MOLECULAR IDENTIFICATION AND CHARACTERIZATION OF *SPIROPLASMA* IN *ANOPHELES ARABIENSIS* FROM KENYA

TOWETT SHARON CHEPKEMOI

156/67604/2013

A THESIS SUBMITTED TO THE UNIVERSITY OF NAIROBI, CENTER FOR BIOTECHNOLOGY AND BIOINFORMATICS IN PARTIAL FULFILLMENT FOR AWARD OF THE DEGREE OF MASTER OF SCIENCE IN BIOTECHNOLOGY

NOVEMBER 2016

DECLARATION AND APPROVAL

I declare that this thesis is my own work, and to the best of my knowledge has never been submitted as proposed work of study or examined for the award of degree in any university

TOWETT SHARON CHEPKEMOI

Registration Number:156/67604/2013

Signature;..... Date:..13th November 2016.....

APPROVAL

This thesis has been submitted with our approval as supervisors

Martin Rono, PhD

KEMRI-Wellcome trust Programme (Kilifi)

Signature; Date;.....14th November 2016.....

.....

Isabella Oyier, PhD

Centre for Biotechnology and Bioinformatics(University of Nairobi)

KEMRI-Wellcome trust Programme (Kilifi)

Signature; Date; .....14/11/16.....

.....

Jeremy Herren, PhD

International Centre for insect physiology and Ecology(ICIPE)

Signature; Date; 14th November 2016

.....

ACKNOWLEDGEMENTS

I thank the Almighty God, in whom I have found guidance, health and strength to make this research a great success. I take this opportunity to express my gratitude to everyone whose support I got through constructive criticism and friendly advice during the project work.

I am greatly indebted to my supervisors Dr Jeremy Herren, Dr Isabella Oyier, Dr Martin Rono, for their continued support and guidance throughout this project. May God bless you in all your future endeavours. To Dr Juan Paredes, thank you for the support and more importantly the time you took to go through my thesis write-up.

This project could not have been successfully accomplished without the support of my colleagues Enock Mararo and Hellen Butungi who were quite helpful during the laboratory work. In addition, the support of other members of Emerging Infectious Diseases Laboratory (EID) cannot go unnoticed.

Many thanks to the University of Nairobi and the Centre for Biotechnology and Bioinformatics (CEBIB) where I was registered as a student for their invaluable support throughout the project period. I take this chance to thank International Centre for Insect Physiology and Ecology (ICIPE) as an institution and especially the capacity building for nurturing and strengthening my role as a research student and future scientist. I acknowledge the EID laboratory for their resources and above all for hosting me as a master's research student.

Finally, to my caring, loving and supportive parents Mr David Towett and Mrs Esther Towett my deepest gratitude. Your encouragement at all times are much appreciated and duly noted.

ABSTRACT

Vector-borne diseases (VBDs) are a devastating global health problem. Mosquitoes are amongst the most important vectors of human VBDs. Several measures have been put in place to manage and eliminate these vectors, however they have all faced a variety of setbacks and raising the need to develop other methods of control. The use of bacterial endosymbionts is a highly promising new method to explore for this purpose. This study aimed to identify and characterize *Spiroplasma,* a maternally transmitted endosymbiont in *Anopheles arabiensis* mosquitoes as a candidate to block vector transmission in Africa. The study involved the development and validation of a PCR-based pan-*Spiroplasma* detection procedure that can be used for the screening of *Spiroplasma* in other mosquitoes as well as other insects/vectors. The *Spiroplasma* detection strategy was utilized for the examination of *Spiroplasma* prevalence in natural *Anopheles arabiensis* mosquito. Miseq illumina sequencing was used for validation of the developed PCR-based method. Moreover, this study also investigated the diversity and prevalence of microsporidian protozoan parasites in natural *Anopheles arabiensis* populations. Microsporidia are amongst the most important mosquito parasites that can be transmitted both vertically and horizontally and studying the infection of microsporidia and *Spiroplasma* has the potential to give insights into the protection of *Spiroplasma* to the mosquito. Two strains of *Spiroplasma* were found in one sampling location (Mwea), that is *Spiroplasma insolitum*-type and *Spiroplasma melliferum*-type while mosquitoes collected from the other sampling site (Mbita) having no *Spiroplasma* infection. In Mwea, the *Spiroplasma insolitum*-type was abundant in females with an overall population prevalence of 2% while the *Spiroplasma melliferum*-type was found only in males with a prevalence of approximately 7%.In addition, Miseq illumina sequencing results showed the prevalence of *Spiroplasma insolitum* to be 0.02%. Analysis on the mosquito *ND5* mitochondrial *DNA* (mt*DNA*) gene showed that the two types of *Spiroplasma* were evenly distributed with the mt*DNA* haplotypes. The microsporidia infection rate varied between sites (a range of 9% to 35%). Notably, no samples had a coinfection of *Spiroplasma* and microsporidia. These results showed two strains of *Spiroplasma* circulating in the Mwea population with the possibility of being transmitted both horizontally and vertically, lack of coinfection with microsporidia suggested that the *Spiroplasma* found in mosquitoes confers protection to the mosquito against microsporidia.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

DNA	Deoxyribonucleic acid
˚C	Degrees Celsius
µl	Microlitres
ABI	ABI instruments
BLAST	Basic Local Alignment
BLASTN	Nucleotide Basic Local Alignment
Ct	Cycle Threshold
CTAB	Cetyl trimethylammonium bromide
DCV	Drosophila C Virus
EDTA	Ethylenediaminetetraacetic acid
EID	Emerging Infectious Diseases
ELISA	Enzyme-linked Immunoassay
hd	Haplotype diversity
hr	Hour
HRM	High Resolution Melting
HRM	High Resolution Melting
ICIPE	International Centre of Insect Physiology and Ecology
idt	Intergrated DNA technology
ITS	Internally Transcribed Spacer
mins	Minutes
ml	Millilitres
MSROs	Melanogaster Sex Ratio Organisms
mt*DNA*	Mitochondrial *DNA*
MUSCLE	Multiple Sequence Comparison by Log-Expectation
Nacl	Sodium Chloride
NADH-ND5	NADH dehydrogenase 5
ND5	NADH dehydrogenase 5
NSROs	Nebulosa Sex Ratio Organisms
PCR	Polymerase Chain reaction
qPCR	Quantitative Polymerase Chain reaction
r*RNA*	ribosomal Ribonucleic Acid
SDS	Sodium dodecyl sulfate
secs	Seconds
SROs	Sex Ratio Organisms
ssr*RNA*	Small subunit ribosomal Ribonucleic Acid
TRIS	Trisaminomethane
UPGMA	Unweighted Pair Group Method with Arithmetic Mean
VBDs	Vector Borne Diseases
WSRO	Willistoni Sex Ratio Organism

CHAPTER 1

1.0 INTRODUCTION

1.1 Mosquitoes as disease vectors

Demographic and climatic changes have been linked to the re-emergence and spread of several vector-borne diseases (VBDs). These diseases are a huge burden globally and in Africa in particular, as evident in Figure 1 showing vector borne diseases hitmap. Insect-transmitted diseases cause over 1 million deaths annually and account for 17% of the infectious diseases [1]. Mosquitoes are amongst the most significant vectors of VBDs affecting nearly 700 million people each year with a mortality rate of 1 million[2].

The most common diseases transmitted by mosquitoes include; Malaria, Dengue fever, West Nile fever, Chikungunya, Zika virus disease and Yellow fever[3–5]. Malaria is the most deadly causing an estimated 400,000 deaths and 214 million new cases in 2015 alone[6]. It is transmitted by anopheline mosquitoes mainly *Anopheles gambiae sensu stricto*, *Anopheles arabiensis and Anopheles funestus*[7]. *Anopheles arabiensis* and *Anopheles gambiae sensu stricto* are morphologically identical members of the *Anopheles gambiae* complex[8]

Figure 1: The Global burden of Vector Borne Diseases: Image adapted from WHO Health report, 2004. Map shows the number of deaths caused by VBDs in world.

1.2 Vector Borne Diseases Control

The elimination/management of VBDs has been an area of intense research focus for many years. Research programmes have developed several methods to control VBDs but most have faced great challenges[1].

Vaccines against most of these diseases have not been successfully developed. For instance, the complex life cycle of *Plasmodium spp.* (malaria causing parasite) have largely hindered the development of an effective vaccine [10]. Vaccines against dengue are also in development with the most advanced showing protection against three of the four dengue serotypes only[1].

Chemical and environmental controls involving clearing mosquito breeding sites, indoor and outdoor spraying and the use of insecticide treated bed-nets (ITNs) have also faced numerous setbacks including the emergence of insecticide resistant vectors and changes in mosquito foraging habits[11]. For these reasons, there is a great need to develop a new method of vector control that is manageable and environment-friendly[12]. One of the most promising new avenues is the use of bacterial endosymbionts[13].

1.3 Endosymbionts

Insects have developed a long-term relationship with bacteria living within them, these have been instrumental in the success of insects including their ability to colonize diverse habitats[14]. In addition to their intestinal microbiota, many insects also harbour endosymbiotic bacteria and depending on their effects on host fitness, endosymbionts can be mutualistic, commensal or parasitic[14,15].

In addition to these distinctions, most endosymbionts fall into one of two broad categories; obligate and facultative. Obligate endosymbionts are those that have a long term obligate relationship with the host, are vertically transmitted and are required to support insect development[16]. For example, *Buchnera* found within aphids[17]. The other class of endosymbionts are facultative endosymbionts which are not essential for host survival, their effects can either be detrimental or advantageous[14].

Wolbachia, a facultative endosymbiont has been studied widely in the context of VBD transmission blocking. This is because *Wolbachia* is known to protect its hosts from infections (primarily viral infection) and also induce a reproductive manipulation known as cytoplasmic incompatibility (CI). CI results in reduced fertility when an infected male mates with an uninfected female ,this manipulation enables spreading of the bacterium through host

populations[5,18]. *Wolbachia*'s ability to cause CI gives a relative advantage to *Wolbachia*-infected females, since they can mate with either infected or uninfected males (*Wolbachia*-uninfected females can only mate with *Wolbachia*-uninfected males). The relative increase in fitness experienced by *Wolbachia*-infected females is dependent on the prevalence of *Wolbachia* within the population. Studies suggest that once *Wolbachia* prevalence crosses a threshold, the relative increase in fitness is sufficient to drive its prevalence to near 100%[19].

