments

Eight different pairs of candidate reference genes were anaylsed in real-time for suitability as reference. Each pair was tested using cDNA from the blood fed and unfed *An. funestus* cohorts used in the microarray study. Figure 3.7 shows the amplification curves generated for seven of the eight candidate genes. *RPL8* could not be included in the analysis as the primer pair targeting this gene yielded more than one amplification product, evidenced by double peaks in the melt curve analysis (see Appendix D).

Visual inspection of the amplification curves in Figure 3.7 can help to confirm the suitability of the candidate genes. A suitable candidate gene, such as *RPL19* and *RPS7*, should have equal amplification for the blood fed and unfed samples, since the starting concentration of

cDNA used is the same for all samples. Hence, the coloured amplification lines should overlap. Poor candidates are typified by separated amplification curves for each treatment type, as can be seen for the GAPDH, CO1 and FUN18S primer pairs. These three primer pairs display differential expression for the blood fed and unfed samples.

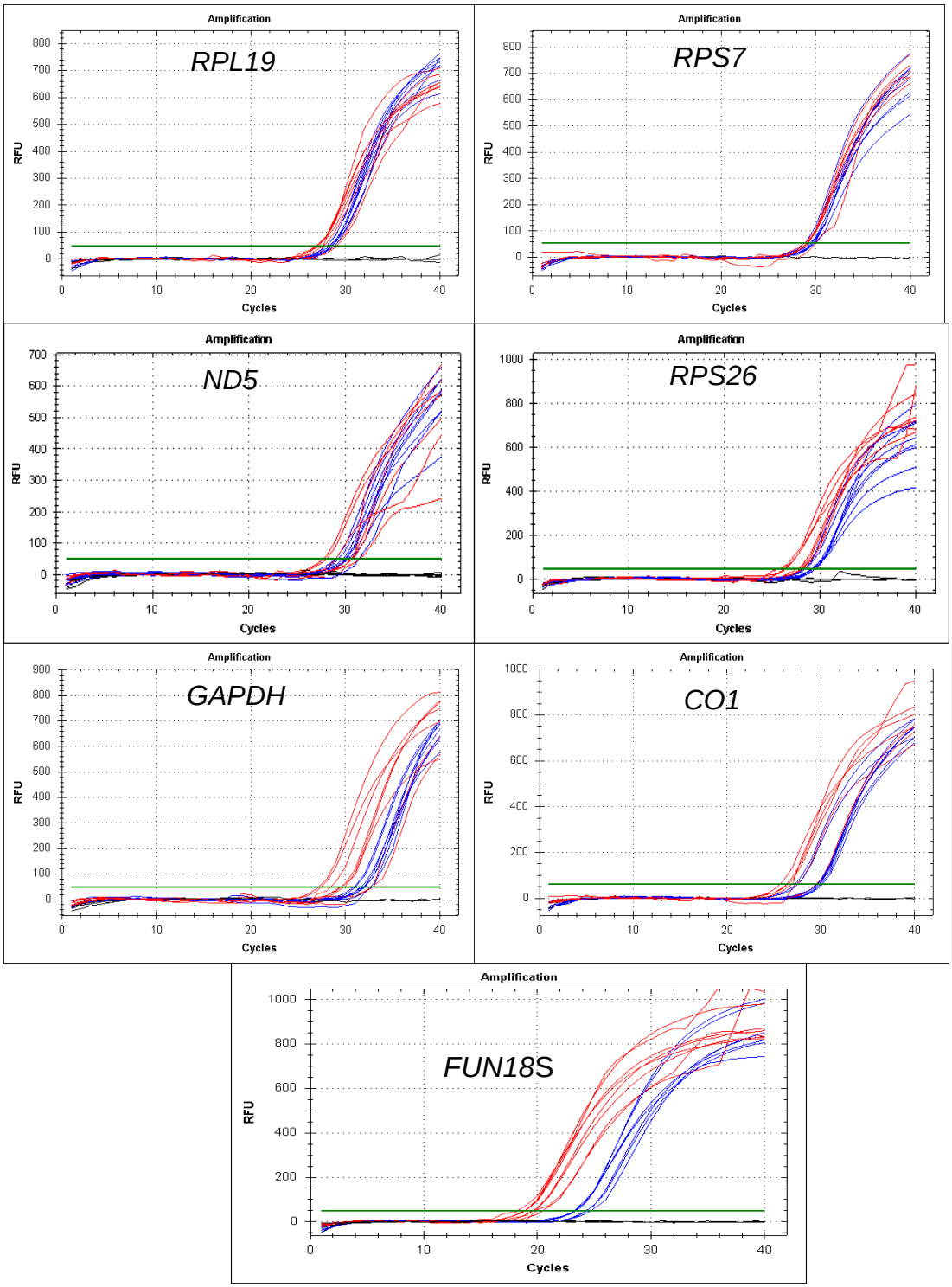


Figure 3.7 Amplification curves using *An. funestus* cDNA, to assess the suitability of the candidate reference genes. Red = blood fed samples, Blue = Unfed Samples.

Since visual inspection of amplification curves cannot be the sole discriminating factor for assessing candidate primer pairs, a box and whisker plot of the Cq (Cycle of quantitation) values was generated (Figure 3.8). This plot depicts the distribution of the Cqs for each primer pair enabling one to graphically assess the variation in expression levels.

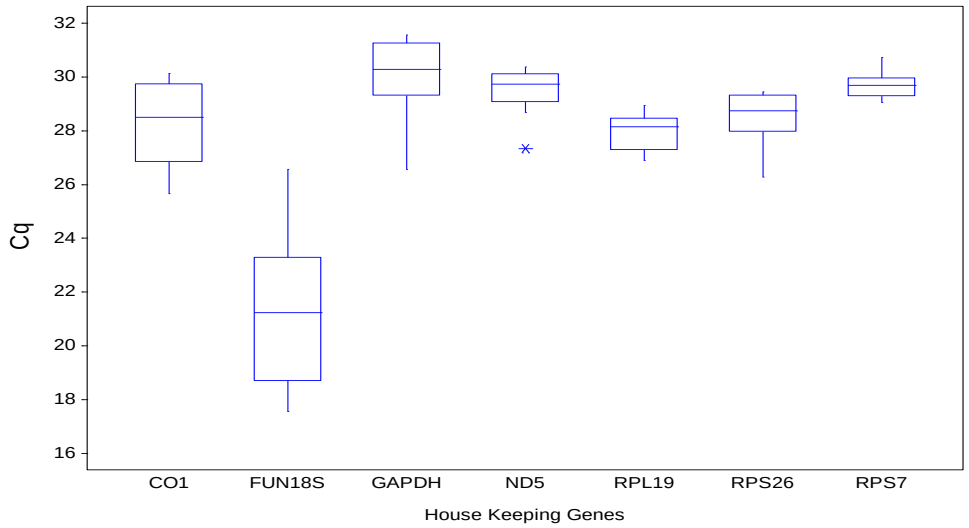
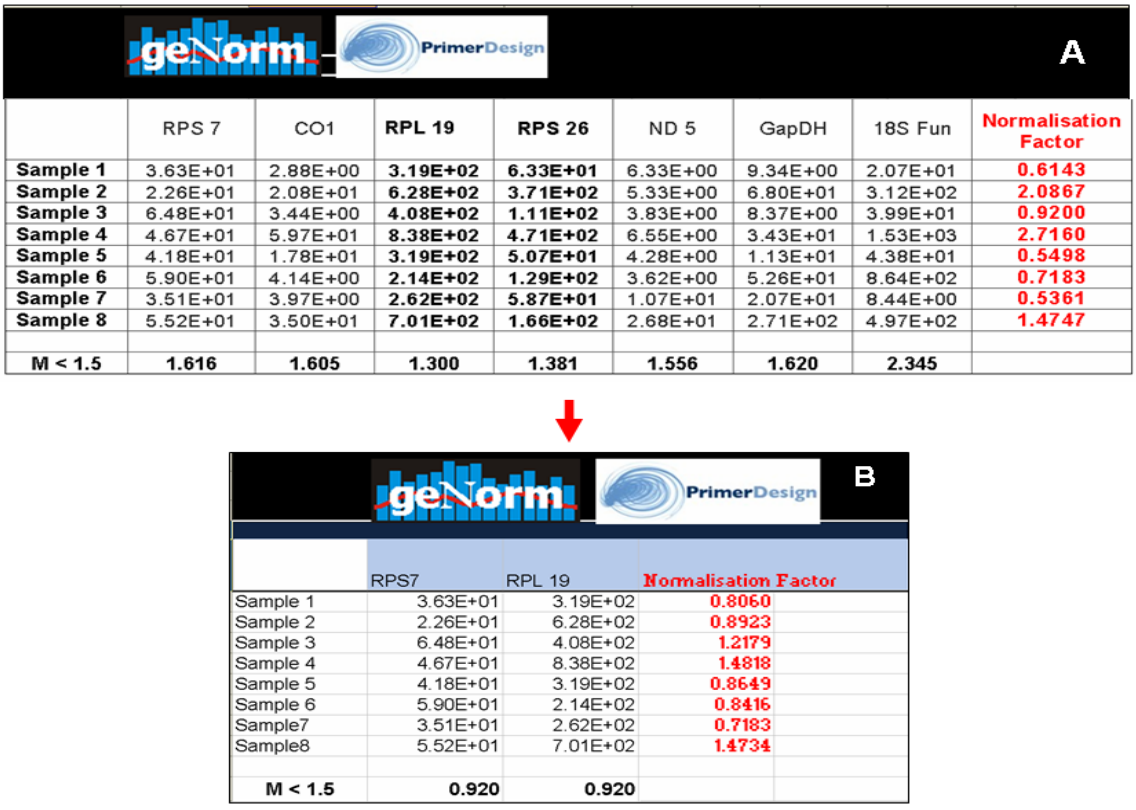


Figure 3.8. Box and Whisker Plot showing the distribution of Cq values for each candidate reference gene. The central line through the boxes indicates the median, extended vertical bars represent the standard deviation of the mean derived from 8 individual samples run in duplicate. * = an outlier for the *ND5* group.

The greatest variation in Cqs is displayed graphically for the *FUN18S* primer pair, where as in contrast, *RPS7* shows the least dispersed Cqs. Although box and whisker plots are suitable for excluding the extremely poor candidate pairs, it is also necessary to use reference gene analysis tools to determine what the best combination of reference genes is. Three different analysis tools (geNorm, NormFinder and Bestkeeper) were used and compared.

The visual basic plugin, geNorm (Vandersompele *et al.*, 2002), requires the user to enter relative starting quantities into the analysis tool (Figure 3.9A). In this screen shot, the first round of analysis has taken place and the two best candidate genes are highlighted in bold. The poorest candidate, *FUN18S*, has the highest M-value and is removed from the data set and the analysis repeated. Figure 3.9B is the last stage of the step-wise analysis of the candidate reference genes, showing that *RPS7* and *RPL19* are the two most suitable reference genes. The M-value is below 1.0, which is acceptable since the data entered into the algorithm was for a heterogenous sample set.



A

	RPS 7	CO1	RPL 19	RPS 26	ND 5	GapDH	18S Fun	Normalisation Factor
Sample 1	3.63E+01	2.88E+00	**3.19E+02**	**6.33E+01**	6.33E+00	9.34E+00	2.07E+01	**0.6143**
Sample 2	2.26E+01	2.08E+01	**6.28E+02**	**3.71E+02**	5.33E+00	6.80E+01	3.12E+02	**2.0867**
Sample 3	6.48E+01	3.44E+00	**4.08E+02**	**1.11E+02**	3.83E+00	8.37E+00	3.99E+01	**0.9200**
Sample 4	4.67E+01	5.97E+01	**8.38E+02**	**4.71E+02**	6.55E+00	3.43E+01	1.53E+03	**2.7160**
Sample 5	4.18E+01	1.78E+01	**3.19E+02**	**5.07E+01**	4.28E+00	1.13E+01	4.38E+01	**0.5498**
Sample 6	5.90E+01	4.14E+00	**2.14E+02**	**1.29E+02**	3.62E+00	5.26E+01	8.64E+02	**0.7183**
Sample 7	3.51E+01	3.97E+00	**2.62E+02**	**5.87E+01**	1.07E+01	2.07E+01	8.44E+00	**0.5361**
Sample 8	5.52E+01	3.50E+01	**7.01E+02**	**1.66E+02**	2.68E+01	2.71E+02	4.97E+02	**1.4747**
M < 1.5	**1.616**	**1.605**	**1.300**	**1.381**	**1.556**	**1.620**	**2.345**	

B

	RPS7	RPL 19	Normalisation Factor
Sample 1	3.63E+01	3.19E+02	**0.8060**
Sample 2	2.26E+01	6.28E+02	**0.8923**
Sample 3	6.48E+01	4.08E+02	**1.2179**
Sample 4	4.67E+01	8.38E+02	**1.4818**
Sample 5	4.18E+01	3.19E+02	**0.8649**
Sample 6	5.90E+01	2.14E+02	**0.8416**
Sample7	3.51E+01	2.62E+02	**0.7183**
Sample8	5.52E+01	7.01E+02	**1.4734**
M < 1.5	**0.920**	**0.920**	

Figure 3.9. Identification of candidate reference genes using geNorm. A represents all seven genes after the first cycle of analysis. Step-wise removal of the most unsuitable gene followed by re-analysis results in the two most suitable genes (B).

NormFinder (Andersen *et al.*, 2004) is a unique program in that the user is able to analyse the expression of the candidate reference genes by using expression quantities combined with the power of choosing to use identifiers or not. Table 3.6 below is the result of the analysis in NormFinder without the use of identifiers. In this mode of analysis, the program suggests that *RPL19* and *RPS26* are the best candidate reference genes as they have the lowest stability values.

Table 3.6 Identification of candidate reference genes using NormFinder without using identifiers.

First Round			**Second Round**		
	Stability			**Stability**	
Gene	**value**	**SE**	**Gene**	**value**	**SE**
RPS7	0.841	0.263	*RPS7*	0.907	0.290
CO1	0.734	0.241	*CO1*	0.736	0.259
RPL19	0.400	0.197	*RPS26*	0.417	0.234
RPS26	0.407	0.198	*ND5*	0.852	0.279
ND5	0.795	0.253	*GAPDH*	0.668	0.249
GAPDH	0.738	0.242	*FUN18S*	1.458	0.409
FUN18S	1.498	0.414			

The result of the NormFinder analysis, with the inclusion of identifiers, alters the outcome of the analysis (Table 3.7). When identifiers are incorporated, the program suggests that the most suitable single reference gene to use is *RPS26* and that the best combination of reference genes is *RPL19* combined with *GAPDH*.

Table 3.7 Identification of candidate reference genes using NormFinder using sample identifiers.

Gene Name	**Stability Value**
RPS7	0.830
CO1	0.613
RPL19	0.468
RPS26	**0.373**
ND5	0.836
GAPDH	0.729
FUN18S	1.353
Best combination of two genes	*RPL19* and *GAPDH*
Stability value for best combination of two genes	0.305

	Intragroup Variation	
Group Identifier	**Unfed**	**Blood Fed**
RPS7	0.019	0.368
CO1	0.680	0.734
RPL19	**0.005**	**0.045**
RPS26	0.128	0.432
ND5	0.419	0.366
GAPDH	**0.248**	**0.823**
FUN18S	0.539	0.707

Analysis of the reference genes in Bestkeeper (Pfaffl *et al.*, 2004) requires the input of raw Cq values. The results of this analysis (Table 3.8) suggest that *RPL19* is the most suitable reference gene, due to its low Power and *p*-value. The second and third best reference genes are *RPS7* and *ND5*. This program does not however yield a result for the best two possible primers to be used in combination.

Table 3.8 Identification of candidate reference genes using Bestkeeper.

	RPS7 vs. BK	*CO1* vs. BK	*RPL19* vs. BK	*RPS26* vs. BK	*ND5* vs. BK	*GAPDH* vs. BK	*FUN18S* vs. BK
p-value	0.628	0.017	0.053	0.031	0.428	0.043	0.002
Power [x-fold]	**1.07**	2.34	**1.36**	1.67	**1.22**	2.11	5.46

Optimization of Primers for GSTD3

Since *GSTD3* was the only gene up-regulated in both the *An. gambiae* and *An. funestus* microarray studies, it was targeted for analysis in real time qPCR. Primers were designed and tested on multiple strains of *An. gambiae* and the resulting fragments were sequenced in order to ensure specificity of the targeted region (Figure 3.10). Multiple strains of *An. gambiae* were used to cover any variation that may be due to locality or M/S status. All of the five *An. gambiae* strains tested yielded the same 603bp fragment which when sequenced confirmed that the primers were targeting the *GSTD3* gene region.

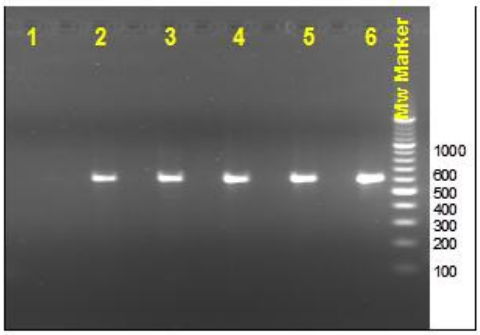


Figure 3.10 Genomic DNA extracted from *An. gambiae* specimens used to test the specificity of the GSTD3 For and Rev primer pair. The length of the amplicon will include the intron (69bp) and is expected to be 603bp in length. Lane 1: PCR –ve control. Lanes 2-6: BOA, COGS, JS3, NAG and SUA strains of *An. gambiae*. (2% TAE agarose gel, Fermentas 100bp O'Range Ruler.)

The *GSTD3* primers were then used to amplify potential *GSTD3* orthologues in *An. funestus* genomic DNA extracts. Three different FUMOZ-RH individuals were used to test the primers, with varying annealing temperatures. Figure 3.11 shows the range of amplicons that resulted from the amplification using the *GSTD3* primers.

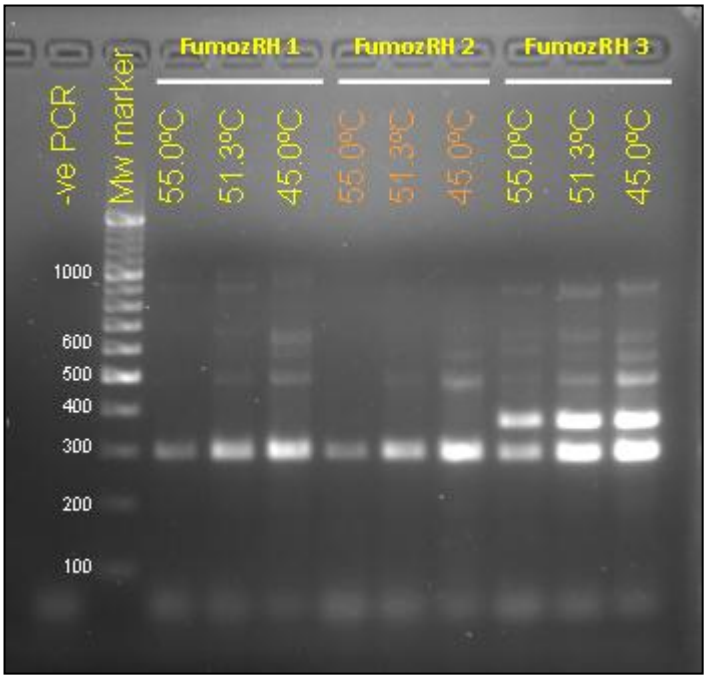


Figure 3.11 PCR products arising from a reaction using the *GSTD3* For and Rev primer pair with three different FUMOZ-RH templates. (2% TAE agarose gel, Fermentas 100bp O'Range Ruler.)

Since the *GSTD3* primers were yielding multiple fragments in *An. funestus*, all the amplicons were cloned for sequencing. Approximately 160 clones were screened and five clones from each representative fragment length were sequenced (Figure 3.12). The contigs generated from the sequencing run were analysed for sequence similarity against the *GSTD3* microarray probe, at both the nucleotide and amino acid level (Table 3.9). Contigs 2_1609; 3_1609 and

3_0110 had the highest % nucleotide sequence similarity to the *GSTD3* probe. These sequences however, when translated to amino acid sequence had low similarity to the probe. The highest level of amino acid sequence similarity achieved was 32.3%.

The G3 primer pair, which was designed to target a smaller portion of the *GSTD3* gene, failed to yield amplicons when applied to *An. funestus*. Attempts to optimize salt concentration and annealing temperature failed to result in amplification.

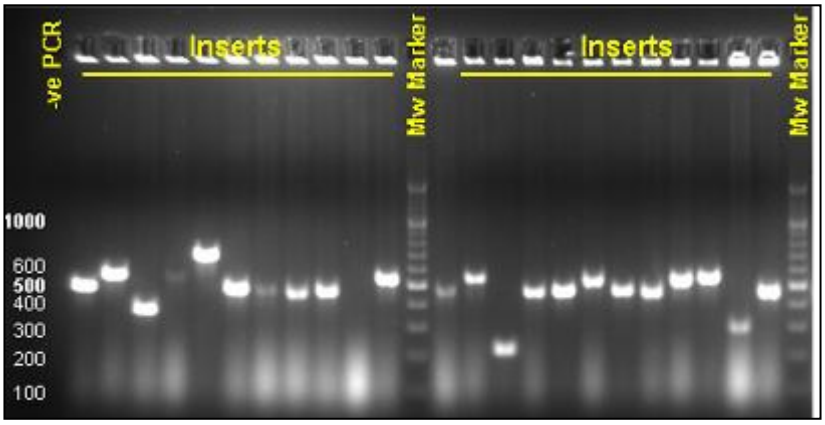


Figure 3.12 *An. funestus GSTD3* inserts cloned into pGem®-T Easy Vectors, depicting the range in size obtained. (2% TAE agarose gel, Fermentas 100bp O'Range Ruler) Screened ~160 clones.

Table 3.9 Contigs resulting from the sequenced clones, aligned to the *GSTD3* probe sequence from the *An. gambiae* detox chip. The three sequences that displayed the greatest sequences similarity have been highlighted in bold.

Contig ID	Contig Length	% Nucleotide Sequence similarity to the GSTD3 probe sequence	% Amino acid similarity to the translated GSTD3 probe sequence
Contig 1_1609	363bp	40.5%	26.9%
Contig 2_1609	385bp	**48.6%**	25.0%
Contig 3_1609	429bp	**44.8%**	27.3%
Contig 4_1609	273bp	38.5%	**32.1%**
Contig 5_1609	225bp	42.9%	**32.3%**
Contig 6_1609	307bp	42.9%	24.7%
Contig 7_1609	285bp	42.2%	**30.6%**
Contig 1_2909	377bp	43.9%	28.6%
Contig 2_2909	418bp	44.1%	21.3%
Contig 1_0110	516bp	41.8%	28.6%
Contig 2_0110	520bp	43.8%	14.8%
Contig 3_0110	513bp	**45.3%**	16.9%
Contig 4_0110	460bp	40.0%	22.4%

Quantitative Real Time PCR for CYP6P9; CYP6P13 and CYP6Gen

Real time qPCR analysis of two cytochrome P450 genes was carried out. Both *CYP6P9* and *CYP6P13* have previously been implicated in the pyrethroid based resistance in *An. funestus* FUMOZ-RH (Amenya *et al.*, 2008; Wondji *et al*., 2009; Matambo *et al*., 2010). These genes,

when analysed in real time (Figures 3.13 and 3.14) to investigate the effect of the presence of a blood meal, showed slight yet significant up-regulation when compared to the unfed cohorts.

