Genetic and ecological processes of speciation in the *Anopheles coluzzii* and *Anopheles gambiae* sibling species

Nancy Naanogot Dawam

Thesis submitted for the Degree of Doctor of Philosophy

June 2019

Keele University

Abstract

Malaria is a leading cause of death, killing thousands every year. Sub-Saharan Africa has the highest mortality rate, with pregnant women and children under the age of five most affected. The culprits responsible for transmission of the disease are mosquitoes of the genus *Anopheles.* Preventative measures using vector control are mainly insecticide based, however, resistance threatens the efficacy of these methods. The spreading of transgenes is considered an alternative vector control method but the success of this depends on knowledge of the *Anopheles* population and cryptic taxa in the wild. The genetic and ecological differences that exist between the populations and how reproductive isolation occurs between the vectors is of great importance in prospective vector control programs which rely on the release of transgenic mosquitoes.

Focusing on *Anopheles coluzzii* and *Anopheles gambiae*, the recently diverged members of the *Anopheles gambiae* complex, morphological and molecular techniques were used to study genetic and ecological differences between the sibling species. Egg morphological differences facilitated by ecological divergences between the two sibling species were studied. Results showed a difference in egg shape and size between the two species and populations within species.

The genetic studies focused on the identification of assortative mating genes. In order to identify candidate mate choice genes, expression levels in 27 putative genes located on the X-island of speciation were investigated. A majority of the genes were over-expressed in virgin males in the samples. An attempt to silence the two top candidate putative assortative mating genes through RNAI using injection as the method of dsRNA delivery led to a surprising outcome, as the mechanical impact of the injections appeared to disrupt the assortative mating pattern.

Table of content

List of Figures

List of Tables

Abbreviations

ACT	Artemisinin-based combination therapy
AgamOBPI	*Anopheles gambiae* odorant protein
AgORs	*Anopheles gambiae* odorant receptors
cDNA	Complementary DNA
CT	Cycle threshold
DDT	Dichlorodiphenyltrichloroethane
DEET	N,N-dielthyl-meta-toluamide
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
dsRNA	Double stranded RNA
IRS	Indoor residual spray
kdr	Knock down resistance
LLINs	Long lasting insecticide-treated nets
mRNA	Messenger RNA
OBP	Odorant binding protein
ONNV	O'nyong-nyoung virus
Orco	Odorant receptor co-receptor
ORN	Olfactory receptor neuron
Ors	Olfactory receptors
PC	Principal component
PCA	Principal component analysis
PCR	Polymerase chain reaction
qRT-PCR	Real-Time quantitative reverse transcription PCR
rDNA	Ribosomal deoxyribonucleic acid
RNA	Ribonucleic acid
RNAi	RNA interference
RT-PCR	Reverse transcription polymerase chain reaction
SINEs	Short interspersed nuclear elements
siRNA	Short interfering RNAs
SIT	Sterile insect technique
SNP	Single nucleotide polymorphism
SSM	Sexual sporogenic, mosquito-stage
VIMT	Vaccine to interrupt malaria transmission
Taq	*Thermus aquaticus*
WHO	World Health Organization

Acknowledgements

In God I live, and move and have my being. To Him be the glory.

I have looked forward with excitement to writing this section of my thesis from my first year at Keele believing it would be the easiest and most exciting section of my thesis, however, it turned out not the case as it involved some reflection.

I wish to first and foremost acknowledge the efforts of my supervisor Professor Frederic Tripet who despite his busy schedule has painstakingly guided me throughout this research. I am grateful to have worked under your supervision. Besides my supervisor, I would like to thank my academic advisor Dr Julien Pelletier for the support I enjoyed from him throughout my studies. Your suggestions, discussions and advice were of immense help to me.

I gratefully acknowledge TET-FUND Nigeria for the funding received towards my PhD training and Plateau State University Bokkos for nominating me for the scholarship.

I would like to thank my fellow lab mates in the Tripet group, Akbar Ganatra, Mauro Pazmino, and Aisha Carmicheal whom I worked with closely during their MSc research. Their contributions towards the optimisation of some techniques used in this research were valuable. Other members of the Tripet lab whose support I enjoyed during my studies include Esther Ekechukwu, Nahla Alhafiz, Holly Mathew, Emmanuel Ottih, Nwamaka Akpodiete, Florian Noulan, Ahbi, Hahni, Mara, Andrea, and Niang.

My appreciation also goes to my friends and colleagues at the school of life sciences Nana Efua Andoh, Adeyemi Akinyemi, Nawroz Kareem, Saja, Hamza, Ibrahim, Linda, Jaksha Chandrathas, Atieme, Mukdad, Lekia Kumbe, Sundus, Tayo, Emma Whittle, Jonathan Swift, Lynne, and James Smallborne for the support I enjoyed from them. Thanks to Ann Underhill, Ron and the other technicians for been very responsive. My

thanks also go to all staff and students of School of Life Sciences, Huxley building Keele University.

Thank you to my friends outside of Keele, Uju, Fidelia, Soonjae, Emmanuel Ofomana, Mulikat, Lilian, Paul Okeny and Mrs Monica Yearsley (My Landlady). Your friendship is greatly appreciated.

My appreciation goes to Prof. Gomwalk, Prof. Sheni, Prof. Matur, Associate Prof. Nvau, Prof. Ekechukwu and Dr. Nangbes for their encouragement through my academic career.

My sincere appreciation goes to Uncle Mathew Mangdong, Mrs Bibi Cassidy, Mrs. Afodia Philip, Prof Theresa Odumu, Roseline Fashe, Mr. Kyauta Damulak, Gapsuk Dashin, Keziah Ayuba, Kim Luka Dung, Mansha Dakup, Tessy Ojeka, Eke Ukwedeh, Mefoni Olofu and Steve Ali for being responsive whenever I needed someone to talk to.

Special thanks to my family, Naanep whom I have always passed some of my responsibilities to knowing he makes me a priority, Naandye Dabuagat who is always there for me, Mr and Mrs Shetak, Mercy Mangdong, Mr and Mrs Dayong, Uncle Kanshal, Emmanuel, Blessing and Philip. Thank you for your love and support.

Lastly, I am indebted to my Father and Mother, Patrick Daduan Dawam and Felicia Gwamni Datong Dawam for their encouragement, support, guidance and above all their unconditional love. Thank you for having faith and believing in me.

Chapter 1 General Introduction

1.1 Introduction

Malaria is an infectious disease caused by a microscopic parasitic protozoan of the genus *Plasmodium* and transmitted by mosquitoes. It is considered by many to be the most important disease of humans globally and affects mostly poor countries (Marquardt *et al*., 2000). Malaria is now known as a tropical disease though until the mid 19th century, 90% of the world's population was at risk because the distribution spanned to the artic circle in the northern hemisphere (Carter and Mendis, 2002) and there were outbreaks in both temperate and Mediterranean regions.

There are currently five species of human malaria parasites, namely; *Plasmodium falciparum, Plasmodium vivax, Plasmodium malariae* and *Plasmodium ovale* and the recently discovered *Plasmodium knowlesi* that infects mainly macaque monkeys in South Asia but can also infect humans (Service, 2012).

Malaria perhaps existed before man and it is assumed to have adapted to the infection of humans in Africa (Knell, 1991). However, *P. falciparum*, the most deadly human malaria parasite (Coulson *et al*., 2004; Pasternak and Dzikowski, 2009) is believed to have adapted to man more recently (Knell, 1991). As individuals travelled from one place to another during the Neolithic era, malaria was introduced into Europe, the Middle East and Asia. Malaria was then possibly introduced to South and Central America from Asia during the pre-Columbian times in the first millennium AD. It is assumed to have reached the Caribbean from the mainland or through slave trade (Knell, 1991).

In ancient times, the Europeans categorized *P. vivax* and *P. malariae* as 'benign tertian and quartan' fevers in reference to their 48 hours and 72 hours periodicity. 'Benign' because *P. vivax* infections were considered mild in their manifestations unlike the continuous sub tertian malignant fever caused by *P. falciparum*, which had severe symptoms. These fevers have

been made reference to in many Northern Europe writings and around the shores of the Mediterranean Sea from about the 5th century B.C. onwards reviewed in (Carter *et al*., 2002). Malaria hindered economic activities and hampered efforts in agriculture and engineering as places with great potentials for industrial activities and also large plantations were barely inhabited because they were considered unsafe (Knell, 1991).

In 1897, Ross discovered the role of mosquitoes in the transmission of malaria. That was when malaria epidemiology became a science. His discovery led to the possibility of controlling and preventing malaria and he went ahead to make plans to eradicate the disease (Knell, 1991).

The malaria menace and burden continues today (Figure 1.1.1), approximately 216 million cases of malaria were recorded globally in 2016, with 90% of cases occurring in Sub-Saharan Africa (World Health Organization, 2017). South-East Asia accounts for 7% and the Eastern Mediterranean 2% of cases (World Health Organization, 2016). Globally, 445,000 deaths were recorded in 2016 with African countries accounting for 91% (World Health Organization, 2017). About 99% of estimated malaria cases in 2016 were caused by *Plasmodium falciparum* malaria (Figure 1.1.1) (World Health Organization, 2016).

Plasmodium falciparum is most prevalent in Africa where the direct and indirect burden of malaria is caused by the pathogen. The pathogen is also found in most parts of South-East Asia. It has been difficult to provide a single figure for mortality or clinical events resulting from *P. falciparum* infection (Snow, 2014). This is because the epidemiology of malaria has made disease surveillance and method of estimating the burden of the disease difficult (Hay *et al*., 2010a) The global map for malaria endemicity should be updated frequently as this would give an idea of the population at risk and also give a direction to tackle the disease. Accurate evaluations of the burden of malaria are important for intervention strategies (Hay *et al*., 2004).

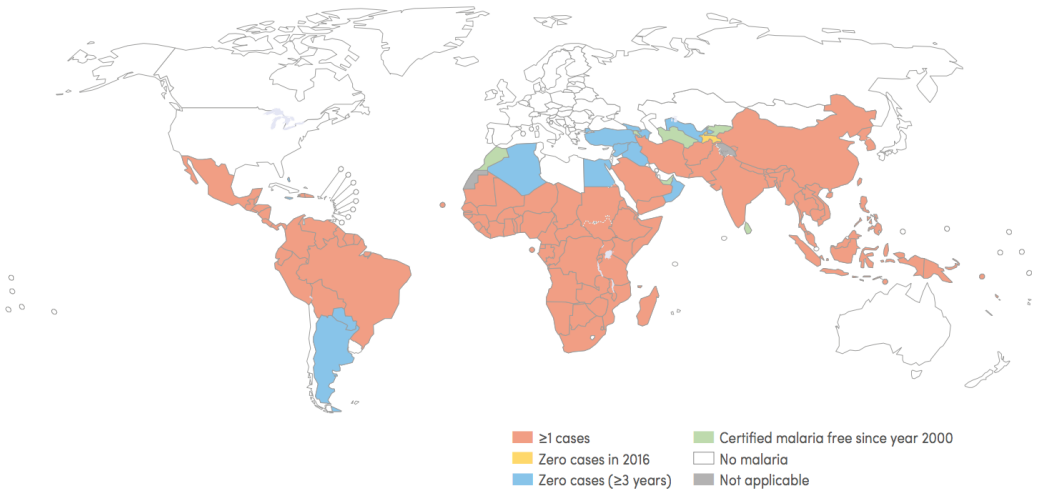

Figure 1.1.1: Countries and territories with indigenous cases of malaria in 2000 and their status by 2016. **Source:** (World Health Organization, 2017)**.**

1.2 The Anopheline Vector

The family Culicidae, commonly known as mosquitoes has 43 genera and more than 3,530 species of mosquitoes (Service, 2012). These are divided into 3 subfamilies namely: Toxorhynchitinae, Anophelinae (Anophelines) and Culicinae (Culicines). Members of the genus *Anopheles* are the only vectors of the *Plasmodium* parasites that cause human malaria (Service, 2012). There are 476 known species of *Anopheles* out of which about 70 can transmit human malaria (Service, 2012). In addition to transmission of malaria, *Anopheles* species are occasional vectors of filariasis (*Wuchereria bancrofti*, *Brugia malayi* and *Brugia timori*) and some arboviruses such as O'Nyong Nyoung virus (ONNV) (Service, 2012).

Mosquitoes of the genus *Anopheles* can be found in both temperate and tropical regions where breeding sites are available. These mosquitoes are quite small in size measuring about 8mm in length. They are not loud while flying and also do not bite so painfully compared to other blood sucking insects therefore they can easily be ignored (Knell, 1991). The adult *Anopheles* has dark dotted wings, its head is usually down, body at an angle

and hind legs raised when either resting or biting (Knell, 1991). *Anopheles* is mostly active at night. Feeding and oviposition takes place in the early hours of the morning, in the evening or late at night. Species that bite humans are called 'anthropophagic' while those species that have preference for animals are called 'zoophagic'. Those that find their way into houses to feed are called 'endophagic' and those that feed outside are called exophagic (Service, 2012). Mosquitos seek shelter during blood digestion and development of eggs, therefore resting before or after blood feeding takes place either inside or outside houses. Those who rest inside houses are called endophilic while those that rest outside are said to be exophilic (Service, 2012). Most Anopheline mosquitoes are not entirely endophagic, exophagic, endophilic or exophilic; they exhibit a combination of the aforementioned extreme cases (Service, 2012).

1.2.1 Life cycle and Ecology of Anopheles

The adult *Anopheles* mosquito's first activity is mating. The female copulates just once in its lifetime (Knell, 1991), though (Tripet *et al*., 2003) observed evidence of multiple inseminations in some field-collected female *An. gambiae* in Mali. During mating, sperm are collected and stored for subsequent use and the last thing the male does after mating is to infuse a sealing substance that prevents the passage of sperm if any other mating does occur (Knell, 1991). A mated and bloodfed *Anopheles* lays about 50 to 200 small brown eggs (Service, 2012) deposited individually on water bodies. The eggs can be characterized by the presence or absence of floats (Hinton, 1968), cannot survive desiccation (Service, 2012) and therefore must remain in water to survive. The eggs hatch within 2 to 6 days into larvae. In the tropics the eggs can hatch within a short period of 2 to 3 days while in cold regions it can take as long as 2 to 3 weeks. The length of time is dependent on temperature (Service, 2012).

Just like all mosquitoes, *Anopheles* has four larval stages called instars. The larvae are filter feeders and are usually found on the surface of the water bodies (Figure. 1.2.1) because

they lack a siphon and must breathe air through their posterior spiracles. When disturbed, they swim to the bottom of the water body but re-emerge almost immediately. They feed on bacteria, protozoa and other microorganisms (Service, 2012). *Anopheles* larvae can be found in different habitats, ranging from freshwater, saltwater marshes, mangrove swamps, rice fields to grassy ditches, wells, edges of streams and rivers as well as ponds and borrow pits. It is also possible to find larvae in water-storage pots and some species can be found in water-filled tree-holes. In Central America, South America and the West Indies, a few *Anopheles* breed in water that collects in the leaf axils of epiphytic plants (Service, 2012). Adaptations vary in *Anopheles* as some species inhabit aquatic vegetation and other waters without vegetation. Some species prefer exposed sunlit waters whereas others more shaded larval habitats. Generally, Anophelines are found in clean unpolluted waters, and are usually absent from habitats containing rotting plants or faeces (Service, 2012). The larval period lasts about 7 days in the tropics while it takes about 2 to 4 weeks in cooler climates (Service, 2012). The larva moults three times and in the forth moult the larvae becomes a locomotive pupae which does not feed (Figure 1.2.1) (Knell, 1991).

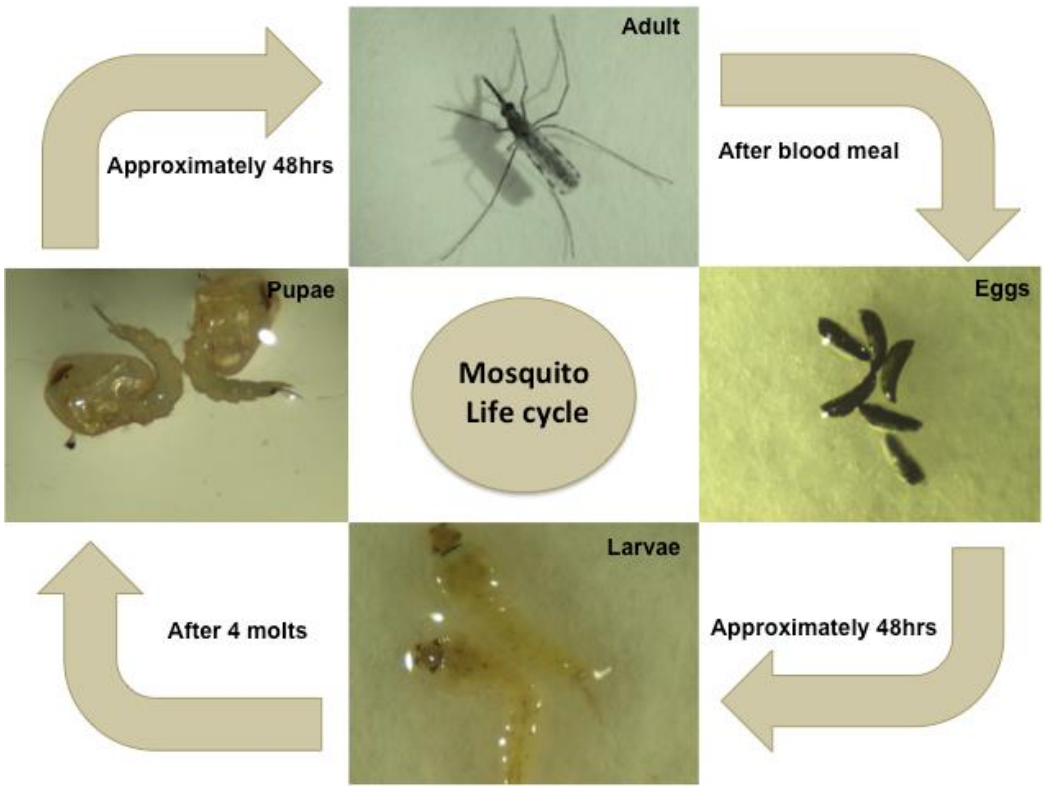

Figure 1.2.1: The mosquito life cycle.

The pupae are comma shaped and are also found on water surfaces, however, they swim to the bottom with a jerking movement when disturbed (Service, 2012). The pupae use two air trumpets for breathing. (Knell, 1991) and pupal period last about 2 to 3 days (Figure 1.2.1) in the tropics and can last about 2 to 3 weeks in cold regions (Service, 2012). When the a pupae splits open, the adult mosquito emerges having tender limbs and stays on the water surface to harden and dry before its first flight. Development of mosquitoes from egg to adult takes about three weeks to one month depending on temperature (Knell, 1991).

1.2.3 Geographical distribution of the Malaria vectors

Anopheles are usually present in most countries be it temperate or tropical provided there are breeding sites to strive (Knell, 1991). Anopheles species are considered important vectors in a region when they are able to spread the human malaria parasite due to their population and

ability to blood feed on humans. The vectors should also have a relatively average lifespan to enable incubation of the malaria parasite (Hay *et al*., 2010b) In order to understand the most suitable control measures to be applied in a given area, it is important to know the types of *Anopheles* vectors that exist in the area, their abundance and also their behaviour (Sinka, 2013).

America has 9 dominant vector species which are: *Anopheles albimanus* Wiedemann, *An. albitarsis, An. aquasalis* Curry, *An. darlingi* Root, *An. freeborni* Aitken, *An. marajoara* Galvão & Damasceno, *An. nuneztovari*, *An. pseudopunctipennis* and *An. quadrimaculatus* (Sinka *et al*., 2010). Only *An. freeborni* is present in northwest America and *An. quadrimaculatus s.l* found in the southeastern part with slight interjection of *An. pseudopunctipennis* in the extreme South of the continent. *Anopheles darlingi* is dominant in South America but *An. Albimanus* and *An. pseudopunctipennis* dominate in Central America even where *An. albimanus* is present. *An. aquasalis* though not an important vector remains dominate because it can strive in some habitats especially in the coastal areas of Central and South America. *An. marajoara* and *An. albitarsis* are emerging vectors in this region (Sinka *et al*., 2012). *An. messeae* occurs in Europe and the Middle East. It occurs in the United Kingdom in the west, spreads to the Eastern part of Europe and then into Asia (Sinka, 2013).

The vector situation in the Asian-Pacific region is complex because the region is characterized by high vector diversity (Figure 1.3.1). Many are sympatric and able to modify their behaviour according to varying environmental conditions resulting in the high prevalence rate of malaria (Sinka *et al*., 2012). Understanding vector transmission in the environment is a colossal task.

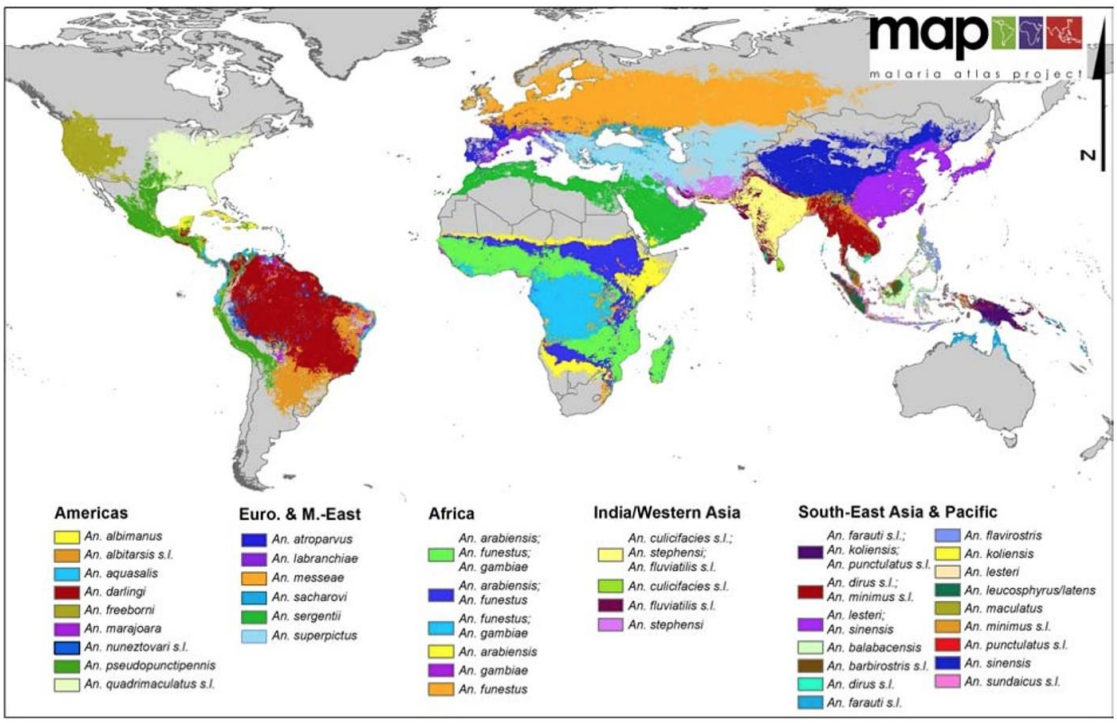

Figure 1.3.1: A global world map of dominant malaria vector species. **Source:** (Sinka *et al.*, 2012).

The dominant species in Northern areas of Asia are *An. sinensis* and *An. lesteri*, which occur in sympatry in China and Korea. Sympatry occurs for *An. stephensi* and *An. fluviatilis* in India (Sinka *et al*., 2012). *An. dirus A* is found to play an important role as a dominant vector in the transmission of malaria in Southeast Asia. This was reported in a study carried out in Vietnam, Laos and Cambodia (Trung *et al*., 2004). *An. minimus* complex also dominates the Southeast, however in Thai/Malaysian peninsula where both species both *An. dirus A* and *An. minimus* occur together they tend to diverge and occupy different niches. On the Pacific islands, vectors include *An. farauti, An. koliensis, An. punctulatus* and *An. farauti*. Great vector diversity is observed in Indonesia and sympatry occurs between *An. sinensis*, *An. barbirostris*, *An. leucosphyrus, An. latens* and the *An. minimus* complex. *An. maculatus* and *An. flavirostris* exist but are not considered as dominant (Sinka *et al*., 2012).

In Africa there is a correlation between malaria transmission and the abundant presence of the most effective malaria vector, *An. gambiae*. Its ability to associate closely

with human habitat, preference for human blood, ability to get into houses at night to feed and its abundance in most parts of the continent enhances its vectoral capability (Hodges *et al*., 2013). A high percentage of the deaths and illnesses resulting from malaria are recorded from Sub-Saharan Africa as a consequence of the widespread of the *An. gambiae* in this region. The vector is a member of the *An. gambiae* complex. Other members of the complex include *An. arabiensis, An. merus,* and *An. melas,* which are also considered as dominant species in the continent. *An. funestus* is also considered an important vector. *An. moucheti* is not widespread but extremely anthropophilic, *Anopheles nili* complex is also anthropophilic with members biting both indoors and outdoors. These are all considered effective vectors contributing to the spread of malaria in Africa (Sinka, 2013).

Malaria vectors have been classified as primary or secondary depending on their importance though this could be misleading as a primary vector in one part could be considered as secondary in another region (Obsomer *et al*., 2013; Service, 2012). Effective vector control therefore varies depending on the vector's epidemiology and distribution (Hay *et al*., 2010a).

1.3 Malaria control and interventions

The first treatment for malaria was recorded in 1600, when Peruvian Indians used the bitter bark of the Cinchona tree. It became available to the English in 1649 as Jesuits powder so that those suffering from Agues (the English name for malaria until the 19th century) would benefit from it. The active ingredient of the powder was its chemical component quinine, used to combat malaria and is still in use today (Reiter, 2000). In the latter part of the 19th century, in 1880 precisely, the protozoan was discovered and by 1897, the *Anopheles* mosquito was associated with the disease (Cox, 2010; Lambert, 2003). This created the opportunity to target the disease via mosquito control. The Romans carried out the first vector control against

malaria. This was done through drainage programs because the disease was associated with stagnant water (Reiter, 2000). Elevation of houses and draining of swamplands was carried out to prevent fever and was practiced among the Greeks and Romans before malaria was discovered (Knell, 1991). Angelo Celli carried out the first trial on vector control against malaria in 1899. In his survey, some families were protected with screened windows and doors to prevent mosquitoes getting indoors while other families were not protected. The result showed that nearly all families that were unprotected contracted malaria. Only 4 out of 24 people who were protected had malaria, coincidentally they all worked at night and did not cover themselves when outdoors (Lindsay *et al*., 2002).

By 1910 the screening of houses against mosquitoes became a practice all over the world. The method was used to protect Europeans living in the Tropics. The British army mosquito-proofed their barracks in Lahore in 1925 and this singular act led to a 68% reduction in malaria cases. This technique also led to the reduction malaria in the USA. Watson in 1949 reported that Screening of houses is a very effective method for malaria control. He advised on the development of low-cost housing designed with the prevention of the disease as priority (as cited in (Lindsay *et al*., 2002). The screening of houses was quite effective against malaria but man wanted a faster way and the use of Dichlorodiphenyltrichloroethane (DDT) was introduced (Lindsay *et al*., 2002).

DDT is a synthetic organic compound used as an insecticide. It was considered cheap and highly effective. Therefore, the WHO malaria eradication program between 1955-1979 was centered on the use of DDT and chloroquine (a synthetic antimalarial based on the structural elements of quinine). Unfortunately this led to the neglect of other controls measures (Lindsay *et al*., 2002). The program was effective as the burden of malaria was further reduced and even eradicated in some parts of the world (Carter *et al*., 2002) Unfortunately, the program was not sustained resulting in an upsurge of malaria cases. The

parasites became resistant to chloroquine and the vectors to DDT (Carter *et al*., 2002).

Over the years funds have been made available for malaria control programs and research. The United States government being the highest contributor to malaria control activities. Countries like Britain, Japan, Germany, France and so many others have invested in the fight against malaria in varying capacity. Individual organizations like the Bill and Melinda Gates foundation are also major contributors to the fight against malaria (World Health Organization, 2016). These funds are channeled towards vector control approaches, prophylaxis, chemotherapy and possible production of malaria vaccines.

1.3.1 Vaccine development

Vaccines have been used as a public health tool in combating diseases (Alonso *et al*., 2011). Malaria vaccine development can be classified into two, the pre-erythrocytic vaccine and the transmission blocking vaccine. The transmission blocking vaccine targets the stage where the parasite is in the mosquito. One of such vaccines is the sexual sporogenic, mosquito-stage vaccine to interrupt malaria transmission (SSM-VIMT). This vaccine is to stop the parasites from developing inside the mosquito (Sauerwein and Bousema, 2015). These vaccines are to induce long-lived anti-bodies aimed at blocking transmission from humans to mosquitoes leading to the interruption of the transmission cycle. Though the possibility of using the transmission blocking technique for the control of malaria seems to have increased over the years, the likelihood of it being accepted anytime soon is slim (Nunes *et al*., 2014). The most promising vaccine is the RTS, S/AS (AS01 or AS02), a pre-erythrocytic malaria vaccine targeting the circumsporozoite surface protein of *Plasmodium falciparum* thereby reducing liver stage infection. In the acronym RTS,S, R stands for repeat region of *P. falciparum* circumsporozoite protein (CSP), T stands for the T-cell epitopes of the CSP while the S stands for hepatitis B surface antigen (HBsAg). This has been found to provide protection

against *P. falciparum* in children and infants in phase 2 trials (Karunamoorthi, 2014). The vaccine was found to be effective in a study carried out in the Gambia where 71% protection was achieved in men within the first 9 weeks but declined to 0% afterward (Bojang *et al*., 2001). The protection acquired through the RTS S/AS02 vaccine has been reported to be short lived (Gosling and Seidlein, 2016), RTS S/AS01 on the other hand was shown to have reduced the clinical incidence of malaria by 39% and severe malaria by 31.5% among children from 5-17 months who completed the four doses treatment (World Health Organization, 2016). This vaccine has completed the phase 3 testing (World Health Organization, 2016). Vaccines should be cheap and affordable and a large number of people should be able to have access to it. Most importantly, it should be a preventive measure against the target disease (Karunamoorthi, 2014). Currently, vaccines are not considered contributors to malaria control.

1.3.2 Antimalaria Drugs

The main malaria drugs are Halofantrine, Artemisinins and sulphadoxine-pyrimethamine. Others are Chloroquine, Amodiaquine, Mefloquine, and quinine. Quinoline has its active ingredient from Cinchona, which originated from Peru and was used as an antimalarial since the 17the century (Ridley, 2002). For decades chloroquine has been the main chemotherapeutic treatment for malaria (Bir *et al*., 2002). It is an amazing antimalarial compound active against the four species of human malaria parasite (White, 2008). However, its immense use led to resistance in *P. falciparum* and *P. vivax* in Colombia in the 1950s (Wellems and Plowe, 2001). Chloroquine resistance started in 1959 from Colombia to eastern Thailand and then to the tropics (Moore and Lanier, 1961). Chloroquine resistance was discovered in Kenya and this led to its withdrawal as a first line drug in 1998. After 10 years, a re-introduction of Chloroquine into the system was still not possible due to resistance

(Mang'era *et al*., 2012). Resistance of *P. falciparum* to chloroquine and sulphadoxine-pyrimethamine resulted in the implementation of the Artemisinin-based combination therapy (ACT) as a first line treatment (Eastman and Fidock, 2009). 'Qinghao' *Artemisia annua*, a Chinese herb used to treat fevers, is the origin of artemisinin derivatives. The use of artemisinin has increased over the last two decades (Ridley, 2002). Resistance to artemisinin has also been observed with *Plasmodium falciparum* and has affected about 5 countries in the greater Mekong sub region. In Cambodia there is a high failure rate of treatments with ACT (World Health Organization, 2016). This is worrying as the consequence of resistant *Plasmodium falciparum* malaria poses a threat to the world, most notably in sub-Saharan Africa, which is worst hit by the burden (Talisuna *et al*., 2012). Drug resistance threatens the fight against malaria as resistance has emerged several times and has spread in endemic regions (Guyant *et al*., 2015).

Diagnosis also poses a challenge to the fight against malaria. Drugs are readily available even in villages, but proper diagnosis can help in providing the right medication (Bell *et al*., 2006; Collin and Fumagalli, 2011). Microscopy is considered a good technique for diagnosis, however, the technique needs trained technologists and the use of modern equipment (Mouatcho and Goldring, 2013). This might not be readily available in endemic regions. It is also time consuming and cannot detect sequestered *Plasmodium falciparum* parasites and less reliable in low parasitaemia (Mouatcho *et al*., 2013). The regions most affected by malaria are in Africa where the poverty rate is high and laboratory diagnosis is not accessible to majority of the population. Therefore diagnosis is symptom based hence, some resolve to self-medication. Illnesses with similar symptoms as malaria are sometimes mistaken and treated as such and the end result can be fatal. There is therefore a need for a diagnostic method that is fast and accurate.

The Rapid diagnostic tests (RDTs) use an immunochromatography process in which

an antibody binds to the parasites antigen resulting in a coloured band on the test strip. The technique requires the use of test kits of high standard which can be used for malaria diagnosis even in field conditions. The test kits can be easily transported, does not need electricity to function and very little training is required (Bell *et al*., 2006). The increase in parasitological screening for malaria from 40% in 2010 to 76% in 2015 (World Health Organization, 2016) was attributed to the use of RDTs. The setbacks for this technique are poor product performance and inability to test the quality of products. If the challenges were overcome, this method would lead to a parasite-based diagnosis and proper treatment that could be readily available and would improve fever management in endemic regions (Bell *et al*., 2006). In view of the challenges facing vaccine production and drug resistance, there is high research interest on development and implementation of alternative vector control programs.

1.3.3 Vector control

Vector control can be said to be any method that is targeted at limiting the ability of a vector to transmit disease (Karunamoorthi, 2011). To date vector control is still considered fundamental to malaria control (Barreaux *et al*., 2017; Karunamoorthi, 2014). Vector control against malaria can target the adult mosquitoes, the aquatic stage and can also be physical which involves source or breeding site reduction (Karunamoorthi, 2011).

Adult

Control measures against adult mosquitoes include the use of indoor residual spray (IRS), Insecticide treated nets (ITNs), long lasting insecticidal nets (LLINs) and repellants. These methods rely on the use of chemical insecticides. There are four classes of chemicals from which these insecticides are derived, they are organochlorine, organophosphates, carbamates and pyrethroids (Kisinza *et al*., 2017). Dichlorodiphenyltrichloroethane (DDT) is an effective

insecticide of the organochlorine class of chemicals. However, residues of DDT have been reported in human tissue and therefore the use of this insecticide has been generally restricted (Beard, 2006; Service, 2012). Organophosphates e.g. Malathion, fenitrothion or pirimiphos-methyl; the carbamates such as propoxur and bendiocarb; and the pyrethroids e.g. deltamethrin, cypermethrin, lambda-cyhalothrin are alternatives to DDT but these are quite expensive and less persistent therefore spraying has to be done more frequently (Hemingway and Ranson, 2000).

The use of indoor residual spray (IRS), involves spraying surfaces of walls, roofs, and ceilings of houses with residual insecticides (Service, 2012). This method is aimed at killing endophilic and endophagic mosquitoes. In terms of immediate results, IRS remains the most effective control method against malaria. The method was key in the eradication of malaria from Southern Europe, Russia, Asia and Latin America and some areas in South Africa (Karunamoorthi, 2011).

Insecticide treated nets (ITNs) and long lasting insecticide treated nets (LLINs) and repellants are personal protection measures. These are very powerful tool against malaria and easy to administer in terms of logistics. Pyrethroids are the only class of insecticides recommended by the World health organization (WHO) for use on the LLINs. This is because they are potent against mosquitoes, they are safe, have a low cost of production and also have the advantage of been able to repel mosquitoes unlike non-pyrethroids (Ngufor *et al*., 2017; World Health Organization, 2016). There is a wide coverage of ITNs and IRS distribution in the sub-Saharan parts of Africa ITNs has about 67% coverage and IRS 59% coverage. This had led to the aversion of about 89% of incidence and mortality rate caused by malaria (World Health Organization, 2015). Both LLINs and indoor residual sprays target endophilic vectors and the methods are used extensively. This has selected for outdoor transmission by exophilic species, those who are adapted to feeding outside (Benelli and Beier, 2017). There

is evidence of avoidance behavior in the Anopheles species as they are reported to avoid insecticides sprayed indoors (Reddy *et al*., 2011). The emergence of insecticide resistance in mosquitoes (Figure 1.3.4) has increased over the years and is becoming a threat to chemical based control programs (Ranson and Lissenden, 2016).

In view of this, there is a need for alternative vector control programs (Hemingway *et al*., 2000; Naqqash *et al*., 2016).

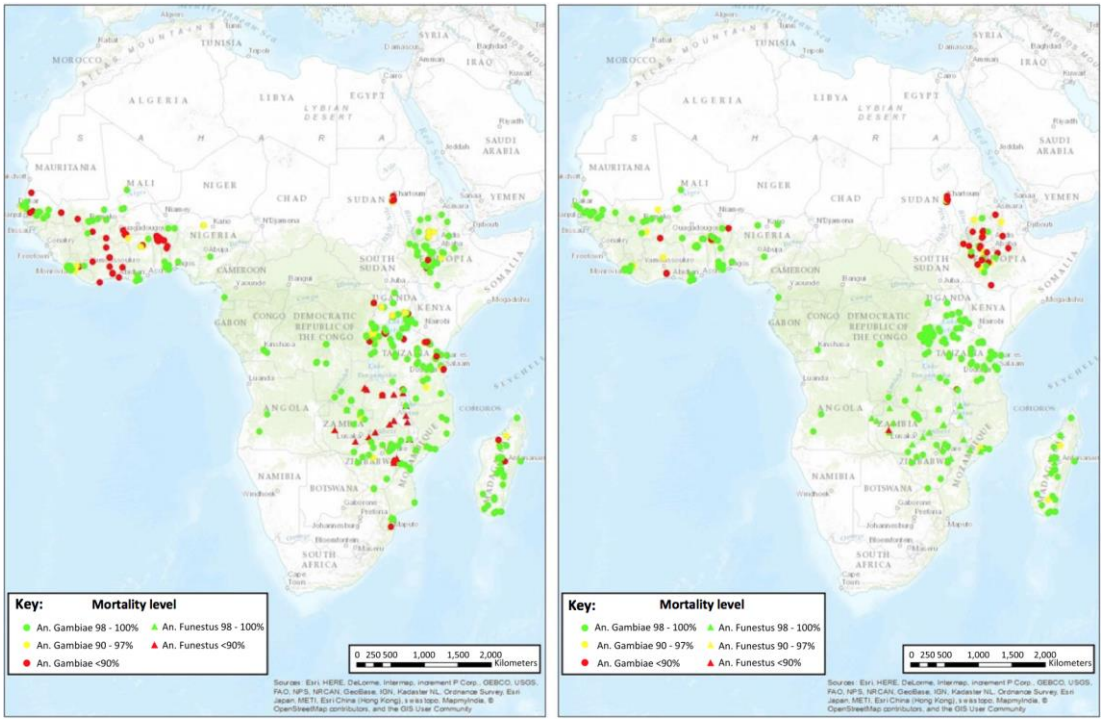

Figure 1.3.4: Map showing insecticide resistance using data collected from 2011 and 2015. Carbamate (diagram A) and organophosphate (diagram B) resistance in Africa. **Source:** (Ranson *et al*., 2016).

Larvae

Several interventions that focus on the larval stages have been in place from the 1920s. This includes draining of swamps, vegetation clearance and modifying river boundaries (Utzinger *et al*., 2001). Disruptions of larval habitats have proven successful in the control of malaria;

indeed, the disease was successfully eradicated in Italy through the drainage of swamps (Lindsay *et al*., 2002).

Larvicides include chemicals used to kill mosquitoes during the aquatic stages. They can be chemicals and oils such as methoprene. The use of active toxic agents from plant extract as larvicides has been considered (Ghosh *et al*., 2012).

Biological agents such as amphibian tadpoles, aquatic insects, larvivorous fish and odonate larvae are also used to control mosquito larvae (Kumar and Hwang, 2014). *Bacillus thuringiensis* can also affect the survival of mosquitoe larvae through the production of toxins (Tchicaya *et al*., 2009). The predator fish *Gambusia* species feed on mosquito larvae and is still used for larval control in some populations (Service, 2012). Most malaria control campaigns however focus on the adults.

1.3.4 Genetic control

New methods in the field of genetic engineering are providing opportunities to create mosquito strains to control natural vector populations. The discovery and use of genetic tools has made it possible to study and edit the mosquito genome (Alphey, 2014). Genetic control can be self-limiting or self-sustaining (Figure 1.3.5). The self-limiting approach involves the reduction of the number of vectors in a target area. In this method, a population of genetically modified mosquitoes are introduced to the population from time to time in order to retain the genetically modified strains in the target population. Sterile-male technique is an example of this approach. Though not entirely new and involves the use of radiation sterilized insects (Alphey, 2014). This method involves the use of modified or sterile mosquitoes to reduce population size to a level whereby the disease cannot be transmitted (Catteruccia, 2007). Population replacement or self-sustaining approach on the other hand is targeted at reducing

vectorial capacity; it may lead to the replacement of the niche and can also lead to new traits in the target population (Alphey, 2014).

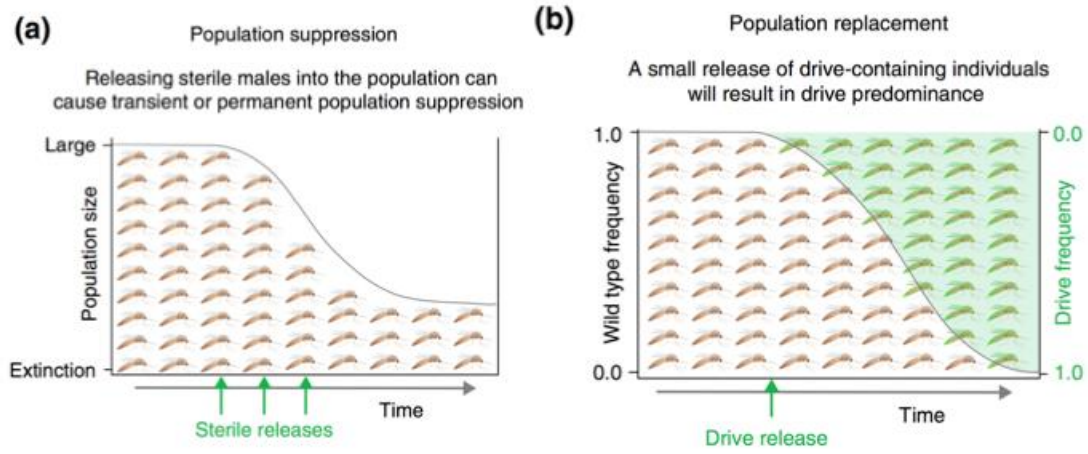

Figure 1.3.5: Methods for genetic control of vector populations. A = Population suppression by the release of large numbers of sterile males which mate with wild females leading to no progeny. B= Population replacement where traits carried by modified species are transferred to already existing species and eventually replacing existing population. **Source:** (Gabrieli *et al*., 2014).

The Self-sustaining approach involves the spread of an element into the target population. Introducing genes with anti-pathogenic properties so that organisms become refractory to disease transmission can modify vector populations as well (Catteruccia, 2007).

The aforementioned approaches involve the modification of organisms, and this may have an impact on the behavioural traits and life span of the vectors (Alphey, 2014). Techniques considered for genetic control include the use of *Wolbachia* endosymbiotic bacteria (Wiwatanaratanabutr *et al*., 2010). *Wolbachia* an endosymbiont of many insects is able to manipulate the reproductive systems of its host. It is also capable of inhibiting virus replication of the host. Its manipulation and blocking of dengue virus replication in *Aedes aegypti* have been studied. *Wolbachia* uses the host microRNA's in manipulating gene expression when the vector *Aedes aegypti* is infected with dengue and *Wolbachia*. The

upregulation of *Aedes aegypti* DNA methyltransferase gene (AaDnmt2) in the mosquitoes led to the inhibition of *Wolbachia* replication which in turn led to the replication of dengue virus. On the other hand, a downregulation of AaDnmt2 led to an increase in *Wolbachia* infection and a reduction in dengue (Zhang *et al*., 2013).

The Sterile insect technique (SIT) was proposed by Knipling in 1937 but it took about 2 decades before the technology was first applied (Knipling, 1955). The technique takes advantage of the ability of x-rays to destroy reproductive cells and cause sterility in insects (Oliva *et al*., 2014). SIT could be said to be a birth control method for insects that aims at reducing the population of target species (Oliva *et al*., 2014). The technique involves the mass production of the species of interest. In some cases sexes are separated and, the insects are sterilized. Sterilization is achieved using ionizing radiation or chemicals on one or both sexes after which they are released into the wild (Knipling, 1955). The possibility of using SIT for the control of mosquitoes is highly considered (Robinson *et al*. 2009) as the screwworm was successfully eradicated from the United States of America (Scott *et al*., 2017). SIT has been used in the eradication of insects of agricultural importance (Bloem *et al*., 2007; Reyes *et al*., 2007) SIT can only be successful when rearing and sterilization of the insect is done in an effective way. This would result in the release of insects that are fit to compete in all life's processes with their counterparts in the wild (Parker and Mehta, 2007). Since SIT involves the release of millions of insects to target areas, the process can be disturbing to the inhabitants; therefore there is a need for proper communication before implementation (Oliva *et al*., 2014). SIT is relatively environmentally friendly compared with other methods (Parker *et al*., 2007) because it is a self-limiting technique. A reduction in mosquito population would go a long way in the reduction of incidence of malaria, as there would be less host vector contact (Anguelov *et al*., 2012). SIT might be the most appropriate technique to curtail invasion of vectors into areas where eradication was successful. It is also ideal for newly

invaded populations (Oliva *et al.*, 2014). The SIT has also been shown to be better when combined with other genetic control methods like the use of *Wolbachia* and RIDL (Release of insects carrying dominant lethal) (Alphey, 2002; Zhang *et al*., 2016). SIT can only be effective if the biology of the mosquitoes is understood. The male genitalia rotation is a process of sexual maturity in male mosquitoes and has been studied in *An. funestus* where the time taken for the genitalia to be fully rotated was compared between laboratory bred and wild males. It was discovered that 36 hours after emergence, the genitalia of both laboratory and wild male mosquitoes attain complete rotation of 135-180^{o} at a temperature of 23±1^{o}C (Dahan and Koekemoer, 2014). This finding can be used in vector control techniques as the optimal timing for male releases can be utilized (Dahan *et al*., 2014). Molecular techniques such as gene editing have led to the improvement of these methods because many mosquitoes can be produced for implementation (Kuzma and Rawls, 2015), however, their success still depends on the ability of the vectors to mate and to transfer the heritable element (Alphey, 2014).

In addition to genetic control, other vector control methods considered are listed below.

Insecticide treated clothing

This method has been used for decades mainly by soldiers and sometimes during recreational activities where there is possible exposure to insect bites. The method is now considered for the control of malaria. There are also novel methods in preventing the loss of efficacy of this technique, which could occur when clothes are washed or when they get old (Banks *et al*., 2014).

Spatial repellants

Spatial repellants are chemicals that work in the vapor state and can help reduce contact with host, as it could be a good outdoor treatment (Achee *et al*., 2012). It would be of advantage to identify breeding sites and target those sites for spraying (Ribiero *et al*., 1996).

Plant based repellents

This has also been demonstrated to increase the efficiency of ITN when used in combination (Hill *et al*., 2007). Plant synthesized nanoparticles have also been used as control against mosquito larvae, pupae and adults (Benelli *et al*., 2017).

Livestock treatment

Mosquitoes that prefer to feed on livestock are likely not affected by control measures. Therefore, livestock shelters can be protected using nets and repellents. This could lead to a reduction in mosquito population (Njoroge *et al*., 2017). Livestock can also be treated directly with insecticides (Hewitt and Rowland, 1999; Mahande *et al*., 2007).

Aquatic predators

Backswimmers and water bugs could be used as biological means of controlling mosquitoes, which is achieved by predators feeding on the larvae (Bowatte *et al*., 2013; Murugan *et al*., 2015).

Ivermectin

Ivermectin is a drug used against onchocerciasis. It could also help in the reduction of malaria as it has been suggested that the reformulating and restructuring of ivermectin could have a residual impact on mosquitoes by causing death of blood feeding mosquitoes (Ōmura and Crump, 2017).

Lure and kill

This involves the use of an attractant e.g. attractive toxic sugar baits. This method will be cost effective as it involves the mosquitoes feeding on toxic substances disguised in the sugar solutions (Benelli *et al*., 2017).

Swarm sprays

Swarm spraying takes advantage of the swarming behavior of mosquitoes (Figure 1.3.6).

Swarm killing or trapping can be used as a vector control method (Diabate and Tripet, 2015).

Figure 1.3.6: A volunteer spraying swarms with a bomb spray. **Source:** (Sawadogo *et al*., 2017).

A research conducted in Bama, Burkina Faso reported a decrease in vector density. The region consists of clusters of villages named VK1-7. Swarms in VK5 village were targeted and sprayed leading to a decrease in vector density (Sawadogo *et al*., 2017). This indicates that targeting male mosquitoes for vector control is a worthwhile addition to the vector control methods already in use (Diabate and Tripet, 2015). In the study, the killing of mostly males and some females in swarms led to an 80% reduction in the vector population compared to the control, which is villageVK7 (Sawadogo *et al*., 2017).

1.3.5 Limitations of Vector control by release of modified insects

Limitations to vector control, aside those already mentioned in sections above, include limited knowledge of mosquito behavior and how environmental changes affect the mosquitos' ecological traits and mating (Benelli and Mehlhorn, 2016). Fitness of released males has been an important factor because *Anopheles* species mate in large swarms and males are subjected to strong competition to find mates (Alphey *et al*., 2010). Colonization has a negative impact on the genotypic and phenotypic quality of male mosquitoes. This can reduce their ability to

compete favorably for mates in the wild (Baeshen *et al*., 2014) though heterosis has been shown to improve reproductive traits in inbred populations (Ekechukwu *et al*., 2015). The role of larval development and pupation on mating behaviour in *Anopheles gambiae* s.s was demonstrated in a study by (Paton *et al*., 2013). They conducted a series of mating assays using laboratory bred and field strain mosquitoes. They discovered a lack of assortative mating in crosses between laboratory-bred males and females while assortative mating was observed when laboratory bred strains were crossed with field strain. The study also shows the impact of colonization on laboratory-bred strains.

Ecological factors such as biodiversity of wild populations and the effect of environmental changes on mosquito ecological traits affect the success of genetically modified mosquitoes (Benelli *et al*., 2016) There are over 30 primary vectors of malaria (Kiszewski *et al*., 2004). The *Anopheles gambiae* complex consists of morphologically indistinguishable species, with some genetic differences (Lanzaro and Lee, 2013). They are sometimes found in sympatry, exhibiting different resting, blood feeding and mating behavior. These complexities possess challenge to the implementation of genetic engineering for malaria control. If the differences between these vectors are not properly understood, the elimination of this disease by transgenes would require the release of all malaria-transmitting species in a given area, this can be a herculean task if not almost impossible (Gabrieli *et al*., 2014).

1.4 *Anopheles gambiae* population structure

Anopheles gambiae sensu lato

The *Anopheles gambiae* complex presently comprises eight sibling species (Coetzee *et al*., 2013). The complex is still undergoing speciation. The *An. gambiae* complex consists of morphologically indistinguishable species that have some differences in their behaviour,

capacity as vectors of malaria and genetic makeup. There are eight sibling species whose status was established by crosses leading to sterile F1 generation hybrids (Davidson, 1964) morphological features and through fixed chromosomal inversion (Coluzzi *et al*., 1979).

Anopheles gambiae was thought to be a single species and the major vector of malaria displaying much variation in its ecology. The larvae were found in different habitats, adults had different feeding behaviour as they fed either on humans or animals, indoor and others outdoor. Some mosquitoes were found to cause malaria while others did not (Mastbaum, 1957). In the 1940s and 50s research revealed variations among the species. The salt-water breeding forms in East Africa were observed to be different from those in West Africa. Marked differences were found in the egg characteristics (Thomson, 1951). In the 1960s, studies first uncovered genetic differences between populations by identifying different crosses that led to the production of sterile F1 males (Davidson and Jackson, 1962; Paterson, 1964). The complex was then known as having three freshwater, one mineral-water breeding species (known as A, B, C, and D) and two salt-water breeders. *An. merus* in East African, and *An. melas* in the west. These species were differentiated by male progeny sterility when crossed with other species within the complex (Davidson, 1964) (Edited in Coetzee *et al*. 2013). Coluzzi & Sabatini (1967) examined the chromosome from 8 species A strains and 6 species B as well as hybrids that were reared from larvae. The autosomes were found to possess some polymorphism, which could be linked to adaptation. The two species were identical in banding patterns on the autosomes but there were differences on the X chromosome of species A and species B (Coluzzi and Sabatini, 1967). Later the banding patterns of the giant polytene chromosome found in the salivary glands and in the female ovaries were used in differentiating species (Coluzzi and Sabatini, 1969).

Several publications were eventually released proposing names for the members of the *An. gambiae* complex, and names were adopted. *An. gambiae* s.s, Giles (originally species A)

which are from specimens collected in 1900 on McCarthy Island, Georgetown, the Gambia, *An. arabiensis* Patton (originally species B) found in more arid climate compared with *An. gambiae*. *An. quadriannulatus* Theobald (originally species C) (Figure 1.4.1) thought to be a highland species because of its distribution in Ethiopia. However, the Ethiopian population was shown to be a separate species based on cross-mating and chromosomal studies and is formally assigned a new name *An. amharicus* (Coetzee *et al*., 2013). *An. melas* Theobald was known as the West African saltwater breeder, *An. merus* Donitz the East African saltwater breeder and the name *An. bwambae* White was given to originally species D.

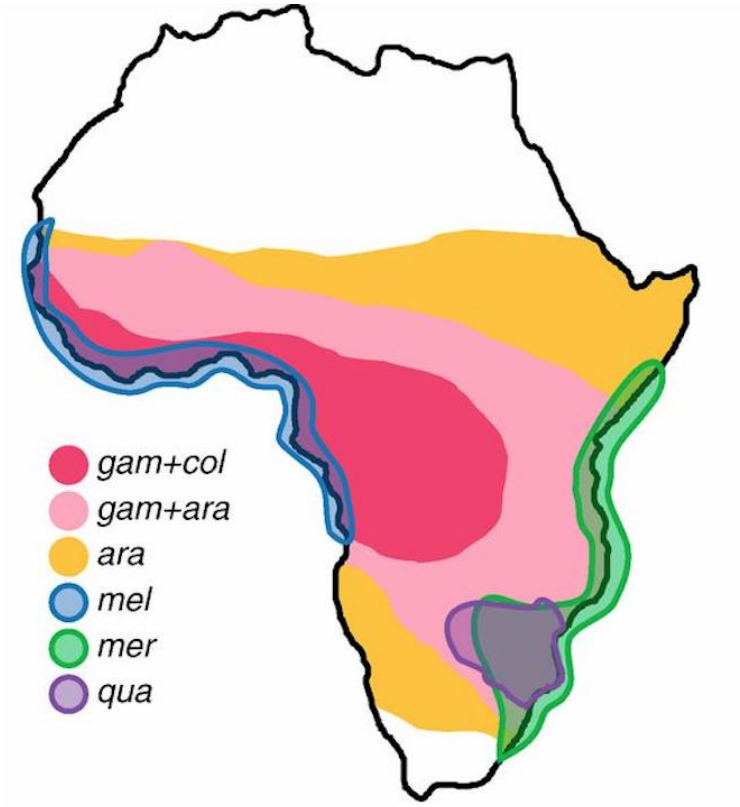

Figure 1.4.1: Distribution of the Anopheles gambiae complex. **Source**: (Fontaine *et al*., 2015).

In addition to the above a species *An. comorensis* Brunhes, Le Goff & Geoffrey was described and named based on the morphology of a single specimen from the Indian Ocean islands of Grande Comoro in the Mozambican channel (Edited from Coetzee *et al.*, 2013).

Several species can be sympatric in some regions and at least two are usually found occurring in malaria endemic regions (Scott *et al*., 1993). *An. arabiensis* and *An. gambiae s.s* have the widest distribution and are very important vectors of malaria (Coetzee *et al*., 2000).

Anopheles gambiae sensu stricto

The most anthropophilic malaria vector in the *An. gambiae* complex is *An. gambiae* s.s which is still adapting to man and his environment (Favia and Louis, 1999). In Africa, malaria transmission is achieved largely by this species. Its ability to transmit malaria is enhanced by its wide distribution, their ability to live around human habitats and even enter houses at night to feed (Hodges *et al*., 2013). *An. gambiae* s.s can be found in semi-arid regions characterized by strong seasonal variations in temperature, humidity and rainfall (Charlwood *et al*., 2000). Many polymorphisms exist within the species including chromosomal inversions, this has led to its subdivision into several chromosomal forms (Coluzzi *et al*., 1985) and also two molecular forms that can be recognized by differences in their rDNA sequences as discussed below (Fanello *et al*., 2002).

Chromosomal and Molecular forms of Anopheles gambiae s.s.

Genetic and ecological divergence among species and the evolution of reproductive isolation among populations within a species can be facilitated by numerous mechanisms that could include chromosomal inversions and pericentric areas characterized by reduced recombination. Paracentric chromosomal inversions occur when a section of the chromosome breaks and rearranges itself, the chromosome fragment that does not include the centromere is reinserted in a reverse order leading to an inverted gene order. When this happens, gene expression can be changed in the area of the inversion, and crossing-over may be reduced during meiosis, which can locally result in reduced recombination (Lee *et al*., 2013a).

The *An. gambiae* genome is organized in three chromosomes, two sub-metacentric autosome, and an X/Ychromosome with males being the heterozygote sex. The autosomes are divided into two arms at the centromere, the longer arm referred to as right arm and the shorter arm, referred to as left arm (Figure 1.4.2) (Lanzaro *et al*., 2013). Inversions are not randomly distributed but occur more often on the right arm of chromosome 2 (Pombi *et al*., 2008) The chromosome arm 2*R* has 18/31 of the common polymorphic inversions, and this represents only about 30 percent of the complement (Pombi *et al*., 2008). The frequencies of inversions differ across geographical regions, suggesting a role in adaptation (Coluzzi *et al*., 2002; Lee *et al*., 2009; Touré *et al*., 1998). In *An. gambiae* sensu stricto, 6/7 common polymorphic inversions occur on 2*R* and most are considered markers of ecological adaptation that increase fitness of the carriers of alternative karyotypes in contrasting habitats (Pombi *et al*., 2008). The studies of *An. gambiae* in West Africa based on chromosomal inversions showed deficits of heterozygotes (with respect to corresponding Hardy-Weinberg expectations) in a number of inversions. The chromosomal forms were defined based on the frequencies of five paracentric chromosome inversions on the right arm of the chromosome 2 (inversions 2*Rj, b, c, d,* and *u*) and also on the left arm of the chromosome 2 (inversion 2*La*) (Coluzzi *et al*., 1985).

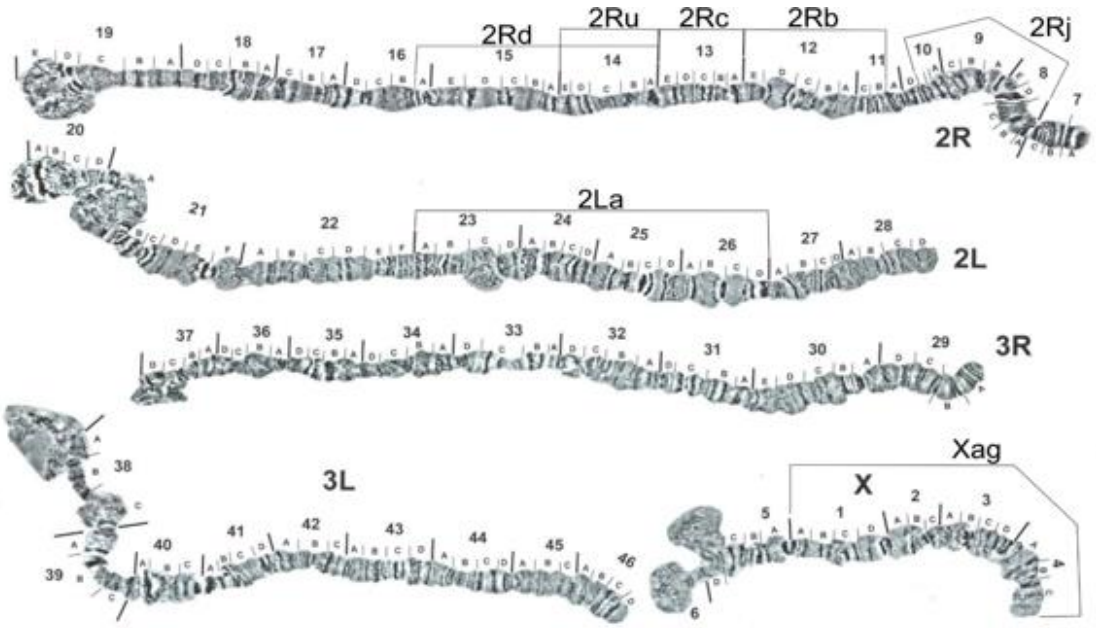

Figure 1.4.2: A photomap showing the polytene chromosome of *Anopheles coluzzii*. Six inversions used in identifying chromosomal forms are shown located on chromosome 2. **Source**: (Lee *et al*., 2013a).