Several strains of *Wolbachia* (*wMel and wMelPop*) have been successfully introduced into *Aedes aegypti* the major Dengue vector. *Wolbachia*-infected *Aedes aegypti* are unable to transmit dengue, this finding is being exploited in the control of Dengue fever in the field[20]. Although the precise mechanism by which *Wolbachia* protects mosquitoes against dengue virus infection is still largely unknown, there is evidence suggesting that *Wolbachia*-pathogen blocking is correlated to bacterial density and tissue distribution[13], which may indicate competitive exclusion. In addition to reducing host longevity *wMelPop* also upregulates the mosquito's innate immune system, which could be a factor that contributes to *Wolbachia* pathogen-blocking[21].

Endosymbionts can also influence evolution of their host by affecting the population genetics of mitochondrial *DNA*[22]. This is because mitochondria and endosymbionts are maternally transmitted in the host egg cytoplasm to its offspring[23]. When endosymbionts increase the fitness of the maternal lineage they inhabit, associated mitochondrial haplotypes can increase in prevalence due to 'hitchhiking'. An example of this is observed in the bird nest blowfly *Protocalliphora sialia,* where mitochondrial haplotypes associated with *Wolbachia* infection 'hitchhike' to high prevalence[24]. Studies have shown that male-killing endosymbionts decrease the diversity of host mitochondrial *DNA* by increasing the frequency of the mitochondrial *DNA* haplotype transmission associated with the endosymbiont[25].

Spiroplasma, another common maternally transmitted endosymbiont, is known to have the ability to protect its hosts against parasite infections and also manipulate host reproduction to enhance their transmission making it a favourable candidate for vector control[26,27].

1.4 Microsporidia

Microsporidia are important parasites of several organisms ranging from animals to arthropods[28]. These obligate parasites cause diverse effects to their hosts for instance it affects host metabolism and distorts the reproductive system[29]. They are highly specialized microorganisms with a unique mechanism of colonizing host cells using their infectious spores[30]. Microsporidia parasites affect a wide range of mosquito species and they cause late male killing in mosquitoes. This is by over proliferating in the host thus killing mosquito larvae before pupating[31]. The study of microsporidia gives an insight into what species of microsporidia are found in *Anopheles arabiensis* and also their prevalence in relation to *Spiroplasma*

1.5 Problem Statement

Sub-Saharan Africa suffers disproportionately from a burden of VBDs. Drug resistant pathogens and insecticide resistant vectors have led to the resurgence and increase of these VBDs[32]. Presently, effective vaccines for most of these VBDs have not been developed, whilst the ones already in place have not been potent, thus the primary tool for intervention is vector control[33]. The current vector-control strategies have proved to be inefficient due to the emergence of insecticide resistant vectors and unpredictable vector feeding behaviours which renders indoor spraying and the use of ITNs less effective. Therefore, a better method that is ecologically and environmentally sustainable is needed[12].

Endosymbiotic bacteria have been studied as candidates for vector control. This is due to their ability to protect their hosts from parasite infection, in addition, their maternal transmission enables sustained presence across host generations[16]. These qualities render endosymbionts a potentially sustainable and effective means to limit the transmission of VBDs. *Spiroplasma* is a potential candidate for vector control and merits further investigation. Identification of the best strategy for detecting the presence of *Spiroplasma* and its strain type in *Anopheles arabiensis* is a key step to achieving the long term goal of using *Spiroplasma* to limit VBD transmission. Furthermore, studying the relationship between *Spiroplasma* and microsporidia will show the effect of *Spiroplasma* on the parasites and pathogens affecting its host, this knowledge can be further applied in studying the relationship between *Spiroplasma* and other protozoan parasites like *Plasmodium*.

1.6 Justification

Due to the heavy burden of VBDs, the development of sustainable methods of control is very important. However, the methods in place to control and eliminate these VBDs have faced a variety of setbacks. For instance, diseases such as dengue, yellow fever and malaria have proven to be very difficult targets for vaccine development. Development of vaccines against dengue has been challenging due to its several serotypes while the development of vaccines to target the different stages of malaria has proved futile. Other methods, including environmental, chemical and physical control have been effective only under certain circumstances that include protection at night when using insecticides treated bed-nets (ITNs). For instance, they only target indoor feeding mosquitoes and none are available to target outdoor mosquitoes, this has therefore necessitated the search for better methods.

The use of endosymbiotic bacteria has been shown to be a successful alternative method for controlling vector-borne diseases[9]. While *Wolbachia* is the well-studied endosymbiont in this context, *Spiroplasma* has also demonstrated potential as a candidate for VBD control. *Wolbachi*a shows protection to its host against viruses[34,35], while endosymbiotic *Spiroplasmas* provide their hosts with protection against parasites, making it an ideal candidate to curb transmission of malaria[36]. These reasons prompted the study of *Spiroplasma* in *Anopheles arabiensis* to determine its prevalence and its relationship with other microorganisms.

A study done in Mbita Kenya reported the presence of *Spiroplasma* in *Anopheles funestus* mosquitoes[37], alluding to the possibility of finding *Spiroplasma* in *Anopheles arabiensis* which is a close relative of *Anopheles funestus* a significant carrier of *Plasmodium,* the causative agent of malaria. Moreover, studying the relationship between *Spiroplasma* and microsporidia can give a better understanding on parasite-protective effects of *Spiroplasma*, this is because microsporidia are amongst the most important natural parasites of mosquitoes[29]. Generally, this study serves as a foundation for the development of a method for screening and characterizing *Spiroplasma* in mosquitoes.

Determining the *Spiroplasma* strains naturally inhabiting *Anopheles arabiensis,* their prevalence and effect to other mosquito microbiota is important in understanding *Spiroplasma* for vector transmission blocking strategies.

1.7 Objectives

1.7.1 Main Objective

- To identify and characterize *Spiroplasma* in *Anopheles arabiensis* mosquitoes

1.7.2 Specific objectives

1) To design and test general PCR-based assays for detecting *Spiroplasma*
2) To determine the prevalence of *Spiroplasma* species in *Anopheles arabiensis* mosquitoes obtained from the field
3) To investigate the population dynamics of *Spiroplasma* species in *Anopheles arabiensis* mosquitoes by determining the mitochondrial *DNA* haplotypes associated with infected and non-infected samples
4) To correlate *Spiroplasma* infections with microsporidia infections in the collected *Anopheles arabiensis* samples

CHAPTER 2

2.0 LITERATURE REVIEW

2.1 Mosquito Vectors and Vector Disease control

Mosquitoes make up a large percentage of vectors of parasitic and viral pathogens. These pathogens include malaria (*Anopheles*), zika virus (*Aedes*), dengue virus (*Aedes*) and filariasis (*Mansonia* and *Culex*). Diseases caused by these pathogens are devastating to the whole world. Africa is majorly affected by these diseases with an annual mortality rate of 100-500 million (Figure1)[38].

Control of these VBDs has been an area of great interest, however, the methods that have been developed are currently facing a variety of setbacks. For instance, development of vaccines against malaria has been slow since it requires multiple vaccines targeting the different stages of the parasite[10]. However, recently the RTS S/AS01 vaccine in phase III is promising showing 50% disease reduction in African children[39]. Vaccines against dengue have shown good progress, that is protection against only three dengue serotypes and no protection to the fourth serotype, this renders the vaccine not effective since it has to be equally successful to all serotypes[1].

These reasons have made vector control to be the only promising strategy nevertheless, this method is also facing challenges. Indoor and outdoor spraying to control mosquitoes has been rendered ineffective due to the emergence of resistant vectors. The use of ITNs has been effective but this is only limited to prevent mosquitoes biting at night and only when one is using the ITNs and in addition mosquitoes have changed their foraging behaviour[11].

Genetic modification of mosquitoes has also been promising but maintenance of the genetically modified adults once released into the field has proved to be challenging since the modified mosquitoes often have reduced fitness[13].These reasons have raised a need to develop an alternative method of vector control. Bacterial endosymbionts have been a promising strategy to control VBDs.

In the context of vector competence, bacterial endosymbionts have the capacity to make their vector hosts resistant to infection by agents of human diseases[18]. In addition, vertical transmission means that bacterial endosymbionts are inherited in insect populations over generations, which renders them a more sustainable approach than many of the currently used control methods due to their efficient transmission and spread[40].

Wolbachia and *Spiroplasma* are the most prevalent and well known bacterial endosymbionts, they affect approximately 40% and 5–10% of insect species, respectively[41,42]. Since they are maternally transmitted in the egg cytoplasm, some facultative endosymbionts have evolved phenotypes to manipulate the host's reproductive system to increase the fitness of infected females at the expense of their male counterparts[43]. These traits are expressed in phylogenetically diverse groups of endosymbiotic bacteria including: *Spiroplasma, Wolbachia, Rickettsia, Arsenophonus* and *Cardinium*[44,45]. Diverse reproductive manipulations induced by endosymbionts include cytoplasmic incompatibility, male-killing, feminization of males and parthenogenesis[44,46].

Some facultative endosymbionts are known to be 'protective'. These protective effects on hosts are diverse and include; i) protection against viruses[34,35] ii) protection against eukaryotic parasites[27,36] and iii) protection against environmental stress[47]. These effects not only favour the spread of the bacteria through insect populations but are also important for insect evolutionary ecology and the dynamics of acquisition of ecologically favourable traits like thermal tolerance[48]. The facultative endosymbionts that protect their hosts against parasites and pathogens are potential candidates for VBD control, these endosymbionts include *Wolbachia* and *Spiroplasma*[15,49]. This study therefore focuses on studying *Spiroplasma* as a vector control candidate.

2.2 Spiroplasma

Spiroplasma are facultative, motile, wall-less bacterium of the family *Spiroplasmataceae* of the *Mollicutes* class, related to *Mycoplasma* and *Phytoplasma*[50]. *Spiroplasma* exploit numerous habitats but are mostly associated with plants and arthropods[51]. *Spiroplasma* are thought to be found in a wide range of insect species, heritable *Spiroplasma* infect approximately 5-10% of insects while non-heritable *Spiroplasma* have a prevalence of >50%[52].