CYP6P9 showed an induction of 5.08. ± 0.95 fold or 2.74 ± 0.70 fold when *RPS7* and *RPL19* were used respectively as reference controls. *CYP6P13* was induced 6.95 ± 0.78 fold and 5.91 ± 3.52 SD fold for the *RPS7* and *RPL19* controlled experiments respectively (Figure 3.15).

The relative expression values for the target genes vary depending on the reference gene used in the analysis. This is due to differences in suitability of these genes themselves. *RPS7* is the most suitable reference gene available for this study. However, since the MIQE guidelines (and many journals) require validation by more than one reference gene, *RPL19* was included in the analysis. Visual inspection of the amplification curves for the *RPS7* and *RPL19* reference genes (Figure 3.13 and 3.14) confirms that the variation in Cq observed for *RPL19* is far greater than that of *RPS7*. One can thus expect a greater variation in the relative expression ratios that this reference gene delivers.

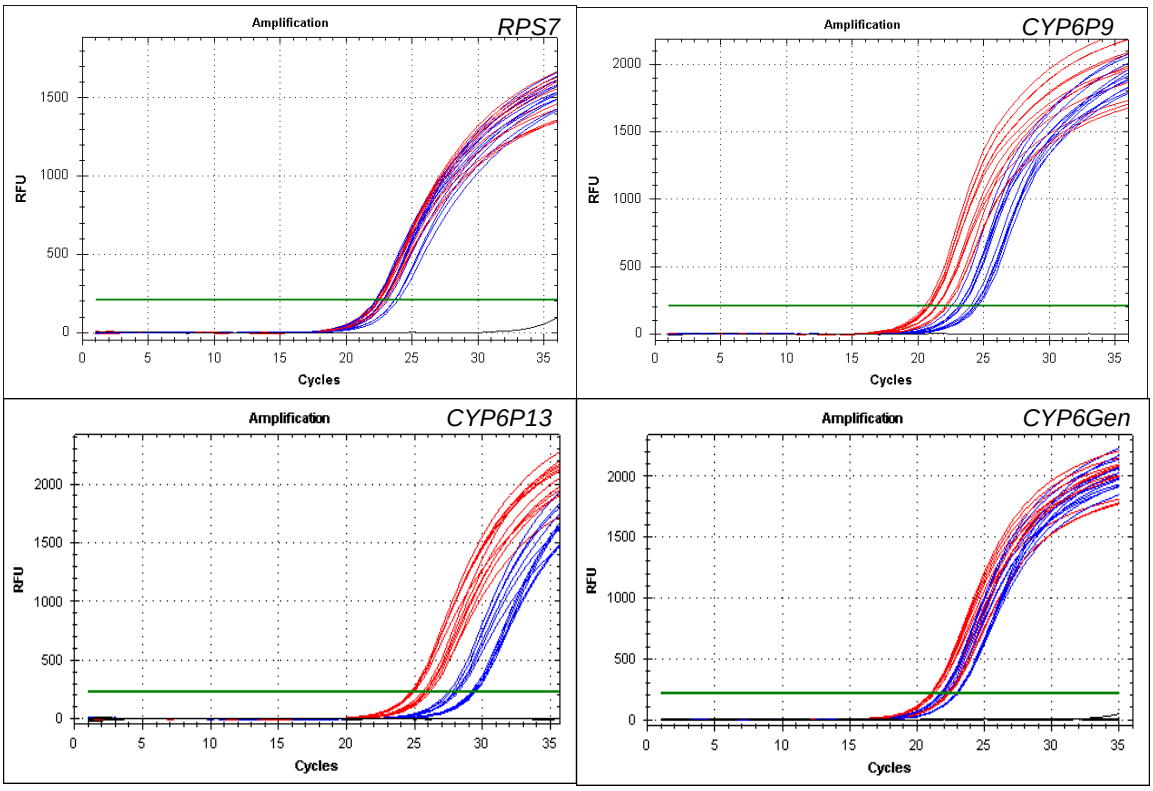


Figure 3.13 Amplification curves for *CYP6P9*, *CYP6P13* and *CYP6Gen*. The Amplification curve for the reference gene *RPS7* has been included for comparison. Red lines depict amplification curves for the blood fed samples, blue lines for the unfed samples.

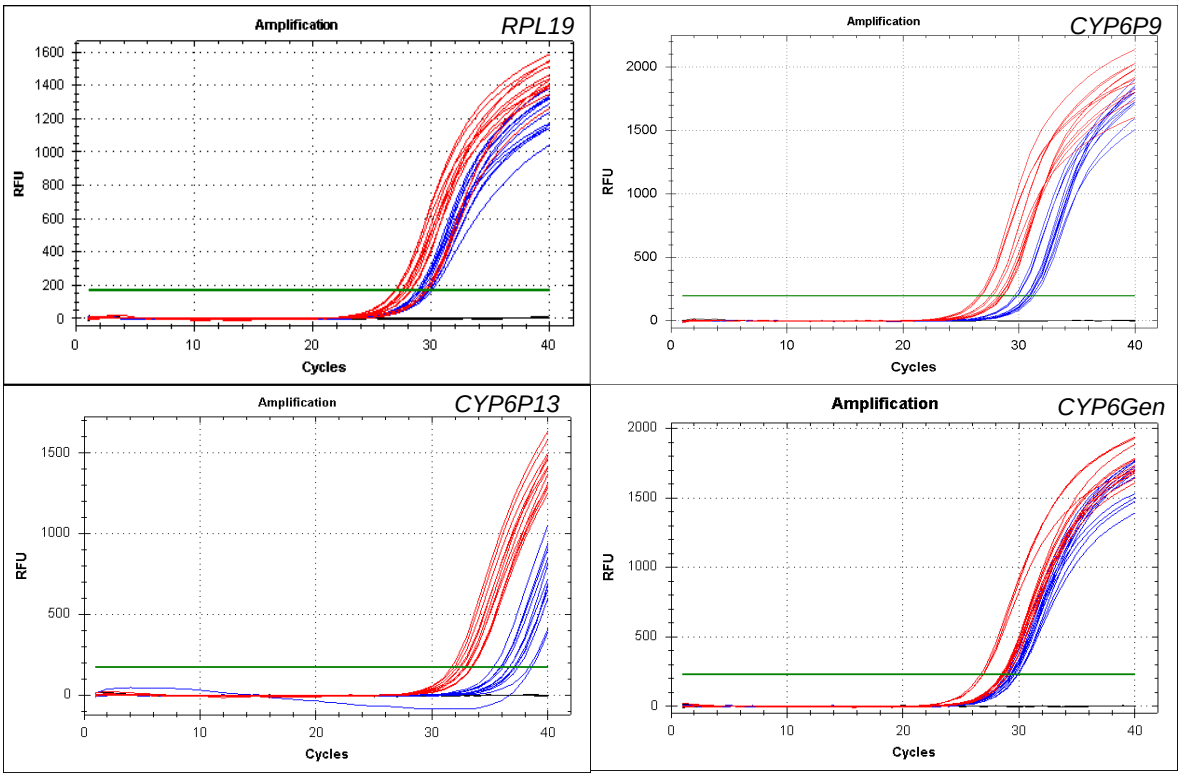


Figure 3.14 Amplification curves for *CYP6P9*, *CYP6P13* and the general *CYP6Gen*. The Amplification curve for the reference gene *RPL19* has been included for comparison. Red lines depict amplification curves for the blood fed samples, blue lines for the unfed samples.

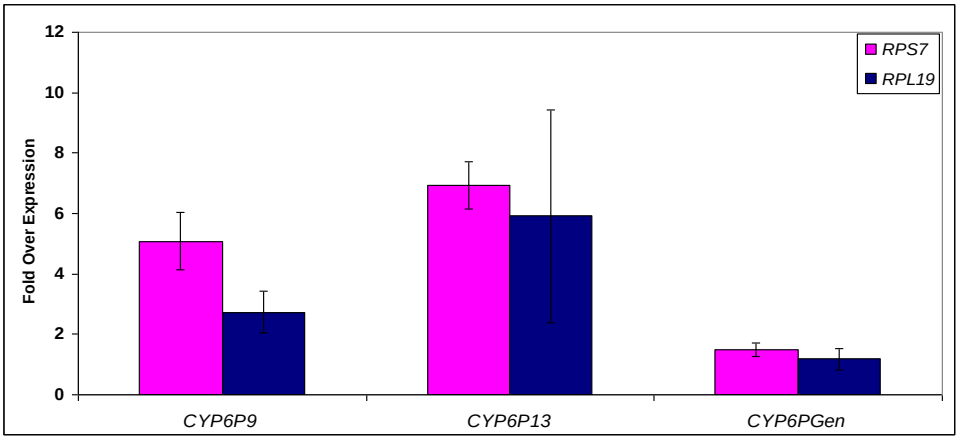


Figure 3.15 Bar chart depicting the fold over expression for *CYP6P9*, *CYP6P13* and *CYP6Gen* with both *RPS7* and *RPL19* as reference genes.

The general CYP6 targeted primer pairs showed no significant change in expression levels, with a fold change of 1.48 ± 0.22 and 1.19 ± 0.35 for the *RPS7* and *RPL19* controlled experiments respectively (Figure 3.15). In order to better understand this result, an analysis of the potential gene targets of this general CYP6 primer pair was done for *An. funestus* (Figure 3.16).

The sequence similarity within the forward primer region between *CYP6P9*, *CYP6P13* and *CYP6P5* is very high, with only one base pair dissimilarity in the *CYP6P5* sequence. The reverse primer region shows very high levels of similarity for all of the *CYP6P4* sequences, *CYP6P13* and *CYP6P9*, with only a single base dissimilarity between each of the sequences. The reverse primer regions of the *CYP6P1* and *CYP6P5* sequences have more than 3bp dissimilarity to the *CYP6P9* primer region.

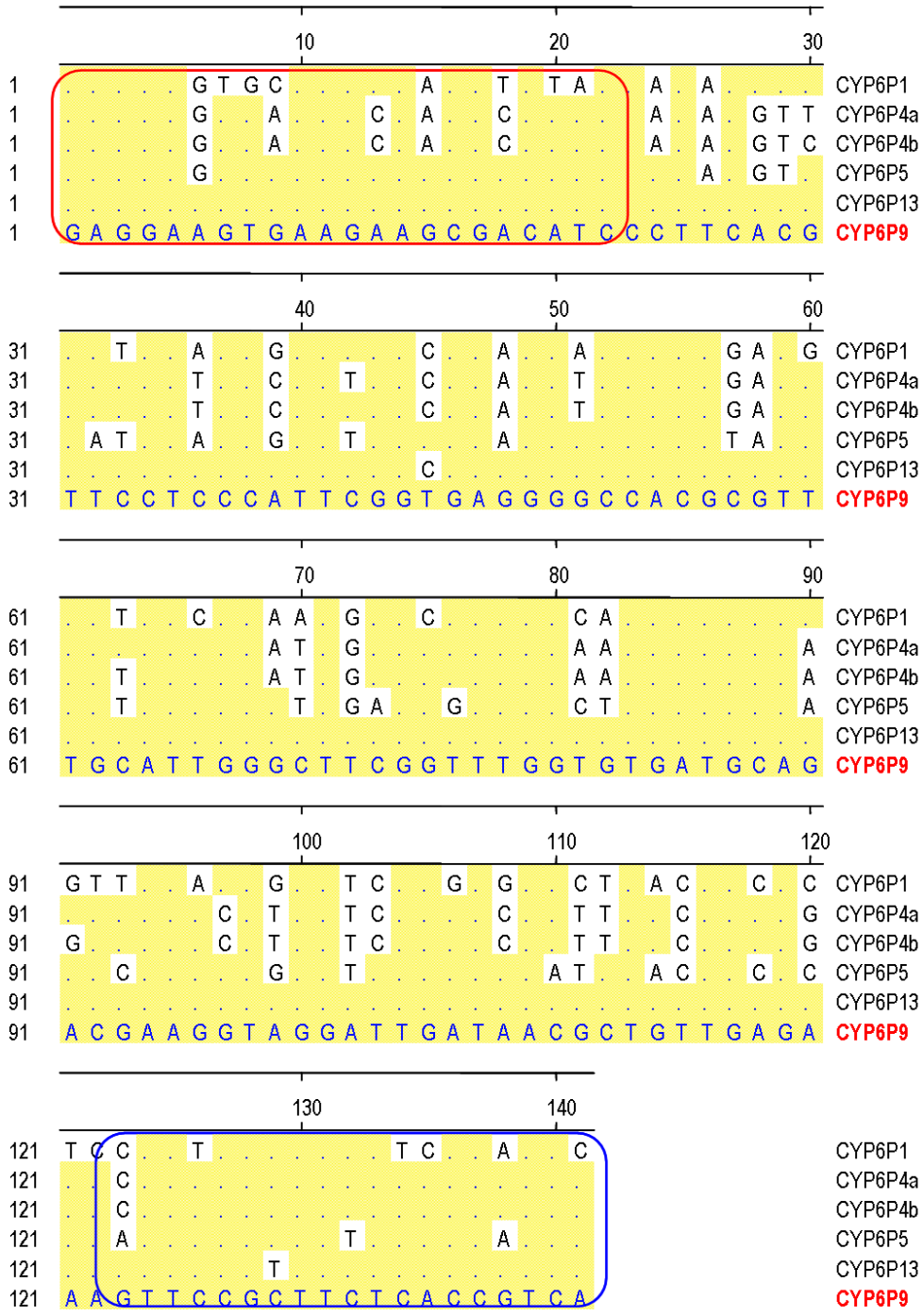


Figure 3.16 Alignment of multiple *An. funestus* CYP6 gene sequences that fall within the amplification region for the CYP6Gen primer pair. (Residues that match the *An. funestus CYP6P9* sequence are shaded in yellow and hidden. Red box denotes forward primer region; blue box denotes reverse primer region.)

3.6 DISCUSSION

Since the quality of the RNA (Vermeulen *et al*., 2011) used in gene transcription studies can affect the performance of both microarray analysis and real time qPCR, all of the RNA used in this study was assessed by formaldehyde gel electrophoresis and the microfluidics capillary system of the Agilent RNA 6000 Nano chips. In both methods the most prominent RNA fragment is the 18S rRNA. The 28S rRNA fragment is much reduced and at times absent altogether. In insects, the large 28S ribosomal subunit is made up of a combination of two smaller, 18S sized subunits (28Sα and 28Sβ linked by a weak hydrogen bond) and a smaller 5.8S subunit (Figure 3.17). When heat is applied to insect RNA the hydrogen bond is irreversibly broken and the RNA appears to have a single 18S peak profile (Winnebeck *et al.*, 2010). Since the RIN values calculated by Agilent's 2100 expert software take into consideration the presence and surface area of the 28S RNA peak, the absence of this peak in insect RNA specimens will always have a negative impact on the resulting RIN values. Currently, the gold standard for RNA integrity analysis is the Agilent chip system. Thus, researchers need to be cognizant of the fact that the resulting RIN values for mosquito RNA samples are merely an indication of quality and cannot be used as an empirical analysis tool. With this in mind, a cutoff value of RIN = 6.0 was deemed acceptable for this study.

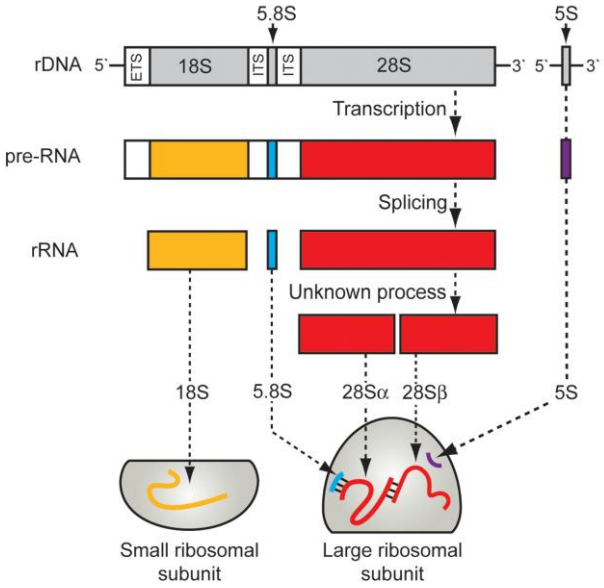


Figure 3.17. Assembly of the rRNA into ribosomal subunits in insects (Winnebeck *et al.*, 2010). Note that the 28S α and β subunits are hydrogen bonded together to form the majority of the large ribosomal subunit.

Microarray analysis of the effect of a blood meal on detoxification gene expression in *An. gambiae* highlighted potential gene targets for further investigation. The most up regulated genes were found to be four delta class GSTs and a glutaredoxin gene. The increase in glutaredoxin expression seen here has also been described in a blood meal experiment in *An. gambiae* by Felix *et al* (2010). The upregulation of GSTs in response to a blood meal has been described in a number of hematophagous arhtropods, including the sandfly *Lutzomyia longipalpis* (Jochim *et al.*, 2008), the common hard-bodied tick *Ixodes ricinus* (Rudenko *et al.*, 2005) and *An. gambiae* (Dana *et al.*, 2006). It has been suggested that GSTs may play an antioxidant role during blood meal digestion, sequestering free radicals, toxic free iron and other break down products that are released during the breakdown of blood proteins and

hemoglobin (Jochim *et al.*, 2008). The down regulation of cytochrome P450s in response to the blood meal observed in this study is in agreement with a genome-wide microarray study in *An. gambiae* carried out by Marinotti *et al* (2005).

Application of *An. funestus* cDNA to the *An. gambiae* detox chip has been optimized previously (Christian *et al.*, 2011b). In this study, the detox chip was used as an investigative tool to identify possible targets involved in blood meal associated insecticide resistance. The most up-regulated genes were a nitrilase, two cytochrome P450s (*CYP6AG2* and *CYP6M3*), the delta class glutathione S-transferase *GSTD3* and tubulin. During blood feeding, a female mosquito can take up her own body weight in blood (Gwadz R, 1969). Changes in midgut structure need to take place in order to accommodate this increase in volume. Hence, structural genes like tubulin may exhibit increased levels of expression in response to the presence of the blood meal. The only gene that was identified as being up-regulated in both the *An. gambiae* and *An. funestus* microarrays was *GSTD3*. This gene has been shown to have variable expression levels following a blood meal in *An. gambiae* (Marinotti *et al.*, 2005), with expression peaking at 3 hours post blood meal and declining thereafter. Since *GSTD3* was highlighted by both species arrays, it served as a good target for quantitative real time PCR analysis. In addition, it has previously been established that *CYP6P9* presents as one of the major effect genes for the production of pyrethroid resistance in FUMOZ-RH (Amenya *et al.*, 2008; Wondji *et al*., 2009; Matambo *et al*., 2010). *CYP6P9* and its related genes were therefore included in the real time analysis even though the probe for *CYP6P3* (the *An.*

gambiae orthologue to *An. funestus CYP6P9*) did not show any significant changes in expression levels in association with blood intake using microarrays.

Prior to initiating real time qPCR analysis of the target genes, suitable reference genes needed to be identified. Analysis of candidate reference genes showed contradicting results between the different analysis tools used. An ideal reference gene could be described as having stable expression levels across all samples and treatments used within the experiment. Visual inspection of the amplification curves allows researchers to determine the initial spread of Cq values obtained for each gene, where the most suitable gene would display the smallest spread of Cqs (such as *RPS7* and *RPL19* in these experiments). A box and whisker plot of the Cq values removes bias from affecting the visual inspection of the amplification curves and from this analysis it was evident that *RPS7*; *RPL19* and *ND5* all held good potential as reference genes.

geNorm is the only program that confirmed the visual inspection of the amplification curves, confirming suitability of *RPS7* and *RPL19* as reference genes. The Normfinder program gave very conflicting results. Initially, when analysing the data without the use of identifiers, the program identified *RPL19* and *RPS26* as the two best candidate reference genes. When repeating the analysis with the inclusion of identifiers, Normfinder suggested that *GAPDH* and *RPL19* were the two most suitable genes to use. The *GAPDH* gene gave a spread of approximately 6 Cq values across the samples. According to the defining quality of a reference gene mentioned above, this range of Cq would be considered far too variable to be

acceptable. It was hence decided that the results of this analysis could not be taken into consideration when choosing the reference genes.

Although the Bestkeeper program identified *RPL19*; *RPS7* and *ND5* as the three most suitable reference genes, the *p*-values for significance would exclude both *ND5* and *RPS7*. Between all three programs the two candidate genes that were most commonly identified as potential reference genes were *RPL19* and *RPS7*. Based on these results *RPL19* and *RPS7* were chosen as the reference genes for the real time qPCR experiments.

Having identified suitable reference genes for the real time qPCR analysis it was then necessary to design and optimize primers targeted at the *GSTD3* gene identified by microarray analysis. Two sets of primers were designed for this region. The GSTD3 primer pair was designed to fall just within the start and stop codons of the gene whereas the G3 primer pair was designed to span the intron, with the forward and reverse primers each falling within the coding exons of the gene. The primers were tested on *An. gambiae* genomic DNA and the resulting amplicons were sequenced thereby confirming specificity for the *GSTD3* gene region.

The GSTD3 specific primers were then applied to *An. funestus* DNA. The reaction failed to yield one specific amplicon, instead giving a 'ladder' of amplicons for each reaction. Salt concentration in the PCR mix and annealing temperature in the cycling protocol were optimized. However, the number of products could not be reduced to a single amplicon. In

order to determine if any of the fragments were possibly the *GSTD3* target, the fragments were cloned and sequenced. Analysis of sequence similarity to the GSTD3 probe sequence from the detox chip was done for both the nucleic acid and translated amino acid sequences. The low level of similarity (48.6% at nucleic acid level; 32.1% at amino acid level) of the amplicons to the GSTD3 probe sequence suggests that these primer pairs are not specific enough to target *GSTD3* in *An. funestus*. Gregory *et al*. (2011) showed that the percentage similarity for a small number of detoxification genes compared between *An. funestus* ESTs and orthologues from *An. gambiae* displayed high levels (>73%) of similarity at the amino acid level. This finding further suggests that the *GSTD3* targeted primers used in this study are not specific enough for *An. funestus*.

It is possible that due to the lowered levels of stringency of the microarrays for the *An. funestus* study, the signal achieved for the GSTD3 probe could be due to false, non-specific binding. Since the full genome of *An. funestus* has yet to be sequenced and annotated, it cannot be ruled out that the level of sequence variation between *An. gambiae* and *An. funestus* GSTs may be greater than first assumed. Even though the possibility of a false signal exists for the *An. funestus* study, the signal achieved in the *An. gambiae* study is likely to be a true reflection since the probe specificities and stringency of the experiment are far greater. The up-regulation of *GSTD3* in response to a blood meal in the *An. gambiae* study lends credence to the microarray results for this gene in the *An. funestus* study.