Anopheles gambiae was subdivided into several chromosomal forms of uncertain taxonomic status named: Savanna, Mopti, Bamako, Forest and Bissau (Coluzzi *et al*., 1985). Savanna has high frequencies of 2*Rb* and 2*La* inversions and also polymorphism involving 2*Rcu* arrangement and also in *j, d* and the rare *k* inversion. Mopti shows high frequencies of 2*Rbc*, 2*Ru* and nearly fixed for 2*La* (Table 1.4.1). Bamako is characterized by fixed 2*Rjcu* arrangement and polymorphism in 2*Rb* inversion, Forest is characterized by the non-inverted arrangement 2*R*+/+, 2*L*+/+, with same polymorphism in the two molecular forms due to inversion 2*Rb* or 2*La* and. Finally Bissau is characterized by high frequencies of the 2*Rd* inversion and standard 2*L*+ arrangement (Della Torre *et al.,* 2002).

Table 1.4.1: Typical and less frequent arrangements of chromosomal forms on the 2R and 2L arms of the second chromosome in the five chromosomal forms described by Coluzzi, Patrarca and Di Deco 1985 **Source:** (Lanzaro and Tripet, 2004).

Chromosomal forms	Typical arrangements		Less frequent arrangements	
	2R	2L	2R	2L
Forest	2R+	2L+	2Rb/d/+	2La
Bissau	2Rd	2L+	2R+	2La
Savanna	2Rb/+	2La/+	2Rcu/bcu/bd/bcd/d/j/jb/jbd/jbcu/jcu/bk/+	
Bamako	2Rjcu/jbcu	2La		
Mopti	2Rbc/u	2La	2R+	2L+

The geographical regions where the forms were first collected were considered while naming the chromosomal forms, putting into consideration the specific habitat in which they were collected (Coluzzi *et al*., 1985).

Efforts made in trying to identify fixed nucleotide differences between chromosomal forms in Mali led to the definition of the molecular forms M and S. Fixed differences were found in the sequence in the X-linked intergenic spacer of the multicopy 28S ribosomal DNA in the pericentromeric region of the X chromosome (Favia *et al*., 1997). In Mali and Burkina Faso, the M molecular form corresponds to the Mopti chromosomal form while the S molecular form corresponds to both the Bamako and the Savanna chromosomal forms. The close correspondence between the chromosomal and molecular forms is lost outside Mali and Burkina Faso as the Savanna chromosomal forms also corresponds to M molecular form and the Forest chromosomal form can be either M or S (Della Torre *et al*., 2001; Wondji *et al*., 2002). The Savanna chromosomal form is dependent on rainfall for breeding, therefore it strives well during the rainy season. In the forest region, the S form corresponds to the forest chromosomal form and is sometimes sympatric with the forest M molecular form. The M molecular form, also characterizes the Mopti chromosomal form, which is found in drier areas, usually where irrigation occurs. The larvae's ability to adapt to different environmental

and climatic conditions explains why they are usually present and active all year round, which then enhances active transmission of malaria even during the dry season (Slotman *et al*., 2006). The S form savanna populations are widely distributed while the M forms are found in Western parts of Africa. The M and S forms are found in sympatry in many places in West Africa. Reproductive isolation still exists and hybrids are rare amongst the M and S forms even in areas where they are found to be sympatric (Della Torre *et al*., 2005).

The chromosomal forms may seem to be disconnected from the molecular forms however, they usually do not belong to a genetically distinct population, but represent polymorphisms within populations of the M and S molecular forms resulting from different selection pressures acting on frequencies of inversion alleles (Della Torre *et al.,* 2002) were given species status in 2013 and named *Anopheles coluzzii* and *Anopheles gambiae* (Coetzee *et al*., 2013).

Recently, the M and S molecular forms of *Anopheles gambiae* s.s have been named as individual sibling species based on molecular and bionomical evidence (Coetzee *et al*., 2013). *Anopheles gambiae* molecular M form is now named *An. coluzzii* Coetzee & Wilkerson, while the S form retains the name *An. gambiae* sensu stricto Giles (Coetzee *et al*., 2013). Henceforth we will refer to the recently diverged molecular forms (M and S molecular forms) as *Anopheles coluzzii* and *Anopheles gambiae* respectively.

1.5 Divergence in the recently named sibling species: *Anopheles gambiae and Anopheles coluzzii*.

Speciation is an important process of evolution and it is thought to happen when complete reproductive isolation occurs between incipient species (Figure 1.4.3). There two theories for speciation, the first happens when gene pools are separated by geographical and reproductive factors and there is no genetic exchange between the populations. The second theory supports

the presence of gene flow in speciation. The presence of gene flow indicates incomplete reproductive isolation therefore some exchange of genetic materials between the incipient species. The theory postulates that speciation in the presence of gene flow results in introgression between gene pool; it also leads to the differentiation of some regions of the genome (Crawford *et al*., 2015).

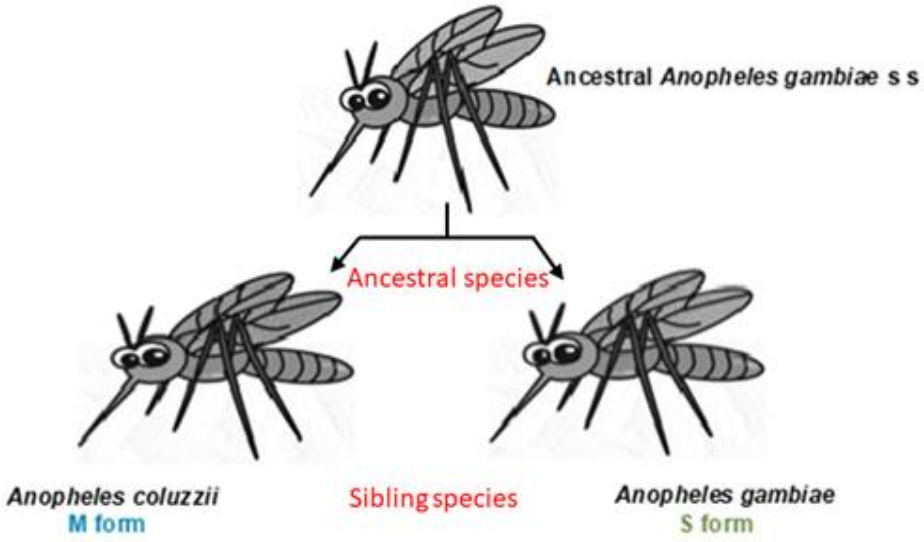

Figure 1.4.3: Sympatric speciation in *Anopheles gambiae* s.s. The sibling species are morphologically identical, often found in sympatry with restricted gene flow and low rates of natural hybridization.

Given that no intrinsic reproductive barriers between *An. coluzzii* and *Anopheles gambiae* have been detected, divergence between the two is thought to be made possible through the joint effect of assortative mating and some level of extrinsic post-mating isolation driven by differential ecological adaptations and genetic factors (Aboagye-Antwi *et al*., 2015).

1.5.1 Ecological differentiation

Coluzzi *et al* (2002) argued that the *An. gambiae* complex is a result of a speciation process caused by the influence of humans on the environment; this includes agricultural activities, which started after the Neolithic revolution where human lifestyle of hunting and gathering

changed to agriculture and settlement. They postulate that the ancestral form of anthropophilic *An. gambiae* associated with the equatorial forest spread from the rainforest into the savannah through close association with humans (Coluzzi *et al*., 2002). However, the exact geographical origin of the *An. coluzzii* and *An. gambiae* remains unclear (Choi and Townson, 2012). An environment either supports or does not support the successful breeding of a given species (Obsomer *et al*., 2013). *An. gambiae* complex can adapt to different micro and macro environmental conditions, this is seen in its wide distribution and its ability to survive in places with unfavourable climatic variations (Coluzzi *et al*., 1985).

An. coluzzii and *An. gambiae* can be found in semi-arid regions where temperature, humidity and rainfall change according to season (Charlwood *et al*., 2000; Toure *et al*., 1996). *An. gambiae* is found all over sub-Saharan Africa while *An. coluzzii* is found only in West Africa (Lehmann and Diabate, 2008). The reason behind the geographical spread remains unknown (Lehmann *et al*., 2008). This implies that populations adapted to one type of condition would be genetically different from populations adapted to different conditions (Lanzaro *et al*., 2004). In places where the sibling species coexist, *An. coluzzii* predominate in higher latitudes while *An. gambiae* predominates in humid lower latitudes. The difference in aridity tolerance between the sibling species agrees with the seasonal fluctuations in numbers *An. coluzzii* and *An. gambiae* (Touré *et al*., 1998).

Adult *An. gambiae* sibling species can be found in sympatry. Larvae are said to inhabit different ecological habitats though they can also be found together in temporary collections of water during the rainy season (Gimonneau *et al*., 2010). *An. coluzzii* larva prefers permanent freshwater habitats, for this reason they are found in urban reservoirs, irrigated areas, etc. and they are found all year round (Della Torre *et al*., 2005; Gimonneau *et al*., 2010). A study carried out in Ghana showed the preference of *An. coluzzii* to permanent breeding sites, which are urban farms, puddles, swamps, and ditches. Larval density was high

during the rainy season compared to the dry season (Mattah *et al*., 2017). Larval adaption to different water bodies is one trait in which ecological differentiation between the sibling species has been shown. The *An. coluzzii* are more, adapted to water bodies that contain predators compared to the *An. gambiae* (Diabaté *et al*., 2008). The sibling species also employ other adaptive strategies for survival.

A study was conducted investigating whether the population of *An. coluzzii* observed during the early rainy season was as a result of estivation or migration. They found that *An. coluzzii* estivate during the dry season and emerge during the early rainy season. The species also migrates during the late rainy season. Morphological changes have also been reported in adults (Lehmann *et al*., 2017). The shape of the wings of *An. coluzzii* was reported to be determined by its geographical and climatic conditions therefore its body size and surface changes in response to its environment (Hidalgo *et al*., 2015). The importance of growth conditions and water availability on adult physiological status and subsequent resistance to desiccation in the Mopti chromosomal forms has been reported, it was discovered that glycogen and water content were most used to avoid desiccation however this depends on the mosquitoes phenotypic quality (Aboagye-Antwi and Tripet, 2010).

Predation on the immature stages of *An. gambiae* complex is said to cause disruptive selection leading to ecological divergences between the *An. coluzzii* and *An. gambiae*, which then promotes speciation (Roux *et al*., 2013). In a study carried out with the progeny of female mosquitoes captured in Burkina Faso, *An. coluzzii* and *An. gambiae* immature forms were exposed to the same predator, *Anisops jakzewskii,* in the same habitat, *An. gambiae* were observed to have a higher predation rate compared to *An. coluzzii*. The reduced activity rate of the *An. coluzzii* when exposed to predators was thought to be an adaptive feature to avoid predation common to its natural habitat (Gimonneau *et al*., 2010). *An. coluzzii* can therefore avoid predators more than their sibling species the *An. gambiae* (Diabaté *et al*. 2008b;

Gimonneau *et al*. 2010).

Strong association of chromosome inversions with dry or wet habitats has also been reported. The Mopti chromosomal forms have high inversion frequencies at *a* and *bc*, associated with its ability to resist drought (Coluzzi *et al*., 1985; Touré *et al*., 1998). The forms are dominant in drier parts of Mali and their distribution increases during the dry seasons (Touré *et al*., 1998). The Bamako and Savannah forms on the other hand do not have the combination of the inversions and are better adapted to wet areas and increase in number during the rainy season (Touré *et al*., 1998).

The choice of an oviposition site is important to the female mosquito for the survival of its offspring. Important factors for an oviposition site include the presence of nutrients, possibility of desiccation, predation and competition (Diabaté *et al*., 2008). The presence of cospecific species in an oviposition site could indicate stability of the site for oviposition (Munga *et al*., 2006) it could also discourage oviposition as competition could occur (Blaustein *et al*., 2004). It has been reported that vegetative cues could be involved in the choice of larval habitats and oviposition sites, as gravid female *An. coluzzii* were found to be attracted to some grass volatile (Asmare *et al*., 2017).

1.5.2 Genetic differentiation

Genetic differentiation between the recently diverged sibling species often found in sympatry, was first identified within the division 6 and part of division 5 of the X chromosome. This region is adjacent to the centromere and the region includes the ribosomal deoxyribonucleic acid (rDNA) that was used in defining the molecular forms. Very little recombination occurs in regions adjacent to the centromere (Slotman *et al*., 2006). It is suggested that low recombination rates enhance the build up and maintenance of genes responsible for reproductive isolation in partially isolated populations (Slotman *et al*., 2006). The control of

malaria through conventional or novel technologies must take the genetic differentiation between the sibling species of *An. gambiae* into consideration (Esnault *et al*., 2008). We need to understand the genetic differentiation that occur between the sibling species as this is very important in vector evolution and can affect malaria control, including resistance to insecticides (Nwakanma *et al*., 2013). These differences reflect reproductive isolation and have consequences on malaria control programs based on the introduction and spread of transgenes in mosquitoes (Tripet *et al*., 2005).

Research over the years has shown that there are differences in some specific parts of the two sibling species genomes where very low recombination occurs. These islands of divergence are thought to contain genes responsible for speciation in the species. The presumed islands of speciation are the 3-pericentromeric islands of divergence on the X, 2L, 3L chromosomes and some other small Islands located in the vicinity of inversion breakpoints (Turner *et al*., 2005; White *et al*., 2010). There is an island of speciation model, which suggest that diverged regions contain genes that are sustained by selection in the presence of gene flow. Another model states that there is no gene flow between *An. coluzzii* and *An. gambiae*, it speculates that divergence islands have nothing to do with speciation this is called the "incidental island" model (Lee *et al.*, 2013b).

1.6 Speciation genomics

Islands of speciation

An. coluzzii and *An. gambiae* are sibling species with some level of gene flow occurring among them, which prevents their genomes from diverging completely, it is assumed that there are small regions of divergence between the species (Lanzaro *et al*., 2013) The *An. gambiae* genome was sequenced in 2002, 14,700 genes and more than 16,100 transcripts have been predicted. About 40% of the gene products have known functions and the 60% still

unknown. Therefore information on the function and expression of these genes is very important (Dana *et al*., 2006) as they could serve as pointers to conditions of speciation.

So far different molecular studies have focused on distinguishing genetic units of *An. gambiae s.s* and to know whether the different units are evolving into individual species (Della Torre *et al*., 2002). Turner *et al* (2005) hybridized population samples of genomic DNA from *An. coluzzii* and *An. gambiae* to Affymetrix GeneChip microarrays and discovered three significantly genetically differentiated regions totaling about 2.8 Mb. Two of the regions on chromosome 2L and X contained 50 and 12 predictable genes, respectively. Sequencing of genes within these loci revealed fixed differences and unique polymorphisms. The third region located on chromosome 2R contained only five predictable genes. These regions are considered as Speciation Islands because they remain different between the sibling species regardless of gene flow and are therefore thought to contain the genes that are responsible for reproductive isolation (Turner *et al*., 2005).

Reproductive isolation is expected to be as a result of changes at small number of loci, and these regions together contain only 67 predictable genes (Turner *et al*., 2005). Research on these regions should reveal genes responsible for reproductive isolation between *An. coluzzii* and *An. gambiae* sibling species (Turner *et al*., 2005). A study in Guinea Bissau showed high hybridization (19-24%) of M/S hybrids in samples collected for 2 consecutive years. The hybridization coincided with very low levels of genetic differentiation between the M and S forms when analyzed using microsatellites mapped at chromosome-3. The X-linked AGX678 locus was an exception because there was high differentiation measured there. This locus maps near the centromere of chromosome X, is a low recombination region in which selection is likely to promote divergence between M and S forms. These results strongly suggest that the degree of isolation between M and S forms considering the units of incipient speciation within *An. gambiae* is not the same throughout the species distribution range

(Oliveira *et al*., 2008). Using microarray-based divergence mapping, a third region was also discovered on chromosome 3 (White *et al*., 2010).

Gene expression profiles were examined in two independent colonies of *An. coluzzii* and *An. gambiae* using oligonucleotide microarrays, with the aim of identifying possible genes involved in the ongoing speciation process. The quantitative real-time reverse transcription polymerase chain reaction of RNA samples from laboratory bred and wild mosquitoes were analyzed, genes that were expressed differently represented about 1-2% of all expressed genes. The genes were not appropriately represented on the X speciation island. Most of the genes were located outside the island of speciation (Cassone *et al*., 2008).

In another study, the standard chromosomal arrangement of the Savanna form was compared with genomes homozygous for *j, b, c* and *u* inversions in the Bamako forms. This was to identify regions where their genomes differ with respect to inversion polymorphism. They found levels of divergence between the two sub-taxa though lower than expected and mostly around inversion breakpoints. They also found that most of the divergent regions were located on the X chromosome, which contains half of all significant diverged regions. Two X-linked genes (a heat shock protein and P450) encoding genes involved in reproductive isolation between *An. coluzzii* and *An. gambiae* were also significantly diverged between the two species. This result showed that the genes responsible for reproductive isolation are likely to be located on the X chromosome and also likely to be responsible for reproductive isolation between the molecular forms and between the Bamako and Savanna chromosomal forms (Lee *et al*., 2013a). The relationship between the X-linked region and assortative mating has been reported, this they did by the introgression of the region from *An. gambiae* into *An. coluzzii*. Several mating assays were carried out and the recombinants mated more assortatively with those with same X-island as theirs (Aboagye-Antwi *et al*., 2015). Therefore genes on the sex chromosome could be the driving force for speciation (Lee *et al*., 2013a).

Gene flow

An. gambiae and *An. coluzzii* are found in sympatry in many parts of West Africa, reproductive isolation still exists and hybrids are rare even in areas where they are found to be sympatric (Della Torre *et al*., 2005). Several reports have shown that there is restricted gene flow among these species in areas where they occur in sympatry (Della Torre *et al*., 2005, 2002). A study in Mali suggested that the molecular forms are almost completely reproductively isolated (Favia *et al*., 1997). However these species mate freely in the laboratory producing offspring's that are viable (Taylor *et al*., 2001). Some level of gene flow was observed between the two molecular forms. Mating was assortative between the two forms except in very few cases (Tripet *et al*., 2001). In the study, sperms were collected from mated females and analysed. All 193 *An. coluzzii* female samples analysed mated assortatively. Mating was almost exact except for samples in *An. gambiae* females analyzed. Their result showed the presence of assortative mating and some level of gene flow between the two species (Tripet *et al*., 2001).

Several other reports have shown that there is restricted gene flow among these species in areas where they occur in Sympatry (Della Torre *et al*., 2005, 2002). Evidence shows that the extent to hybridization varies significantly across the geographic range in which the two species are sympatric (Della Torre *et al*., 2005). In Guinea-Bissau 19-24% hybridization was observed (Oliveira *et al*., 2008). In a study in the Gambia, Senegal and Republic of Guinea, different frequencies of *An. coluzzi* and *An. gambiae* hybrids were observed ranging from 5% to 42%. In the Gambia, the hybrids had same frequencies throughout periods of marked seasonal variations during the year round sampling over 2 years (Nwakanma *et al*., 2013). The chromosomal forms Bamako and Mopti rarely mate in the wild even though both forms have been seen to hybridize with the Savanna form. It was observed that the percentage of hybrids in the population range from 0 to 11% and this depends on

location, time of the year and the form (Touré *et al*., 1998). Patterns of genetic divergence (F_{ST}) between sympatric populations of Bamako and Mopti chromosomal forms from five sites in Mali were compared using microsatellite loci within the *j* inversion of chromosome 2, which is fixed for Bamako but absent in Mopti forms, and also microsatellite on chromosome 3 a region without inversions. The estimates of genetic diversity and F_{ST}'s show the occurrence of genetic exchanges between forms for chromosome 3 and very little for the *j* inversion. This shows the role of inversions in speciation (Tripet *et al*., 2005). These patterns of gene flow infer that mosquito genetic control project should target areas of the genome situated outside inversions if they are to spread effectively (Tripet *et al*., 2005). Studies on levels and pattern of gene flow within and among populations is important as it would provide information about conditions of gene flow in the wild and this knowledge would be useful for the application of new molecular level approaches (Lanzaro *et al*., 2004).

1.7 Mating and mechanism of reproductive isolation

Sexual selection has now attained widespread recognition as an engine of speciation, perhaps the most important of the forces that generates new species (Ritchie, 2007). Assortative mating can be said to be the ability of an organism to recognize and mate with its own kind. Assortative mating has been recorded in insects such as the soldier beetle *chauliognathu spp*, a cryptic taxon (Bernstein and Bernstein, 2003). Very strong assortative mating occurs between *An. coluzzii* and *An. gambiae* (Tripet *et al*., 2001). Premating isolation is not complete, as a number of hybrids have been seen to occur in nature showing some level of gene flow (Taylor *et al*., 2001; Tripet *et al*., 2001). Hybrid frequencies have been reported to be low in many studies around endemic regions in West Africa (Della Torre *et al*., 2005; Esnault *et al*., 2008) and Burkina Faso (Costantini *et al*., 2009), these shows that assortative mating leads to pre-reproductive barriers. Low hybridization rate of (2.16% and 1.86%)

among *An. coluzzii* and *An. gambiae* from two inland transects which had different ecological characteristics in South-eastern Senegal was reported (Niang *et al*., 2014). Such results signify reasonable reproductive isolation between the sibling species and support the process of ongoing speciation in these inland areas. The result shows a possibility of a post-mating process, which acts at the larval stage. Post-mating reproductive isolation was not observed in a study where mating assays was carried out on same species and on cross-mated individuals of *An. coluzzii* and *An. gambiae*. Cross-mated females differed from those mated with same species by the number of eggs hatched and larval survival. This difference could be due to hybrid vigour (Tripet *et al*., 2005). The results from the study suggest that the mechanism responsible for reproductive isolation is pre-zygotic (Pennetier *et al*., 2010; Sanford *et al*., 2011).

Post-zygotic isolation between the forms have not been demonstrated (Diabaté *et al*., 2007) though low level of hybridization was reported in the larval stage and also in adults collected indoors. Hybridization was low compared to cross-mating rates and this shows a possibility of a post mating process that acts at the larval stage (Niang *et al*., 2015). A divergence island SNP assay was used by Lee *et al*. (2013b) to look at the temporal distribution of hybrid genotypes. It was discovered that assortative mating is not stable and sometimes the level of assortative mating decreases leading to widespread hybridization. The results showed that hybrids have fitness disadvantage but some hybrid genotypes are viable. Through a temporal analysis, Lee *et al.* (2013b) showed that assortative mating is not constant and sometimes random mating occurs which leads to large numbers of hybrids with most hybrids having a fitness disadvantage.

The question then is how reproductive isolation is able to occur despite gene flow (Slotman *et al*., 2006). If strong assortative mating also occurs in nature, then there must be a way in which cryptic species recognize their own kind. Knowledge of the signals leading to

recognition would be important in unraveling phenotypic differences between forms, which can then lead to the precise areas in the genome responsible for assortative mating (Lanzaro *et al*., 2004). In a view to find answers to such questions, Tripet and colleagues (Aboagye-Antwi *et al*., 2015) investigated how the divergence islands influence sympatric speciation. This was done using an experimental functional genomics approach. They hypothesized that the largest presumed speciation island, which is on the X chromosome, was responsible for protecting assortative mating genes despite on-going gene flow. This is because the island combines pericentromeric recombination suppression with the hemizygosity and decreased recombination typically associated with the X chromosome.

The next step was to introgress the *An. gambiae* X-linked Island of speciation into an *An. coluzzii* genetic background. This led to the creation of recombinant strains (RbSS), which had the same genetic background with the *An. coluzzii*-type (RbMM) but differed only in the X-chromosome islands of speciation. A series of assortative mating experiments to test the association of the X-island molecular type with mating preferences of recombinants and parental strains were carried out. The mating choice preference of females and males were tested between mates with the same and also a different type of X-Island. Females from the RbMM were seen to mate highly assortatively. RbSS females were also observed to mate entirely with males with the same X-type Island. Their results showed close association between pre-mating isolation genes and the X-island, which supports the hypothesis that pericentric regions can protect genes responsible for pre-mating isolation and this in turn can lead to sympatric speciation. It also shows that hemizygosity and low recombination rates of sex chromosomes encourages the build up of pre and postmating isolation genes. The results suggest that the low-recombining pericentromeric X-island enables these incipient species to maintain their genetic integrity in parts of Africa where they are sympatric. The X-island is larger and contains more genes compared to the other two-pericentromeric islands described

between the *An. coluzzii* and *An. gambiae* (Turner *et al*., 2005). A comparison of sympatric populations from Ghana and the recombinant strains was done to estimate the size of island, and it was estimated to be 6Mb long extending from positions approximately 18.1 to 24.2Mb (Aboagye-Antwi *et al*., 2015). Full-genome sequencing of the RbSS, RbMM and Mopti strains were carried out. They also compared protein-coding differences identified between the recombinants with those from two sympatric populations of *An. coluzzii* and *An. gambiae* from Southern Ghana. Some interesting protein-coding changes were observed between the two recombinant strains and also nearly fixed in the two sibling species found in sympatry. These changes affected 12 genes in the cryptic taxa. Six out of the 12 genes that were identified have been assumed to perform certain biological functions but the function of the others is not yet known (Aboagye-Antwi *et al*., 2015). Therefore the genes that have been seen to differ between the recombinant strains, and also between *An. coluzzii* and *An. gambiae*, could be responsible for some mating behaviour in these species. They could also be contributors to the ecological speciation, such as larval adaptation of the sibling species (Aboagye-Antwi *et al*., 2015; Diabate *et al*., 2005). It is therefore important to study the role of these reproductive isolation genes in the two sympatric sibling species. This could reveal the genetic basis of assortative mating, lead to a better understanding of speciation and also the production of release males that effectively mate with the intended target populations (Tripet *et al*., 2005).

1.7.1 Swarming in *Anopheles gambiae* and *Anopheles coluzzii*

Mosquitoes usually mate in flights over landmarks. The mating structure has been described, as lek-like where female choice seems to be absent. Mating has been observed to occur in aggregations called swarms (Diabaté *et al*., 2011) Swarming is a mating system based on aerial aggregations that serve as a meeting place for mate-seeking females (Diabaté *et al*.,

2006). Swarms are important because they provide opportunity for mate choice. In *An. gambiae* swarming is an avenue for premating reproductive isolation between sibling species. It gives individuals, both males and females the opportunity to choose the right mates. Swarms are solely for mating, as they do not contain any resources (Diabate and Tripet, 2015).

Female *An. gambiae* mate only once in a lifetime (Tripet *et al*., 2003) that invariably would reduce the number of females available for mating in relation to males. Males mate multiple times therefore it is likely to find more males than females available for mating. It is necessary for both sexes to arrive to the swarm in good time, though most importantly for males. Arriving the swarms in good time gives the male the opportunity of attracting available females (Diabate and Tripet, 2015). Most swarms are composed of males, with females approaching the swarm, acquiring a mate and leaving in-copula. This selection of mates usually occurs through swarm segregation evidenced by specific site selection by mating individuals, over certain landmarks called swarm markers (Charlwood *et al*., 2002; Diabate and Tripet, 2015).

Swarm markers are used in the location and creation of swarms (Diabaté *et al*., 2011) In the field, swarms include stacks of wood and trash, wells, footpath junctions and grasses. Having an illusion of landscape discontinuity seems to attract the males (Figure 1.7.1).

Figure 1.7.1: Pictures of different swarm markers. Arrows point to positions where swarms are situated. **Source:** (Diabaté *et al.,* 2009).

An. coluzzii is apparently more attracted to these varied markers while *An. gambiae* is mostly found on bare ground (Diabate and Tripet, 2015). A study carried out in Dongola northern Sudan by Hassan *et al*. (2014) reported that swarms were seen at sunset around larval habitats in irrigated areas. The report also states that swarming stopped when it was dark and the time used for swarming was about 21-25 minutes. Palm trees, bare ground and manure were used as swarm markers and mosquitoes were observed leaving in copula going in the direction of sunlight. They reported that copulation occurred about 12-15 minutes after the onset of swarming. Swarming behavior, swarm initiation, the peak of swarming and the termination time was observed in semi-field enclosures and it coincided with that in the wild. They suggested that these enclosures could be used to study mating ability in strains (Hassan *et al*., 2014). Strong correlation between swarm size and mating success between swarms has been reported (Diabaté *et al*., 2011). Females were reported to be attracted to large swarms. Swarm

markers did not have a role to play regarding the size of the swarm and did not affect mating success; this could mean that they are attracted to the markers but respond to some other factors in the sites. Cues suggested to be involved in swarming include visual and acoustic. Visual cues are said to be important in locating landmarks and swarm formation. Males are attracted to landmarks used as swarm sites. These markers can be dark light contrast on the ground or landscapes that are discontinued (Diabate and Tripet, 2015; Facchinelli *et al*., 2015). Swarming was stimulated among *An. gambiae* G3 strain in large indoor cages. They observed that fading ceiling lights at dusk alone could not initiate swarming but when a dark foreground, contrasting illuminated background and a contrasting landmark was included, swarming was initiated. Swarming was initiated using this setting in mosquitoes that have been colonized for 30years and insemination frequency was reported to have greatly increased (Facchinelli *et al*., 2015).

Swarms are spherical with most individuals at the centre. It is assumed that the density of individuals at the centre could be beneficial for accessing females (Diabaté *et al.,* 2009). It could also be that male's position themselves in areas (the centre) where they have proximity to all sides of the swarm, which increases their chances of mating with females approaching the swarm since they rely on cues (Diabate and Tripet, 2015). Patterns of attraction may add to specific mate recognition systems. This segregation is one of the factors leading to reproductive isolation between the sibling species (Diabaté *et al.,* 2009). Among cryptic taxa, spatial swarm segregation serves as a barrier to hybridization, therefore gene flow, though not a major mechanism behind reproductive isolation (Diabate and Tripet, 2015).

1.7.2 Inter-swarm conspecific recognition

Spatial Swarm Segregation

Swarm site segregation is thought to contribute to assortative mating though the occurrence of

mixed swarms in some regions (Diabaté *et al*., 2008) could indicate the use of signals e.g. flight tones (Pennetier *et al*., 2010). The role of swarm segregation in facilitating assortative mating was report by Diabate *et al*. (2009). The study, done in Mali, revealed a strict pattern of spatial segregation, which resulted in almost exclusive monotypic swarms according to species. They found out that clusters of swarms were composed of individuals of the same species. But when *An. coluzzii* and *An. gambiae* females were introduced into *An. coluzzii* swarms, both species were inseminated regardless. This suggests that there was no within-mate recognition. They argue that their results provide evidence that swarm spatial segregation is a strong contributor to reproductive isolation between molecular forms. Spatial swarm segregation, though crucial to reproductive isolation is not solely responsible. Other factors such as the production of pheromones by males have been shown to attract females to the swarms, though this requires more research (Diabate and Tripet, 2015).

Flight tone and harmonic convergence

Flight tones produced in the process of courtship have been suggested as a cue employed in conspecific recognition within swarms. Detection of flight tones as the basis of audiomotor interactions occurring between male and virgin females mosquitoes of same species to present quantifiable means of discriminating the sibling species was studied (Pennetier *et al*., 2010). In their report, the mosquitoes maintained a constant ration, between their fundamental wingbeat, when flying with the conspecific individuals only, which aided in sexual recognition within same species. These contributions though not concrete in their depiction of the speciation process are definitely contributory (Pennetier *et al*., 2010).

The use of acoustic signals is very important in mating behavior as the early identification of a mate reduces an unproductive chase. This could be the reason why males use different signals to help in decision-making during swarming (Puckett *et al*., 2015). *An.*

coluzzii and *An. gambiae* both display acoustic behavior response to frequencies range of female flight-tone (Simões *et al*., 2017).

Flight-tones have been shown as a way of mate recognition in sympatric population. Flight tone matching confers the capability of mate recognition and it's a precursor for assortative mating. In a study between the *An. coluzzii* and *An. gambiae*, frequency matching of flight tones was more consistent within pairs of the same species than pairs of different species (Pennetier *et al*., 2010). In another study, a photosensor was used to measure the transient wavelengths generated by individuals of *An. arabiensis, An. coluzzii* and *An. gambiae*. They found no significant difference between and within species and also between males and females in a population in Mali (Tripet *et al*., 2004). Flight initiation involves a high increase in wingbeat followed by a rapid frequency modulation (RFM) when the mosquito is close to the sound source. It was noticed that RFM seem to occur without any signal or the presence of a female. It seems to be a behavior that occurs before mating (Simões *et al*., 2017). It is also said wingbeat frequency (WBF) and RFM might not be a factor for assortative mating as no difference in WBF and RFM was found in research conducted by (Simões *et al*., 2017). They reported that acoustic signals only inform the male mosquito of the presence of a flying female and female flight tunes are not for identification of mates of the same kind (Simões *et al*., 2017). Though a study in Mali and Guinea-Bissau showed a relationship between wing size and assortative mating (Sanford *et al*., 2011). They collected and measured the wings of *An. coluzzii* and *An. gambiae.* The wing length and wing width differed with respect to the strength of assortative mating. Those in Mali where assortative mating was strong due to the low rates of hybridization had larger mean wing length and wing width. Those of Guinea- Bissau where assortative mating seemed to be reduced, though wing length and width did not differ significantly (Sanford *et al*., 2011).

Flight tone stimulus has been reported to exert harmonic divergence in both male and female mosquitoes. Individuals were able to identify frequencies from small mates and from large ones. Flight tone frequency is probably used for mate assessment in mosquitoes (Cator *et al*., 2010). Harmonic convergence does not occur by chance, it occurs in tethered females and in single individuals as well (Aldersley *et al*., 2016). Both tethered and single males were discovered to respond to pre-recorded flight tones from same sex or opposite sex. A wide range of harmonic combinations elicited male-female response and it involved both sexes' participation. Male-male interaction showed frequency avoidance (Aldersley *et al*., 2016). Harmonic convergence before mating led to increased mating success and male off springs from such mating were likely able to converge before mating (Cator and Harrington, 2011). Cator and Zanti (2016) reported a relationship between body size and harmonic convergence by carrying out a study on *Ae. aegypti*. They reported that convergence is a prerequisite for a successful mating but body size influences the outcome of mating (Cator and Zanti, 2016).

The cues used by both males and female *Anopheles gambiae* complex for mating is yet to be elucidated. The success of malaria control projects, which includes the spread of transgenes into mosquito populations, depends on how isolated populations are to one another (Tripet *et al*., 2005). Studies of the mechanisms of reproductive isolation between forms of *An. gambiae* are very important in the development of reproductive competitive laboratory strains (Lanzaro *et al*., 2004).

1.8 Aim and objectives

Speciation is a process by which new species evolve from existing ones and the process has a major implication on vector control of malaria. Cryptic populations that evolve as a result of speciation can be similar in many aspects yet differ remarkably in ways such as behaviour and ecological preferences. It is therefore important to understand the mechanisms of speciation and the population structure of the vectors. Assortative mating and ecological divergence are considered mechanisms of divergence between *Anopheles coluzzii* and *Anopheles gambiae* the recently diverged members of the *Anopheles gambiae* complex. These processes can impact on the success of vector control measures focusing on the spread of transgenes and also the development of novel vector control methods.

This research examined the ecological and genetic factors influencing speciation between the *Anopheles coluzzii* and *Anopheles gambiae* sibling species. The genetic studies focused on the X-island of speciation. The X-island of speciation has been predicted to contain genes of assortative mating and ecological adaptation in *Anopheles coluzzii* and *Anopheles gambiae* sibling species. Morphological differences facilitated by ecological divergence between the two species were also studied.

Specific Aims

1. Investigate possible egg morphological differences facilitated by ecological adaptation in *An. coluzzii* and *An. gambiae* eggs using egg morphometry.
2. Investigate differentially expressed genes located on the X-island of speciation between *Anopheles coluzzii* and *Anopheles gambiae* at the different developmental stages.
3. Investigate the functions of candidate assortative mating genes through gene knockdown experiments and behavioral assays.

Chapter 2 General methods

2.1 Mosquito species

Four different wild type strains of *Anopheles gambiae* were used for the experiments reported in this thesis. These are Mopti (MRA-763), a strain originally from the N'Gabacoro Droit, area of Mali western Africa where Professors' Lanzaro and Tripet collected viable eggs in 2003. The eggs were gotten from adult females collected from the field. The strains have the +bc and +u inversions on the 2R chromosome and are of the Mopti chromosomal forms. The second is Pimperena (MRA-861) from Mali as well and was originally collected by M.Coulibaly in 2005. Akron (MRA-913), one of the strains used in this research was isolated in Akron, Benin. The fourth strain used is Kisumu (MRA-762), with origin from Kenya and established by Davidson in 1975. This strain has the b inversions on chromosome 2R and is of the Savanna chromosomal forms. Mopti and Akron are strains of *Anopheles coluzzii* species while Kisumu and Pimperena are both *Anopheles gambiae*.

The mosquitoes were maintained in the Patrick Manson insectary situated in Huxley building, Keele University. Breeding of the mosquitoes was done under optimal controlled conditions.

2.2 Insectary conditions

Mosquitoes can be kept in a ventilated room having the right humidity. An environment with ideal conditions is very important, as this would have an effect on the size and general development of the mosquito (Benedict, 1997; Lyimo *et al*., 1992). The mosquitoes were kept in an insectary, each insectary unit had an inbuilt thermostat, with temperature set at 26 ± 1^{o}C, and a 505-model defensor hydrostat (Klima international, Germany) which regulates the humidity to about 70% to 80%. The light in the insectary was set at a 12:12 light:dark cycle.

Mosquitoes go through four distinct stages in their life cycle (complete metamorphosis) therefore, rearing is based on its needs at each stage. The stages are egg, pupae, Larvae and Adult. The outcome of a rearing method can be observed not only from the size of the larvae but how consistent the larvae are in size. Growth conditions have an impact on the physiological status of adults (Aboagye-Antwi *et al*., 2010). Adults reared in good conditions will most likely live longer than poorly reared ones (Benedict, 1997).

2.2.1 Rearing

Adult rearing

Adult mosquitoes in the insectary emerge in prepared cages, which are white and cylindrical in shape. The cages are plastic buckets of 20.5cm in height and 20cm in width. Polyester nets were used to cover the opening at the top and the side of the cage (Figure 2.2.1). Pots containing about 800-1000 pupae were put in a cage. As a food source, a bottle containing 30ml of 10% glucose solution was placed in each cage. A sheet of brown paper rolled into a tube was put into the bottle containing the glucose solution, this was to enable the mosquitoes perch and feed on the glucose-saturated paper. The glucose solution was changed twice every week. Cotton wool soaked in deionized water was placed on the cage and covered with a square white antistatic weigh boat (Fisher brand, UK) to keep the mosquitoes hydrated. Each cage contained 800-1000 mosquitoes.

It is important that cages are not overcrowded as this could have an impact on the quality of breeding. Competition and males bumping into a pair in copula can prevent the complete transfer of sperm into females during mating (Ponlawat and Harrington, 2009).

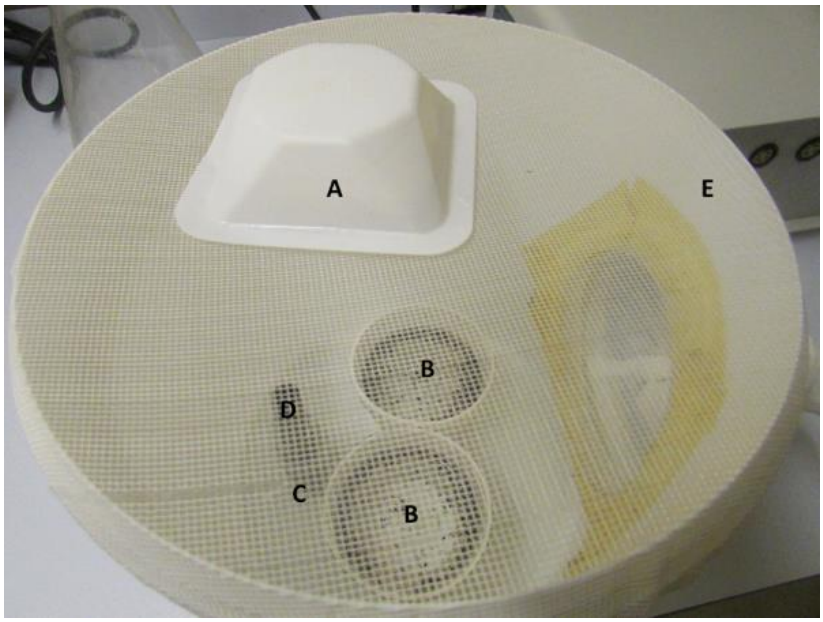

Figure 2.2.1: A prepared emergence cage A= Cotton wool soaked in deionized water and covered with a weighing boat. B= Pots containing pupae. C=Bottle containing 10% glucose solution D=Paper soaked in glucose solution, E= plastic cage**.**

At 3-7 days old, adult mosquitoes were bloodfed twice on defibrinated horse blood (TCS Biosciences, UK). Blood feeding was done using artificial feeders (Hemotek® membrane feeding system, Discovery workshops, UK) (Figure 2.2.2). The feeders were lined with parafilm and 10ml of blood put into the feeders. The feeders were set at 37±1°C temperature to heat up the blood. A brown paper was placed on the bottom of the cage before blood feeding. The amount of droppings on the paper gave an idea as to whether or not the mosquitoes fed properly. This was in addition to observing engorged abdomens in well-fed females.

Figure 2.2.2: Blood feeding of mosquitoes. A= Cages containing adult mosquitoes, B= Hemotex feeders, C= Heating source.

Gravid females were provided with oviposition pots 48 to 72 hours after blood feeding. These were polyester pots lined with whatman filter paper (cat no. 1001125) and deionized water added to the pot. The prepared pots were placed in the cage to mimic oviposition sites for gravid females. After 2-3 days, eggs laid in the insectary begin to hatch into first instar larvae.

Larval rearing

Two hundred first instar larvae were put into white rectangular trays (33 x 24 x 5cm). This was done using a 3ml pipette after which 500ml of deionized water was put into each tray and 2 drops of liquifry (Interpet ltd., Dorking UK) added to the trays. The trays were covered with clear acrylic sheet (Figure 2.2.3), which have been cut to size and perforated to allow for passage of air. After 24 hours, 10mg of grounded flakes tetramin flakes (Tetra werk, Melle, Germany) was added into each tray; the quantity was increased slightly after 48 hours. On the

fourth day 30mg of tetraMinBaby (Tetra werk, Melle, Germany) was added. The amount of food fed to the larvae was slightly increased, as the larvae would be seen to have increased in size.

Figure 2.2.3: Insectary setting of trays for experiments. Colours are used to differentiate between strains. Trays are arranged interchangeably to ensure same rearing conditions for all larvae in order to avoid variations.

On the fifth day, 500ml of deionized water was added to the trays making a total volume of 1liter per tray and 30mg of TetraMinBaby was added. After 6 hours four sinking carnivore pellets (Tetra werk, Melle, Germany) and six tetra pond sticks (Tetra werk, Melle, Germany) were distributed evenly in the trays. These pellets dissolve slowly leading to constant availability of food, as the larvae are voracious eaters at this stage. It takes about 7-9 days before pupation begins.

Pupae

The pupae were picked out of the tray using a suction pump (model no. FB 70155, Fisher brand, UK) (Figure 2.2.4) and transferred into white polystyrene pots containing deionized water. Pupae were split into several polystyrene cups of about 200 pupae per cup, this was to prevent overcrowding, which could lead to mortality.

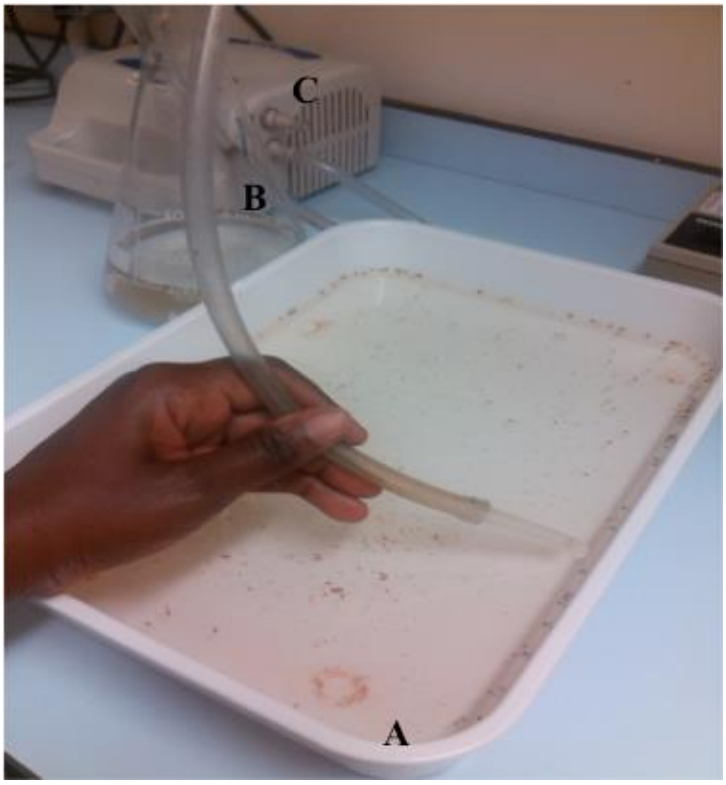

Figure 2.2.4: Picking of pupae in the insectary. A= Tray containing pupae and larvae, B=Conical flask with a tube attached to the pump C= Aspirator vacuum pump.

Picking of pupae was continuous, as all larvae do not pupate the same time. Pots containing pupae were put into dry and clean emergence cages. After 1-2 days adult mosquitoes emerged and the rearing of a new generation began. Two or more strains of Anopheles were sometimes reared in same insectary therefore, the integrity of the mosquitoes is of great importance as contamination can occur in such situations. To prevent contamination, different pipettes were

used for each strain during traying out of larvae. Suction tubes were flushed thoroughly with hot water after picking of pupae. Oviposition cups, emergence cups, cages, trays and all equipment used in the insectary were labelled appropriately and colours were used to differentiate between species (Figure 2.2.2 and 2.2.3). The possibility of unintentional contamination cannot be completely ruled out; therefore, the authenticity of the strains was checked before embarking on any experiment.

2.3 Genomic DNA extraction

This technique was employed for checking the integrity of strains. Genomic DNA was extracted from a sub population followed by a polymerase chain reaction for species identification. This technique was also used during some mating experiments where genomic DNA was extracted from sperm bundle removed from dissected mated female mosquitoes. Genomic DNA was also used during double stranded RNA synthesis. The amount of tissue used in the different experiments was dependant on the purpose of the extraction.

DNAzol method

This extraction method was carried out using the DNAzol (Invitrogen, USA) extraction protocol. Subpopulations of the mosquitoes were picked using a mouth aspirator and put into tubes containing 75% ethanol. Individual mosquitoes were. Individual mosquitoes were placed in properly labelled 1.5ml eppendorf tubes, 100µl of DNAzol was put into the tubes and homogenised using a pestle until body parts were not recognized. The homogenate was then spun for 10 minutes at 10000g. The supernatant was transferred into a clean 1.5ml eppendorf tube and residue discarded after which 50µl of 100% ethanol was added to each tube and mixed by inverting the tube 5 to 8 times. It was then left to sit at room temperature for 3 minutes after which it was spun for 10 minutes at 7000g and supernatant discarded. The

pellet was re-suspended by adding 750µl of 75% ethanol to each tube and inverted about 5 to 8 times and spun for 10 minutes at 7000g and supernatant discarded. This step was repeated for an extra clean DNA. The sample was again short spun, any leftover ethanol removed using a micropipette and 200µl of Tris-EDTA (TE) buffer was added to the sample. To dissolve DNA pellets, samples were vortexed and placed on the heat block at 80°C for 60-120 minutes before transferring to permanent storage tubes and stored at -20°C until needed.

2.4 Polymerase chain reaction

Three different methods of DNA amplification were used frequently in the course of this research. They are, the Restriction fragment length polymorphism (RFLP) for species identification, Short Interspaced Elements *S200X6.1* PCR for species identification as well and the Y-chromosome specific PCR for sex differentiation.

Restriction Fragment Length Polymorphism (RFLP) PCR

The morphologically indistinguishable sibling species of *Anopheles gambiae* and *Anopheles coluzzii* can be identified using the ribosomal DNA-polymerase chain reaction (PCR) method. The difference in the intergenic spacer regions of their ribosomal DNA is used to distinguish them (Scott *et al*., 2013). To identify different sibling species among the *Anopheles gambiae* complex, and also simultaneously distinguish between *An. coluzzii* and *An. gambiae*, the PCR-Restriction fragment length polymorphism (RFLP) is efficient and fast (Fanello *et al*., 2002) (Scott *et al*., 1993). Two controls were always used, the *An. coluzzii* and *An. gambiae* control. Primers that have been designed for the DNA sequences were used. They are abbreviated as GA for *Anopheles gambiae* with the sequence (5'-CTG GTT TGG TGG GCA CGT TT -3'), Primer UN (5'-GTG TGC CCC TTC CTC GAT GT-3') as universal and it anneals to the same position on the rDNA for all Anopheles species. Primer for *Anopheles*

arabiensis is abbreviated as AR, however, this was not used, as the differentiation was between the two molecular forms of *Anopheles gambiae s s* now sibling species. The master mix for the PCR consisted of the primers mentioned above. For a 25µl reaction, 18.975µl of H_2O, 0.5µl of Deoxyribonucleotide triphosphate (DNTPs) (10mM), 2.5µl x10 dream taq buffer (by fermentas) and 0.125µl of dream Taq polymerase (by fermentas) was mixed in a tube, distributed into PCR plate and 2µl of the DNA template was pipetted into the PCR plates. The master mix was put into Peltier thermal cycle 200 DNA engine thermo cycler (Bio rad) and optimised thermal conditions were used for amplification (Table 2.1). The PCR product was then digested using 0.25µl of Hha1 enzyme, 2ul of 10x reaction 2 buffer and de-ionized water was added to make a total volume of 10µl. The PCR product was added to the enzyme and put in an incubator (Inca Mikura ltd UK) set at 37°C and left over night. The digestion was to differentiate between DNA fragments for the PCR-RFLP for the two sibling species. Some setbacks were observed with this method, as bands on gel were sometimes difficult to interpret because of improper digestion. The bands appeared like hybrids and digestion was often repeated. The sine200 PCR was then used instead as it does not require digestion.

SINE-PCR

SINEs (Short Interspaced Elements) are homoplasy-free genetic markers that can be used for population genetics studies. The PCR is based on the analysis of insertion polymorphism of nearly 200 bp-long SINE (SINE200) within the islands of speciation of the *An. coluzzii* and *An. gambiae*. The insertion of *S200* X6.1 was found to be fixed in all *An. coluzzii* (M-specimens) and absent in all *An. gambiae* (S-specimens) and this led to the development the PCR for the identification of *A. gambiae* molecular forms. Since it is based on a single copy and irreversible SINE200 insertion, it is not affected by evolutionary patterns that affect

rDNA markers (Santolamazza *et al*., 2008). The PCR reaction was carried out in a 25µl reaction. The primer used was 0.15µl at 100pm/µl S200X6-1F (5'-TCGCCTTAGACCTTGCGTTA-3'); R (5'-CGCTTCAAGAATTCGAGATAC-3') was used. 2.5µl of 10x buffer, 0.5µl DNTP, 0.125µl dreamtaq, 20.375µl nuclease free water and 1ul of DNA template in the case of whole mosquitoes body and 2-3µl when extraction when done on mosquitoes sperm bundle. The SINE-PCR was done for species identification (Figure 2.4.1).

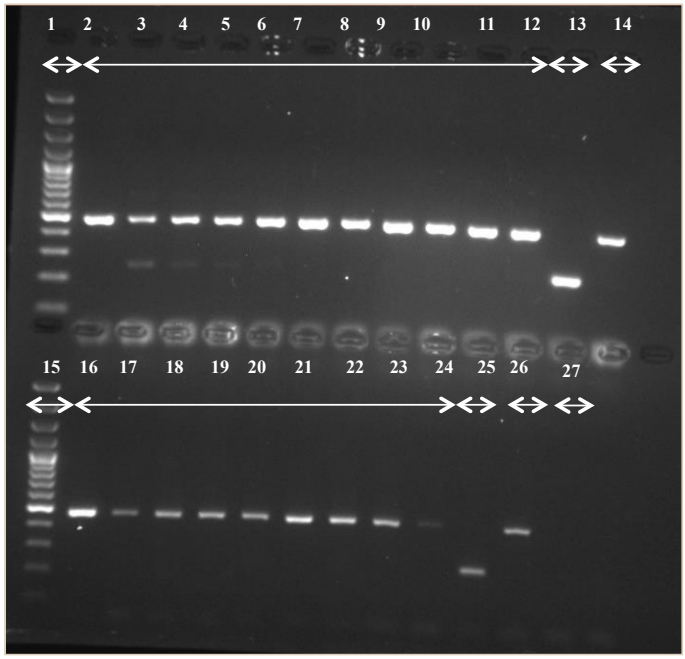

Figure 2.4.1: SINE-PCR. Lanes 1 and 15=DNA marker, Lanes 2-12 and 16-24= *Anopheles coluzzii* (PCR product of 479bp amplified). Lanes 14 and 26=*An. coluzzii* (Positive control). Lane 13 and 25= *An. gambiae* (Positive control, PCR product of 249bp amplified).

Y-chromosome specific PCR

The third PCR used in the course of this research was the Y chromosome-specific polymerase chain reaction (PCR) in which makers are amplified and used to differentiate between mated and unmated females (Ng'habi *et al*., 2007). The PCR was done to correct errors that might have occurred during dissection of the spermatheca. Samples were sometimes checked to

ascertain the status as male sperm and not tissue mistakenly taken from the female mosquito's body. Anopheles has 3 sets of chromosomes, two autosomes and a sex chromosome. The Y-chromosome is found only in males but is also found in females when males have transferred sperm during mating. The technique is based on the amplification of a multi-copy locus (ribosomal DNA) in mated females. The amplification is possible because the Y-chromosome has been characterised and Y-chromosome specific polymerase chain reaction markers have been developed (Ng'habi *et al*., 2007). Prior to the PCR method, checking for mating status in females was a tedious task which involved a lot of time because the spermatheca in females were dissected and checked for sperm bundle or mating plugs (Ng'habi *et al*., 2007) and the samples had to be fresh. On the other hand the Y-chromosome PCR can be done using both dried and fresh samples, mated and virgin females. The primer sequence used to amplify the DNA was S23F (5'CAAAACGACAGCAGTTCC3') R:(5'TAAACCAAGTCCGTCGCT-3'). It consists of a 2.5µl 10x reaction buffer containing $MgCl_2$, 0.25ul of 100pm/µl primers, 0.5µl DNTP, and 0.125µl dreamtaq and 18.875µl volume of water. Amplification was carried out using optimised thermal cycling conditions (Table 2.4.1).

Table 2.4.1: Thermal cycling conditions for the restriction fragment length polymorphism (Scott fanello PCR) and Sine-200 PCR (for Species identification), Y-chromosome specific PCR (for Sex differentiation). **Source:** (Fanello *et al*., 2002; Ng'habi *et al*., 2007; Santolamazza *et al*., 2008).

	Scott Fanello PCR		**Sine-200 PCR**		**Y-chromosome specific PCR (S23)**	
	PCR Temp.	**Time**	**PCR Temp.**	**Time**	**PCR Temp.**	**Time**
Denaturing	95°C	3 minutes	94°C	10 minute	94°C	3 minutes
	95°C	15 seconds	94°C	30 Seconds	94°C	20 seconds
Annealing	52°C- 65°C	30 seconds	58°C	30 Seconds	55°C	30 seconds
	72°C	1 minute	72°C	1 minute	72°C	1 minute
Extension	72°C	1 minute	72°C	10 minutes	72°C	10 minute
Hold	4°C	1 hour	4°C	1 hour	4°C	1 hour

Gel electrophoresis was done using 1.8% agarose gel in 100ml of 1xTE buffer and it was stained with 1µl of ethidium bromide so as to identify the DNA fragments. The gel was put into mini gel system (MP-250) and allowed to solidify after which 5µl of each sample was loaded into the gel and ran at 60 volts using a gel electrophoresis machine. There were cases where 6x loading dye was added to the sample before loading on the gel. DNA molecular weight ladder (Ambion) was used to estimate the size of the PCR product. The results were checked using the Bio Imaging system (SynGene). The bands produced due to the digestion were interpreted using 100bp ladders.

2.5 RNA extraction using trizol

RNA extraction was done according RNeasy Mini Kit QIAGEN manufacturers manual with some modifications. This procedure was used for different experiments reported in this thesis. Before the start of extraction, work area was cleaned with RNase Zap. Clean gloves were used because contamination is likely to occur from contact with tube caps. Inner part of the

tube caps were avoided when discarding flow through and filter tips were used.

Total RNA was extracted from a pool of mosquitoes depending on the life stage (Table 2.5.1). The appropriate volume of trizol was put into a 1.5ml Eppendorf tubes (Tables 2.5.1) and the starting material, which was larvae, pupae or adult, was added to the tube. The volume of trizol added depended on the amount of tissue (1ml of trizol to 0.1g of tissue) used.

Table 2.5.1: Amount of tissue and appropriate volume of trizol used during RNA extraction

Sample	Number	Amount of trizol in tube before homogenizing (µl)	Amount of trizol added after homogenizing (µl)	Total volume of trizol	Notes
Larvae	20	100µl	400µl	500µl	
Pupae	10	400µl	100µl	500µl	200µl of trizol was put into two separate tubes. The number of pupae was spilt in two i.e. five (5) pupae in each tube and homogenized. Homogenate was combined in a single tube and 100µl of trizol added.
Adult	10	600µl	0µl	600µl	300µl of trizol into two tubes. Number of mosquitoes was spilt in two i.e. Five (5) mosquitoes in each tube. Homogenate was combined in a single tube.

Samples were homogenised in trizol and the homogenate left to sit at room temperature for 5 minutes. A 400µl of trizol was added to the homogenate to give a total volume of 500µl of trizol and 100µl of chloroform was then added (0.2µl of chloroform per 1µl trizol used). It was shaken vigorously for 20 seconds, vortexed slightly and allowed to sit at room temperature for 2-3 minutes. The tube was then spun at 10000g for 18 minutes at 14^{o}C for larvae or pupae and 11000g for 19 minutes for adults. The aqueous solution (topmost clear phase) was carefully collected and transferred to a new sterile RNase-free tube (1.5 ml tube). Equal volume of 100% RNA-free EtOH was added to the tube and mixed thoroughly by pipetting up and down. The sample was loaded into RNeasy column Qiagen kit seated in a collection tube and spun for 30 seconds at 8000g. The flow through was discarded and 700µl buffer RW1 was put onto column spun for 30 seconds at 8000g and flow-through discarded again. The column was then transferred into a new collection tube and 500µl of buffer RPE was added and spun for 30 seconds at 8000g, and flow-through discarded. The step was repeated and the column spun for 2 minutes and the flow through discarded. This is to wash the membrane. The column was then spun for 1 min at 8000g to get rid of leftover buffer. The column was transferred to a new 1.5µl collection tube after which 50µl of TE was put directly onto the column membrane and allowed to sit at room temperature for 2 minutes. The tube was spun for 1 minute at 8000g to elute RNA. The concentration and quality of RNA was recorded using the Nanodrop ND-1000 spectrophotometer (Nanodrop technologies) at wavelengths of 230nm, 260nm and 280nm. The samples were then stored in at -80^{o}C.

2.6 Sexing of pupae

Mating occurs quite early in a mosquitoes adult life, hence separating males and females before they emerge as adult is very important where virgins are needed. Sexing is done at the pupal stage. A clean petri dish was put under the microscope (Leica Microsystems GmbH,

Germany). Low light intensity was used as very high intensity could kill the pupae. A pasture pipette was used to pick some Pupae and put in the dish adding a drop of water. Using a sexing iron rod, and looking through the microscope, male and females were separated (Figure 2.6.1).

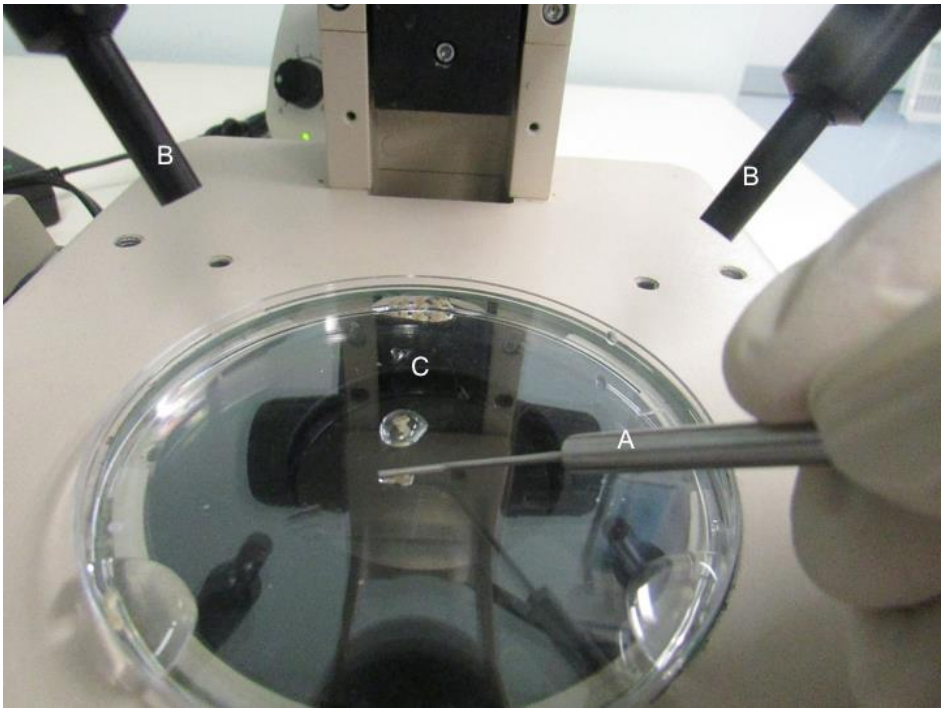

Figure 2.6.1: Sexing of pupae. A= Iron rod for separation of males and females, B= Microscope illuminator, C= Petri dish containing pupae.