Spiroplasma have different modes of survival ranging from mutualism to parasitism, however, most of the characterized *Spiroplasma* appear to be insect commensals. Some *Spiroplasma* can be harmful to their hosts, notable examples include *Spiroplasma citri* and *Spiroplasma phoenecium* that are pathogens of citrus plants and periwinkles, respectively[53]. Additionally, the species *Spiroplasma culicicola* and *Spiroplasma taiwanese* are pathogenic to mosquitoes[54].

2.3 Endosymbiotic *Spiroplasma*

Heritable *Spiroplasma* strains have been discovered in numerous insect orders including Diptera, Lepidoptera and Coleoptera[55–57]. *Spiroplasma* are found in 17 species of *Drosophila* and one of the most well-studied heritable *Spiroplasma* strains is the MSRO *Spiroplasma poulsonii* strain harboured by *Drosophila melanogaster*[58]. This strain of *Spiroplasma* was isolated from wild *Drosophila melanogaster* obtained in Uganda[58]. Vertically transmitted *Spiroplasma* associated with *Drosophila* as well as other insects are fastidious and difficult to culture outside of their hosts[59] and can protect the host and manipulate host reproduction[27,59].

Endosymbiotic *Spiroplasma* are transmitted maternally and hence their survival entirely depends on their host. They persist in the hemolymph and achieve maternal transmission by getting into the cytoplasm of nascent oocytes[60]. During maternal transmission the endosymbiont colonizes the germ line and takes advantage of the yolk uptake machinery to reach the oocyte. *Spiroplasma* goes through the intercellular space that surrounds the ovarian follicles and is subsequently taken up together with the yolk granules[60].

While vertical transmission of *Spiroplasma* is highly efficient, it remains imperfect. Therefore, *Spiroplasma* must have additional strategies to compensate for this imperfect transmission and in order to maintain its prevalence in insect populations. *Spiroplasma* protects against macro-parasites and manipulates its host's reproductive system (e.g. male-killing [27,61]). This selective pathogenicity by *Spiroplasma* reduces the number of males in brood, and theories suggest that this reduces sibling competition for resources and inbreeding hence resulting in competent female offspring[23]. Since *Spiroplasma* are only transmitted down the female lineage, this re-partitioning of fitness from males to females can increase their prevalence in subsequent generations[62]. The mechanistic basis of male-killing by *Spiroplasma* has not been fully elucidated. However a recent study in *Drosophila melanogaster* suggests that *Spiroplasma* targets the dosage compensation system that leads to increased epithelial cell death and distorted central nervous system development[63].

Male-killing *Spiroplasma* have been found in different insect species, for instance ladybird beetles[56] and *Drosophila*. Notable examples include; NSRO (Nebulosa sex ratio organism) from *D. nebulosa*, MSRO (Melanogaster sex ratio organism) from *D. melanogaster* and WSRO (Willistoni sex ratio organism) from *D. willistoni*[52]. Some *Spiroplasma* strains harboured by members of the genus *Drosophila*, and other insects, do not cause male-killing[64].

Endosymbiotic *Spiroplasmas* are known to confer their hosts with protection against parasites and pathogens [65]. *Spiroplasma* SPHY confers on its host, *Drosophila hydei* increased resistance to two common parasitic wasps *Leptopilina heterotoma* and *Leptopilina boulardi*[27]. In other species of *Drosophila* like *Drosophila neotestacea*, *Spiroplasma* protects against *Howardula aoronymphium,* a nematode that causes sterilization[49]. The mechanism of *Spiroplasma* protection has not been fully understood,but possible explanations include i) presence of *Spiroplasma*-encoded substance that is toxic to the nematode[66] and ii) competition for nutrients and resources that are important for nematodes survival as for the case with protection of *Hamiltonella defensa* to aphids against parasitoids [67]

The possibility of *Spiroplasma* to be maternally transmitted and its ability to protect its hosts against pathogens and parasites makes it a good candidate to control VBDs[68,69]. In addition, maternal transmission of mitochondrial *DNA* has served as an excellent tool for studying evolutionary processes at the host population level especially hosts that harbour endosymbiotic bacteria. Mitochondrial *DNA* is strictly vertically transmitted, whereas endosymbionts are generally vertically transmitted and on occasion they can be horizontally transmitted[70], this can hence be used to study the effect of an endosymbiont on its host by studying its mitochondrial *DNA*.

In this study, mitochondrial *DNA* haplotypes were correlated with *Spiroplasma* infection to infer the likely importance of horizontal and vertical transmission of *Spiroplasma*. This is because in most cases strictly vertically transmitted endosymbionts tend to be confined in only one haplotype due to hitchhiking that can lead to fixation[22,24].

2.4 *Spiroplasma* and Mosquito vector diseases

Mosquitoes constitute the most important group of insect vectors of human diseases. Mosquito vectors are hosts of various types of *Spiroplasma*[54]. *Spiroplasma culicicola* was the first to be isolated, from *Aedes sollicitans* a salt marsh mosquito collected in New Jersey, USA[71]. Others include; *Spiroplasma sabaudiense* isolated from a mixed pool of *Aedes sticticus* and *Aedes vexans* collected in the French Alps [72], *Spiroplasma taiwanense* from *Culex tritaeniorhynchus*[73], *Spiroplasma cantharicola* and *Spiroplasma diminutum* were isolated from *Culex annulus* and *Culex tritaeniorhynchus* respectively[74].

Thus far, most of the *Spiroplasma* strains isolated from mosquitoes are pathogenic and unlikely to be vertically transmitted[75] . Infection of *Aedes albopictus* with *Spiroplasma diminutum* results in bacterial proliferation, although in this case there is no detrimental effect on

mosquitoes lifespan under laboratory conditions[76]. In contrast, *Spiroplasma taiwaniese* infection of *Anopheles albopictus* reduces the survival of larvae and reduces the lifespan of female adults[68].

It was also demonstrated that *Spiroplasma. taiwanese* replicates both intra- and extracellularly in *Aedes aegypti* and *Anopheles stephensi*. In this system, *Spiroplasma* can be observed in the hemolymph, thoracic flight muscles, hemocytes and the neural system[77]. Replication in the thoracic flight muscles leads to impaired mobility and loss of flight ability associated with excessive cell lysis and polysaccharide depletion and this eventually shortens the host's lifespan[58]

Another *Spiroplasma* strain was identified in *Anopheles funestus* one of the malaria vectors in a study done in Lwanda, East of the ICIPE Thomas Odhiambo Campus, Mbita, in Western Kenya[37]. Based on the 16S ribosomal *DNA* sequence, this strain appears to be closely related to *Spiroplasma ixodetis*.

2.5 Microsporidia

Microsporidia are a diverse group of single celled eukaryotic intracellular parasites with approximately 200 characterized genera[78]. They are highly prevalent in various animal groups, including fish and arthropods and are obligate parasites, which likely explains their fast evolving and highly reduced genome[28].

The microsporidian spore is the driving force of infection and is the only distinct stage of microsporidia that can survive outside of the host cell[79]. The spores have a small distinct size (2-20µm) with thick walls made up of exospore and endospore. The three principal spore structures of infection are: the posterior vacuole, polar filament and polaroplast that occupies the anterior part of the spore[28].

In invertebrates, microsporidia transmission is either horizontal or a combination of horizontal and vertical. *Nosema apis* is an example of a horizontally transmitted microsporidia and its transmission depends on the release of spores into the environment for ingestion by the next host[30]. Horizontally transmitted microsporidia tend to be more virulent to their hosts[80].

In contrast, vertically transmitted strains are less virulent to the host and (much like bacterial endosymbionts) have complex strategies to colonise the host's germ line. For instance, *Nosema granulosis* is transmitted vertically *in Gammarus duebeni* [81]. *Nosema granulosis* is less virulent to its host and is primarily localized to the host's gonads where it is vertically transmitted to

oocytes during vitellogenesis. In addition, *Nosema granulosis* can manipulate its host's reproduction by causing feminization[82].

2.6 Microsporidia and Mosquitoes

Microsporidia in mosquitoes can be classified into two categories based on their lifestyle and interaction with their host[80] Approximately 90 isolates have been characterized in 79 different mosquito species, mainly affecting the following genera: *Aedes, Aedeomyia, Anopheles, Coquilletidia, Culex, Culiseta, Mansonia, Ochlerotatus* and *Psorophora*[83]. In mosquitoes, microsporidia can be transmitted horizontally or both horizontally and vertically. Vertically transmitted microsporidia have been associated with late male-killing and feminization of their mosquito hosts[84].

Some species of microsporidia affect one generation of mosquitoes and are not host or tissue specific. These species have a simple life cycle that involves the release of one spore that takes part in horizontal transmission**.** *Vavraia culicis* is an example of this type of microsporidia that infects a wide range of mosquito species including *Anopheles* and *Culex*[85]. *Takaokaspora nipponicus* is another species of microsporidia isolated from *Ochlerotatus japonicus japonicus* mosquito and is transmitted both vertically and horizontally[86].

CHAPTER 3

3.0 MATERIALS AND METHODS

3.1 Study Site and Sample collection

This study was conducted at the Emerging infectious Diseases (EID) laboratory at ICIPE Duduville Campus. As depicted in Figure 2, samples were collected in Karima and Mbui-njeru Villages in Mwea (Central Kenya) 100km northeast of Kenya and Kirindo and Kinyege villages in Mbita Point, Western Kenya. During sampling, geographical co-ordinates were recorded and used to plot the map in QGIS v2.8.9[87]. Samples in the two sites were collected at different times, these was due to the fact that high mosquito prevalence times in the two regions are different. Methods used during mosquito collections were different in the two, they were selected depending on the best method that could catch a large number of mosquitoes in the specific site.

The Mwea region produces over 50% of Kenya's rice. Rice paddies and associated irrigation canals provide suitable breeding habitats for mosquitoes. Karima and Mbui-njeru villages are surrounded by rice paddies. Although both anopheline and culicine mosquitoes are prevalent in Mwea, the most abundant species is apparently *Anopheles arabiensis,* which represent greater than 53% of the mosquitoes[88]. Prevalence of malaria in Mwea is relatively low for reasons that are still not entirely clear. It has been suggested that this might be due to the abundance of *Anopheles arabiensis* which is known to feed preferentially on livestock rather than humans[89]. The area has relatively hot climate with temperatures ranging from 16-26°C with an average humidly of 50-66% and characteristic long rainfall in April to May and short rains in October and December. The season dictates mosquito species composition and abundance. *Anopheles arabiensis* is abundant during seedling transplantation and land preparation while C*ulex quinquefasciatus* are more predominant during short rains and the final stages of rice maturation [90]. Samples were collected from this site on 11th and 21st April and later on 11th May 2016 by aspiration of resting mosquitoes in houses.