Quantitative real time analysis of the *CYP6P9* and *CYP6P13* genes revealed slight, yet significantly elevated levels of expression in the blood fed cohorts of *An. funestus.* The *CYP6Gen* expression level remained unchanged in response to the blood meal. Since this primer pair may target multiple Cyp6 sequences, it is possible that any great increase in expression of a particular Cyp6 may have been masked by a significant down regulation of a different Cyp6. The CYP6Gen primer pair can only be used to gain a general overview of Cyp6 expression levels and its application is thus very limited.

Since it is established that the *CYP6P9* and *CYP6P13* genes are already up-regulated in pyrethroid resistant *An. funestus* (Amenya *et al*., 2008; Matambo *et al*., 2010), further elevation after a blood meal could be expected to increase an individual's insecticide detoxification capability during subsequent exposure to insecticide. The concomitant increase in GST activity in association with blood feeding could play one of two roles. Firstly, GSTs have been shown to have antioxidant properties and function in response to oxidative stress. The presence of toxic free iron and other breakdown products released during blood meal digestion could induce oxidative stress pathways in the mosquito (Kumar *et al.,* 2003). Since some GSTs have peroxidase activity (Vontas *et al*., 2001), an increase in their expression may alleviate the oxidative stress associated with the digestion of the blood meal. Secondly, GSTs have been tentatively associated with pyrethroid resistance owing to the sequestering of insecticide in a non-catalytic manner (Kostaropoulos *et al.,* 2001). Sequestration is basically the process of binding insecticide for subsequent metabolic degradation and excretion before the insecticide is able to interact with its target.

GSTD3, which was upregulated in association with blood feeding in *An. funestus* and *An. gambiae*, may either function as an antioxidant or as an insecticide sequester, or perhaps both. Vontas *et al.* (2001) suggest that GSTs may assist with resistance to pyrethroids by alleviating the effects of oxidative stress induced by pyrethroid exposure. This is because the lethal effect of pyrethroid induced lipid peroxides may be abrogated by the activity of GSTs.

3.7 CONCLUSION

Acquiring and digesting a blood meal likely involves a suite of physiological adjustments in female mosquitoes. The ability of blood fed, pyrethroid resistant *An. funestus* to tolerate increased doses of insecticide has been phenotypically demonstrated previously (see chapter 2). Here it is shown that the effect of a blood meal at the molecular level causes increased expression of the *CYP6P9* and *CYP6P13* detoxification genes. This increase in gene expression, albeit small, is postulated to enhance the detoxifying capability of those *An. funestus* females already carrying the resistance haplotype, enabling them to tolerate higher doses of pyrethroid insecticide than unfed females. Although the functional activity of GSTD3 in the context of pyrethroid resistance is unknown, it is postulated that it may play a supportive role by alleviating the effects of oxidative stress in blood fed, pyrethroid exposed mosquitoes.

3.8 Christian RN, Matambo TS, Spillings BL, Brooke BD, Coetzee M and Koekemoer LL. 2011. Age-related pyrethroid resistance is not a major function of P450 gene expression in the major African malaria vector, *Anopheles funestus* (Diptera: Culicidae). *Genetics and Molecular Research* **10**: 3220 – 3229.

Contribution to publication:

Belinda Spillings carried the *CYP6P9* and *CYP6P13* sequence analysis and design of the *CYP6P9* primers for qPCR. The protocol used for the analysis of the *CYP6P9* expression was also optimised.

CHAPTER 4 – *ANOPHELES FUNESTUS-LIKE*: THE DISCOVERY OF A NEW MEMBER SPECIES OF THE *ANOPHELES FUNESTUS* GROUP

4.1 INTRODUCTION

The *Anopheles funestus* group consists of nine African species (Gillies and De Meillon, 1968; Gillies and Coetzee, 1987; Harbach, 2004), five of which belong to the Funestus subgroup (Harbach, 2004; Garros *et al*., 2005). Phylogenetic analyses suggest that *An. funestus*, *An. vaneedeni* Gillies and Coetzee, *An. parensis* Gillies, *An. aruni* Sobti and *An. confusus* Evans & Leeson be grouped within the Funestus subgroup (Harbach, 2004; Garros *et al.* 2005). The remaining four members (*An. rivulorum* Leeson, *An. brucei* Service, *An. fuscivenosus* Leeson and the *An. rivulorum*-like species (Couhet *et al*. 2003)) have been placed in the Rivulorum subgroup.

Although the adult stage of *An. leesoni* Evans is similar in morphology to the *An. funestus* subgroup members, it has very distinct eggs and larvae (Gillies and De Meillon, 1968). Cytogenetic (Green, 1982) and molecular (Garros *et al*., 2005) evidence suggest that this species is more closely related to the Asian *An. minimus* subgroup even though geographically, it is an African species. These findings have resulted in *An. leesoni* being removed from the *An. funestus* subgroup (Gillies and De Meillon, 1968; Gillies and Coetzee, 1987) and placed as a member of the *An. minimus* subgroup (Harbach, 2004).

All of the members of the *An. funestus* subgroup are morphologically similar at the adult stage and accurate identification often requires the availability of material at multiple life stages (egg, larva or adult). Although they may be similar in morphology, their efficiencies as malaria vectors vary greatly. Due to its highly anthropophilic nature and its tendency to rest indoors, *An. funestus s.s.* is one of the most successful vectors of malaria in sub-Saharan Africa (Gillies and De Meillon, 1968). *Anopheles rivulorum* has only once been implicated in malaria transmission in Tanzania (Wilkes *et al*., 1996) but this species generally elects to blood feed on domestic animals rather than humans. The remaining members of the *An. funestus* group have never been shown to be malaria vectors in nature (Gillies and De Meillon, 1968; Gillies and Coetzee, 1987), although *An. vaneedeni* has been infected experimentally in the laboratory (De Meillon *et al*., 1977).

Owing to differences in vector capacity, biting and resting behaviours as well as the very close morphological similarity of the members of the *An. funestus* group, accurate identification of field caught material is critical for vector control programmes. Early identification methods relied solely on morphological characters which detailed minor differences between the members of the group (Evans and Symes, 1937; Evans, 1938; De Meillon 1947; Gillies and de Meillon 1968; Gillies and Coetzee 1987). This process of identification relied on the availability of multiple life stages and required a high level of expertise. A further drawback of this method of identification is that the morphological characters used for the species identifications overlap between species.

The success of using polytene chromosome morphology and banding sequences from fourth instar larvae and adult females to distinguish between members of the *Anopheles gambiae* complex was a remarkable advance in determining the specific status of sibling species (Coluzzi and Sabatini, 1967). Green and Hunt (1980) demonstrated that such cytogenetic analysis can also be used to distinguish *An. parensis* from *An. funestus*. However, *An. vaneedeni* (formerly *An. aruni?*) has homosequential chromosomal banding patterns with *An. funestus.* Cross mating studies between *An. vaneedeni* and *An. funestus* produced sterile male hybrids and asynapsis in hybrid polytene chromosomes (Green and Hunt, 1980), thereby confirming the specific status of *An. vaneedeni* which lead to it being formally named by Gillies and Coetzee (1987). Green (1982) published chromosome maps for three more members of the *An. funestus* group: *An. rivulorum*, *An. leesoni* and *An. fuscivenosus*. More recent cytogenetic studies of West African *An. funestus* have shown evidence of genetic differentiation within populations, suggesting that this taxon may consist of a complex of cryptic species (Lochouarn *et al.*, 1998; Costantini *et al*., 1999; Michel *et al*, 2005; Guelbeogo *et al*, 2005).

Although cytogenetic maps have been published for most of the members of the *An. funestus* group (Green, 1982), this method of species identification requires considerable skill in order to interpret variation in the banding patterns on the chromosomes. Furthermore, polytene chromosomes are only found in specific mosquito tissues such as larval malpighian tubules and salivary glands or in the ovarian nurse cells of adult half gravid females. The need for simpler identification methods drove the development of molecular based approaches. One of

the first molecular techniques was a PCR-RFLP method designed to distinguish between *An. funestus* and *An. vaneedeni* (Koekemoer *et al.*, 1998). Although this method was successful, it only targeted two members of the Funestus subgroup. Shortly after this, a PCR-SSCP method was developed that could distinguish four of the member species (Koekemoer *et al.*, 1999). In 2002 the development of a multiplex PCR assay targeting the variable ITS2 regions allowed for the identification of the five most prevalent species in the *An. funestus* group (Koekemoer *et al*., 2002). There is no doubt that the relative simplicity of these DNA based assays has resulted in the rapid development of our knowledge of the *An. funestus* group and has aided in the discovery of new species within this group.

Recent discoveries of cryptic and incipient species have been aided by the development of molecular tools, specifically ITS2 sequence analyses. Hackett *et al.* (2000) utilised ITS2 sequence divergence to reveal differences between the southeastern African *An. rivulorum* populations from those that occur in West and Central Africa. Failure of the species-specific multiplex PCR (Koekemoer *et al*., 2002) to detect *An. rivulorum* from Cameroon (Couhet *et al*., 2003) led to investigations into this population of mosquitoes. The levels of sequence variation between the *An. rivulorum* populations are significant and suggest that the West African population is a new species (Couhet *et al*., 2003) that has been provisionally named *An. rivulorum*-like.

Although thousands of new plant and animal species are described annually worldwide, the description of new species within the genus *Anopheles* is extremely rare. Intense attention has

been paid to this group due to its importance in malaria transmission. This chapter will describe the discovery of yet another new species of anopheline, closely related to the Funestus subgroup.

4.2 RATIONALE

It is shown in Chapter 2 that an increase in pyrethroid resistance, in a strain of *An. funestus* exhibiting a resistance phenotype, can be linked to the presence of a blood meal. It has further been shown that the presence of a blood meal exerts an effect at the transcriptional level within these mosquitoes, causing a small but significant elevation in *CYP6P9* and *CYP6P13* transcription.

Since the effects described above have been studied in laboratory reared strains of anophelines, it was decided to determine whether this phenomenon was seen in the progeny of wild caught *An. funestus* specimens. Although this objective was not completed owing to difficulties in identifying the wild-caught specimens, it remains as a future project. This chapter details the discovery of a new species of anopheline, belonging to the Funestus subgroup.

4.3 AIMS AND OBJECTIVES

The initial aim of this study was to investigate the effects of blood feeding on insecticide tolerance using the progeny of wild caught *An. funestus*. This necessitated the collection of wild *An. funestus* specimens and their transport to the NICD laboratories for analysis. However, the field material could not be identified as *An. funestus sensu stricto* resulting in the following aims:

- To identify the field caught *An. funestus* group specimens
- To investigate nucleotide sequence similarity/divergence in the wild caught specimens compared to *An. funestus* laboratory colony material
- To assess the possibility that the material collected may be a new species of anopheline, closely related to *An. funestus*

4.4 MATERIALS AND METHODS

4.4.1 Study site and collection method. Field collections of anopheline mosquitoes were conducted in the rural villages of Karonga, in the Northern District of Malawi (S 10° 18.627' E 34° 07.901'), (Figure 4.1). The villages are all in close proximity to the shores of Lake Malawi, where the predominant activities are fishing and maize cultivation. Rice paddies were also visible from the main road in the area surrounding Karonga (Figure 4.2A).
Malawi has seasonal rainfall, starting November and ending in April. Mosquito collections were carried out in December 2007. It was hoped that the rainy season would have started

prior to our arrival. However, Malawi experienced late rains during this period. The conditions were very arid and hot (Figure 4.2B).

Figure 4.1. The field collections in northern Malawi were carried out in the villages surrounding Karonga - marked by *.

Successful catches were predominantly daytime, indoor resting catches with the exception of samples that were collected from abandoned tyres placed at the entrance to a reed hut (Figure 4.3). All adults collected were identified using the morphological keys of Gillies and De Meillon (1968) and Gillies and Coetzee (1987). *Anopheles funestus* group males were dry-preserved on silica gel and the females transported live, with access to 10% sucrose solution, to the insectaries of the Vector Control Reference Unit, NICD, (Johannesburg, South Africa).

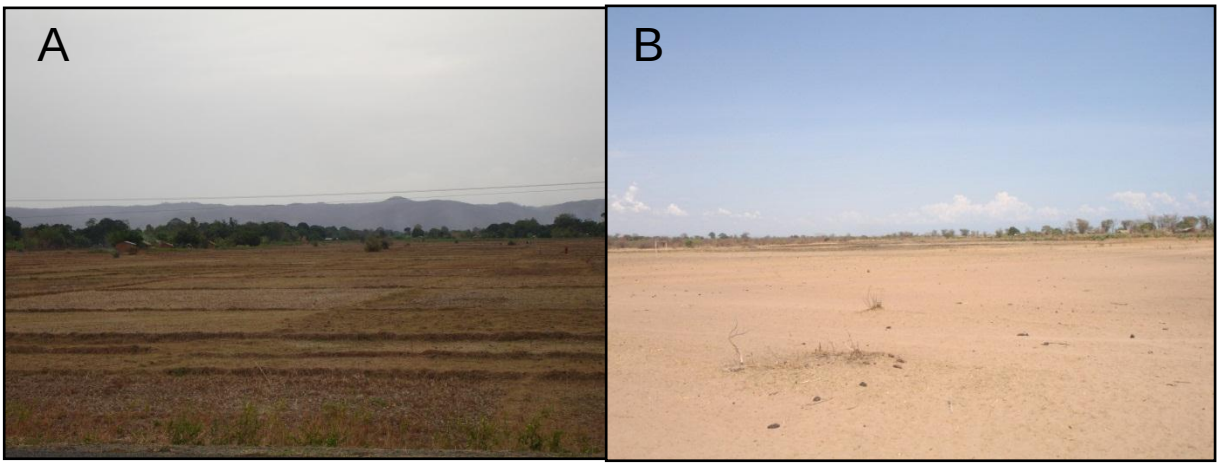


Figure 4.2. A. Dry rice paddies alongside the main road leading into Karonga town from the South. B. Karonga homestead area, approximately 20km South of the town, December 2007.

Figure 4.3. A successful collection point at a reed hut. Abandoned tyres were stacked at the entrance and proved to be very productive.

4.4.2 Laboratory rearing of wild caught specimens. A total of 63 females morphologically identified as belonging to the *An. funestus* group were placed in glass vials for egg laying. The females were offered a blood meal every alternate day. When not being fed they had constant access to 10% sucrose. Egg morphology confirmed the samples belonged to the Funestus subgroup (Gillies and De Meillon, 1968). Isofemale lines from each egg batch were reared to

adulthood and used for molecular identifications, WHO susceptibility assays and cross-mating experiments. Females that survived egg laying were offered another blood meal and at the half gravid stage their ovaries were dissected for cytogenetic analysis of their polytene chromosomes.

4.4.3 WHO Susceptibility assays. A mix of F1 progeny from the isofemale lines were used to determine the presence of insecticide resistance using the standard World Health Organization (WHO) Bioassay test. Briefly, approximately 20 to 25 adults (2-3 days old, mixed sex) were placed in holding tubes and allowed to rest for one hour. After this the adults were blown through to an exposure tube containing an insecticide treated paper. Adults were exposed for one hour, after which they were blown through to the holding tube. The adults were allowed to rest for 24 hours with access to 10% sugar solution. The percentage mortality was scored for each insecticide tested: 0.75% permethrin; 4% DDT; 5% malathion and 0.1% bendiocarb. Each exposure was run in parallel with an unexposed control.

4.4.4 Species identification of field material. One specimen from each isofemale line was sacrificed and DNA was extracted (see Appendix E for full protocol) using the salt precipitation method of Collins *et al*. (1987). The multiplex PCR for the identification of *An. funestus*, *An. rivulorum*, *An. vaneedeni*, *An. parensis* and *An. leesoni* (Koekemoer *et al*., 2002) was performed on all the DNA samples. A negative extraction control, a no DNA-template control and two positive controls were included in the PCR. The positive controls were *An. funestus* from colony material and *An. leesoni* from a previous field collection. Each of the samples was tested 2 to 3 times to ensure accuracy of the PCR identification.

4.4.5 *ITS1, ITS2* and *D3* Sequencing for unidentified Malawi samples. From four of the unidentified Malawian specimens (MalaF) and an *An. funestus s.s.* control, a partial region of the *ITS2* gene was amplified. The primers used were: ITS2A (5'- TGT GAA CTG CAG GAC ACA T- 3') and ITS2B (5' -TAT GCT TAA ATT CAG GGG GT- 3') (Koekemoer *et al*., 2002). The PCR reaction was carried out in a volume of 25-µL and was adapted from Hackett *et al*. (2000). The reaction contained: 50pmol of each primer, 1.5mM $MgCl_2$, 200µM of each dNTP, 1.25 units of *Taq* DNA polymerase and 1µl of extracted DNA. Both annealing temperature and $MgCl_2$ gradients were attempted in an effort to reduce the non-specific amplicon, which ultimately could not be completely eliminated. The thermal cycling conditions are detailed in Table 4.1 below. The resulting PCR products were electrophoresed on a 1.8% low melting temperature Tris-acetate-EDTA (TAE) agarose gel stained with ethidium bromide (0.3 $\mu g.ml^{-1}$).

The ~850bp fragment from each sample was excised from the gel and cleaned using the Qiaquick® Gel Extraction kit (Qiagen; Valencia, CA cat 28704). Direct sequencing was carried out by Inqaba Biotechnical Industries (Pretoria, South Africa), using both the ITS2A and ITS2B primers. Sequence alignment and analysis was carried out using DNASTAR® (Lasergene version 6; Wisconsin, USA). Sequence comparisons were carried out by creating a consensus sequence for the four MalaF samples and aligning it to the *An. funestus s.s.* control sequence.

Table 4.1. Thermal cycling conditions for the *ITS2*, *D3,* Untargeted *An. funestus* group and *ITS1* PCRs.

***ITS2* PCR:**

Initial Denaturation	94°C	2 minutes
Cycle 40x	94°C	30 seconds
	50°C	30 seconds
	72°C	40 seconds
Final Extension	72°C	10 minutes

***D3* PCR:**

Initial Denaturation	94°C	3 minutes
Cycle 30x	94°C	30 seconds
	63°C	40 seconds
	72°C	40 seconds
Final Extension	72°C	10 minutes

Un-targeted *An. funestus* group PCR:

Initial Denaturation	94°C	2 minutes
Cycle 35x	94°C	30 seconds
	45°C	30 seconds
	72°C	30 seconds
Final Extension	72°C	10 minutes

***ITS1* Touchdown PCR:**

Initial Denaturation	94°C	2 minutes
Touch Down Cycles:	94°C	30 seconds
Cycle 20x	65°C - 55°C	30 seconds
-1°C every 2 cycles	72°C	40 seconds
Cycle 15x	94°C	30 seconds
	55°C	30 seconds
	72°C	40 seconds
Final Extension	72°C	10 minutes

The same four MalaF samples used above were used for the Domain 3 (D3) sequence analysis. The primers D3A (5' - GAC CCG TCT TGA AAC ACG GA - 3') and D3B (5' – TCG GAA GGA ACC AGC TAC TA – 3') (Koekemoer *et al.*, 1999) were used to amplify a region approximately 400bp in length. The 25 µL reaction contained: 25 pmol of each primer, 1.5 mM $MgCl_2$, 200 µM each dNTP, 2 units of *Taq* DNA polymerase and 1µl DNA. The thermal cycling conditions are detailed in table 3.1 above. In an attempt to minimise non-specific amplification, a temperature gradient between 40 to 63°C for the annealing temperature was carried out. The non-specific product could not be reduced completely and hence a final annealing temperature of 63°C was used for subsequent PCRs. All the *D3* amplicons were electrophoresed and sequenced as for the *ITS2* amplicons. Sequencing was carried out for both strands, using the D3A and D3B primers. Sequence alignment and analysis was carried out using DNASTAR® (Lasergene v6; Wisconsin, USA). A consensus sequence was created for the four MalaF samples and this was then aligned to the *An. funestus s.s.* control sequence for sequence analysis.

Amplification of the *ITS1* region was attempted using the following primers: 18SFor (5' – CCT TTG TAC ACA CCG CCC GT – 3') and 5.8sRev (5' – CAT GTG TCC TGC AGT TCA CA – 3') adapted from Sharpe *et al.* (2000). The 25 µL reaction contained: 25 pmol of each primer, 250 µM each dNTP, 0.5 units of *Taq* DNA polymerase and 1µl DNA. Optimizations of the annealing temperature using a temperature gradient (40˚C to 55˚C) and salt concentrations (1.5mM to 3mM) were carried out. In an effort to reduce the number of non-

specific products, touchdown PCRs in combination with varying salt concentrations were attempted. The cycling conditions for the touchdown PCR are detailed in Table 4.1.

4.4.6 Design and application of the MalaF specific PCR primer. Primer annealing sites, specific to the MalaF samples, were identified within the *ITS*2 gene region. MalaF specific primers were designed to yield an amplicon with a different size to the amplicon yielded for *An. funestus s.s.* in the species-specific PCR of Koekemoer *et al*. (2002). These primers were designed manually and analysed for melting temperature, GC content, self-complementarity and 3' stability using the primer check function in Primer3Plus (Untergasser *et al*., 2007). The two potential primers (MalaFA: 5' – CCT GCG TCC CAA GGT T- 3'; MalaFB: (5' –GTT TTC AAT TGA ATT CAC CAT T- 3') were individually tested for their efficiency in the species-specific PCR for the identification of *An. funestus* group members. Each of the newly designed primers was included in the reaction mixture and the products run on a 3% Tris-Borate-EDTA (TBE) agarose gel stained with ethidium bromide (0.3 $\mu g.ml^{-1}$). All 61 of the unidentifiable *An. funestus*-like samples were then tested with the MalaFB primer.

Samples that failed to amplify with the new MalaFB primer were confirmed for the presence of nucleic acids using a Nanodrop spectrophotometer (Appendix F; NanoDrop Technologies Inc., Wilmington, DE). These samples were then subjected to an in-house PCR to detect the presence of DNA. This PCR was designed to detect members of the *An. funestus* group, but not to identify them to species level. The primers used were: UF (5' - TGT GAA CTG CAG GAC ACA T -3') and LRev (5' – CCA AGC ACG TTG ATC CAG TAT TAC - 3')

(Koekemoer *et al*., 2008). The 25-µl PCR reaction mixture contained: 6.6pmol of each primer, 1.5mM $MgCl_2$, 200µM each dNTP, 1 unit of *Taq* DNA polymerase and 1µl DNA. The cycling conditions are listed in Table 4.1 and the PCR is referred to as the Un-targeted *An. funestus* group PCR. The resulting amplicons were electrophoresed on a 2.5% TAE gel stained with ethidium bromide (0.3 $\mu g.ml^{-1}$). The presence of a ~440bp fragment confirmed the presence of DNA for these samples.

4.4.7 ITS2 Secondary Sequence Structure. The ITS2 secondary structure of the MalaF specimens and *Anopheles funestus* group species were generated using the mfold program (Zuker *et al*., 2003). This program can be run through the UNAfold web server (http://mfold.rna.albany.edu/) or downloaded and run directly on a Linux platform. Mfold calculates the free energies for the predicted RNA structures (Mathews *et al.*, 1999) which allows for comparisons between species. Initially, all sequences are trimmed to the same start codon as this could affect the folding of the structure. Sequences were entered individually into the program and with default constraints as the settings for analysis. The best structure, with the lowest deltaG, was chosen and used for comparison between species.