Differentiation is done based on the last segment of the pupae's abdomen, which is the genital lobe. The males have a more pointed genital lobe while that of the female is not prominent. Male pupae were moved to one side of the dish and female pupae are moved towards the opposite direction. Pupae were then collected and put into separate pots containing deionized water. This was put into separate emergence cages. The cages were labelled according to species and sex.

Identification

Male and female pupae are differentiated using their posterior end called bristles, tracheal tube or pedals. The males posterior end is hook like in shape while it is straight in females (Fig 2.6.2).

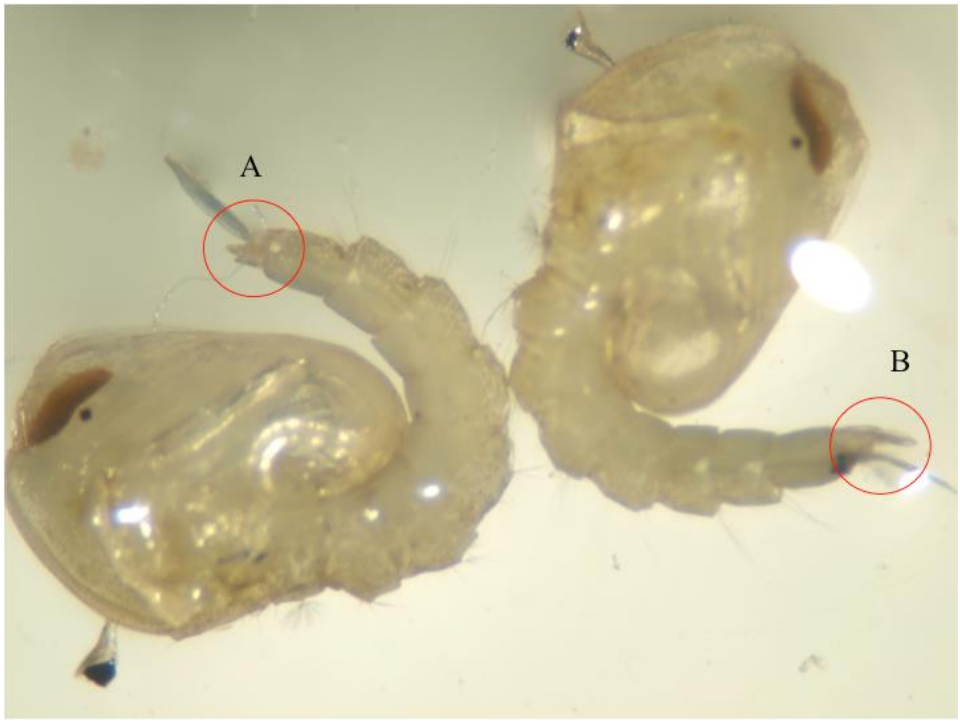

Figure 2.6.2: *Anopheles gambiae* pupae A = Female, B=Male.

2.7 Sperm bundle dissection

Mosquitoes stored in 70% ethanol were collected for dissection. Females were separated using forceps and each female picked and put on a clean slide and placed under the microscope (Leica Microsystems GmbH, Germany). A drop of deionized water is added so as to prevent the spermatheca from drying up. The posterior region containing the spermatheca was detached from the body. Using fine thin dissecting pins, the spermatheca was pulled out from the abdomen and tissues around it cleaned using the pins. The cleaned spermatheca is then cut open and the sperm bundle pulled out careful. The sperm bundle is usually seen as a

cream round bundle of tissue (Figure 2.7.1). If the sperm bundle is mistakenly disintegrated at the process of breaking open the spermatheca, the sperm bundle looks like a cluster of thread. The sperm bundle in most cases stays stuck to the pin; the pin was then dipped into the prepared lysis buffer and moved from side to side in the tube to ensure the sperm bundle dislodges from the pin and remained in the buffer. To ensure the sperm bundle has been transferred into the buffer, the pin was checked under the microscope.

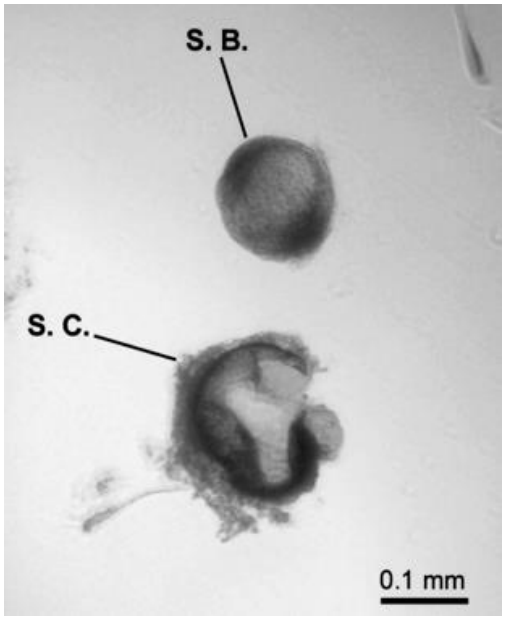

Figure 2.7.1: Sperm bundle dissected from a spermatheca S.C =*An.gambiae* spermatheca capsule, S.B= Sperm bundle. **Source:** (Tripet *et al*., 2001).

2.8 Wing measurement

The size and weight of a mosquito can be deduced from the size of its wings. Wings were measured to check for variation between the populations used during mating experiments. Mating experiments were done using two strains, Mopti and Kisumu. In most cases, females were given a choice of males of both species. Therefore, wing size measurements would show whether or not mates were of same size, which can alter chances of mating in both species. During all assays, a subset of the population was taken and stored in 70% ethanol. Wings were carefully detached from the body placed on a microscope slide and put under a

microscope. An Olympus camera model E-520 was used to take individual wing pictures (Figure 2.8.1). A picture of a stage graticule was also taken using the same objective as that used for the wings. The stage graticule was used for calibration. Image J was used to analyse the length of wings.

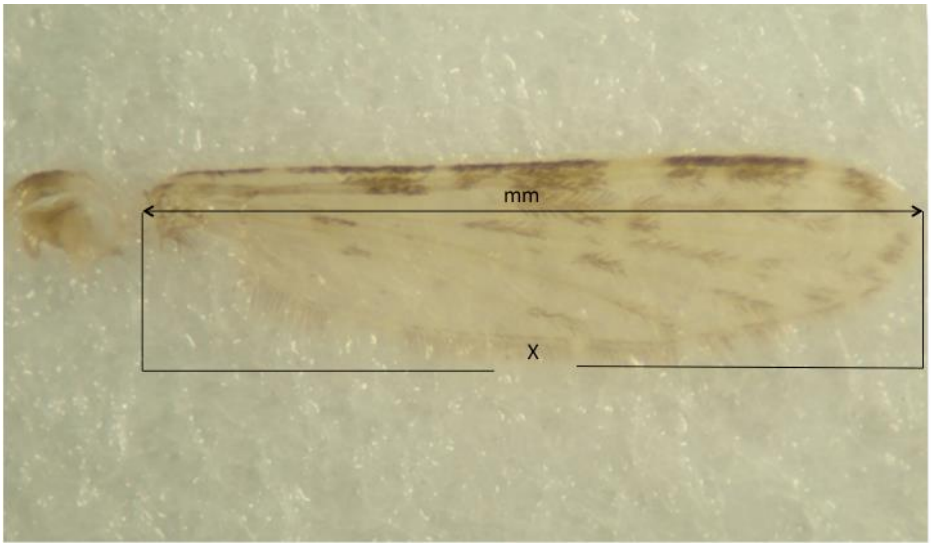

Figure 2.8.1: *Anopheles gambiae* wing. X shows the wing length calculated in millimetres.

Chapter 3 Egg morphological adaptation

3.1 Introduction

Oviposition preference is an important part of a female mosquito's life. Members of the genus *Culicinae* deposit egg rafts on water surfaces and eggs hatch almost immediately (Michael and Day, 1989). *Aedes* lay individual eggs on substrate or above water line, eggs are sometimes laid on soil or in containers and the eggs do not possess floats. The eggs are resistant to desiccation therefore survive months or even years of dryness until they are stimulated to hatch through flooding (Gonzalez *et al*., 2016). Anophelines on the other hand lay hundreds of single eggs on water surfaces; the eggs do not survive desiccation and consist of floats (Pratt *et al.,* 1963).

Differentiation of members of the genus *Anopheles* is difficult due the presence of large numbers of species that cannot be differentiated using morphological characteristics (White, 1977). Coluzzi (1964) in a WHO report mentioned some characteristics used in the identification and naming of the early members of the *Anopheles gambiae* complex. These consist of genetic incompatibilities between the sister taxa, combined with differences in habitat preferences. For example *An. melas* and *An. merus* larvae were found in typical coastal environments and resistant to salinity while those of *An. gambiae* were found mostly in freshwater habitats and do not survive high salinity (Coluzzi, 1964). Most studies focus on location and insecticide resistance, and not much attention has been given to studies on the phenotypic differences between the molecular forms (Lehmann *et al*., 2008). Several attempts have been made in order to distinguish the members of the complex morphologically (Table 3.1.1), but so far this has remained unsuccessful (Lanzaro *et al*., 2013). Important discoveries have however been made regarding marked differences in the egg structures among members of the complex (Coluzzi, 1964). *An. melas* have been reported to have eggs different from the typical *An. gambiae* forms and egg characteristics have been used to differentiate between

species (Thomson, 1948).

Table 3.1.1: Phenotypic comparisons between *Anopheles coluzzii* and *Anopheles gambiae* formerly called the M and S molecular forms of *Anopheles gambiae* s.s.

Trait	M	S	CO[a]	Reference[b]
Adaptation to larval & adult habitat				
Geographical range	West + Central	Continental	H	(della Torre et al., 2005)
Arid-mesic gradient (spatial + season)	Dry	Wet	H	(Coluzzi et al., 1979, 1985)
Adaptation to larval habitat				
Egg hatch timing & responses	Fast	Fast[c]	L	(Yaro et al., 2006a)
Egg desiccation tolerance	Low	Low[c]	H	Dao et al. unpublished
Larval habitat (preferred)	Permanent (rice)	Temporary (puddle)[c]	H	(Diabate et al., 2002, 2005; Robert et al., 1988; Toure et al., 1998)
Larva predator avoidance	Higher	Lower[c]	M	(Diabate et al., 2008)
Larva competitiveness (no predators)	Lower	Higher[c]	M	(Diabate et al., 2005)
Larval developmental time	Slower	Faster[c]	M	(Diabate et al., 2008)
Adaptation to adult environment				
Longevity (adult)	Longer	Shorter[c]	L	Dao et al.: unpublished
Body size (adult)	Larger	Smaller[c]	L	(Yaro et al., 2006b)
Reproductive output	Larger	Smaller[c]	L	(Yaro et al., 2006b)
Anthropophily	High	High[c]	L	(Wondji et al., 2005a)
Endophily	High	High	L	(Wondji et al., 2005a)
Plasmodium susceptibility	High	High	M	(Wondji et al., 2005b; Yaro et al., 2006b)
Insecticide resistance (kdr)	Low	High	H	(Chandre et al., 1999; Tripet et al., 2007; Yawson et al., 2004)
Mating Behavior				
Flight tone	492 Herz	493 Herz[c]	M	(Tripet et al., 2004)
Swarm landmark type	High contrast	Bare ground[c]	H	(Diabate et al. unpublished)
Indoor mating	Low frequency	None[c]	M	(Dao et al., 2008)

Traits in which only minimal differences were found between forms are coloured grey.

[a] Confidence level: high (H), moderate (M) and low (L) refers to the uncertainty in the generality of the result due to possible confounding factor, e.g., variation due to within form inversion karyotype, locality and time. Repeated patterns in independent studies were considered as highly confident, whereas low confidence score was assigned to a single study on a single population.

[b] Reference of studies comparing between molecular forms.

[c] Comparisons include only West African S populations and may not apply for East African populations. **Source:** (Lehmann *et al.*, 2008).

In the study, the upper surface of *An. gambiae* egg was reported to be narrow but broad in *An. melas*, space between frill and explain floats (Figure 3.2.2) were wide in *An. gambiae* but narrow in *An. melas* (Thomson, 1948). All eggs laid in places populated by *An. melas* were similar and no intermediate group was found (Ribbands, 1944), thus, egg morphology became a major way of identifying *An. melas* as it became possible to separate *An. gambiae* from *An. melas*. The discovery paved way for further research into egg morphology as a means of

species identification (Linley *et al*., 1993). Up to now, egg morphology is being used to differentiate between various species of mosquitoes and morphological characteristics of eggs have also been used to differentiate between several sibling species of mosquitoes (Tyagi *et al*., 2017).

Anopheles eggs are boat shaped consisting of floats situated on both sides of the egg. The floats are filled with air hence also called air chambers. They are balancing agents to keep the egg afloat (Hinton, 1968). The surface of the egg has an outer membrane called the exochorion. This is the outermost layer of the mosquito eggshell that protects the developing embryo from external stress and also helps maintain water balance (Figure 3.1.1) (Farnesi *et al*., 2015).

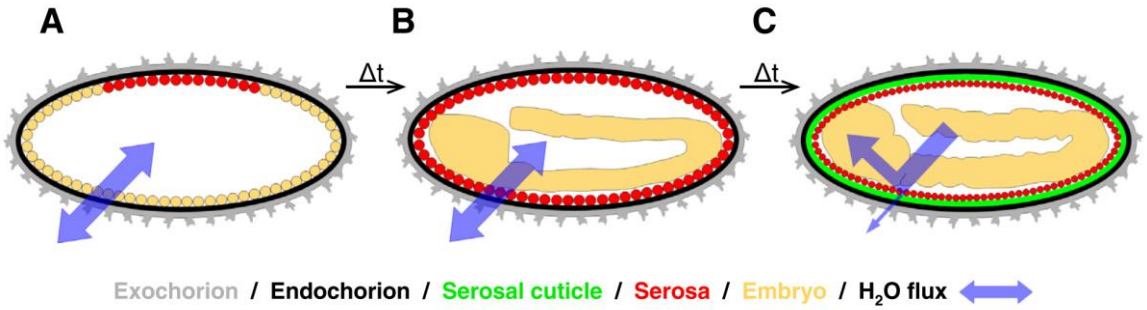

Figure 3.1.1: Mosquito eggshell layers. (A) Immediately after oviposition egg shells comprise of maternally produced exochorion and endochorion (B) During embryogenesis, serosal cells surround the embryo and subsequently (C) secrete the serosal cuticle that considerably decreases water flow. **Source:** (Farnesi *et al*., 2015).

The exochorion on the dorsal surface of the egg is finely marked in different patterns (Farnesi *et al*., 2015). Apart from the exochorion, the other two layers of the chorion are the endochorion and the serosal cuticle (Hinton, 1968). The egg also consists of a deck bordered by the frill (Figure 3.2.2). In most species, the frill is just as extended as the float. The deck could be said to be the upper surface of the egg and are bounded in the middle by upper parts

of the floats (Hinton, 1968). There are lobed tubercles located on either end of the deck. The function has not been fully described but when the egg is in water, the lobed tubercles are seen to hold a film of air (Hinton, 1968). The egg has a micropylar disc located at the anterior end. The micropylar disc harbours the micropyle, which is a tiny opening on the chorion. Sperms penetrate the micropyle during oviposition and the micropyle appears to be traversable by sperm during the process (Clements and Kerkut, 2013).

Some features used to differentiate between different species of *Anopheles* are floats (present or absent; length and width, number of ribs in the floats), micropylar collar, lobed tubercles and frills (Hinton, 1968).

The first species complex of Anopheles, the *An. maculipennis* complex (White, 1978) was identified using egg morphological characteristics (Linley *et al*., 1993; Sedaghat *et al*., 2003). Bates (1940) relied mostly on morphological studies to name the *maculipennis* complex in Europe (Bates, 1940) and the results were further confirmed by hybridization and cytogenetical analysis performed on the cryptic taxa (White, 1977). In the past, egg differentiation depended basically on the use of light microscope and drawing of micrographs, but in recent times, the scanning electron microscope has sometimes been used to describe relationships between eggs (Linley *et al*., 1996, 1993). Despite these discoveries, a lot remains to be done in studying the morphological characteristics of *Anopheline* eggs particularly using recent advances in the analysis of morphometric studies.

Egg shapes, like any biological shape, can be studied using morphometrics. Morphometrics include the analysis of shape variation and covariation of shapes (Webster and Sheets, 2010). Morphometry was used to measure shapes even in the 18th century but has undergone some modifications over time (Mitteroecker and Gunz, 2009). There are 2 broad types of morphometric studies: *Traditional morphometrics* and *Landmark-based geometric morphometrics*. Traditional morphometrics is widely used in egg morphological studies and

involves the measurements of length, angles and rations. It is the most convenient way to measure distances that are linear (Tatsuta *et al*., 2018) and has been reported in several egg morphological studies. Traditional morphometrics was used to distinguish eggs of species A, B, C and D of *An. dirus* complex (Damrongphol and Baimai, 1989). In the study, eggs of species A and C were found to be similar in shape while Species B was the largest and species D the smallest. Patterns of the exochorion between the frills and float differed between species. The deck tubercles arrangements, which are in aggregates, were more spaced in species A than in B and species C had larger aggregates. These differences in structures were important in distinguishing the different species of the *An. dirus* complex. In a study conducted in India, the eggs of *An. fluviatilis* a major vector of malaria responsible for about 15% of malaria cases in the country was compared with eggs of other Anopheline species. The shape and size of the tubercle in *An. fluviatilis* was found to differ from those of *An. culicifacies*, *An. nyssorhynchus*, *An. nuneztovary* and *An. apicimacula*. Some similarities were observed at the micropylar region between *An. fluviatilis* and *An. culicifacies* but not with *An. darlingi*, *An. rangeli* and *An. dunhami* (Sehrawat, 2014).

In another study, *An. culicifacies* A, B, C, D and E sibling species were differentiated using PCR assay (Tyagi *et al*., 2016), the morphometric studies of the sibling species showed differences in the egg attributes. Species D had narrower deck and smaller floats compared to A while species A had a bigger micropyle compared to D and E. The number of ribs was reduced in species A and E. The study suggests that the dissimilarity observed in the eggs could be used in differentiating the sibling species (Tyagi *et al*., 2016). Jukum *et al* (2004) studied the morphological characteristics of *An. aconitus* form B and form C eggs. Variations were observed in the float width and the number of posterior tubercles on egg deck was also different between the different forms. The samples used in the study were collected from the malaria endemic region of Thailand (Junkum *et al*., 2004).

Adult *An. costai* and *An. mediopunctatus* have similar male genitalia (Sallum and Flores, 2004). The species are similar at the larval, pupal and adult stages but differences were observed in the eggs (Sallum *et al*., 2004). These findings were carried out using eggs collected from Shannon trap in Icapara, Iguape municipality Brazil. The eggs of eight different species of *Anopheles hyrcanus* group (*An. argyropus, An. crawfordi, An. nigerrimus, An. nitidus, An. paraliae, An. peditaeniatus, An. pursati, and An. sinensis*) were described using scanning electron micrographs. Among the eight species studied, five showed differences in deck characteristics (*An. argyropus*, *An. nigerrimus*, *An. paraliae*, *An. peditaeniatus*, and *An. sinensis*). The ratio of egg length per maximal deck width differed between the species and differences were found in the numbers of float ribs (Saeung *et al*., 2014). *An. quadrimaculatus* complex comprises of species A, B, C (C1 and C2) and D. Forty-two egg attributes were used to provide interspecific differences between the complex. Thirteen characters were selected and used for multivariate analysis. The first 7 principal components (which shows major features of shape variation in a data set) accounted for about 91.18% variation and the discriminant function analysis showed 90.97% differences between species. The results revealed that species C1 and C2 are closely related followed by Species A and B. Species D was found to be closely related to species C1 and C2 (Linley *et al*., 1993). A recent study described the morphological and morphometric characters of *An. stephensi* 'type' form eggs extensively. The findings provided some important information on the species egg morphology that could lead to the identification of the different variants of the species (Tyagi *et al*., 2017).

The *Anopheles gambiae* complex has also been studied using traditional morphometric technique. The scanning electron microscope was used to check for differences in egg characteristics of six species of the complex (*An. arabiensis*, *An. gambiae*, *An. quadriannulatus*, *An. bwambae*, *An. merus*, and *An. melas*) (Lounibos *et al*., 1999). The

morphometric measurements of these eggs showed no obvious morphological differences. *An. melas* and *An. merus* were noticed to have wider decks and shorter floats. An overlap between species in the principal component analysis and the discriminant function analysis was observed between the sibling species (Lounibos *et al*., 1999).

The second type of morphometric studies is the Landmark-based geometric morphometrics. This involves the use of landmarks to summarize shapes. It is important in biological studies as it enables the physical presentation of information after removing irrelevant information such as the position and orientation of specimens (Tatsuta *et al*., 2018). The use of landmarks is employed in morphometric studies because it enables the extraction of shape information thereby visualizing different shapes changes (Klingenberg, 2011). It is based on the visualization of shape change by showing relative displacement of landmarks or by showing deformation of regular grids (Klingenberg, 2013). This morphometric approach is called geometric morphometrics because the geometry of the landmarks configurations is preserved throughout the analysis therefore results can be presented as actual shape or forms (Mitteroecker and Gunz, 2009). The landmark-based technique is used frequently in geometric morphometric analysis (Tatsuta *et al*., 2018). Landmark morphometrics was used to study wing shape in male and female *An. superpictus*. Twenty-two landmarks were used to study the wings in relation to ecological parameters and development rate. A decrease in wing size was observed at higher temperatures (Aytekin *et al*., 2009). Wing geometric morphometry has also been used for the differentiation of *Cx. coronator* from those of *Cx. usquatus* (Demari-Silva *et al*., 2017). However, landmark geometric morphometry has not been reported in egg morphological studies.

Egg morphological characteristics have been associated with factors such as geographical regions and drought. Phylogenetic analyses of geographically distinct *Aedes notoscriptus* populations have recently suggested the species as a complex of genetic lineages

(Endersby *et al*., 2013). Surface morphology of eggs was correlated with geographical regions where forty-four attributes of *Culex quinquefasciatus* eggs from Jodhpur, Bikaner, Jamnagar and Bathinda were differentiated based on the different regions (Suman *et al*., 2009). Eggs from Jodhpur, Bikaner and Bathinda were seen more similar than those of Jamnagar strain. A correlation of (r= 0.95) shows how geographical distribution affects the different egg attributes. This suggests that ecological factors impact on egg structures of mosquitos (Suman *et al*., 2009). In another study *Anopheles nuneztovari* eggs were collected from 3 different locations in Venezuela, 4 in Brazil and one location in Suriname. The Venezuelan and Suriname eggs were similar morphologically using features as anterior deck region and the pores in the dorsal plastron. The eggs from Brazil did not show any similarity with the eggs from Suriname. The differences noticed between eggs from Venezuela and Suriname was related with chromosomal, ecological and molecular evidence for regional genetic differentiation in these species (Linley *et al*., 1996).

Egg morphological characteristics have also been associated with drought resistance. Drought resistance in mosquito eggs can improve the survival of species and the spread of malaria. Mosquito eggs are liable to dehydration when surroundings dry out immediately after the eggs are laid (Farnesi *et al*., 2015). The serosal cuticle produced during embryogenesis in most mosquito species wraps the entire embryo and becomes the eggshell. The serosal cuticle is very important in egg resistance to desiccation (ERD) (Farnesi *et al*., 2015). The relationship between eggshell attributes and egg resistance to desiccation (ERD) has been investigated. Chitin content, eggshell surface density in addition to unidentified eggshell attributes were observed to be responsible for ERD in *Ae. aegypti*, *An. aquasalis* and *Cx. quinquefasciatus* (Farnesi *et al*., 2015). ERD depends on the mosquito species and seasons. *Aedes* can survive several months of drought while *Anopheles* and *Culex* can only survive a few hours when subjected to the same conditions. The difference in ERD among mosquitoes

could be a result of difference in traits such as length, width and weight area etc. (Farnesi *et al*., 2015).

An. punctulatus found in the Island of New Guinea are adapted to drought conditions. They are found in disturbed areas with limited availability of water. Eggs of *An. punctulatus* can survive drought conditions and larvae can survive several days in damp mud during desiccation. This contributes to the species efficiency as vectors of malaria in areas where they exist (Sinka *et al*., 2011). Mosquito egg size has been related to its ability to withstand desiccation. *Ae. aegypti*, which has larger eggs, were noticed to survive long periods of desiccation while *Ae. albopictus* eggs were less resistant to desiccation (Sota and Mogi, 1992). *Ae. aegypti* and *Ae. albopictus* eggs have also been compared using morphological and morphometric characteristics. The study was carried out using a scanning microscope. Morphometrically, 48.48% significant difference was found between eggs of the two species (Suman *et al*., 2011). *Ae. albopictus* eggs were significantly smaller and more reduced at the posterior end. The exochorionic networks in *Ae. aegypti* were interwoven, reticulated and wide while in *Ae. Albopictus,* the exochronic network is narrow and like solid-wall. This could account for the ability of *Ae. albopictus* to withstand desiccation when laid in containers leading to its ability to do well in artificial breeding sites and invariably in different environments (Suman *et al*., 2011).

In *Anopheles*, the Mopti chromosomal form, a strain of *An. coluzzii* sibling species has high frequencies of a and bc chromosomal inversions, the inversions are said to be responsible for drought resistance. These forms are found in large numbers in dry areas and are prevalent during dry seasons. Bamako and Savannah forms on the other hand lack such combination of inversions thereby found in wet areas and are abundant during the rainy season (Touré *et al*., 1998, 1994). Aboagye-Antwi and Tripet (2010) studied the impact of water and food availability on phenotypic quality of larvae and adult female *An. coluzzii*

Mopti chromosomal forms and how this affects their ability to cope with desiccation. The result revealed a carry over effect of water and food availability contributing to survival when females were faced with desiccation challenge. The carryover effect could be a survival strategy employed by this species (Aboagye-Antwi *et al*., 2010).

Apart from differences in their ability to withstand arid conditions, *An. coluzzii* and *An. gambiae* though morphologically identical have different oviposition and larval habitat preferences as well. This ecological flexibility does not only affect the mosquito's adult and larval stages but could have an impact on egg behavioural and ecological characteristics.

The primary aim of this study was to investigate possible morphological differences facilitated by ecological adaptation in *An. coluzzii* and *An. gambiae* eggs using egg morphometry. Egg shape variations between species and populations within species were studied. Two approaches were used in the study, traditional and landmark-based geometric morphometrics. *An. coluzzii* and *An. gambiae* are important vectors of malaria therefore egg morphological differentiation between this recently diverged species could offer an alternative method of species identification; it could lead to breeding sites identification. Egg attributes contributing to drought resistance in *An. coluzzii* could also be discovered, leading to a better understanding of malaria transmission. Egg morphological differentiation can also contribute to our understanding of the mechanism of speciation between *An. coluzzii* and *An. gambiae.*

3.2 Methods

3.2.1 Sample used

Egg source

Four populations of laboratory-bred strains were used. The populations are Akron and Mopti, populations of *Anopheles coluzzii* species and Pimperena and Kisumu strains of *Anopheles gambiae ss*. The mosquitoes were bred in the Manson's insectary at the centre for applied entomology and parasitology, Keele University. The mosquitoes were bred under controlled conditions as described in Chapter 2.

Sample collection and Preparation

Three-to-five days old gravid females were fed on horse blood and oviposition pots placed in the cages 4 days post blood feeding. The oviposition pots containing water were lined with filter paper to obtain batches of eggs. The oviposition pots were removed from the cages after 24 hours. In order to slow down development and prevent hatching, eggs were left in oviposition pots and stored in $4^{o}C$ until microscopy. Egg collections were done in three independent cohorts at different time periods.

3.2.2 Microscopy

A moist filter paper was placed in a petri dish and using a soft fine brush, individual eggs were removed from oviposition pots and placed on the middle of the filter paper. Picking of eggs from oviposition pots was done at random to ensure selected eggs were from a number of different female mosquitoes. A photograph of the dorsal, lateral and ventral view of each egg was taken. Ten eggs were used for each cohort and a total of 120 eggs were sampled. Therefore a total of 30 eggs were sampled for each population studied (Figure 3.2.1). The pictures of each egg were taken and landmarks allocated.

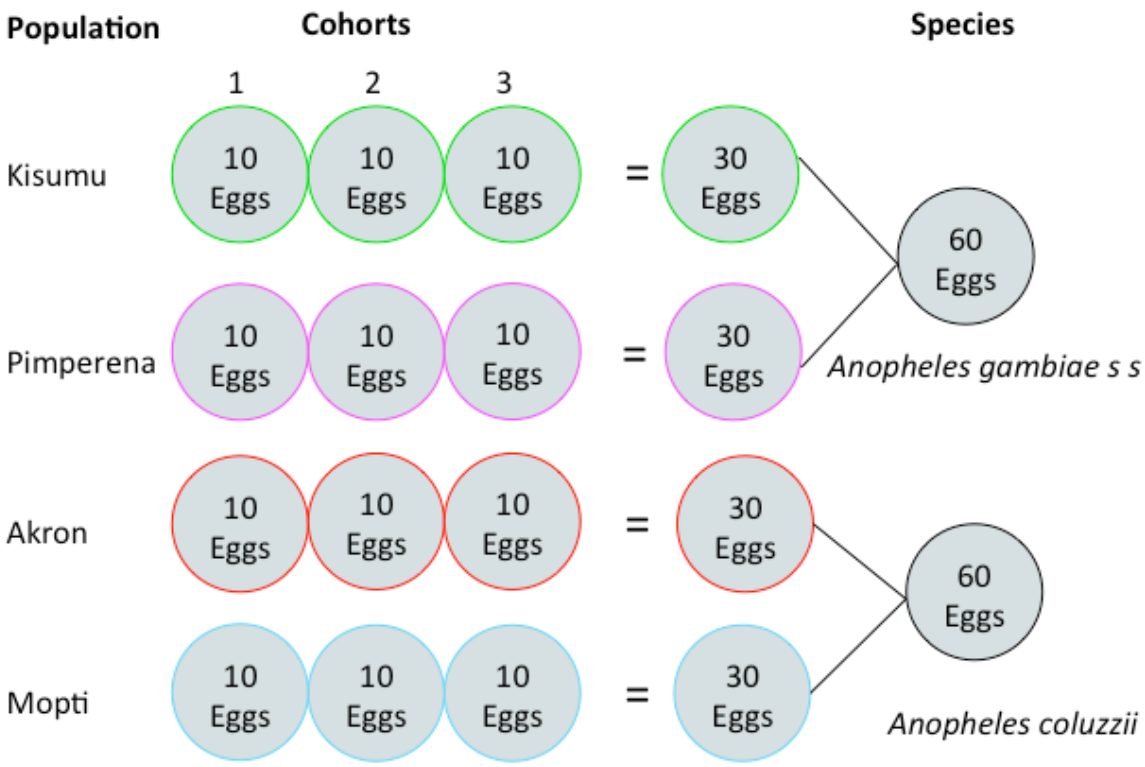

Figure 3.2.1: A Schematic representation of the experimental design and sample size used. A total number of 120 eggs were used. Two species and four different populations within species were used. Each population comprise 3 cohorts, 10 eggs per cohort, 30 eggs per population and 60 eggs for each species.

An Olympus digital camera model E-520 (10.0 megapixel) attached to a Leica microscope (Microsystems GmbH, Germany) was used to take photographs of individual eggs. The Same camera settings and microscope magnification of X63 was used for all photographs to prevent variation. A photograph of a 1mm length stage micrometre was also taken at the same magnification and this was used for calibration and conversion of units for measurements.

3.2.3 Egg Morphology and Morphometry

Egg morphology (*under the microscope*)

The egg appears boat-shaped (Figure 3.2.2A and B). The anterior and posterior end of the egg looks pointed. The anterior end appears to be broader than the posterior end. The float looks wide and spread out when egg is placed on the dorsal or the ventral side (Figure 3.3.2B) but appears short and clustered together when on the lateral side (Figure 3.2.2A). The frills are clearly visible on both ventrolateral and the lateral view (Figure 3.2.2 A and B) but not seen when egg is placed on the dorsal side. The deck runs from the anterior pole to the posterior end of the egg and looks narrow at the middle part of the egg.

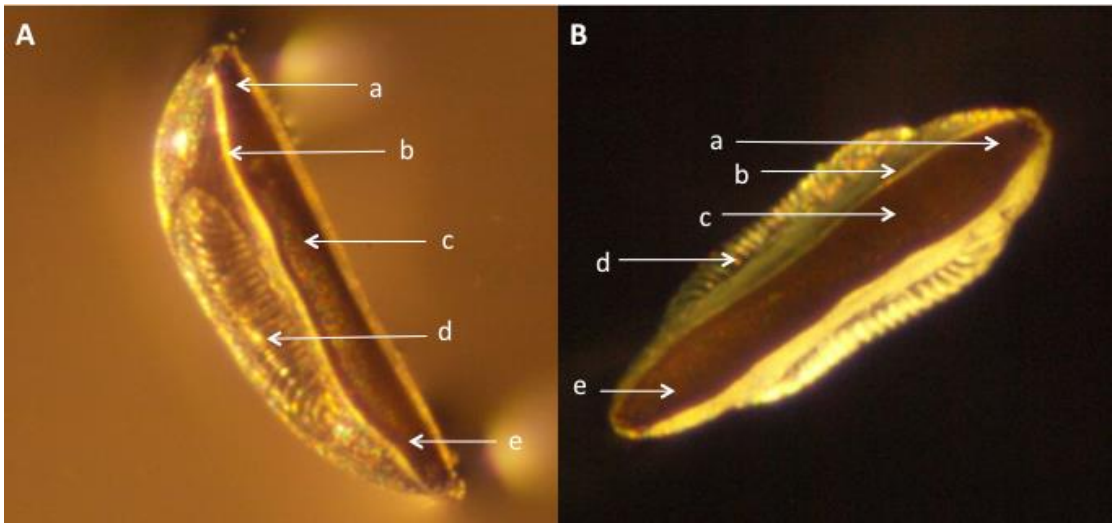

Figure 3.2.2: Eggs of *Anopheles gambiae.* **A** is the ventrolateral view showing the egg as boat shaped and **B** is the Ventral view. The anterior pole of the egg is at the top. **a**= Anterior deck region (micropylar region) **b=** Frills which encloses elongated rows of lobe tubercles. **c**= Deck area which runs from the anterior to the posterior of the egg. **d**= Float present on both sides of the egg surface**. e**= Posterior deck region (lobe tubercles situated).

3.2.4 Morphometrics analysis

Two methods of morphometrics were used to analyse the data for this study. The traditional morphometrics was used to compare lengths and width of eggs (Figure 3.2.3) and the landmark-based morphometrics was used to carry out multivariate analyses of egg shape.

Traditional morphometrics

The egg length (EL) was measured from the anterior to the posterior end, Egg width including floats (EFW), the width of mid-deck/dorsal (MDW) region, posterior deck/dorsal (PDW) region, and anterior deck/dorsal (ADW) region of the egg were also measured.

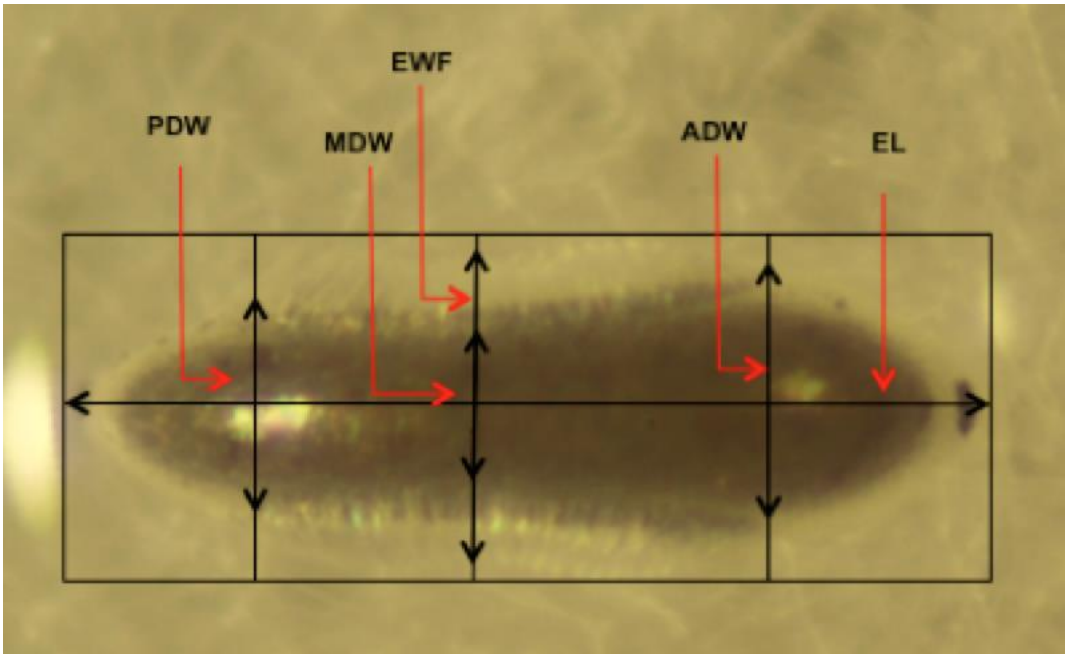

Figure 3.2.3: *Anopheles gambiae s.s* egg dorsal view showing attributes (lengths and widths) used for the traditional morphometrics analysis.

The measurements were taken from the dorsal view using ImageJ (version 1.50 c4) https://imagej.nih.gov/ij and the result generated was used to check for differences between species and populations. JMP® 13.0.0 www.jmp.com was used for this analysis.

Landmark-based geometric *morphometrics*

Photographs of eggs were imported to PowerPoint and landmarks created on the photographs using red dots (Figure 3.2.2 and 3.2.3). The photographs were again imported to ImageJ (version 1.50 c4) software. Using a point picker, landmarks were selected and the measurements for corresponding landmarks were produced using ImageJ. MorphoJ (version 1.06d) software was used for the landmark-based morphometrics analysis because the software offers a wide range of options of analysing shapes. It executes high quality multivariate techniques widely used in morphometrics and also has a number of specialized new methods of studying shapes (Klingenberg *et al*., 2010). ImageJ and MorphoJ download link is http://www.java.com.

MorphoJ, the program used in this chapter is one of the few program packages that take into account the symmetry of landmark configurations throughout the analyses. It also contains advanced tools for analysing how shapes are separated or related (Klingenberg, 2011). MorphoJ provides different options of result presentation. The graphical outputs can be scatter plots, or standard type of graphs. It also provides graphs for the visualization of shape changes (Klingenberg, 2011).

3.2.5 Egg Landmarks (LM) acquisition

Eighteen landmarks taken from the dorsal view (Figure 3.2.4) and 14 landmarks for the lateral view (Figure 3.2.5) were analysed. Egg attributes used were those considered for the traditional morphometrics in addition to others. Attributes considered were Egg length (EL), Egg width including floats (EFW), float length (FL), middle deck/dorsal minimum width (MDW). This is also the egg width-float. Anterior deck/dorsal width (ADW) and posterior deck/dorsal width (PDW). Additional egg attributes considered just for the lateral morphology include anterior pole apex (APP), posterior pole apex (PPA) and base of the egg

(EB) (Table 3.2.1 and Figure 3.2.6). Acronyms were given for each attribute and most structures were named based on (Linley *et al*., 1993; Malhotra *et al*., 2000)

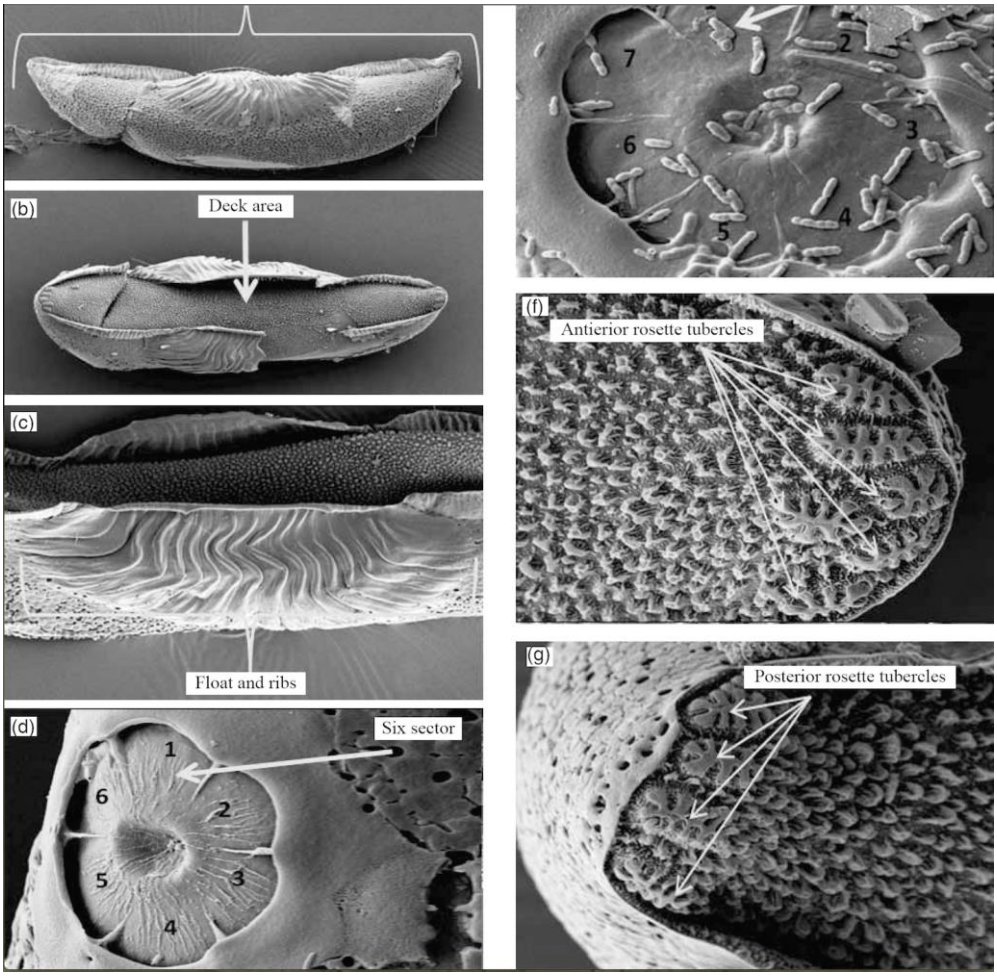

Figure 3.2.4: Scanning electron micrograph of *Anopheles stephensi* egg: (a) Lateral aspect; (b) Ventral aspect showing deck area; (c) Lateral aspect showing floats and ribs; (d) Micropylar disc showing 'six' sectors; (e) Micropylar disc showing 'seven' sectors; (f) Anterior; and (g) Posterior rosette tubercles **Source:** (Tyagi *et al*., 2017).

The data collated from ImageJ was used to build a Tps file. The Tps files were opened in MorphoJ software and the classifier variables (species, population and view) for all samples were imported following the software instructions.

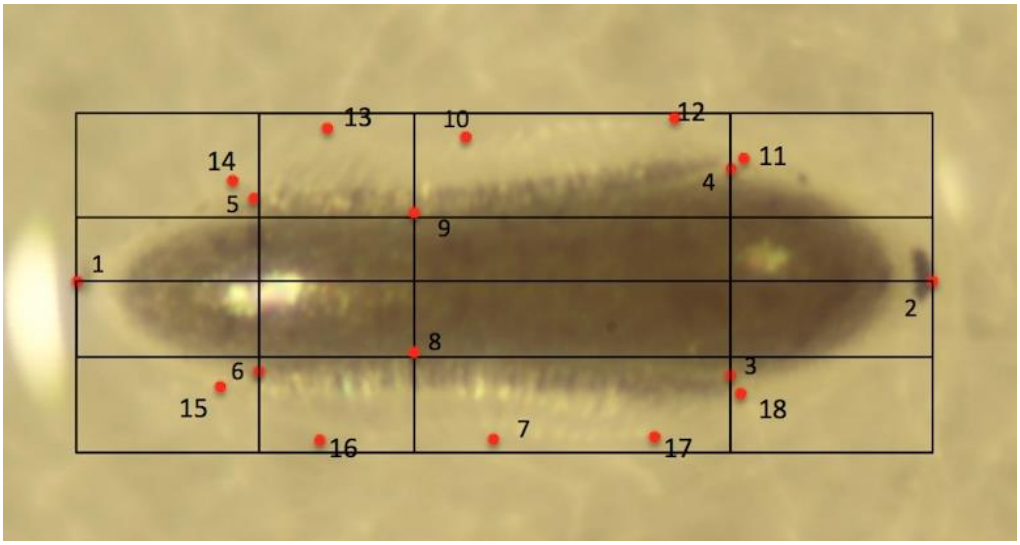

Figure 3.2.5: *Anopheles gambiae* egg as seen from the dorsal view. Eighteen landmarks (LM) were used (LM 1 to 18) were used to describe egg structures (Table 3.2.1). LM 1 and 2= EL, LM 3 and 4= ADW, LM 4 and 5 = FW, LM 5 and 6= PDW, 8 and 9= MDW. LM 7 to 10 = EFW. LM 11, 12, 13,14, 15,16,17 and 18 all describe the floats (Table 3.2.1).

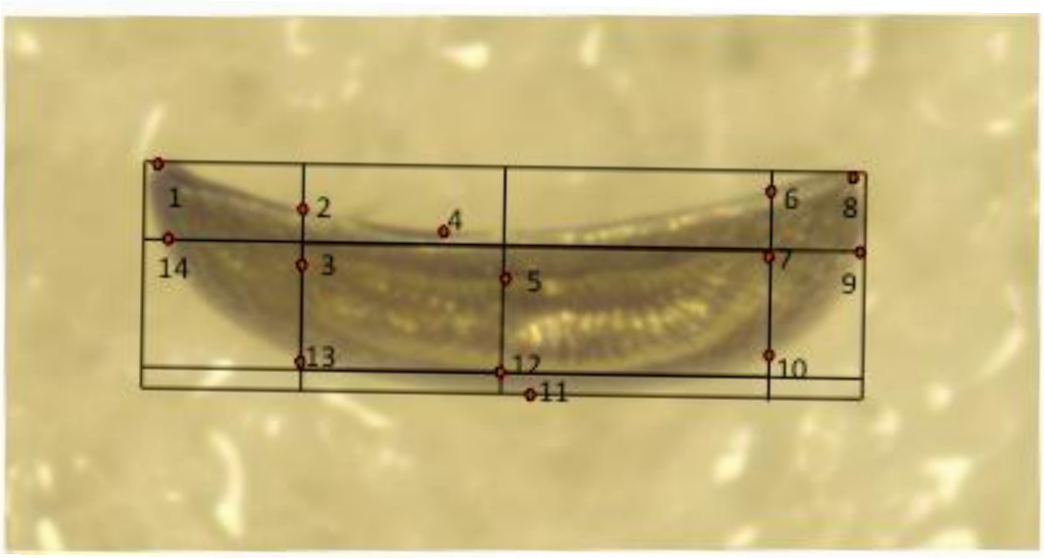

Figure 3.2.6: *Anopheles gambiae* egg as seen from the lateral view. Fourteen landmarks were used to describe egg structure (Table 3.2.1). LM1= PPA, LM 8=APP, LM 9 and 14= EL, LM 6 and 10= ADW, LM 5 and 12 = FW, LM 2 and 13= PDW, LM 4 and 11 = EFW (LM 4 shows the most interior point of the egg, LM 11 also shows the lowest base of the egg (EB). Landmarks 3,7,5 and 12 describe the floats (Table 3.2.1).

Table 3.2.1: Egg attributes and landmarks.

Dorsal view Landmark	Lateral view Landmark	Landmark title	Acronym	Further description
1,2	9,14	Egg Length	EL	Egg length from anterior micropylar region to the posterior apex.
7,10	4,11	Egg Width + floats	EFW	Width of narrowest point of entire egg with floats.
	11	Egg base	EB	Lowest point of the egg when on lateral side.
9,10	5,12	Float length	FL	Float length excluding deck.
8,9		Middeck minimum width	MDW	Most interior point of the deck.
3,4	6,10	Anterior deck width	ADW	Width of Deck around the egg anterior region.
5,6	2,13	Posterior deck width	PDW	Width of Deck around the egg posterior region.
	8	Anterior pole apex	APP	Most anterior part of the egg.
	1	Posterior pole apex	PPA	Most posterior part of the egg.
	3,7	Float	F	Most anterior and posterior part of float cluster.
11,12,13,14, 15,16,17,18		Float	F	The landmarks all describe the float in different ways.

3.2.6 Outlier detection

Outlier detection was the first step carried out to correct abnormalities in the data set. To detect outliers, a 'find outliers' test was performed using MorphoJ. This test shows the average egg shape as a configuration of blue dots. Outliers were seen as red lines shown as deviation of individual samples from the list of samples selected. All points considered as outliers were noted and the eggs measured again or substituted when appropriate.

3.2.7 Statistical test

Shapes can be compared by superimposition and comparing the difference in the position of landmark points (Rohlf, 1999). A procrustes test was performed for superimposition of the landmarks. This is usually the first step after identification of outliers. Procrustes fit aligns landmarks and centres them on one point. It eliminates variation due to differences in scale, orientation or position from coordinates. The data produced from the procrustes fit was used for subsequent analysis.

A covariance matrix was performed for the modularity and integration of the samples. This was performed on all the landmarks after superimposition to prevent differences from overall orientation for all analyses.

Procrustes ANOVA was performed to quantify relative amounts of variation at different levels in relation to size and shape in the four populations studied. The analysis resulted in the sums of squares (SS), Individual shape variation and residual, which is the variability among replicate.

A Principal Component Analysis (PCA) was performed to enhance chances of finding differences. The first few principal components (PCs) usually account for most variation in a data set leading to data reduction. Another importance of PCA it that it shows graphical structural differences among the populations studied and PCA can also show specific features

of variation for comparison (Klingenberg and McIntyre, 1998). In order to analyse the data for population and species discrimination, Canonical variance analysis (CVA) and Discrimination analysis (DFA) were used. They were used to test differences between groups, plot differences and also predict affiliation (Viscosi, 2015). Canonical variance analysis was used in determining important features that best distinguish between populations. The Discriminant function analysis (DFA) was conducted to observe separation between paired groups and a cross validation to assess the reliability of the discriminant function. Accuracy of species and population classification was tested using the Mahalanobis distance by comparing individual egg with the overall mean egg size for each population and also species.

3.3 Results

3.3.1 Traditional morphometrics

An. coluzzii and *An. gambiae* eggs were observed to have similar shape. The difference in mean egg length (EL) between both species was not significant (Tables 3.3.1). Length was measured from the tip of the anterior end of the egg to the tip of the posterior end. The egg width including floats (EFW) was significantly different between the two species (Tables 3.3.1). This measurement was the full width of the egg including the floats. The difference in middeck width (MDW) between the two sibling species was highly significant. Anterior deck width (ADW) and posterior deck width (PDW) were significant between the species (Table 3.3.1).

**Table 3.3.1: ***Comparative dimensions of morphometrics attributes of *Anopheles coluzzii* and *Anopheles gambiae* (one way ANOVA).

Egg attributes (µm)	*Anopheles coluzzii*	*Anopheles gambiae*	*F*-value	Prob> *F*
EL	514.70 (509.28-520.13)	517.98 (512.56-523.42)	0.72	0.3982
ADW	132.44 (129.9-134.98)	128.143 (125.6-130.68)	5.61	0.0195
PDW	102.58 (100.09-105.08)	98.53 (96.04-101.02)	5.18	0.0247
EFW	182.05 (179.16-184.94)	174.85 (171.96-177.74)	12.17	0.0007
MDW	91.08 (88.49-93.68)	78.50 (75.91-81.1)	46.14	<0.0001
ADW/EL	0.26 (0.25-0.26)	0.25 (0.24-0.25)	8.65	0.0039
PDW/EL	0.20 (0.19-0.20)	0.19 (0.19-0.19)	7.94	0.0057
EFW/EL	0.35 (0.35-0.36)	0.33 (0.33-0.34)	15.69	0.0001
MDW/EL	0.18 (0.17-0.18)	0.15 (0.15-0.16)	56.46	<0.0001

*Sample size is 60 for the two species, populations were pooled. Means within rows with 95% confidence intervals of the mean in bracket Measurements are in (µm). Full meaning of acronyms (Table 3.2.1).

All egg attributes were corrected for by the egg length ADW/EL, PDW/EL, EFW/EL and MDW/EL. There was a significant difference in the ratios of egg attributes between the two species studied. The ratios were also significant in the four populations studied. The *df*= 1 for species and *df*= 3 for the populations. The results show variations between species and between populations within species (Figure 3.3.1, Table 3.3.2). The mid-deck region (MDW) was 78.50 ± 1.30μm for *An. gambiae*, this is the width of the egg without the floats, and the EFW/EL ratio was 0.33 ± 0.01μm for *An. gambiae* (Table 3.3.1). A significant difference was observed in MDW between the two species $P< 0.001$, there was also a difference in the egg length and egg width ratio EFW/EL $P< 0.001$. The results revealed variations between populations (Figure 3.3.1, Table 3.3.2). EL and MDW varied between the four populations. $P< 0.001$ for the two attributes.

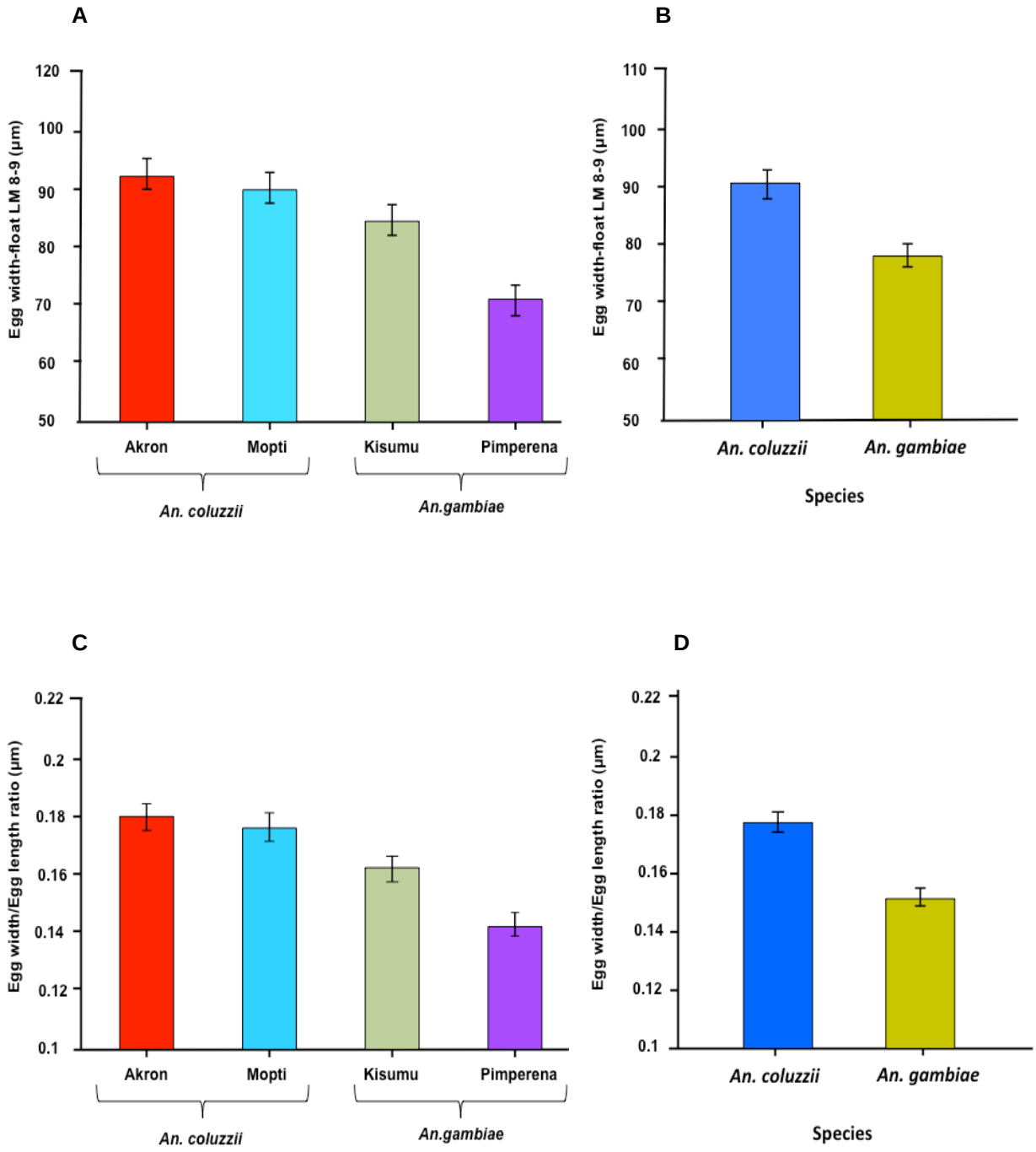

Figure 3.3.1: Bar plots showing egg attributes. Egg width –floats (A= Population, B= species) Egg length and width ratio (C shows the ratio for population and D shows the species)

Table 3.3.2: Comparative account of egg morphometrics attributes of four populations of *Anopheles coluzzii* and *Anopheles gambiae**

	An. coluzzii		***An. gambiae***			
Egg attributes	**Akron Mean**	**Mopti**	**Kisumu**	**Pimperena**	***F*-value**	**Prob > *F***
EL	517.46 (510.45-524.47)	511.95 (504.94- 518.96)	530.33 (523.32-537.34)	505.65 (498.64-512.66)	8.79	<0.0001
ADW	136.69 (133.88-139.5)	128.19 (125.39-131)	135.88 (133.07-138.68)	120.41 (117.6-123.21)	28.94	<0.0001
PDW	102.72 (99.28-106.16)	102.45 (99.00 -105.89)	101.97 (98.526-105.41)	95.09 (91.647-98.53)	4.42	0.0055
EFW	186.99 (183.35-190.63)	177.11 (173.47-180.75)	180.38 (176.74-184.03)	169.32 (165.68-172.96)	15.96	<0.0001
MDW	92.51 (89.28-95.74)	89.66 (86.43-92.89)	85.31 (82.08-88.535)	71.71 (68.479-74.935)	12.47	<0.0001
ADW/EL	0.26 (0.26-0.27)	0.25 (0.24-0.26)	0.26 (0.25-0.26)	0.24 (0.23-0.24)	12.47	<0.0001
PDW/EL	0.20 (0.19-0.21)	0.20 (0.19-0.21)	0.19 (0.19-0.20)	0.19 (0.18-0.19)	2.91	0.0375
EFW/EL	0.36 (0.35-0.37)	0.35 (0.34-0.35)	0.34 (0.33-0.35)	0.34 (0.33-0.34)	8.28	<0.0001
MDW/EL	0.18 (0.17-0.19)	0.18 (0.17-0.18)	0.16 (0.15-0.17)	0.14 (0.14-0.15)	27.44	<0.0001

*Sample size is 30 for each population. Means within rows with 95% confidence intervals of the mean in bracket Measurements are in (µm). Full meaning of acronyms (Table 3.2.1).

Pimperena was observed to be generally smaller than the other three populations (Fig 3.3.1). The population was observed to have smaller mean values of EL, ADW, PDW, EFW and MDW smaller than all three populations (Table 3.3.2).

Table 3.3.3: *Effect of test for simple (traditional) morphometrics comparison of egg attributes for sum of squares and populations.

	Population				Species			
Egg attributes	***DF***	**Sum of Squares**	***F*-Value**	**Prob > *F***	***DF***	**Sum of Squares**	***F*-Value**	**Prob > *F***
EL	2	9585.71	12.75	<0.001	1	324.03	0.86	0.355
ADW	2	4673.48	38.80	<0.001	1	554.36	9.20	0.003
PDW	2	710.81	3.91	0.023	1	492.74	5.43	0.022
EFW	2	3301.49	16.27	<0.001	1	1554.86	15.33	0.0002
MDW	2	2896.19	18.18	<0.001	1	4745.52	59.57	<0.001
ADW/EL	2	0.008	13.46	<0.001	1	0.003	10.48	0.002
PDW/EL	2	0.0003	0.43	0.648	1	0.003	7.86	0.006
EFW/EL	2	0.0041	4.16	0.0180	1	0.008	16.53	<0.001
MDW/EL	2	0.0057	9.08	0.0002	1	0.020	64.18	<0.001

*Total number of samples per population was 30 and 60 per species. Measurements were taken in µm. Full meaning of acronyms (Table 3.2.1)

Akron eggs were wider in width compared to the other three populations (Fig 3.3.1) and when the two species were compared, *An. coluzzii* was observed to have a wider egg width compared to *An. gambiae*. The effect test for species was significant in all attributes measured except EL. An effect of population was also significant in all attributes except PDW/EL (Table 3.3.3).