The Mbita region lies along the shores of Lake Victoria in Homa Bay County. Studies conducted in Rusinga Island in the same county reported 10.9% malaria prevalence with characteristic infection of *Plasmodium falciparum, Plasmodium malariae* and *Plasmodium ovale*[91]. Up to 90% of the malaria vectors in this region are *Anopheles gambiae s.s* but studies show that this species is being replaced by *Anopheles arabiensis*[7,92]. In addition, *Anopheles*

funestus mosquitoes are also significant vectors of malaria in this region[92]. The current study was conducted in Kirindo and Kinyege villages located 5km from Mbita point where most people are fishermen and practise subsistence farming. Mosquito collections were done from 1st to 5th June 2015 using cattle baited-traps, where mosquitoes are lured into the trap using cow odour and trapped using a net. CDC-light traps were hanged indoors to trap mosquitoes flying towards the light and the mosquitoes were trapped in the collection container.

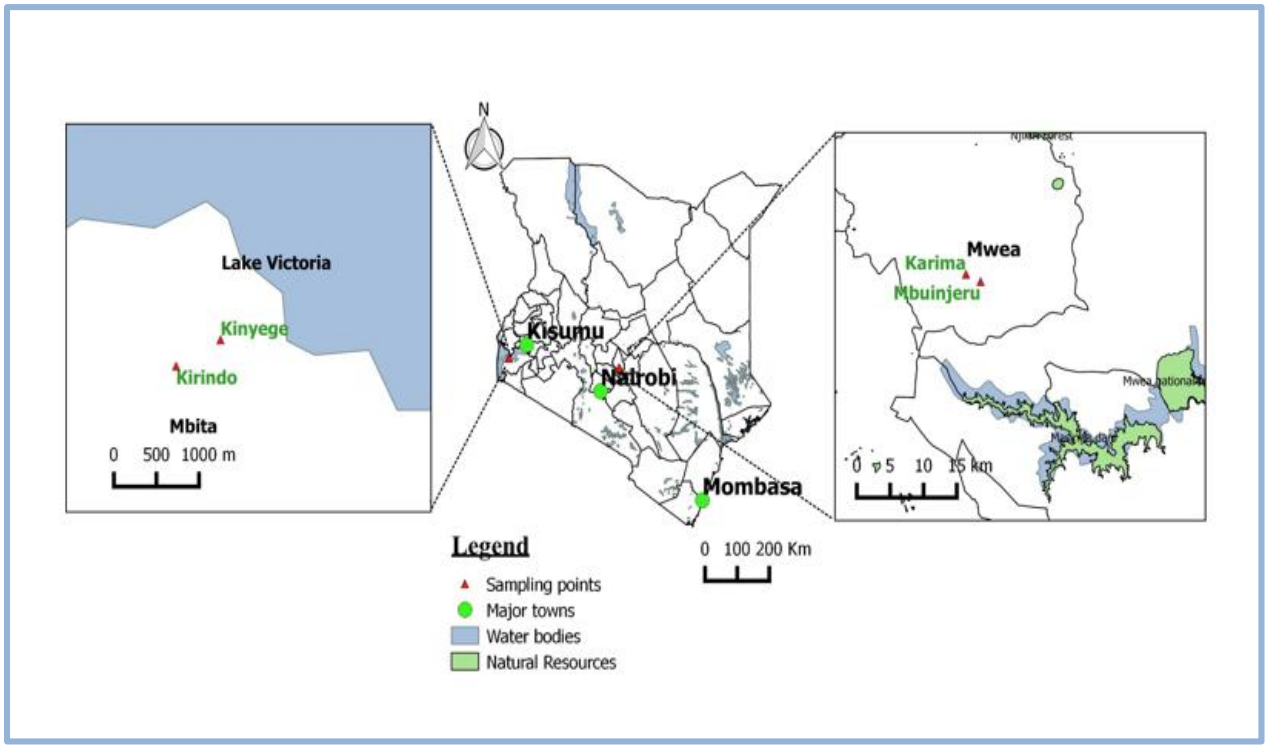


Figure 2: Map of Kenya showing mosquito sampling sites and locations

3.2 Mosquito identification and *DNA* extraction

Samples collected from the two sites, Mwea (*n*=385) and Mbita (*n*=357) were first morphologically identified using a key by Gillies and Coetze that guides in the identification of mosquito genera , sub genera and species[93].To differentiate between *Anopheles gambiae s.s* and *Anopheles arabiensis* the mosquitoes were further identified molecularly. *DNA* extraction was done using two methods; 1) modified protein precipitation method [94] and 2) modified CTAB (cetyl trimethylammonium bromide) protocol [95]. This was to determine which of the two methods would produce good quality DNA for detection of *Spiroplasma*.

The protein precipitation method involved grinding one adult mosquito in 300µl of lysis buffer (10mM trisaminomethane (TRIS), 0.5% sodium dodecyl sulphate (SDS), 1mM Ethylenediamine tetraacetic acid (EDTA) in a 1.5ml Eppendorf tube then incubated at 65°C

for 30mins to allow for lysis to occur. 100µl of protein precipitate solution (8M ammonium acetate, 1mM EDTA) was added, the mixture was vortexed and centrifuged at 19000 x g. The supernatant was transferred to a new 1.5ml tube and mixed with isopropanol (2:1), which results in *DNA* precipitation. An additional centrifugation step results in *DNA* forming a pellet at the bottom of the tube. The supernatant is then pipetted off and 300µl of 70% ethanol added. The mixture was once more centrifuged at max speed of 19000 x g and supernatant pipetted off. The samples were air dried by inverting the tube for 10mins after which 100µl of de-ionized water was added and stored at -20°C.

For the modified CTAB method whole mosquitoes were ground in 250µl of Tris EDTA buffer. 15µl of SDS and 1.5µl of 25mg/ml of Proteinase K was added and mixed thoroughly. After 1hr of incubation 25µl of 5M Sodium Chloride (NaCl) was added and thoroughly mixed, 20µl of CTAB/NaCl solution was then added and incubated for 30mins at 65°C. An equal volume of chloroform/isoamyl alcohol (24:1) was added, mixed and centrifuged at maximum speed of 19000 x g at 4°C for 5mins. The top aqueous layer was transferred into a new tube, an equal volume of phenol/chloroform /isoamyl alcohol (25:24:1) was then added, mixed and centrifuged at max speed of 19000 x g. The supernatant was transferred to a new tube and 1 volume of isopropanol was added and incubated for 5mins. After incubation the mixture was centrifuged at maximum speed of 19000 x g for 30mins, isopropanol was removed and 70% ethanol added and centrifuged at 4°C for 15mins, residual ethanol was removed and tubes left to air dry. Thereafter the pellet was re-suspended in 50µl-100µl of water and stored at -20°C.

Four mosquitoes from each site (Mwea and Mbita) were dissected under a dissecting microscope (Leica) and *DNA* extracted from their ovaries. These *DNA* was pooled together with *DNA* extracted from 6 whole mosquitoes collected from respective sites into distinct (Mwea and Mbita pools)

3.3 Primer design for all *Spiroplasma* detection

rpoB gene, encoding the beta sub-unit of the *RNA* polymerase was chosen for the detection of *Spiroplasma*. This is because it occurs in single copies and contains both highly conserved and variable regions[96]. As illustrated in Figure 3, universal primers targeting 313bp were designed by first aligning the partial *rpoB* sequences of *Spiroplasma poulsonii, Spiroplasma ixodetis, Spiroplasma taiwanese, Spiroplasma syrphidicola, Spiroplasma melliferum*. *Spiroplasma. apis, Spiroplasma citri, Spiroplasma chrysipicola* and *Spiroplasma diminitum.* Thereafter, highly conserved regions of these sequence were manually selected for the design of the primers and the 9 internal mismatches were made degenerate to ensure universality while maintaining annealing specificity and sensitivity. This was carried out in Geneious sequence analysis software, v8.05 [97]. As illustrated in table 1, a set of primers RPOB3044F_ALL and RPOB3380R_ALL were designed, these primers were checked for dimer and hair-pin formation using online algorithms Integrated Device Technology (www.idt.com).