4.4.8 Cytogenetics, cross-mating studies and ELISA for sporozoite detection:

The cytogenetics and cross-mating experiments were carried out by Professor Hunt and Dr Brooke respectively. Hybrid status of the F1 progeny was tested using the *An. funestus* specific and the universal forward primer of Koekemoer *et al*. (2002) in combination with the MalaFB primer designed above. ELISA (Wirtz *et al.*,1987) was kindly carried out by Mr Vezegneho. Details of the methodology used are outlined in Spillings *et al*. (2009).

4.5 RESULTS

Species Identification and WHO Bioassays

The multiplex PCR assay (Koekemoer *et al*., 2002) for the identification members of the *Anopheles funestus* group was used to molecularly identify the wild caught specimens. Of the 63 females that were brought back to the laboratory, only two could be positively identified. They were found to be *An. rivulorum.* The remaining samples failed to amplify even though the PCR was repeated 2 to 3 times (Figure 4.4), hence sequence analysis was carried out. Samples of F1 males and females from 17 of the isolines were pooled together and used for the WHO susceptibility assays. All of the F1 adults were found to be fully susceptible to the four insecticides tested (Table 4.2).

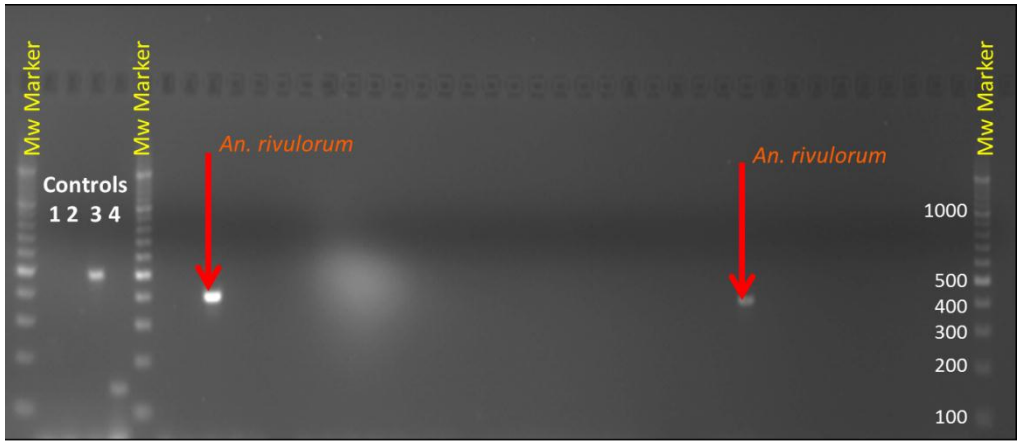


Figure 4.4. PCR identification of the Malawi field samples using the multiplex PCR assay of Koekemoer *et al*. (2002). Control 1, no DNA template; Control 2, negative extraction control; Control 3, *An. funestus* positive control, Control 4, *An. leesoni* positive control.

Table 4.2. Insecticide susceptibility status of the F1 adults (males and females) of the wild caught *An. funestus*–like adults. Positive controls of insecticide susceptible *An. funestus* exhibited 100% mortality on all the insecticide papers. Negative controls of the F1 adults showed 100% survival on untreated papers.

Insecticide	n	♀ : ♂	% Mortality
0.75% Permethrin	132	69 : 63	100%
4% DDT	105	57 : 48	100%
5% Malathion	131	66 : 65	100%
0.1% Bendiocarb	101	60 : 41	100%

Sequence Analysis

The *ITS2* and *D3* gene regions of four Malawian specimens were amplified and sequenced. Both the *ITS2* and *D3* PCRs were optimised to reduce the non-specific amplification. However small non-specific bands could not be eliminated (Figure 4.5A). Gel extraction was necessary to prepare the targets for sequencing to ensure that only a single fragment was sequenced (Figure 4.5B). The *ITS1* amplification was unsuccessful as the number of non-specific products could not be reduced to an acceptable level (Figure 4.5C). Attempts were made to gel extract specific bands for sequencing, however background contamination from the non-specific bands was always present.

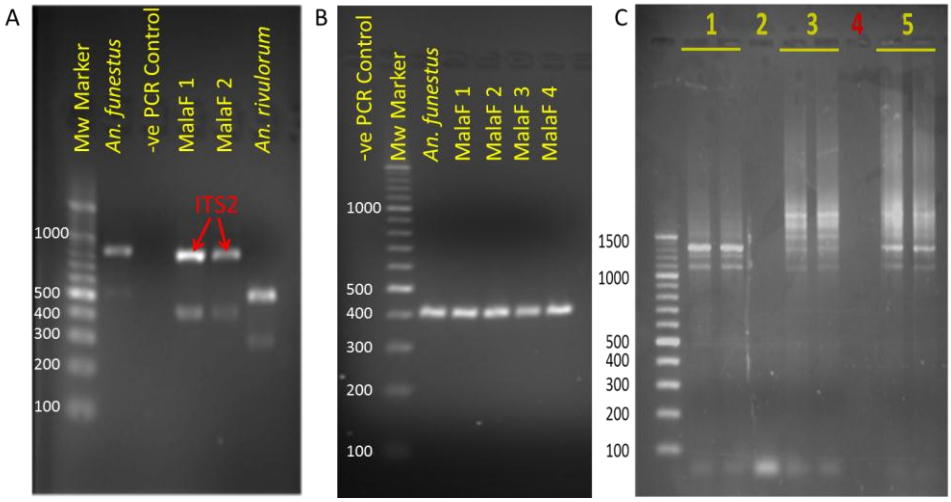


Figure 4.5. A. *ITS2* PCR of MalaF specimens with *An. funestus* and *An. rivulorum* as positive controls, prior to gel extraction. Non-specific amplification can be seen, with these bands being approximately half the length of the target *ITS2* fragments. B. *D3* PCR amplicons of MalaF specimens, with *An. funestus* as a positive control, after gel extraction. C. *ITS1* Touchdown PCR using *An. funestus* genomic DNA and 1.5 mM $MgCl_2$. The combined lanes 1, 3 and 5 contain PCR products obtained from amplifications using different specimens of *An. funestus*. Non-specific amplification could not be reduced further than what is seen above.

Sequencing of the 740bp *ITS2* gene regions of the MalaF specimens and the standard *An. funestus* colony allowed for comparison between the two sequences (Figure 4.6). Across the 740bp fragment, the MalaF specimens varied by 4.5% (33/740bp). These changes include a 3 bp insertion; a 4 bp deletion; 10 transitions and 16 transversions. It was also noted that the *An. funestus* specific primer binding region on the MalaF sequence had three deletions and a T to C transition.

Two MalaF specific primers were designed to anneal to the most variable regions in the MalaF *ITS2* sequence (Figure 4.6). These primers (MalaFA and MalaFB) were each tested on

extracted MalaF DNA in order to assess their suitability as a species specific primer. The MalaFA primer was inconsistent and did not always result in an amplicon even though the same MalaF template DNA was used. The second primer, MalaFB, displayed consistent amplification on the same, as well as different, MalaF template DNAs. The inclusion of the MalaFB primer into the *An. funestus* group identification cocktail (Koekemoer *et al*., 2002) resulted in good amplicon yield, with a product of 390 basepairs (Figure 4.7A). All of the 61 unidentifiable specimens were tested using the MalaFB primer and amplicons were obtained for 54 (88.5%) of the samples. The remaining seven specimens failed to amplify even after repeated PCR attempts. The presence of DNA in these samples was confirmed spectrophotomically (Appendix F) using a NanoDrop (NanoDrop Technologies Inc., Wilmington, DE) as well as by PCR using the untargeted *An. funestus* group primers (Figure 4.7B).

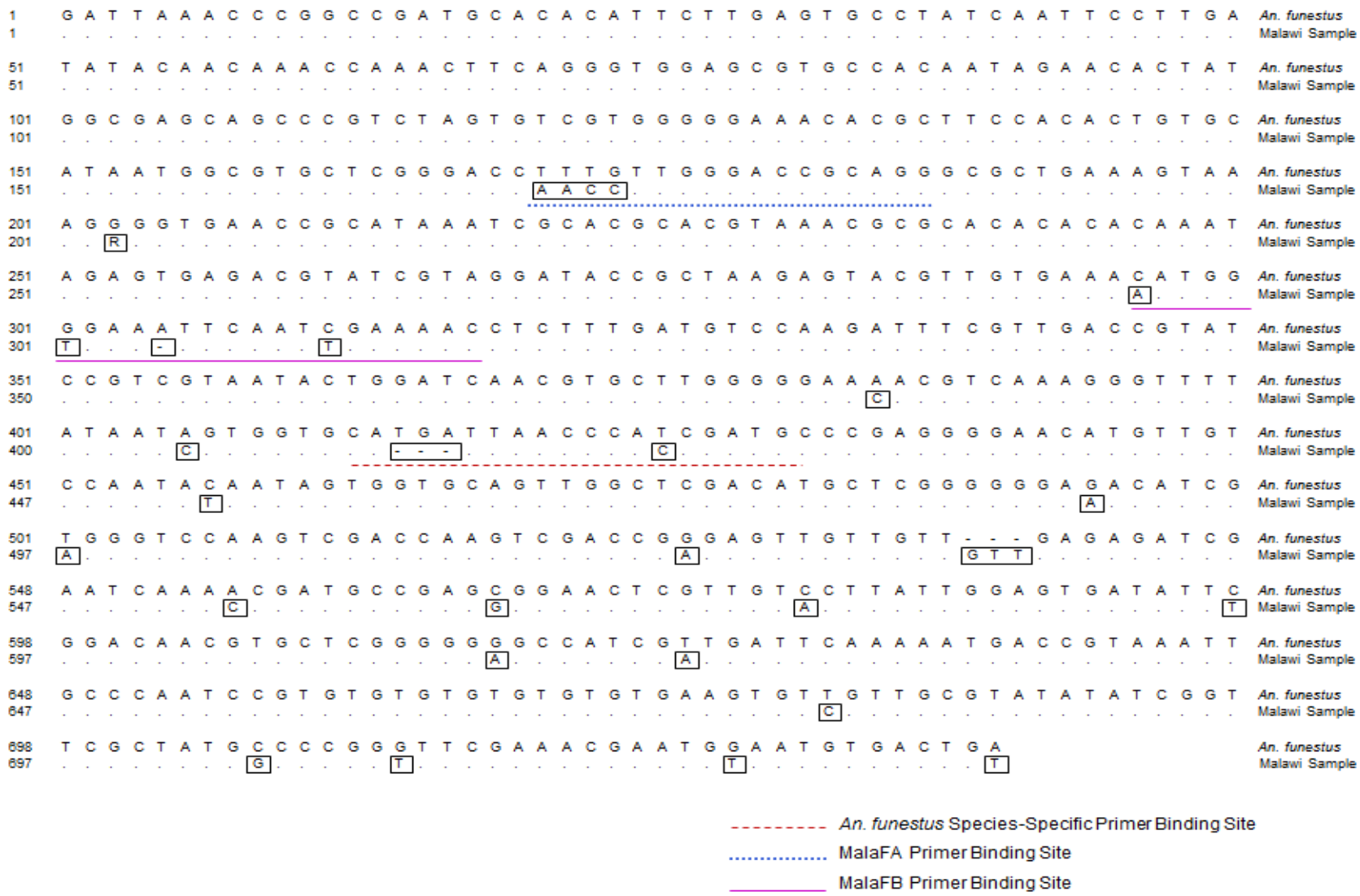


Figure 4.6. Alignment of the *ITS2* region in *Anopheles funestus* and the Malawian specimens. (•) indicates similar base pairs, (-) indicates deletions.

A 300bp portion of the *D3* region was sequenced and analysed for both the standard *An. funestus* colony and the MalaF specimens. Alignment of the sequences revealed a 1.5% change (5/300bp) which consisted of a 2bp deletion, 2 base transversions and a base transition (Figure 4.8).

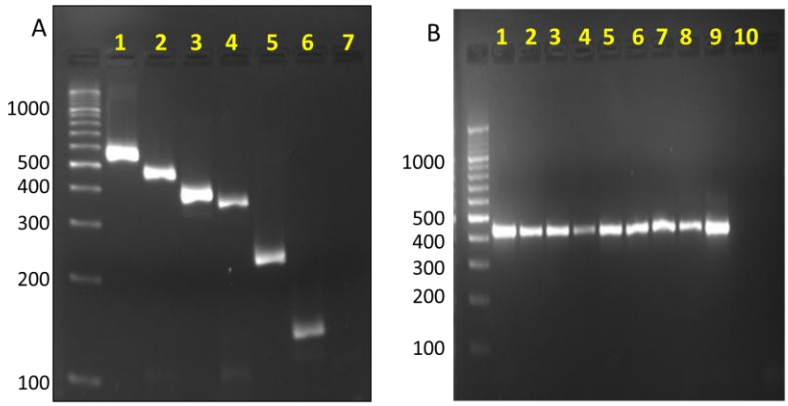


Figure 4.7. A. MalaFB primer included in the PCR cocktail mix for *An. funestus* group identifications. Lanes 1-3: Positive Controls *An. vaneedeni*, *An. funestus* and *An. rivulorum* respectively; Lane 4: MalaF sample; Lanes 5-6: Positive Controls *An. parensis* and *An. leesoni* respectively; Lane 7: PCR Negative Control. **B**. Lanes 1-7: Seven MalaF specimens that failed to amplify using the MalaFB primer, amplified here with the untargeted *An. funestus* group primers. Lanes 8-9: Positive Controls. Lane 10: No DNA PCR control.

```
1    TTATAACCATTAAACCCACAGGCGAAGACAACTCGATTGTCACGG  An. funestus
1    .....................................          Malawi Sample

46   GATTACGGGCACGGATAGGTGGCGCAAGCCCCTTATAGAACCGAG  An. funestus
46   ..........G..--.........A...T.....             Malawi Sample

91   CCCCTCCATCCCAGGGTGCTCCGTCACGGGTGCTTGCACCCAGCG  An. funestus
89   .....................................          Malawi Sample
```

Figure 4.8. Sequence alignment of the partial D3 region of the unidentifiable Malawian (MalaF) samples (133 basepairs) and *Anopheles funestus* (135 basepairs). Blocks highlight sequence variation. − indicates deletions. The forward D3A primer binds approximately 90 basepairs upstream from the start of this sequence

The secondary structure of the *ITS2* regions was analysed using the freeware program, mFold (Zucker, 2003). The sequences for the MalaF specimens as well as *An. funestus* group members were folded in their respective RNA formats (Figure 4.9). It is clear from the predicted structures that the ITS2 fragments from each of the *An. funestus* group members folds with a distinct pattern. The MalaF samples fold similarly to *An. funestus* although a fair amount of structural change can be seen. The same can be seen for the structures created for the *An. rivulorum* and *An. rivulorum*-like sequences.

Lastly, in order to test specific status of the MalaF specimens, species crossing experiments were carried out between F1 MalaF specimens and a laboratory reared strain of *An. funestus.* The resulting hybrids were scored for asynapsis between homologous chromosomes and the testes were dissected to score hybrid infertility. Detailed results of this work are published in Spillings *et al*. (2009). Briefly, the hybrid polytene chromosomes (MalaF ♀ X *An. funestus* ♂) displayed consistent asynapsis whilst the testes in hybrid males appeared normal. The reciprocal cross produced a large number of eggs with an extremely low hatch rate of 0.2%. The two eggs that did hatch yielded two females that died prior to taking a blood meal.

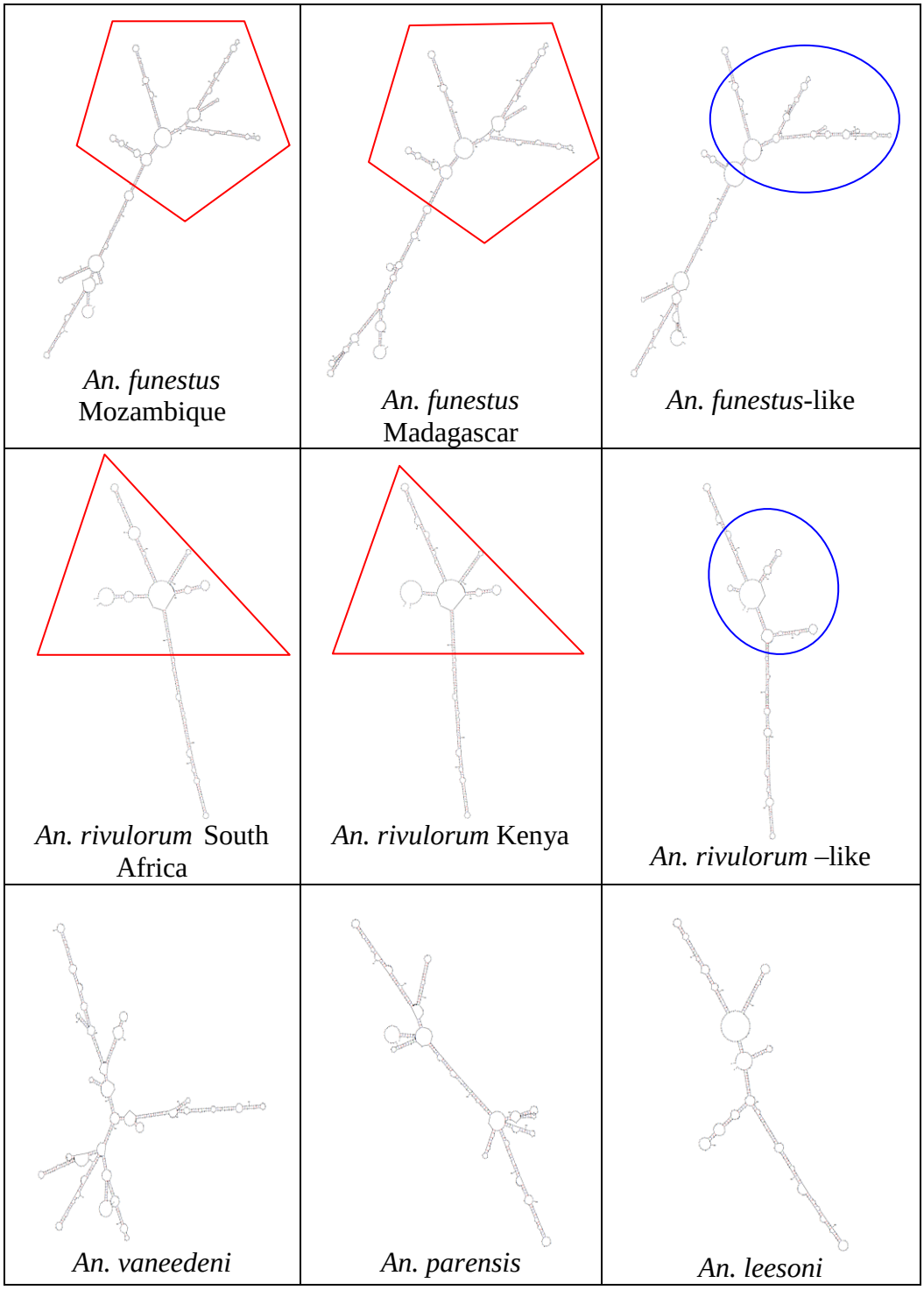


Figure 4.9. Predicted ITS2 secondary structures for members of the *An. funestus* group. Structures highlighted in red are conserved structures within a species from two different localities. Structures highlighted in blue show the areas which vary from the conserved structure.

The hybrid status of the progeny resulting from the crossing experiments was tested using the *An. funestus* species identification PCR with the MalaFB primer included in the PCR master mix. All the hybrids gave fragments for both the *An. funestus* and MalaF species (Figure 4.10) irrespective of hybrid gender, thereby confirming their hybrid status.

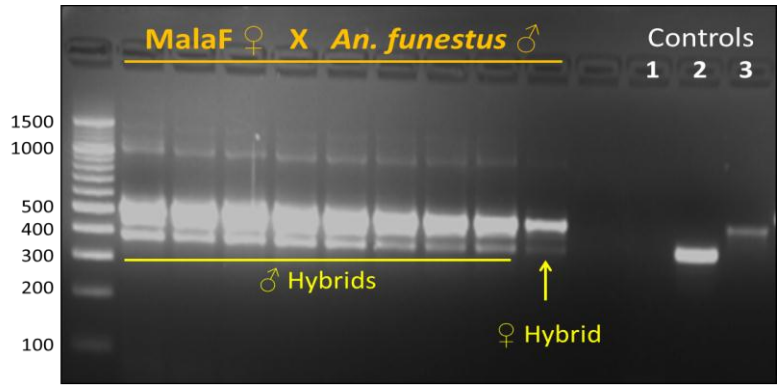


Figure 4.10. PCR confirmation of hybrid status. Control 1: no DNA PCR Control; Control 2: MalaF Positive Control; Control 3: *An. funestus* Positive Control.

4.6 DISCUSSION

Malawi's malaria control and prevention programme relies heavily on the use of insecticide treated bednets (ITN) to prevent the transmission of malaria (www.cdc.gov/Malaria/control-prevention/malawi.htm). In Karonga, bed-nets were often seen hanging in the traditional houses. Most of the household owners claimed that their bednets had never been retreated with insecticide. It is thought that this may be due the fact the bednet owner's bear the costs of retreatment since it is not provided for by the control programme. This locality has not been targeted for indoor residual spraying and farming is limited to a subsistence/feudal system, where few insecticides are used. The predominant malaria vector species in Malawi are *An. gambiae*, *An. arabiensis* and *An. funestus* (Spiers *et al*., 2002; Hunt *et al*., 2010). Based on

this, collections were carried out in Karonga, northern Malawi in December 2007. The mosquitoes caught during this field trip were identified as belonging to the Funestus subgroup, based on the adult, larval and egg morphology (Gillies and De Meillon, 1968). The sensitivity of the F1 progeny of these Malawian specimens to all classes of insecticide is not surprising since the use of insecticides in this locality appears to be limited.

Repeated attempts to identify these mosquitoes to species level using the multiplex PCR of Koekemoer *et al*. (2002) failed for 61/63 of the specimens. Previously, in August 2007, a collection carried out in the same locality exhibited an identification failure rate of 80%. Initially it was suspected that the DNA in these specimens may have degraded thereby resulting in the high failure rate observed. However, the specimens used for the PCR identification in this study, were freshly euthanized F1 progeny and DNA extraction controls were used in all experimental procedures. The failure to amplify the species targets for these specimens could not be explained by DNA degradation and thus the *ITS1*, *ITS2* and *D3* regions were investigated.

The reasons for investigating the *ITS2* region were two fold. Firstly, the species identification PCR used in this study targets the variable regions in the *ITS2* gene. Secondly, *ITS2* tends to be less functionally restricted than other barcoding type genes. The nucleotide sequence of the *ITS2* region may be free to evolve more rapidly whilst managing to conserve structural functionality (Müller T *et al*., 2007). Mutations could thus accumulate relatively quickly over

an evolutionarily short period of time. This suggests that the *ITS2* region may potentially be a good indicator of early genetic discontinuity between organisms (Coleman, 2009).

Sequencing of the *ITS2* region for the Malawian specimens showed the presence of a base pair deletion and a T to C transition within in the region of the *An. funestus* specific primer binding site. These alterations resulted in the species-specific primer being unable to anneal and hence caused the failed amplification reactions. A similar failure of annealing in the species multiplex PCR was observed by Couhet *et al*. (2003) during the discovery of *An. rivulorum*-like in Cameroon.