3.3.2 Landmarks geometric morphometrics (Dorsal morphology)

Superimposition of data

Procrustes test was performed for superimposition of data. This was done to remove any variation due to size, position or orientation. The coordinates of the superimposed landmarks (Table 3.3.4) consist of the 18 landmarks and the X and Y-axis, which are the mean coordinates of corresponding landmarks. These coordinates were used for the multivariate analysis. The graphical output of the procrustes fit (Figure 3.3.2) further revealed some outliers in the data set and this was corrected.

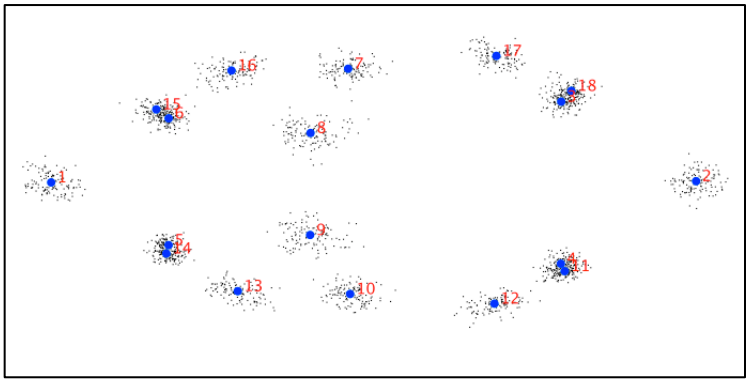

Figure 3.3.2: Procrustes fit for *Anopheles gambiae* and *Anopheles coluzzii* eggs. Data includes four different populations (Akron, Mopti, Kisumu and Pimperena). The data set contains 18 landmarks obtained from the dorsal view of a 120 eggs. The procrustes fit shows the landmark configurations of all eggs in the data set after superimposition. The blue circle represents the location of each landmark with numbers in red and the black dots surrounding the circle represents the location individual eggs.

Outliers were seen as black dots away from similar clusters representing individual eggs around landmark locations. Dots considered as outliers were checked thoroughly for possible errors because superimposition standardizes every specimen to a unit centroid size resulting to an estimated size of the structure studied.

Table 3.3.4: Dorsal morphology of mean Procrustes Coordinates.

Landmark (L)	Axis 1 (x)	Axis 2 (y)
1	-0.3578	-0.0009
2	0.3943	0.0005
3	0.2369	0.0935
4	0.2366	-0.0959
5	-0.2209	-0.0742
6	-0.2208	0.0740
7	-0.0116	0.1320
8	-0.0554	0.0569
9	-0.0558	-0.0624
10	-0.0092	-0.1314
11	0.2410	-0.1047
12	0.1591	-0.1427
13	-0.1402	-0.1281
14	-0.2235	-0.0841
15	-0.2352	0.0844
16	-0.1475	0.1299
17	0.1614	0.1470
18	0.2488	0.1061

3.3.3 Morphometrics analysis of egg Shape variation between populations of *Anopheles coluzzii* and *Anopheles gambiae* (dorsal morphology).

The morphometric analysis of eggs for the dorsal morphology was conducted to identify differences in egg shape between the two sibling species.

Procrustes ANOVA

Procrustes ANOVA assesses the relative amount of variation among individuals and this was computed for both species and populations (Table 3.3.5; Table 3.3.6). A significant difference was found for effect of species ($P< 0.001$) on shape and ($P= 0.008$) for centroid size (Table 3.3.5). The results showed variation in shape and centroid size among populations ($P< 0.001$) (Table 3.3.6).

Table 3.3.5: Procrustes ANOVA for the effect of species showing the sum of squares (SS) and mean sum of squares (MS).

Effect	SS	MS	*DF*	*F*	*P*-value
Centroid size					
Species	7862.06	7862.06	1	7.16	0.008
Residual	129621.80	1098.48	118		
Shape					
Species	0.05	0.001	32	8.96	<0.001
Residual	0.64	0.0002	3776		

Table 3.3.6: Procrustes ANOVA for the effect of population showing the sum of squares (SS) and mean sum of squares (MS).

Effect	SS	MS	*DF*	*F*	*P*-value
Centroid size					
Populations	48005.31	16001.77	3	20.74	<0.001
Residual	89478.55	771.36	116		
Shape					
Populations	0.10	0.001064	96	6.67	<0.001
Residual	0.59	0.000159	3712		

Principle component analysis

The principal component (PC) analysis for the dorsal morphology revealed some shape variation between the data set of 120 eggs pooled from the 4 populations. The first four PCs accounted for 58.95% of the total variation and PC1 to PC10 accounts for 85.98% of the total variation (Table 3.3.7 and Figure 3.3.3).

Table 3.3.7: Egg shape variables in *Anopheles coluzzii* and *Anopheles gambiae* Populations. Principal Component Analysis: PCA: CovMatrix, Dorsal Morphology, Procrustes coordinates of 120 samples with 18 landmarks configuration of eggs.

Principal components	Eigenvalues	% Variance	Cumulative %
PC1	0.00137	23.62	23.62
PC2	0.00106	18.25	41.87
PC3	0.00058	10.07	51.94
PC4	0.00041	7.02	58.96
PC5	0.00034	5.98	64.93
PC6	0.00034	5.90	70.84
PC7	0.00028	4.85	75.69
PC8	0.00022	3.76	79.45
PC9	0.00021	3.66	83.11
PC10	0.00017	2.88	85.98

A total of 32 principal components were described to account for100% of the total variation. PC11 to PC32 were not included in the table because PC1 to PC10 accounts for most variation. PC11 to 32 accounts for less than 2% variation each (Figure 3.3.3). PC1 is the most important with a % variance of 23.62 (Figure 3.3.3).

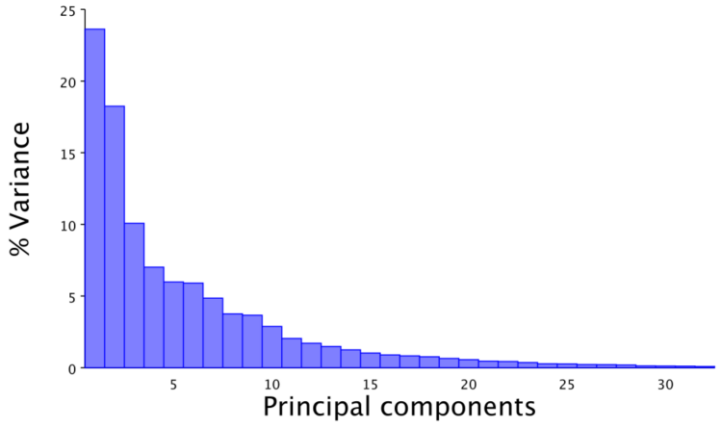

Figure 3.3.3: Histogram of the Eigenvalues of the 32 principal components (PC) based on the 18 landmarks for the dorsal morphology of the four populations of *Anopheles coluzzii* and *Anopheles gambiae*.

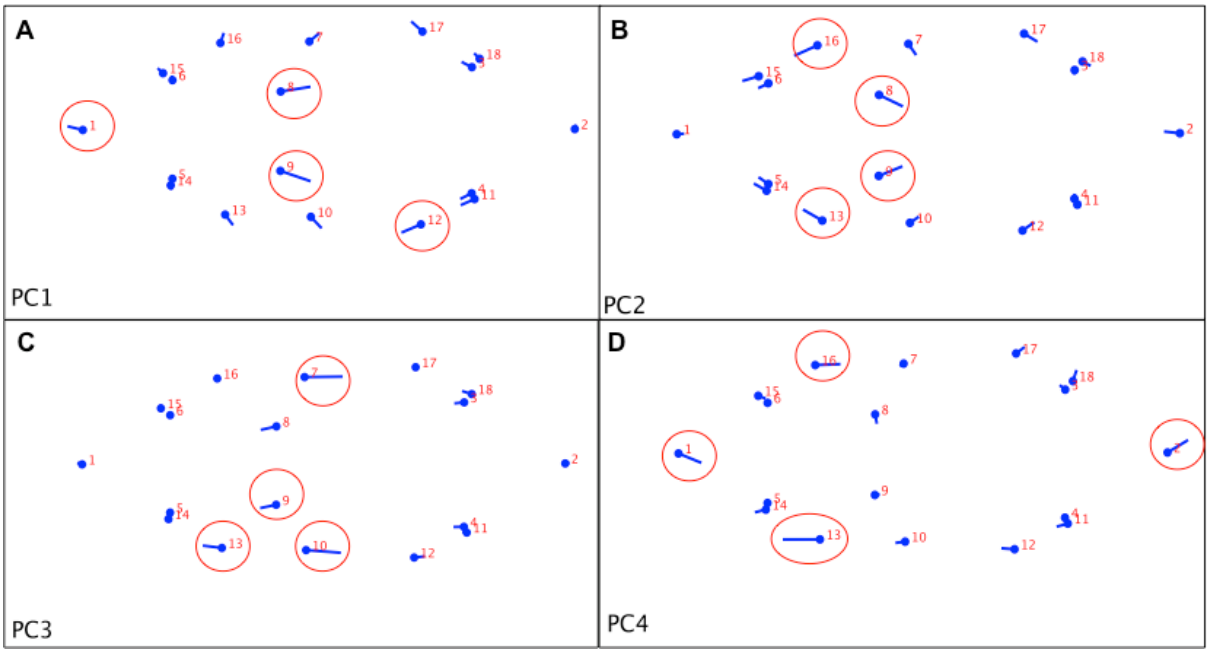

Figure 3.3.4: A lollipop graph showing shape changes associated with PC1 (A), PC2 (B), PC3 (C) and PC4 (D) for the dorsal morphology. Lines stretching out from landmarks moving in a particular direction indicate movement of landmarks from the starting shape to the target shape. Reds boxes contain the 4 most significant landmarks on each PC.

The first Principal Component showed several differences in the structural components of the egg. Lines in red circles are those with the most variation and corresponding high coefficient are coloured red (Figure 3.3.4 and Table 3.3.8). Variation at LM8 and LM9 was observed (Figure 3.3.4A) this is the mid-deck/dorsal width (MDW). The variation corresponds to the high values observed at landmark x8 and x9 (Table 3.3.8) for the principal component coefficient. High coefficient values are also recorded at Landmark x8 and x9 for PC2 (Table 3.3.8) hence corresponding directional movements at L8 and L9 (Figure 3.3.4B). PC2 axis also shows shape variation at LM13 and LM16, which is a part of the float. A deviation at LM13 and LM16 was also observed in PC4. The principal component coefficient (Table 3.3.8) showed high PC values at landmark 7, 9, 10 and 13 for PC3 corresponding to the

deviations observed in the lollipop graphs (Figure 3.3.4). Landmarks 7 and 10 are the width of the egg (EFW).

Table 3.3.8: Principal component coefficient for the 18 landmarks used for the dorsal morphology.

Landmark	PC1	PC2	PC3	PC4
x1	-0.234	0.112	-0.077	0.352
y1	0.056	0.008	0.024	-0.139
x2	0.018	-0.237	0.067	0.316
y2	0.067	0.025	0.008	0.184
x3	-0.158	-0.010	-0.154	-0.085
y3	0.079	-0.068	-0.0153	0.066
x4	-0.174	0.004	-0.162	-0.034
y4	-0.086	0.075	-0.002	-0.016
x5	-0.012	-0.145	0.025	-0.003
y5	-0.075	0.101	-0.049	-0.038
x6	-0.010	-0.153	0.034	-0.009
y6	0.068	-0.073	0.041	-0.033
x7	0.153	0.116	0.593	-0.061
y7	0.127	-0.165	0.006	-0.034
x8	0.456	0.356	-0.244	0.026
y8	0.071	-0.169	-0.052	-0.139
x9	0.462	0.354	-0.249	0.076
y9	-0.155	0.141	-0.051	0.006
x10	0.164	0.134	0.541	-0.149
y10	-0.172	0.091	-0.039	-0.018
x11	-0.212	0.041	-0.067	-0.172
y11	-0.092	0.050	0.005	-0.045
x12	-0.297	0.176	0.164	-0.202

y12	-0.124	0.122	0.017	0.014
x13	0.120	-0.288	-0.297	-0.571
y13	-0.159	0.162	0.035	-0.001
x14	0.008	-0.191	0.046	-0.169
y14	-0.071	0.105	-0.044	-0.039
x15	-0.082	-0.244	-0.005	0.112
y15	0.075	-0.070	0.057	-0.041
x16	0.053	-0.347	-0.016	0.387
y16	0.153	-0.154	0.014	0.013
x17	-0.171	0.199	-0.048	0.131
y17	0.152	-0.118	-0.003	0.096
x18	-0.083	0.122	-0.151	0.057
y18	0.0840	-0.065	0.045	0.166

*Numbers in red indicate the first four landmarks with the most variation

Scatter plots consisting of individual eggs showed slight dissociation between populations and species along the first two principal components PC1 and PC2 axis (Figure 3.3.5A) while PC3 and PC4 did not show any dissociation between populations (Figure 3.3.5B).

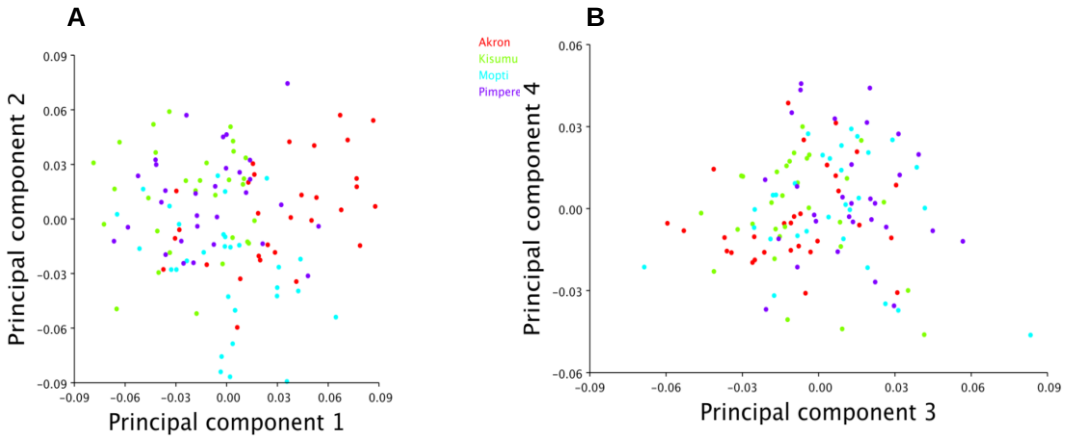

Figure 3.3.5: Scatter plot of Principal component scores for four populations of *Anopheles coluzzii* and *Anopheles gambiae ss* species (Akron, Kisumu, Mopti and Pimperana).

3.3.4 Egg shape discrimination among populations and species

Canonical variate analysis (CV)

The canonical variate analysis showed shape variation among species. The graph shows changes in association to the canonical variate (CV). The deviations on the landmarks are an outcome of the regression of shape onto the scores for the respective CV (Rohlf, 1996) The analysis was carried out between the two species and between populations. Shape variations between species occurred in landmarks 8 and 9 the mid-deck/dorsal region (MDW). Variations also occurred at landmarks 12,13,16 and 17 that describe the floats. The CV scores showed partial overlap between species (Figure 3.3.6) however; grouping within species was quite distinct (Figure 3.3.7).

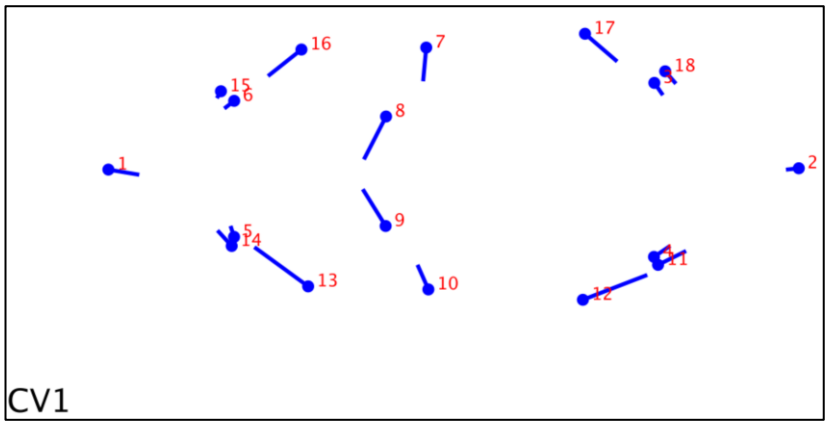

Figure 3.3.6: Canonical Variate analysis showing changes in shape in dorsal morphology between species (*Anopheles coluzzii* and *Anopheles gambiae).*

The canonical variance analysis for *An. coluzzii* and *An. gambiae* showed a Mahalanobis distance of (2.46), the procrustes distance of (0.04). There was a significant difference at *P*< 0.001 between species.

The canonical variate analysis showed significant difference (*P*< 0.001) for mahalanobis and procrustes distance for all pairwise analysis for the four populations studied (Table 3.3.10) Pimperena and Kisumu pairwise analysis was (*P*= 0.001). This was examined

after permutation test (10.000 permutation runs). Variations between populations were observed at Landmark 8, 13, 15 and 17 for CV1, landmark 4, 5, 6, and 11 for CV2 and landmark 5, 6, 8 and 13 in CV3 (Figure 3.3.8; Table 3.3.9).

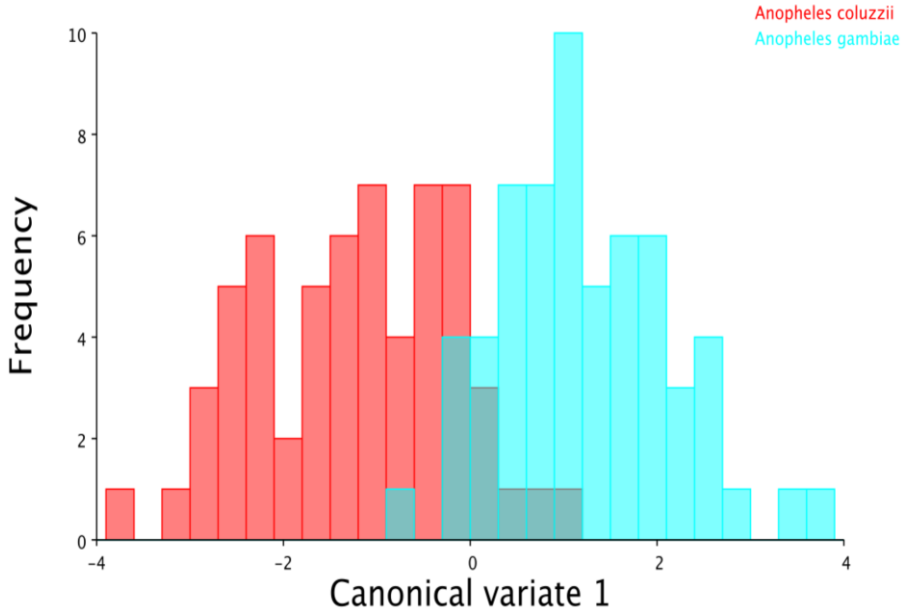

Figure 3.3.7: Scatter plot of CVA for species (*Anopheles coluzzii* and *Anopheles gambiae)*.

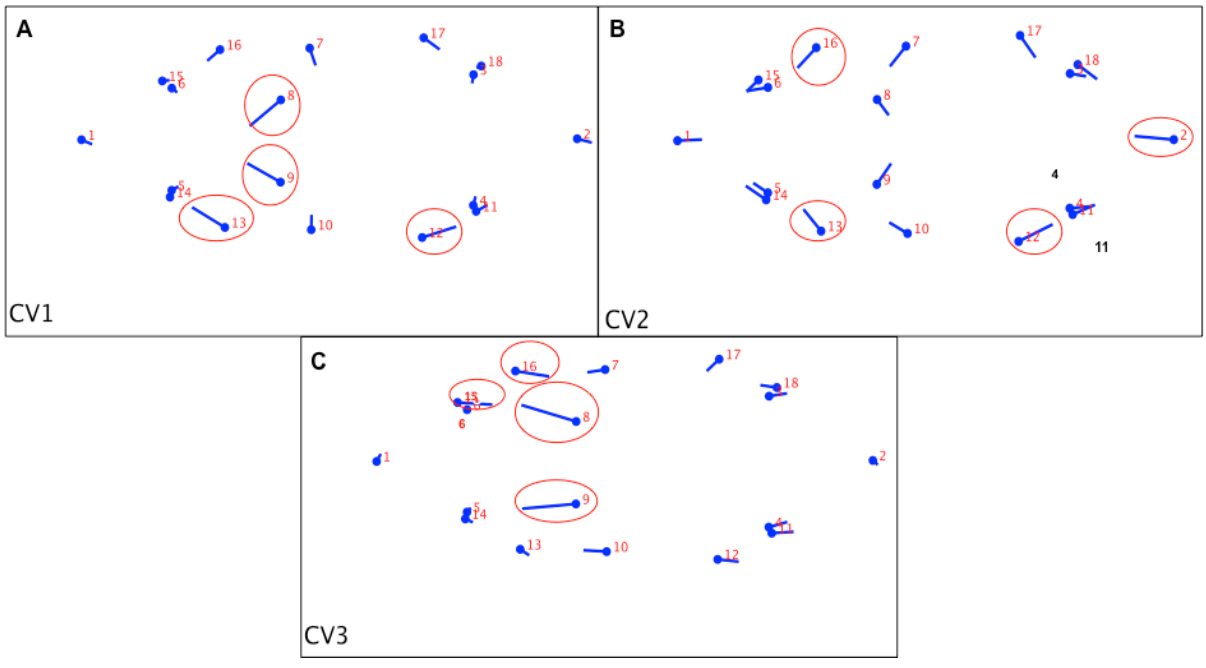

Figure 3.3.8: Canonical Variate analysis showing changes in shape dorsal morphology of the different populations (Akron, Kisumu, Mopti and Pimperena).

**Table 3.3.9: ***Canonical coefficient for the 18 landmarks for the dorsal Morphology. CV1, CV2 and CV3 show canonical coefficient for the four different populations (Akron, Kisumu, Mopti and Pimperena) and CV (for species) shows the canonical coefficient for the two species (*Anopheles coluzzii* and *Anopheles gambiae*).

Landmark	CV1	CV2	CV3	CV1 (For species)*
x1	-13.501	9.617	-1.333	-3.131
y1	6.943	10.294	46.708	6.620
x2	25.645	-14.736	6.370	7.891
y2	-16.119	-14.003	-39.996	-16.535
x3	-29.977	-10.857	-12.332	-27.612
y3	-26.187	59.151	33.739	17.089
x4	44.050	24.249	46.929	42.872
y4	30.790	-103.035	24.099	-49.990
x5	43.573	73.225	62.080	73.835
y5	-51.301	-2.567	72.828	-47.210
x6	-23.565	-66.899	-108.596	-49.755
y6	7.010	79.638	-27.400	61.946
x7	11.255	2.236	2.081	9.417
y7	-52.191	-19.435	-18.252	-48.801
x8	-14.630	12.040	-61.879	4.621
y8	-72.253	-15.128	35.550	-66.581
x9	-2.508	1.404	44.706	-5.990
y9	39.622	25.043	-13.761	47.154
x10	-13.402	-23.957	-5.331	-25.255
y10	0.484	-72.972	-3.339	-48.501
x11	-19.202	3.701	-2.798	-11.082
y11	-15.909	103.887	-25.713	61.522
x12	23.965	16.137	6.768	27.460
y12	-32.014	30.769	14.722	-4.100

x13	-1.319	2.934	-19.815	3.295
y13	74.214	13.947	-83.804	72.731
x14	0.159	-27.783	33.874	-22.510
y14	-8.662	16.343	17.399	2.763
x15	7.489	-11.501	31.572	-5.955
y15	62.341	-43.876	-2.214	15.777
x16	-6.733	5.416	9.895	-2.352
y16	-14.868	-14.347	-12.461	-19.015
x17	15.221	-5.073	-2.881	7.925
y17	58.211	-27.917	-42.698	28.192
x18	-46.519	9.846	-29.309	-23.673
y18	9.888	-25.791	24.593	-13.060

*Numbers in red indicate the first four landmarks with the most variation

CV1 makes up for 51.67% of the total variation and shows same trend in shape change as PC1 (Figure 3.3.8A). CV2 makes for 27.35% of variation and CV3 20.97% of variation where 100% cumulative variance was attained (Figure 3.3.8 B and C). There was indication of clustering in the canonical variate analysis. Scatter plots of CV1 and CV2 axis showed clear clusters of populations with very little overlapping (Figure 3.3.9A) same effect was observed between species (Figure 3.3.9C). CV1 and CV3 axis also showed evidence of clustering between populations (Figure 3.3.9B) and species (Figure 3.3.9D).

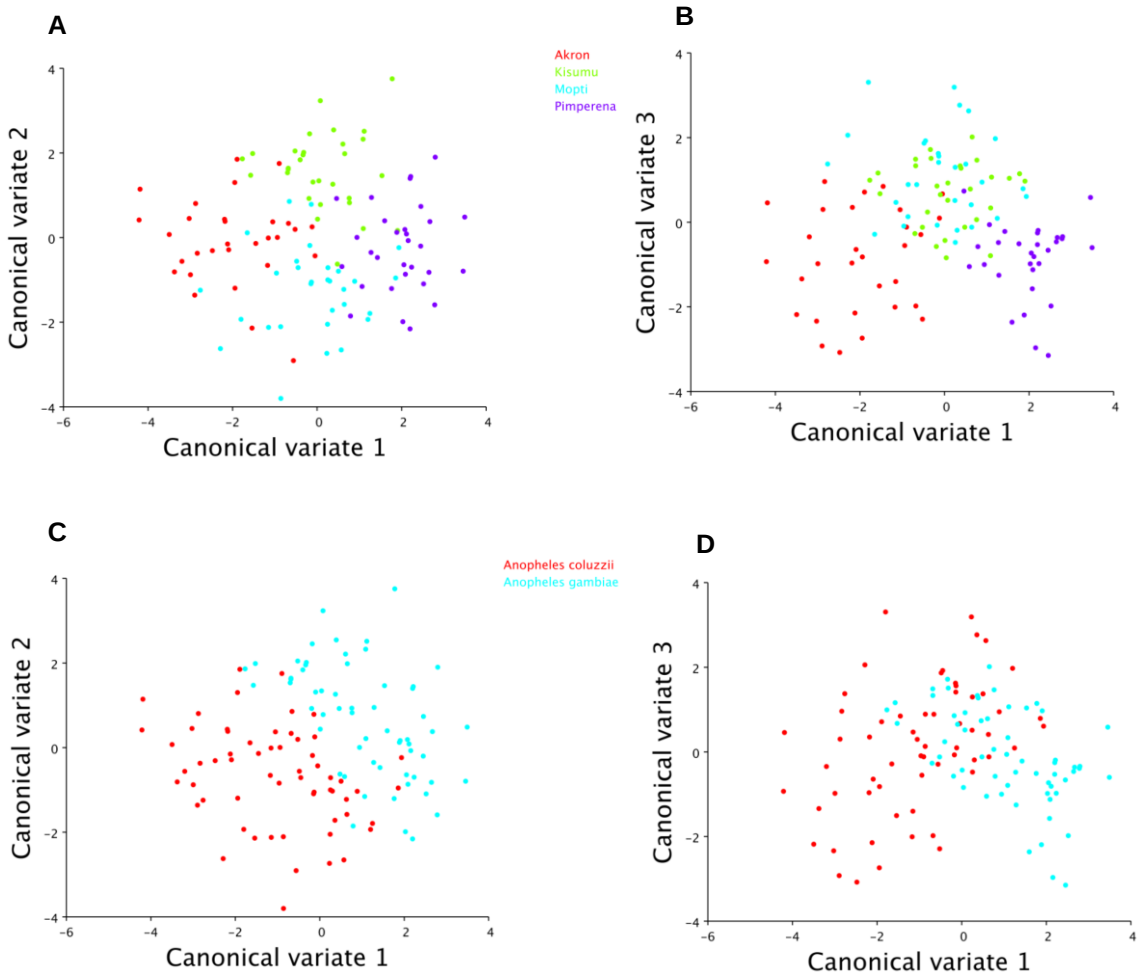

Figure 3.3.9: Morphological space of two canonical variates. A and C show the first two canonical variate components axis. B and D show the first and third canonical variate component axis. A and B = population. C and D = species.

Table 3.3.10: Results from Canonical variate analysis showing Mahalanobis and procrustes distances and *P*-values for the four populations in dorsal morphology.

Population	Mahalanobis distances	*P*-values for Mahalanobis distances	Procrustes distances	*P*-values for Procrustes distances
Kisumu - Akron	3.137	<0.001	0.056	<0.001
Mopti - Akron	2.99	<0.001	0.049	<0.001
Mopti - Kisumu	2.905	<0.001	0.047	<0.001
Pimperena - Akron	3.995	<0.001	0.052	<0.001
Pimperena - Kisumu	3.059	<0.001	0.032	0.001
Pimperena - Mopti	3.059	<0.001	0.042	< 0.001

Discriminant function analysis

The Discriminant function analysis for the two species (Figure 3.3.10) showed a *p*-value of (*P*< 0.001) for both mahalanobis and procrustes distance. The procrustes distance was 0.04 and a Mahalanobis distance of 2.46.

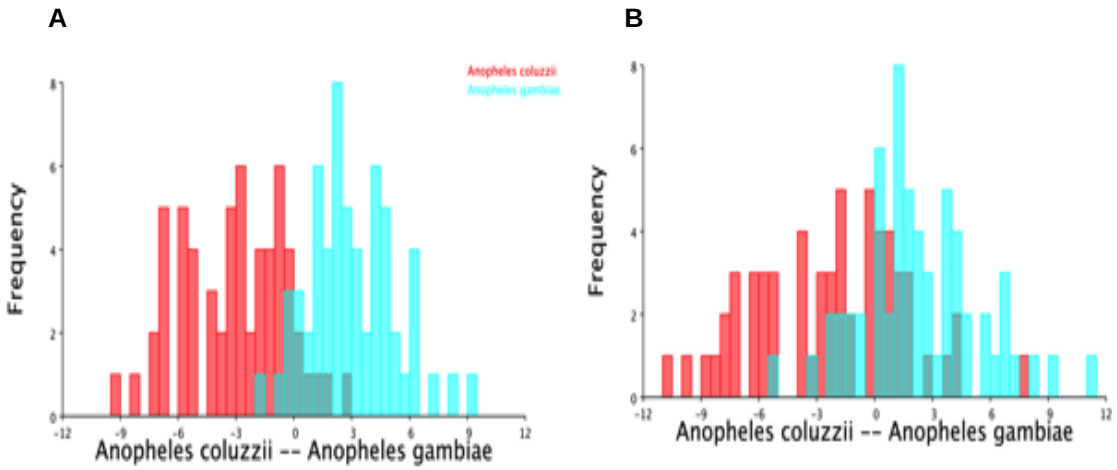

Figure 3.3.10: Dorsal morphology discriminant function (A) and cross-validation (B) for pairwise analysis between the two species (*Anopheles coluzzii* and *Anopheles gambiae*).

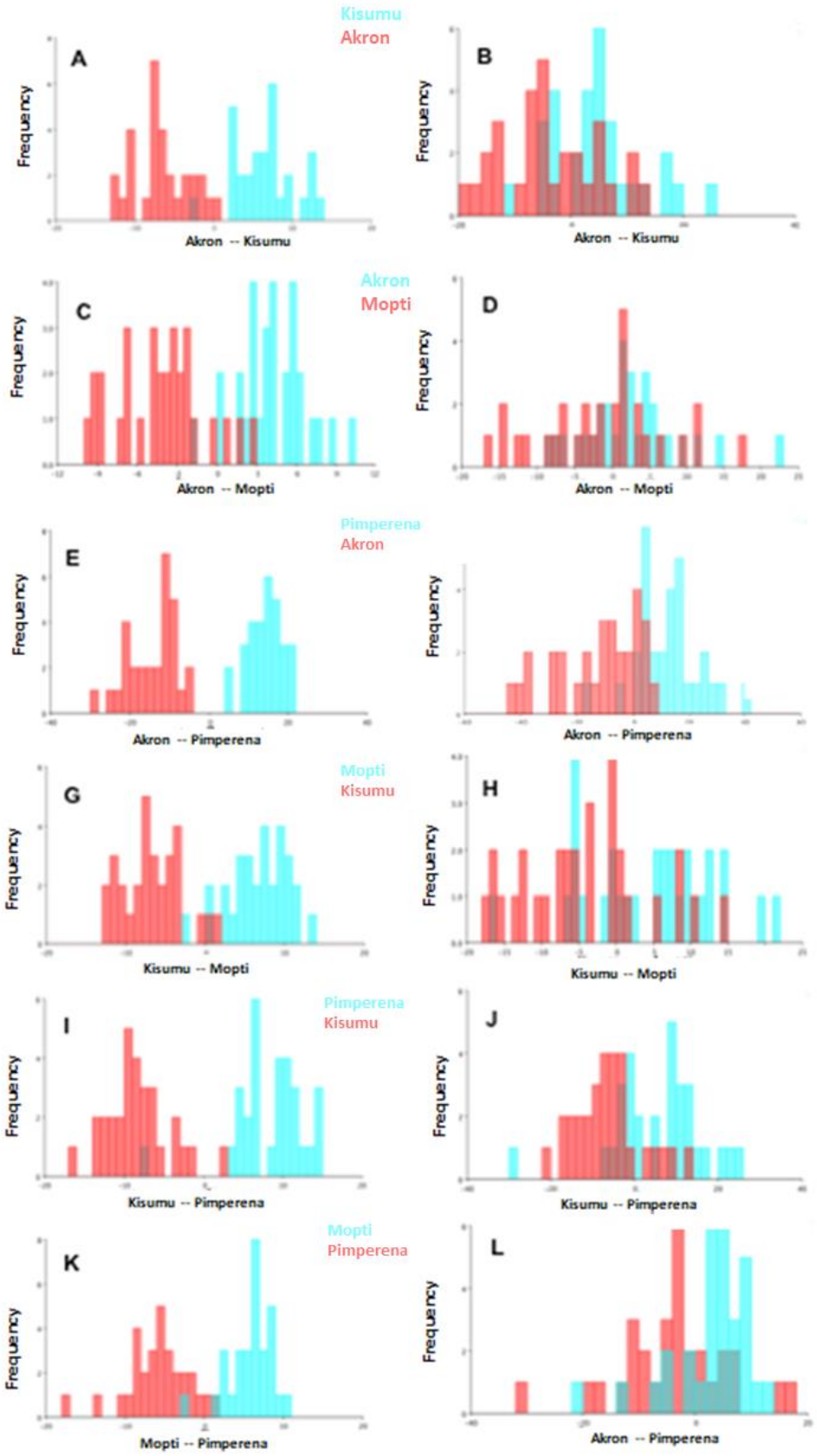

Figure 3.3.11: Dorsal morphology discriminant function and cross-validation for pairwise analysis between the four populations (Akron, Kisumu, Mopti and Pimperena) of *An. coluzzii* and *An. gambiae*.

A, C, E, G, I and K are the discriminant analysis graphs which B, D, F, H, J and L are the cross-validation graphs assessing the reliability of the discrimination.

A cross validation was done for the discriminant function and resulted in only 10% miscalculation in *An. coluzzii* and 8.3% in *An. gambiae*. The result for pairwise comparison cross-validation showed a misclassification of 31.7% for *An. coluzzii* and 16.7% for *An. gambiae*. The cross-validation classification for the different populations (Figure 3.3.11) ranged between 6 to 33%. Akron-Pimperena had 6.7% misclassifications for Pimperena and 26% for Akron. This is indicative of the clear separation devoid of overlapping (Figure 3.3.11).

Table 3.3.11: Difference in egg morphology among four populations of *Anopheles coluzzii* and *Anopheles gambiae* Mahalanobis and procrustes distance computed from the pairwise discriminant function analysis of the dorsal morphology.

Population	**Mahalanobis distances**	***P*-values for Mahalanobis distances**	**Procrustes distances**	***P*-values for Procrustes distances**
Akron - Kisumu	3.625	0.006	0.056	<0.001
Akron - Mopti	2.917	0.048	0.049	<0.001
Akron - Pimperena	5.344	<0.001	0.052	<0.001
Kisumu - Mopti	3.658	0.002	0.047	<0.001
Kisumu -Pimperena	4.096	<0.001	0.032	0.001
Mopti - Pimperena	3.424	0.01	0.042	<0.001
Anopheles coluzzii - *Anopheles gambiae*	2.462	<0.001	0.040	<0.001

3.3.5 Landmark morphometrics (Lateral morphology)

Superimposition of data

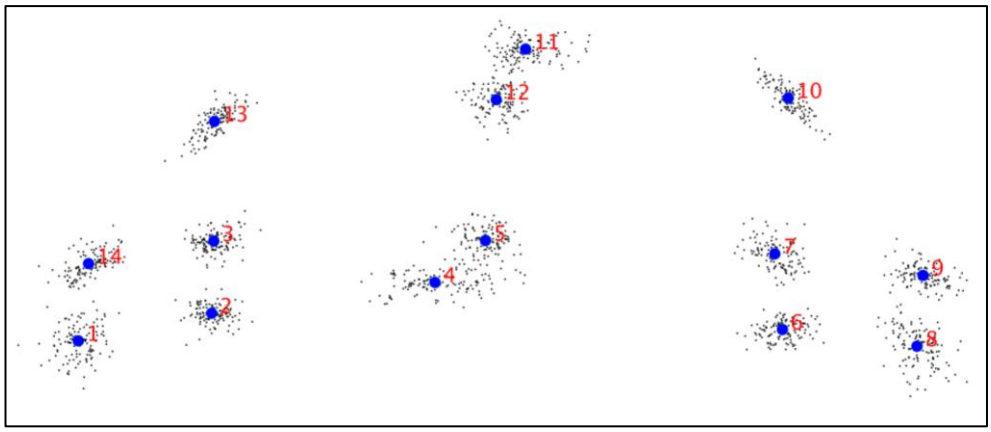

Figure 3.3.12: Procrustes fit for *Anopheles gambiae* and *Anopheles coluzzii* eggs. Data includes four different populations (Akron, Mopti, Kisumu and Pimperena). The data set contains 14 landmarks obtained from the lateral view of a 120 eggs. The procrustes fit shows the landmark configurations of all eggs in the data set after superimposition. The blue circle represents the location for each landmark and the black dots surrounding the circle represents the location individual eggs.

The Procrustes sum of square for the lateral morphology was 0.629 and the tangent sums of squares are 0.624. The procrustes fit was performed after identification and correction for outliers. The 14 landmarks used for the lateral morphology are shown as the blue dots.

Table 3.3.12: Lateral morphology of mean Procrustes Coordinates.

Landmark	Axis 1 (x)	Axis 2 (y)
1	-0.363	-0.094
2	-0.247	-0.070
3	-0.245	-0.007
4	-0.052	-0.043
5	-0.008	-0.006
6	0.250	-0.084
7	0.243	-0.018
8	0.367	-0.098
9	0.372	-0.037
10	0.255	0.116
11	0.026	0.159
12	0.001	0.115
13	-0.244	0.096
14	-0.354	-0.027

3.3.6 Morphometrics analysis of egg shape variation between populations of *Anopheles coluzzii* and *Anopheles gambiae* (lateral morphology)

The morphometrics analysis was conducted to access shape variation between populations of *An. gambiae* and *An. coluzzii* for the lateral morphology.

Procrustes ANOVA

Procrustes ANOVA were computed and the results showed variation in shape between species $P< 0.0001$. There was no variation in centroid size between the two species $P= 0.099$ (Table 3.3.13). Procrustes ANOVA for both shape and centroid size for population $P< 0.0001$ (Table 3.3.14).

Table 3.3.13: Procrustes ANOVA for the effect of species on size and shape showing the sum of squares (SS) and mean sum of squares (MS).

Effect	SS	MS	*DF*	*F*	*P*-value
Centroid size					
Species	3237.826	3237.826	1	2.77	0.099
Residual	137904.945	1168.685	118		
Shape					
Species	0.016	0.0006	24	3.19	<0.001
Residual	0.608	0.0002	2832		

Table 3.3.14: Procrustes ANOVA for the effect of population on size and shape showing the sum of squares (SS) and mean sum of squares (MS).

Effect	SS	MS	*DF*	*F*	*P*-value
Centroid size					
Individual	52728.050	17576.016	3	23.06	<0.001
Residual	888414.721	762.195	116		
Shape					
Individual	0.083	0.0012	72	5.98	<0.001
Residual	0.541	0.0002	2784		

Principle component analysis

The data set for the lateral morphology was pooled from 120 eggs with 14 landmarks each from two populations of *An. coluzzii* and two populations of *An. gambiae*. PC1 accounts for 23.737% of the total variation and PC4 68.203% of the total variation (Table 3.3.15).

Table 3.3.15: Egg shape variables in *Anopheles coluzzii* and *Anopheles gambiae*. Populations. Principal Component Analysis: PCA: CovMatrix, lateral Morphology, Procrustes coordinates of 120 samples with 14 landmarks configuration of eggs.

Principal components	Eigenvalues	% Variance	Cumulative %
1	0.00124	23.737	23.737
2	0.00089	16.962	40.699
3	0.00080	15.255	55.953
4	0.00064	12.25	68.203
5	0.00035	6.774	74.977
6	0.00027	5.279	80.256
7	0.00021	4.182	84.438
8	0.00015	2.876	87.314
9	0.00013	2.556	89.87
10	0.00008	1.654	91.524

At PC10, a cumulative of 91.524% of the total variation was accounted for (Figure 3.3.13) Since PC1-10 explains most of the variation (Table 3.3.15; Figure 3.3.13), PC 11-24 were not included in the table as they account for less than 9% of the total variation. PC1 shows variations at landmark 7,8, and 9 and 10. This is the anterior region of the egg (micropyle region) (Figure 3.3.14A).

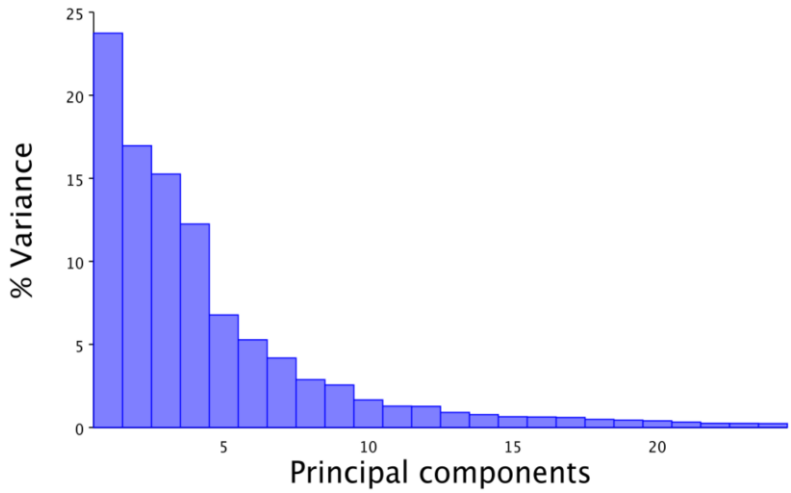

Figure 3.3.13: Histogram of the Eigenvalues of the 32 principal components (PC) based on the 18 landmarks for the dorsal morphology of the four populations of *Anopheles coluzzii* and *Anopheles gambiae.*

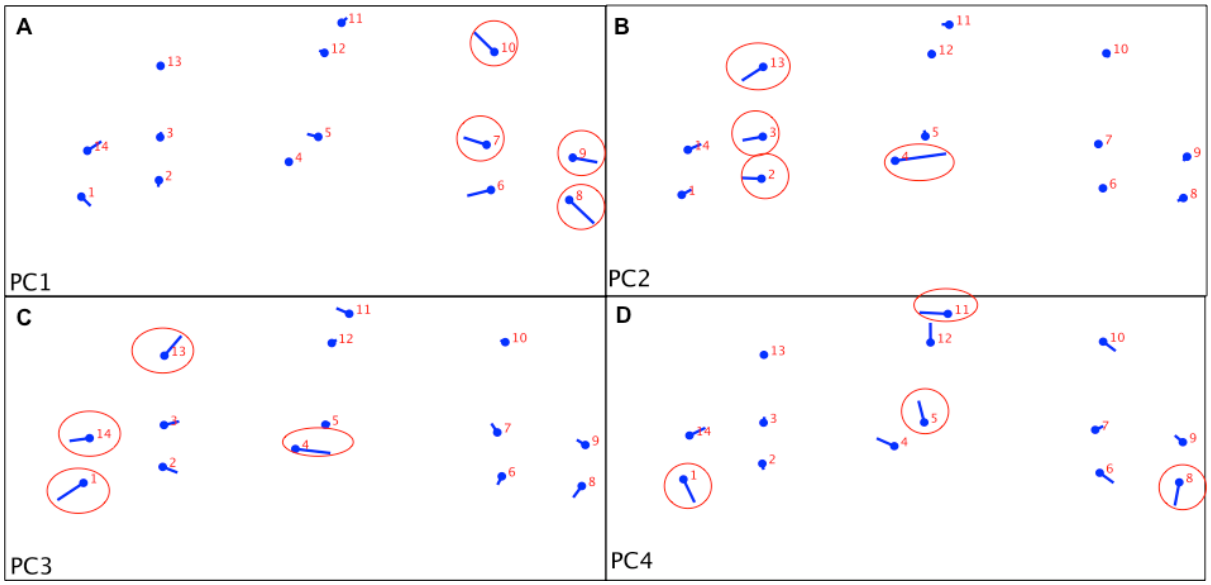

Figure 3.3.14: A lollipop graph showing shape changes associated with PC1 (A), PC2 (B), PC3 (C) and PC4 (D). Lines stretching out from landmarks moving in a particular direction indicate movement of landmarks from the starting shape to the target shape.

PC2 has a very high value of coefficient at landmark 4 (Table 3.3.16) this is seen by the corresponding variation in landmark 4 (Figure 3.3.14B). PC3 shows variation at landmark 1, 4, landmark 13 and 14 (Table 3.3.16; Figure 3.3.14C). Landmarks 1, 5, 8 and 11 show slight

variations in PC4 (Figure 3.3.14D). Scatter plot for the four different principal components does not show any striking clustering among populations (Figure 3.3.15A and B).

Table 3.3.16: Principal component coefficient for the lateral morphology

Landmark	PC1	PC2	PC3	PC4
x1	0.138	0.142	-0.382	0.164
y1	-0.135	0.079	-0.257	-0.354
x2	-0.010	-0.293	0.216	0.017
y2	-0.097	0.012	-0.085	-0.095
x3	0.014	-0.295	0.222	0.013
y3	0.075	-0.054	0.054	0.089
x4	0.048	0.743	0.513	-0.259
y4	-0.016	0.105	-0.061	0.119
x5	-0.167	-0.014	0.068	-0.083
y5	0.041	0.091	0.019	0.331
x6	-0.355	0.042	-0.063	0.204
y6	-0.085	0.017	-0.126	-0.153
x7	-0.343	0.008	-0.087	0.119
y7	0.105	-0.031	0.133	0.057
x8	0.371	-0.087	-0.125	-0.065
y8	-0.342	-0.046	-0.171	-0.357
x9	0.368	-0.054	-0.126	-0.118
y9	-0.066	-0.062	0.078	0.103
x10	-0.300	0.038	-0.071	0.184
y10	0.286	-0.061	0.027	-0.143
x11	0.082	-0.110	-0.189	-0.418
y11	0.071	0.005	0.083	0.021
x12	-0.079	-0.006	0.069	-0.002
y12	0.034	0.062	0.049	0.302

x13	0.021	-0.309	0.249	0.010
y13	-0.006	-0.204	0.297	-0.039
x14	0.210	0.195	-0.294	0.233
y14	0.135	0.086	-0.042	0.117

*Numbers in red indicate the first four landmarks with the most variation

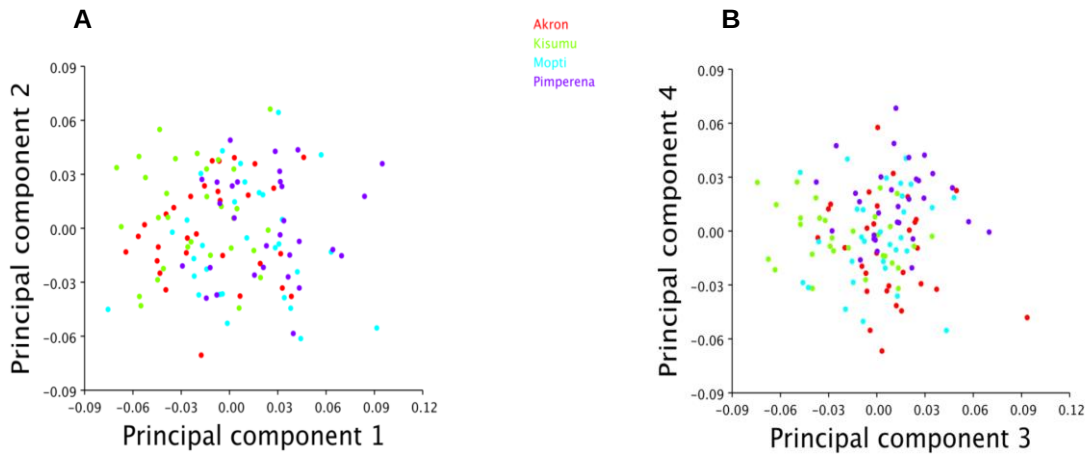

Figure 3.3.15: Scatter plot of Principal component scores for four populations *Anopheles coluzzii* and *Anopheles gambiae* species (Akron, Kisumu, Mopti and Pimperana) A= Principal component 1 VS 2. B= Principal component 3 VS 4.

3.3.7 Egg shape discrimination among populations and species

Canonical variate analysis

The canonical variate analysis for species (*An. coluzzii* and *An. gambiae*) showed a Mahalanobis distance of (2.040), the procrustes distance of (0.023). There was a significant difference ($P< 0.001$) between species. The eigenvalues for CV1 was 1.058 where 100% variance was achieved. Variations at landmark 1,7,12 and 14 was observed between species were observed (Figure 3.3.16) and the scatter plot for species (Figure 3.3.17) showed an overlap between the two species.

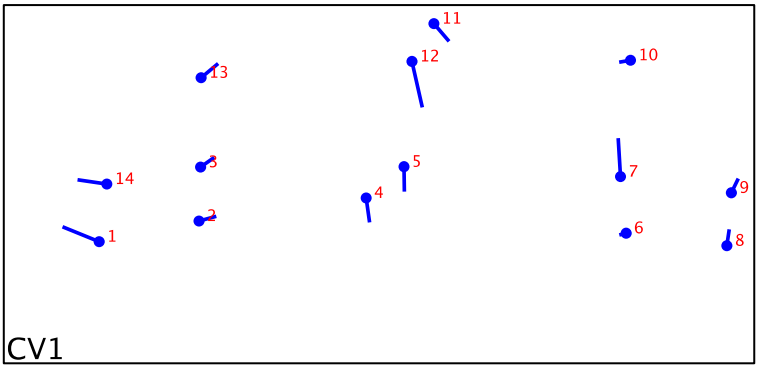

Figure 3.3.16: Canonical Variate analysis showing changes in shape in dorsal morphology between species of *Anopheles coluzzii* and *Anopheles gambiae.*

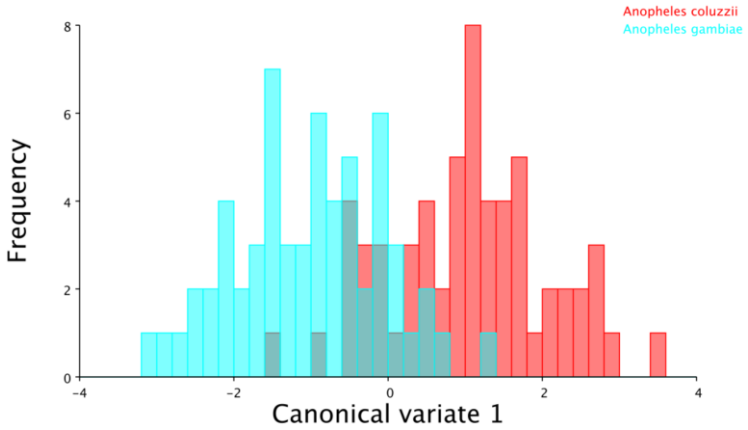

Figure 3.3.17: Scatter plot of CVA for *Anopheles coluzzii* and *Anopheles gambiae*.

Canonical Variate (CV) for the four populations had a variance of 59.370% at CV1, and an eigenvalues of 2.143. CV2 accounts for 31.081% variation, eigenvalues of 1.122 and a cumulative % variance of 90.460. CV3 accounts for 9.54% variance, eigenvalues of 0.344 and a 100% cumulative variance. Shape discriminations for CV1 were at landmarks 1, 7, 8, and 10, CV2 were landmarks 1, 11,12 and landmark 14. CV3 had variations at landmark 1, 2, 4 and 13 (Figure 3.3.18). High canonical coefficient values were observed for landmarks with deviations (Table 3.3.17). Variations at CV1 were on the anterior part of the egg, and CV3 had most variations on the posterior part of the egg.

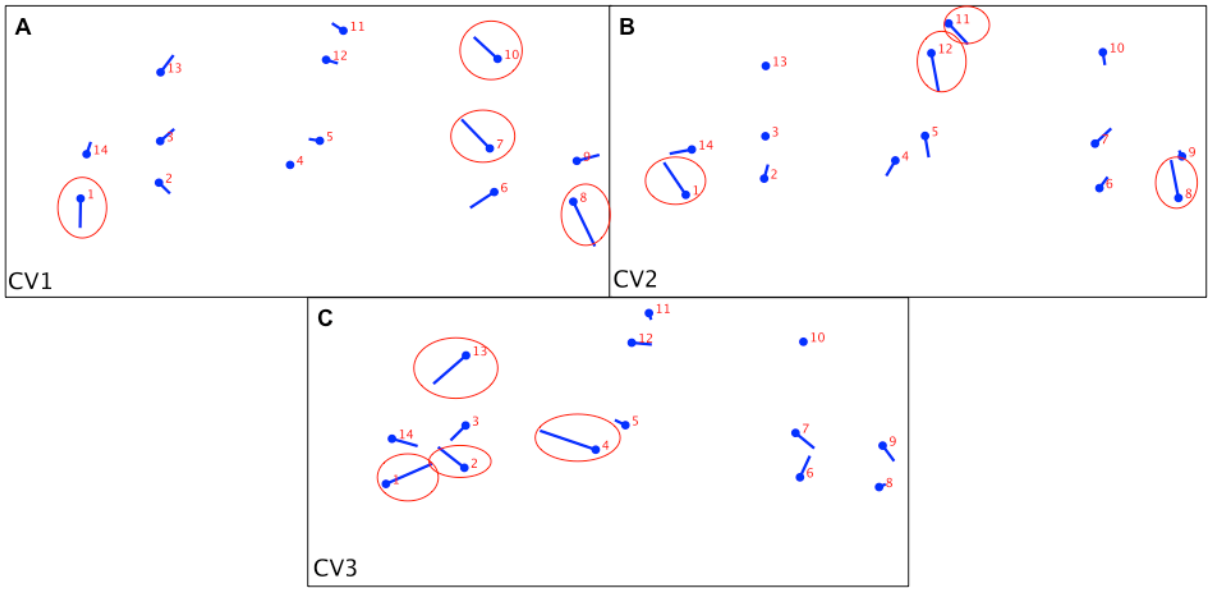

Figure 3.3.18: Canonical Variate analysis showing changes in shape lateral morphology of the different populations of *Anopheles coluzzii* and *Anopheles gambiae*.

Table 3.3.17: Canonical coefficient of the 14 landmarks for the lateral Morphology. CV1, CV2 and CV3 show canonical coefficient for the four different populations (Akron, Kisumu, Mopti and Pimperena) and CV (for species) shows the canonical coefficient for the two species (*Anopheles coluzzii* and *Anopheles gambiae*).

Landmarks	CV1	CV2	CV3	CV1 (For Species)
x1	54.658	-51.372	32.901	-38.861
y1	-14.829	11.244	6.219	3.668
x2	4.362	-15.573	-19.997	-6.641
y2	-0.818	36.821	37.401	21.513
x3	10.124	-1.689	35.898	-9.029
y3	31.007	32.616	-36.886	49.402
x4	-2.556	-3.812	-15.149	0.300
y4	16.900	-60.734	16.227	-53.771
x5	-9.978	-7.491	-2.466	-8.984
y5	-11.698	25.636	-9.011	21.946

x6	16.388	20.751	2.825	22.646
y6	-27.769	2.555	34.480	-16.362
x7	4.246	55.642	32.414	41.299
y7	58.417	47.205	9.255	57.114
x8	22.818	-7.079	14.035	-3.508
y8	-49.614	7.411	-35.559	2.007
x9	-24.544	10.292	-0.668	1.898
y9	11.881	2.594	-4.092	7.137
x10	-32.360	-57.917	-42.364	-48.946
y10	7.972	-20.367	-14.193	-11.529
x11	-11.456	10.241	-2.804	6.467
y11	34.562	15.797	-23.898	31.658
x12	36.166	5.034	26.414	7.691
y12	-70.139	-78.781	18.579	-97.044
x13	-20.523	4.816	-45.743	11.559
y13	10.527	-18.153	16.934	-17.975
x14	-47.347	38.157	-15.294	24.108
y14	3.598	-3.846	-15.456	2.234

*Numbers in red indicate the first four landmarks with the most variation

CV1 and CV2 axis showed clusters of the different populations (Fig 3.3.19A). There was overlap between populations in CV1 and CV3 axis (Fig 3.3.19B). Points were coloured by species to show clustering for CV1 and CV2 axis (Fig 3.3.19C) and CV1 and CV3 axis (Figure 3.3.19D).

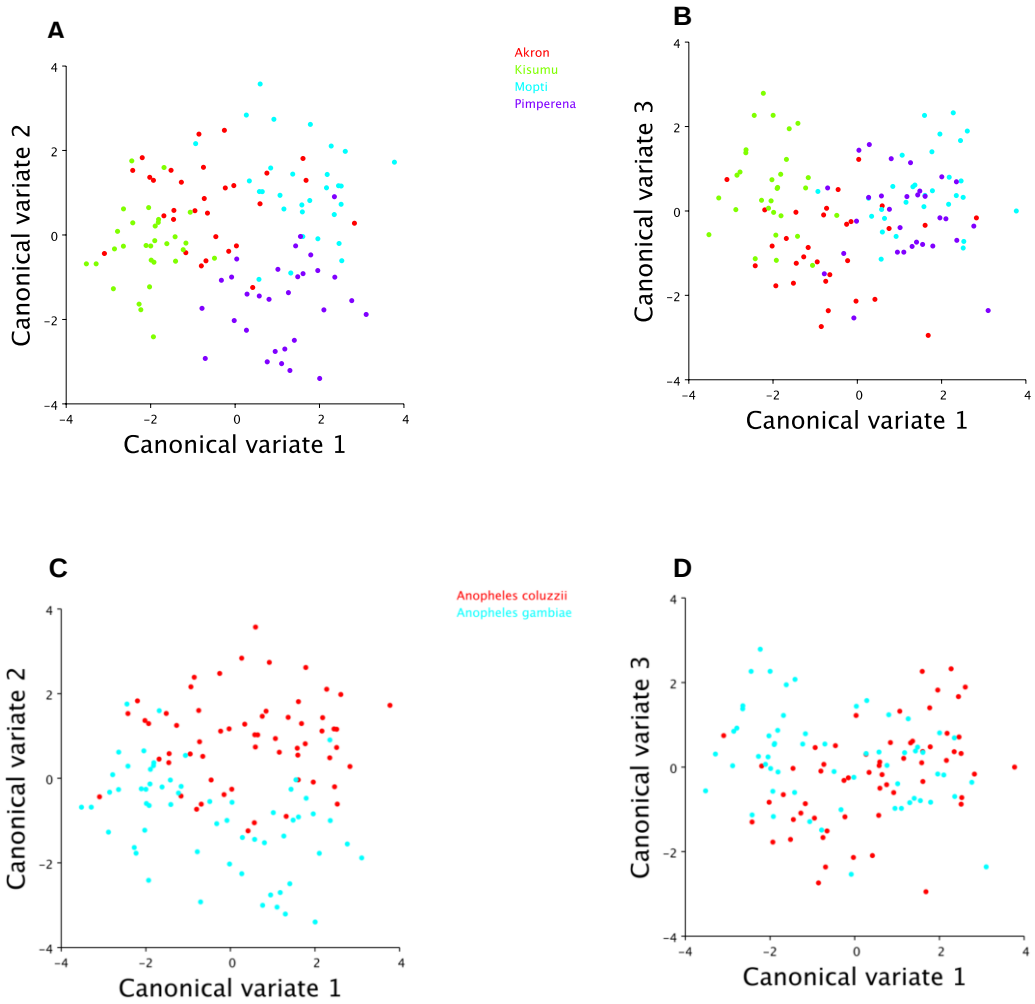

Figure 3.3.19: Morphological space of two canonical variants. A and C show the first two canonical variate components axis. B and D show the first and third canonical variate component axis. A and B = population. C and D = species.