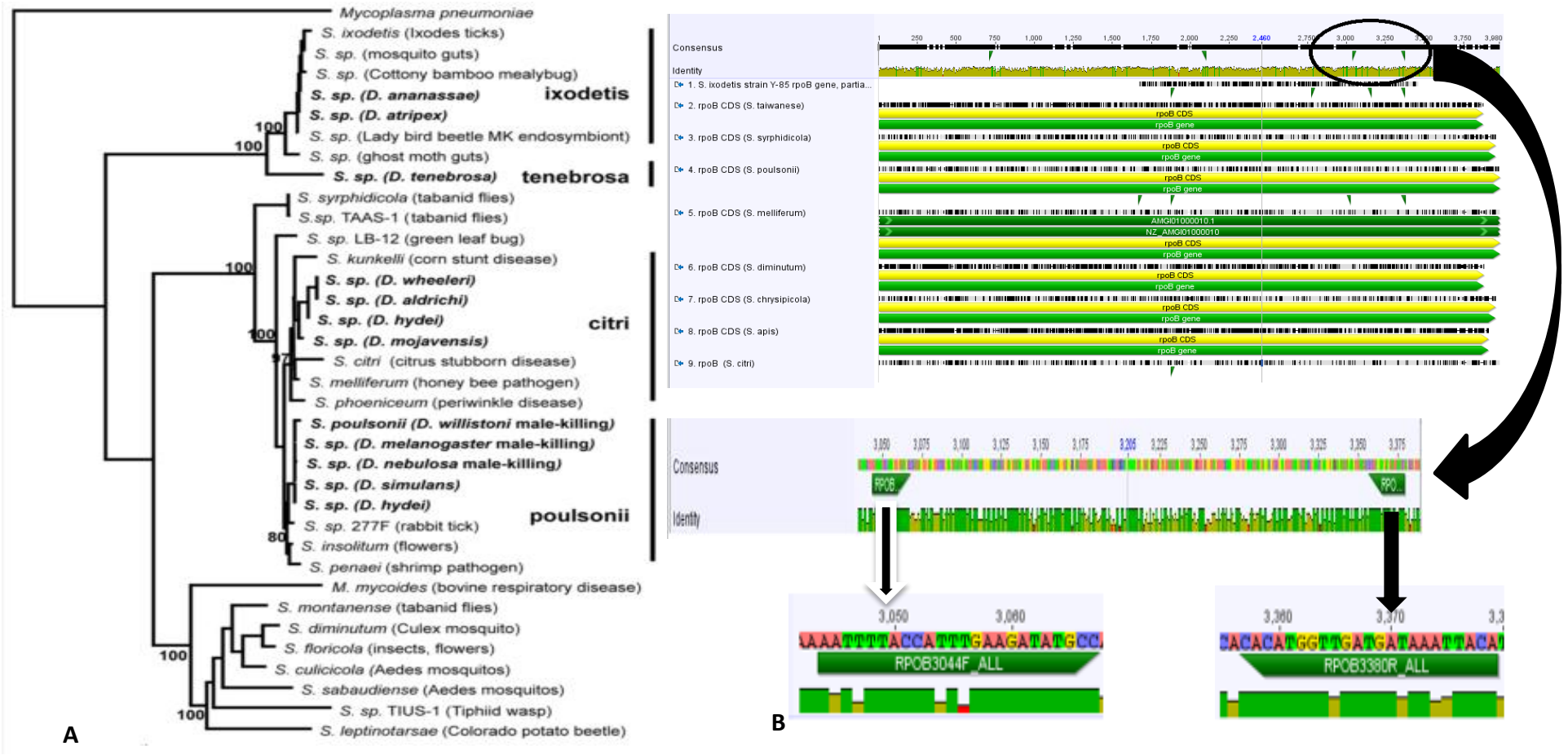


Figure 3: Primer design: A, General *Spiroplasma* phylogenetic tree showing *Spiroplasma* Clades used as a reference for primer design ([107], B Universal primer design diagram based on the phylogenetic tree

Table 1: Primer Table: Shows the primers used in the PCR procedures

Primers	Sequence	Target Species	Reaction	Annealing T_m	Reference
27F **519R**	AGAGTTTGATCCTGGCTCAG GWATTACCGCGGCKGCTG	16S rRNA (Bacteria)	454 MiSeq illumina sequencing	53°C	[98]
RPOB3044F_ALL **RPOB3380R_ALL**	ARTHTTACCADTDGAAGATATGCC TGTARYTTRTCATCWACCATGTG	*rpob* (*Spiroplasma)*	PCR	53°C	This study
FTSZIXOF **FTSZIXOR**	TGTTGCTAATACTGATGCACAAG AATGTCATTGTTGTTCCACCAGTAAC	*ftsz* (*Spiroplasma ixodetis*)	PCR	56°C	This study
RPOBINSPOUL **RPOB3380INS**	AATTTAACCATTAGAAGATATGCC TGTAATTTATCATCAACCATGTG	*rpob* (*Spiroplasma.poulsonii*)	PCR	59°C	This study
RPOB3044FINSPOU **RPOB3380CITRI**	AATTTACCATTAGAAGATATGCC AATTTTACCATTGGATATGCC	*rpob* *(Spiroplasma citri)*	PCR	58°C	This study
19CL **DMP3A**	CTCCACCAATTACTATAACAG AGGATGAGATGGCTTAGGTT	*ND5* *(*mosquito*)*	*PCR*	55°C	[99]
ss18sf **ss1492r**	GTTGATTCTGCCTGACGT GGTTACCTTGTTACGACTT	*ssrRNA* *(*microsporidia*)*	PCR	50°C	[100]

3.4 16S ribosomal RNA High-throughput screening

The pooled *DNA* extracted from whole mosquitoes and mosquito ovaries (Mwea and Mbita) were sent to the Research and Testing Laboratory in Lubbock, Texas for amplification with universal 16S primers followed by High throughput MiSeq illumina sequencing. The primers used for amplification of 16S ribosomal *RNA* genes are listed in Table 1.

Samples were amplified in a two-step process in a total volume of 25µl using Qiagen Hotstart Taq mastermix mix (Qiagen Inc, Valencia, California) containing DNA Polymerase, dNTPs, MgCl2, KCl and stabilizers, 1µl of each 5µM primer, and 1µl of template. Reactions were performed on ABI Veriti thermocyclers (Applied Biosytems, Carlsbad, California). The following PCR cycling conditions were used 95°C for 5 min, then 25 cycles of 94°C for 30 sec, 54°C for 40 sec, 72°C for 1 min, followed by one cycle of 72°C for 10 min and 4°C hold.

Amplification products were visualized with eGels (Life Technologies, Grand Island, New York). Products were then pooled in equimolar concentrations and each pool was selected using Agencourt AMPure XP (BeckmanCoulter, Indianapolis, Indiana). The selected pools were then quantified using the Qubit 2.0 fluorometer (Life Technologies) and loaded on a MiSeq Illumina (Illumina, Inc. San Diego, California) 2x300 flow cell at 10pM.

3.5 Polymerase chain reaction

3.5.1 Mosquito identification

PCR was used to determine the *Anopheles gambiae* subspecies since they are morphologically identical. This was achieved using a high resolution melting PCR as previously described[101]. Briefly, we used a 10µl reaction volume and primers targeting the internally transcribed Spacer region (ITS) ITS_Zianni_F (5'-GTG AAG CTT GGT GCG TGC T-3') and ITS_Zianni_R (5'-GCA CGC CGA CAA GCT CA-3'). We used *Anopheles gambiae s.s* and *Anopheles arabiensis* positive controls obtained from ICIPE insectary that indicted the samples' subspecies. The PCR cycling conditions included initial enzyme activation at 95°C for 15mins, followed by 35cycles of denaturation at 95°C for 30secs, annealing at 57°C for 30secs, elongation at 72°C for 30secs and finally a hold temperature of 72°C for 1min.

3.5.2 Universal *Spiroplasma* detection

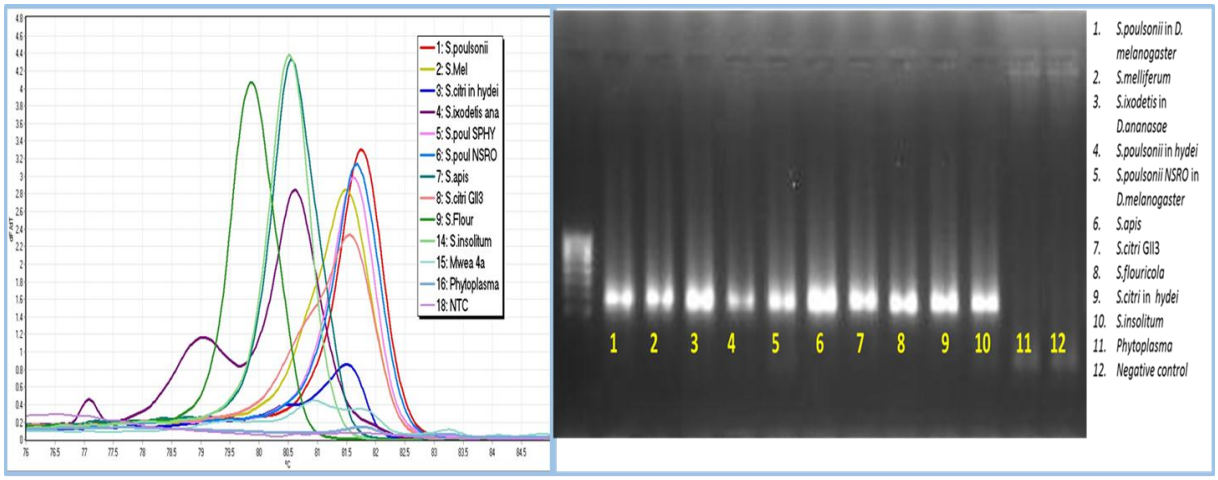
The annealing temperature was established using a gradient PCR with the universal primers. To determine the sensitivity and specificity of these universal primers a conventional and real-time PCR was performed on all the *Spiroplasma* controls available and including *Phytoplasma* a close relative of *Spiroplasma* (Illustrated in Figure 4). Calibration was done using *Spiroplasma poulsonii* control, six-fold serial dilutions of the control was subjected to a real-time PCR with the universal primers for quantitative analysis where the cycle threshold (Ct) values (number of cycles it takes for a signal to be detected from a sample) were used to plot a standard curve and determine the standard primer Ct value for screening.

Spiroplasma detection was initially carried out by screening the samples with the universal primers (RPOB3044F_RPOB3380R, Table 1 and Figure 5). This was performed in a 10µl reaction volume that included 5X Hot Firepol Evagreen HRM Mix (Solis BioDyne, Tartu, Estonia) and 1µl of *DNA* template. The Rotor Gene Q cycler (Qiagen) quantitative PCR (qPCR) machine was used. The cycling conditions included initial enzyme activation at 95°C for 15mins, followed by 35 cycles of denaturation at 95°C for 30secs, annealing at 53°C for 30secs, elongation at 72°C for 30secs then hold temperature of 72°C for 10min.

Subsequently, PCR amplicons were subjected to melting by gradually increasing the temperature in 0.1°C increments from 65°C to 90°C and recording and plotting changes in the fluorescent intensity with changes in temperature (dF/dT) and the melting profiles were assessed using Rotor-Gene Q series software 2.1.0 (Build 9). Melting curves are unique based on the *DNA* sequence of the amplicon, and therefore enable us to infer *Spiroplasma* strain types and also the mean melting temperature.