Further sequence analysis of the Malawian *ITS2* region showed a high level of sequence variation (4.5%) when compared to the *An. funestus* control. In the *An. gambiae* complex the levels of inter-specific divergence in this region can range from lows of 0.4% to 1.6% (Paskewitz *et al*., 1993). The sequence divergence between *An. rivulorum* in eastern and southern Africa and *An. rivulorum*-like in west and central Africa was found to be 19% (Hackett *et al*., 2000).

Analysis of the *D3* sequence of the Malawian specimens showed 1.5% variation in comparison to the *An. funestus* control. This low level of variation is significant since differences as small as 2-3 basepairs within *D3* have been used to differentiate the members of the *An. fluviatilis* complex (Singh *et al*., 2004). A small difference of 5 substitutions has been observed for this region in *An. mimimus* species A and C (Sharpe *et al*., 2000). Combined, the

levels of sequence variation seen in both the *ITS2* and *D3* regions in the Malawian specimens confirm, at the molecular level, the results obtained from the cytogenetics indicating that these specimens are genetically distinct from *An. funestus*.

The ITS2 secondary structure for the Malawian specimens shows close similarity to the *An. funestus* structure. This level of similarity is echoed in the structures derived for *An. rivulorum* and *An. rivulorum*-like. It was not the purpose of this research to develop a primary ITS2 structure for the *An. funestus* group. However, the structures do have regions of similarity between the species. Further work on this region would enable one to generate a primary sequence where compensatory base changes (CBCs) could then be identified for each species. Since almost all eukaryotic ITS2 structures have a typical four helix shape (Coleman, 2003), the presence and positioning of CBCs can help to elucidate evolutionary pathways as well as assist in predicting sexual compatibilities within and between clades and groups (Coleman, 2009).

In order to be able to identify these Malawian specimens, a PCR primer was designed that was specific to their *ITS2* region. Although two potential primers were tested for suitability, the most suitable primer (MalaFB) yielded an amplicon of approximately 390 basepairs in length. Unfortunately this amplicon is too close in size to that of *An. rivulorum* (~411bp) to be incorporated in the multiplex PCR mix, as seen by the species product "ladder" created in Figure 4.8A. Koekemoer *et al*. (2002) based the design of their multiplex PCR on the basis that the optimal difference in amplicon size for easy visualisation on an agarose gel is

approximately 50 basepairs. The amplicon derived from the Malawian specimens is only 21 basepairs shorter than the *An. rivulorum* amplicon. This small difference in length increases the risk of misidentifications especially when separation between the amplicons is not fully achieved. Factors such as agarose concentration and the length of time spent electrophoresing the amplicons will affect the success of the electrophoretic separation. With this in mind, the MalaFB primer does however allow for the identification of unidentified specimens where the initial multiplex PCR has failed. The MalaFB has been tested on the 24 specimens that failed to amplify in the collections of August 2007 (data not shown) and a high rate of amplification was achieved. We will continue to test this primer on field-caught specimens from other regions that fail to amplify using the *An. funestus* multiplex PCR assay of Koekemoer *et al*. (2002). The seven samples from this study that failed to amplify using both the multiplex PCR assay and the MalaFB primer were confirmed for DNA integrity through the success of the in-house Funestus subgroup PCR. Thus, DNA degradation could not be the cause of the failed amplifications and further molecular analyses are necessary for these specimens.

Before molecular tools became available for distinguishing morphologically similar species, cytogenetics and cross-mating studies were used to discriminate between members of sibling species complexes. Members of the European *An. maculipenis* complex were distinguished by the chromosomal banding patterns seen in salivary gland polytene chromosomes (Frizzi, 1947 and 1953). This success in cytotaxonomy was quickly followed by the cytogenetic description of the members of the *An. gambiae* complex (Coluzzi and Sabatini, 1967, 1968 and 1969). More recently, in 2001, a cytogenetic map for *An. funestus* was published giving details and

locations of inversion break points and inversion frequencies (Sharakhov *et al*., 2001). The chromosomal banding patterns of the Malawian specimens displayed homosequential banding arrangements with *An. funestus*, but were fixed for the inverted arrangements of 3a, 3b and 5a inversions, which are normally polymorphic in *An. funestus*. Although *An. vaneedeni* also has homosequential chromosomes with *An. funestus* (Green and Hunt, 1980), the fixed inverted arrangements on arms 3 and 5 of MalaF distinguish it from *An. vaneedeni.*

Cytogenetic studies on West African *An. funestus* have provided evidence of potential species differentiation within these populations (Lochouarn *et al*., 1998; Constantini *et al*., 1999). Unfortunately, these findings could not be confirmed by cross-mating studies, since at that time *An. funestus* had not been successfully colonised. Cross-mating experiments have been widely used to establish genetic discontinuities between sibling species within anopheline complexes (*An. maculatus* form K by Thongwat *et al*., 2008; *An. pseudopunctipennis* species C by Coetzee *et al.,* 1999; *An. minimus* species E by Somboon *et al.,* 2005; *An. quadrimaculatus* types A and B by Kaiser, 1988; *An. annulipes* species A and G by Foley and Bryan 1991) where hybrids were scored for asynapsis between homologous chromosomes as well as hybrid infertility. In this study, the hybrid chromosomes resulting from the MalaF females × *An. funestus* males showed consistent asynapsis between homologous chromosomes, typical of inter-species crosses (Davidson *et al*., 1967; Green and Hunt, 1980; Hunt *et al*., 1998). Dissections of the resulting male hybrids showed normal testis morphology with the possible exception that the head region of the spermatozoa appeared narrower. Unfortunately, the effect of this narrower morphology in terms of male fertility is unknown because we were unable to carry out back crosses. Eggs produced from the reciprocal crosses

displayed a very low level of viability (< 0.2% hatch rate), suggesting a genetic discontinuity between the parental samples.

The status of the resulting hybrids was confirmed using the *An. funestus* multiplex PCR assay (Koekemoer *et al*., 2002) combined with the MalaFB primer. In the past, it was generally believed that rDNA was restricted to the X-chromosome in anopheline species and that recombination between the X and Y chromosomes did not occur. If this was the case, one would expect male hybrids to only carry the *ITS2* genotype of their mother. In this study, the hybrids generated from crossing the Malawian specimens with *An. funestus* carry the *ITS2* genotypes of both of their parents. This suggests that the rDNA in these species cannot be restricted to the X chromosome. It has been suggested that rDNA might not be restricted to the X chromosome in some members of the *An. gambiae* complex, namely in *An. quadriannulatus, An. merus* and *An. melas* (cited as unpublished data by S. Paskewitz in Collins *et al*., 1989). Further, the possibility of recombination between the X and Y chromosomes has been suggested for *An. quadrimaculatus* (Mitchell and Seawright, 1989); *An. gambiae* (Krzywinski *et al*., 2005; Wilkins *et al*., 2007) and *An. culicifacies* (Sakai *et al*., 1979).

4.7 CONCLUSION

Based on the combined molecular, cytogenetic and cross-mating evidence, it is concluded that the Malawi population is a new member of the Funestus Subgroup. It is provisionally designated "*An. funestus*-like" until such time as a formal species description is published.

Further molecular investigations are needed to determine how this new species impacts on the variation seen in RFLP (Garros *et al.*, 2004) and mitochondrial DNA (Michel *et al.*, 2005) analyses of *An. funestus* populations from the southern African region. More detailed studies need to be carried out to generate a primary *ITS2* secondary structure for the Funestus Subgroup as well as to clarify the potential for X and Y chromosomal recombination in these species.

Further investigations into the basic biology of this new species are also required. Although none of the 61 specimens examined for malaria parasites during this study were positive for *P. falciparum*, the fact that these mosquitoes are common inside houses makes them potential vectors. Future collections at different times of the year are needed to clarify the vector status, abundance and distribution of this new species. These studies could also provide data on the interactions between this new species and *An. funestus s.s.* in areas where they occur in sympatry.

4.8 Spillings BL, Brooke BD, Koekemoer LL, Chiphwanya J, Coetzee M and Hunt RH. 2009. A new species concealed by *Anopheles funestus* Giles, a major malaria vector in Africa. *American Journal of Tropical Medicine and Hygiene* **81**: 510 – 515.

Contribution to publication:

Belinda Spillings assisted with the field collections in Karonga, Malawi; carried out the WHO exposures; the molecular work and the data analysis; wrote the first and subsequent drafts of the manuscript.

CHAPTER 5 – GENERAL DISUCSSION, FUTURE RESEARCH AND CONCLUDING REMARKS

5.1 GENERAL DISCUSSION

The outbreak of malaria in South Africa during 1999/2000, and the discovery of pyrethroid resistant *An. funestus* in KwaZulu-Natal, highlighted the need to explore this resistance phenotype. The threat of pyrethroid resistant *An. funestus* populations from Mozambique encroaching into South Africa and the potential resurgence of *An. funestus* mediated malaria infections pose a serious obstacle to South Africa's stated goal of eliminating malaria by 2015.

The pyrethroid resistant phenotype of southern African *An. funestus* has been linked to increased P450-mediated metabolism of the insecticide, specifically through the over expression of *CYP6P9* (Amenya *et al*., 2008; Wondji *et al*., 2009). Further, this major effect gene is located within a QTL (*rp1*) which has been linked to pyrethroid resistance in *An. funestus* (Wondji *et al*., 2009). Given that the malaria transmitting portion of mosquito populations enter houses to blood feed and then rest, it was necessary to consider whether a blood meal would confer enhanced protection from the pyrethroid insecticide used in the IRS campaigns. The effect of a blood meal on insecticide tolerance in both pyrethroid resistant and susceptible strains of *An. funestus* was investigated in order to determine if a genetic mechanism was involved.

Through the use of CDC bottle bioassays it was demonstrated that blood fed pyrethroid resistant *An. funestus* displayed higher levels of insecticide tolerance than their unfed counterparts. In contrast, insecticide susceptible *An. funestus* were equally susceptible to pyrethroid intoxication, regardless of whether or not a blood meal had been taken. This suggests that the increased body mass of blood fed individuals does not significantly dilute internalized insecticide and thus increased vigour tolerance is not a suitable explanation for the increase in insecticide tolerance recorded in blood fed pyrethroid resistant *An. funestus*. Consequently, these results suggest that the presence of a blood meal enhances the expression of pyrethroid resistance only in those females that already carry a pyrethroid resistant phenotype, inadvertently priming them against the effects of insecticide intoxication.

In order to determine whether this blood meal 'priming mechanism' is linked to detoxification gene expression, microarray analysis using the *An. gambiae* detox chip was performed. A comparison between blood fed and unfed cohorts, for both insecticide resistant *An. gambiae* and *An. funestus* was then undertaken. The data generated allowed for the comparison of homologous and heterologous hybridisation arrays as an investigative tool to highlight the potential gene targets common to both species.

Microarray analysis revealed that blood fed *An. gambiae* display elevated delta-class GST expression and concomitant down regulation of cytochrome P450s. Previous studies have shown that GSTs may be upregulated in response to a blood meal in hematophagous arthropods (Rudenko *et al*., 2005; Dana *et al.*, 2006; Jochim *et al*., 2008). Expression analysis

of blood fed *An. funestus* did not reveal any significantly down regulated genes, but did reveal a commonly up-regulated GST (*GSTD3*) in both *An. gambiae* and *An. funestus.* Although an increase in *GSTD3* expression following a blood meal in *An. gambiae* has been shown previously (Marinotti *et al*., 2005), this is the first study to demonstrate the increased expression of *GSTD3* in blood fed *An. funestus*. Since *GSTD3* is over expressed in both *An. gambiae* and *An. funestus*, it served as a good target for validation through q-PCR.

Since *CYP6P9* and *CYP6P13* have been implicated as the major effect genes for pyrethroid resistance in *An. funestus* FUMOZ-RH (Amenya *et al*., 2008; Wondji *et al*., 2009; Matambo *et al.*, 2010), they were also included in the real time qPCR investigation. Each of these genes, when assayed independently, exhibited slightly elevated levels of expression in association with blood feeding in pyrethroid resistant *An. funestus*. Although these genes normally occur in an overexpressed state in *An. funestus* FUMOZ-RH, the slight elevation in expression following a blood meal is likely to be responsible for the increased pyrethroid tolerance recorded in the bottle bioassay experiments. Unfortunately, due to the inability to target *GSTD3* in *An. funestus*, expression levels of this gene in response to a blood meal could not be validated.

Many biological experiments carried out using laboratory reared organisms are considered predictive of the functions and responses of the same organisms under field conditions. The sole use of laboratory reared organisms in experiments has been criticised, especially where parallels need to be drawn between laboratory and field organisms. The physiology and fitness

of field caught *An. gambiae* in comparison to laboratory reared strains has been found to differ significantly (Day and Van Handel, 1986; Huho *et al*., 2007). With this in mind, it was intended to test whether the findings of the blood meal investigations would hold true for wild caught, blood fed *An. funestus*.

Field collections of mosquitoes were carried out in Northern Malawi and specimens were morphologically identified as *An. funestus* group based on the keys of Gillies and Coetzee (1987). Upon attempts to molecularly identify the species it was determined that these Malawian specimens were in fact not *An. funestus*. Through the use of molecular, cytogenetic and cross-mating studies we were able to show that these specimens are a new member of the Funestus Subgroup and have thus been provisionally named *An. funestus*-like.

5.2 FUTURE RESEARCH

Since the *An. funestus* genome has as yet not been sequenced, the *An. gambiae* detox chip was used as an investigative tool to identify detoxification genes that play a role in insecticide tolerance following a blood meal in *An. funestus*. Transcriptional studies of the resistance profiles of *An. stephensi* (Vontas *et al*., 2007), *An. arabiensis* (Müller *et al*., 2008) and *An. funestus* (Christian *et al*., 2011b) have been successfully investigated using the *An. gambiae* detox chip. Although heterologous hybridisations are possible (Moody *et al*., 2002; Nuzhdin *et al*., 2004; Sartor *et al*., 2005), caution should be exercised during experimental design and the downstream analysis of the results (Lu *et al*., 2009). Validation of expression data from a heterologous hybridisation is essential and can be done through the use of real time qPCR.

This serves to ensure that differential gene transcription reported in microarray experiments are the result of true target-probe binding and are not false signals.

Recent advances in the study of *An. funestus* have seen the development of an EST and SNP database (Serazin *et al*., 2009; Gregory *et al*., 2011). This represents a huge gain in information, a fair amount of which still requires confirmation at the level of the functional annotation. It is hoped that high throughput analyses and technological advances such as direct RNA sequencing will start to close the gaps in *An. funestus* sequence information.

Since *GSTD3* was also upregulated in the *An. gambiae* microarrays it is still considered an important target for future studies, especially when there is evidence to suggest that GSTs may play a role in minimizing the impact of oxidative stress brought about by the digestion of blood (Jochim *et al*., 2008) and pyrethroid intoxication (Vontas et al., 2001). Since GSTs function to protect cells by removing reactive oxygen species, it is suggested that *GSTD3* involvement in blood meal digestion and the subsequent increase in insecticide tolerance could be a secondary process to that of the major effect resistance genes.

Ultimately, the availability of the *An. funestus* genome will enable greater accuracy in gene transcription profiling. The creation of an *An. funestus* microarray will allow far more accurate investigations into gene expression within this organism. The application of RNA interference will allow researchers to investigate and confirm gene function through knock-down assays.

Lastly, research is required to determine the disease vector capacity (if any), distribution and basic biology of *An. funestus*-like. Further sampling of this species will elucidate the natural level of population variation as well as clarify issues such as host choice, seasonal preferences etc. Since species can be defined by the limits of gene exchange, investigations into the level of sequence divergence between the mitochondrial and genomic genes of the members of the Funestus Subgroup will assist in determining these limits of gene exchange and create more detailed phylogenies.

5.3 CONCLUDING REMARKS

An effective and successful vector control programme has, at its foundation, an accurate knowledge base which is constantly updated with recent information. This knowledge base includes basic information such as: mosquito species identity; disease vector incrimination; mosquito species composition by region; and insecticide susceptibility status of vector species and closely related non-vectors. Strong collaborations with the communities involved assist in information on bednet advocacy and amenability to IRS.

Vector incrimination of African malaria vectors is confounded by overlapping morphology between vector and non-vector species. The accurate identification of members of the *An. gambiae* complex is only possible through cytogenetic and molecular based tools. Although there are small differences in the morphological characters between members of the *An. funestus* group, discrimination between the species on morphology alone is extremely difficult and accurate identification also relies heavily on the use of molecular tools. Added to this, the

discovery of a new species, which occurs in sympatry with and is morphologically identical to the major malaria vector *An. funestus*, confounds the situation further. Clarity on the basic biology and bionomics of both vector and non-vector species is thus becoming increasingly important.

The detection of insecticide resistance in malaria vector populations is primarily based on WHO bioassays. Given that blood feeding enhances the expression of insecticide resistance in young (3 day old) genotypically resistant females depending on the resistance mechanism, it is recommended that young, blood fed females are included in samples during the routine insecticide susceptibility testing of target vector populations. This is important because blood-feeding may enhance the expression of resistance to a level where it is detectable using the WHO bioassay system, whereas testing non blood fed females only could allow newly emerging resistance to go undetected for a longer period, facilitating the spread of resistance alleles through affected populations.

The development of insecticide resistance in southern African malaria vectors serves to undermine the goal of malaria elimination in the region. The design of resistance management tools relies on an understanding of how resistance develops in field populations. The data presented here highlight the effect that a blood meal has on insecticide tolerance in resistant *An. funestus* and reinforces the necessity of accurate vector species identification.

APPENDIX A

Linear Regression Analysis of Bottle Bioassay Data

Mortality data from replicate one of the FUMOZ-RH 18 hour, blood-fed exposures is shown below in Table A1. This data was used to plot the mortality versus dosage graph in Figure A1. From the equation of the regression line, the LD50 can be determined as below:

$$y = 58.191x - 73.346$$

Where y is the percentage mortality and x the log (dosage)

If you want to calculate the LD_{50}, enter 50 for y and then solve for x as follows:

$$50 = 58.191x - 73.346$$
$$58.191x = 123.346$$
$$x = 123.346/58.191$$
$$x = 2.119$$

But x is the log (dosage), therefore you must antilog x to get the dosage:

$$x = 10^{2.119}$$
$$x = 131.5$$
$$LD_{50} = 131.5\ \mu g$$

After calculating the LD_{50} for each replicate, the mean LD_{50} and the standard error can be calculated (Table A2).

Table A1. Percentage mortality data for a replicate of blood-fed FUMOZ-RH mosquitoes, exposed to varying doses of Permethrin, 18 hours after feeding.

Dosage (µg)	Log (Dosage)	Percentage Mortality
10	1.00	0.00
25	1.40	4.20
50	1.70	9.50
100	2.00	34.80
250	2.40	68.20
500	2.70	95.83
1000	3.00	100.0

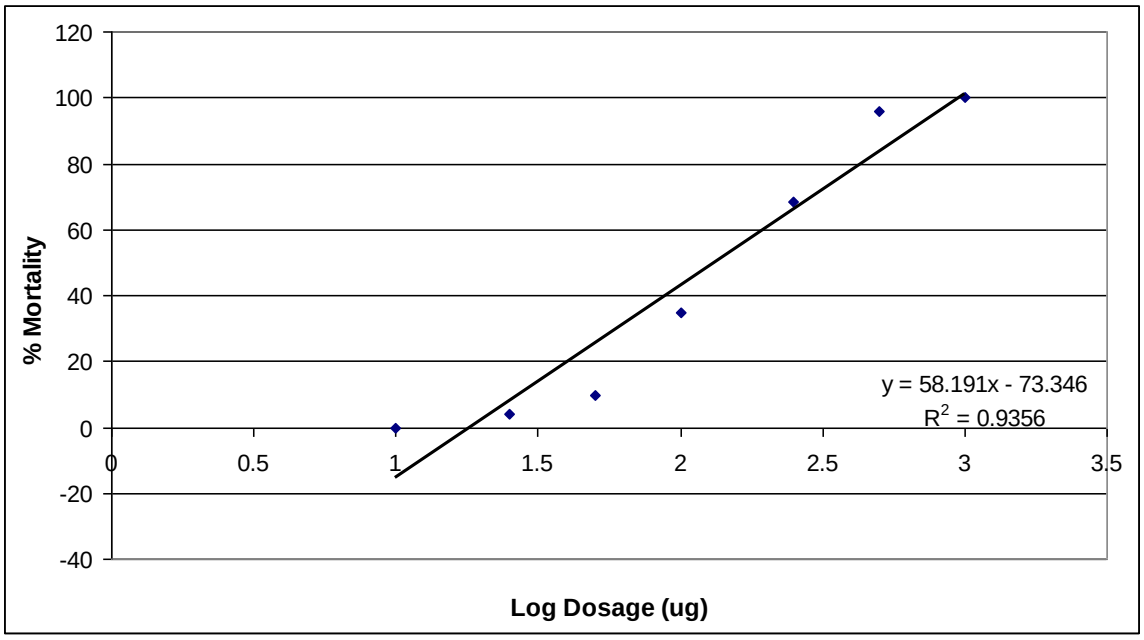


Figure A1. Percentage mortality plotted against log (dosage) for the blood-fed Fumoz-RH mosquitoes exposed to Permethrin 18 hours post blood feeding. The equation of the regression line and the R^2 value are shown.

Table A2. LD_{50} values for the FUMOZ-RH unfed cohort used in the 18 hour exposures.

	LD_{50}
Replicate 1	131.51
Replicate 2	97.53
Replicate 3	115.03
Replicate 4	90.60
Replicate 5	129.97
Replicate 6	156.66
Replicate 7	174.24
Replicate 8	125.73
Replicate 9	135.93
Mean	128.58
SE	8.77

APPENDIX B

RNA Extraction and Amplification for Microarray Analysis

RNA was extracted from 10 to 15 mosquitoes using the PicoPure™ RNA Isolation Kit (Arcturus Bioscience, Inc.) following manufacturer's guidelines. Briefly, the mosquitoes were placed in a 1.5ml microcentrifuge tube and then placed on ice. Extraction buffer (100 µl) was added to the tubes and the mosquitoes were ground using a sterile pestle. After the mosquitoes were finely ground, the tubes were incubated at 42˚C for 30 minutes. During the waiting period an RNA purification column was pre-conditioned by pipetting 250 µl conditioning buffer onto the column and incubating for 5 minutes at room temperature. The column was then centrifuged at 16 000g for 1 minute.

After incubating the RNA extractions at 42˚C, the samples were centrifuged at 12 000g for 1 minute. The supernatant was transferred to a clean 1.5ml microcentrifuge tube and 100 µl 70% ethanol added. The mixture was then pipetted onto the pre-conditioned column and centrifuged at 6g for 2 minutes. A second centrifuge step was carried out at 16 000g for 30 seconds. Wash buffer 1 (100 µl) was pipetted onto the column which was then centrifuged at 8 000g for 1 minute. In order to limit DNA carry over, a DNase I treatment was included in the protocol. DNase I solution (5 µl) was mixed with 35 µl RDD buffer (RNase-Free DNase Set, Qiagen) and then pipetted onto the column. The column was incubated at room temperature for 15 minutes.