Pairwise analyses for populations were all significantly different. Mopti- Kisumu and Pimperena - Akron were also highly significant (Table 3.3.18).

Table 3.3.18: Canonical variate analysis showing Mahalanobis and Procrustes distances and P-values for the four populations in lateral morphology.

Population	Mahalanobis distances	*P*-values for Mahalanobis distances	Procrustes distances	*P*-values for Procrustes distances
Kisumu - Akron	2.248	<0.001	0.034	0.003
Mopti - Akron	2.602	<0.001	0.029	0.008
Mopti - Kisumu	3.858	<0.001	0.046	<0.001
Pimperena - Akron	2.998	<0.001	0.047	<0.001
Pimperena - Kisumu	3.509	<0.001	0.06	<0.001
Pimperena - Mopti	2.795	<0.001	0.033	0 .001

Discriminant function analysis

The discriminant function comparison between the two species *An. coluzzii* and *An. gambiae* has a Procrustes distance of 0.023, Mahalanobis distance 2.040, *T*-square 124.878, *P*< 0.001. *P*-values for permutation test *T*-square for permutation test was *P*< 0.001.

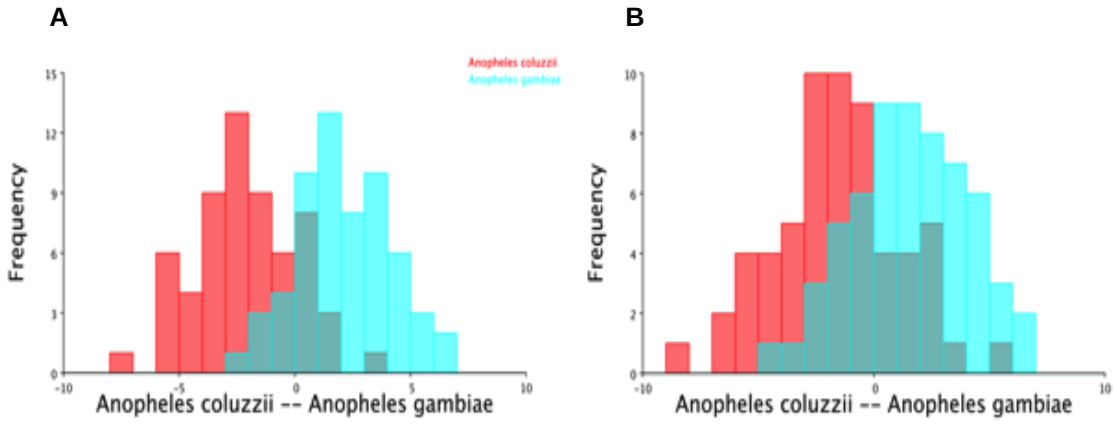

Figure 3.3.20: Lateral morphology discriminant function (A) and cross-validation (B) for pairwise analysis between the two species *Anopheles gambiae* and *Anopheles coluzzii*.

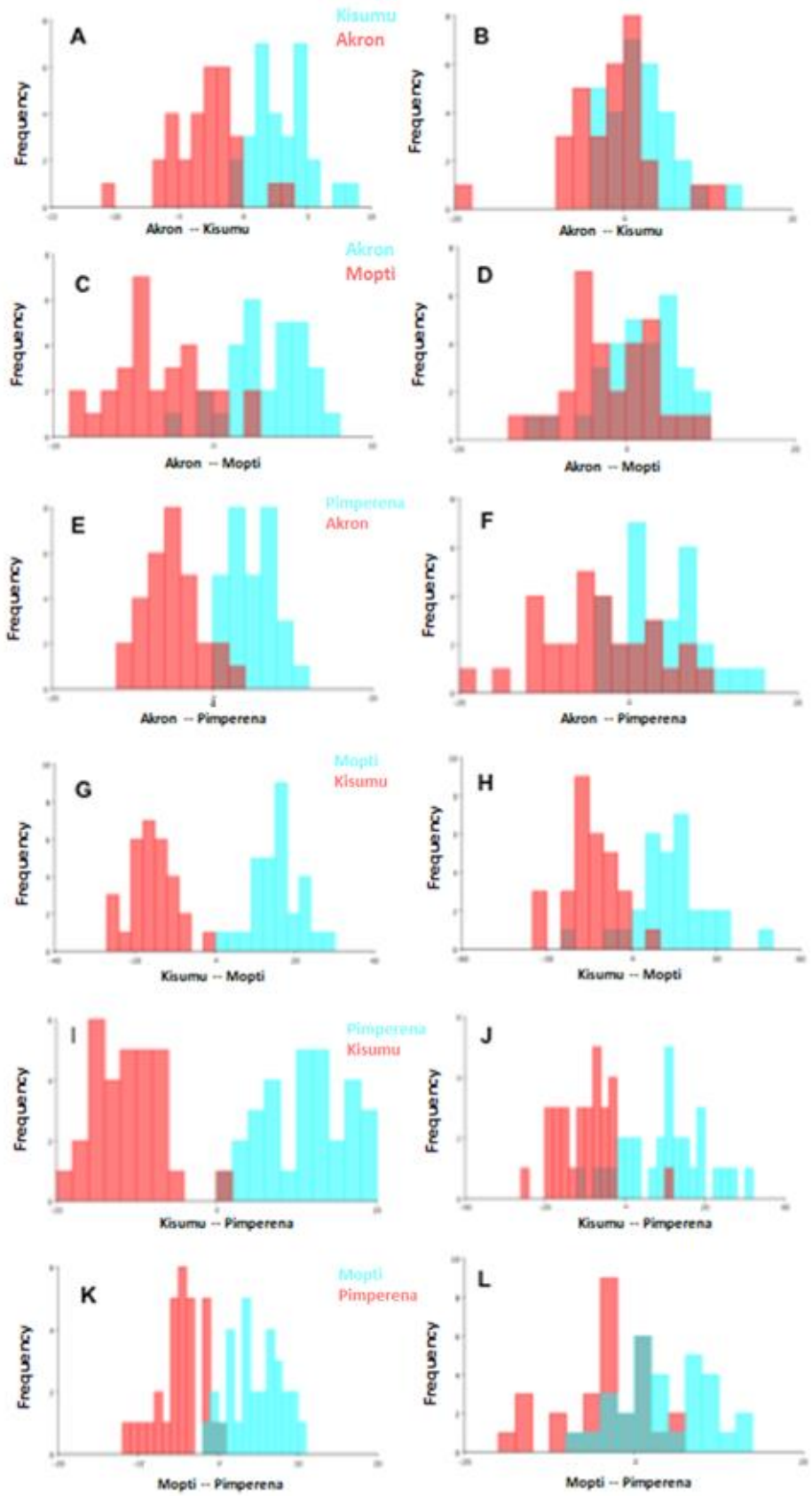

Figure 3.3.21: Lateral morphology discriminant function and cross-validation for pairwise analysis between the four populations (Akron, Kisumu, Mopti and Pimperena) studied. A, C, E, G, I and K are the discriminant analysis graphs which B, D, F, H, J and L are the cross-validation graphs assessing the reliability of the discrimination.

The cross-validation results for the pairwise comparison discriminant function analysis showed misclassification ranging from 3.3% to 40%. Akron - Kisumu had a misclassification of 40% and 30% respectively, Kisumu - Mopti 3.3% and 10% misclassifications indicative of the separation observed between the populations (Figure 3.3.21). High value was observed for the Mahalanobis distance for the Kisumu - Mopti pairwise comparison (Table 3.3.19).

Table 3.3.19: Difference in lateral morphology among the four populations of *Anopheles coluzzii* and *Anopheles gambiae* Mahalanobis and procrustes distances computed from the discriminant function analysis of the lateral morphology. The *P*-values were computed using permutation test (10 000 replications).

Population	**Mahalanobis distances**	***T*-square**	**Procrustes distances**	***P*-values for Procrustes distances**
Akron - Kisumu	2.451	0.016	0.034	<0.001
Akron - Mopti	2.592	0.003	0.029	0.005
Akron - Pimperena	3.173	<0.001	0.046	<0.001
Kisumu - Mopti	5.574	<0.001	0.046	<0.001
Kisumu - Pimperena	4.723	<0.001	0.059	<0.001
Mopti - Pimperena	3.059	0.001	0.033	<0.001
An. coluzzii - An. gambiae	2.040	0.001	0.023	0.001

Surprisingly Kisumu - Pimperena were separated with a misclassification of 3.3% (Figure 3.3.21 I and J). All other pairwise cross-validation between the four populations ranged from 30% to 40%. The cross-validation for the two species *An. coluzzii* and *An. gambiae* was 25%.

The procrustes and mahalanobis distance for all pairwise comparisons were significant $P< 0.001$ except Kisumu --Pimperena $P= 0.001$(Table 3.3.19). The T-square value varied between pairs (Table 3.3.19).

3.4 Discussion

Several reports on egg morphology of the genus *Anopheles* does exist however this is the first study focusing on the two recently diverged sibling species *Anopheles gambiae* and *Anopheles coluzzii*. We looked at possible differences in the egg structure between the sibling species. The eggs of both species were observed to be boat like with the anterior end broader than the posterior end. This observation reflects those described in several Anopheline egg morphology studies (Linley *et al*., 1993; Malhotra *et al*., 2000; Rueda *et al*., 2009; Tyagi *et al*., 2017). No egg morphological differences were noticed between the species by observation.

The results of the traditional morphometrics analysis revealed differences in the mid-deck/dorsal width (MDW) of *An. coluzzii* and *An. gambiae* eggs. Pimperena, a population of *An. gambiae* species was observed to be generally smaller having the minimum lengths and widths (EL, ADW, PDW, EFW and MDW) when compared to the other three populations studied. The floats of both *An. gambiae* and *An. coluzzii* were observed to be about half the width of the entire egg as reported by (Malhotra *et al*., 2000). The mean length of *An. gambiae* was 505.65µm and *An. coluzzii* was 517.46µm, this falls within the range of 398-649µm reported in several studies, *An. nuneztovari* (Linley *et al*., 1996) *An albimanus* (Lounibos *et al*., 1997; Rodriguez *et al*., 1992) and *An. gambiae* complex (Lounibos *et al*., 1999). Egg length and width ratio was 2.83µm for *An. coluzzii* and 2.97µm for *An. gambiae*. A length/width ratio of 2.63µm was observed in *An. fluviatilis* (Sehrawat, 2014). Linley *et al* (1993) did not observe a significant difference in the egg length- width ratio in strains of *An. quadrimaculatus*. The anterior dorsal width (ADW) of the egg was observed to be wider than the posterior end for all four populations studied (Linley *et al*., 1993). The mean ADW and posterior dorsal width (PDW) was 132.44µm and 102.58µm for *An. coluzzii*, 128.143µm and 98.53µm for *An. gambiae*. A similar outcome with wider anterior width (Malhotra *et al*.,

2000; Sehrawat, 2014) has been reported. A possible explanation for the difference observed in egg length and width among the different species could be attributed to ecological adaptations. *An. coluzzii* was observed to have slightly larger lengths and widths than *An. gambiae* eggs. The mopti chromosomal form a strain of *An. coluzzii* is found in large numbers in dry areas and is prevalent during dry seasons (Touré *et al*., 1998, 1994). *An. gambiae* species on the other hand are found in abundance in wet areas and during the rainy season (Touré *et al*., 1998, 1994). Thus, Mopti strain had slightly larger eggs and this may be due to its habitat preference. *An. coluzzii* are exposed to drought conditions thereby larger eggs size could be a strategy for drought resistance as the eggs are likely to survive longer in arid conditions. A somewhat similar observation was made in a study on *Aedes* mosquitoes where humidity conditions were compared to the survival of eggs (Sota *et al*., 1992). *Aedes aegypti*, which has large eggs, survived the longest periods for the different humidity conditions. *Aedes aegypti* is associated with arid region and human disturbed regions (Sota *et al*., 1992).

An. gambiae prefer temporary water bodies such as rain pools and thereby are exposed to limited predation. *An. coluzzii* inhabit permanent water bodies prone to predation. It has been reported that increase in predation increase development rate in *An. coluzzii* (Diabate *et al*., 2005). The difference observed in egg size could be an adaptive feature leading in a faster rate of development into adulthood thereby avoiding aquatic predators.

Geometric morphometrics showed the possibility of distinguishing *An. coluzzii* and *An. gambiae* based on egg characteristics. Eighteen Landmarks for the dorsal morphology and 14 for the lateral morphology were used to investigate certain egg attributes as explained in section 3b. Some of these landmarks were found to be important in separating between the four populations studied. Landmarks 8 and 9, mid-deck/dorsal width (MDW) of the dorsal morphology were different between populations and species in most analysis. These landmarks could be important in distinguishing between the two species.

Principal component analysis (PCA), which checks for variation among all samples, showed variation along some axes. The dorsal morphology revealed clustering among populations along PC1 and PC2 axis (Figure 3.3.4). The first two principal components all showed a graphical representation of divergence along landmark 8 and 9, the mid-deck region (MDW). Other important Landmarks were 7 and 10 in PC 3, which is the egg width and landmark 13 and 16 in PC2 and PC4, which are parts of the floats. Half of the cumulative variation was explained at PC3 with 51.94%, this means that the overall egg shape variation is concentrated in the dimension of that shape space and changes are likely to occur in that direction. A scatter plot of PC1 vs PC2 axis (Figure 3.3.5) showed some clustering among the different populations. The use of principal component analyses has been adopted for the separation of eggs of different members of the genus *Anopheles* (Linley *et al*., 1996, 1993; Rueda *et al*., 2009).

Canonical variate analysis of egg morphology shows similarity to what was observed in the PCA. Differences between populations can be related to the changes on mid-deck region (Landmarks 7 and 8). The mid-deck/dorsal region was statistically significant between populations and species for the dorsal morphology based on the landmarks chosen during the study. These results are in line with that of Tyagi *et al* (2016) were a difference in deck between sibling species of *An. culicifacies* was observed (Tyagi *et al*., 2016). A more recent study also showed differences in deck width in *An. stephensi* variants (Tyagi *et al*., 2017). Landmarks 12, 13, 16 and 17 describing the floats could be important in the separation of the species (Figure 3.3.6; Figure 3.3.8), as a directional change in shape was observed at those landmarks. Eight different Anopheles species were studied and floats were found to differ between the species (Saeung *et al*., 2013). The canonical variate showed clustering of populations on CV1 and CV2 axis, and CV1 and CV3 axis. This further support the idea that the populations can be separated based the egg attributes used for the studies.

Results of the discriminant analysis further showed that populations within the species could be separated using egg morphology. The pairwise discriminant function had an overall 6%-10% misclassification for all pair groups for the dorsal morphology. The test revealed certain pairwise distinctions as being strong. However the cross-validation test overlapped for most pairwise analysis. The cross-validation test misclassifications ranged from 6-50%. For the dorsal morphology, a clear separation was observed between Akron and Pimperena, which suggest a separation between the 2 populations. The low misclassification value of 6.6% on the cross-validation test further supports the results. Akron-Mopti on the hand had 50% misclassifications leading to an overlap (Figure 3.3.11). This finding is consistent with other research where closely related species overlap. Akron and Mopti are populations of the same species, *An. coluzzii.* Overlapping in closely related species was observed in a study on *An. quadrimaculatus* (Linley *et al*., 1993). Most cross-validation pairwise analysis showed overlapping among populations indicating incomplete separation between the populations. Overlapping was not far-fetched as the four populations used in this study are morphologically identical at the other life stages and can only be differentiated through molecular techniques (Scott *et al*., 1993; Townson and Onapa, 1994). The principal component for the lateral morphology showed a range of variation in most axes. Shape changes were observed at landmark 6,7,8 and 9. This is the anterior region where the micropyle is situated. PC2 had changes on landmark 2, 3, 4 and 13, PC3 landmark 1, 4 and 13 most of which are situated at the posterior end of the egg. This shows that variation is concentrated on this space though no striking clustering was observed among populations.

The canonical variate analysis showed separation between populations. Populations of same species were seen to cluster together. Clustering of populations was identified in CV1 and CV2 axis. There was no clear separation in the CV1 and CV3 axis.

Discriminant function analysis showed a negligible variation among species and

populations. Most pairwise groups were classified correctly with a few misclassifications but the cross-validation test for pairwise distinctions overlapped. Kisumu-Mopti was well separated (Figure 3.3.21) and a very low misclassification value of 3.3%-10% on the pairwise cross-validation was observed. This outcome may be explained by the fact that two populations are strains of different species. Kisumu is a strain of *An. gambiae* species and Mopti a strain of *An. coluzzii* species. Surprisingly a strong distinction was also observed for Kisumu-Pimperena with 3.3% and 20% misclassification. This difference was observed in the traditional morphometrics as well. The separation was not expected because the two populations are of the same species i.e. *An. gambiae* therefore expected to be closely related. This could suggest the possibility of intraspecific egg morphology variation as reported in a study on species of *Ae. aegypti* where variations were observed in eggs belonging to same species though obtained from different locations (Faull and Williams, 2016). The other pairwise distinctions showed misclassification of 23.3% to 40% that translated into overlapping in the cross-validation test. The result for the lateral morphology still does not give a clear distinction between eggs of *An. coluzzii* and *An. gambiae***.**

In conclusion, there appears to be more differences between populations than between species and a considerable amount of overlapping was observed between the populations studied suggesting that the eggs of the two species *An. coluzzii* and *An. gambiae* are similar. Similarity was also observed in an earlier study describing the egg shape of six members of the *An. gambiae* complex where no species differentiation in morphological character was found (Lounibos *et al*., 1999). However, there appears to be apparent morphological variation among the four populations studied. The findings therefore highlight some morphological features that could be used to distinguish the two species. The findings are not sufficient to suggest a difference in the egg morphology between the two species. Landmark-based analysis can be important in revealing morphological variations between the species.

Further studies using the scanning electron microscope can be used to investigate possible differences in features such as micropyle, float numbers etc. A study on, *Aedes aegypti* and *Aedes albopictus* eggs showed *Ae. aegypti* had wider eggs compared to *Ae. albopictus*. A difference was also observed in the egg exochorionic networks of the two species. *Aedes albopictus* had a more interwoven wall like exochorionic networks though smaller in size. The interwoven network probably aid in its ability to withstand desiccation in containers (Suman *et al*., 2011). Therefore further studies could reveal similar outcomes. Though (Saeung *et al*., 2014) compared egg characteristics of eight anopheline species, *An. argyropus, An. crawfordi, An. nigerrimus, An. nitidus, An. paraliae, An. peditaeniatu, An. pursati* and *An. sinensis* using the scanning microscope. They discovered that their description of the eggs was similar to that given by Ried in a similar study in 1968 using the light microscope (Saeung *et al*., 2014).

If the eggs do not differ morphologically, they could react differently at different conditions and this could be used in distinguishing between *Anopheles coluzzii* and *An. gambiae* sibling species. Factors influenced by ecological adaptation such as egg hatchability, and survival, oviposition site preferences and predation avoidance behaviour in the sibling species can be studied to reveal differences between the sibling species.

Chapter 4 Differential gene expression in populations of *Anopheles gambiae* and *Anopheles coluzzii*

4.1 Introduction

The success of any malaria vector control program such as genetically modified sterility-inducing males, Wolbachia-carrying males and males with gene drive effector genes are reliant on male mosquito mating behaviour. Unfortunately, the success of these programmes are daunted by inadequate knowledge of the vectors (Diabate and Tripet, 2015). Insect releases depend on the ability of males to compete favorably and mate with females in the wild (Harris *et al*., 2011) however, laboratory colonization, and genetic manipulations affect the ability of males to compete and mate (Ekechukwu *et al*., 2015). This invariably affects the success of vector control programmes dependant on vectors ability to mate.

Mating behavior in mosquitoes is complex involving behavioral, physiological and developmental factors influenced by genetic and ecological influences. Our understanding of this mechanism is vital to malaria control.

Assortative mating is considered a driving force of divergence between members of the *Anopheles gambiae* complex; other factors are ecological adaptation and genetic differentiation (Figure 4.1.1). These factors are evident in the recently diverged members of the *An. gambiae* species complex *Anopheles coluzzii* and *Anopheles gambiae*. They have different breeding site preferences (Cassone *et al*., 2008). The existence of strong assortative mating between the sibling species (Tripet *et al*., 2001) and limited differentiation in certain regions of the species genome is termed islands of speciation (Cassone *et al*., 2008).

Postzygotic reproductive isolation has not been observed between *An. coluzzii* and *An. gambiae* (Lanzaro *et al*., 2004) supporting assortative mating as a mechanism for divergence. Mate recognition cues such as, swarming (Diabaté *et al*., 2006; Sawadogo *et al*., 2013) wingbeat, flight tunes, (Cator *et al*., 2010; Pennetier *et al*., 2010) have been reported but to

date the mechanism for assortative mating remains unclear (Aboagye-Antwi *et al*., 2015). No morphological difference has been found between *An. gambiae* and *An. coluzzii* nor has any phenotypic difference been understood (Lehmann *et al*., 2008). This has heightened interest in discovering genomic signatures that can unravel the genetic and ecological conditions aiding the emergence and divergence of these species (Aboagye-Antwi *et al*., 2015).

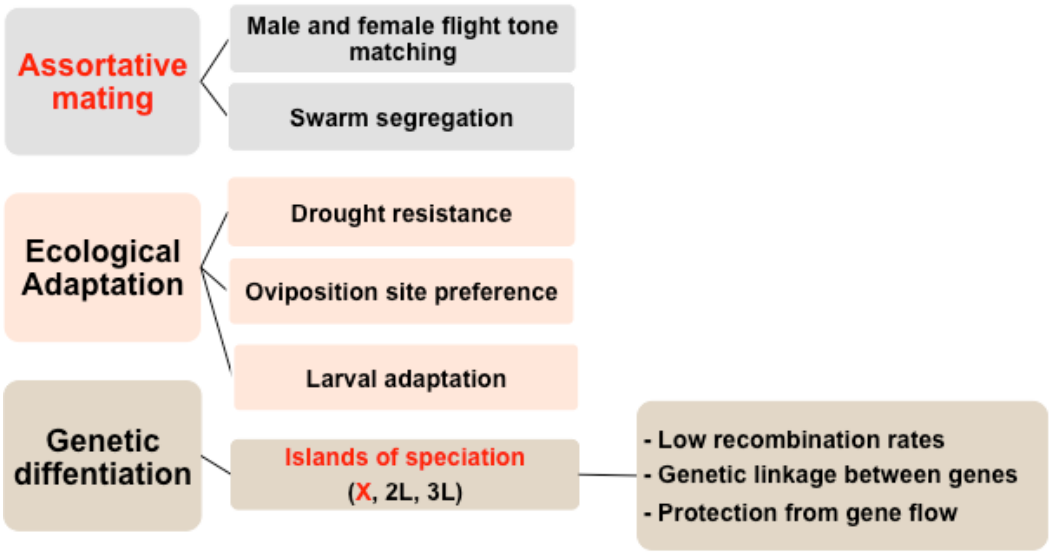

Figure 4.1.1: Mechanism of divergence and speciation in *An. coluzzii* and *An. gambiae* sibling species. Assortative mating as a mechanism of divergence employs mate recognition cues such as swarm segregation and flight tones. Differences in oviposition site preferences, ability to avoid predation and drought resistance are considered traits for ecological adaptation. Those regions in the genome of the two sibling species that are different are considered to harbor genes responsible for speciation. The regions are called the islands of speciation.

Genome scans focusing on differences between *An. coluzzii* and *An. gambiae* have revealed regions involved with natural selection but not much is known regarding the functional importance of their divergence (Cassone *et al*., 2014). Turner *et al*, (2005) discovered three significant genetically differentiated regions totaling about 2.8 Mb in *An. coluzzii* and *An. gambiae*. Two of these regions on chromosome 2L and X contain 50 genes and 12 predictable genes, respectively. Sequence comparison of genes within these loci

revealed fixed differences and unique polymorphisms between the species (Turner *et al*., 2005). The third region located on chromosome 2R contained only five predictable genes (Turner *et al*., 2005), however, differentiation and reduced variation between the two sibling species was not complete in all populations (Turner and Hahn, 2007) hence, not considered a strong contributor to reproductive isolation.. The two regions on chromosome 2L and X are considered as Speciation Island because they remain different between *An. coluzzii* and *An. gambiae* regardless of considerable gene flow and are therefore thought to contain genes responsible for reproductive isolation (Turner *et al*., 2005). Using microarray-based divergence mapping, a third region was discovered on chromosome 3 (White *et al*., 2010). Therefore, the presumed islands of speciation are the 3-pericentromeric islands of divergence on the X, 2L, 3L chromosomes (Turner *et al*., 2005; White *et al*., 2010) and some other small Islands situated in the area of inversion breakpoints (Aboagye-Antwi *et al*., 2015). A recent study estimated the X island of speciation to be over 6MB-long starting from position ~18.1 to 24.2Mb containing about 153 genes (Aboagye-Antwi *et al*., 2015).

In an attempt to identify genes responsible for assortative mating between *An. coluzzii* and *An. gambiae*, Tripet, and colleagues (Aboagye-Antwi *et al*., 2015) investigated how the divergence islands influence sympatric speciation. This was done using an experimental functional genomics approach. It was hypothesized that the largest presumed speciation island, which is on the X chromosome, was responsible for protecting assortative mating genes despite on-going gene flow. This is because the island combines pericentromeric recombination suppression with the hemizygosity and decreased recombination typically associated with the X chromosome (Aboagye-Antwi *et al*., 2015).

The *An. gambiae* (S-form) X-linked Island of speciation was introgressed into the *An. coluzzii* (M-form) genetic background. This led to the creation of recombinant strains (RbSS) having the same genetic background as the *An. coluzzii* type recombinant (RbMM) but

differing only on the X-chromosome islands of speciation. The recombinants had 96.88% genetic similarity with the Mopti M-form strain. A series of assortative mating experiments to test the association of the X-island molecular type with mating preferences of recombinants and parental strains was carried out (Figure 4.1.2).

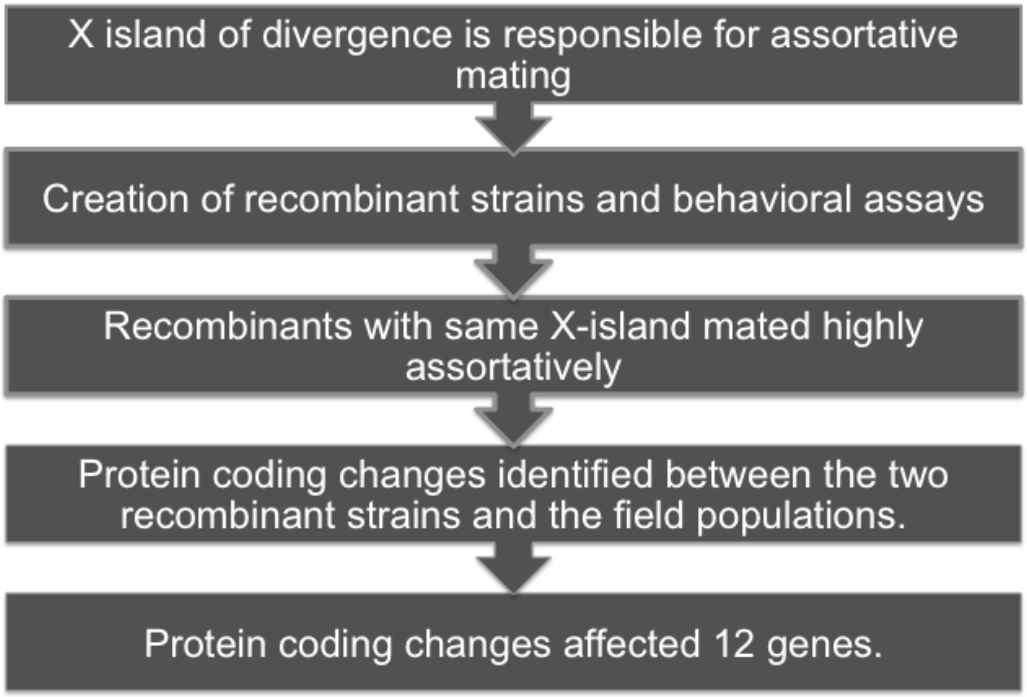

Figure 4.1.2: Flow chart showing the summary of the Experimental swap of *Anopheles gambiae's* assortative mating preference experiment demonstrating key role of the X-chromosome divergence island in incipient sympatric speciation (**Adapted from** Aboagye-Antwi *et al.,* 2015).

The mating choice preference of females and males was tested between mates with the same and with a different type of X-Island. Recombinant RbMM females were seen to mate highly assortatively while RbSS recombinant females mated entirely with males with the same X-type Island (Aboagye-Antwi *et al*., 2015). These results showed a close association between pre-mating isolation genes and the X-island thereby supporting the hypothesis that pericentric regions can protect genes responsible for pre-mating isolation leading to sympatric

speciation. It also shows that hemizygosity and low recombination rates of sex chromosomes encourage the build-up of pre and postmating isolation genes. The results suggest that the low-recombining pericentromeric X-island enables these incipient species to maintain their genetic integrity in parts of Africa where they are sympatric.

A comparison of sympatric populations from Ghana and the recombinant strains was done to estimate the size of the island, and the X-island of speciation was estimated to be 6Mb long extending from positions approximately 18.1 to 24.2Mb. Full-genome sequencing of the RbSS, RbMM and Mopti strains was carried out (Aboagye-Antwi *et al*., 2015).

Protein-coding differences was assessed between the two recombinants and between the two sympatric field populations of *An. coluzzii* and *An. gambiae* from Southern Ghana. Some interesting protein-coding changes observed between the two recombinant strains were fixed or nearly fixed in the sympatric populations of *An. coluzzii* and *An. gambiae* from Ghana (Table 4.1.1) (Aboagye-Antwi *et al*., 2015).

Table 4.1.1: Protein coding changes identified between the RbMM and RbSS strains confirmed in sympatric field populations of *An. coluzzii* and *An. gambiae* **Source**: (Aboagye-Antwi *et al*., 2015).

Reference position	M-form allele	S-form allele	Codon (protein change)	Gene (putative function)
18314527	G	A	gGg/gAg (G74E)	AGAP013136
19052774	T	G	aAc/aCc (N25T)	AGAP001002 (Toll protein)
19114172	C	G	gGg/gCg (G232A)	AGAP001009
19114460	A	C	Tgg/Ggg (W160G)	AGAP001009
19114646	G	C	Ctg/Gtg (L98V)	AGAP001009
19116606	T	C	cAc/cGc (H197R)	AGAP013526
19636043	A	G	aAg/aGg (K272R)	AGAP001022 (gastrin/cholecystokinin receptor)
19636265	C	T	gCc/gTc (A346V)	AGAP001022
19637489	T	C	Tcg/Ccg (S355P)	AGAP001022
19714995	C	T	cGc/cAc (R491H)	AGAP001025 (protein msta)
19815325	A	T	Acg/Tcg (T339S)	AGAP001033 (mab-21 like protein)
20955148	G	A	aCg/aTg (T281M)	AGAP001050 (chondroitin polymerizing factor)
21093789	A	C	Ttt/Gtt (F288V)	AGAP001052 (ubiquitin carboxyl-terminal hydrolase)
22104989	T	C	Aat/Gat (N1296D)	AGAP001061
22159522	T	C	Atg/Gtg (M514V)	AGAP001061
22751398	A	G	Act/Gct (T252A)	AGAP001073
23799338	T	A	gaA/gaT (E879D)	AGAP013341
23799431	T	G	gaA/gaC (E848D)	AGAP013341
23799525	T	G	aAt/aCt (N817T)	AGAP013341
23799541	T	G	Att/Ctt (I812L)	AGAP013341

Twenty differences identified between the RbMM and RbSS recombinant strains were also fixed or nearly-fixed (freq >0.95) in sympatric field populations. These were distributed over 12 genes located within the ~ 6MB pericentromeric island of speciation.

It was discovered that 114 out of 117 coding differences observed in the recombinants were fixed or nearly fixed (>0.95) for the M-form in the field populations of *An. coluzzii*, 61 out of the 114 coding differences found in *An. coluzzii* were also found to be fixed or nearly fixed for the S-type allele in *An. gambiae* at a frequencies of >0.8. Twenty of the fixed or nearly fixed (freq >0.95) coding changes were conserved between the two sibling species and these changes affected 12 genes in the cryptic taxa. Six out the 12 genes were identified to perform biological functions such as development and immunity (AGAP001002) and neural and sensory organ development (AGAP001033) but the function of the others is not yet known (Aboagye-Antwi *et al*., 2015).

Many genes code for polypeptide chains of proteins, this is through a transcription of RNA into mRNA. The polypeptides are produced by translation of mRNA and this is done in a triplet genetic code (codon) that specifies the amino acid attached to the growing chain.

Alleles usually have different nucleotide sequence. If a nucleotide substitution occurs in the first two position of a codon, it results in amino acid replacement (Hartl and Clark, 1997). The nucleotide sequence of a gene can be used to deduce the protein sequence that the gene encodes for (Old and Primose, 1994). There is no role for predicting the proteins three-dimensional structure from the amino sequence nor does amino acid sequence on its own show gene function. Amino acid sequence can be compared with that of a better-characterized protein and a high degree of homology suggest similarity in function. Two sequences can be compared using central storage banks from which data can be obtained (Old *et al*., 1994). The genes found to differ between the recombinant strains and also between the sympatric fields populations could have an impact on mating behaviour in these sibling species. The genes could be contributors to ecological speciation such as larval adaptation in the sibling species (Aboagye-Antwi *et al*., 2015; Diabate *et al*., 2005). Their findings, therefore, suggest the possibility of the genes responsible for assortative mating being on the X-island of speciation. Further investigation into the possible role of the speciation genes discovered in the studies could lead the discovery of phenotypic differences between the species.

Aside from fixed coding differences that can influence the phenotype of an organism through changes in protein sequence, changes in the amount of protein produced through translational modification also play an important role in the phenotypic divergence between the two sibling species. Reverse ecology e.g. gene expression studies can be used to measure differences between species using the hypothesis that the difference in transcript can be a representation of some phenotypic traits (Li *et al.,* 2008). Genes are activated or inactivated to regulate the amounts of RNA and proteins produced for a specific signal. This difference in expression is termed gene expression (Reddy Palli, 2012). Different traits in insects are as a result of how differently the genes responsible for the traits are expressed hence transcriptome analysis experiments and gene expression analysis are important and carried out in many

studies (Oppenheim *et al*., 2015). The ability to study gene expression in different species under different conditions alongside studying the different developmental stages has led to the knowledge of evolution and molecular basis of insect development, physiology and behavior (Oppenheim *et al*., 2015).

As far as transcriptomics involving *An. coluzzii* and *An. gambiae* are concerned, Cassone *et al* (2008) carried out the first study investigating gene expression profiles between *An. coluzzii* (M-form) and *An. gambiae* (S-form). They examined the different developmental stages, larvae, Virgin females and gravid females using whole-genome oligonucleotide microarrays. Approximately 1-2% of all the genes studied where expressed differently between the two species. A 19% (n=164) difference in the transcript was found in virgin females (Cassone *et al*., 2008). At the late larval stage, 61% of genes were overexpressed in *An. gambiae* in comparison with *An. coluzzii*, there was a 48% overexpression in virgin females and 42% in gravid females when compared with *An. coluzzii*. In virgin females, differentially expressed genes between the two species were those containing nucleic acid-binding proteins for transcription and splicing. Genes that could be involved in olfaction and might have a role in mate recognition were also identified. The majority of the genes were expressed outside of the X-island of speciation (Cassone *et al*., 2008).

Many studies focusing on differential gene expression have revealed genes with stage-specific or sex-specific differences associated to mating in species. Differences in sexual traits between male and females are due to the different selective pressures (Lande, 1980). The most rapidly evolving loci in the genome are those genes that are responsible for mating and reproduction, they show positive selection (Haerty *et al*., 2007; Singh and Artieri, 2010). Due to their adaptive significance, proteins involved in reproduction are the subjects of many evolutionary studies (Oppenheim *et al*., 2015). Genome-wide transcriptional analysis was used to assess sex-regulated genes at different developmental stages of *An. gambiae*. The

genes were categorized as male or female bias based on male/female expression ratio. An increase in sex-bias genes was observed at the adult stage. The results showed that female-biased genes were abundant at the adult stage while male-biased genes were in abundance at the larval and pupae stage. More variable expression profile was reported in males suggesting its role in the differences in expression observed between males and females in the different life stages. The function of previously unknown genes on the testis and ovaries of *An. gambiae* were identified (Magnusson *et al*., 2011). In another study, difference in gene expression between sexes in *An. gambiae* was also examined using Affymetrix GeneChip microarrays. A greater than fourfold difference in 10% of the genome were expressed differently between males and females. In the study, 71% of the genes studied were female biased (Hahn and Lanzaro, 2005). Global gene expression in *An. gambiae* revealed a mild female bias through a gene-by-gene comparison between males and females. In larvae, ratios of X chromosome expression levels were mostly female bias when male-female ratios were compared (Rose *et al*., 2016).

Global transcriptional profiles were compared between field and laboratory populations of *An. coluzzii* and *An. gambiae* using microarrays. *An. coluzzii* had an overexpression of genes involved in detoxification and immunity compared to *An. gambiae*. Overexpression of those genes could be due to the difference in larval habitat preference as *An. coluzzii* larvae inhabit mostly polluted environments while *An. gambiae* are found in temporary pools and puddles. The difference in expression was said to be an adaptive response to environmental conditions (Cassone *et al*., 2014). The 2nd instar larvae were observed to contain 46 differentially expressed genes between the two species of which 89% was overexpressed in *An. coluzzii*. At the 4th instar, 71% of the immune-related genes were upregulated in *An. coluzzii* (Cassone *et al*., 2014).

An. gambiae s.s. and *An. quadriannulatus* have different host preference. The later feeds mostly on vertebrates and mammalian host while *An. gambiae* prefers human blood. A study was carried out on the different host seeking phenotypes, it was hypothesized that the transcriptome profiles of the antennae are responsible for host-seeking behaviour because the behaviour is driven by the sensory modality of olfactory (Rinker *et al*., 2013). The antennal mRNA of nonblood-fed females was sequenced and observed for quantitative and qualitative differences in the chemosensory gene. There were high rates of sequence polymorphism in the transcriptomes of both *An. gambiae* and *An. quadriannulatus* potentially important in differentiating between the species. The analysis revealed that the *An. gambiae* s.s behaviour as anthropophagic is reflected in the different distribution of olfactory receptors in its antenna (Rinker *et al*., 2013). The different studies on *An. gambiae* discussed above have led to uncovering genes associated with different phenotypic and behavioural characteristics of the species. This shows the importance of gene expression studies.

Here we decided to take advantage of gene expression comparisons as a prerequisite to identifying putative assortative mating genes in *An. coluzzii* and *An. gambiae.* Differential expression of genes located on the X-Island of speciation with fixed or frequent coding differences between *An. coluzzii* and *An. gambiae* as reported in (Aboagye-Antwi *et al*., 2015) was investigated. The analysis involved stage-specific expression profiling at different developmental stages of the mosquitos' life with an interest in differential expression in young adult mosquitoes. Virgin females have been reported to mate more assortatively compared to males (Aboagye-Antwi *et al*., 2015) therefore differential expression of genes in males could allow identification of putative assortative mating genes involved in certain mate recognition cues. Identification of differentially expressed genes may not only reveal the genetic basis of assortative mating and ecological adaptation but also lead to a better understanding of speciation. This knowledge would have a positive impact on the production

and release of males that effectively mate with the intended target populations in male release vector control programs. It could also lead to the discovery of new vector control tools taking advantage of mating cues such as visual, auditory and olfactory**.**

4.2 Methods

4.2.1 Mosquito colony

The experiment was carried out using four different populations of *Anopheles coluzzii* (Akron and Mopti), and *Anopheles gambiae* (Kisumu and Pimperena) species. The mosquitoes were maintained in the Mason insectary Centre for parasitology and entomology Keele University insectary under ideal conditions as described in chapter 2.

4.2.2 Experiment set up

Gene expression patterns were investigated at different developmental stages of the mosquitos' life cycle. Stages considered were larvae, pupae, and adults. Fourth instar unsexed larvae, sexed pupae (male and females); unmated adults (Males and females) and mated females (Figure 4.2.1) were used in this experiment.

Gravid females were provided oviposition pots three days after blood feeding. Oviposition pots were removed from the cages after two days and two hundred first instar larvae put in trays. This was done in triplicate for each population. Individual trays containing 200 larvae were considered as biological replicates. Three biological replicates were used for each population in this study. Larvae were given the standard feeding regime (See general methods chapter 2 section 2.2.1).

4.2.3 RNA isolation

Total RNA was extracted from groups of 5-20 individual mosquitoes depending on the stage. Details are provided in Table 2.2.

Larvae

Total RNA was extracted from 20 larvae. Larvae were picked from tray using a pasteur pipette. This was put on a filter paper placed in a petri dish. The filter paper soaks up water collected while picking out the larvae, this was to prevent escape of larvae. Forceps were used to quickly transfer larvae from the filter paper into a 1.5 ml eppendorf tubes containing appropriate volume TRIzol (Table 2.5.1). Larvae were killed immediately by homogenization using a pestle. Those larvae left in the trays were fed appropriately so as to evolve into good-sized pupae.

Pupae

Pupae were picked and separated into males and females. Total RNA was obtained from ten males and from 10 females. Those that had just pupated were used for the experiment. Those pupae already separated according to sex but not used for extractions were put in emergence cages so as to emerge into virgin adults. Some pupae were picked from the tray and put in another emergence cages without separation to constitute the mated group.

Virgin females and virgin males

The separated pupae (males and females) were put into different properly labelled emergence cages. This was to ensure that the mosquitoes used for the unmated group were virgins. Ten percent glucose solution was provided and wet cotton pad placed on the cages. Cages were checked early next morning to confirm if sexing was devoid of error. If an error was observed, by having a male in the female cage or vice versa, a mouth aspirator was used to pick them out of the cage and discarded. Total RNA was obtained from ten 1-day-old virgin male and female pupae for each replicate.

Mated females

Emergence cages were prepared and pupae picked directly from trays were put in cups and placed in the cages without sexing. This was to produce mated females. After 5 days, females were removed from the cage (three replicates per population). They were dissected and spermatheca checked for presence or absence of sperm. Unmated females were discarded while mated females were put in a 1.5ml eppendorf tubes containing TRIzol for extraction. Ten females were used for each replicate.

RNA isolation was done as described in chapter 2 (Table 2.2 of general methods). Samples were homogenised in TRIzol (Invitrogen, USA) and total RNA was extracted using RNeasy Mini Kit (Qiagen). The number of Adults and pupae used were split into 2 two individual tubes, therefore 5 adult mosquitoes were homogenised in 300µl of TRIzol after which the samples were combined in a single tube to make a total of 10 adult mosquitoes in 600µl of TRIzol. Ten pupae were split into equal numbers and homogenized in two separate tubes containing 200µl of TRIzol. This was combined and 100µl of TRIzol was added to the homogenate resulting in a total of 10 pupae in 500µl to TRIzol. Twenty larvae were homogenised in 100µl of TRIzol after which 400µl of TRIzol was added. The concentration, integrity and quantity of RNA were checked using a ND-1000 Spectrophotometer (NanoDrop Technologies). Wavelengths of 230nm, 260nm, and 280nm with values ranging from 1.8 to 2.2 were used. Samples were preserved immediately after extraction at -80°C until needed.

4.2.4 Quantitative RT-PCR validation

cDNA synthesis

The Quantitect Reverse Transcription Kit (Qiagen) was used for cDNA synthesis of 500-1µg of total RNA for each sample. The procedure was carried out following the manufacturers manual. RNA template was thawed on ice. gDNA Wipeout Buffer, Quantiscript Reverse

Transcriptase, Quantiscript reverse transcriptase Buffer, Primer Mix and RNase-free water were thawed at room temperature. A 14µl reaction comprising of 2µl gDNA wipeout buffer, template RNA, and RNase-free water was carried out to remove genomic DNA. The volume of template RNA used was dependant on the RNA concentration and therefore RNAse free water was used to scale up the volume. The sample was put in the thermocycler and incubated for 7min at 42˚C and 5 minutes at 4˚C.

Reverse-transcription reaction

After the removal of genomic DNA, the reverse-transcription (RT) master mix was prepared on ice. A 20µl reaction comprising of 1µl RT primer mix, 1µl Quantiscript Reverse Transcriptase, 4µl Quantiscript RT buffer 5x and 14µl Template RNA from the initial step was prepared. This was incubated in a thermocycler for 15 minutes at 42ºC and 3 minutes at 95ºC to inactivate Quantiscript Reverse Transcriptase. cDNA was Stored at -20ºC until use.

Quantitative RT-PCR

The cDNA was diluted in a 7:1 dilution ratio. This was to ensure enough for the number of reactions required per sample. For a 10µl reaction, 5µl of SYBR green was put in a 96 well reaction plate (Applied Biosystem), 1µl of the cDNA, 2µl of 10pmol forward primer and 2µl of the reverse primer for individual targets were added to the mix. The reaction was performed for three biological replicates consisting of two technical duplicates. Two sets of independent primers were used for all target genes to increase the reliability of the data.(Appendix 1).

The quantitative RT-PCR was performed using the stepOne plus real-time PCR machine. The cycling conditions were pre-initial heat activation at 95˚C for 5 minutes followed by 40 cycles of denaturation at 95˚C for 10 seconds, then annealing and extension at

60˚C for the 30 seconds for an extension. The melting temperature was at 95˚C for 15 seconds, 60˚C for 1 minute and then 95˚C for 15 seconds. Gene expression level was then quantified. The expression levels were measured for target genes and controls.

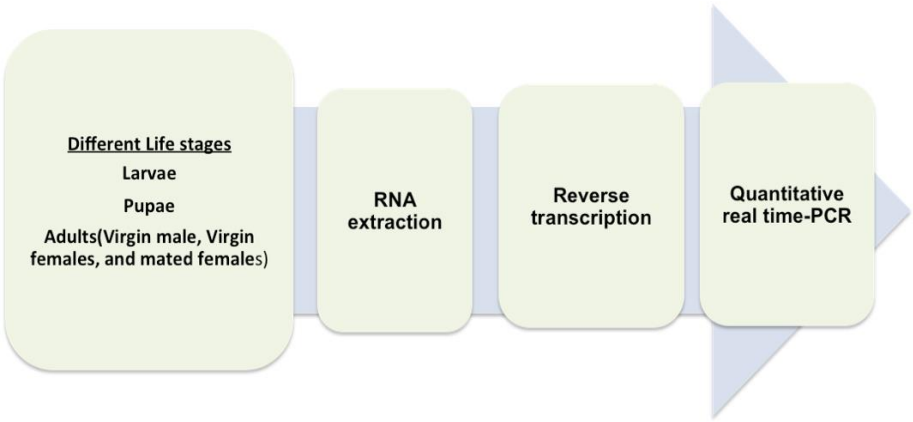

Figure 4.2.1: Illustration of the different steps taken for the gene expression studies.

4.2.5 Target gene selection and Primer design

The target genes were selected based on results from Aboagye-Antwi *et al.*, (2015). Genes that were found to have protein coding changes between the recombinants RBMM and RBSS that were fixed (>0.95) or frequent (0.8-0.95) in field populations of *An. coluzzii* and *An. gambiae* from Ghana were considered as targets. Twenty-seven genes and 30 transcripts were studied (Appendix 1). Three housekeeping genes were selected based on previously published literature on *An. gambiae*. The stability of these genes was tested among samples. Two out of the three housekeeping genes were used 40S ribosomal protein S7 (AGAP010592) (Cook and Sinkins, 2010) and 60S ribosomal protein L1 (AGAP004422) as positive controls to normalize the CT (cycle threshold). Two independent Primer sets were designed and optimized for each target and the two controls. The two sets of independent primers were used to enhance reliability of the data. Primers were designed using PrimerBLAST (www.ncbi.nlm.nih.go) and Oligo Prefect Designer tools. Primers are shown in (Appendix 1).

Vector base (Vectorbase.org) was used to check for gene ontology (GO) based on homology to previously annotated *An. gambiae* genes (Tables 4.3.5).

4.2.6 Statistical analysis

Both controls and target genes were measured based on three biological replicates and two technical replicates; the threshold cycle (CT) values were obtained. Data were normalized and JMP statistical discovery software (SAS) was used to analyse data.

Outliers

Primer mismatches were identified as outliers. A scatter plot was plotted for CT values of the two independent primers sets CT1 versus CT2. A line of 95% confidence limit was plotted and targets that were outside the confidence limits were considered as outliers. Where one CT for a target was an outlier, the second primer was used since two independent primers were used for each target. Where both CT values for a target were outliers, samples were re-ran. Reruns were carried out for targets where both primer sets failed or where the endogenous controls failed to be transcribed.

Normalizing of CT value

The CT values were normalized to control for any effects of technical error. The method for analysing relative gene expression data using real-time quantitative PCR as explained in (Livak and Schmittgen, 2001) was used with some modification. Three biological and two technical replicates were used therefore the average CT value for each target was calculated. The mean CT value of primer 1 and 2 for each target gene and for the Endogenous reference genes (Controls) was calculated:

(CT value Primer 1 + CT value Primer 2)/2 = Mean CT value for target gene.

Two (2) controls (endogenous reference genes) were used and the means was calculated:

(CT value of control 1 + CT value of control 2)/2 = Control CT value for control.

The normalized value of the target gene was calculated by deducting the mean control CT value from corresponding mean CT value of the target gene:

Mean CT value of target – Mean CT value control = Normalized CT value of the target gene

This was done for each technical replicate and the means of normalized CT values between technical replicates was calculated to give the CT values for biological replicates (ΔCT). Some CT values were in the negative and 1 was added to all values for the gene to give a positive value.

4.3 Results

Differential gene expression was examined between *An. coluzzii* and *An. gambiae* using two strains from each species. The normalized CT values were used for all analysis. Differential expression was also compared between life stages (larvae, male and female pupae, virgin males and females and mated females). Data collected were subjected to a multivariate analysis (General linear model) using 3 models. The best model for each transcript was adopted (Table 4.3.1). Expression patterns were compared between species, *An. coluzzii* and *An. gambiae* and between populations within species (Akron, Kisumu, Mopti and Pimperena). JMP was used for the statistical analysis.

Table 4.3.1: Test of interaction effects of putative speciation genes with fixed of frequent differences in *An. coluzzii* and *An. gambiae.*

Transcript	Chromosome location	Fixed (>0.95) or Frequent (>0.8)	Species		Population [species]§		Stage		Stage* Species		Population* Stage [Species]§	
			df =1		*df* = 2		*df* = 5		*df* = 5		*df* = 10	
			F-Ratio	Prob	*F-Ratio*	Prob	*F-Ratio*	Prob	*F-Ratio*	Prob	*F*-Ratio	Prob
AGAP000940	18135303-18239622	Frequent	0.06	ns	8.19	0.001	19.93	<0.001				
AGAP013136	18313949-18315389	Fixed	4.58	0.036	5.37	0.007	7.08	<0.001				
AGAP000998	19000322-19100511	Frequent	11.99	0.001	0.45	ns	4.86	0.001				
AGAP001009	19113598-19115474	Fixed	7.48	0.008	11.25	<0.001	12.98	<0.001				
AGAP013526	19115743-19117278	Fixed	0.22	ns	8.9	<0.001	8.51	<0.001				
AGAP001022RA	19629896-19640578	Fixed	0.58	ns	0.6	ns	2.95	0.019				
AGAP001022RB	19629896-19640578	Fixed	0.26	ns	3.27	0.045	10.22	<0.001				
AGAP001025	19714751-19718165	Fixed	0.93	ns	10.85	<0.001	5.7	<0.001				
AGAP001026	19721319-19723125	Frequent	6.92	0.011	0.51	ns	3.76	0.005				
AGAP001031	19776387-19778243	Frequent	2.91	ns	1.73	ns	0.77	ns				
AGAP001033	19814213-19815888	Fixed	1.71	ns	9.89	<0.001	12.28	<0.001				
AGAP001035	19923244-19929520	Frequent	0.18	ns	2.84	ns	9.5	<0.001				

AGAP001040	20051826-20053170	Frequent	0.67	ns	9.28	<0.001	5.54	<0.001				
AGAP001047	20424066-20427902	Frequent	6.72	0.012	4.93	0.01	20.19	<0.001				
AGAP001048	20588919-20748760	Frequent	0.32	ns	9.05	<0.001	16.92	<0.001				
AGAP001050	20944180-20955989	Fixed	1.39	ns	12.89	<0.001	4.6	0.001				
AGAP001052RC	20969559-21116818	Fixed	10.16	0.003	20.01	<0.001	36.44	<0.001				
AGAP001052RB	20969559-21116865	Fixed	4.71	0.034	4.33	0.017	16.84	<0.001	3.56	0.008	3.81	0.001
AGAP001061	22104915-22212108	Fixed	3.37	ns	3.69	0.031	17.55	<0.001				
AGAP001070	22496341-22563331	Frequent	1.66	ns	3.07	ns	5.62	<0.001				
AGAP001073RA	22747835-22776731	Fixed	7.79	0.007	8.97	<0.001	10.25	<0.001				
AGAP001073RB	22747835-22818204	Fixed	2.53	ns	6.86	0.002	6.83	<0.001				
AGAP001076	22937392-22947129	Frequent	3.27	ns	29.89	<0.001	32.67	<0.001				
AGAP001082	23455889-23493707	Frequent	12.21	0.001	33.08	<0.001	12.27	<0.001				
AGAP001083	23498802-23501429	Frequent	1.52	ns	3.93	0.025	8.78	<0.001				
AGAP001084	23511522-23634351	Frequent	4.07	0.049	5.89	0.005	28.84	<0.001				
AGAP013341	23797952-23801974	Fixed	0.84	ns	6.57	0.003	19.64	<0.001	16.48	<0.001	5.1	<0.001
AGAP001090	23834053-23862691	Frequent	0.71	ns	20.39	<0.001	25.16	<0.001				

AGAP001091	23957193-23998340	Frequent	17.52	<0.001	24.82	<0.001	7.87	<0.001
AGAP001094	24233179-24254341	Frequent	0.9	ns	6.89	0.002	10.02	<0.001

§ Interactions nested within species when appropriate.
*Different colours for Prob indicating level of significance. Red= Highly significant <0.0001, Blue=Significant <0.005, Green=<0.0

4.3.1 Differential gene expression in species

Differentially expressed genes between *An. coluzzii* and *An. gambiae* species were identified. The data was pooled from all life stages. A total of 36.7% (n=11) transcripts were differentially expressed between *An. coluzzii* and *An. gambiae*. Out of the 11 differentially expressed genes 54.5 % were upregulated in *An. coluzzii* and 45.5% in *An. gambiae* (Table 4.3.2; Figure 4.3.1).

Table 4.3.2: Differentially expressed genes in *An. coluzzii* and *An. gambiae*.

Species	***An. coluzzii***	***An. gambiae***
Transcript	AGAP001026	AGAP001009
	AGAP001047	AGAP000998
	AGAP001052RB	AGAP001073RA
	AGAP001052RC	AGAP001091
	AGAP001082	AGAP013136
	AGAP001084	

§ Gene were listed based on unconnected letters between species using the LSTukey HDS comparisons

No significant difference was found in the number of differentially expressed genes with fixed or frequent protein coding changes between species. Pearson chi-square (df = 1, χ^2 = 0.010, Prob> = 0.9193) (Table 4.3.3).

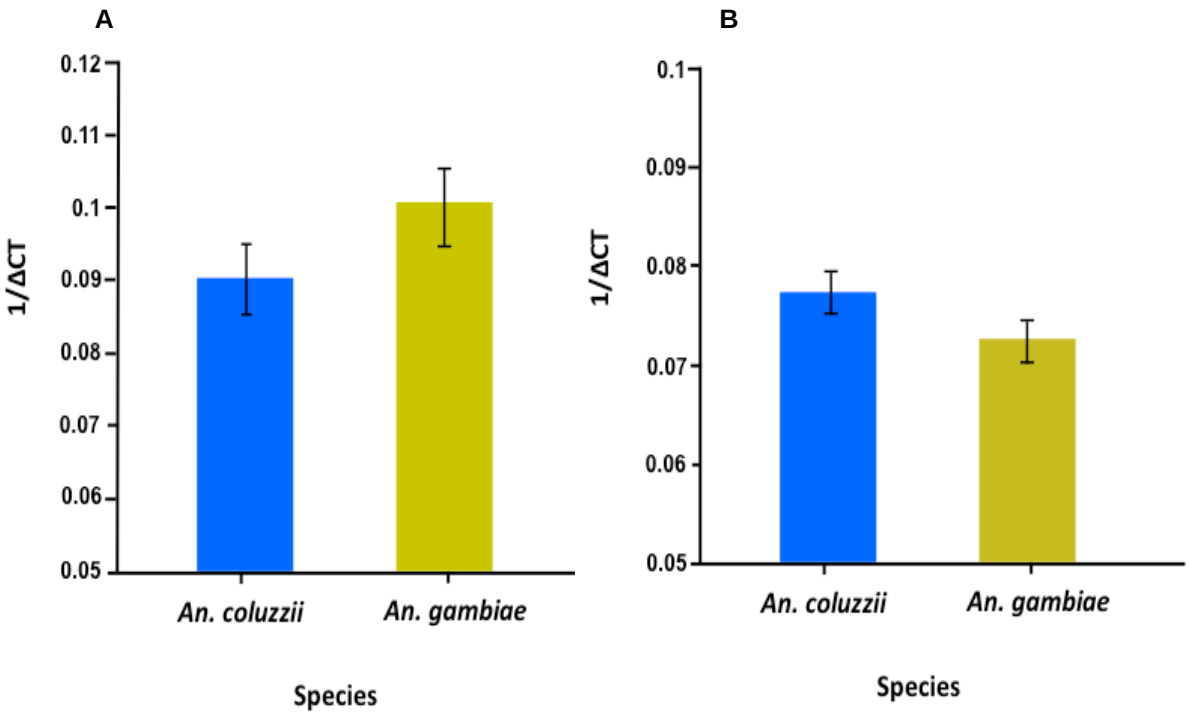

Figure 4.3.1: Differentially expressed genes between *An. coluzzii* and *An. gambiae*. A) AGAP001091 showing overexpression in the *An. gambiae* species. B) AGAP001026 showing an overexpression in the *An. coluzzii* species. Significance was determined by LSMeans difference Tukey HSD. Error bars show 95% confidence interval.

Table 4.3.3: Number of transcripts with fixed or frequent protein coding changes and differentially expressed in *An. gambiae* and *An. coluzzii*.

		Fixed coding difference(s) between *An. coluzzii* and *An. gambiae?*	
Transcripts		YES	NO
differentially expressed		(Freq >0.95)	(Freq >0.8-0.95)
between *An. coluzzii*	**YES**	**5**	**6**
and *An. gambiae*	**NO**	**9**	**10**

4.3.2 Differential gene expression in populations

Four populations were studied, Akron, Mopti, Kisumu, and Pimperena. Population effect of differentially expressed genes was 80% (n= 24). Using Tukey HDS comparison, 20.8% (n=5) of the differentially expressed genes were downregulated in Mopti strain (Figure 4.3.2; Figure 4.3.3) while 20.8% (n=5) were overexpressed in Akron compared to the other 3 populations. Two genes were overexpressed in Pimperena with a downregulation of AGA001025 (Table 4.3.4). Other differentially expressed genes 45.8% (n=11) were not assigned to any of the populations using the tukey HDS comparisons. No gene was over or underexpressed in the Kisumu strain when compared to the other 3 populations (Table 4.3.4).

Table 4.3.4: Overexpressed and underexpressed genes in four populations of *An. coluzzii* and *An. gambiae*.

Species	*An. coluzzii*		*An. gambiae*	
Populations	**Mopti**	**Akron**	**Kisumu**	**Pimperena**
Transcripts	AGAP001009 (UE)[§]	AGAP001052RB (OE)		AGAP001025 (UE)
	AGAP001033 (UE)	AGAP001052RC (OE)		AGAP001076 (OE)
	AGAP001091 (UE)	AGAP001076 (OE)		AGAP001082 (OE)
	AGAP001073RA (UE)	AGAP001082 (OE)		
	AGAP001090 (UE)	AGAP001084 (OE)		

[§]Genes were considered as over expressed (OE) and underexpressed (UE) based on unconnected letters of levels(population) of the LSTukey HDS comparisons

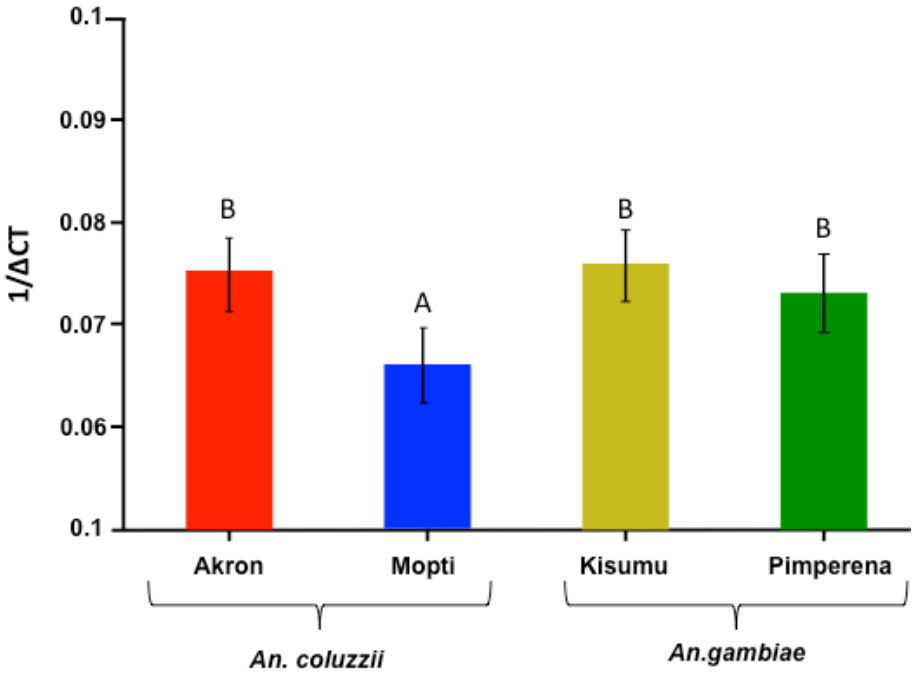

Figure 4.3.2: Differential gene expression between populations for AGAP001009. Mopti strain is underexpressed.