3.5.3 Specific *Spiroplasma* detection

Secondly, as stated in table 1, positive samples were then amplified using other primers that target specific *Spiroplasma* clades (Citri-Poulsonii clade and Ixodetis clade). Samples that amplified using Citri-Poulsonii clade cocktail primers were subsequently tested using primers specific to Citri-Melliferum clade and Citri-Poulsonii clade (Figure 5). Regular PCR was used with the following the cycling conditions; initial enzyme activation at 95°C for 15mins, followed by 35cycles of denaturation at 95°C for 30secs, annealing (at a temperature specific to primers being utilized) for 30secs, elongation at 72°C for 30secs then hold temperature at 72°C for 10 min. PCR products were visualized on 1% agarose gels, along with a 100bp *DNA* ladder



A B

Figure 4: Optimization of the *rpoB* universal primers. (A) and (B) Shows melting curves and gel electrophoresis image of the different *Spiroplasma* strains and *phytoplasma*. The universal primers are specific to *Spiroplasma*, as demonstrated by the absence of amplification of *phytoplasma* a close relative of *Spiroplasma.*

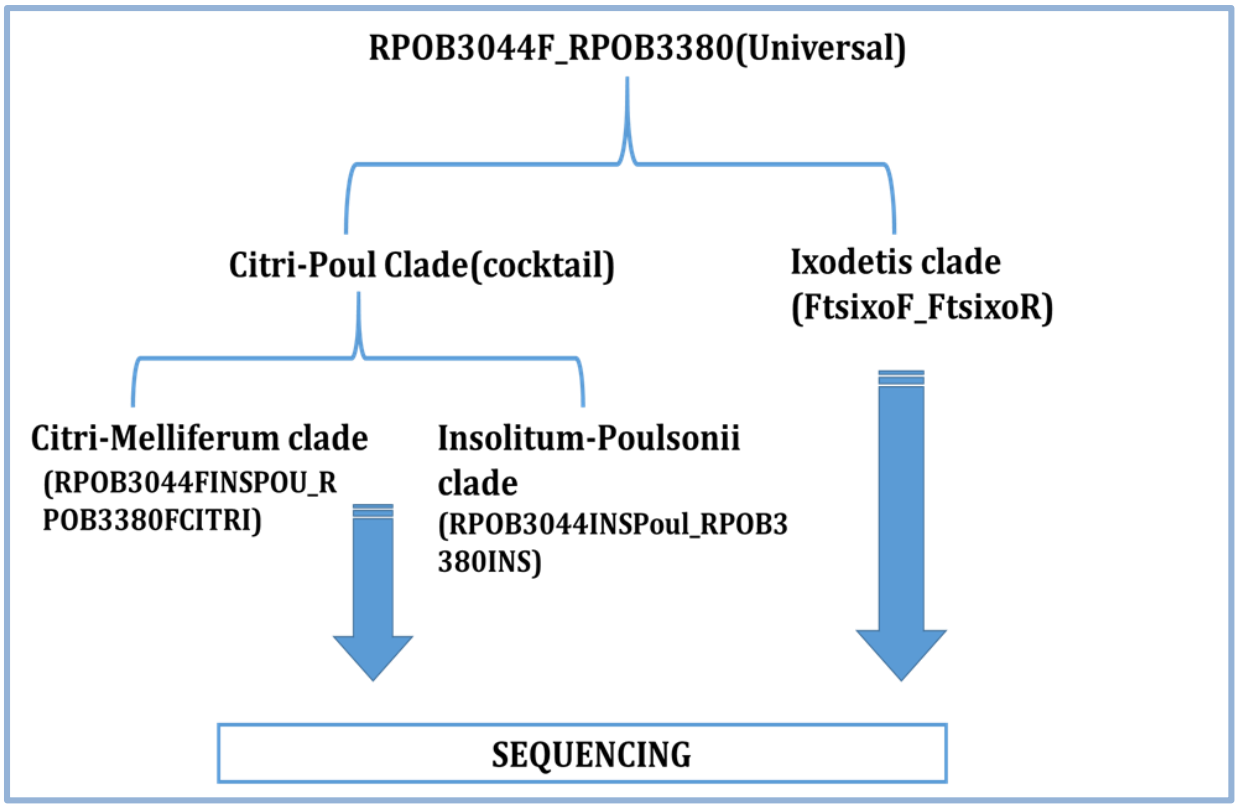


Figure 5: *Spiroplasma* detection pipeline: Screening samples using the universal primers indicates *Spiroplasma* infection. Specific primers narrow down to clade-level. Clades are indicated by cocktail primers for (Citri-Poulsonii clade) and Ftsixo_FtsixoR for ixodetis clades. For Citri-Poulsonii clade the specific primers named discriminate the different strains.

3.5.4 Sequencing and Phylogenetic analysis of *Spiroplasma* positive samples

Positive PCR products were cleaned prior to Sanger sequencing using ExoSap-IT purification protocol. 5µl of the PCR product was mixed with 0.5µl of Exonuclease 1 (Thermoscientific) and 1µl of FASTAP™ Alkaline Phosphatase (Thermoscientific) in a 0.2ml PCR tube. The mixture was centrifuged then incubated at 37°C for 15mins. The reaction mixture was then heated at 85°C for 15mins. The samples were submitted to Macrogen Inc. (Amsterdam) for Sanger sequencing

Sequences obtained were cleaned and aligned, manual corrections of bases caused by sequencing error was performed by comparing them to known *Spiroplasma* strains was done in Geneious v8.05. A phylogenetic tree of the sequences was constructed using UPGMA (Unweighted Pair Group Method with Arithmetic) method and evaluated by bootstrapping using 100 replicates and supported by 70% confidence value.

3.5.5 Mitochondrial *DNA* Analysis

To determine the diversity of mosquito mitochondrial *DNA*, the Nicotinamide adenine dinucleotide dehydrogenase 5 gene (*NADH-ND5*) was amplified using the primers described by Besansky,1997 and as listed in Table 1[99]. Single PCR reactions were performed on the Veriti Thermal Cycler (Applied Biosystems, Carlsbad, CA). PCR cycling conditions included initial denaturation at 95°C for 15mins, followed by 35 cycles of denaturation at 95°C for 30secs, annealing at 55°C for 30secs, elongation at 72°C for 30secs then hold temperature of 72°C for 10mins. PCR products were visualized on 1% agarose gels, along with a 1kb *DNA* ladder

PCR products were purified prior to Sanger sequencing using ExoSap-IT purification protocol. 5µl of the PCR product was mixed with 0.5µl of Exonuclease 1 (Thermoscientific) and 1µl of FASTAP™ Alkaline Phosphatase (Thermoscientific) in a 0.2ml PCR tube. The mixture was centrifuged then incubated at 37°C for 15mins. The reaction mixture was then heated at 85°C for 15mins. The samples were submitted to Macrogen Inc. (Amsterdam) for Sanger sequencing

Sequences obtained were cleaned and aligned using MUSCLE algorithm in Geneious v8.05. Haplotype generation, number of polymorphic sites, nucleotide diversity and haplotype diversity (Hd) was performed in DNAsp v2.0[102]. A haplotype tree was constructed using UPGMA basing on pairwise similarity, tree robustness was evaluated by bootstrapping (100 replicates) and 95% confidence. Statistical analysis of mitochondrial variability was deduced by a Tajimas D-test calculated in DNAsp v2.0[102]

3.5.6 Microsporidia *DNA* analysis

A PCR based microsporidia-screening strategy was used[100]. Specifically, primers targeting the small subunit ribosomal *RNA* (ssr*RNA*) region (approximately 1200bp) were used for detection and sequencing of microsporidia *DNA*. For an initial characterization of microsporidia in our *Anopheles arabiensis* population, universal primers SSR218F_SSR1492R were used (Table1). PCR cycling conditions included: initial denaturation at 95°C for 15mins, 40 cycles of denaturation at 95°C for 30secs, annealing at 50°C for 1min, elongation at 72°C for 45secs then hold temperature of 72°C for 10mins. PCR products were visualized on 1% agarose gels, along with a 1kb *DNA* ladder. After an initial characterization of microsporidia diversity, we established and tested a rapid High resolution melting PCR based screening procedure that enables the identification of all strains of observed microsporidia [103]. We used SSR218F_SSR378R primers in this PCR procedure. Following amplification, melting-curve analysis of the amplicons was performed by plotting a curve of changes in florescence against changes in temperature (dF/dT) which showed different melting curves for specific strains of microsporidia. To confirm the validity of this assay, representative peaks were sequenced.

3.6 Sequencing

Microsporidia Positive PCR products were cleaned prior to Sanger sequencing using ExoSap-IT purification protocol. 5µl of the PCR product was mixed with 0.5µl of Exonuclease 1 (Thermoscientific) and 1µl of FASTAP™ Alkaline Phosphatase (Thermoscientific) in a 0.2ml PCR tube. The mixture was centrifuged then incubated at 37°C for 15mins. The reaction mixture was then heated at 85°C for 15mins. The samples were submitted to Macrogen Inc. (Amsterdam) for Sanger sequencing.

The sequences were cleaned and aligned using the Multiple Sequence Comparison by Log-Expectation (MUSCLE) algorithm to reference sequences of accession numbers (*Y00266, JF826421, KF110990, JF826420, JF826419, JF826402, HM594267, AY090067, AY090065, AY090045, AJ252961, AF069063, AF027685, AF027684, AF027683, AF027682, JH370132.1, AY326269, AY305325, AY090043, AY013359, DQ641245, EU664450*) This was done in Geneious v8.05. Unweighted Pair-Group Method with Arithmetic mean (UPGMA) phylogenetic tree was constructed using 100 replicates and supported by 70% bootstrapping

CHAPTER 4

4.0 RESULTS

4.1 Anopheles Species identification

Amplification using the sub-species identification primers showed that 337 (87.5%) of the Mwea samples were *Anopheles arabiensis,* 29 (0.5%) *were Anopheles gambiae* while 19 (1.8%) of the samples did not amplify. In Mbita 355 were *Anopheles arabiensis,* 1 was *Anopheles gambiae* and 1 did not amplify. Only *Anopheles arabiensis* samples were used for subsequent experiments.

4.2 *Spiroplasma* Prevalence

Spiroplasma prevalence in Mwea was approximately 4% with 6 males and 7 females of the total 337 mosquitoes (total Mwea samples were composed of 250 females and 87 males mosquitoes) were infected. As demonstrated in figure 6 no mosquitoes from Mbita tested positive for *Spiroplasma*. Prevalence was significantly higher in Mwea than in Mbita using 95% confidence interval Chi-square values were (χ^2=13.41, df =1, p-value=0.0002. Sequencing results indicated that there were two strains of *Spiroplasma* in the samples. These strains included the *Spiroplasma insolitum* and *Spiroplasma melliferum* types. Notably, the *Spiroplasma melliferum* type seemed to infect only male mosquitoes with a prevalence of 7% (6 of 87 males) while the *Spiroplasma insolitum* type was predominant in females with a prevalence of 2.14%.