The column was then rinsed with 40 µl wash buffer 1, followed by centrifugation at 8000g for 15 seconds. Wash buffer 2 (100 µl) was pipetted onto the column which was then centrifuged

at 8 000g for 1 minute. A second wash of 100 µl wash buffer 2 was pipetted onto the column and then centrifuged at 16 000g for 2 minutes. The column was centrifuged twice more at 16 000g for 1 minute, to remove the remaining traces of the wash buffers. The column was transferred into a clean 0.5 ml collection tube.

By gently touching the surface of the column with the pipette tip, an aliquot of 30 µl elution buffer was pipetted onto the column. The column was incubated at room temperature for 1 minute after which it was centrifuged 100g for 1 minute. This was immediately followed by 16 000g for another minute. The RNA extraction passed through the column and collected in the 0.5ml collection tube. The total RNA quantity and quality was assessed by spectrophotometry using the RNA-40 setting on the Nanodrop machine (Nanodrop Technologies, UK). Samples with less than 200 $ng.\mu l^{-1}$ of RNA were excluded from the following steps.

Approximately five micrograms of total RNA was amplified using the RiboAmp™ amplification kit (Arcturus Bioscience, Inc.) according to manufacturer's instructions. Briefly, to each RNA sample 1 µl of Primer A was added. The RNA samples were then incubated at 65˚C for 5 minutes. After incubation the tubes were placed onto ice for at least 1 minute. A first strand amplification mix was made using 7 µl of first strand master mix and 2 µl first strand enzyme. Once the RNA samples were cooled, the 9 µl of first strand master mix was added. The samples were then incubated at 42˚C for 45 minutes. After incubation the samples were placed on ice to cool. First strand nuclease mix (2 µl) was added to the samples which were then incubated at 37˚C for 20 minutes. This was followed by an enzyme denaturation step at 95˚C for 5 minutes. The samples were then placed on ice to cool. To each of the first

strand sample tubes, 1 µl of Primer B was added. The samples were briefly denatured at 95˚C for 2 minutes and then chilled on ice for at least 2 minutes. A second strand amplification mix was made using 29 µl second strand master mix and 1 µl second strand enzyme. Once the samples were cooled, the 30 µl second strand master mix was added. The samples were then incubated at 25˚C for 5 minutes; 37˚C for 10 minutes; 70˚C for 5 minutes. After incubation the samples were placed on ice to cool.

Whilst the samples cooled on ice, the nucleic acid purification column was prepared. Room temperature DNA binding buffer (250 µl) was pipetted onto the column and allowed to stand at room temperature for 10 minutes. The column was then centrifuged at 16 000g for 1 minute. The second strand synthesis sample was ten mixed with 200 µl of DNA binding buffer and pipetted onto the column. The column was centrifuged briefly at 6g for 2 minutes, followed by 10 00g for 30 seconds. The column was then rinsed with 250 µl DNA wash buffer and then centrifuged at 16 000g for 2 minutes. The column was then placed in a clean 0.5 ml collection tube and 16 µl of elution buffer pipetted onto the surface of the column. The buffer was distributed by gently tapping the column and tube. This was followed by a five minute incubation at room temperature. The cDNA was then eluted by centrifuging the column at 1 000g for 1 minute, immediately followed by 16 000g for 1 minute.

For each sample, an IVT reaction mix was created by mixing the following: 8 µl IVT buffer, 12 µl IVT master mix and 4 µl IVT enzyme mix. This 24 µl IVT mixture was combined with the eluant above. The samples were mixed well and incubated at 42˚C for 6 hours. After incubation the samples were treated with DNase in order to remove the remaining DNA. The

antisense RNA (aRNA) was then purified using a freshly prepared nucleic acid purification column. The column was prepared by pipetting 250 µl RNA binding buffer onto the column and allowing it to incubate at room temperature for 10 minutes. The column was the centrifuged at 16 000g for 1 minute. To each sample 200 µl of RNA binding buffer was added and the mixture was pipetted onto the column. The column was centrifuged at 6g for 2 minutes, immediately followed by 10 000g for 30 seconds. RNA wash buffer (200 µl) was pipetted onto the column and centrifuged at 16 000g for 2 minutes, immediately followed by 16 000g for 1 minute. The column was placed into a clean 0.5 ml collection tube and 30 µl elution buffer added directly onto the column. The column was tapped gently to distribute the elution buffer and then incubated at room temperature for 5 minutes. The column was then centrifuged at 1 000g for 1 minute, immediately followed by 16 000g for 1 minute. The quantity and quality of product was confirmed using the Nanodrop machine (minimum cutoff values: 1000 $ng.\mu l^{-1}$ and a 260/280 ratio of at least 1.7).

Cy Dye Labeling of RNA for Microarrays

The resulting aRNA (8 µg) was placed in 0.2 ml PCR tubes with 1.3 µl random hexamers (50 µM, Invitrogen Life Technologies) and 2 µl RNA spike-in control from the Lucidea Universal Score Card (Amersham Biosciences Ltd, UK). These were incubated at 70˚C for 5 minutes followed by 4˚C for at least 2 minutes. The samples were then reverse transcribed to cDNA using Superscript III reverse transcriptase kit (Invitrogen) and fluorescently tagged Cy3- and Cy5-dUTPs (Amersham Biosciences Ltd, UK). Reverse transcription was carried out at 50˚C for 2½ hours.

To stop the reaction and degrade the original aRNA template, 1 µl of a 1M NaOH, 20 mM EDTA solution was added. The Cy-labeled targets were then combined and cleaned using the illustra^{TM} CyScribe^{TM} GFX^{TM} Purification Kit (GE Healthcare, USA).

Briefly, to each Cyscribe column 500 µl of capture buffer was added. The labeled cDNAs were then added to the column and mixed with the capture buffer. Each cDNA set consisted of a Cy3 labeled cDNA and it corresponding Cy5 labeled sample. Since Cy dyes are light sensitive, the columns were kept under foil at all times. The columns were then centrifuged at 13 000 rpm for 30 seconds. The flow through was discarded and the columns placed in clean collection tubes. To each column, 600 µl of wash buffer (containing ethanol) was added and the column centrifuged again at 13 000 rpm for 30 seconds. The flow through was discarded and the columns placed in clean collection tubes. The wash step was then repeated twice more. A final spin at 13 000 rpm for 10 seconds was carried out in order to remove any remaining ethanol. The columns were then placed in clean collection tubes. Elution buffer was pre-warmed to 37˚C and 60 µl added to each column. The columns were allowed to incubate at room temperature for 2 – 4 minutes and then centrifuged at 13 000 rpm for 2 minutes. A further 40 µl of elution buffer was added to the columns and allowed to incubate at room temperature for 2 to 4 minutes. A final centrifuge at 13 000 rpm for 2 minutes eluted the remaining labeled sample.

After purification, dye incorporation was assessed using a Nanodrop machine (Nanodrop Technologies, UK) and the microarray (33-factor) settings. A minimum cDNA yield of 15 ng.µl^{-1}; minimum dye incorporation of 0.1 pmol.µl^{-1} for each Cy dye and a 260/280 ratio of

less than two, were the cutoff values used to control the efficiency of labeling and purification. Finally, 5 µg of poly dA oligo (Sigma) was added to each target set, combined and then vacuum dried using an Eppendorf Concentrator 5301 (Eppendorf, Hamburg) set at 45°C for 30 minutes.

Preparation of the Detox Chip and Pre-Hybridisation

The detox chips used in this study were prepared for hybridisation using the Pronto! Universal Microarray Hybridisation Kit (Corning Inc, ref 40026). The following solutions were prepared in advance for the *An. funestus* hybridisations. These solutions were at 3 x recommended concentrations. The *An. gambiae* hybridisations used the solutions made up according to manufacturer's guidelines. If two slides needed to be prepared at the same time, they could be prepared back to back in 50 ml Falcon tubes.

Wash Solution 1: 75 ml Universal Wash Reagent A, 3.75 ml Universal Wash Reagent B and 171.25 ml deionized water.

Wash Solution 2: 75 ml Universal Wash Reagent A and 425 ml deionized water.

Wash Solution 3: 300 ml Wash Solution 2 and 200ml deionized water.

The pre-soak steps are started by warming the Pronto! Universal Pre-Soak Solution and the Pronto! Universal Pre-Hybridisation Solution to 42°C for at least 30 minutes. After the solutions were warm, 500 µl sodium borohydride solution was added to 50 ml Universal Pre-Soak Solution in a 50ml Falcon Tube. The microarray slide was then immersed into the sodium borohydride-Pre-soak solution and incubated at 42°C for 20 minutes. The slide was

then transferred to Wash Solution 2 and incubated at 25°C for 30 seconds. The slide was then transferred to a fresh container of Wash Solution 2 for a further 30 seconds. Next the slide was transferred to the Universal Pre-Hybridisation Solution and incubated at 42°C for 15 minutes. The slide was then transferred back to a fresh container of Wash Solution 2 and incubated at 25°C for 1 minute. Lastly, the slide was rinsed in two rinses of Wash Solution 3 at 25°C for 30 seconds. The slide was dipped briefly into nuclease-free water (at 25°C) and then placed in a clean 50ml Falcon tube. The slide was then dried by centrifugation: 2 500 rpm for 2½ minutes. At his stage the slide is ready for hybridisation and should remain in a dust free, dark container until ready for use.

Target preparation and Hybridisation onto the Detox Chip

The dried pellets resulting from the cy dye labeling were resuspended in 15.5 µl of long hybridisation buffer (Corning Inc.) and denatured at 95°C for 5 minutes. The samples were then centrifuged at 13 000 rpm for two minutes and then kept in the dark at room temperature until they were loaded onto the array. Targets were pipetted onto the center of a 22x22 hybrislip (Grace Biolabs). The microarray slide was then lowered, DNA side down, onto the coverslips. Care was taken not to allow the coverslips to move once they were on the array. The microarray slides were then placed in a hybridisation chamber. Hybridisation was carried out in a water bath at 38°C and 42°C for the *An. funestus* and the *An. gambiae* experiments respectively (Christian *et al*., 2011b; David *et al*, 2005). Hybridisation was carried out for 16 hours in hybridisation chambers that had approximately 40 µl of 3x SSC buffer distributed around the inner edges.

Post-Hybridisation Washing of the Detox Chip and Scanning

For the *An. funestus* hybridisations, post-hybridisation washes were carried out at 38°C for half the specified times (Christian *et al*., 2011b). The *An. gambiae* hybridisations on the detox chip were washed at 42°C for the recommended wash times, according to manufacturer's guidelines for the Pronto! Universal Microarray Hybridisation Kit (Corning Inc.). Wash steps requiring room temperature were carried out as close to 25°C as possible. Wash solutions were prepared as for the pre-soak protocol above.

Prior to starting the post hybridisation washes, Wash Buffer 1 was warmed to 38°C (*An. gambiae* hybridisations) or 42°C (*An. funestus*) for at least 30 minutes. The incubation times set out below are those for *An. funestus* hybridisations.

After removing the slide from the hybridisation chamber it was immersed in Wash Solution 1 for 1 to 2 minutes. As soon as the both coverslips lifted off the slide, it was transferred to a fresh container of Wash Solution 1 and incubated for 2½ minutes. Directly after this, the slide was transferred to a container of Wash Solution 2, at room temperature, and incubated for 5 minutes. The slide was then transferred to Wash Solution 3, again at room temperature, for 1 minute. The wash step using Wash Solution 3 was repeated twice more. The slide was dried by centrifugation at 2 500 rpm for 2 minutes.

All the slides were scanned immediately after drying. Scanning was carried out using a GenePix 4000B scanner (Axon Instruments, Molecular Devices, USA). The photomultplier (PMT) range was adjusted to fall within 500 – 750V for each channel of detection thereby resulting in an overall intensity ratio as close to 1 as possible. The resulting TIFF files were analysed using the GenePix Pro 6.0 software (Axon Instruments). Preprocessing according to

the following cutoff values was used to determine spot quality. If a spot met either of the following criteria, it was excluded from analysis: a signal to noise ratio of less than 3 and or spot intensity values of greater than 65,000. The raw intensity values were then imported into the Limma 2.4 software running in the R 2.9.0 package of Bioconductor for analysis (Gentleman *et al.*, 2004).

APPENDIX C

Programming Code Used for Microarray Analysis

Programming code for analysing the data generated for the Detox Chip hybridisations. Lines preceded by "#" and in blue font are descriptions of the programming code below and would not be entered into the program.

```
#Data Analysis of the microarray slides hybridised with blood fed and unfed An. gambiae
GAH #cDNA. This was done in order to determine which genes are differentially expressed in
#response to the blood meal.
setwd("C:/Documents and Settings/belindas/Desktop/GAH GPR Files for Analysis")
library(limma)
utils:::menuInstallLocal()
#Select the Statmod package in its zipped format
#Read in the list of slides and spot types for the analysis
targets<-readTargets()
table<-read.table("spotfile.txt", header = TRUE, sep = "\t", as.is = TRUE, nrows = 8)
spottypes<-table
targets
spottypes
#Read in the image data for each microarray slide
RG<-read.maimages(targets$FileName, source = "genepix",
wt.fun=wtflags(weight=0,cutoff=0))
#Read in the gal file. This file gives information relating to the location of each probe on the
#array and what the target is
RG$genes<-readGAL("New_An gambiae detox gal May 2007.gal")
#Confirm that the printer layout has been read correctly
names(RG$printer)
RG$printer
#Read in the status of each spot in the arrays
RG$genes$Status<-controlStatus(spottypes,RG, spottypecol = "Spottype", verbose = TRUE)

#Plot the MA plots of the raw data. Bear in mind that the outliers will not be seen.
#Command x11 below opens a graphics window for the graphs to be viewed
x11()
plotMA(RG, array=1, legend = T, ylim=c(-10,10))
x11()
plotMA(RG, array=2, legend = T, ylim=c(-10,10))
```

```
x11()
plotMA(RG, array=3, legend = T, ylim=c(-10,10))
x11()
plotMA(RG, array=4, legend = T, ylim=c(-10,10))
x11()
plotMA(RG, array=5, legend = T, ylim=c(-10,10))
x11()
plotMA(RG, array=6, legend = T, ylim=c(-10,10))
x11()
plotMA(RG, array=7, legend = T, ylim=c(-10,10))
x11()
plotMA(RG, array=8, legend = T, ylim=c(-10,10))
x11()
plotMA(RG, array=9, legend = T, ylim=c(-10,10))
x11()
plotMA(RG, array=10, legend = T, ylim=c(-10,10))
x11()
plotMA(RG, array=11, legend = T, ylim=c(-10,10))

x11()
imageplot(log2(RG$Rb[,1]),RG$printer,low="white", high="red")
x11()
imageplot(log2(RG$Gb[,1]),RG$printer,low="white", high="green")
x11()
imageplot(log2(RG$Rb[,2]),RG$printer,low="white", high="red")
x11()
imageplot(log2(RG$Gb[,2]),RG$printer,low="white", high="green")
x11()
imageplot(log2(RG$Rb[,3]),RG$printer,low="white", high="red")
x11()
imageplot(log2(RG$Gb[,3]),RG$printer,low="white", high="green")
x11()
imageplot(log2(RG$Rb[,4]),RG$printer,low="white", high="red")
x11()
imageplot(log2(RG$Gb[,4]),RG$printer,low="white", high="green")
x11()
imageplot(log2(RG$Rb[,5]),RG$printer,low="white", high="red")
x11()
imageplot(log2(RG$Gb[,5]),RG$printer,low="white", high="green")
x11()
imageplot(log2(RG$Rb[,6]),RG$printer,low="white", high="red")
x11()
imageplot(log2(RG$Gb[,6]),RG$printer,low="white", high="green")
x11()
```

```
imageplot(log2(RG$Rb[,7]),RG$printer,low="white", high="red")
x11()
imageplot(log2(RG$Gb[,7]),RG$printer,low="white", high="green")
x11()
imageplot(log2(RG$Rb[,8]),RG$printer,low="white", high="red")
x11()
imageplot(log2(RG$Gb[,8]),RG$printer,low="white", high="green")
x11()
imageplot(log2(RG$Rb[,9]),RG$printer,low="white", high="red")
x11()
imageplot(log2(RG$Gb[,9]),RG$printer,low="white", high="green")
x11()
imageplot(log2(RG$Rb[,10]),RG$printer,low="white", high="red")
x11()
imageplot(log2(RG$Gb[,10]),RG$printer,low="white", high="green")
x11()
imageplot(log2(RG$Rb[,11]),RG$printer,low="white", high="red")
x11()
imageplot(log2(RG$Gb[,11]),RG$printer,low="white", high="green")

#The following lines of code are the background correction for each of the channels. The
#normexp method for correction is the recommended method. The offset adds a factor of 50 to
#each of the values so that the log transformed values at the lower ranges are shrunk towards
#zero, giving a better looking fit.
RGb<-backgroundCorrect(RG, method="normexp",offset=50)
x11()
plotMA(RGb,array=1,legend=T,ylim=c(-6,6))
plotMA(RGb,array=2,legend=T,ylim=c(-6,6))
plotMA(RGb,array=3,legend=T,ylim=c(-6,6))
plotMA(RGb,array=4,legend=T,ylim=c(-6,6))
plotMA(RGb,array=5,legend=T,ylim=c(-6,6))
plotMA(RGb,array=6,legend=T,ylim=c(-6,6))
plotMA(RGb,array=7,legend=T,ylim=c(-6,6))
plotMA(RGb,array=8,legend=T,ylim=c(-6,6))
plotMA(RGb,array=9,legend=T,ylim=c(-6,6))
plotMA(RGb,array=10,legend=T,ylim=c(-6,6))
plotMA(RGb,array=11,legend=T,ylim=c(-6,6))

#Normalize within arrays
nonDE<-grep("Calib*",RG$genes[,"Name"])
MA<-normalizeWithinArrays(RGb,
method="loess",iterations=5,bc.method="subtract",offset=0)
x11()
plotMA(RGb,array=1,legend=T,ylim=c(-5,5))
```

```
plotMA(RGb,array=2,legend=T,ylim=c(-6,6))
plotMA(RGb,array=3,legend=T,ylim=c(-5,5))
plotMA(RGb,array=4,legend=T,ylim=c(-5,5))
plotMA(RGb,array=5,legend=T,ylim=c(-5,5))
plotMA(RGb,array=6,legend=T,ylim=c(-6,6))
plotMA(RGb,array=7,legend=T,ylim=c(-5,5))
plotMA(RGb,array=8,legend=T,ylim=c(-5,5))
plotMA(RGb,array=9,legend=T,ylim=c(-5,5))
plotMA(RGb,array=10,legend=T,ylim=c(-6,6))
plotMA(RGb,array=11,legend=T,ylim=c(-5,5))


#Draw boxplots for the background distribution
x11()
boxplot(data.frame(log2(RG$Rb)),main="Red background",notch=TRUE)
x11()
boxplot(data.frame(log2(RG$Gb)),main="Green background",notch=TRUE)
#Normalize between arrays
MAbet<-normalizeBetweenArrays(MA,method='Aquantile')
x11()
plotDensities(MAbet)
#Start to do Bayesian stats and draw top table
design<-modelMatrix(targets,ref='Unfed')
design
MAbet$M<-MAbet$M[MAbet$genes$Status=="cDNA",]
MAbet$A<-MAbet$A[MAbet$genes$Status=="cDNA",]
MAbet$genes<-MAbet$genes[MAbet$genes$Status=="cDNA",]
i<-order(MAbet$genes$ID)
MAbet<-MAbet[i,]
#Calculate the consensus correlation between the arrays
cor<-duplicateCorrelation(MAbet, ndups=4, design)
cor
cor$concensus.correlation
#start to fit Bayesian stats to the data and create the top table
fit<-lmFit(MAbet,design,ndups=4,correlation=cor$consensus.correlation)
fit<-eBayes(fit)
UnfedvsBloodfed<-(topTable(fit,adjust="fdr",number="all"))
UnfedvsBloodfed
#Save the toptable which is a file of all the data for each target
write.table(UnfedvsBloodfed,file="C:/Documents and Settings/belindas/Desktop/Gambiae
GPR Files for Analysis/toptable.txt",sep="\t",row.names=FALSE)
volcanoplot(fit,coef=1, highlight=15,names=fit$genes$Name, main= "Bloodfed vs Unfed")
```

APPENDIX D

Optimisation of *RPL8* for Use as a Reference Gene

Although amplification using the *RPL8* primer resulted in good amplification curves (Figure D1), this primer pair could not be included in the analysis. The primer pair targeting the *RPL8* gene yielded more than one amplification product, evidenced by double peaks in the melt curve analysis (Figure D2).

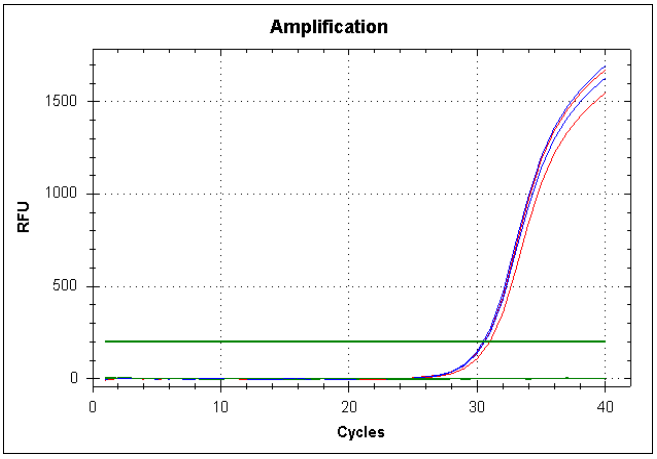


Figure D1. Amplification curves for the *RPL8* primer pair. (Red lines = blood fed *An. funestus*; blue lines = unfed *An. funestus*; green lines = no template control).

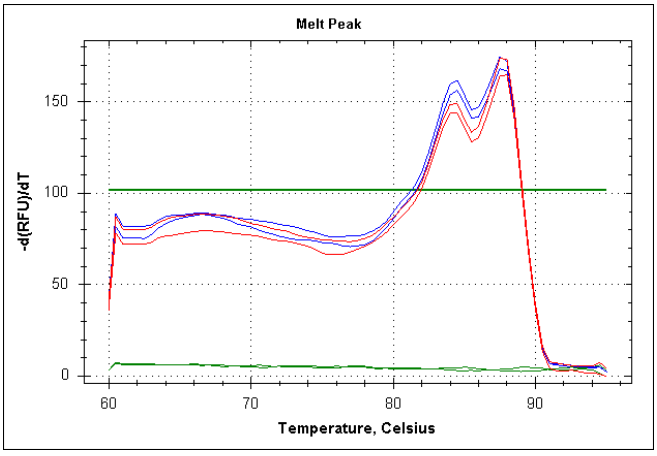


Figure D2. Melt curve for the *RPL8* primer pair. Note the double peaks indicating the presence of two amplification products. (Red lines = blood fed *An. funestus*; blue lines = unfed *An. funestus*; green lines = no template control).

APPENDIX E

Collins DNA Extraction Method

Detailed below is the DNA extraction protocol used in Chapter 4. This protocol is described in Collins *et al*. (1987).