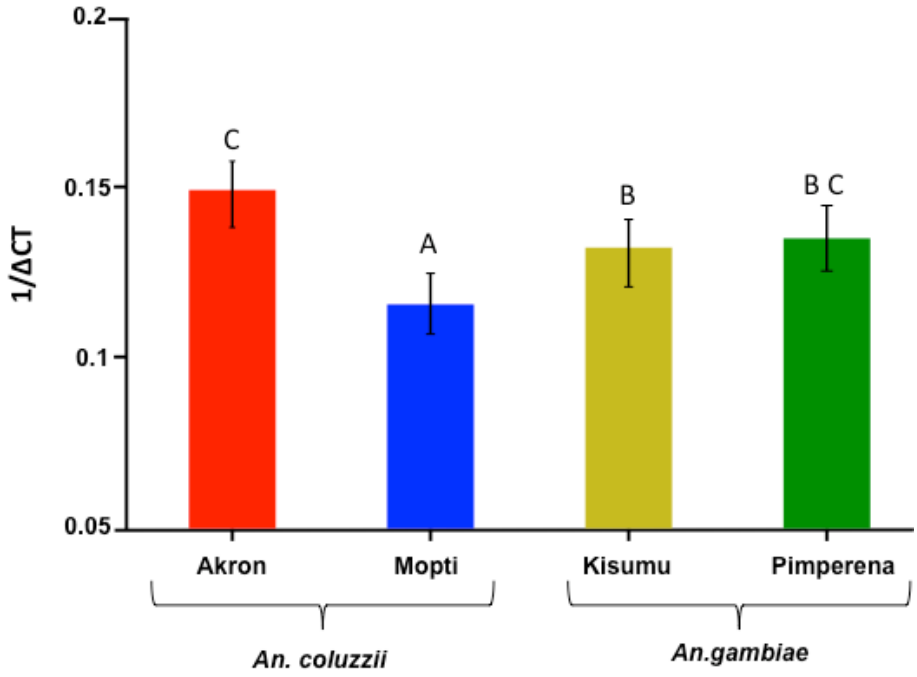

Figure 4.3.3: Differential gene expression between populations for gene AGAP010190. Mopti strain is underexpressed.

4.3.3 Differential gene expression between life stages

Genes differentially expressed at the life stages were 96.7 % (n=29). All genes except AGAP001031 were differentially expressed at the stage level (Table 4.3.1 and 4.3.2). Differentially expressed genes were male bias as majority of the genes studied were overexpressed in virgin males. Forty percent (n=12) of the stage specific differentially expressed genes were overexpressed in males. Twenty percent of the genes were downregulated at the larval stage (Table 4.3.5 and 4.3.6). AGAP001052RC was differentially expressed in mated females though downregulated (Table 4.3.5). None of the genes was found to be overexpressed/underexpressed in virgin females in comparison to the other life stages. An up-regulation of genes AGAP001052RB, AGAP001052RC, AGAP001061, AGAP001070, AGAP001076 and AGAP001084 was observed in male pupae when compared to the other life stages though not statistically significant (Table 4.3.5). AGAP001026 had an equal expression in Larvae and Virgin males. AGAP001009 and AGAP0013526 were male biased (Figure 4.3.4 and 4.3.5).

Table 4.3.5: Overexpressed and underexpressed genes in *Anopheles coluzzii* and *Anopheles gambiae* with stage specific differences.

Stage	Virgin Males	Virgin females	Larvae	Mated Female
Transcripts	AGAP000940 (OE) [§]		AGAP000940 (UE)	AGAP001052RC (UE)
	AGAP001009 (OE)		AGAP001047 (UE)	
	AGAP001033 (OE)		AGAP001048 (UE)	
	AGAP001035 (OE)		AGAP001076 (UE)	
	AGAP001047 (OE)		AGAP001090 (UE)	
	AGAP001048 (OE)		AGAP013341 (UE)	
	AGAP001076 (OE)			
	AGAP001083 (OE)			
	AGAP001090 (OE)			
	AGAP001094 (OE)			
	AGAP013341 (OE)			
	AGAP013526 (OE)			

*Staged specific gene expression. Significance between life stages was tested using the LS Means Tukey HDS comparison following the model fitting in Table1.
[§]Genes were considered as over expressed (OE) and underexpressed (UE) based on unconnected letters of levels(stage) of the LSTukey HDS comparisons.

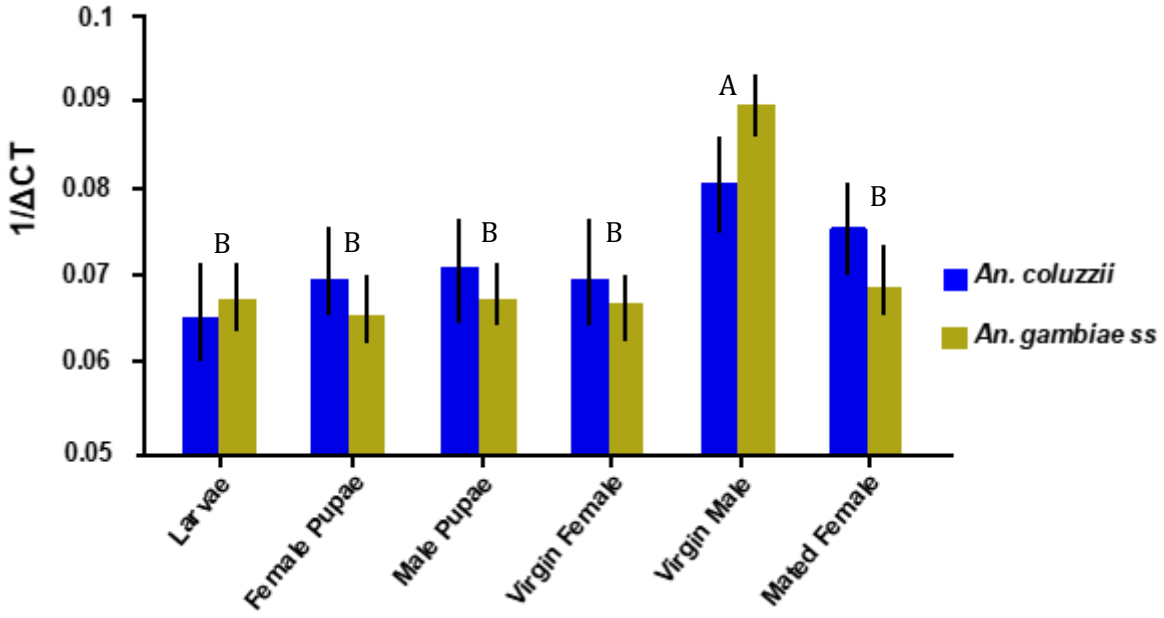

Figure 4.3.4: Stage specific differential expression for transcript AGAP0013526 showing overexpression in males. Significance was determined by LSMeans difference Tukey HSD. Error bars show 95% confidence interval.

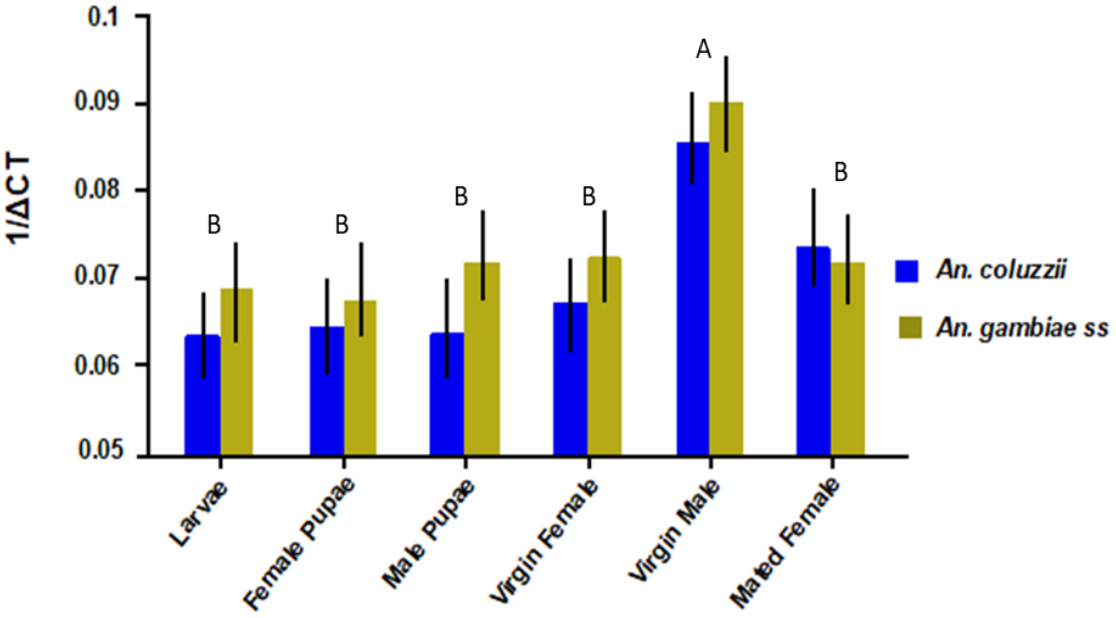

Figure 4.3.5: Stage specific differential expression for transcript AGAP001009 showing overexpression in males Significance was determined by LSMeans difference Tukey HSD. Error bars show 95% confidence interval.

Table 4.3.6: Least squares mean of ΔCT for the different life stages of *An. coluzzii* and *An. gambiae.*

Transcript	Larvae	Male Pupae	Female Pupae	Virgin Female	Mated Female	Virgin Male
AGAP000940	15.72 (14.85-16.58)	12.64 (11.78-13.5)	13.21 (12.35-14.07)	13.34 (12.48-14.2)	12.58 (11.72-13.45)	9.71 (8.85-10.57)
AGAP000998	12.75 (11.73-11.73)	11.19 (10.17-12.22)	10.47 (9.44-11.49)	11.18 (10.15-12.2)	10.81 (9.78-11.83)	9.28 (8.25-10.31)
AGAP001009	15.21 (14.44-15.98)	14.78 (14.01-15.55)	15.11 (14.34-15.88)	14.31 (13.54-15.08)	13.84 (13.07-14.61)	11.49 (10.72-12.26)
AGAP001022RA	16.36 (14.72-18.02)	15.51 (13.87-17.15)	14.73 (13.09-16.38)	15.35 (13.71-16.99)	14.59 (12.96-16.24)	12.21 (10.59-13.87)
AGAP001022RB	15.93 (15.1-16.76)	14.95 (14.12-15.78)	15.22 (16.38-16.05)	14.31 (13.48-15.14)	13.78 (12.96-14.61)	12.15 (11.32-12.98)
AGAP001025	12.48 (11.6-13.35)	11.13 (10.26-12.01)	11.39 (10.52-12.27)	11.64 (10.76-12.51)	11.97 (11.09-12.84)	9.43 (8.56-10.31)
AGAP001026	12.93 (12.1-13.76)	13.82 (12.99-14.65)	13.28 (12.45-14.11)	14.66 (13.83-15.49)	14.15 (13.32-14.98)	12.49 (11.66-13.32)
AGAP001031	10.1 (8.83-11.374)	9.65 (8.37-10.92)	9.6 (8.33-10.87)	11.01 (9.74-12.28)	9.84 (8.57-11.12)	9.51 (8.24-10.79)
AGAP001033	13.48 (12.74-14.23)	11.26 (10.51-12)	11.86 (11.11-12.6)	12.45 (11.71-13.19)	12.33 (11.58-13.07)	9.63 (8.88-10.37)

AGAP001035	13.03	11.37	12.03	12.25	12.17	9.81
	(12.32-13.75)	(10.65-12.08)	(11.32-12.75)	(11.53-12.96)	(11.45-12.88)	(9.09-10.52
AGAP001040	13.06	11.21	11.16	11.21	10.58	9.53
	(12.08-14.04)	(10.23-12.19)	(10.18-12.14)	(10.23-12.18)	(9.6-11.55)	(8.55-10.51)
AGAP001047	12.39	10.26	10.41	9.38	9.54	7.02
	(11.61-13.17)	(9.48-11.04)	(9.64-11.19)	(8.61-10.16)	(8.77-10.32)	(6.24-7.79)
AGAP001048	13.96	11.81	11.75	12.36	12.41	9.41
	(13.24-14.68)	(11.09-12.53)	(11.03-12.46)	(11.64-13.08)	(11.69-13.13)	(8.69-10.13)
AGAP001050	8.76	7.89	7.17	7.69	7.66	6.39
	(8.03-9.47)	(7.16-8.62)	(6.44-7.88)	(6.96-8.42)	(6.93-8.39)	(5.67-7.13)
AGAP001052RB	10.51	6.66	7.58	10.48	11.75	8.11
	(9.53-11.49)	(5.69-7.64)	(6.61-8.56)	(9.5-11.46)	(10.78-12.73)	(7.13-9.08)
AGAP001052RC	8.048	5.08	5.89	8.14	9.77	6.23
	(7.46-8.63)	(4.5-5.67)	(5.3-6.47)	(7.56-8.73)	(9.19-10.36)	(5.65-6.81)
AGAP001061	13.78	9.65	10.09	13.86	13.21	11.97
	(12.89-14.68)	(8.76-10.54)	(9.2-10.97)	(14.75-14.75)	(12.33-14.09)	(11.09-12.86)
AGAP001070	10.1	8.29	8.37	8.65	8.39	7.05
	(9.28-10.92)	(7.48-9.12)	(7.55-9.19)	(7.83-9.47)	(7.57-9.22)	(6.22-7.87)
AGAP001073RA	10.93	9.79	9.68	9.37	8.91	7.55
	(10.23-11.62	(9.1-10.49)	(8.99-10.38)	(8.67-10.06)	(8.21-9.6)	(6.86-8.25)

AGAP001073RB	9.65	7.28	8.54	7.6	7.02	6.02
AGAP001073RB	(8.67-10.62)	(6.32-8.25)	(7.57-9.5)	(6.63-8.57)	(6.05-7.99)	(5.05-6.98)
AGAP001076	6.96	4	4.34	4.54	5.13	2.32
AGAP001076	(6.43-7.49)	(3.47-4.53)	(3.81-4.87)	(4.01-5.07)	(4.59-5.66)	(1.79-2.85)
AGAP001082	4.45	2.53	2.71	3.42	3.34	1.92
AGAP001082	(3.95-4.95)	(2.03-3.03)	(2.21-3.21)	(2.92-3.92)	(2.84-3.84)	(1.42-2.42)
AGAP001083	8.68	6.84	6.94	7.26	7.01	4.75
AGAP001083	(7.83-9.53)	(5.99-7.69)	(6.09-7.79)	(6.41-8.11)	(6.16-7.86)	(3.9-5.6)
AGAP001084	11.2	9.23	7.88	13.49	12.75	10.27
AGAP001084	(10.41-11.99)	(8.43-10.02)	(7.08-8.67)	(12.67-14.28)	(11.96-13.55)	(9.48-11.06)
AGAP001090	10.61	7.45	8.23	8.09	7.73	5.73
AGAP001090	(9.98-11.24)	(6.82-8.08)	(7.6-8.86)	(7.47-8.72)	(7.1-8.36)	(5.1-6.36)
AGAP001091	9.423	8.35	8.06	8.51	7.9	7.37
AGAP001091	(8.93-9.92)	(7.86-8.84)	(7.57-8.55)	(8.02-8.99)	(7.41-8.39)	(6.88-7.86)
AGAP001094	11.85	10.19	10.94	11.13	11.06	8.48
AGAP001094	(11.11-12.59)	(9.45-10.93)	(10.19-11.68)	(10.39-11.87)	(10.32-11.8)	(7.74-9.22)
AGAP013136	9.5	9.17	9.21	8.89	7.99	7.42
AGAP013136	(8.89-10.12)	(8.56-9.79)	(8.6-9.82)	(8.28-9.49)	(7.39-8.61)	(6.8-8.03)
AGAP013341	10.14	7.17	7.16	7.19	7.54	4.29
AGAP013341	(9.31-10.98)	(6.33-8.01)	(6.32-7.99)	(6.36-8.037)	(6.7-8.38)	(3.46-5.13)
AGAP013526	15.09	14.41	14.67	14.65	13.98	11.77

*Ninety-five percent confidence intervals in bracket for each gene. Values were obtained from the multivariate analysis using 3 models.

4.3.4 Identification of candidate genes for RNAi

In other to identify putative assortative mating genes in the two sibling species, candidate genes for the RNAi studies were selected based on the following criteria.

- *Transcripts with fixed or frequent protein-coding differences between Anopheles coluzzii (M) and Anopheles gambiae (S).*
- *Transcripts differentially expressed between An. coluzzii and An. gambiae.*
- *Transcripts with stage specific differences (Male bias).*
- *Transcipts such as AGAP001009 and AGAP0013526 overexpressed in males and have been reported to be involved in olfaction* (Pitts *et al.*, 2011).

Some of the overexpressed genes in males have been reported to have functions in olfaction (Pitts *et al.*, 2011) (Figure 4.3.4 and 4.3.5). This was also considered an important factor in choosing our top candidate genes of speciation. Gene AGAP001009 was differentially expressed between species; differentially expressed within stage and has a fixed protein coding difference between the species. This gene has also been reported to be involved in olfaction (Figure 4.3.6 and 4.5.7) therefore a top candidate gene for RNAI. We chose two candidate genes for RNAi, AGAP001009 and AGAP0013526 that may have essential roles in assortative mating and speciation.

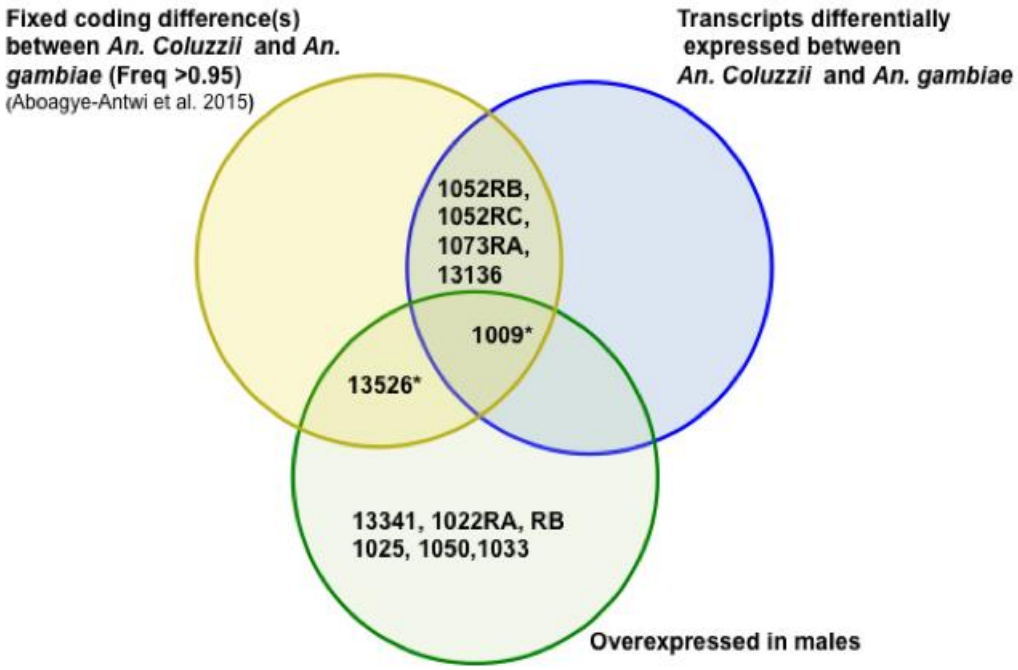

Figure 4.3.6: Venn diagram showing genes with fixed protein-coding differences between *An. coluzzii* and *An. gambiae (orange circle)*, differentially expressed genes between the species (*blue circle*) and genes overexpressed in males (*Green circle*) and genes involve in olfaction.

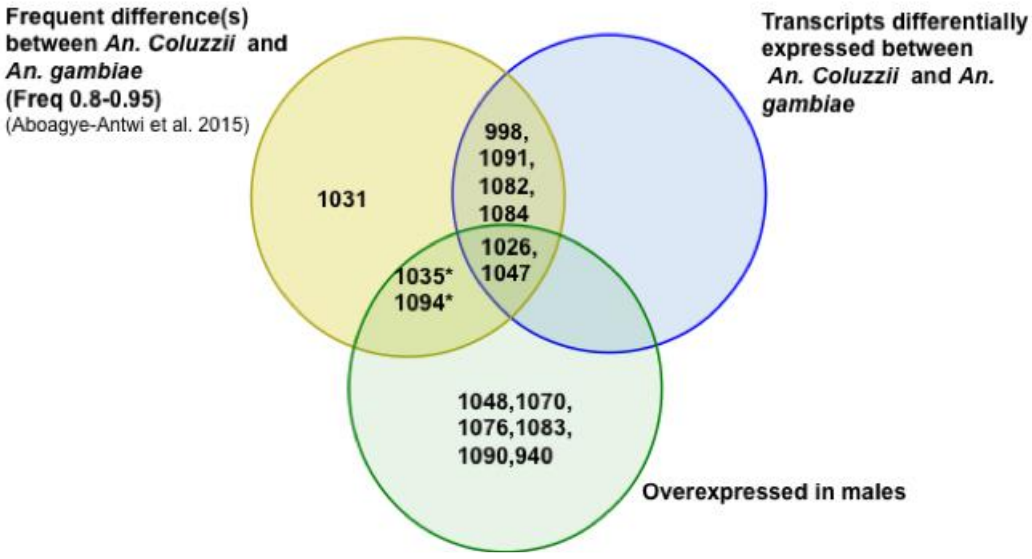

Figure 4.3.7: Venn diagram showing genes with frequent protein coding differences between *An. coluzzii* and *An. gambiae (orange circle)*, differentially expressed genes between the species (*blue circle*) and overexpressed in males (*Green circle*) and genes involve in olfaction.

4.4 Discussion

Differences in gene expression among populations could be as a result of random neutral changes and evolution by natural selection (Whitehead and Crawford, 2006). Gene expression patterns may be consistent with Stabilizing selection (when expression variance within and between species is small), neutral evolution (large expression variance within and between species), directional selection (small expression variance within and large variance between species) or balancing selection (small expression variance between-species and large variance within-species) (Whitehead and Crawford, 2006). Not much is known about gene expression evolution during speciation (Martin *et al*., 2016) nevertheless, physiological and behavioural differences between species have been correlated with gene expression (Nowick *et al*., 2009).

Speciation between *An. coluzzii* and *An. gambiae* is not fully understood because cues involved in assortative mating and ecological divergence between the sibling species are yet to be elucidated. A number of studies have shown that gene expression can be used to identify the genetic basis of various organisms through the study of gene activation and how corresponding protein perform its function. In this chapter, differential expression of 27 genes and 30 transcripts located on the X-island of speciation with fixed or nearly fixed protein coding differences between *An. coluzzii* and *An. gambiae* was studied using high throughput qPCR technique. This method is quite sensitive and has the capacity of screening one gene at a time and does not require a pre-knowledge about the differences in phenotype (Cassone *et al*., 2008). Expression levels of the 30 transcripts were compared between species, populations within species and between life stages and putative gene function discussed. Assigning of putative function of the genes was done based on gene ontology (GO) and by using putative functions already assigned to genes. This is because functional annotation of the *An. gambiae* genome is incomplete, therefore limited information is

available about possible functions and interaction of genes and how they influence ecological and behavioural differences between the two sibling species (Cassone *et al*., 2008).

In this study, differential expression between *An. coluzzii* and *An. gambiae* was observed in 11 out of the 30 transcripts studied. An up-regulation of AGAP001009 was observed in *An. gambiae* compared to *An. coluzzii* though the gene has no known function. AGAP001091 and AGAP000998 were also differentially expressed between the two sibling species with overexpression in *An. gambiae*. AGAP000998 encodes for a receptor activity protein and has a biological process of lysosome transport. An overexpression of the gene, AGAP000998 has been reported in *An gambiae* (S form) male reproductive tissues (Papa *et al*., 2017) and in the *An. gambiae* (S-form) larval stage (Cassone *et al*., 2008).

Another gene with differential expression between species in this chapter is AGAP001073. Cassone *et al* (2008) reported a differential expression of AGAP001073 between *An. coluzzii* (M) and *An. gambiae* (S) with an overexpression in *An. gambiae* (S) forms gravid females and downregulated in *An. coluzzii* (M). An upregulation of the gene was also reported in *An. gambiae* (S) virgin females (Cassone *et al*., 2008). Our study revealed an upregulation of AGAP001073RA in *An. gambiae.* The gene is found in the nucleus and encodes for protein binding with a biological process of regulation of transcription from RNA polymerase II promoter (Appendix 2).

An upregulation of AGAP001082 in *An. coluzzii* was observed in our study and the gene is involved in enzyme activator activity. The gene has been reported to be significantly different between the sibling species with overexpression in *An. coluzzii* gravid females and virgin females and downregulated in *An gambiae* (Cassone *et al*., 2008). Another gene with overexpression in *An. coluzzii* is AGAP001084. This gene encodes for alcohol-forming fatty acyl-CoA reductase activity and a biological process of lipid metabolic process. A significant

downregulation of the gene in *An. coluzzii* form 4th instar larvae and a non-significant upregulation in *An. gambiae* virgin females has been reported (Cassone *et al*., 2008).

Differential expression of genes between species could have many implications. Genes overexpressed in *An. coluzzii* could be contributory to the species ability to adapt to different environmental conditions. The species are known to withstand drought conditions (Touré *et al*., 1998), the larvae strive in polluted water bodies and are also better at avoiding predators (Diabaté *et al*., 2008). These differences in life processes could generally translate to differential gene expression between *An. coluzzii* and *An. gambiae*. The sibling species are morphologically identical but are able to identify mates of their own kind even in sympatric populations (Tripet *et al*., 2001). Mate recognition cues in *An. coluzzii* and *An. gambiae* such as swarm segregation (Diabaté *et al*., 2006; Diabaté and Tripet, 2015; Sawadogo *et al*., 2013). Wingbeat, flight tones (Cator *et al*., 2010; Pennetier *et al*., 2010) and contact pheromone (Pitts *et al*., 2014; Tripet *et al*., 2004) could translate into differential expression of genes since physiological and behavioural differences between species have been correlated with gene expression (Nowick *et al*. 2009).

Some genes with differential expression between species were observed to be upregulated in males. These genes could have an implication on mating as well. AGAP001009 was differentially expressed between *An. gambiae* and *An. coluzzii* and the gene was upregulated in males. A non-significant differential expression of AGAP001009 with an upregulation in the carcass of male *An. gambiae* and downregulated in the reproductive tissues of females have been reported (Papa *et al*., 2017). AGAP013526 is a gene present in the peroxisome and has no known function. It has been associated with overexpression in the testis of male *An. gambiae* when compared to female reproductive tissues (Papa *et al*., 2017). Our data revealed a significant difference in AGAP013526

between life stages with an overexpression in virgin males. The gene was not differentially expressed between species.

Other genes with overexpression in virgin males include AGAP001035, AGAP001094, AGAP013341 and AGAP001025. AGAP001035 encodes for enzyme activator activity and AGAP001094 encodes for DNA binding, and a biological process of regulation of transcription. An upregulation of AGAP001094 and AGAP013341 have been reported in male carcass and downregulated in female tissues (Papa *et al*., 2017). AGAP001025 has no known function and has been reported to be upregulated in male tissues as well when compared to female tissues (Papa *et al*., 2017). Other differentially expressed genes with overexpression in males are AGAP001033 and AGAP001026. AGAP001033 has no known function. AGAP001026 encodes for Carboxypeptidase activity and a biological process of proteolysis, and has been reported to be overexpressed in the malphigian tubules in males (Baker *et al*., 2011).

An overexpression of AGAP001076 was observed in our data, the gene is involved in Oxidoreduction activity and has no known function. The gene was reported to be upregulated in male tissues when compared to female tissues (Papa *et al*., 2017). A significant differential expression of AGAP001076 between males and females with an upregulation in males has also been reported (Marinotti *et al*., 2006). Energy is released through oxidation-reduction in living organisms and AGAP001076 encodes for oxidation-reduction. An upregulation of AGAP001076 in virgin males could be attributed to the vigorous processes involved in life activities of adults such as searching for food and locating mates.

Our result showed an overexpression of AGAP013526, AGAP001009 in males, and these genes have been reported to involved in olfaction (Pitts *et al*., 2011). Mosquitoes rely on their sense of smell for most of life's choices. They rely on olfactory cues to locate human

host during blood feeding, location of oviposition sites and even location of nectar (Potter, 2014). The genes could have an implication on the life processes aforementioned.

Some differentially expressed genes between species have been associated with DDT resistance from previous studies. These genes are AGAP001025, AGAP000998, AGAP000940 AGAP001048, AGAP001061 AGAP001048 and AGAP001061. Knockdown resistance is associated to a mutation occurring in the gene encoding a voltage gated sodium channel (Wilson *et al*., 2018). Knockdown resistance to pyrethroid and dichlorodiphenyltrichloroethane (DDT) occurs in *An. gambiae* (S) individuals (Esnault *et al*., 2008). Kdr resistance was not found in the *An. coluzzii* (M forms) even in sympatric populations (Tripet *et al*., 2007). When Kdr resistance is present in the *An. coluzzii* (M forms), it is said to be through introgression from the *An. gambiae* (S) taxon (Weill *et al*., 2000). AGAP001025, AGAP000998 and AGAP000940 were associated with DDT resistance in field samples (Tene *et al*., 2013). AGAP000998 encodes for a receptor activity protein, and has a biological process of lysosome transport. An overexpression of these genes were observed in the *An. gambiae* when to compared to the *An. coluzzii* (Tene *et al*., 2013). AGAP001048, AGAP001061 (encoding for Protein binding), were overexpressed in the *An. coluzzii* (M forms) in DDT resistant field samples (Tene *et al*., 2013). AGAP001091 is involved in protein phosphorylation, and has also been associated with DDT resistance in field isolates with an upregulation in *An. coluzzii* (M forms) (Tene *et al*., 2013). The differential expression of these genes could play a role in the differences in insecticide resistance between the two species. Insecticides could be a selective factor that leads to differences in gene expression between the two species.

Some differentially expressed genes in this study have also been associated with desiccation stress effect. These genes include AGAP000998, AGAP013136, AGAP001050, AGAP001070, AGAP001073, AGAP001083 (nucleic acid binding), AGAP001084,

AGAP001090 and AGAP001091 (Wang *et al*., 2011). Differential expressions of these genes were observed in our data. *An. coluzzii* and *An. gambiae* have different tolerance to drought. *An. coluzzii* can withstand drought and this is associated with some chromosomal inversions therefore they are dominant in dry regions and present during the dry season. *An. gambiae* on the other hand are found mostly during the rainy season (Touré *et al*., 1998). This ecological characteristic is contributory to all year round transmission of malaria therefore the differential gene expression could translate to ecological adaptation in the two sibling species leading to speciation.

When gene expression levels were compared between life stages, our findings revealed genes with upregulation in virgin males and a downregulated in larvae, these genes are: AGAP013341, AGAP001090 (protein binding), AGAP001048, AGAP001047, AGAP000940, and AGAP001076. An overexpression of AGAP000940 in *An. coluzzii* (M form) adults and underexpressed at the larval stage have been reported (Marinotti *et al*., 2006). Genes with differential expression between adults and larvae might have an implication on ecological divergence. The adult and larval life stages inhabit separate habitats leading to difference in their biotic and abiotic features. The difference in larval habitat preference exposes *An. coluzzii* to predation and pollution. This is because *An. coluzzii* have a preference for permanent water bodies e.g. rice fields, which are usually prone to pollution. Therefore, overexpression in larvae could involve transcripts involved in detoxification.

We anticipated more genes with fixed protein coding changes (>95) between the species would be differentially expressed compared to those with nearly fixed protein coding changes (> 80) between the two species. However, no difference was found in the number of differential expression genes with fixed and nearly fixed protein coding changes. This is an indication that genes without fixed protein coding differences could be involved in speciation in the two sibling species.

The naming of *An. coluzzii* and *An. gambiae* as two species may be arguable because of on-going hybridization and introgression. However, differences in larval habitat preferences, predator avoidance behaviour, adult mate recognition cues such as swarming exist between the species and this indicates selection on genetic targets such as chemosensory genes and regulators of development (Cassone *et al*., 2014).

Chapter 5 Putative assortative mating gene knockdown and its impact on *An. coluzzii* and *An. gambiae*

5.1 Introduction

Very strong assortative mating occurs between *An. coluzzii* and *An. gambiae* (Tripet *et al*., 2001), even in mixed swarms, the two sibling species are able to choose potential mates of their own kind (Diabate and Tripet, 2015). Swarms present an opportunity for aggregation and mating (Diabaté *et al*., 2006) and the creation of swarms involves the use of visual cues necessary for the identification of different markers used for swarm site selection. Till date, the environmental and behavioural cues used during swarming is not fully understood (Diabate and Tripet, 2015). Male aggregation pheromones have been suggested to be responsible for attracting females to swarms (Pitts *et al*., 2014). However, in *An. coluzzii* and *An. gambiae*, aggregation pheromone has been questioned due to the presence of mixed swarms (Diabate *et al*., 2015). It has also been suggested that *An. gambiae* species depend on contact pheromone for mate recognition (Tripet *et al*., 2004). Contact pheromones have been observed in the mosquito *Culiseta inornata*, where the presence of pheromones on the legs of females aid in recognition of conspecific females by males through receptors located on their tarsi (Lang, 1977). *Aedes aegypti* females are also thought to have pheromones on their tarsi that could be used in conspecific mate recognition (Nijhout and Craig, 1971).

The creation of swarms by onset of photophase was reported in *Aedes aegypti*. Swarm creation was triggered by odours from rats that are potential source of blood meal for the females. Both males and females were attracted to the swarm and olfactometry, revealed a volatile pheromone produced by swarming males which triggered flying activities of females at a distance (Cabrera and Jaffe, 2007). Females were found to also produce a volatile attractant. These results, therefore suggest the possibility of aggregation pheromone produced by both males and females, which attracts other individuals towards swarms. These findings

suggest that pheromone based vector control is possible (Cabrera *et al*., 2007).

Chemical cues involved in swarming were also incriminated in another study carried out on *Aedes aegypti* (Fawaz *et al*., 2014). Three aggregation pheromones that could be involved in swarming were isolated (Fawaz *et al*., 2014). The study reported that males depend on acoustic signals for swarming while females depend on olfactory cues (Fawaz *et al*., 2014). The results suggested that *Aedes aegypti* females use olfactory cues such as aggregation pheromones.

Apart from the possible use of olfaction in mate recognition, mosquitoes rely on olfaction for most of life's choices and this can be considered a weakness that can be harnessed (Carey and Carlson, 2011). They rely on olfactory cues to locate human hosts during blood feeding; this is because humans secrete several attractants in sweat and CO2 while breathing. Their sense of smell is also needed in locating nectar and for the location of appropriate oviposition sites (Carey *et al*., 2011).

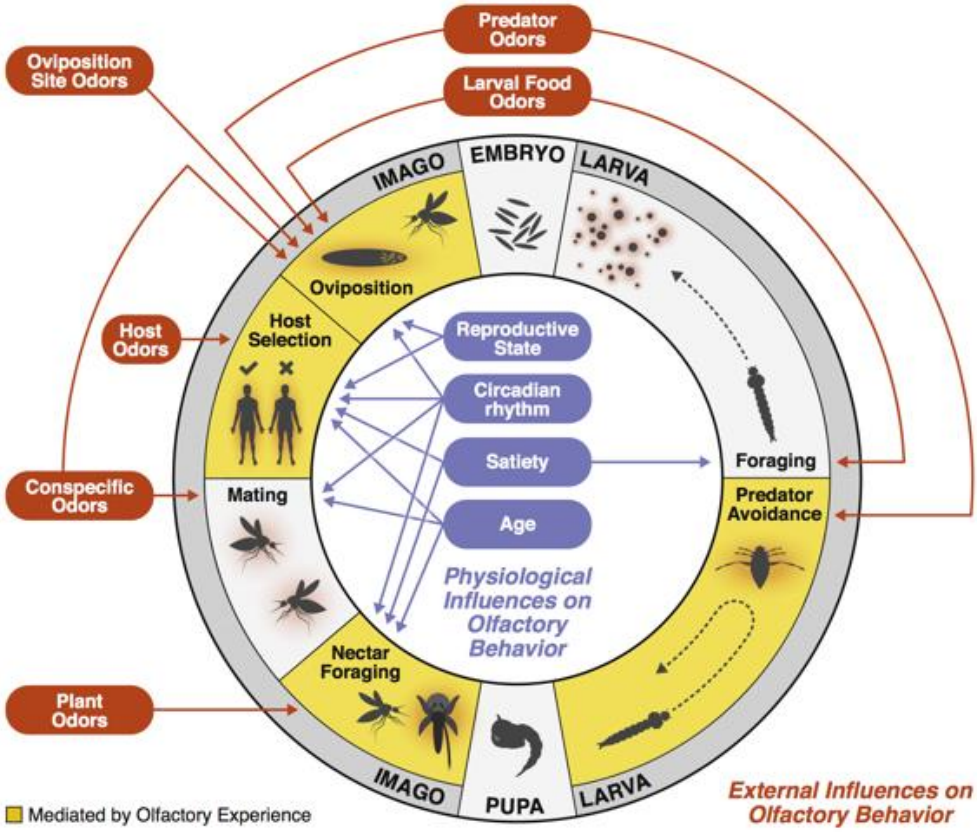

Figure 5.1.1: Known physiological and external influences on mosquito olfactory behaviour. Behaviours marked in yellow are mediated by olfactory experience. These behaviours can also be modulated by physiological factors (blue), **Source:** (Lutz *et al*., 2017).

The host seeking phenotype of the highly anthropophilic *An. gambiae* and *An. quadriannulatus*, a related but zoophilic species, were studied using transcriptome profiles of the antennae. Messenger RNA from bloodfed females were sequenced and compared. The study revealed differences in the chemosensory genes of both species suggesting divergence between the species (Rinker *et al*., 2013). A more recent study was carried out on *An. coluzzii* which is also highly anthropophilic, and on the zoophilic *An. quadriannulatus* which prefers to feed on bovids and is therefore not considered as contributors to the prevalence of malaria (Athrey *et al*., 2017). In the study, transcriptomes of the antennae and maxillary palps in *An. coluzzii* and *An. quadriannulatus* were compared and differences in the expression of

chemosensory genes were observed. *An. coluzzii* had an expression of 6 olfactory receptors (Ors), seven ionotropic (Irs), while 11 Ors and 3 Irs were upregulated in *An. quadriannulatus*. There was a general overexpression of odorant binding protein (OBPs) in the antennae and the palps of *An. coluzzii*. The chemosensory genes were expressed differently between the two species and these genes could be involved in the difference in host preference in *An. coluzzii* (Athrey *et al*., 2017). Studies on the characterisation of large families of *An. gambiae* odorant receptors (AgORs), gene silencing and behavioural analyses are now used to examine the roles of AgORs (Liu *et al*., 2010).

Gene silencing, also known as RNA interference (RNAi) is a method in which the expression of a specific gene sequence is suppressed (Figure 5.1.2) (Airs and Bartholomay, 2017). This has been used to study the biological function of genes as well as used in pest management and in the reduction of disease pathogens (Airs *et al*., 2017). RNAi is a gene specific silencing mechanism using dsRNA synthesized with sequences for a specific target. In 1998, RNAi was reported to have reduced gene expression in *Caenorhabditis elegans* through the injection of dsRNA (Fire *et al*., 1998). After that discovery, RNAi has been widely used in research on functional genomics because researchers can now investigate the functional roles of genes of interest (Regna *et al*., 2016). The ability to silence individual genes has led to the discovery of the roles of many genes in an organism. This is usually identified by the loss of functions after silencing (Regna *et al*., 2016).

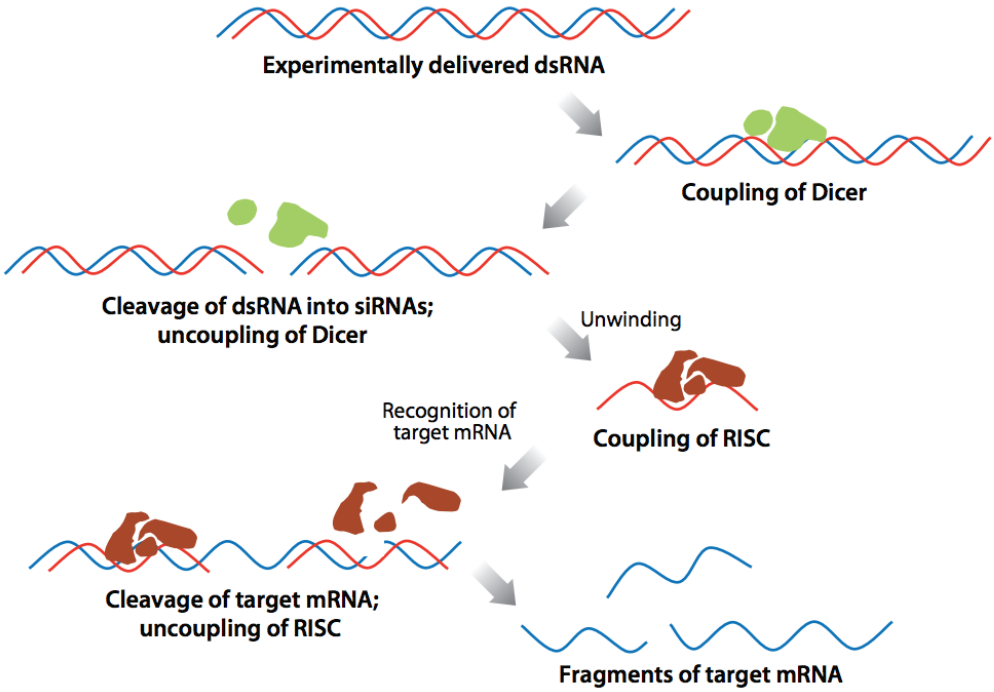

Figure 5.1.2: Basic mechanism of RNA interference (RNAi). The double-stranded RNA (dsRNA) is cleaved into fragments of ~21 nucleotides (the short interfering RNAs, or siRNAs) by the enzyme Dicer. The siRNAs unwind, and the antisense strand couples to the RNA-induced silencing complex (RISC) and conveys it to the target mRNA. Then RISC couples to the target mRNA, blocking and degrading it. **Source**: (Bellés, 2010).

There are different methods for dsRNA delivery. These include microinjection, ingestion, oral feeding and lipid and nanoparticle-coated delivery (Figure 5.1.3). However, these methods have advantages and limitations (Pillai *et al*., 2017). Microinjection was used with success on RNAi firstly on Drosophila melanogaster. It has now been applied on several insects such as the red flower beetle *Tribolium castaneum* (Brown *et al*., 1999), *An.gambiae* (Blandin *et al*., 2002). In *An. gambiae,* injection of dsRNA into the haemocoel is the most common method used for gene silencing (Regna *et al*., 2016). Mosquito tissue can be reached by injections of dsRNA, however, the success of the knockdown varies between genes and the concentration of dsRNA injected (Biessmann *et al*., 2010).

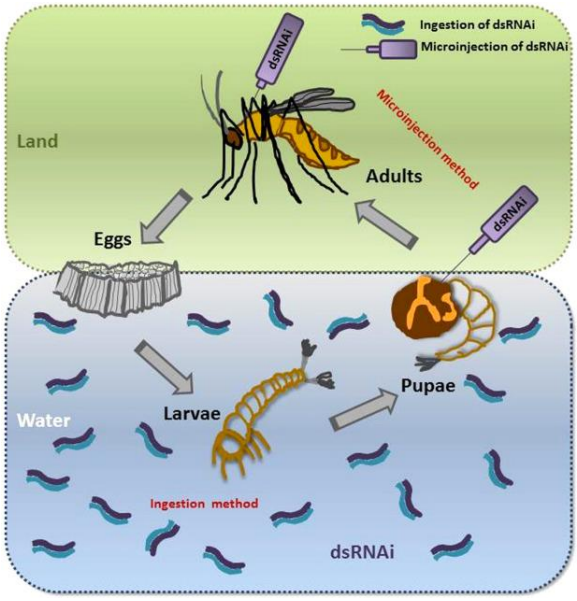

Figure 5.1.3: Possible modes of double stranded RNA (dsRNA) uptakes in different stages of mosquito. Various developmental stages live either in water or land to complete their life cycle. The technique employed would depend on the developmental stage. In embryonic, larval and pupal stages, the method of dsRNA employed in ingestion while in pupae and adult stage microinjection is utilized. **Source:** (Pillai *et al*., 2017).

The efficiency of RNAi and how long the transcript is suppressed differs between species (Scott *et al*., 2013). There is no specific strategy when designing RNAi for a particular gene or insect. This is because each species is different and a lot is yet to be understood about RNAi signal and how it is amplified and spread in the insect's cell (Scott *et al*., 2013).

RNAi mediated gene silencing has been used to demonstrate the significance of the interaction of *An. gambiae* odorant protein AgamOBP1 and a ligand, indole in OBPs odor recognition (Biessmann *et al*., 2010) The transcription of odorant binding protein (OBP) gene CquiOBPI was reduced in the antennae of *Culex quinquefasciatus* using RNAi. This led to a reduction in responses to oviposition attractants compared to those injected with water (Pelletier *et al*., 2010). Studies on the characterisation of large families of *An. gambiae* odorant receptors (AgORs), gene silencing and behavioural analyses are now used to examine

the roles of AgORs (Liu *et al*., 2010) even though the olfactory system is involved in the coordination of a lot of behaviour in the mosquito, its involvement in sex specific regulation is not clear (Das De *et al*., 2017). The transcriptomic data annotation of bloodfed adult female olfactory tissue of *An. culicifacies* revealed 383 bp transcripts previously found to be involved in courtship behaviour in male *Drosophila*. The study revealed that *Ac-qtc*, (quick-to-court) a protein whose increased expression was found to correlate with the maturation of the olfactory system could play a role in sex conflicting demand of mosquito courtship. Gene silencing using dsRNA injected in the thorax of male mosquitoes was unsuccessful therefore they used transcriptional modulation for the investigation (Das De *et al*., 2017).

To identify chemosensory-enhanced genes in male and female *An. gambiae*, chemosensory tissue transcriptomes and whole body transcriptome were compared in both male and female *An. gambiae* (Pitts *et al*., 2011). In this extensive study, all known chemosensory genes were enriched in both male and female olfactory tissues. Gene expression patterns were evaluated between sexes. Differences were observed in sensory specialization between male and female antennae. Chemosensory genes were highly expressed in female antenna an indicator that the organ is used for chemosensation. These genes were underexpressed in males. The males on the other hand had an enhancement of genes involved in audition and these genes were observed to a lesser degree in females. It was concluded that male and female antennae have inverse prioritizations and sensitivities (Pitts *et al.*, 2011) as female antennae are more developed for chemoreception while in males; it is more developed for sensing auditory cues (Pennetier *et al*., 2010). Differences in gene expression in the palps had 778 genes common between males and females, 1906 were enhanced in females and 2284 in males. Enhanced gene overlap was observed as 61% of female antennal-enhanced genes were shared with males, 41% of the total female palp

enhanced genes were shared with males. It was suggested the palps could have cryptic sex-specific specializations (Pitts *et al*., 2011).

In a bit to narrow down on assortative mating genes, the extensive study by (Pitts *et al*., 2011) on olfaction was used as an additional filter to select our best candidate genes for gene knockdown experiments. Genes involved in olfaction were considered, as they could also be involved in host finding and mating. This is because chemosensory signal guides majority of insect behaviour (Ihara *et al*., 2013). Filters used were genes in *An. coluzzii* and *An. gambiae* (from Chapter 4) with fixed or frequent protein coding changes between the sibling species (Aboagye-Antwi *et al*., 2015) differentially expressed between species, and differential expression between life stages (overexpression in males) (Figure 4.3.6; Figure 4.3.7). Genes that passed through the filters and reported to be involved in olfaction by (Pitts *et al*., 2011) were considered as top candidate genes. AGAP001009 was found to be differentially expressed between *An. coluzzii* and *An. gambiae* and has fixed protein coding changes between the two species. The gene was also overexpressed in males and has been linked to olfaction by (Pitts *et al*., 2011). Therefore, the gene was considered our top candidate gene. AGAP013526 has fixed protein-coding changes between the two species, overexpressed in males and has been reported in olfaction (Pitts *et al*., 2011). This gene was not differentially expressed between the two species (chapter 4) but was still considered a candidate gene. This study is important as it could lead to the identification of assortative mating genes involved in mate recognition cues such as the use of pheromones and wingbeats u by *An. coluzzii* and *An. gambiae*. Elucidating these mating cues will not only contribute to vector control programmes focusing on mosquito releases but can also lead to development of alternative vector control tools that takes advantage the mosquitoes mate and host finding behaviours.

5.2 Methods

The method used for dsRNA delivery was injections and adult *An. coluzzi* and *An. gambiae* were used for the gene knockdown experiments. Double stranded RNA was synthesized targeting the two candidate genes AGAP001009 and AGAP013526. dsRNA delivery in *An. coluzzii* and *An. gambiae* was done using microinjection and behavioural assays conducted.

5.2.1 Mosquito rearing

The mosquitoes were breed in optimal conditions in the Keele University insectary. *An. coluzzii* (Mopti strain) and *An. gambiae* (Kisumu strain) were used for this study. Two hundred first instar larvae were trayed out and fed using the standard feeding regime described in chapter 2. The two strains were reared at the same time, trays containing larvae where placed interchangeably on the shelves to prevent variations occurring between species due to temperature or any other environmental factors. After 7 days, pupae were picked and sexing was done separating males and females of the two strains. They were put into properly labelled emergence cages. Ten percent glucose solution was provided and a wet cotton pad was placed on the cage. The next morning, cages were properly checked for possible errors in sexing. This was done to ensure all mosquitoes used were virgins. The male genitalia rotate at maturity (Dahan *et al*., 2014) and only after then does mating occur. Therefore, checking for errors was done within 24 hours after adults emerged. Gene silencing was attempted on three days old mosquitoes by intra-thoracic injection as a method of dsRNA delivery.

5.2.2 dsRNA synthesis

The candidate genes for silencing were selected based on the outcome of results from Chapter 4 and primers were designed for genes of interest. dsRNA was synthesised using the MEGAscript RNAi Kit (Life Technologies, USA). The kit required a template DNA for

synthesis and gene specific primers (with T7 promoter).

Primer Design

Primers were designed using two primer design software, vector base (vectorbase.org blast) and primer blast (ncbi.nlm.nih.gov/tools/primer-blast). The sequence for each gene was obtained from vector base. Transcripts with multiple splice variants were encountered and the appropriate exons that overlapped on most or all the splice variants were selected. Exons of approximately 500base pairs were selected. These exons were then put into primer blast where options of possible primers were gotten. The forward and reverse primer pair for the different genes of interest were chosen based on the GC content 50%, melting temperature 60^{o}C, and the size of the product. Primer choice was done considering factors such as melting temperature, which was 59^{o}C-60^{o}C. GC content of 50% and the lengths of the primers were between 20-22 base pairs. The primers were then checked with the Mopti (*Anopheles coluzzii* strain) and BSS (recombinant strain) SNPs database (database source: Pitts, 2011). This was done to check that primers do not bind in regions with SNPs, to ensure that the chosen primers would work for all strains used in the study**.** A T7 promoter sequence was added to the primers. This was for the dsRNA transcription process. T7 promoter sequence was (5'-TAA TAC GAC TCA CTA TAG GG-3'). This was added to the 5' ends alongside a 6bp promoter as started in the Promega (2015) manual. Since there were different genes of interest, primers were designed for the different target genes (Table 5.2.1) and the PCR conditions for amplification of each transcript was optimized (Table 5.2.3). Two controls were used at the course of this research; these were LacZ from *Escherichia coli* and Normal saline.

Table 5.2.1: List of primers for RNAi candidate genes

GENE	Length (bp)	Forward primer	Reverse primer
AGAP001002	1245	AACGCAGTGAACAGCTCTAATC	TTTTGAGTTGCGACTGCCAC
AGAP001009	624	GATTTACGGAGAGCCTGCCC	GATGGGAGTGGAAGACCAGG
AGAP001022	450	TACGGTATCGAGCTGAGGGAT	GACGACCGACAGCAAGAAGAT
AGAP001025	1147	TGAAGCTGAGTGCCATCCTG	AAACGGTCCGGTAGTTGGTC
AGAP001033	1375	TCCAGATTACACCCGCCTTC	TTGGTACACCGGGCAATTCT
AGAP001050	616	CTTCGTTTTGGGCGAAGTGTA	TTCCTCTCGGTTCCTGACAC
AGAP001052	457	CTTGGGCTACAGCCGTTCAG	AGGCGTTCGAGCGTAATTTAGTC
AGAP001061	583	GCTACGACGGACTGACAAGA	GTAAACTTCAGTCGCTTGGCG
AGAP001073	1096	ACCTACCGAAATGCCCTCCT	CACCAGATTGCCCTGTCGTC
AGAP013136	663	CGCAAACGGGGTCGAGTATT	ATCCGGTGATACTCCCGGTC
AGAP013341	4023	CGGACTCAGGGAAAGCATGT	ACTGACGGAGGAGCATCTGA
AGAP013526	438	CATGGGGACGGTCTGCTTTA	AGCTCTAGAATCGCGCCAAA
LacZ (*E. Coli*)	3075	TCTGGCGGAAAACCTCAGTG	TCGTAATCAGCACCGCATCA

Template DNA

To generate the template for the dsRNA synthesis, we extracted genomic DNA from a pool of five *An. coluzzii* mosquitoes; DNA extraction was carried out using the DNAzol extraction method (see chapter 2). A minimum concentration of 250ng/µl of DNA was used for a 25µl PCR reaction. This was to ensure a high concentration of dsRNA. Double stranded RNA was routinely synthesized through vitro transcription of a PCR generated DNA template containing T7 promoter sequence on both ends.

PCR

A master mix was prepared for a 25µl reaction using appropriate volumes of reagents for each

gene of interest (Table 5.2.2). The genomic DNA was added to the master mix and amplified using optimised thermal cycling conditions (Table 5.2.3). The volume of components used was same for all target genes. The PCR conditions for amplification differed between genes (Table 5.3.3). The PCR product was run on gel to check if amplicon was the exact size. DNA ladder were used to check the product size (Figure 5.2.2).

Table 5.2.2: PCR Components for dsRNA synthesis.

Components	**Volume for 25µl reaction**
10X PCR Buffer with $MgCl_2$	2.5
DNTPs (10x dilution)	0.5
Forward primer (*Gene specific*)	0.25
Reverse Primer (*Gene specific*)	0.25
Taq polymerase	0.25
Template DNA	1
(*Use concentration of 250ng/µl*)	
Nuclease-free water	20.375 (add up to total volume)
Total	25

Table 5.2.3: PCR conditions for dsRNA synthesis.

PCR Cycles	Periods
95^{o}C	5 minutes
95^{o}C	30 seconds
57^{o}C (Temp. changed for each gene)	30 seconds
72^{o}C	1minute
Repeat cycle 4, 5 times	
95^{o}C	30 seconds
72^{o}C (Temp. changed for each gene)	1 minute
Repeat the cycle 25 times	
72^{o}C	5 minutes
4^{o}C	1 hour to allow for storage

Purification of amplified PCR product

Purification of the amplicon was done using Wizard SV Gel and PCR clean-up system (Promega, USA) following the manufacturers manual. Several reactions of amplified PCR product were combined. A maximum of 100µl was used per column and the volume at which the column is most efficient in terms of yielding the highest concentration was checked (Figure 5.2.1). An equal volume of membrane binding solution was added to the amplified product. SV minicolumn was inserted into a collection tube and the mixture was passed through the column and incubated at room temperature for 1minute.

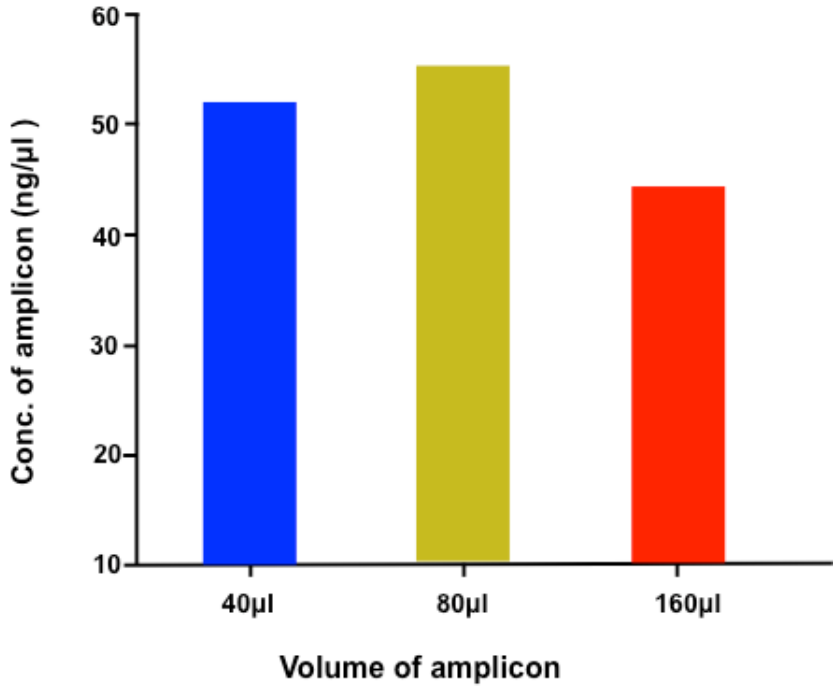

Figure 5.2.1: Volume with the best yield of PCR amplicon after purification using.

This was done to bind the DNA to the column. It was centrifuged at 16000xg for 1 minute, flow through discarded and column inserted into a new tube. Washing was done by adding 700µl membrane wash solution (ethanol added) through the column and spun at 16000xg for 1 minute. Flow through was discarded and minicolumn reinserted into the collection tube. This was repeated with 500µl membrane wash solution and centrifuged at 16000xg for 5 minutes. The collection tube was emptied and the column assembly was spun for 1 minute with the microcentrifuge lid open (or off) to allow evaporation of any residual ethanol. The minicolumn was carefully transferred to a clean 1.5ml microcentrifuge tube and 30µl of nuclease free water was added to the minicolumn. This was incubated at room temperature for 1 minute and centrifuged at 16000xg for 1 minute to elute DNA. A concentration of 300ng/µl amplicon was needed so as to have a final concentration of about 3µg synthesized dsRNA. When the concentration of purified DNA was low, a freeze dryer was used to dehydrate samples. The samples were re-eluted to the desired concentration using less amount of RNase

free water.

Transcription reaction assembly

The purified amplicon was used for the transcription reaction using a MEGAscript RNAi Kit (Thermofisher Scientific) and following manufacturers manual. A 20µl reaction was assembled in a 0.5ml PCR Eppendorf tube (Table 5.2.4).

Table 5.2.4: Transcription reaction.

Component	**Volume (µl)**
10x T7 reaction buffer	2µl
ATP Solution	2µl
CTP solution	2µl
GTP Solution	2µl
UTP Solution	2µl
T7 enzyme mix	2µl
Linear template DNA	1µl-8µl (Total of 2µg purified PCR product per reaction)
Nuclease-free water	0-7µl (add to make up total volume per reaction)
Total	20µl

The reagents were mixed by pipetting up and down and incubated at 37^{o}C for 4 hours. After incubation, nuclease digestion to get rid of any DNA was done by mixing 5µl 10x digestion buffer, 2µl DNase 1, 2µl RNase, 21µl nuclease-free water and the 20µl of dsRNA from the transcription step and incubated at 37^{o}C for 1 hour. dsRNA purification was done by putting the dsRNA from the previous step into a 1.5ml Eppendorf tube and adding 50µl of 10X binding buffer, 250µl of 100% ethanol and 150µl of nuclease free water.

Table 5.2.5: dsRNA purification.