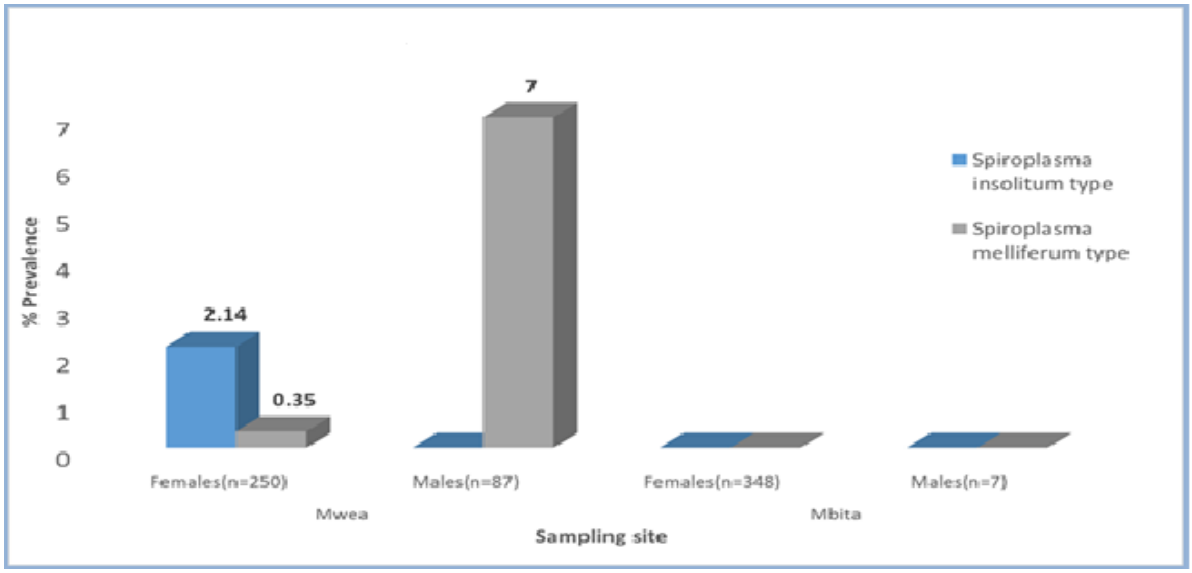


Figure 6: *Spiroplasma* prevalence bar graph: The bar graph shows the prevalence of *Spiroplasma insolitum*-type and *Spiroplasma melliferum* type relative to mosquito sexes in Mwea and Mbita. Female mosquitoes collected from Mwea had a prevalence of 2.14% of *Spiroplasma.insolitum*-type and 0.35% of *Spiroplasma melliferum* type while the males had Spiroplasma melliferum type only with frequency of 7% .There were no *Spiroplasma* in mosquitoes from Mbita .

A phylogenetic tree constructed using sequences from this study together with *Spiroplasma* sequences obtained from NCBI reveals the two strains of *Spiroplasma* circulating in mosquitoes collected in Mwea, Kenya. As illustrated in Figure 7, *Mwea_66* represents the *S. melliferum* type while *Mwea_61* represents the *S. insolitium* type *Spiroplasma*.

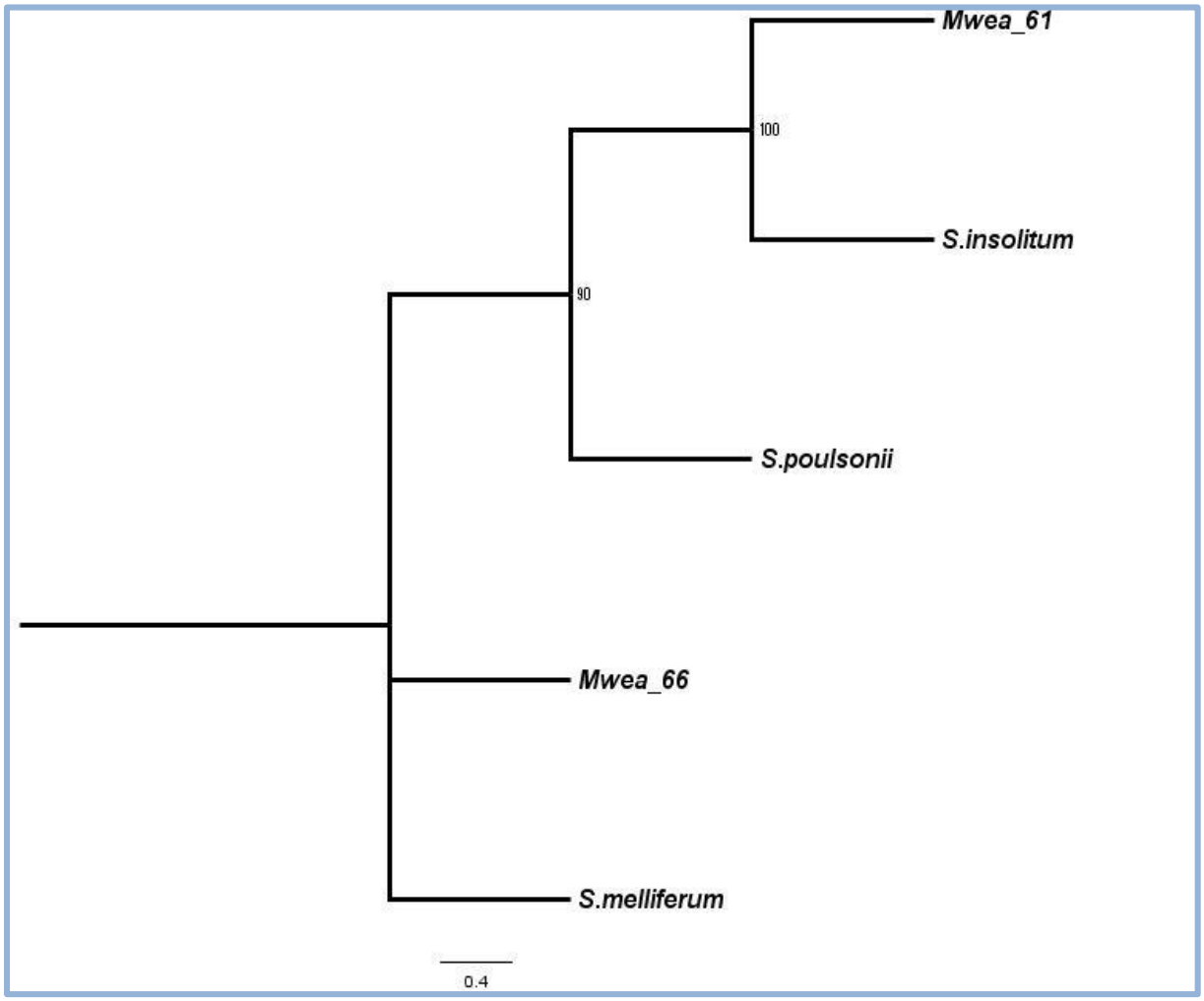


Figure 7: A *Spiroplasma* phylogenetic tree. The rooted UPGMA tree shows mosquito infection with *Spiroplasma insolitum*-type and *Spiroplasma melliferum*- type. Numbers at tree nodes represent bootstrap support values (100 replicates). The bar at the bottom of the figure shows molecular clock scale. In this case the line segment with 0.4 represents an amount genetic change of 0.4

4.3 MiSeq illumina sequencing analysis

Interestingly, Miseq Illumina data demonstrates that mosquito samples contain a variety of bacteria including *Enterobateriaceae, Protobacteria, Gammaproteobacteria, Actinobacteria* and *Cyanobacteria. Proteobacteria* was the most abundant family in both sites, *Actinobacteria* and *Cyanobacteria* were abundant in Mbita compared to Mwea while those that were found in small frequency in *Mwea* include *Bacterioides* and *Tenericutes.* Samples from Mwea showed *Spiroplasma* prevalence of 0.02% while in Mbita no individuals were infected with *Spiroplasma* (Figures 8 and 9). This results show the bacteria present in the mosquitoes and also suggests the species that can co-exist with *Spiroplasma*.

To confirm the specific *Spiroplasma* strain a local BLASTN analysis was done using specific *Spiroplasma* sequence queries of accession numbers (AJ579919.1, NR_025705.1 and AJ631998.1). As illustrated in table 2, cleaning and merging of the MiSeq reads resulted in 25,347 and 25,254 high-quality 16S ribosomal *RNA* sequences of mosquitoes from Mwea and Mbita, respectively. This were used to perform the BLAST. The local BLAST search on the Mwea sequences showed a prevalence of approximately 0.02% of the *Spiroplasma insolitum* strain (Appendix 1), this was calculated as the number of *Spiroplasma insolitum* hits found out of the total number of sequences obtained from MiSeq Illumina. In contrast none of the merged Mbita sequences was identified as *Spiroplasma*

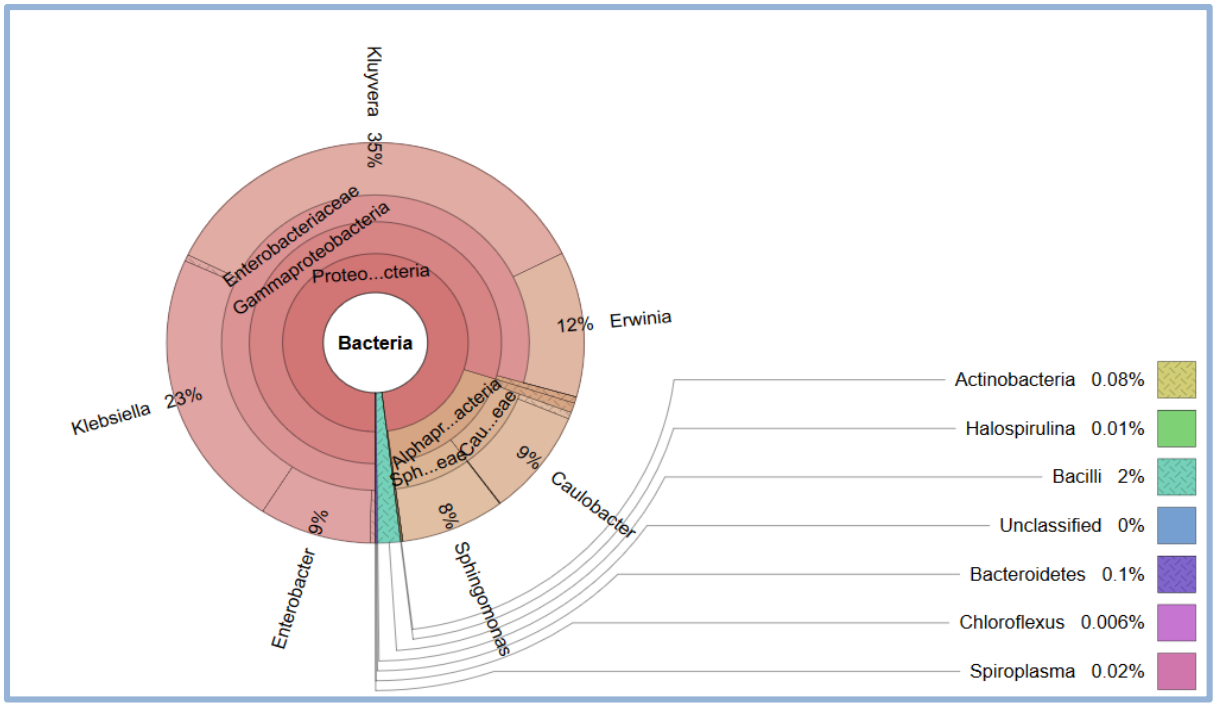


Figure 8: Mwea MiSeq illumina sequencing chart. A chart showing the presence of *Spiroplasma* and the diverse bacteria in the pooled mosquito samples. The percentage values at beside the bacterial names denote their prevalence