All pipette tips, 1.5ml microcentrifuge tubes and pestles for grinding were autoclaved for 20 minutes at 121°C. A heating block was set 70°C and allowed to reach the temperature at least half an hour before use. Grinding buffer was prepared by mixing 1.6 ml 1M NaCl, 1.095 g sucrose, 2.4 ml 0.5 M EDTA, 1 ml 10% SDS, 2 ml 1 M Tris-Cl (pH 8.6). The final volume was made up to 20ml using dH_2O.

Using sterile forceps, the specimen placed into a 1.5ml microcentrifuge tube. The mosquito was ground dry, as finely as possible. The pestle was not removed from the tube, whilst 200 µl grinding buffer was added. The specimen was ground again until very fine. The pestle was removed, ensuring that no mosquito parts remained on it. Whilst removing the pestle, it was ensured that not too much liquid was removed at the same time.

The closed tubes were then incubated in the heating block at 70°C for 30 minutes, never longer than 1 hour. After incubating, the tubes were removed from the heating block and 28 µl 8 M KAc was added. The mixtures were vortex by tapping the tube. The tubes were then placed on ice for 30 minutes, followed by centrifugation for 20 minutes at 13 000rpm. The

supernatant was pipetted off into a clean 1.5ml microcentrifuge tube taking care not to disturb the pellet. The pellets were then discarded.

Ice cold 100% ethanol (400 µl) was added to the supernatent and mixed by inverting the tube. The tubes were incubated overnight at -20°C. The samples were then centrifuged for 30 minutes, 13 000 rpm. The 100% ethanol was then pipetted off and discarded, taking care not to disturb the pellet. Traces of salt were removed by adding 200 µl ice cold 70% ethanol to the tube containing the pellet. The tubes were then centrifuged at 13 000 rpm for 30 minutes. The 70% ethanol supernatant was then pipetted off and the pellet allowed to air dry on the lab bench. The extracted DNA was then re-suspended in 200 µl of 1 x TE.

An extraction positive and negative control was included in every extraction. Positive controls were specimens of known identity, harvested from laboratory colonies.

Multiplex PCR for the Identification of *An. funestus* Group

This PCR was run according to the method of Koekemoer *et al*. (2002). The primers used in the PCR are detailed in Table E1 below. The 12.5 µl PCR reaction mixtures were made up as follows: 1.25 µl of 10x PCR buffer (500 mM KCl, 100mM Tris-HCl, pH 8.3), 1.5mM $MgCl_2$, 3.3 pmol/primer of each primer, 200µM of each dNTP, and 0.5 units taq DNA polymerase. Approximately 0.5 to 1 µl of DNA was added to each reaction tube. Positive controls for each species were included, using specimens that have been previously identified. A no template control was included as well as the DNA extraction controls.

The PCR cycles were programmed as follow: initial denaturation at 94°C for 2 minutes; followed by 30 cycles of 94°C for 30 seconds, 45°C for 30 seconds, and 72°C for 40 seconds. A final extension step was included at 72°C for 5 minutes. The PCR products were electrophoresed on a 2.5% TAE agarose gel stained with ethidium bromide.

Table E1. Table of primers used in the *An. funestus* multiplex PCR (Koekemoer *et al*., 2002).

Primer	Sequence	Tm (°C)
UV	TGT GAA CTG CAG GAC ACA T	55.3
FUN	GCA TCG ATG GGT TAA TCA TG	52.4
VAN	TGT CGA CTT GGT AGC CGA AC	58.0
RIV	CAA GCC GTT CGA CCC TGA TT	58.8
PAR	TGC GGT CCC AAG CTA GGT TC	60.5
LEES	TAC ACG GGC GCC ATG TAG TT	60.2

Plasmodium falciparum **Sporozoite ELISA**

The ELISA protocol used to detect *P. falciparum* sporozoites is that of Wirtz *et al*. (1987). Briefly, the following solutions were prepared in advance: blocking buffer (0.5% casein boiled in 100 ml 0.1N NaOH until dissolved, 0.002% phenol red, made up in 1 x Phosphate Buffered Saline, pH 7.4); grinding buffer (blocking buffer containing 0.5% Nonidet P-40 (NP-40, Sigma); PBS-Tween (1x PBS with 0.005% Tween-20 (Sigma)).

The head and thorax of each mosquito was placed in a clean 1.5ml microcentrifuge tube. Using a clean pestle for each specimen, the samples were ground in 50 µl of grinding buffer. Any mosquito parts remaining on the pestle were rinsed off into the microcentrifuge tube by pipetting a further 150 µl grinding buffer over the pestle. Homogenates were frozen at -70°C until ready for use. An antibody plate was prepared the day before the ELISA was run using a

U-shaped 96 well ELISA plate. For each 96 well plate, fresh antibody solution was prepared (5 ml PBS with 40 µl MAb PF 2A10 (KPL, Maryland USA, Cat # 3701242)). Using a multichannel pipette, 50 µl of antibody solution was dispensed into each well on the plate. The plate was wrapped with plastic wrap and foil and incubated at 4°C overnight. Following this, the wells were cleared of antibody solution and filled with blocking buffer and incubated at room temperature for 1 hour.

During this incubation step, the mosquito homogenates were thawed on ice and the positive and negative controls prepared. The seven negative controls consisted of freshly ground mosquitoes acquired from colony material from the VCRU insectary. Positive control was prepared by mixing 50 µl blocking buffer with 1 µl of Pf2+ antigen (donation from RA Wirtz).

After the 1 hour incubation, the wells were aspirated. Positive control was added to the first well on the top row of the plate (A1). Negative controls were added to the last seven wells, on the last row of the plate (H6 to H12). Samples (50 µl) were added from well A2 onwards. The plate was allowed to incubate for 2 hours at room temperature.

The wells were then washed twice using PBS-Tween. A peroxidase mix of 5.6 ml blocking buffer and 10 µl peroxidase Pf2A10 (KPL, Maryland USA Cat # 3700242) was prepared and 50 µl aliquoted into each well. The plate was then incubated for 1 hour at room temperature. The wells were then washed with PBS-Tween solution, three to four times. Following this, 100 µl of ABTS peroxidase substrate (KPL, Maryland USA, Cat # 506601) was added to each well. The plate was covered in foil and allowed to incubate at room temperature for 30 to 60

minutes. The absorbance of the solutions was then read at a wavelength of 414 nm, using an ELISA plate scanner (Multiskan Ascent, Thermo Electron Corp.).

The cutoff value for the absorbance of reactive samples was calculated by finding twice the average absorbance of the negative controls. Samples that displayed absorbances equal to or greater than this cutoff value were flagged as potential positive samples. These specimens were then run through a second ELISA to confirm positivity for *P. falciparum* parasites.

APPENDIX F

Nanodrop results of MalaF DNA

Sample ID	ng/ul	A260	A280	260/280
Unidentified 1	20.16	0.403	0.203	1.99
Unidentified 2	25.20	0.504	0.249	2.02
Unidentified 3	21.85	0.437	0.196	2.23
Unidentified 4	16.04	0.321	0.121	2.66
Unidentified 5	49.85	0.997	0.709	1.41
Unidentified 6	18.45	0.369	0.164	2.26
Unidentified 7	10.39	0.208	0.124	1.68
dH2O	-3.76	-0.075	-0.049	1.53
ID 1	9.58	0.192	0.094	2.04
ID 2	81.26	1.625	0.791	2.05
ID 3	9.16	0.183	0.102	1.80
ID 4	13.74	0.275	0.144	1.90
ID 5	19.94	0.399	0.221	1.81
ID 6	61.14	1.223	0.604	2.02

Figure F1. Nanodrop results of the seven unamplified MalaF specimens and a set of randomly chosen identifiable MalaF specimens.

The DNA extractions of the unidentifiable MalaF specimens fell within the range of 10 to 50ng.ul^{-1}. This is a good yield of DNA which should have given amplicon in the MalaF PCR. A small number of identified MalaF specimens were randomly chosen and nanodropped. These samples exhibited a variation in DNA concentrations from approximately 9 to 80ng.ul^{-1}.

The absorbance ratio of a DNA sample measured at 260nm/280nm, gives a good indication of the purity of the sample. A ratio above 1.8 is considered acceptable, where as a ratio below 1.8 is considered to be contaminated with protein and/or phenol (Sambrook *et al*., 1989). Only two of the unidentified specimens had low 260/280 ratios: specimen 5 and 7. Both of these specimens did however result in amplicon when tested using the in-house PCR.

REFERENCES

Amenya DA, Naguran R, Lo TC, Ranson H, Spillings BL, Wood OR, Brooke BD, Coetzee M and Koekemoer LL. 2008. Overexpression of a cytochrome P450 (*CYP6P9*) in a major African malaria vector, Anopheles funestus, resistant to pyrethroids. *Insect Molecular Biology* **17**: 19 – 25.

Andersen CL, Jensen JL and Ørntoft TF. 2004. Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. *Cancer Research* **64**: 5245 – 5250.

Baker M. 2011. qPCR: Quicker and Easier but Don't be Sloppy. *Nature Methods* **8:** 207 – 212.

Baldridge GD and Feyereisen R. 1989. Blood meal and cytochrome P-450 monooxygenases in the Northern house mosquito, *Culex pipiens*. *Pesticide Biochemistry and Physiology* **25**: 407 – 413.

Barbulovic-Nad I, Lucente M, Sun Y, Zhang M, Wheeler AR and Bussman M. 2006. Bio-Microarray Fabrication Techniques – A Review. *Critical Reviews in Biotechnology* **26**: 237-259.

Bass C, Williamson MS and Field LM. 2008. Development of a multiplex real-time PCR assay for identification of members of the *Anopheles gambiae* species complex. *Acta Tropica* **107**: 50 - 53.

Bowler, PJ. 1992. *The Earth Encompassed: A History of the Environmental Sciences*. WW Norton & Company, New York.

Boyce R. 1910. *Mosquito or Man? The Conquest of the Tropical World.* (2nd Edition) John Murray, Arlbemarle St, London.

Brazma A, Hingamp P, Quackenbush J, Sherlock G, Spellman P, Stoeckert C, Aach J, Ansorge W, Ball C, Causton H, Gaasterland T, Glenisson P, Holstege F, Kim I, Markowitz V, Matese J, Parkinson H, Robinson A, Sarkans U, Schulze-Kremer S, Stewart J, Taylor R, Vilo J and Vingron M. 2001. Minimum information about a microarray experiment (MIAME) – toward standards for microarray data. *Nature Genetics* **29**: 365 - 371.

Brogdon WG and McAllister JC. 1998. Simplification of adult mosquito bioassays through use of time-mortality determinations in glass bottles. *Journal of the American Mosquito Control Association* **14**: 159 – 167.

Brooke BD, Kloke G, Hunt RH, Koekemoer LL, Temu EA, Taylor ME, 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.

Bustin S, Benes V, Garson J, Hellemans J, Huggett J, Kubista M, Mueller R, Nolan T, Pfaffl M, Shipley G, Vandesompele J and Wittwer C. 2009. The MIQE Guidelines: *M*inimum *I*nformation for Pulication of *Q*uantitative Real-Time PCR *E*xperiments. *Clinical Chemistry* **55**: 611 - 622.

Bustin SA. 2010. Why the need for qPCR publication guidelines?-The case for MIQE. *Methods* **50**: 217 - 226.

Casimiro SL, Hemingway J, Sharp BL and Coleman M. 2007. Monitoring the Operational Impact of Insecticide Usage for Malaria Control on *Anopheles funestus* from Mozambique. *Malaria Journal* **6**: 142.

Chervoneva I, Li Y, Schulz S, Croker S, Wilson C, Waldman SA, Hyslop T. 2010. Selection of Optimal Reference Genes for Normalization in Quantitative RT-PCR. *BMC Bioinformatics* **11**: 253.

Chima RI, Goodman CA and Mills A. 2003. The Economic Impact of Malaria in Africa: a Critical Review of the Evidence. *Health Policy* **63**: 17 – 36.

Christian RN, Matambo TS, Spillings BL, Brooke BD, Coetzee M and Koekemoer LL. 2011a. Age Related Pyrethroid Resistance is not a Function of P450 gene expression in the Major African malaria vector, *Anopheles funestus* (Diptera: Culicidae). *Genetics and Molecular Research* **10**: 3220 – 3229.

Christian RN, Strode C, Ranson H, Coetzer N, Coetzee M and Koekemoer LL. 2011b. Microarray analysis of a pyrethroid resistant African malaria vector, *Anopheles funestus*, from southern Africa. *Pesticide Biochemistry and Physiology* **99**: 140 – 147.

Coosemans M and Carnevale P. 1995. Malaria Vector Control: A Critical Review on Chemical Methods and Insecticides. *Annales de la Société Belge de Médecine Tropicale* **75**: 13 – 31.

Coetzee M, Estrada-Franco JG, Wunderlich CA and Hunt RH. 1999. Cytogenetic evidence for a species complex within *Anopheles pseudopunctipennis* Theobald (Diptera: Culicidae). *American Journal of Tropical Medicine and Hygiene* **60**: 649 – 653.

Coetzee M and Fontenille D. 2004. Advances in the Study of *Anopheles funestus*, Major Vector of Malaria in Africa. *Insect Biochemistry and Molecular Biology* **34**: 599 – 605.

Coetzee M and Hunt RH. 1998. Malaria at its Southern-Most Fringe in Africa. *Research and Reviews in Parasitology* **58**: 175 – 179.

Coleman AW. 2003. ITS2 is a Double-edged Tool for Eukaryote Evolutionary Comparisons. *Trends in Genetics* **19**: 370 – 375.

Coleman AW. 2009. Is there a Molecular Key to the Level of "Biological Species" in Eukaryotes? A DNA Guide. *Molecular Phylogenetics and Evolution* **50**: 197 – 203.

Collins FH, Mendez MA, Rasmussen MO, Meheffey PC, Besansky NJ and Finnerty V. 1987. A ribosomal RNA gene probes differentiates members of the *Anopheles gambiae* complex. *American Journal of Tropical Medicine and Hygiene* **37**: 37 – 41.

Collins FH, Paskewitz SM and Finnerty V. 1989. *Advances in Disease Vector Research: Ribosomal Genes of the Anopheles Gambiae Species* Complex. Volume 6. Springer-Verlag, New York.

Coluzzi M and Sabatini A. 1967. Cytogenetic observations on species A and B of the *Anopheles gambiae* complex. *Parassitologia* **9**: 73 – 88.

Coluzzi M and Sabatini A. 1968. Cytogenetic observations on species C of the *Anopheles gambiae* complex. *Parassitologia* **10**: 155 – 166.

Coluzzi M and Sabatini A. 1969. Cytogenetic observations on the salt water species *Anopheles merus* and *Anopheles melas*, of the Gambiae complex. *Parassitologia* **11**: 155 – 166.

Costantini C, Sagnon, N'F, Ilboudo-Sanogo E, Coluzzi M and Boccolini D. 1999. Chromosomal and bionomic heterogeneities suggest incipient speciation in *Anopheles funestus* from Burkina Faso. *Parassitologia* **41**: 595 – 611.

Couhet A, Simard F, Toto J-C, Kengne P, Coetzee M and Fontenille D. 2003. Species Identification within the *Anopheles funestus* group of malaria vectors in Cameroon and evidence for a new species. *American Journal of Tropical Medicine and Hygiene* **69**: 200 – 205.

Crossley SA. 1974. Changes in Mating Behaviour Produced by Selection for Ethologiccal Isolation Between Ebony and Vestigial Mutants of *Drosophila melanogaster*. *Evolution* **28**: 631 – 647.

Cuamba N, Morgan J, Irving H, Steven A and Wondji CS. 2010. High Level of Pyrethroid Resistance in an *Anopheles funestus* population of the Chokwe District in Mozambique. *PLoS One* **5**: e11010.

David JP, Strode C, Vontas J, Nikou D, Vaughn A, Pignatelli PM, Louis C, Hemingway J and Ranson H. 2005. The *Anopheles gambiae* detoxification chip: A highly specific microarray to study metabolic-based insecticide resistance in malaria vectors. *Proceedings of the National Academy of Sciences USA* **102**: 4080 – 4084.

Davidson G. 1956. Insecticide Resistance in *Anopheles gambiae* Giles. *Nature* **178**: 705-706.

Davidson G, Paterson H, Coluzzi M, Mason G and Micks D. 1967. *Genetics of Insect Vectors of Disease: The Anopheles gambiae complex*. Wright J, Pal R, eds. Elsevier Publishing Company, London.

Day JF and Van Handel E. 1986. Differences between the nutritional reserves of laboratory-maintained and field collected adult mosquitoes. *Journal of the American Mosquito Control Association* **2**: 154 – 157.

De Meillon B. 1947. New records and species of biting insects from the Ethiopian Region II. *Journal of the Entomological Society of Southern Africa* **10**: 110 – 124.

De Meillon B. 1986. The Control of Malaria with Special Reference to the Contributions made by the Staff of the South African Institute of Medical Research. *South African Medical Journal* **76**: 67 – 69.

De Meillon B, Van Eeden GJ, Coetzee L, Coetzee M, Meiswinkel R, Du Toit CLN and Hansford CF. 1977. Observations on a species of the *Anopheles funestus* Subgroup, a suspected exophilic vector of malaria parasites in North Eastern Transvaal, South Africa. *Mosquito News* **37**: 657 – 661.

Derveaux S, Vandesompele J and Hellemans J. 2010. How to do Successful Gene Expression Analysis using Real-time PCR. *Methods* **50**: 227 – 230.

Dimopoulos G, Christophides GK, Meister S, Schultz J, White KP, Barillas-Mury C and Kafatos FC. 2002. Genome Expression Analysis of *Anopheles gambiae*: Responses to injury, bacterial challenge and malaria infection. *Proceedings of the National Academy of Sciences USA* **99**: 8814 – 8819.

Djouaka RF, Bakare AA, Coulibaly ON, Akogbeto MC, Ranson H, Hemingway J and Strode C. 2008. Expression of the cytochrome P450s, CYP6P3 and CYP6M2 are significantly elevated in multiple pyrethroid resistant populations of *Anopheles gambiae s.s.* from Southern Benin and Nigeria. *BMC Genomics* **9**: 538.

Dobzhansky T. 1951. *Genetics and the Origin of Species*. Columbia University Press, New York.

Dobzhansky T. 1937. Genetic Nature of Species Differences. *American Nauralist* **71**: 404 – 420.

Evans AM and Symes CB. 1937. *Anopheles funestus* and its allies in Kenya. *Annals of Tropical Medicine and Parasitology* **31:** 105 – 111.

Evans, AM, 1938. *Mosquitoes of the Ethiopian Region*. Bernard Quartich Ltd; Dulau & Co Ltd; and The Oxford University Press, London.

ffrench-Constant RH, Steichen JC, Rocheleau TA, Aronstein K and Roush RT. 1993. A Single-amino Acid Substitution in a γ-Aminobutyric Acid Subtype A Receptor Locus is Associated with Cyclodiene Insecticide Resistance in *Drosophila* populations. *Proceedings of the National Academy of Science, USA* **90**: 1957 – 1961.

Foley DH and Bryan JH. 1991. *Anopheles annulipes* Walker (Diptera: Culicidae) at Griffith New South Wales. 1. Two sibling species in sympatry. *Australian Journal of Entomology* **30**: 109 – 112.

Frizzi G. 1947. Salivary gland chromosomes of *Anopheles. Nature* **160**: 226 – 227.

Frizzi G. 1953. Étude cytog'enetique d'anopheles maculipennis en Italie. Extension des recherches 'a d'autres esp'eces d'anoph'eles. *Bulletin of the World Health Organization* **9**: 335 – 344.

Garnham PCC. 1966. *Malaria Parasites and Other Haemosporidia*. Blackwell, Oxford England.

Garros C, Harbach RE and Manguin S. 2005. Morphological Assessment and Molecular Phylogenetics of the Funestus and Minimus groups of *Anopheles* (*Cellia*). *Morphology, Systematics, Evolution* **42**: 522 – 536.

Garros C, Koekemoer LL, Kamau L, Awolola TS, Van Bortel W, Coetzee M, Coosemans M and Manguin S. 2004. Restriction fragment length polymorphism method for the identification of major African and Asian malaria vectors within the *Anopheles funestus* and *An. minimus* groups. *American Journal of Tropical Medicine and Hygiene* **70**: 260 – 265.

Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J, Hornik K, Hothorn T, Huber W, Iacus S, Irizarry R, Leisch F, Li C, Maechler M, Rossini AJ, Sawitzki G, Smith C, Smyth GK, Tierney L, Yang JY and Zhang J. 2004. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biology* **5**: R80.

Gillies MT and Coetzee M. 1987. *A Supplement to the Anophelinae of Africa South of the Sahara (Afrotropical Region).* Publications of the South African Institute for Medical Research, Johannesburg, no.55.

Gillies MT and De Meillon B. 1968. *The Anophelinae of Africa South of the Sahara.* Publications of the South African Institute for Medical Research, Johannesburg, no.54.

Govere JM, Durrheim DN and Kunene S. 2002. Malaria trends in South Africa and Swaziland and the introduction of synthetic pyrethroids to replace DDT for malaria vector control. *South African Journal of Science* **98**: 19 – 21.

Green CA, 1982. Cladistic analysis of chromosome data (*Anopheles* (*Cellia*) Myzomia). *Journal of Heredity* **73**: 2 – 11.

Green CA and Hunt RH. 1980. Interpretations of variation in ovarian polytene chromosomes of *Anopheles funestus* Giles, *A. parensis* Gillies and *A. aruni*? *Genetica* **51**: 187 – 195.

Gregory R, Darby AC, Irving H, Coulibaly MB, Hughes M, Koekemoer LL, Coetzee M, Ranson H, Hemingway J, Hall N and Wondji CS. 2011. A *De novo* Expression Profiling of *Anopheles funestus*, Malaria Vector in Africa, Using 454 Pyrosequencing. *PLoS ONE* **6**: e17418.

Guelbeogo WM, Grushko O, Boccolini D, Ouédraogo PA, Besansky NJ, Sagnon NF and Constantini C. 2005. Chromosomal evidence of incipient speciation in the Afrotropical malaria mosquito *Anopheles funestus*. *Medical and Veterinary Entomology* **19**: 458 – 469.

Gwadz R. 1969. Regulation of Blood Meal Size in the Mosquito. *Journal of Insect Physiology* **15**: 2039 – 2044.

Harbach RE. 2004. The classification of genus *Anopheles* (Diptera: Culicidae): a working hypothesis of phylogenetic relationships. *Bulletin of Entomological Research* **94**: 537 – 553.

Hargreaves K, Hunt RH, Brooke BD, Mthembu J, Weeto MM, Awolola TS 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 LL, Brooke BD, Hunt RH, Mthembu J and Coetzee M. 2000. *Anopheles funestus* resistant to pyrethroid insecticides in South Africa. *Medical and Veterinary Entomology* **14**: 181 – 189.

Hackett BJ, Gimnig J, Guelbeogo W, Costantini C, Koekemoer LL, Coetzee M, Collins FH and Besansky NJ. 2000. Ribosomal DNA internal transcribed spacer (ITS2) sequences differentiate *Anopheles funestus* and *An. rivulorum*, and uncover a cryptic taxon. *Insect Molecular Biology* **9**: 369 – 374.

Hemingway J and Ranson H. 2000. Insecticide Resistance in Insect Vectors of Human Disease. *Annual Review of Entomology* **45**: 371 – 391.