Component	Volume (µl)
dsRNA from previous step (From step 4)	50µl
10X binding buffer	50µl
100% ethanol	250µl
Nuclease free water	150µl
Total	500µl

The total volume of 500µl dsRNA binding mix was passed through a filter cartridge and centrifuged at max speed for 2 minutes twice. This was done twice. Flow through was discarded and centrifugation was repeated for 1 second to get rid of any wash solution. dsRNA was eluted dsRNA using 30µl ultrapure Nuclease free water. This was incubated in a heat block set to 65^{o}C for 2 minutes after which it was centrifuged for 2 minutes at Max speed to elute dsRNA. The column was transferred to a new tube and the process repeated again to elute any dsRNA left in the column. dsRNA was then quantified using a nanodrop and the integrity of dsRNA was checked by running on gel. When dsRNA concentration was low, the sample was freeze-dried and re-suspended in a lower volume of water. For injection into mosquitoes using 69nl per injection dsRNA was adjusted to a concentration of 3µg/µl.

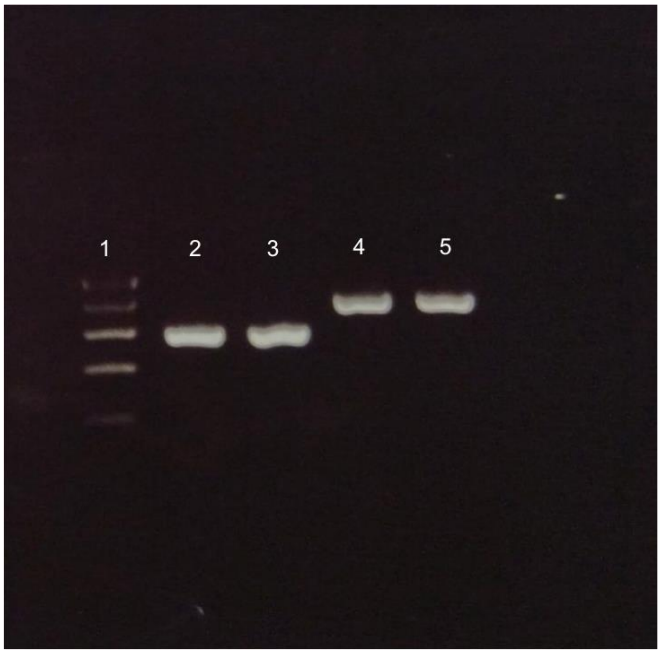

Figure 5.2.2: dsRNA synthesis. Lane 1= DNA marker, Lane 2 and 3= Gene AGAP013526 expected amplicon size=249. Lane 4 and 5 = Gene AGAP001009 expected amplicon size =440.

5.2.3 Gene Knockdown

Mosquito collection and preparation

Two days after emerging, the mosquitoes were prepared for injections. A clean cage was prepared for injected mosquitoes. A 10% glucose source was provided, wet cotton bud was placed on the cage and a wet blue roll was also placed on the cage to improve humidity. Using a mouth aspirator, individual mosquitoes were collected from the cage and transferred into a test-tube lined with cotton wool and buried into a bucket of ice. This was because cold temperatures anaesthetise mosquitoes. The test tubes were usually prepared and placed on ice before mosquitoes were transferred so that anaesthetising can be quick to ensure mosquitoes spend less time on ice to prevent mortality. The waiting time for anaesthetising is 1-5 minutes. When the mosquitoes were anaesthetised and ready for injections, a petri dish

(60mm) filled with ice and lid placed on was improvised as a cold block. This was wrapped with foil paper to prevent the mosquitoes from sticking to it during injection. The petri dish was placed under a dissecting microscope. Forceps and a soft paintbrush were used to transfer mosquitoes onto the cold block and they were injected using a Nanoject II Auto-Liter micro-injector (Drummer scientific, USA).

Preparation of microinjection

Borosilicate capillary tubes were used for injections. The 3.5mm tubes were pulled to produce a thin fine needle using a pipette-puller (P-2000, sutter instrument company, Novato, CA). The conditions for the puller was set at Heat= 350, Fil=4, Vel-50, Del=225, Pull=150. The tip of the micropipette was then cut off using forceps. This was to ensure the right diameter that would allow for the passage of the viscous dsRNA with minimal damage to the mosquito. It was observed that when micropipette were too thin, it was difficult to puncture the thorax and the dsRNA did not flow easily through the pipette. And when the diameter is too wide, the mortality rate increased due to the impact of injection. The capillaries tubes were filled with mineral oil and fixed to the Nanoject II Auto-Liter micro-injector (Drummer scientific, USA) as instructed in the manual.

dsRNA delivery

The anaesthetized mosquitoes were picked and kept on the prepared cold block and injected with 69nl of 1500ng/µl or 3000ng/µl dsRNA. Normal saline or LacZ were used as control. Male or female (depending on experiments) mosquitoes were injected with double stranded RNA synthesised for the genes of interest 2 days after emergence. After injections, mosquitoes were transferred into the prepared cages and a wet roll used to cover the cage to

keep the conditions humid. They were left for 2 days before mating assays was carried out with five days old mosquitoes.

Behavioural assay

Mating assays were done as described in (Aboagye-Antwi *et al*., 2015). They were carried out to assess the level of assortative mating between the sibling species when RNAi was employed. Females were given a choice of mates of their own kind and those of another kind. This was done to test the effect of assortative mating on both strains. The experiment was carried for injected females and injected males as well. During this experiment, it was key that all strains are raised in the same conditions to prevent any variations. Mosquitoes of same age were used. Mating cages were created at 6 pm at dusk. This was to mimic what happens in nature. The mating cages comprised of thirty females and 60 males. Thirty females of same species were put in the cage with 30 males of same species and 30 of the other sibling species. For instance, Thirty *An. coluzzii* females were put in a cage, 30 *An. gambiae* males and 30 *An. coluzzii* males were added to the cage. Therefore, mating cages comprised of 90 individuals. Males were put into the cages before the females were introduced. This was also to mimic swarming, because males create swarms and females fly into such swarms to copulate. Mating cages were left for 24 hours. After 24 hours, all individuals in the cage were collected transferred into a tube containing 70% ethanol for preservation until dissection. Sperm bundle dissection was carried out as explained in Chapter 3. This was to check for assortative mating. The sperm bundle was dissected from females and DNA was extracted from the sperm. PCR was performed for species identification (Chapter 2).

Size determination

In order to ascertain that there was no variation between the two species used for mating, wing measurements were taken from a subset of mosquitoes from the groups studied. Wings were cut and placed on a microscope slide and measured. Pictures were taken and a stage micrometre was used for calibration. Large variation in size could potentially affect female choice and result in non-equal opportunity to both species to be mated.

5.2.4 Determining the efficacy of dsRNA- mediated gene silencing

The double-stranded RNA (dsRNA) targeting the two genes of interest AGAP001009 and AGAP001326 were synthesized. Mosquitoes were bred using the standard breeding methods. Larvae were trayed out, fed appropriately and pupae were picked, separated into male and female and put in different emergence cages. One day and 3 day old *An. coluzzii* adult male and female mosquitoes were injected. Sixty males and 60 female mosquitoes were injected with dsRNA for each gene of interest. Normal saline was injected in 60 mosquitoes as positive control. These were put in respective cages after injection. Groups of 10 mosquitoes were removed from the cages to test for silencing efficacy. This was carried out using qRT-PCR. RNA was extracted at 24, 42 and 72 hours post injection. RNA was extracted from whole bodies as described in chapter 2. Collection of samples and RNA extraction was done at fixed times for all groups. This was to maintain consistency. RNA was stored at -80^{o}C until needed.

Quantitative Real-Time PCR (qRT-PCR)

Verification of the efficacy of RNAi was done using total RNA extracted from the whole body. The process was carried out as described in chapter 2 and 4. qRT-PCR was carried out using SYBR green. A 1000ng of RNA was used for cDNA synthesis for all samples and 40S

ribosomal protein S7 (AGAP010592) and 60S ribosomal protein L1 (AGAP004422) were used as positive controls to normalize the CT values.

Statistics

JMP software was used to perform all statistical analysis. All variables were checked for normality.

5.3 Results

To investigate the effect of RNAi on assortative mating, we measured the level of assortative mating by providing *An. coluzzii* females and *An. gambiae* females a choice of males of their own species and those of the sibling species. This was done in both directions and percentage assortative mating measured.

5.3.1 Microinjections and Survival of Adults

Preliminary injections were done to identify the best way of injecting the mosquitoes with minimal damage and mortality since mating requires mosquitoes that are fit and able to swarm. Several groups of mosquitoes were injection with normal saline and monitored for survival rates (Figure 5.3.1). Example presented below, no significant difference in survival rate was found between males and females, DF =1 F-value=0.02 p-value=0.889. A significant difference was found between survival rate and time of injection DF=2, F-value=13.37, P-value=0.001. In an attempt to increase survival rate, injections were done bearing in mind the tip of the capillary tube used for injection was the right size. Mating assays were carried out when about 70% post survival rate was attained consistently.

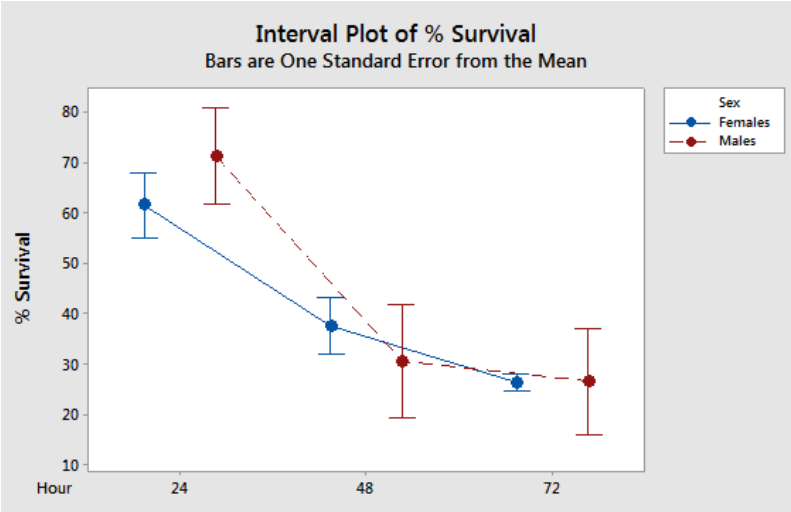

Figure 5.3.1: Post injection survival rate in males and female *Anopheles coluzzii.*

Injection was tried on both males and females and compared. Post injection survival rate was lower in males for majority of the experimental groups except a few exceptions where female survival rates were observed to be lower.

5.3.2 Correlation between number of females and percentage assortative mating

A nonparametric test was carried out using Spearman's *p* to check the correlation between the number of females and the percentage assortative mating (Figure 5.3.2). No correlation was observed (Spearman correlation: $p = -0.04$ and a $p > 0.85$).

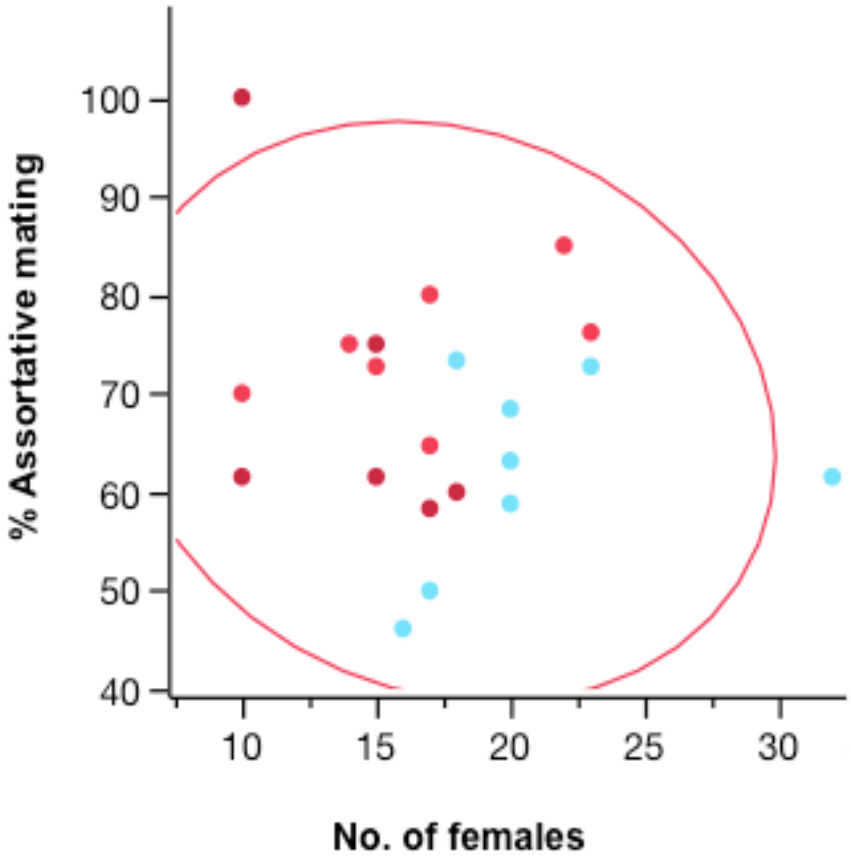

Figure 5.3.2: Correlation between the numbers of inseminated females and % assortative mating in experiments where *An. gambiae* or *An. coluzzii* females were given a choice of males of both species. Males were injected with 1500ng/µl (blue dots) in one set of experiments and 3000ng/µl (red dots).

5.3.3 Assortative mating in relation to female species and injected groups

The impact of dsRNA injection on assortative mating was studied to ascertain whether or not dsRNA injections of our target genes interfered with assortative mating. Two sets of controls were used, negative control which are the non-injected and those injected with normal saline were used as another set of control. There was no significant difference in percentage assortative mating in females of both species (Logistic regression Likelihood ratio: $df = 1$, $\chi^2 = 1.47$ and $p = 0.22$) and between the experimental groups (Likelihood ratio: $df = 3$, $\chi^2 = 2.07$, $p = 0.56$). Evidence of assortative mating was tested for *An. coluzzii* and *An. gambiae* females. Assortative mating was not disrupted in the dsRNA injected mosquitoes when *An. gambiae* females were provided options of *An. coluzzii* and *An. gambiae* males (Figure 5.3.3; Table 5.3.1). The proportion of females mating with males of their own species in the untreated control groups were 75% for non-injected and 59.46% for those injected with normal saline. The RNAi treated groups were 70.45% for AGAP001009 and 60% for AGAP013526.

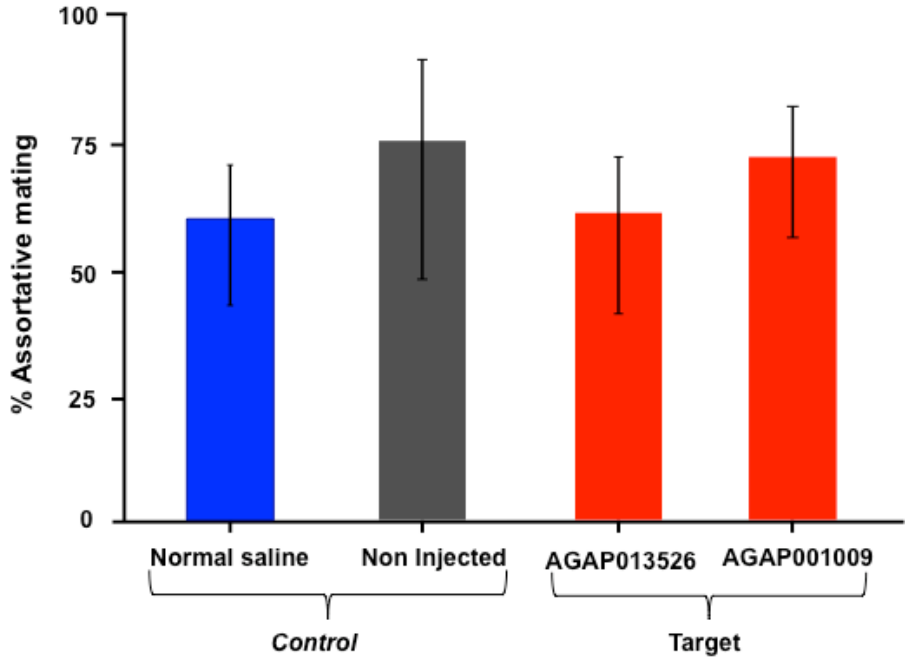

Figure 5.3.3: Impact of dsRNA for genes AGAP013526 and AGAP001009 on *An. gambiae* species.

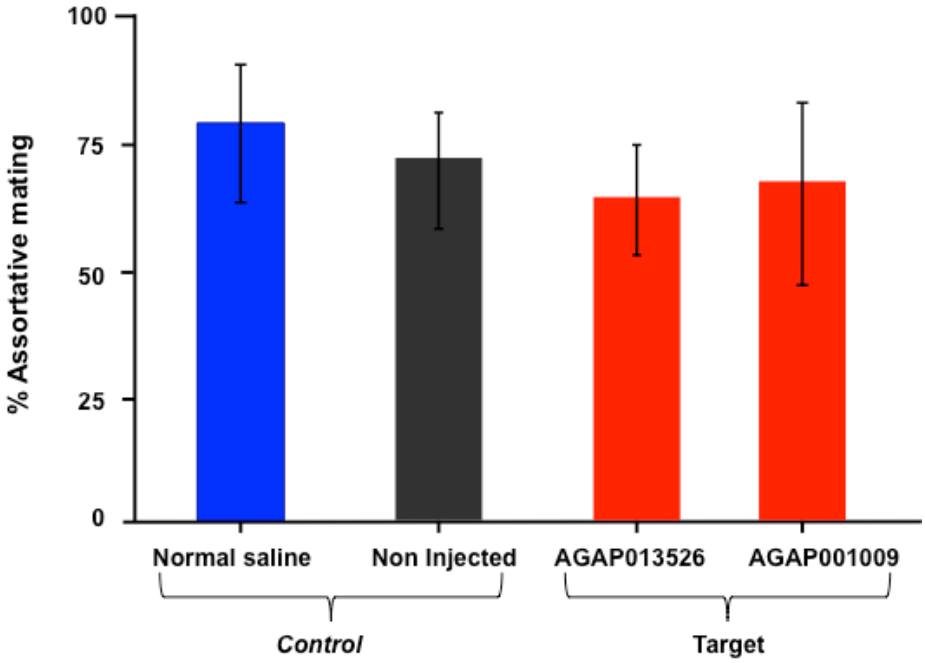

Figure 5.3.4: Impact of dsRNA injections for genes AGAP013526 and AGAP001009 on *An. coluzzii*.

Disruption of mating was also not observed in *An. coluzzii* females when provided with males of the two species (Figure 5.3.4; Table 5.3.2). The non-injected group had a 71.70%

assortative mating, and the controls injected with normal saline were 78.72%. The RNAi treated group were 65.28% and 68.42% for AGAP013526 and AGAP001009 respectively.

Table 5.3.1: Mean assortative mating in *Anopheles coluzzii* adult females.

Gene of Interest	Sample size	% Assortative mating	Confidence interval	Lower confidence interval	Upper confidence interval
AGAP001009	19	68.42	0.68	0.46	0.85
AGAP013526	72	65.28	0.65	0.54	0.75
Non-injected	53	71.70	0.72	0.58	0.82
Normal saline	47	78.72	0.79	0.65	0.88

Table 5.3.2: Mean assortative mating in *Anopheles gambiae* adult females.

Gene of Interest	Sample size	% Assortative mating	Confidence interval	Lower confidence interval	Upper confidence interval
AGAP001009	44	70.45	0.70	0.56	0.82
AGAP013526	35	60	0.6	0.44	0.74
Non-injected	12	75	0.75	0.47	0.91
Normal saline	37	59.46	0.59	0.43	0.74

5.3.4 Specific Gene Knockdown in adult mosquitoes

To investigate the efficacy of RNAi, we injected dsRNA for our target genes into adult male and female *An. coluzzii*. The expression levels of the targets, AGAP001009 and AGAP013526 were determined over a period of 72h. RNA was extracted from whole bodies and expression levels of the targets and controls, which were the Non-injected, and those injected with normal saline were determined using qRT-PCR. No striking reduction in transcript level was observed in both controls and target genes. dsRNA-injected mosquitoes did not display any reduction in AGAP001009 and AGAP013526 transcript levels. The mean

values for the non-injected, saline-injected and dsRNA injected did not differ for the various groups (Table 5.3.3; Table 5.3.5).

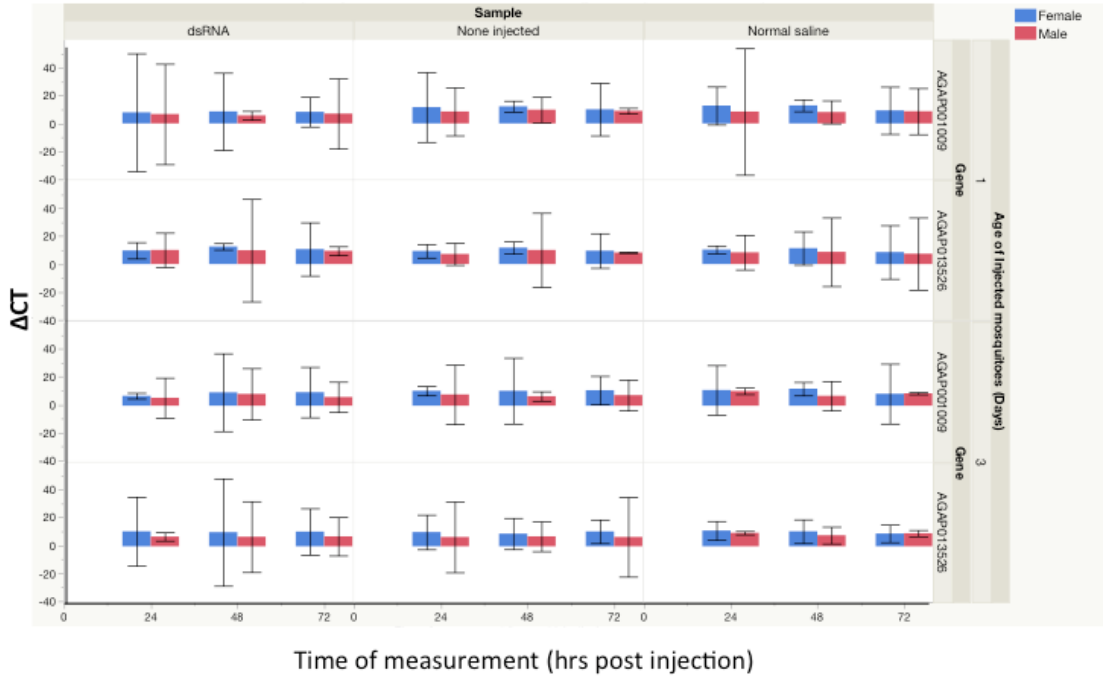

Figure 5.3.5: Mean ΔCT vs. time of measurement (Post injection hours).

Table 5.3.3: Mean of ΔCT for RNAi impact for gene AGAP001009

Sex	Age	Post Injection time	Non-Injected			Saline Injected			dsRNA injected		
			Mean	Lower CI	Upper CI	Mean	Lower CI	Upper CI	Mean	Lower CI	Upper CI
Females	1day	24h	11.73	-13.27	36.73	12.83	-0.81	26.47	8.02	-34.19	50.23
		48h	12.31	8.38	16.26	12.98	8.73	17.24	8.68	-18.98	36.33
		72h	10.2	-8.64	29.04	9.43	-7.51	26.37	8.38	-2.28	19.04
Females	3days	24h	10.77	7.43	14.1	11.3	-6.47	29.07	7.14	4.99	9.29
		48h	10.74	-12.85	34.33	12.2	7.39	17.01	9.61	-18.25	37.48
		72h	11.15	0.96	21.34	8.62	-12.83	30.08	9.72	-8.32	27.77
Males	1day	24h	8.61	-8.6	25.82	8.55	-36.76	53.89	6.79	-29.20	42.77
		48h	9.91	0.77	19.05	8.24	0.01	16.47	5.84	2.68	8.99
		72h	9.13	7.1	11.17	8.8	-7.73	25.33	7.13	-18.01	32.27
Males	3days	24h	8.1	-13.2	29.4	10.63	8.28	12.98	5.7	-8.65	20.05
		48h	6.89	3.51	10.27	7.14	-3.35	17.64	8.51	-9.86	26.88
		72h	7.6	-3.43	18.62	8.95	8.13	9.76	6.3	-4.53	17.14

* Ninety-five percent confidence intervals.

Table 5.3.4: Mean of ΔCT for RNAi impact for gene AGAP013526

Sex	Age	Post Injection time	Non-Injected			Saline Injected			dsRNA injected		
			Mean	Lower CI	Upper CI	Mean	Lower CI	Upper CI	Mean	Lower CI	Upper CI
Females	1day	24h	9.44	4.6	14.29	10.47	7.65	13.28	9.81	3.97	15.66
		48h	11.85	7.45	16.24	11.44	-0.52	23.39	12.64	10.28	15.01
		72h	9.66	-2.72	22.05	8.65	-10.53	27.84	10.77	-8.31	29.86
Females	3days	24h	10.26	-2.12	22.63	11.36	4.62	18.1	10.79	-13.90	35.48
		48h	9.13	-2.03	20.3	10.81	2.31	19.3	10.17	-28.26	48.61
		72h	10.71	2.26	19.16	9.22	2.93	15.5	10.58	-6.05	27.22
Males	1day	24h	7.29	-0.62	15.19	8.46	-3.95	20.87	10.16	-2.17	22.49
		48h	10.15	-16.41	36.72	8.79	-15.78	33.35	9.95	-26.78	46.68
		72h	8.32	7.91	8.73	7.48	-18.37	33.33	9.57	6.41	12.73
Males	3days	24h	6.73	-18.62	32.08	9.73	8.49	10.98	7.03	4.02	10.05
		48h	7.19	-3.5	17.89	7.97	1.91	14.04	6.92	-18.44	32.28
		72h	6.76	-21.77	35.29	9.42	7.14	11.71	7.18	-6.59	20.95

Ninety-five percent confidence intervals

5.4 Discussion

There are potential hurdles for novel malaria vector control methods. Amongst these hurdles is mating. Strong assortative mating exists between *An. coluzzii* and *An. gambiae* (Tripet *et al*., 2001). Assortative mating can lead to the splitting of a population even when natural selection is absent. If assortative mating on its own cannot cause the separation of species, it is enough to prevent gene flow between them (Kondrashov and Shpak, 1998). This can hinder the spread of transgenes, which is the concept in most genetic methods. Genetic control of vectors depend on mating (Alphey, 2014). One of the challenges of strong assortative mating on mosquito release programmes is that each new species would need to genetically modified to ensure mating and this may not be an easy task (Diabaté and Tripet, 2015). Interventions that focus on mosquito release programmes such as SIT, Wolbachia carrying males and males with the genetic drive mechanisms also depend heavily on the mating success of the transgenes (Alphey, 2014). Other potential vector control methods that depend on the knowledge mosquito mating behaviour are the use of chemical traps targeting cues used in swarming, contact pheromones, sound traps and visual cues (Diabaté and Tripet, 2015).

In this study, attempted gene knockdown experiments to investigate the role of genes located on the X Island of speciation and to determine whether these genes have any role in conspecific mating between *An. coluzzii* and *An. gambiae* was carried out. Identifying those genes will be very useful for the spread of transgenes; it could also lead to the exact phenotypic qualities involved in mate recognition and these phenotypes could be considered for the development of new vector control tools.

Genes with differential expression in males were of interest in this study because difference in the expression of chemosensory genes between males and females has been related to their role in mating (Pitts *et al*., 2011). Chemical ecology of mosquitoes is also an

area needing to be explored because a number of mosquito's behaviour is mediated by olfaction (Zwiebel and Takken, 2004).

In this study, microinjection was used for dsRNA delivery to attempt gene knockdown of targets after which behavioral assays were conducted. Adult *An. coluzzii* females were first injected with dsRNA but unfortunately mating was inconsistent between the control and dsRNA injected groups. Hence, it was concluded that the impact of injection on the females might have led to the inability of females to accurately choose males of their own kind and as a result, males were injected. Injection of males was quite challenging, as males are smaller than females. This was observed by measuring the dry weight mass of males and female *An. coluzzii*. The mortality rate was higher in males for majority of the experimental groups with a few exceptions. More injection time was needed for males to attain sufficient number for the mating crosses. Another constraint with injection of males was more numbers were needed since males of both sibling species were needed for the crosses because females were given a choice of males of both species.

The use of microinjection has been reported to be successful in a study where spermless *An. gambiae* s s males were developed using RNAi silencing of a germ cell differentiation gene through injection. The males mated successfully and females who mated with the spermless males showed normal postcopulatory responses. They observed that sperms are not needed to oviposit; they also reported that females mated by sperm less males resisted mating again (Thailayil *et al*., 2011). Microinjection for dsRNA delivery has also been reported to be successful and gene knockdown achieved in *An. gambiae* (Suzuki *et al*., 2014).

In this experiment, 30 females and 60 males i.e. 30 males of each sibling were used for mating assays. Assortative mating was inconsistent in the control and target groups. A likely explanation for this inconsistency could be mechanical damage caused my injections

therefore survivors were immobilized. The volume and site of injection can also affect the impact of injections (Scott *et al*., 2013). Male fitness and swarming are crucial for mating in *An. coluzzii* and *An. gambiae* (Baeshen *et al*., 2014), therefore in this study large numbers of mosquitoes were used to trigger mating by having more swarming males in mating cages. Forty females and 80 males were used so as to increase swarming activity as swarming is necessary for mating to occur (Diabate and Tripet, 2015). The increase in number of mosquitoes in mating cages did not improve mating neither was mating consistent. Mating assays were performed for the two candidate genes AGAP001009 and AGAP013526. Over 70% assortative mating was observed in the non-injected control group for *An. coluzzii* and *An. gambiae* females. However, there was a high rate of assortative mating in the normal saline injected group for *An. coluzzii* females and AGAP001009 injected group for *An. gambiae* females. Assortative mating was reduced in the normal saline group, and gene AGAP013526 injected group for *An. gambiae* females. The inconsistency of mating data shows that gene knockdown was not successful. The attempted knockdown of AGAP001009 and AGAP013526 seems to have failed. Gene silencing carried out by injecting dsRNA into male mosquitoes has also been reported to be unsuccessful in a behavioural studies carried out to investigate whether the quick- to-court protein (*qtc*) play a role in the regulation of sex-specific conflicting demand of mosquito courtship and blood feeding behaviour in the adult female mosquitoes (Das De *et al*., 2017). The failure was due to high mortality rate of injected mosquitoes. In the study, the quick- to-court protein (*qtc*) found to regulate courtship behaviour in adult drosophila was analysed in *An. culicifacies* (*Ac-qtc*). It was found that the level of expression was dependent on the maturation of the olfactory system. Possible function of male Ac-qtc in mating success was also investigated (Das De *et al*., 2017).

RT-qPCR was used to examine the efficacy of RNAi. Reduction in transcript level was quantified in dsRNA-injected male and female *An. coluzzii* species. The transcript levels

of the dsRNA-injected mosquitoes were compared with the control groups that were the non-injected and the normal saline injected mosquitoes. dsRNA-injected mosquitoes did not display any reduction in transcript level for both genes and for the different groups tested. RNAi could fail for several reasons one of which is the concentration of dsRNA used. In insects, RNAi can differ depending on tissue due to differences in dsRNA uptake however, most mosquitoes tissues are reached by injections though knockdown in the nervous system depends on the gene (Biessmann *et al*., 2010). A concentration of 69nl of 1500 ng/µl to 3000 ng/µl dsRNA was used for our knockdown experiments. These concentrations have been used with success in several studies. The required dose of RNAi molecules depends on the target gene. High concentrations of dsRNA are usually viscous and that can obstruct flow through the injections (Scott *et al*., 2013). Another factor that can affect RNAi success is the choice of sequence for dsRNA preparation (Scott *et al*., 2013). If the gene of interest had a known function, then the RT-qPCR to measure the success of reduction in transcript would be carried out using specific body parts e.g. the head, legs etc. The protein levels of the gene of interest could also be measured, as it is important to measure the relative protein concentration. This is because the effect of RNAi may not correlate to the level of transcript suppression (Scott *et al*., 2013).

In conclusion, the use of alternative methods of dsRNA delivery should be employed for behavioural studies. Though the use of injection for dsRNA delivery has been successful in mosquitoes (Campbell *et al*., 2008) and multiple genes have been targeted in *An. gambiae* at high frequency using the microinjection (Suzuki *et al*., 2014). Failed attempts due to injection impact have been reported (Das De *et al*., 2017). Major setbacks to this method includes, training time, injecting time, injecting at the right site, needle size, and dsRNA concentration (Kumar and Puttaraju, 2012). Injections sometimes result in poor reproducibility and inconsistency in results (Kumar *et al*., 2012). Microinjections are time

consuming and not suitable for high-throughput studies neither are they applicable for mosquito control (Gu *et al*., 2011). Alternative dsRNA techniques are the uptake or the ingestion of dsRNA. It has been shown to be effective in *Aedes* (Singh *et al*., 2013), and in *Culex* larvae (Lopez-Martinez *et al*., 2012). Oral delivery targets several insects at the same time and is devoid of mechanical harm. It is important as it can be used for insects that are not tolerant to injection (Scott *et al*., 2013). Oral delivery of dsRNA was used to target testis genes and female sex determination gene (doublesex) to inhibit sterility and female development in *Aedes aegypti*. Treated males were found to be less competitive for mates and fertility was reduced. The genes were found to be very effective in competing for mates. The gene knockdown in females led to a male-biased progeny of mosquitoes (Whyard *et al*., 2015). The study shows the possibility of feeding larvae to achieve a knockdown effect. Though knockdown of our candidate genes was not achieved, the data suggest that mechanical impact of injection has a negative effect on the mating behaviour of *An. coluzzii* and *An. gambiae*. Therefore, limitations can be closely looked at and other less invasive RNAi delivery methods could be implemented to achieve the desired outcome. The success of silencing those genes is very crucial because it could reveal the role of those genes in mating and ecological adaption. It will also lead to a better understanding of speciation.

Chapter 6 General discussion

The burden of malaria to man is enormous as it negatively affects every sphere of human life. The vectors *Anopheles* transmit the disease, their bites and the noise they make during the process causes discomfort to the human host. In endemic regions, people's way of life is altered as adjustments are made to prevent being bitten by mosquitoes. Success has been recorded in the fight against malaria due to the interventions available, which are mainly insecticide base; however, parasite resistance to drugs and vectors resistance to insecticides threaten this success. Hence, the need for new control measures against both parasites and vectors in order to tackle this menace.

New techniques considered for vector control are genetically modified mosquitoes (GMM), Sterile insect technique (SIT) and gene drive constructs. The success of these methods is highly dependent on the mosquito's mating ability. Any transgenes created would need to spread into the ecosystem and this can only be successful if we understand the vector's mating behaviour. *Anopheles gambiae* is currently considered the most deadly of the anthropophilic mosquito species (Potter, 2014) and the mating behaviour of members of this species is complex and yet to be properly understood. Strong assortative mating exists between *An. coluzzii* and *An. gambiae* the two recently diverged members of the *An. gambiae* complex (Tripet *et al*., 2001). This in addition to ecological and genetic differences is believed to have contributed to the ongoing speciation in the two sibling species. Speciation a process in which new species are formed from previously existing ones is detrimental to vector control as it gives rise to cryptic populations. These populations may differ in behaviour, ecology and their capacity as vectors (Fouet *et al*., 2017). This would mean studying many more species and designing vector control methods effective against each species. This would be an enormous if not impossible task. It is therefore important to understand the population structure and mechanism of differentiation among species.

Understanding the population structure of species is important for the development of large-scale vector control measures.

Apart from genetic manipulations, other novel methods considered for malaria vector control are the use of biocontrols, and the use of behavioural cues. Behaviours such as swarming, feeding and mating involve the use of cues. Visual, auditory and olfactory cues are considered possible areas to be harnessed for vector control. The success of these methods also depends on knowledge of the vectors behaviour, the genetic and ecological differences that exist between the different species.

In this thesis, molecular techniques and behavioural studies were explored to gain new insight and into the ecological and genetic differences that exist between the recently diverged members of the *Anopheles gambiae* complex namely *Anopheles coluzzii* and *Anopheles gambiae*.

The first study investigated ecological divergence between the recently diverged sibling species *An. coluzzii* and *An. gambiae*. The prediction was that ecological differences between the species would relate to difference in egg morphology between the species. Understanding the ecological factors aiding divergent selection between *An. coluzzii* and *An. gambiae* is complex because the different life stages inhabit different ecological settings and species ecological preferences does exist. Studying phenotypic variation between the sibling species at different life stages could help unravel factors involved in speciation (Lehmann *et al*., 2008). This study focused on differences in egg morphological characteristics between the two sibling species *An. coluzzii* and *An. gambiae*. Analysis was conducted using two approaches; traditional morphometrics and landmark geometric morphometrics analyses and these were useful in exploring possible differences in egg structure between the two species. The prediction made for the study was not completely met as eggs were found to be quite similar between the two sibling species though *An. coluzzii* eggs were observed have slightly

larger egg width compared to *An. gambiae*. The landmark geometric morphometry also showed shape differences along the egg mid-deck region corresponding to the difference in egg width observed between the two species using the traditional morphometrics.

There are different potential reasons for the differences found in the egg structure. The first could be as a result of *An. coluzzii* ability to withstand drought conditions (Hidalgo *et al*., 2015). *An. coluzzii* strives during the dry season. The maximal survival in *An. coluzzii* is usually increased from 4 weeks to 7months (Yaro *et al*. 2012). Aestivation is said to account for the increase in survival rate. According to Yaro *et al* (2012), oviposition response decreased from 70 to 20% during the dry season, mean egg batch is also reduced and gonotrophic dissociation in females increased during the dry season leading to depressed reproduction (Yaro *et al*. 2012).

In addition to the survival strategies mention above, larger egg size in *An. coluzzii* could translate to larger adults that are able to withstand drought conditions. The difference in egg size could a result of differences in larval adaptation. *An. gambiae* strives in rain pools while *An. coluzzii* prefers permanent water bodies (Mattah *et al*., 2017). These permanent water bodies are prone to high level of predation and *An. coluzzii* is known to be better at avoiding predation compared to *An. gambiae* (Diabaté *et al*., 2008). *An. gambiae* have been reported to suffer high predation rates when exposed to predation (Diabaté *et al*., 2008). Behavioural plasticity caused by exposure to predation has been reported in *An. coluzzii* (Roux *et al*., 2013) a reason they are able to inhabit different habitat. A study conducted by Gimonneau *et al* (2014) reported a high level of larval fitness in *Anopheles coluzzii* compared to *An gambiae* when both species were subjected to competition in different environmental conditions (Gimonneau *et al*., 2014). Our findings could have an implication on predation avoidance behaviour between the two species. Bigger egg size in *An. coluzzii* could be an adaptive feature leading to faster development to adult stage in other to escape predation.

Apart from shape differences observed between species, comparison between populations revealed similarities between the Kisumu and Mopti eggs although they are populations of different species. These similarities potentially indicates that speciation between the two sibling species is not complete as there is a considerable amount of gene flow between the species (Della Torre *et al*., 2002; Lee *et al*., 2013b).

There are several studies comparing between the two sibling species, however, the studies fail to lay emphasis on difference between populations within species (Lehmann *et al*., 2008). In this study, differences and similarities were found between populations. This just indicates that speciation between the sibling species is incomplete.

The second study, evaluated genetic differences between the two sibling species. We speculated that differentially expressed genes on the X-island of speciation could harbour genes responsible for assortative mating and ecological divergence. The island is hypothesised to harbour genes involved in speciation (Turner *et al*., 2005). In order to narrow in on candidate mate choice genes, 27 putative mating genes located on the X-island of speciation were investigated. Twelve genes with fixed protein coding (>95%) and 15 genes with frequent (>80%) protein coding changes between *Anopheles coluzzii* and *Anopheles gambiae* (Aboagye-Antwi *et al*., 2015).

Quantitative real time PCR was used to study differential gene expression between species, populations within species and life stages. The majority of the genes were overexpressed in males. An upregulation of genes in males could suggest the roles of those genes in mate recognition. Some genes were overexpressed in virgin males and downregulated at the larval stage. This could suggest a transcriptional turnover of male biased genes during the developmental stages (Magnusson *et al*., 2011). In this study it was presumed that most genes with fixed protein coding changes between the two species would be differentially expressed compared to genes with nearly fixed protein coding changes.

Surprisingly, genes with nearly fixed protein coding changes were also differentially expressed at equal proportion as transcripts with fixed protein-coding changes. This finding suggests that genes with frequent protein coding changes are equally worth studying, as they could also be involved in speciation. Unravelling the role of the genes could lead to the identification phenotypic characteristics involved in mate recognition and ecological adaptation.

Finally, the third study (chapter 5) presented in this thesis was an attempt at gene knockdown of selected top two candidate putative assortative genes. The top candidate genes were identified using different filters. Several behavioural assays were carried out to ascertain the impact of the dsRNA delivery and gene silencing on assortative mating in the two sibling species. Injection was used as the method for dsRNA delivery. This method has been reported to be successful in several studies on mosquitoes (Pillai *et al*., 2017). However our data showed disruption of mating even in the controlled groups. Possible explanation to this outcome is that injections into the haemocoel of individual mosquitoes as the method of dsRNA delivery might not have been ideal for this study. The mechanical impact of the injection could have had an effect on the mosquitoes thereby affecting their ability to mate. *Anopheles* mating behaviour is complex as it involves the use of cues which requires mosquitoes that are fit (Diabate and Tripet, 2015). If injections on its own could potentially alter the mosquito's behaviour, then mate recognition and ultimately mating is altered. The injection technique is was time consuming as only one mosquito is injected at a time. A high throughput technique where many mosquitoes can be affected at a time is ideal, as mating cages require a certain number of mosquitoes. A high throughput technique will reduce the dsRNA delivery time thereby mating assays can be carried out more efficiently. Overall, there was no clear pattern of mating in the dsRNA injected mosquitoes and control group. We therefore suggest the use of alternative methods of dsRNA delivery for behavioural studies.

In conclusion, speciation cannot be overlooked in the fight against malaria. The differences that exist between various vector species are numerous and unless we understand them, potential control measures might be compromised. Knowledge of the genetic and ecological difference that exist between *An. coluzzii* and *An. gambiae* will not only broaden our understanding of speciation, it will aid in the proper implementation of vector control measures focusing on the use of transgenes. It could also lead to the development of novel vector control tools taking advantage of the vectors preferences and behaviours. Genetic studies could lead to the identification of specific phenotypes associated with certain behaviours and this can be manipulated to achieve an effective spread of transgenes during vector control. It could also lead to the identification of new cues associated with mating and ecological adaptation.

This thesis hopes to contribute to the fight against malaria by adding to our knowledge of the ongoing speciation process in the most important vectors of malaria *An. coluzzii* and *An. gambiae*. The result of this thesis demonstrates that even though *An. coluzzii* and *An. gambiae* are very similar, there are some differences that exist between them, which can be studied extensively and harnessed for vector control.

For future research, it would be essential to explore the feasibility of using mate recognition cues for vector control. Other dsRNA delivery methods can be explored for mosquito behavioural studies.

References

Aboagye-Antwi, F., Alhafez, N., Weedall, G.D., Brothwood, J., Kandola, S., Paton, D., Fofana, A., Olohan, L., Betancourth, M.P., Ekechukwu, N.E., Baeshen, R., Traorè, S.F., Diabate, A., Tripet, F., 2015. Experimental Swap of Anopheles gambiae's Assortative Mating Preferences Demonstrates Key Role of X-Chromosome Divergence Island in Incipient Sympatric Speciation. PLoS Genet. 11, 1–19.

Aboagye-Antwi, F., Tripet, F., 2010. Effects of larval growth condition and water availability on desiccation resistance and its physiological basis in adult Anopheles gambiae sensu stricto. Malar. J. 9, 225.

Achee, N.L., Bangs, M.J., Farlow, R., Killeen, G.F., Lindsay, S., Logan, J.G., Moore, S.J., Rowland, M., Sweeney, K., Torr, S.J., Zwiebel, L.J., Grieco, J.P., 2012. Spatial repellents: From discovery and development to evidence-based validation. Malar. J. 11, 1–9.

Airs, P., Bartholomay, L., 2017. RNA Interference for Mosquito and Mosquito-Borne Disease Control. Insects 8, 4.

Aldersley, A., Champneys, A., Homer, M., Robert, D., 2016. Quantitative analysis of harmonic convergence in mosquito auditory interactions. J. R. Soc. Interface 13, 20151007.

Alonso, P.L., Brown, G., Arevalo-Herrera, M., Binka, F., Chitnis, C., Collins, F., Doumbo, O.K., Greenwood, B., Fenton Hall, B., Levine, M.M., Mendis, K., Newman, R.D., Plowe, C. V, Rodríguez, M.H., Sinden, R., Slutsker, L., Tanner, M., 2011. A Research Agenda to Underpin Malaria Eradication. PLoS Med. 8, e1000406.

Alphey, L., 2002. Re-engineering the sterile insect technique. Insect Biochem. Mol. Biol. 32, 1243–1247.

Alphey, L., 2014. Genetic Control of Mosquitoes.

Alphey, L., Benedict, M., Bellini, R., Clark, G.G., Dame, D.A., Service, M.W., Dobson, S.L., 2010. Sterile-Insect Methods for Control of Mosquito-Borne Diseases: An Analysis.

Anguelov, R., Dumont, Y., Lubuma, J., 2012. Mathematical modeling of sterile insect technology for control of anopheles mosquito. Comput. Math. with Appl. 64, 374–389.

Asmare, Y., Hill, S.R., Hopkins, R.J., Tekie, H., Ignell, R., 2017. The role of grass volatiles on oviposition site selection by Anopheles arabiensis and Anopheles coluzzii. Malar. J. 16, 65.

Athrey, G., Cosme, L. V, Popkin-Hall, Z., Pathikonda, S., Takken, W., Slotman, M.A., 2017. Chemosensory gene expression in olfactory organs of the anthropophilic Anopheles coluzzii and zoophilic Anopheles quadriannulatus. BMC Genomics 18.

Aytekin, S., Aytekin, A.M., Alten, B., 2009. Effect of different larval rearing temperatures on the productivity (Ro) and morphology of the malaria vector Anopheles superpictus Grassi (Diptera: Culicidae) using geometric morphometrics. J. Vector Ecol. 34, 32–42.

Baeshen, R., Ekechukwu, N.E., Toure, M., Paton, D., Coulibaly, M., Traoré, S.F., Tripet, F., 2014. Differential effects of inbreeding and selection on male reproductive phenotype associated with the colonization and laboratory maintenance of Anopheles gambiae. Malar. J. 13, 19.

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

Banks, S.D., Murray, N., Wilder-Smith, A., Logan, J.G., 2014. Insecticide-treated clothes for the control of vector-borne diseases: a review on effectiveness and safety. Med. Vet. Entomol. 28, 14–25.

Barreaux, P., Barreaux, A.M.G., Sternberg, E.D., Suh, E., Waite, J.L., Whitehead, S.A.,

Thomas, M.B., 2017. Priorities for Broadening the Malaria Vector Control Tool Kit. Trends Parasitol. 33, 763–774.
Bates, M., 1940. The Nomenclature and Taxonomic Status of the Mosquitoes of the Anopheles Maculipennis Complex1. Ann. Entomol. Soc. Am. 33, 343–356.
Beard, J., 2006. DDT and human health. Sci. Total Environ. 355, 78–89.
Bell, D., Wongsrichanalai, C., Barnwell, J.W., 2006. Ensuring quality and access for malaria diagnosis: How can it be achieved? Nat. Rev. Microbiol. 4, 682–695.
Bellés, X., 2010. Downloaded from www.annualreviews.org Access provided by University of Keele on 07/23/18. For personal use only. Annu. Rev. Entomol 55, 111–128.
Benedict, M.Q., 1997. Care and maintenance of anopheline mosquito colonies. In: The Molecular Biology of Insect Disease Vectors. Springer Netherlands, Dordrecht.
Benelli, G., Beier, J.C., 2017. Current vector control challenges in the fight against malaria. Acta Trop. 174, 91–96.
Benelli, G., Mehlhorn, H., 2016. Declining malaria, rising of dengue and Zika virus: insights for mosquito vector control. Parasitol. Res. 115, 1747–1754.
Bentley M.D., Day, J.F., 1989. Chemical ecology and behavioral aspects of mosquito oviposition. Ann. Rev. Entomol. 34, 401-421.
Bernstein, R., Bernstein, S., 2003. Assortative mating in insects: sexual selection vs. sampling artifact. West. North Am. Nat. 63, 374–381.
Biessmann, H., Andronopoulou, E., Biessmann, M.R., Douris, V., Dimitratos, S.D., Eliopoulos, E., Guerin, P.M., Iatrou, K., Justice, R.W., Kröber, T., Marinotti, O., Tsitoura, P., Woods, D.F., Walter, M.F., 2010. The Anopheles gambiae Odorant Binding Protein 1 (AgamOBP1) Mediates Indole Recognition in the Antennae of Female Mosquitoes. PLoS One 5, e9471.
Bir, A., Sidhu, S., Verdier-pinard, D., Fidock, D.A., 2002. Chloroquine Resistance in Plasmodium falciparum Malaria Parasites Conferred by pfcrt Mutations 298.
Blandin, S., Moita, L.F., Köcher, T., Wilm, M., Kafatos, F.C., Levashina, E.A., Cornel, A.J., 2002. Reverse genetics in the mosquito Anopheles gambiae: targeted disruption of the Defensin gene. EMBO Rep. 3, 852–856.
Blaustein, L., Kiflawi, M., Eitam, A., Mangel, M., Cohen, J.E., 2004. Oviposition habitat selection in response to risk of predation in temporary pools: mode of detection and consistency across experimental venue. Oecologia 138, 300–305.
Bloem, S., Carpenter, J., Mccluskey, A., Fugger, R., Arthur, S., Wood, S., 2007. Suppression of the Codling Moth Cydia pomonella in British Columbia, Canada Using an Area-Wide Integrated Approach with an SIT Components. Area-Wide Control Insect Pests From Res. to F. Implement. 591–601.
Bojang, K.A., Milligan, P.J., Pinder, M., Vigneron, L., Alloueche, A., Kester, K.E., Ballou, W.R., Conway, D.J., Reece, W.H., Gothard, P., others, 2001. Efficacy of RTS, S/AS02 malaria vaccine against Plasmodium falciparum infection in semi-immune adult men in The Gambia: a randomised trial. Lancet 358, 1927–1934.
Bowatte, G., Perera, P., Senevirathne, G., Meegaskumbura, S., Meegaskumbura, M., 2013. Tadpoles as dengue mosquito (Aedes aegypti) egg predators. Biol. Control 67, 469–474.
Brown, S.J., Mahaffey, J.P., Lorenzen, M.D., Denell, R.E., Mahaffey, J.W., 1999. Using RNAi to investigate orthologous homeotic gene function during development of distantly related insects. Evol. Dev. 1, 11–15.
Cabrera, M., Jaffe, K., 2007. An aggregation pheromone modulates lekking behavior in the vector mosquito Aedes aegypti (Diptera: Culicidae). J. Am. Mosq. Control Assoc. 23, 1–10.
Campbell, C.L., Keene, K.M., Brackney, D.E., Olson, K.E., Blair, C.D., Wilusz, J., Foy, B.D., 2008. Aedes aegypti uses RNA interference in defense against Sindbis virus

infection.
Carey, A.F., Carlson, J.R., 2011. Insect olfaction from model systems to disease control. Proc. Natl. Acad. Sci. U. S. A. 108, 12987–95.
Carter, R., Mendis, K.N., 2002. Evolutionary and Historical Aspects of the Burden of Malaria Evolutionary and Historical Aspects of the Burden of Malaria. Clin. Microbiol. Rev. 15.
Cassone, B.J., Kamdem, C., Cheng, C., Tan, J.C., Hahn, M.W., Costantini, C., Besansky, N.J., 2014. Gene expression divergence between malaria vector sibling species Anopheles gambiae and An. coluzzii from rural and urban Yaoundé Cameroon. Mol. Ecol. 23, 2242–2259.
Cassone, B.J., Mouline, K., Hahn, M.W., White, B.J., Pombi, M., Simard, F., Costantini, C., Besansky, N.J., 2008. Differential gene expression in incipient species of Anopheles gambiae. Mol. Ecol. 17, 2491–504.
Cator, L.J., Harrington, L.C., 2011. The harmonic convergence of fathers predicts the mating success of sons in Aedes aegypti. Anim. Behav. 82, 627–633.
Cator, L.J., Ng'Habi, K.R., Hoy, R.R., Harrington, L.C., 2010. Sizing up a mate: Variation in production and response to acoustic signals in Anopheles gambiae. Behav. Ecol. 21, 1033–1039.
Cator, L.J., Zanti, Z., 2016. Size, sounds and sex: interactions between body size and harmonic convergence signals determine mating success in Aedes aegypti. Parasit. Vectors 9, 622.
Catteruccia, F., 2007. Malaria vector control in the third millennium: progress and perspectives of molecular approaches. Pest Manag. Sci. 63, 634–40.
Charlwood, J.D., Pinto, J., Sousa, C.A., Madsen, H., Ferreira, C., Do Rosario, V.E., 2002. The swarming and mating behaviour of Anopheles gambiae s.s. (Diptera: Culicidae) from São Tomé Island, Journal of Vector Ecology.
Charlwood, J.D., Vij, R., Billingsley, P.F., 2000. Dry season refugia of malaria-transmitting mosquitoes in a dry savannah zone of east Africa, Am. J. Trop. Med. Hyg.
Choi, K.S., Townson, H., 2012. Evidence for X-linked introgression between molecular forms of Anopheles gambiae from Angola. Med. Vet. Entomol. 26, 218–27.
Clements, A.N. (Alan N., Kerkut, G.A., 2013. The physiology of mosquitoes. Elsevier Science.
Coetzee, M., Craig, M., le Sueur, D., 2000. Distribution of African Malaria Mosquitoes Belonging to the Anopheles gambiae Complex. Parasitol. Today 16, 74–77.
Coetzee, M., Hunt, R.H., Wilkerson, R., Della Torre, A., Coulibaly, M.B., Besansky, N.J., 2013. Anopheles coluzzii and Anopheles amharicus, new members of the Anopheles gambiae complex. Zootaxa 3619, 246–74.
Collin, H., Fumagalli, L., 2011. Evidence for morphological and adaptive genetic divergence between lake and stream habitats in European minnows (Phoxinus phoxinus, Cyprinidae). Mol. Ecol. 20, 4490–4502.
Coluzzi, M., 1964. Morphological divergences in the Anopheles Gambiae Giles complex.
Coluzzi, M., Petrarca, V., di Deco, M. a., 1985. Chromosomal inversion intergradation and incipient speciation in Anopheles gambiae. Bolletino di Zool. 52, 45–63.
Coluzzi, M., Sabatini, A., 1967. Cytogenetic observations on species A and B of the Anopheles gambiae complex. Parassitologia 9, 73–88.
Coluzzi, M., Sabatini, A., 1969. Cytogenetic observations on the salt water species, Anopheles merus and Anopheles melas, of the gambiae complex. Parassitologia 11, 177–87.
Coluzzi, M., Sabatini, A., della Torre, A., Di Deco, M.A., Petrarca, V., 2002. A polytene chromosome analysis of the Anopheles gambiae species complex. Science 298, 1415–8.
Coluzzi, M., Sabatini, A., Petrarca, V., Di Deco, M.A., 1979. Chromosomal differentiation

and adaptation to human environments in the Anopheles gambiae complex. Trans. R. Soc. Trop. Med. Hyg. 73, 483–497.
Cook, P.E., Sinkins, S.P., 2010. Transcriptional profiling of Anopheles gambiae mosquitoes for adult age estimation 19, 745–751.
Costantini, C., Ayala, D., Guelbeogo, W.M., Pombi, M., Some, C.Y., Bassole, I.H., Ose, K., Fotsing, J.-M., Sagnon, N., Fontenille, D., Besansky, N.J., Simard, F., 2009. Living at the edge: biogeographic patterns of habitat segregation conform to speciation by niche expansion in Anopheles gambiae. BMC Ecol. 9, 16.
Coulson, R.M.R., Hall, N., Ouzounis, C.A., 2004. Comparative genomics of transcriptional control in the human malaria parasite Plasmodium falciparum. Genome Res. 14, 1548–54.
Cox, F.E.G., 2010. History of the discovery of the malaria parasites and their vectors 1–9.
Crawford, J., Riehle, M.M., Guelbeogo, W.M., Gneme, A., Sagnon, N., Vernick, K.D., Nielsen, R., Lazzaro, B.P., 2015. Reticulate Speciation and Barriers to Introgression in the Anopheles gambiae Species Complex. Genome Biol. Evol. 7, 3116–3131.
Dahan, Y.L., Koekemoer, L.L., 2014. Analysis of the genitalia rotation in the male Anopheles funestus (Diptera: Culicidae). Acta Trop. 132, S20–S25.
Damrongphol, P., Baimai, V., 1989. Scanning electron microscopic observations and differentiation of eggs of the Anopheles dirus complex. J. Am. Mosq. Control Assoc. 5, 563–8.
Dana, A.N., Hillenmeyer, M.E., Lobo, N.F., Kern, M.K., Romans, P.A., Collins, F.H., 2006. Differential gene expression in abdomens of the malaria vector mosquito, Anopheles gambiae, after sugar feeding, blood feeding and Plasmodium berghei infection. BMC Genomics 7, 119.
Das De, T., Sharma, P., Rawal, C., Kumari, S., Tavetiya, S., Yadav, J., Hasija, Y., Dixit, R., 2017. Sex specific molecular responses of quick-to-court protein in Indian malarial vector Anopheles culicifacies: conflict of mating versus blood feeding behaviour. Heliyon 3, 361.
Davidson, G., 1964. Anopheles gambiae, a complex of species. Bull. World Health Organ. 31, 625–34.
Davidson, G., Jackson, C., 1962. Incipient speciation in Anopheles gambiae Giles. Bull. World Health Organization. 303–305.
Della Torre, A., Costantini, C., Besansky, N.J., Caccone, A., Petrarca, V., Powell, J.R., Coluzzi, M., 2002. Speciation Within Anopheles gambiae-the Glass Is Half Full. Science 298, 115–7.
Della Torre, A., Tu, Z., Petrarca, V., 2005. On the distribution and genetic differentiation of Anopheles gambiae s.s. molecular forms. Insect Biochem. Mol. Biol. 35, 755–769.
Della Torre, A., Fanello, C., Akogbeto, M., Dossou-yovo, J., Favia, G., Petrarca, V., Coluzzi, M., 2001. Molecular evidence of incipient speciation within Anopheles gambiae s.s. in West Africa. Insect Mol. Biol. 10, 9–18.
Demari-Silva, B., Multini, L.C., Suesdek, L., Oliveira, T.M.P., Sallum, M.A.M., Marrelli, M.T., 2017. Wing Morphometry and Genetic Variability Between Culex coronator and Culex usquatus (Diptera: Culicidae), Two Sibling Species of the Coronator Group. J. Med. Entomol. 54, 901–908.
Diabaté, A., Dabire, K.R., Kim, E.H., Dalton, R., Millogo, N., Baldet, T., Simard, F., Gimnig, J.E., Hawley, W., Lehmann, T., 2005. Larval development of the molecular forms of Anopheles gambiae (Diptera: Culicidae) in different habitats: a transplantation experiment. J. Med. Entomol. 42, 548–553.
Diabaté, A., Dabiré, R.K., Heidenberger, K., Crawford, J., Lamp, W.O., Culler, L.E., Lehmann, T., 2008. Evidence for divergent selection between the molecular forms of

Anopheles gambiae: role of predation. BMC Evol. Biol. 8, 5.
Diabaté, A., Dabire, R.K., Kengne, P., Brengues, C., Baldet, T., Ouari, A., Simard, F., Lehmann, T., 2006. Mixed Swarms of the Molecular M and S Forms of Anopheles gambiae (Diptera: Culicidae) in Sympatric Area from Burkina Faso. J. Med. Entomol. 43, 480–483.
Diabaté, A., Dabire, R.K., Millogo, N., Lehmann, T., 2007. Evaluating the effect of postmating isolation between molecular forms of Anopheles gambiae (Diptera: Culicidae). J. Med. Entomol. 44, 60–4.
Diabaté, A., Dao, A., Yaro, A.S., Adamou, A., Gonzalez, R., Manoukis, N.C., Traoré, S.F., Gwadz, R.W., Lehmann, T., 2009. Spatial swarm segregation and reproductive isolation between the molecular forms of Anopheles gambiae. Proc. R. Soc. B Biol. Sci. 276, 4215–4222.
Diabaté, A., Tripet, F., 2015. Targeting male mosquito mating behaviour for malaria control. Parasites and Vectors 8, 1–46.
Diabaté, A., Yaro, A.S., Dao, A., Diallo, M., Huestis, D.L., Lehmann, T., 2011. Spatial distribution and male mating success of Anopheles gambiaeswarms. BMC Evol. Biol. 11, 184.
Eastman, R.T., Fidock, D.A., 2009. Artemisinin-based combination therapies: a vital tool in efforts to eliminate malaria. Nat. Rev. Microbiol. 7, 864–874.
Ekechukwu, N.E., Baeshen, R., Traorè, S.F., Coulibaly, M., Diabate, A., Catteruccia, F., Tripet, F., 2015. Heterosis Increases Fertility, Fecundity, and Survival of Laboratory-Produced F1 Hybrid Males of the Malaria Mosquito Anopheles coluzzii. G3 Genes Genomes Genet. 5, 2693–709.
Endersby, N.M., White, V.L., Chan, J., Hurst, T., Rašić, G., Miller, A., Hoffmann, A.A., 2013. Evidence of cryptic genetic lineages within Aedes notoscriptus (Skuse). Infect. Genet. Evol. 18, 191–201.
Esnault, C., Boulesteix, M., Duchemin, J.B., Koffi, A. a, Chandre, F., Dabiré, R., Robert, V., Simard, F., Tripet, F., Donnelly, M.J., Fontenille, D., Biémont, C., 2008. High genetic differentiation between the M and S molecular forms of Anopheles gambiae in Africa. PLoS One 3, e1968.
Facchinelli, L., Valerio, L., Lees, R.S., Oliva, C.F., Persampieri, T., Collins, C.M., Crisanti, A., Spaccapelo, R., Benedict, M.Q., 2015. Stimulating Anopheles gambiae swarms in the laboratory: application for behavioural and fitness studies. Malar. J. 14, 271.
Fanello, C., Santolamazza, F., Della Torre, A., 2002. Simultaneous identification of species and molecular forms of the Anopheles gambiae complex by PCR-RFLP. Med. Vet. Entomol. 16, 461–464.
Farnesi, L.C., Figueiredo, R., Menna-Barreto, S., Martins, J., Valle, D., Lazzaro Rezende, G., 2015. Physical features and chitin content of eggs from the mosquito vectors Aedes aegypti, Anopheles aquasalis and Culex quinquefasciatus: Connection with distinct levels of resistance to desiccation. J. Insect Physiol. 83, 43–52.
Faull, K.J., Williams, C.R., 2016. Differentiation of Aedes aegypti and Aedes notoscriptus (Diptera: Culicidae) eggs using scanning electron microscopy. Arthropod Struct. Dev. 45, 273–280.
Favia, G., della Torre, A., Bagayoko, M., Lanfrancotti, A., Sagnon, N., Touré, Y.T., Coluzzi, M., 1997. Molecular identification of sympatric chromosomal forms of Anopheles gambiae and further evidence of their reproductive isolation. Insect Mol. Biol. 6, 377–383.
Favia, G., Louis, C., 1999. Molecular identification of chromosomal forms of Anopheles gambiae sensu stricto. Parassitologia 41, 115–8.
Fawaz, E.Y., Allan, S.A., Bernier, U.R., Obenauer, P.J., Diclaro, J.W., 2014. Swarming

mechanisms in the yellow fever mosquito: aggregation pheromones are involved in the mating behavior of Aedes aegypti. J. Vector Ecol. 39, 347–354.
Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E., Mello, C.C., 1998. Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature 391, 806–811.
Fontaine, M.C., Pease, J.B., Steele, A., Waterhouse, R.M., Neafsey, D.E., Sharakhov, I. V, Jiang, X., Hall, A.B., Catteruccia, F., Kakani, E., Mitchell, S.N., Wu, Y.-C., Smith, H.A., Love, R.R., Lawniczak, M.K., Slotman, M.A., Emrich, S.J., Hahn, M.W., Besansky, N.J., 2015. Mosquito genomics. Extensive introgression in a malaria vector species complex revealed by phylogenomics. Science 347, 1258524.
Fouet, C., Kamdem, C., Gamez, S., White, B.J., 2017. Genomic insights into adaptive divergence and speciation among malaria vectors of the Anopheles nili group 897–906.
Gabrieli, P., Smidler, A., Catteruccia, F., 2014. Engineering the control of mosquito-borne infectious diseases. Genome Biol. 15, 535.
Ghosh, A., Chowdhury, N., Chandra, G., 2012. Plant extracts as potential mosquito larvicides. Indian J. Med. Res. 135, 581–98.
Gimonneau, G., Bouyer, J., Morand, S., Besansky, N.J., Diabate, A., Simard, F., 2010. A behavioral mechanism underlying ecological divergence in the malaria mosquito Anopheles gambiae. Behav. Ecol. 21, 1087–1092.
Gimonneau, G., Brossette, L., Mamaï, W., Dabiré, R.K., Simard, F., 2014. Larval competition between An. coluzzii and An. gambiae in insectary and semi-field conditions in Burkina Faso. Acta Trop. 130, 155–61.
Gonzalez, P. V, Audino, P.A.G., Masuh, H.M., 2016. Oviposition behavior in Aedes aegypti and Aedes albopictus (Diptera: Culicidae) in response to the presence of heterospecific and conspecific larvae. J. Med. Entomol. 53, 268–272.
Gosling, R., Seidlein, L. Von, 2016. The Future of the RTS , S / AS01 Malaria Vaccine : An Alternative Development Plan 1–6.
Gu, J., Liu, M., Deng, Y., Peng, H., Chen, X., 2011. Development of an Efficient Recombinant Mosquito Densovirus-Mediated RNA Interference System and Its Preliminary Application in Mosquito Control. PLoS One 6, e21329.
Guyant, P., Corbel, V., Guérin, P.J., Lautissier, A., Nosten, F., Boyer, S., Coosemans, M., 2015. Past and new challenges for malaria control and elimination: the role of operational research for innovation in designing interventions. Malar. J. 14.
Haerty, W., Jagadeeshan, S., Kulathinal, R.J., Wong, A., Ravi Ram, K., Sirot, L.K., Levesque, L., Artieri, C.G., Wolfner, M.F., Civetta, A., Singh, R.S., 2007. Evolution in the fast lane: rapidly evolving sex-related genes in Drosophila. Genetics 177, 1321–35.
Harris, A.F., Nimmo, D., McKemey, A.R., Kelly, N., Scaife, S., Donnelly, C.A., Beech, C., Petrie, W.D., Alphey, L., 2011. Field performance of engineered male mosquitoes. Nat. Biotechnol. 29, 1034–1037.
Hartl, D., Clark, G., 1997. Principles of population genetics. Sinauer Associates, Sunderland.
Hassan, M.M., Zain, H.M., Basheer, M.A., Elhaj, H.E.F., El-Sayed, B.B., 2014. Swarming and mating behavior of male Anopheles arabiensis Patton (Diptera: Culicidae) in an area of the Sterile Insect Technique Project in Dongola, northern Sudan. Acta Trop. 132, S64–S69.
Hay, S.I., Guerra, C.A., Tatem, A.J., Noor, A.M., Snow, R.W., 2004. The global distribution and population at risk of malaria: past, present, and future. Lancet. Infect. Dis. 4, 327–36.
Hay, S.I., Okiro, E.A., Gething, P.W., Patil, A.P., Tatem, A.J., Guerra, C.A., Snow, R.W., 2010a. Estimating the Global Clinical Burden of Plasmodium falciparum Malaria in 2007. PLoS Med. 7, e1000290.