Hemingway J, Hawkes NJ, McCarrol L and Ranson H. 2004. The molecular basis of insecticide resistance in mosquitoes. *Insect Biochemistry and Molecular Biology* **34**: 653 – 665.

Hey J. 2001. The mind of the species problem. *Trends in Ecology and Evolution* **16**: 326 – 329.

Huho BJ, Ng'habi KR, Killeen GR, Nkwegulila G, Knols BGJ, and Ferguson HM. 2007. Nature beats nurture: a case study of the physiological fitness of free-living and laboratory-reared male *Anopheles gambiae s.l. The Journal of Experimental Biology* **210**: 2939 – 2947.

Hunt RH, Coetzee M and Fettene M. 1998. The *Anopheles gambiae* complex: a new species from Ethiopia. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **92**: 231 – 235.

Hunt RH, Brooke BD, Pillay C, Koekemoer LL and Coetzee M. 2005. Laboratory selection for and characteristics of pyrethroid resistance in the malaria vector *Anopheles funestus*. *Medical and Veterinary Entomology* **19**: 271 – 275.

Hunt RH, Edwardes M and Coetzee M. 2010. Pyrethroid resestance in southern African *Anopheles funestus* extends to Likoma Island in Lake Malawi. *Parasites and Vectors* **3**: 122.

Jochim RC, Teixeira CR, Laughinghouse A, Mu J, Oliveira F, Gomes RB, Elnaiem DE and Valenzuela JG. 2008. The midgut transcriptome of *Lutzomyia longipalpis*: comparative analysis of cDNA libraries from sugar-fed, blood-fed, post-digested and *Leishmania infantum chagasi*-infected sand flies. *BMC Genomics* **9**: 15.

Kaiser PE. 1988. Cytotaxonomy as a tool for identification of siblings of the *Anopheles quadrimaculatus* complex. *Florida Entomologist* **71**: 311 – 323.

Kaiser ML, Koekemoer LL, Coetzee M, Hunt RH and Brooke BD. 2010 Staggered larval time-to-hatch and insecticide resistance in the major malaria vector *Anopheles gambiae* S form. *Malaria Journal* **9**: 360.

Knight GR, Robertson A and Waddington CH. 1956. Selection for Sexual Isolation Within a Species. *Evolution* **10**: 14 – 22.

Klug WS and Cummings MR. *Concepts of Genetics*. 5th Edition. Prentice Hall International, Inc. USA. 1997.

Koekemoer LL, Coetzee M and Hunt RH. 1998. HpaII endonuclease distinguishes between two species in the *Anopheles funestus* group. *Insect Molecular Biology* **7**: 273 – 277.

Koekemoer LL, Kamau L, Hunt RH and Coetzee M. 2002. A cocktail polymerase chain reaction assay to identify members of the *Anopheles funestus* (Diptera: Culicidae) group. *American Journal of Tropical Medicine and Hygiene* **6**: 804 – 811.

Koekemoer LL, Lochouarn L, Hunt RH and Coetzee M. 1999. Single-strand conformation polymorphism analysis for identification of four members of the *Anopheles funestus* (Diptera: Culicidae) group. *Journal of Medical Entomology* **36**: 125 – 130.

Koekemoer LL, Misiani EA, Hunt RH, Kent RJ, Norris DE and Coetzee M. 2008. Cryptic species within *Anopheles longipalpis* from southern Africa and phylogenetic comparison with members of the *An. funestus* group. *Bulletin of Entomological Research* **99**: 41 – 49.

Koopman KF. 1950. Natural selection for Reproductive Isolation Between *Drosophila pseudoobscura* and *Drosophila persimilis*. *Evolution* **4:** 135 – 148.

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

Krzywinski J, Sangaré D and Besansky NJ. 2005. Satellite DNA From the Y Chromosome of the Malaria Vector *Anopheles gambiae*. *Genetics* **169**: 185 - 196.

le Sueur D, Sharp BL, Appleton CC. 1993. Historical perspective of the malaria problem in Natal with emphasis on the period 1928 – 1932. *South African Journal of Science* **89**: 232 – 239.

Lochouarn L, Dia I, Boccolini D, Coluzzi M and Fontenille D. 1998. Bionomical and cytogenetic heterogeneities of *Anopheles funestus* in Senegal. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **92**: 607 – 612.

Lyell C. 1832. *Principles of Geology, being an attempt to explain the former changes of the Earth's surface, by reference to causes now in operation*. Volume 2. William Clawes, Stamford Street, London.

Lu Y, Huggins P and Bar-Joseph Z. 2009. Cross species analysis of microarray expression data. *Bioinformatics* **25**: 1476 – 1483.

Maharaj R, Mthembu DJ and Sharp BL. 2005. Impact of DDT re-introduction on malaria transmission on KwaZulu-Natal. *South African Medical Journal* **95**: 871 – 874.

Manga L, Toure A and Shililu J. 2004. Implementation of Integrated Vector Management in the WHO-Africa Region: Progress Report 2000 – 2003. *Joint Publication 10 of the Vector Biology and Control Unit, Division of Prevention and Control of Communicable Diseases, WHO Regional Office for Africa and the Environmental Health Project of USAID*.

Marinotti O, Calvo E, Nguyen QK, Dissanayake S, Ribeiro JMC and James AA. 2006. Genome-wide analysis of gene expression in adult *Anopheles gambiae*. *Insect Molecular Biology* **15**: 1 – 12.

Martinez-Torres D, Devonshire AL and Williamson MS**.** 1997. Molecular Studies of Knockdown Resistance to Pyrethroids: Cloning of Domain II Sodium Channel Gene Sequences from Insects. *Pesticide Science* **51**: 265 – 270.

Martinez-Torres D, Chandre F, Williamson MS, Darriet F, Bergé JB, Devonshire AL, Guillet P, Pasteur N and Pauron D. 1998. Molecular Characterization of Pyrethroid Knockdown Resistance (*kdr*) in the Major Malaria Vector *Anopheles gambiae s.s. Insect Molecular Biology* **7**: 179 – 184.

Matambo TS, Paine MJ, Coetzee M and Koekoemoer LL. 2010. Sequence characterization of cytochrome P450 *CYP6P9* in pyrethroid resistant and susceptible *Anopheles funestus* (Diptera: Culicidae). *Genetics and Molecular Research* **9**: 554 – 564.

Mathews DH, Sabina J, Zuker M and Turner DH. 1999. Expanded Sequence Dependence of Thermodynamic Parameters Improves Prediction of RNA Secondary Structure. *Journal of Molecular Biology*. **288**: 911 – 940.

Mayr E. 1963. *Animal Species and Evolution*. Harvard University Press, Cambridge, Massachusetts.

Mayr E. 1969. *Principles of Systematic Zoology*. McGraw-Hill Book Company, New York.

McAbee R, Kang K-D, Stanich M, Christiansen J, Wheelock C, Inman A, Hammock B and Cornel AJ. 2003. Pyrethroid tolerance in *Culex pipiens pipiens* var *molestus* from Marin County, California. *Pesticide Management Science* **60**: 359 – 368.

Michel AP, Guelbeogo WM, Grushko O, Schemerhorn BJ, Kern, M, Willard MB, Sagnon N'F, Constantini C and Besansky NJ. 2005. Molecular differentiation between chromosomally defined incipient species of *Anopheles funestus*. *Insect Molecular Biology* **14**: 375 – 387.

Michel AP, Ingracsi MJ, Schemerhorn BJ, Kern M, Le Goff G, Coetzee M, Elissa N, Fontenille D, Vulule J, Lehmann T, Sagnon N'F, Costantini C and Besansky NJ. 2005. Rangewise population genetics structure of the African malaria vector *Anopheles funestus*. *Molecular Ecology* **14**: 4235 – 4248.

Miller LH and Greenwood B. 2002. Viewpoint: Malaria – a Shadow over Africa. *Science* **298**: 121 – 122.

Mitchell SE and Seawright JA. 1989. Recombination between the X and Y chromosomes in *Anopheles quadrimaculatus* Species A. *Journal of Heredity* **80**: 496 – 499.

Moody DE, Zou Z and McIntyre LM. 2002. Cross-species hybridisation of pig RNA to human nylon microarrays. *BMC Genomics* **3**: 27.

Mouatcho JC, Munhenga G, Hargreaves K, Brooke BB, Coetzee M and Koekemoer LL. 2009. Pyrethroid resistance in a major African malaria vector *Anopheles arabiensis* from Mamfene, northern KwaZulu-Natal, South Africa. *South African Journal of Science* **105**: 127 – 131.

Müller T, Philippi N, Dandekar T, Schultz J and Wolf M. 2007. Distinguishing Species. *RNA* **13**: 1469 – 1472.

Müller P, Donnelly MJ and Ranson H. 2007. Transcription profiling of a recently colonised pyrethroid resistant *Anopheles gambiae* strain from Ghana. *BMC Genomics* **8**: 36.

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

Munhenga G and Koekemoer LL. 2011. Differential Expression of cytochrome P450 gene in a laboratory selected *Anopheles arabiensis* colony. *African Journal of Biotechnology* **10**: 12711 – 12716.

Mutero A, Pralavorio M, Bride J-M and Fournier D. 1994. Resistance-associated Point Mutations in Insecticide-Insecticide Acetylcholinesterase. *Proceedings of the National Academy of Science, USA* **91**: 5922 – 5926.

Muyal JP, Singh SK and Fehrenbach H. (2008). DNA-Microarray Techonology: Comparison of methodological factors of recent technique towards gene expression profiling. *Critical Reviews in Biotechnology* **28**: 239 - 251

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.

N' Guessan RN, Darriet F, Guillet P, Carnevale P, Traore-Lamizana M, Corbel V, Koffi AA 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.

Nuzhdin SV, Wayne ML, Harmon KL and McIntyre LM. 2004. Common pattern of evolution of gene expression level and protein sequence in *Drosophila*. *Molecular Biology and Evolution* **21**: 1308 – 1317.

Oaks SC, Mitchell VS, Pearson GW and Carpenter CC. 1991. *Malaria: Obstacles and Opportunities. A Report of the Committee for the Study on Malaria Prevention and Control: Status Review and Alternative Strategies.* Division of International Health, National Academy Press, Washington DC.

Okoye PN, Brooke BD, Koekemoer LL, Hunt RH and Coetzee M. 2008. Inheritance of pyrethroid resistance in the major southern African malaria vector *Anopheles funestus*. *Annals of Tropical Medicine and Parasitology* **102**: 275 – 281.

Okoye PN, Brooke BD, Hunt RH and Coetzee M. 2007. Relative developmental and reproductive fitness associated with pyrethroid resistance in the major southern African malaria vector *Anopheles funestus. Bulletin of Entomological Research* **97**: 599 – 605.

Paterniani E. 1969. Selection for Reproductive Isolation between Two Populations of Maize, *Zea mays* L. *Evolution* **23**: 534 – 547.

Paterson HE. 1978. More Evidence against speciation by reinforecement. *South African Journal of Science* **74**: 369 – 371.

Paterson HE. 1980. A Comment on "Mate Recognition Systems". *Evolution* **34**: 330 – 331.

Paterson HE. 1985. *The Recognition Concept of Species: Species and Speciation*. Ed: Vrba ES. Transvaal Museum Monograph 4, Pretoria.

Paterson HE. 1988. On Defining Species in Terms of Sterility: Problems and Alternatives. *Pacific Science* **42**: 65 – 71.

Pfaffl MW. 2001. A new mathematical model for relative quantification in real-time RT–PCR. *Nucleic Acids Research* **29**: e45.

Pfaffl MW, Tichopad A, Prgomet C and Neuvians TP. 2004. Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper--Excel-based tool using pair-wise correlations. *Biotechnology Letters*. **26**: 509 – 515.

Paskewitz SM, Wesson DM and Collins FH. 1993. The internal transcribed spacers of ribosomal DNA in five members of the *Anopheles gambiae* species complex. *Insect*

Molecular Biology **2**: 247 – 257.

Ranson H, Jensen B, Vulule JM, Wang X, Hemingway J and Collins FH. 2000. Identification of a Point Mutation in the Voltage-Gated Sodium Channel Gene of Kenyan Anopheles gambiae Associated with Resistance to DDT and Pyrethroids. *Insect Molecular Biology* **9**: 491 – 497.

Ranson H, Nikou D, Hutchinson M, Wang X, Roth C, Hemingway J and Collins FH. 2002. Molecular Analysis of Multiple Cytochrome P450 genes from the Malaria Vector, *Anopheles gambiae*. *Insect Molecular Biology* **11**: 409 – 418.

Ranson H, Paton MG, Jensen B, McCarroll L, Vaughan A, Hogan JR, Hemingway J and Collins FH. 2004. Genetic Mapping of Genes Conferring Permethrin Resistance in the Malaria Vector, *Anopheles gambiae. Insect Molecular Biology* **13**: 379 – 386.

Ritchie ME, Silver J, Oshlack A, Holmes M, Diyagama D, Holloway A and Smyth GK. 2007. A comparison of background correction methods for two-colour microarrays. *Bioinformatics* **23**: 2700 – 2707.

Rudenko N, Golovchenko M, Edwards MJ and Grubhoffer L. 2005. Differential expression of *Ixodes ricinus* tick genes induced by blood feeding or *Borrelia burgdorferi* infection. *Journal of Medical Entomology* **42**: 36 – 41.

Sanders HR, Evans AM, Ross LR and Gill SS. 2003. Blood meal induces global changes in midgut gene expression in the disease vector, *Aedes aegypti*. *Insect Biochemistry and Molecular Biology* **33**: 1105 – 1122.

Sakai RK, Baker RH, Raana K and Hassan M. 1979. Crossing Over in the Long Arms of the X and Y Chromosomes in *Anopheles culicifacies*. *Chromosoma* **74**: 209 – 218.

Sambrook J, Fritsch AF and Maniatis T. 1989. *Molecular Cloning. A Laboratory Manual*. Second Edition. Ed: Nolan C. Cold Spring Harbor Laboratory Press, United States of America.

Sanders HR, Evans AM, Ross LR and Gill SS. 2003. Blood Meal Induces Global Changes in Midgut Gene Expression in the Disease Vector, *Aedes aegypti*. *Insect Biochemistry and Molecular Biology* **33**: 1105 – 1122.

Sartor MA, Zorn AM, Schwanekamp JA, Halbleib D, Katyala S, Howell ML, Dean GE, Medvedovic M and Tomlinson CR. 2005. A new method to remove hybridisation bias for interspecies comparison of global gene expression profiles uncovers an association between mRNA sequence divergence and differential gene expression in *Xenopus*. *Nucleic Acids Research* **43**: 185 – 200.

Schena M. 2002. *DNA Microarrays. A Practical Approach*. Oxford University Press, New York.

Schena M, Shalon D, Davis RW and Brown PO. 1995. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* **270**: 467 – 470.

Schulze A and Downward J. 2001. Navigating gene expression using microarrays – a technology review. *Nature Cell Biology* **3**: 190 – 195.

Scott JA, Brogdon WG, Collins FH. 1993. Identification of single species of the *Anopheles gambiae* complex by the polymerase chain reaction. *American Journal of Tropical Medicine and Hygiene* **49**: 520 – 529.

Scott JG. 1999. Review: Cytochromes P450 and Insecticide Resistance. *Insect Biochemistry and Molecular Biology* **29**: 757 – 777.

Serazin AC, Dana AN, Hillenmeyer ME, Lobo NF, Coulibaly MB, Willard MB, Harker BW, Sharakov IV, Collins FH, Ribiero JM and Besansky NJ. 2009. Comparative Analysis of the Global Transcriptome of *Anopheles funestus* from Mali, West Africa. *PLos ON*E **4:** e7976.

Sharakhov IV, Sharakhov MV, Mbogo CM, Koekemoer LL and Yan G. 2001. Linear and Spatial Organization of Polytene Chromosomes of the African Malaria Mosquito *Anopheles funestus*. *Genetics* **159**: 211 – 218.

Sharp BL, Kleinschmidt I, Streat E, Maharaj R, Barnes KI, Durrheim DN, Ridl FC, Morris N, Seocharan I, Kunene S, La Grange JJP, Mthembu JD, Maartens F, Martin CL and Barreto A. 2007. Seven Years of Regional Malaria Control Collaboration – Mozambique, South Africa and Swaziland. *American Journal of Tropical Medicine and Hygiene* **76**: 42 – 47.

Sharp BL and le Sueur D. 1996. Malaria in South Africa – the Past, the Present and Selected Implications for the Future. *South African Medical Journal* **86**: 83 – 89.

Sharpe RG, Harbach RE and Butlin RK. 2000. Molecular variation and phylogeny of the members of the Minimus group of *Anopheles* subgenus *Cellia* (Diptera: Culicidae). *Systematic Entomology* **25**: 263 – 272.

Singh OP, Chandra D, Nanda N, Raghavendra K, Sunil S, Sharma SK, Dua VK and Subbarao SK. 2004. Differentiation of members of the *Anopheles fluviatilis* species complex by an allele-specific polymerase chain reaction based on 28S ribosomal DNA sequences. *American Journal of Tropical Medicine and Hygiene* **70**: 27 – 32.

Smyth GK. 2004. Linear models and empirical Bayes methods for assessing differential expression in microarray experiments. *Statistical Applications in Genetics and Molecular Biology* **3**: 1.

Smyth GK, Michaud J and Scott H. 2005. The use of within-array replicate spots for assessing differential expression in microarray experiments. *Bioinformatics* **21**: 2067 – 2075.

Smyth GK and Speed TP. 2003. Normalization of cDNA microarray data. *Methods* **31**: 265 – 273.

Somboon P, Thongwat D, Choochote W, Walton C and Takagi M. 2005. Crossing experiments of *Anopheles minimus* species A and putative species E. *Journal of the American Mosquito Control Association* **21**: 5 – 9.

Spiers AA, Mzilahowa T, Atkinson D and McCall PJ. 2002. The malaria vectors of the Lower Shire valley, Malawi. *Malawi Medical Journal* **14**: 4 – 7.

Spillings BL, Brooke BD, Koekemoer LL, Chiphwanya J, Coetzee M and Hunt RH. 2009. A New Species Concealed by *Anopheles funestus* Giles, a Major Malaria Vector in Africa. *American Journal of Tropical Medicine and Hygiene* **81**: 510 – 515.

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.

Taylor S, Wakem N, Dijkman G, Alsarraj M and Nguyen M. 2010. A practical approach to RT-qPCR – Publishing data that conform to the MIQE guidelines. *Methods* **50**: S1 - S5.

Thongwat D, Morgan K, O'Loughlin SM, Walton C, Choochote W and Somboon P. 2008. Crossing experiments supporting the specific status of *Anopheles maculatus* chromosomal form K. *Journal of the American Mosquito Control Association* **24**: 194 – 202.

Untergasser A, Nivjeen H, Rao X, Bisseling T, Geurts R and Leunissen J. 2007. Primer3Plus, an enhanced web interface to Primer3. *Nucleic Acids Research* **35**: W71 - W74.

Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A and Speleman F. 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biology* **3**: research0034.1 - 0034.11.

Vermeulen J, De Preter K, Lefever S, Nuytens J, De Vloed F, Derveaux S, Hellemans J, Speleman F and Vandesompele J. 2011. Measurable Impact of RNA Quality on Gene Expression Results from Quantitative PCR. *Nucleic Acids Research* **39**: e63.

Vontas JG, Small GJ and Hemingway J. 2001. Glutathione S-tranferases as antioxidant defense agents confer pyrethroid resistance in *Niliparvata lugens*. *Biochemical Journal* **357**: 65 – 72.

Vontas J, Blass C**,** Koutsos AC, David J-P, Kafatos FC, Louis C, Hemingway J, Christophides GK and Ranson H**.** 2005. Gene Expression in Insecticide Resistant and Susceptible *Anopheles gambiae* Strains Constitutively or After Insecticide Exposure. *Insect Molecular Biology* **14**: 509 – 521.

Weill M, Fort P, Berthomieu A, Dubois MP, Pasteur N and Raymond M. 2002. Novel Acetylcholinesterase Gene in Mosquitoes Codes for the Insecticide Target and is Non-homologous to the *ace* gene in *Drosphila*. *Proceedings of the Royal Society of London Series B-Biological Sciences* **269**: 2007 – 2016.

Wilkes TJ, Matola YG and Charlwood JD. 1996. *Anopheles rivulorum*, a vector of human malaria in Africa. *Medical and Veterinary Entomology* **10***:* 108 – 110.

Wilkins EE, Howell PI and Benedict MQ. 2007. X and Y Chromosome Inheritance and Mixtures of rDNA Intergenic Spacer Regions in *Anopheles gambiae*. *Insect Molecular Biology* **16**: 735 – 741.

Williamson MS, Martinez-Torres D, Hick CA and Devonshire AL. 1996. Identification of Mutations in the Housefly para-type Sodium Channel Gene Associated with Knockdown Resistance (*kdr*) to Pyrethroid Insecticides. *Molecular and General Genetics* **252**: 51 – 60.

Winnebeck EC, Millar CD and Warman Gr. 2010. Why does insect RNA look degraded? *Journal of Insect Science* **10**: 159.

Wirtz RA, Zavala F, Charoenvit Y, Cambell GH, Burkot TR, Schneider I, Esser KM, Beaudoin RL and Andre GR. 1987. Comparative testing of *Plasmodium falciparum* sporozoite monoclonal antibodies for ELISA development. *Bulletin of the World Health Organization* **65***:* 39 – 45.

Wondji CS, Morgan J, Coetzee M, Hunt RH, Steen K, Black WC, 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**: 34.

Wondji CS, Irving H, Morgan J, Lobo NF, Collins FH, Hunt RH, Coetzee M, Hemingway J and Ranson H. 2009. Two Duplicated P450 Genes are Associated with Pyrethroid Resistance in *Anopheles funestus*, a Major Malaria Vector. *Genome Research* **19**: 452 – 459.

World Health Organization, 1998. Test procedures for insecticide resistance monitoring in malaria vectors, bio-efficacy and persistence of insecticides on treated surfaces. Document *WHO/CDS/CPC/MAL/98.12*. Geneva, Switzerland.

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

Xu Q, Wang H, Zhang L and Liu N. 2006. *Kdr* Allelic Variation in Pyrethroid Resistant Mosquitoes, *Culex quinquefasciatus* (S). *Biochemical and Biophysical Research Communications* **345**: 774 – 780.

Zlotkin E. 1999. The insect voltage-gated sodium channel as a target of insecticides. *Annual Review of Entomology* **44**: 429 – 455.

Zuker, M. 2003. Mfold web server for nucleic acid folding and hybridisation prediction. *Nucleic Acids Research* **31:** 3406 – 3416.

WEBSITES All websites were last accessed 16 January 2012

http://blast.ncbi.nlm.nih.gov

www.premierbiosoft.com

www.bioconductor.org

www.fidssa.co.za/B_FIDSSA2011Presen.asp

www.nimblegen.com/products/expression/eukarya/index.html

www.affymetrix.com

WHO/CDS/MAL/98.12: Discriminating Concentrations of Insecticides for Adult Mosquitoes. World Health Organization, Geneva, Switzerland, 1998 available at: http://www.who.int/whopes/resistance/en/

http://www.who.int/malaria/world_malaria_report_2010

www.malariaeliminationgroup.org

www.oxforddictionaries.com

http://en.wikipedia.or/wiki/File:Speciation_modes.svg