Hay, S.I., Sinka, M.E., Okara, R.M., Kabaria, C.W., Mbithi, P.M., Tago, C.C., Benz, D., Gething, P.W., Howes, R.E., Patil, A.P., Temperley, W.H., Bangs, M.J., Chareonviriyaphap, T., Elyazar, I.R.F., Harbach, R.E., Hemingway, J., Manguin, S., Mbogo, C.M., Rubio-Palis, Y., Godfray, H.C.J., 2010b. Developing Global Maps of the Dominant Anopheles Vectors of Human Malaria. PLoS Med. 7, e1000209.

Hemingway, J., Ranson, H., 2000. Insecticide resistance in insect vectors of human disease. Annu. Rev. Entomol. 45, 371–91.

Hewitt, S., Rowland, M., 1999. Control of zoophilic malaria vectors by applying pyrethroid insecticides to cattle. Trop. Med. Int. Heal. 4, 481–486.

Hidalgo, K., Dujardin, J.P., Mouline, K., Dabiré, R.K., Renault, D., Simard, F., 2015. Seasonal variation in wing size and shape between geographic populations of the malaria vector, Anopheles coluzzii in Burkina Faso (West Africa). Acta Trop. 143, 79–88.

Hill, N., Lenglet, A., Arnéz, A.M., Carneiro, I., 2007. Plant based insect repellent and insecticide treated bed nets to protect against malaria in areas of early evening biting vectors: Double blind randomised placebo controlled clinical trial in the Bolivian Amazon. Br. Med. J. 335, 1023–1025.

Hinton, H.E., 1968. Observations on the biology and taxonomy of the eggs of Anopheles mosquitos. Bull. Entomol. Res. 57, 495–508.

Hodges, T.K., Athrey, G., Deitz, K.C., Overgaard, H.J., Matias, A., Caccone, A., Slotman, M.A., 2013. Large fluctuations in the effective population size of the malaria mosquito Anopheles gambiae s.s. during vector control cycle. Evol. Appl. 6, 1171–83.

Ihara, S., Yoshikawa, K., Touhara, K., 2013. Chemosensory signals and their receptors in the olfactory neural system. Neuroscience 254, 45–60.

Junkum, A., Jitpakdi, A., Komalamisra, N., Jariyapan, N., Somboon, P., Bates, P.A., Choochote, W., 2004. Comparative morphometry and morphology of Anopheles aconitus Form B and C eggs under scanning electron microscope. Rev. Inst. Med. Trop. Sao Paulo 46, 257–262.

Karunamoorthi, K., 2011. Vector control: a cornerstone in the malaria elimination campaign. Clin Microbiol Infect 17, 1608–1616.

Karunamoorthi, K., 2014. Malaria vaccine: A future hope to curtail the global malaria burden. Int. J. Prev. Med. 5, 529–538.

Kisinza, W.N., Nkya, T.E., Kabula, B., Overgaard, H.J., Massue, D.J., Mageni, Z., Greer, G., Kaspar, N., Mohamed, M., Reithinger, R., Moore, S., Lorenz, L.M., Magesa, S., 2017. Multiple insecticide resistance in Anopheles gambiae from Tanzania: A major concern for malaria vector control. Malar. J. 16, 1–10.

Kiszewski, A., Mellinger, A., Spielman, A., Malaney, P., Sachs, S. ehrlich, Sachs, J., 2004. A global index representing the stability of malaria transmission. Am J Trop Med Hyg 70, 486–498.

Klingenberg, C.P., 2011. MorphoJ: An integrated software package for geometric morphometrics. Mol. Ecol. Resour. 11, 353–357.

Klingenberg, C.P., 2013. Visualizations in geometric morphometrics: How to read and how to make graphs showing shape changes. Hystrix 24, 15–24.

Klingenberg, C.P., Debat, V., Roff, D.A., 2010. Quantitative genetics of shape in cricket wings: developmental integration in a functional structure. Evolution . 64, 2935-2951.

Klingenberg, C.P., McIntyre, G.S., 1998. Geometric Morphometrics of Developmental Instability: Analyzing Patterns of Fluctuating Asymmetry with Procrustes Methods. Evolution . 52, 1363.

Knell, A.J., 1991. Malaria : a publication of the tropical programme of the Wellcome Trust / edited by A.J. Knell. - Version details - Trove.

Knipling, E.F., 1955. Possibilities of Insect Control or Eradication Through the Use of

Sexually Sterile Males1. J. Econ. Entomol. 48, 459–462.
Kondrashov, A.S., Shpak, M., 1998. On the origin of species by means of assortative mating. Proc. R. Soc. London. Ser. B Biol. Sci. 265, 2273-2278.
Kumar, R., Hwang, J.-S., 2014. Larvicidal efficiency of aquatic predators: A perspective for mosquito biocontrol Potential of cyclopoid copepod to control disease vectors View project Anti-predator Behavioral Adaptations of Mosquitoes Immatures View project.
Kumar, S.S., Puttaraju, H.P., 2012. Improvised microinjection technique for mosquito vectors. Indian J. Med. Res. 136, 971–8.
Kuzma, J., Rawls, L., 2015. Engineering the Wild: Gene Drives and Intergenerational Equity ; Engineering the Wild: Gene Drives and Intergenerational Equity.
Lambert, P., 2003. Malaria - History of Malaria [WWW Document]. URL https://www.nobelprize.org/educational/medicine/malaria/readmore/history.html (accessed 7.30.18).
Lang, J.T., 1977. Contact Sex Pheromone in the Mosquito Cullseta lNornata (Diptera: Culicidae)1. J. Med. Entomol. 14, 448–454.
Lanzaro, G., Lee, Y., 2013. Speciation in Anopheles gambiae. The Distribution of Genetic Polymorphism and Patterns of Reproductive Isolation Among Natural Populations. In: Anopheles Mosquitoes- New Insights into Malaria Vectors. Sylvie Manguin, IntechOpen.
Lanzaro, G., Tripet, F., 2004. Gene flow among populations of Anopheles gambiae: a critical review. Frontis 109–132.
Lee, Y., Collier, T.C., Sanford, M.R., Marsden, C.D., Fofana, A., Cornel, A.J., Lanzaro, G.C., 2013a. Chromosome inversions, genomic differentiation and speciation in the African malaria mosquito Anopheles gambiae. PLoS One 8, 1–12.
Lee, Y., Cornel, A.J., Meneses, C.R., Fofana, A., Andrianarivo, A.G., McAbee, R.D., Fondjo, E., Traoré, S.F., Lanzaro, G.C., 2009. Ecological and genetic relationships of the Forest-M form among chromosomal and molecular forms of the malaria vector Anopheles gambiae sensu stricto. Malar. J. 8, 75.
Lee, Y., Marsden, C., Norris, L., Collier, T., Main, B., Fofana, A., Cornel, A., Lanzaro, G., 2013b. Spatiotemporal dynamics of gene flow and hybrid fitness between the M and S forms of the malaria mosquito, Anopheles gambiae. Proc. Natl. Acad. Sci. U. S. A. 110, 19854–9.
Lehmann, T., Diabate, A., 2008. The molecular forms of Anopheles gambiae: a phenotypic perspective. Infect. Genet. Evol. 8, 737–46.
Lehmann, T., Weetman, D., Huestis, D.L., Yaro, A.S., Kassogue, Y., Diallo, M., Donnelly, M.J., Dao, A., 2017. Tracing the origin of the early wet-season Anopheles coluzzii in the Sahel. Evol. Appl. 10, 704–717.
Lindsay, S.W., Emerson, P.M., Charlwood, J.D., 2002. Reducing malaria by mosquito-proofing houses. Trends Parasitol. 18, 510–514.
Linley, J.R., Kaiser, P.E., Cockburn, A.F., 1993. A description and morphometric study of the eggs of species of the _Anopheles quadrimaculatus_ complex (Diptera: Culicidae). Mosq. Syst. 25, 124–147.
Linley, J.R., Lounibos, L.P., Conn, J., Duzak, D., Nishimura, N., 1996. A description and morphometric comparison of eggs from eight geographic populations of the South American malaria vector Anopheles (Nyssorhynchus) nuneztovari (Diptera:Culicidae). J. Am. Mosq. Control Assoc. 12, 275–92.
Liu, C., Pitts, R.J., Bohbot, J.D., Jones, P.L., Wang, G., Zwiebel, L.J., 2010. Distinct Olfactory Signaling Mechanisms in the Malaria Vector Mosquito Anopheles gambiae. PLOS Biol. 8, e1000467.
Livak, K.J., Schmittgen, T.D., 2001. Analysis of Relative Gene Expression Data Using Real-

Time Quantitative PCR and the 2−ΔΔCT Method. Methods 25, 402–408.
Lopez-Martinez, G., Meuti, M., Denlinger, D.L., 2012. Rehydration Driven RNAi: A Novel Approach for Effectively Delivering dsRNA to Mosquito Larvae. J. Med. Entomol 49, 215–218.
Lounibos, L.P., Coetzee, M., Duzak, D., Nishimura, N., Linley, J.R., Service, M.W., Cornel, A.J., Fontenille, D., Mukwaya, L.G., 1999. A description and morphometric comparison of eggs of species of the Anopheles gambiae complex. J. Am. Mosq. Control Assoc. 15, 157–85.
Lounibos, L.P., Duzak, D., Linley, J.R., 1997. Comparative Egg Morphology of Six Species of the Albimanus Section of Anopheles (Nyssorhynchus) (Diptera: Culicidae). J. Med. Entomol. 34, 136–155.
Lutz, E.K., Lahondère, C., Vinauger, C., Riffell, J.A., 2017. Olfactory learning and chemical ecology of olfaction in disease vector mosquitoes: a life history perspective. Curr. Opin. Insect Sci. 20, 75–83.
Lyimo, E.O., Takken, W., Koella, J.C., 1992. Effect of rearing temperature and larval density on larval survival, age at pupation and adult size of Anopheles gambiae. Entomol. Exp. Appl. 63, 265–271.
Magnusson, K., Mendes, A.M., Windbichler, N., Papathanos, P.A., Nolan, T., Dottorini, T., Rizzi, E., Christophides, G.K., Crisanti, A., 2011. Transcription regulation of sex-biased genes during ontogeny in the Malaria vector Anopheles gambiae. PLoS One 6.
Mahande, A.M., Mosha, F.W., Mahande, J.M., Kweka, E.J., 2007. Role of cattle treated with deltamethrine in areas with a high population of Anopheles arabiensis in Moshi, Northern Tanzania. Malar. J. 6, 109.
Malhotra, P.R., Chandra Jatav, P., Singh Chauhan, R., Chauhan, R.S., 2000. Surface morphology of the egg of anopheles stephensi stephensi sensu stricto (diptera, culicidae). Ital. J. Zool. 67, 147–151.
Mang'era, C.M., Mbai, F.N., Omedo, I.A., Mireji, P.O., Omar, S.A., 2012. Changes in genotypes of Plasmodium falciparum human malaria parasite following withdrawal of chloroquine in Tiwi, Kenya. Acta Trop. 123, 202–207.
Marinotti, O., Calvo, E., Nguyen, Q.K., Dissanayake, S., Ribeiro, J.M.C., James, A.A., 2006. Genome-wide analysis of gene expression in adult Anopheles gambiae. Insect Mol. Biol. 15, 1–12.
Marquardt, W.C., Demaree, R.S., Grieve, R.B., 2000. Parasitology and vector biology. Academic Press.
Martin, F.J., Flicek, P., Mugal, C.F., Nater, A., Künstner, A., Burri, R., Ellegren, H., Mäkinen, H., Uebbing, S., Dutoit, L., Aken, B., Searle, S.M.J., Backström, N., Bolivar, P., 2016. Divergence in gene expression within and between two closely related flycatcher species. Mol. Ecol. 25, 2015–2028.
Mastbaum, O., 1957. Past and present position of malaria in Swaziland. J. Trop. Med. Hyg. 60, 119–127.
Mattah, P.A.D., Futagbi, G., Amekudzi, L.K., Mattah, M.M., de Souza, D.K., Kartey-Attipoe, W.D., Bimi, L., Wilson, M.D., 2017. Diversity in breeding sites and distribution of Anopheles mosquitoes in selected urban areas of southern Ghana. Parasit. Vectors 10, 25.
Mitteroecker, P., Gunz, P., 2009. Advances in Geometric morphometrics. Evol. Biol. 36, 235–247.
Moore, D. V., Lanier, J.E., 1961. Observations on Two Plasmodium Falciparum Infections with an Abnormal Response to Chloroquine. Am. J. Trop. Med. Hyg. 10, 5–9.
Mouatcho, J.C., Goldring, D.J.P., 2013. Malaria rapid diagnostic tests: Challenges and prospects. J. Med. Microbiol. 62, 1491–1505.

Munga, S., Minakawa, N., Zhou, G., Barrack, O.J., Githeko, A.K., Yan, G., 2006. Effects of Larval Competitors and Predators on Oviposition Site Selection of Anopheles gambiae Sensu Stricto, Chemical Ecology J. Med. Entomol. 43, 221-224.
Murugan, K., Priyanka, V., Dinesh, D., Madhiyazhagan, P., Panneerselvam, C., Subramaniam, J., Suresh, U., Chandramohan, B., Roni, M., Nicoletti, M., Alarfaj, A.A., Higuchi, A., Munusamy, M.A., Khater, H.F., Messing, R.H., Benelli, G., 2015. Predation by Asian bullfrog tadpoles, Hoplobatrachus tigerinus, against the dengue vector, Aedes aegypti, in an aquatic environment treated with mosquitocidal nanoparticles. Parasitol. Res. 114, 3601–3610.
Naqqash, M.N., Gökçe, A., Bakhsh, A., Salim, M., 2016. Insecticide resistance and its molecular basis in urban insect pests. Parasitol. Res. 115, 1363–1373.
Ng'habi, K.R., Horton, A., Knols, B.G.J., Lanzaro, G.C., 2007. A New Robust Diagnostic Polymerase Chain Reaction for Determining the Mating Status of Female Anopheles gambiae Mosquitoes. ASTMH. 77, 485-487.
Ngufor, C., Fagbohoun, J., Critchley, J., N'Guessan, R., Todjinou, D., Malone, D., Akogbeto, M., Rowland, M., 2017. Which intervention is better for malaria vector control: Insecticide mixture long-lasting insecticidal nets or standard pyrethroid nets combined with indoor residual spraying? Malar. J. 16, 1–9.
Niang, A., Epopa, P.S., Sawadogo, S.P., Maïga, H., Konaté, L., Faye, O., Dabiré, R.K., Tripet, F., Diabaté, A., 2015. Does extreme asymmetric dominance promote hybridization between Anopheles coluzzii and Anopheles gambiae s.s. in seasonal malaria mosquito communities of West Africa? Parasites and Vectors 8, 1–10.
Niang, E.H.A., Konaté, L., Diallo, M., Faye, O., Dia, I., 2014. Reproductive isolation among sympatric molecular forms of An. gambiae from inland areas of south-eastern Senegal. PLoS One 9, e104622.
Nijhout, F.H., Craig, G.B., 1971. Reproductive isolation in Stegomyia mosquitoes. III evidence for a sexual pheromone. Entomol. Exp. Appl. 14, 399–412.
Njoroge, M.M., Tirados, I., Lindsay, S.W., Vale, G.A., Torr, S.J., Fillinger, U., 2017. Exploring the potential of using cattle for malaria vector surveillance and control: a pilot study in western Kenya. Parasit. Vectors 10, 18.
Nowick, K., Gernat, T., Almaas, E., Stubbs, L., 2009. Differences in human and chimpanzee gene expression patterns define an evolving network of transcription factors in brain. Proc. Natl. Acad. Sci. 106, 22358–22363.
Nunes, J.K., Woods, C., Carter, T., Raphael, T., Morin, M.J., Diallo, D., Leboulleux, D., Jain, S., Loucq, C., Kaslow, D.C., Birkett, A.J., 2014. Development of a transmission-blocking malaria vaccine: Progress, challenges, and the path forward. Vaccine 32, 5531–5539.
Nwakanma, D.C., Neafsey, D.E., Jawara, M., Adiamoh, M., Lund, E., Rodrigues, A., Loua, K.M., Konate, L., Sy, N., Dia, I., Awolola, T.S., Muskavitch, M.A.T., Conway, D.J., 2013. Breakdown in the process of incipient speciation in Anopheles gambiae. Genetics 193, 1221–31.
Obsomer, V., Titeux, N., Vancustem, C., Duveiller, G., Pekel, J.-F., Connor, S., Ceccato, P., Coosemans, M., 2013. From Anopheles to Spatial Surveillance: A Roadmap Through a Multidisciplinary Challenge. In: Anopheles Mosquitoes - New Insights into Malaria Vectors.
Old, R.., Primose, S., 1994. Principles of gene manipulation. Wiley & sons Incorporated.
Oliva, C.F., Vreysen, M.J.B., Dupé, S., Lees, R.S., Gilles, J.R.L., Gouagna, L.C., Chhem, R., 2014. Current status and future challenges for controlling malaria with the sterile insect technique: Technical and social perspectives. Acta Trop. 132, S130–S139.
Oliveira, E., Salgueiro, P., Palsson, K., Vicente, J.L., Arez, P., Jaenson, T.G., Caccone, A.,

Pinto, J., 2008. High Levels of Hybridization between Molecular Forms of Anopheles gambiae from Guinea Bissau. J. Med. Entomol. 45, 1057–1063.
Ōmura, S., Crump, A., 2017. Ivermectin and malaria control. Malar. J. 16, 172.
Oppenheim, S.J., Baker, R.H., Simon, S., Desalle, R., 2015. We can't all be supermodels: The value of comparative transcriptomics to the study of non-model insects. Insect Mol. Biol. 24, 139–154.
Papa, F., Windbichler, N., Waterhouse, R.M., Cagnetti, A., Amato, R.D., Persampieri, T., Lawniczak, M.K.N., Nolan, T., Papathanos, P.A., 2017. Rapid evolution of female-biased genes among four species of Anopheles malaria mosquitoes 1–13.
Parker, A., Mehta, K., 2007. Sterile insect technique: a model for dose optimization for improved sterile insect quality. https://doi.org/10.1653/0015-4040(2007)90[88:SITAMF]2.0.CO;2 90, 88–96.
Pasternak, N.D., Dzikowski, R., 2009. PfEMP1: an antigen that plays a key role in the pathogenicity and immune evasion of the malaria parasite Plasmodium falciparum. Int. J. Biochem. Cell Biol. 41, 1463–6.
Paterson, H.E., 1964. Direct evidence for the specific distinctness of forms A, B and C of the Anopheles gambiae complex. Riv. Malariol. 43.
Paton, D., Touré, M., Sacko, A., Coulibaly, M.B., Traoré, S.F., Tripet, F., 2013. Genetic and environmental factors associated with laboratory rearing affect survival and assortative mating but not overall mating success in Anopheles gambiae sensu stricto. PLoS One 8.
Pelletier, J., Guidolin, A., Syed, Z., Cornel, A.J., Leal, W.S., 2010. Knockdown of a Mosquito Odorant-binding Protein Involved in the Sensitive Detection of Oviposition Attractants. J. Chem. Ecol. 36, 245–248.
Pennetier, C., Warren, B., Dabiré, K.R., Russell, I.J., Gibson, G., 2010. "Singing on the wing" as a mechanism for species recognition in the malarial mosquito Anopheles gambiae. Curr. Biol. 20, 131–6.
Pillai, A.B., Nagarajan, U., Mitra, A., Krishnan, U., Rajendran, S., Hoti, S.L., Mishra, R.K., 2017. RNA interference in mosquito: understanding immune responses, double-stranded RNA delivery systems and potential applications in vector control. Insect Mol. Biol. 26, 127–139.
Pitts, R.J., Mozūraitis, R., Gauvin-Bialecki, A., Lempérière, G., 2014. The roles of kairomones, synomones and pheromones in the chemically-mediated behaviour of male mosquitoes. Acta Trop. 132, S26–S34.
Pitts, R.J., Rinker, D.C., Jones, P.L., Rokas, A., Zwiebel, L.J., 2011. Transcriptome profiling of chemosensory appendages in the malaria vector Anopheles gambiae reveals tissue- and sex-specific signatures of odor coding. BMC Genomics 12, 271.
Pombi, M., Caputo, B., Simard, F., Di Deco, M., Coluzzi, M., della Torre, A., Costantini, C., Besansky, N.J., Petrarca, V., 2008. Chromosomal plasticity and evolutionary potential in the malaria vector Anopheles gambiae sensu stricto: insights from three decades of rare paracentric inversions. BMC Evol. Biol. 8, 309.
Ponlawat, A., Harrington, L.C., 2009. Factors associated with male mating success of the dengue vector mosquito, Aedes aegypti. Am. J. Trop. Med. Hyg. 80, 395–400.
Potter, C.J., 2014. Stop the Biting: Targeting a Mosquito's Sense of Smell. Cell 156, 878–881.
Pratt, D.H., Barnes, C.R., Litting, S.K., 1963. Mosquitoes of public health importance and their control. Public Health Service Publication.
Puckett, J.G., Ni, R., Ouellette, N.T., 2015. Time-Frequency Analysis Reveals Pairwise Interactions in Insect Swarms. Phys Rev Letts. 114, 258103.
Ranson, H., Lissenden, N., 2016. Insecticide Resistance in African Anopheles Mosquitoes: A Worsening Situation that Needs Urgent Action to Maintain Malaria Control. Trends

Parasitol. 32, 187–196.
Reddy, M.R., Overgaard, H.J., Abaga, S., Reddy, V.P., Caccone, A., Kiszewski, A.E., Slotman, M.A., 2011. Outdoor host seeking behaviour of Anopheles gambiae mosquitoes following initiation of malaria vector control on Bioko Island, Equatorial Guinea. Malar. J. 10, 184.
Reddy Palli, S., 2012. RNAi methods for management of insects and their pathogens. CAB Rev. 7.
Regna, K., Harrison, R.M., Heyse, S.A., Chiles, T.C., Michel, K., Muskavitch, M.A.T., 2016. RNAi Trigger Delivery into Anopheles gambiae Pupae Video Link. J. Vis. Exp 109, 53738.
Reiter, P., 2000. From Shakespeare to Defore: malaria in England in the little Ice Age. Emerg. Infect. Dis. 6, 1–11.
Reyes, J., Carro, X., Hernandez, J., Méndez, W., Agricultura, M. De, Rica, C., Agropecuario, M., Carretera, K., Agricultura, M. De, Matasano, C. El, Salvador, S., Salvador, E., 2007. A Multi-Institutional Approach to Create Fruit Fly-Low Prevalence and Fly-Free. Area-Wide Control Insect Pests 627–628.
Ribbands, C.R., 1944. Differences Between Anopheles Melas and Anopheles Gambiae. Ann. Trop. Med. Parasitol. 38, 87–99.
Ribiero, J.M.C., Seulu, F., Abose, T., Kidane, G., Teklehaimanot, A., 1996. Temporal and spatial distribution of anopheline mosquitoes in an Ethiopian village: implications for malaria control strategies. Bull. World Health Organ. 74, 299–305.
Ridley, R.G., 2002. Medical need, scientific opportunity and the drive for antimalarial drugs. Nature 415, 686–693.
Rinker, D.C., Zhou, X., Jason Pitts, R., AGC Consortium, T., Rokas, A., Zwiebel, L.J., 2013. Antennal transcriptome profiles of anopheline mosquitoes reveal human host olfactory specialization in Anopheles gambiae. BMC Genomics 14, 749.
Ritchie, M.G., 2007. Sexual Selection and Speciation. Annu. Rev. Ecol. Evol. Syst. 38, 79–102.
Rodriguez, M.H., Chavez, B., Orozco, A., Loyola, E.G., Martinez-Palomo, A., 1992. Scanning Electron Microscopic Observations of Anopheles albimanus (Diptera: Culicidae) Eggs. J. Med. Entomol. 29, 400–406.
Rohlf, F.J., 1996. Morphometric Spaces, Shape Components and the Effects of Linear Transformations. In: Advances in Morphometrics. Springer US, Boston, MA, pp. 117–129.
Rohlf, F.J., 1999. Shape statistics: Procrustes superimpositions and tangent spaces. J. Classif.
Rose, G., Krzywinska, E., Kim, J., Revuelta, L., Ferretti, L., Krzywinski, J., 2016. Dosage Compensation in the African Malaria Mosquito Anopheles gambiae. Genome Biol. Evol. 8, 411–425.
Roux, O., Diabaté, A., Simard, F., 2013. Larvae of cryptic species of Anopheles gambiae respond differently to cues of predation risk. Freshw. Biol. 58, 1178–1189.
Rueda, L.M., Brown, T.L., Kim, H.C., Klein, T.A., Thongkukiatkul, A., Van, S., 2009. Description and comparison of morphological structures of the eggs of Anopheles hyrcanus group and related species (Diptera: Culicidae) from the Republic of Korea. Zootaxa 40, 23–40.
Saeung, A., Hempolchom, C., Baimai, V., Thongsahuan, S., Taai, K., Jariyapan, N., Chaithong, U., Choochote, W., 2013. Susceptibility of eight species members in the Anopheles hyrcanus group to nocturnally subperiodic Brugia malayi. Parasit. Vectors 6, 5.
Saeung, A., Hempolchom, C., Yasanga, T., Otsuka, Y., Thongsahuan, S., Srisuka, W., Chaithong, U., Taai, K., Somboon, P., Choochote, W., 2014. Scanning electron

microscopy of Anopheles hyrcanus group (Diptera: Culicidae) eggs in Thailand and an ultrastructural key for species identification. Parasitol. Res. 113, 973–981.
Sallum, M.A.M., Flores, D.C., 2004. Ultrastructure of the eggs of two species of *Anopheles (Anopheles)* Meigen. Rev. Bras. Entomol. 48, 185–192.
Sanford, M.R., Demirci, B., Marsden, C.D., Lee, Y., Cornel, A.J., Lanzaro, G.C., 2011. Morphological differentiation may mediate mate-choice between incipient species of Anopheles gambiae s.s. PLoS One 6, e27920.
Santolamazza, F., Mancini, E., Simard, F., Qi, Y., Tu, Z., della Torre, A., 2008. Insertion polymorphisms of SINE200 retrotransposons within speciation islands of Anopheles gambiae molecular forms. Malar. J. 7, 163.
Sauerwein, ₹ R W, Bousema, T., 2015. Transmission blocking malaria vaccines: Assays and candidates in clinical development. Vaccine 33, 7476–7482.
Sawadogo, S.P., Diabaté, A., Toé, H.K., Sanon, A., Lefevre, T., Baldet, T., Gilles, J., Simard, F., Gibson, G., Sinkins, S., Dabiré, R.K., 2013. Effects of Age and Size on Anopheles gambiae s.s. Male Mosquito Mating Success. J. Med. Entomol. 50, 285–293.
Sawadogo, S.P., Niang, A., Bilgo, E., Millogo, A., 2017. Targeting male mosquito swarms to control malaria vector density 1–11.
Scott, J.A., Brogdon, W.G., Collins, F.H., 1993. Identification of single specimens of the Anopheles gambiae complex by the polymerase chain reaction, Am. J. Trop. Med. Hyg. 49, 520-529.
Scott, J.G., Michel, K., Bartholomay, L.C., Siegfried, B.D., Hunter, W.B., Smagghe, G., Zhu, K.Y., Douglas, A.E., 2013. Towards the elements of successful insect RNAi. J. Insect Physiol. 59, 1212–1221.
Scott, M.J., Concha, C., Welch, J.B., Phillips, P.L., Skoda, S.R., 2017. Review of research advances in the screwworm eradication program over the past 25 years. Entomol. Exp. Appl. 164, 226–236.
Sedaghat, M.M., Linton, Y.-M., Oshaghi, M.A., Vatandoost, H., Harbach, R.E., 2003. The Anopheles maculipennis complex (Diptera: Culicidae) in Iran: molecular characterization and recognition of a new species. Bull. Entomol. Res. 93, 527–535.
Sehrawat, N., 2014. Scanning Electron Microscopic Observations of Eggs of Anopheles Fluviatilis (T) Mosquito (Diptera: Culicidae). Res. Rev. J. Zool. Sci. 2, 6–12.
Service, M., 2012. Medical Entomology for Students, 5th Edition. New York, Cambridge University press.
Simões, P.M. V., Gibson, G., Russell, I.J., 2017. Pre-copula acoustic behaviour of males in the malarial mosquitoes *Anopheles coluzzii* and *Anopheles gambiae s.s.* does not contribute to reproductive isolation. J. Exp. Biol. 220, 379–385.
Singh, A.D., Wong, S., Ryan, C.P., Whyard, S., 2013. Oral Delivery of Double-Stranded RNA in Larvae of the Yellow Fever Mosquito, *Aedes aegypti:* Implications for Pest Mosquito Control. J. Insect Sci. 13, 1–18.
Singh, R.S., Artieri, C.G., 2010. Male Sex Drive and the Maintenance of Sex: Evidence from Drosophila. J. Hered. 101, S100–S106.
Sinka, M., Rubio-Palis, Y., Manguin, S., Patil, A.P., Temperley, W.H., Gething, P.W., Van Boeckel, T., Kabaria, C.W., Harbach, R.E., Hay, S.I., 2010. The dominant Anopheles vectors of human malaria in the Americas: occurrence data, distribution maps and bionomic précis. Parasit. Vectors 3, 72.
Sinka, M.E., Bangs, M.J., Manguin, S., Chareonviriyaphap, T., Patil, A.P., Temperley, W.H., Gething, P.W., Elyazar, I.R., Kabaria, C.W., Harbach, R.E., Hay, S.I., 2011. The dominant Anopheles vectors of human malaria in the Asia-Pacific region: occurrence data, distribution maps and bionomic précis. Parasit. Vectors 4, 89.
Sinka, M.E., Bangs, M.J., Manguin, S., Rubio-Palis, Y., Chareonviriyaphap, T., Coetzee, M.,

Mbogo, C.M., Hemingway, J., Patil, A.P., Temperley, W.H., Gething, P.W., Kabaria, C.W., Burkot, T.R., Harbach, R.E., Hay, S.I., 2012. A global map of dominant malaria vectors. Parasit. Vectors 5, 69.

Sinka, M.L., 2013. Global Distribution of the Dominant Vector Species of Malaria. In: Anopheles Mosquitoes- New Insights into Malaria Vectors. Sylvie Manguin, IntechOpen.

Slotman, M. a, Reimer, L.J., Thiemann, T., Dolo, G., Fondjo, E., Lanzaro, G.C., 2006. Reduced recombination rate and genetic differentiation between the M and S forms of Anopheles gambiae s.s. Genetics 174, 2081–93.

Snow, R.W., 2014. Sixty years trying to define the malaria burden in Africa: Have we made any progress? BMC Med. 12, 1–6.

Sota, T., Mogi, M., 1992. Interspecific variation in desiccation survival time of Aedes (Stegomyia) mosquito eggs is correlated with habitat and egg size. Oecologia 90, 353–358.

Suman, D.S., Shrivastava, A.R., Pant, S.C., Parashar, B.D., 2011. Differentiation of Aedes aegypti and Aedes albopictus (Diptera: Culicidae) with egg surface morphology and morphometrics using scanning electron microscopy. Arthropod Struct. Dev. 40, 479–483.

Suman, D.S., Shrivastava, a R., Parashar, B.D., Pant, S.C., Agrawal, O.P., Prakash, S., 2009. Variation in morphology and morphometrics of eggs of Culex quinquefasciatus mosquitoes from different ecological regions of India. J. Vector Ecol. 34, 191–199.

Suzuki, Y., Niu, G., Hughes, G.L., Rasgon, J.L., 2014. A viral over-expression system for the major malaria mosquito Anopheles gambiae. Sci. Rep. 4, 5127.

Talisuna, A.O., Karema, C., Ogutu, B., Juma, E., Logedi, J., Nyandigisi, A., Mulenga, M., Mbacham, W.F., Roper, C., Guerin, P.J., D'Alessandro, U., Snow, R.W., 2012. Mitigating the threat of artemisinin resistance in Africa: improvement of drug-resistance surveillance and response systems. Lancet Infect. Dis. 12, 888–896.

Tatsuta, H., Takahashi, K.H., Sakamaki, Y., 2018. Geometric morphometrics in entomology: Basics and applications. Entomol. Sci. 21, 164–184.

Taylor, C., Toure, Y.T., Carnahan, J., Norris, D.E., Dolo, G., Traore, S.F., Edillo, F.E., Lanzaro, G.C., 2001. Gene Flow Among Populations of the Malaria Vector, Anopheles gambiae, in Mali, West Africa. Genetics 157, 743–750.

Tchicaya, E.S., Koudou, B.G., Keiser, J., Adja, A.M., Cissé, G., Tanner, M., Tano, Y., Utzinger, J., 2009. Effect of Repeated Application of Microbial Larvicides on Malaria Transmission in Central Côte d'Ivoire. J. Am. Mosq. Control Assoc. 25, 382–385.

Tene, B.F., Poupardin, R., Costantini, C., Awono-Ambene, P., Wondji, C.S., Ranson, H., Antonio-Nkondjio, C., 2013. Resistance to DDT in an Urban Setting: Common Mechanisms Implicated in Both M and S Forms of Anopheles gambiae in the City of Yaoundé Cameroon. PLoS One 8, e61408.

Thailayil, J., Magnusson, K., Godfray, H.C.J., Crisanti, A., Catteruccia, F., 2011. Spermless males elicit large-scale female responses to mating in the malaria mosquito Anopheles gambiae. Proc. Natl. Acad. Sci. 108, 13677–13681.

Thomson, R., 1948. Studies on Anopheles gambiae and A. melas in and around Lagos. Bull. Entomol. Res. 527–558.

Thomson, R.C.M., 1951. Studies on Salt-Water and Fresh-Water Anopheles gambiae on the East African Coast. Bull. Entomol. Res. 41, 487.

Touré, Y.T., Petrarca, V., Traoré, S.F., Coulibaly, A., Maiga, H.M., Sankaré, O., Sow, M., Di Deco, M.A., Coluzzi, M., 1998. The distribution and inversion polymorphism of chromosomally recognized taxa of the Anopheles gambiae complex in Mali, West Africa. Parassitologia 40, 477–511.

Touré, Y.T., Petrarca, V., Traoré, S.F., Coulibaly, A., Maïga, H.M., Sankaré, O., Sow, M., Di Deco, M.A., Coluzzi, M., 1994. Ecological genetic studies in the chromosomal form Mopti of Anopheles gambiae s.str. in Mali, West Africa. Genetica 94, 213–223.
Toure, Y.T., Traore, S.F., Sankare, O., Sow, M.Y., Coulibaly, A., Esposito, F., Petrarca, V., 1996. Perennial transmission of malaria by the Anopheles gambiae complex in a North Sudan Savanna area of Mali. Med. Vet. Entomol. 10, 197–199.
Townson, H., Onapa, A.W., 1994. Identification by rDNA-PCR of Anopheles bwambae, a geothermal spring species of the An. gambiae complex. Insect Mol. Biol. 3, 279–282.
Tripet, F., Dolo, G., Lanzaro, G.C., 2005. Multilevel analyses of genetic differentiation in Anopheles gambiae s.s. reveal patterns of gene flow important for malaria-fighting mosquito projects. Genetics 169, 313–24.
Tripet, F., Dolo, G., Traoré, S., Lanzaro, G.C., 2004. The "wingbeat hypothesis" of reproductive isolation between members of the Anopheles gambiae complex (Diptera: Culicidae) does not fly. J. Med. Entomol. 41, 375–384.
Tripet, F., Touré, Y.T., Dolo, G., Lanzaro, G.C., 2003. Frequency of multiple inseminations in field-collected Anopheles gambiae females revealed by DNA analysis of transferred sperm. Am. J. Trop. Med. Hyg. 68, 1–5.
Tripet, F., Touré, Y.T., Taylor, C.E., Norris, D.E., Dolo, G., Lanzaro, G.C., 2001. DNA analysis of transferred sperm reveals significant levels of gene flow between molecular forms of *Anopheles gambiae*. Mol. Ecol. 10, 1725–1732.
Tripet, F., Wright, J., Cornel, A., Fofana, A., Mcabee, R., Meneses, C., Reimer, L., Slotman, M., Thiemann, T., Dolo, G., Traoré, S., Lanzaro, G., 2007. Longitudinal survey of knockdown resistance to pyrethroid (kdr) in Mali, West Africa, and evidence of its emergence in the Bamako form of Anopheles gambiae s.s. Am. J. Trop. Med. Hyg. 76, 81-87.
Trung, H.D., Van Bortel, W., Sochantha, T., Keokenchanh, K., Quang, N.T., Cong, L.D., Coosemans, M., 2004. Malaria transmission and major malaria vectors in different geographical areas of Southeast Asia. Trop. Med. Int. Heal. 9, 230–237.
Turner, T.L., Hahn, M.W., 2007. Locus- and Population-Specific Selection and Differentiation between Incipient Species of Anopheles gambiae. Mol. Biol. Evol. 24, 2132–2138.
Turner, T.L., Hahn, M.W., Nuzhdin, S. V, 2005. Genomic islands of speciation in Anopheles gambiae. PLoS Biol. 3, e285.
Tyagi, V., Dhiman, S., Sharma, A.K., Srivastava, A.R., Rabha, B., Sukumaran, D., Veer, V., 2017. Morphometric and morphological appraisal of the eggs of Anopheles stephensi (Diptera: Culicidae) from India. J. Vector Borne Dis. 54, 151–156.
Tyagi, V., Sharma, A.K., Dhiman, S., Srivastava, A.R., Yadav, R., Sukumaran, D., Agrawal, O.P., Veer, V., 2016. Malaria vector Anopheles culicifacies sibling species differentiation using egg morphometry and morphology. Parasites and Vectors 9, 1–13.
Utzinger, J., Tozan, Y., Singer, B.H., Utzinger, È., Tozan, Y., Singer, B.H., 2001. Efficacy and cost-effectiveness of environmental management for malaria control. Trop. Med. Int. Heal. 6, 677–87.
Viscosi, V., 2015. Geometric morphometrics and leaf phenotypic plasticity: Assessing fluctuating asymmetry and allometry in European white oaks (Quercus). Bot. J. Linn. Soc. 179, 335–348.
Wang, D., Pan, Y., Zhao, X., Zhu, L., Fu, B., Li, Z., 2011. Genome-wide temporal-spatial gene expression profiling of drought responsiveness in rice. BMC Genomics. 12, 149.
Webster, M., Sheets, H.D., 2010. A practical introduction to landmark-based geometric morphometrics. Paleontol. Soc. Pap. 16, 163–188.
Weill, M., Chandre, F., Brengues, C., Manguin, S., Akogbeto, M., Pasteur, N., Guillet, P.,

Raymond, M., 2000. The kdr mutation occurs in the Mopti form of Anopheles gambiae s.s. through introgression. Insect Mol. Biol. 9, 451–455.
Wellems, T.E., Plowe, C. V, 2001. Chloroquine-resistant malaria. J Infect Dis 184, 770–776.
White, B.J., Cheng, C., Simard, F., Costantini, C., Besansky, N.J., 2010. Genetic association of physically unlinked islands of genomic divergence in incipient species of Anopheles gambiae. Mol. Ecol. 19, 925–939.
White, G.B., 1977. The Place of Morphological Studies in the Investigation Anopheles Species Complexes. Mosq. Syst. 9.
White, G.B., 1978. Systematic Reappraisal of the Anopheles maculipennis Complex., Mosquito Systematics.
White, N.J., 2008. The role of anti-malarial drugs in eliminating malaria. Malar. J. 7, S8.
Whitehead, A., Crawford, D.L., 2006. Neutral and adaptive variation in gene expression. Proc. Natl. Acad. Sci. 103, 5425–5430.
Whyard, S., Erdelyan, C.N., Partridge, A.L., Singh, A.D., Beebe, N.W., Capina, R., 2015. Silencing the buzz: a new approach to population suppression of mosquitoes by feeding larvae double-stranded RNAs. Parasites & Vectors 8.
Wilson, A.L., Pinder, M., Bradley, J., Donnelly, M.J., Hamid-Adiamoh, M., Jarju, L.B.S., Jawara, M., Jeffries, D., Kandeh, B., Rippon, E.J., Salami, K., D'Alessandro, U., Lindsay, S.W., 2018. Emergence of knock-down resistance in the Anopheles gambiae complex in the Upper River Region, The Gambia, and its relationship with malaria infection in children. Malar. J. 17, 205.
Wiwatanaratanabutr, I., Allan, S., Linthicum, K., Kittayapong, P., 2010. Strain-Specific Differences in Mating, Oviposition, and Host-Seeking Behavior Between Wolbachia-Infected And Uninfected Aedes albopictus. J. Am. Mosq. Control Assoc. 26, 265–273.
Wondji, C., Simard, F., Fontenille, D., 2002. Evidence for genetic differentiation between the molecular forms M and S within the Forest chromosomal form of Anopheles gambiae in an area of sympatry. Insect Mol. Biol. 11, 11–19.
World Health Organization, 2015. World malaria report 2015.
World Health Organization, 2016. World Malaria Report 2016.
World Health Organization, 2017. World malaria report 2017.
Zhang, D., Lees, R.S., Xi, Z., Bourtzis, K., Gilles, J.R.L., 2016. Combining the sterile insect technique with the incompatible insect technique: III-robust mating competitiveness of irradiated triple wolbachia-infected aedes albopictus males under semi-field conditions. PLoS One 11, 1–15.
Zhang, G., Hussain, M., Neill, S.L.O., Asgari, S., 2013. Wolbachia uses a host microRNA to regulate transcripts of a methyltransferase , contributing to dengue virus inhibition in Aedes aegypti. Proc Natl Acad Sci. 110, 1276-10281.
Zwiebel, L., Takken, W., 2004. Olfactory regulation of mosquito–host interactions. Insect Biochem. Mol. Biol. 34, 645–652.

Appendices

Appendix 1. List of primers used for gene expression studies (Chapter 4). Each gene has two independent sets of primers.

Gene Target	Primers sequence	Gene Target	Primers sequence
AGAP000940_F	**TGCCAGTGGATCTGGATGTA**	AGAP001052RB_R	**TGTCGTTGAACATACTGTTGCG**
AGAP000940_R	**GTTG CAATTGAAGCTGACGA**	AGAP001052RB_2F	**TCAATGCCTCAGTCACACCC**
AGAP000940_2F	**AAGTCAGGCTACGCCAAGAA**	AGAP001052RB_2R	**CTCCGTTGTGTTCCACTCCA**
AGAP000940_2R	**CATCGGTTCGTCTTCATCCT**	AGAP001052RC_F	**CATTCAACCACCTGCAAGCC**
AGAP013136_F	**AATCAACCGGCACCTGATAG**	AGAP001052RC_R	**CTTGTCGTTGAACTGTTGCGA**
AGAP013136_R	**CGGAAGCCAAAATCTAGCTG**	AGAP001052RC_2F	**CATACAGTCGTCCTCTGCGA**
AGAP013136_2F	**ATCACCGGATGACGAAGAAC**	AGAP001052RC_2R	**TTCGCATTGGCATTGGCATC**
AGAP013136_2R	**CGGAAGCCAAAATCTAGCTG**	AGAP001061_F	**GCAAATGGGTTGTCTCGATT**
AGAP000998_F	**TCTTCTCCCTCCATCCATTG**	AGAP001061_R	**GCACCATCACCATCACGTAG**
AGAP000998_R	**GTCGTCGCACACGTTAAAGA**	AGAP001061_2F	**CGATTCCGAAATGGAAAAGA**
AGAP000998_2F	**AGCTGTCGTATGTGCACCAG**	AGAP001061_2R	**GACATGGCCCTTTGAGTGAT**
AGAP000998_2R	**GATCGGTTGGAACTGCTGAT**	AGAP001070_F	**ACGGAAGCGCTTAGTTTTGA**
AGAP001009_F	**CAGAACAATGTCCGAGCTGA**	AGAP001070_R	**TTAGCATGCGCATAATCGAG**
AGAP001009_R	**GTACTCGGGATGGCACAGAT**	AGAP001070_2F	**AAATCGAGGAGCAAAGCGTA**
AGAP001009_2F	**GGAAGTACGCACTCGTAGGC**	AGAP001070_2R	**AACGGATGTTTCGGAATCTG**
AGAP001009_2R	**GCAGGCTCTCCGTAAATCAG**	AGAP001073RA_F	**CCAGGGATCAACTGATGAGCA**
AGAP013526_F	**CCGGATCCATGACAACTTCT**	AGAP001073RA_R	**ACTGCGAGAGAAAGAACGCA**
AGAP013526_R	**ACTTTGCTTTCGCACGACTT**	AGAP001073RA_2F	**CACTTGCAGTGCCAATTGCT**
AGAP013526_2F	**TGTGCACTTACGCAGACACA**	AGAP001073RA_2R	**GGATTGCGATTGAGTCTGGGA**
AGAP013526_2R	**CGATAAACTTGTGGGCGAAT**	AGAP001073RB_F	**CTTGGTTGCTACCTGCCCTT**
AGAP001022RB_F	**CGCTCCACCTACATGGACAA**	AGAP001073RB_R	**TTTTGGTTCGAGGCGGCTAT**
AGAP001022RB_R	**TGAGAATGTGTAGCGGTGCC**	AGAP001073RB_2F	**TGGCTGCCTCATCAACACTT**
AGAP001022RB_2F	**AGCGCTACTTTGCCATCTGT**	AGAP001073RB_2R	**AGCTTGTTGGGATTGCGATTG**
AGAP001022RB_2R	**GTTCGCGAGAAAACTGACCG**	AGAP001076_F	**AGCAACAATTGCGCATACAG**
AGAP001022RA_F	**AGGCACCGTTCAAATCGTCT**	AGAP001076_R	**CATGTGGCGTCGTGATAAAC**
AGAP001022RA_R	**GCACCGCTGAAGTTGCTTTT**	AGAP001076_2F	**GCAAGCAAATCGCCATTATT**
AGAP001022RA_2F	**CGGAGGTTCGTCTTCGGTAG**	AGAP001076_2R	**CCTGCAGCTTGAATTCTTCC**
AGAP001022RA_2R	**AAACTGACCGTCCACACCAG**	AGAP001082_F	**CGGTTAGTCAGCTGGAGGAG**
AGAP001025_F	**CTGGACACGAACTGCTACGA**	AGAP001082_R	**TCAGATGCAAGCATTTCGAC**
AGAP001025_R	**CCGGTCGTCAAAATAGTGCT**	AGAP001082_2F	**CAACCAGCTCAGTGTGCAAT**
AGAP001025_2F	**GGAAAATGGCTCGTCGTAAA**	AGAP001082_2R	**CAATTTGTCGCATTGTTTGG**
AGAP001025_2R	**GGGTTGGCTTACATCGGTTA**	AGAP001083_F	**TTCTAGCCTGCATGTGCAAC**
AGAP001026_F	**GGTCAATCACCGAGAAAGGA**	AGAP001083_R	**TGCACAGTCCATTCTTCTCG**
AGAP001026_R	**ACCACTCGATGGGTAGATCG**	AGAP001083_2F	**CATCACCAATGCAGACCAAC**
AGAP001026_2F	**ATTACCGGCTGTACCGTGTC**	AGAP001083_2R	**GTTGCACATGCAGGCTAGAA**
AGAP001026_2R	**CGAGCAGATCCTCAAAGTCC**	AGAP001084_F	**AATTTTAATGGCCCCGTAGG**
AGAP001031_F	**CACTTCGCAAAGTACGGTCA**	AGAP001084_R	**ATCACTTTGGTCCGGTTCAG**
AGAP001031_R	**CATCATGGATAGCGGGAACT**	AGAP001084_2F	**ATGTATCTGGGCTCCTGGTG**
AGAP001031_2F	**TGGAGATCACCAAGGAGGAG**	AGAP001084_2R	**GAGGGCAATATTCGCTGTGT**

AGAP001031_2R	**TTTAAATCGGCCAGCGATAC**	AGAP013341_F	**GTGTTTGACGAGCCCAAAAT**
AGAP001033_F	**CGACGGTTTTGGATAATGCT**	AGAP013341_R	**CACTTGTGACATTCCGGTTG**
AGAP001033_R	**CCTTTGATCGTCCACCACTT**	AGAP013341_2F	**CCACCATTGCTTCATCACAG**
AGAP001033_2F	**ATTACACCCGCCTTCAAGTG**	AGAP013341_2R	**CATTACCACTCCCTGCTGGT**
AGAP001033_2R	**GGGCCAAGGTATGTTGAGAA**	AGAP001090_F	**TCAAGCAGGCGTTGTTACAG**
AGAP001035_F	**CCAAAGGCACGGACTACAAT**	AGAP001090_R	**AGCAACGATGGGGTCACTAC**
AGAP001035_R	**AGCGCATGGTGACTTCTTCT**	AGAP001090_2F	**GATAGGAAGCGGGGTAGGAG**
AGAP001035_2F	**AAAACGCGAACCAAAAACAC**	AGAP001090_2R	**GATGACCTCCGCAGATTGAT**
AGAP001035_2R	**ATTGTGTTGGGAGGAAGACG**	AGAP001091_F	**CAGGGAACAAACCGTGAAGT**
AGAP001040_F	**TGCTGATCGGAGTGTTTCTG**	AGAP001091_R	**TGAAGTCGAGCAAAATGACG**
AGAP001040_R	**CAAATATTCCCGCAGATCGT**	AGAP001091_2F	**AGCCTGCTTCTTTGGAACAA**
AGAP001040_2F	**AGCACAGTTCGTCACAATGC**	AGAP001091_2R	**GGATCCTGAGACGTGTCCAT**
AGAP001040_2R	**TGCCTTGTTGATACGGTGAA**	AGAP001094_F	**ATCCACCACTATCGCACTCC**
AGAP001047_F	**TCCGGAATTACGTGACAACA**	AGAP001094_R	**GTTAGAATTGCGCGAAGAGG**
AGAP001047_R	**TGGATCGTCCGTCACTGATA**	AGAP001094_2F	**AAGGTCTGCCTAGTCCAGCA**
AGAP001047_2F	**AAATTTTCTGGGGCCTATCG**	AGAP001094_2R	**GAAATGGCTCTTGGTGGGTA**
AGAP001047_2R	**GCATCAGTTCCCTGAGAAGC**	AGAP010592_F	**ATCGCTATGGTGTTCGGTTC**
AGAP001048_F	**TGCATTTCAGCCAGATTTTG**	AGAP010592_R	**TTGTTGAACTCGACCTCACG**
AGAP001048_R	**TGTCGAACATGGCAGGTAAA**	AGAP010592_2F	**GACGGATCCCAGCTGATAAA**
AGAP001048_2F	**GAAGACGTGGGCTCTTATCG**	AGAP010592_2R	**GCTGCAAACTTCGGCTATTC**
AGAP001048_2R	**GCGACCAGGAATTTCGTAGA**	AGAP004422_F	**CCAACTCGCGACAAAACATTC**
AGAP001050_F	**ATACACGCGTAAAGGCCATC**	AGAP004422_R	**ACCGGCTTCTTGATGATCAGA**
AGAP001050_R	**ACAACCATGCCTGTTTCCTC**	AGAP004422_2F	**TGGTAAGCGAAAGGGTACGG**
AGAP001050_2F	**GCGGAAACCCCTAGTAAAGG**	AGAP004422_2R	**CAGGTCGTGGTACAGGTGAC**
AGAP001050_2R	**GGATTGATACCCATGCAACC**	AGAP000651_F	**ACGCAGTGCTTGCATTAACC**
AGAP001052RB_F	**ACCACCTGCAAGCCAATTCT**	AGAP000651_R	**AACCTCTTCGTCGCACATCT**
		AGAP000651_2F	**CACACCACATCTCTCGGAGG**
		AGAP000651_2R	**ATGAGGGCACAACACGTCAT**

Appendix 2

Gene ontology (Adapted from vector base, https://www.vectorbase.org/ontology-browser)

Genes	Cellular component	Biological process	Molecular function
AGAP000940	NIL	NIL	NIL
AGAP013136	Integral component of membrane	vacuolar proton-transporting V-type ATPase complex assembly	NIL
AGAP000998	trans-Golgi network, integral component of membrane	transport, lysosomal transport	receptor activity, phosphoprotein binding
AGAP001009	peroxisome	NIL	NIL
AGAP013526	peroxisome, integral component of membrane	NIL	NIL
AGAP001022RA	integral component of membrane	G-protein coupled receptor signalling pathway	G-protein coupled receptor activity
AGAP001022RB	NIL	NIL	NIL
AGAP001025	protein binding	NIL	NIL
AGAP001026	extracellular space	proteolysis	carboxypeptidase activity, metallocarboxypeptidase activity, zinc ion binding
AGAP001031	nucleus,Ino80 complex	chromatin remodelling	NIL
AGAP001033	NIL	NIL	NIL
AGAP001035	TORC2 complex	establishment of cell polarity, actin cytoskeleton reorganization, TOR signalling	binding, enzyme activator activity
AGAP001040	integral component of membrane	NIL	NIL
AGAP001047	NIL	phosphatidylinositol phosphorylation	NIL
AGAP001048	NIL	NIL	NIL
AGAP001050	Golgi cisterna membrane	NIL	acetylgalactosaminyltransferase activity
AGAP001052RC	cytoplasm	ubiquitin-dependent protein catabolic process, protein deubiquitination	thiol-dependent ubiquitinyl hydrolase activity
AGAP001052RB	NIL	NIL	NIL

AGAP001061	NIL	NIL	Protein binding
AGAP001070	cell, integral component of membrane	cell redox homeostasis	NIL
AGAP001073RA	nucleus	regulation of transcription from RNA polymerase II promoter	nucleic acid binding, protein binding
AGAP001073RB	NIL	NIL	NIL
AGAP001076	Integral component of membrane	oxidation-reduction process	monooxygenase activity, iron ion binding, oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen, heme binding
AGAP001082	cytoplasm, lysosome	lipid metabolic process, sphingolipid metabolic process	G-protein coupled receptor binding, enzyme activator activity
AGAP001083	NIL	NIL	nucleic acid binding, protein binding
AGAP001084	peroxisome, integral component of membrane	lipid metabolic process, wax biosynthetic process, long-chain fatty-acyl-CoA metabolic process	fatty-acyl-CoA reductase (alcohol-forming) activity, alcohol-forming fatty acyl-CoA reductase activity
AGAP013341	NIL	NIL	NIL
AGAP001090	cytoplasm	vesicle-mediated transport, response to starvation	protein binding
AGAP001091	NIL	protein phosphorylation	protein kinase activity, binding, ATP binding
AGAP001094	nucleus	regulation of transcription, DNA-templated	DNA binding, transcription factor activity, sequence-specific DNA binding, ATP binding