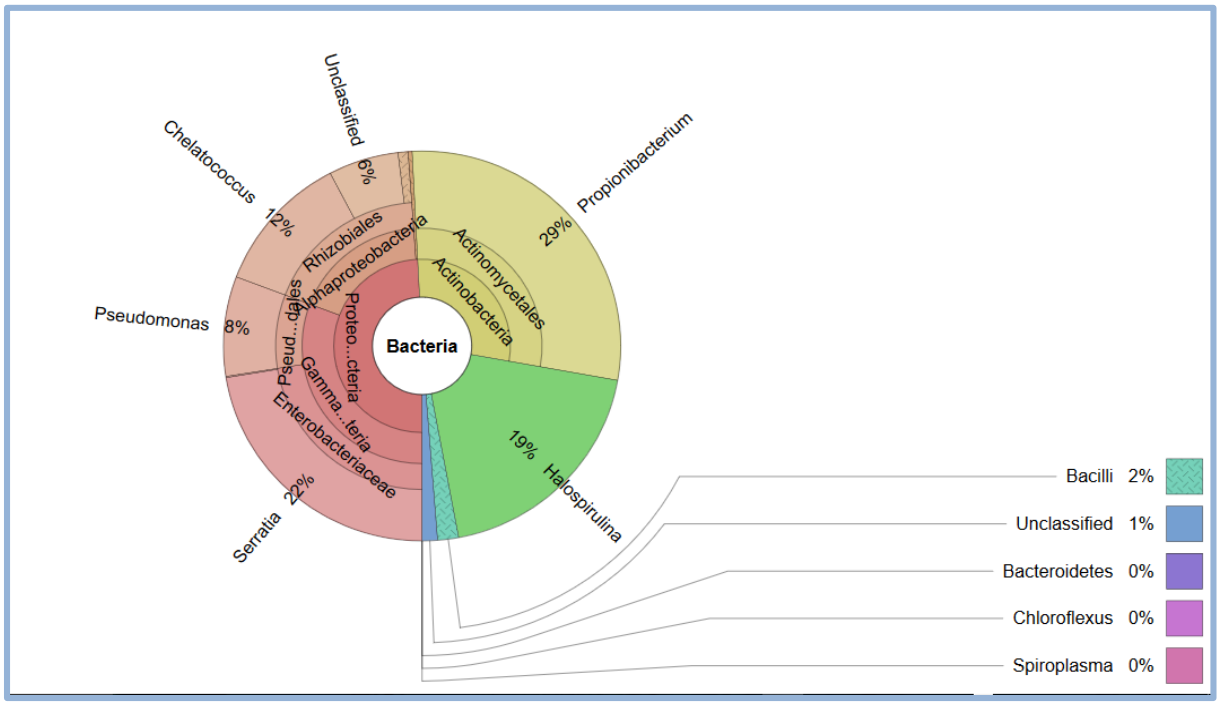


Figure 9: Mbita MiSeq illumina sequencing chart. A chart showing the presence of *Spiroplasma* and the diverse bacteria in the pooled mosquito samples. The percentage values at beside the bacterial names denote their prevalence.

Table 2: MiSeq Illumina sequencing Table: Table shows the prevalence of *Spiroplasma* in the sequences obtained from MiSeq illumina sequencing. It shows the total number of sequences obtained from Mwea and Mbita and Spiroplasma insolitum sequences found and its prevalence in Mwea and Mbita

Site	*Total number of Sequenes*	*S.insolitum Sequences*	*% Prevalence*
Mwea	**225,347**	***40***	***0.02%***
Mbita	**25,254**	***0***	***0***

4.4 Association of mitochondrial *DNA* and *Spiroplasma* infection

The mitochondrial *DNA*, NADH dehydrogenase 5 (ND5) gene responsible for oxidative phosphorylation was sequenced in a subset of samples that had been screened for *Spiroplasma* using primers [99]. This showed the distribution of *Spiroplasma* infection with respect to haplotypes. The ND5 gene was sequenced in a total of 22 samples (13 *Spiroplasma* positive samples and 9 non-infected samples). Only 22 samples were selected for this experiment since they were available and they were therefore used to infer the relationship between *Spiroplasma* and host mitochondrial *DNA.*

A total of 6 distinct haplotypes were observed in the total ND5 sequences. *Spiroplasma* infections were distributed across the haplotypes (see Figure. 9). Notably, three novel haplotypes, (Hap_HMW1, Hap_HMW2 and Hap_HMW3) were observed in this study. The others are identical to haplotypes already observed in previous studies [99]. The two strains of *Spiroplasma* in our samples also indicated an even distribution in the haplotypes.

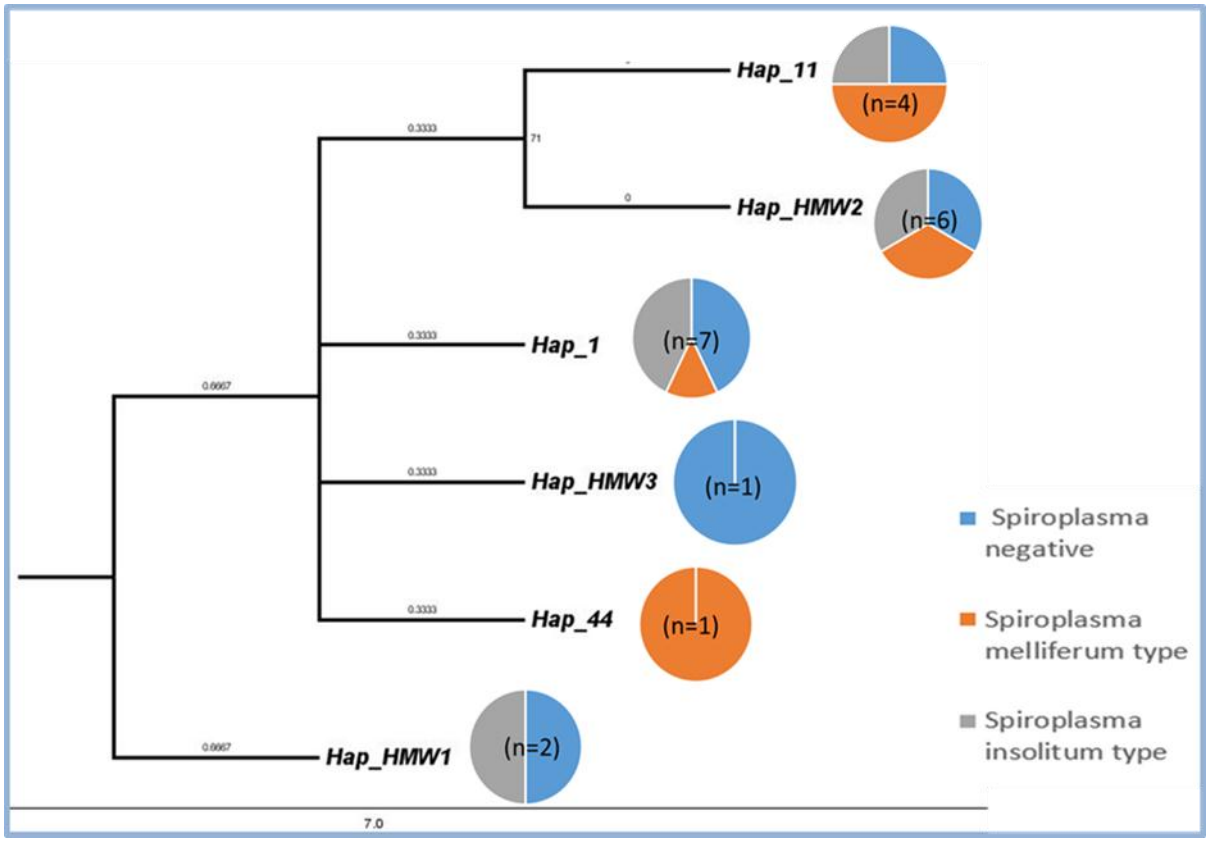


Figure 10: mitochondrial *DNA* phylogenetic tree: Neighbour joining tree of mosquito mitochondrial *DNA* (*ND5*) haplotypes. *n* represents the number of individuals harbouring the particular haplotype while the pie-chart represents *Spiroplasma* strain composition of particular haplotype. Numbers at tree nodes represent bootstrap support values (100 replicates). The bar at the bottom of the figure shows molecular clock scale. In this case the line segment with 7.0 represents an amount genetic change of 7.0

Sequences	**n**	**S**	**H**	**π**	**Tajima's D test**
Total(ND5 sequences)	**21**	**12**	**0.778**	**0.00520**	**0.46187 (Not significant, *P > 0.10*)**
***Spiroplasma* Positive(ND5 sequences)**	**13**	**10**	**0.775**	**0.00475**	**0.35471 Not significant, *P > 0.10*)**
***Spiroplasma* Negative(ND5 sequences)**	**9**	**9**	**0.758**	**0.00490**	**0.84671 (Not significant, *P > 0.10*)**

Table 3: Statistical analysis of mitochondrial *DNA* polymorphisms. n, is the number of sequences; S, number of polymorphic sites; h, haplotype diversity and as defined in the materials and methods are π, Tajima's D test.The table shows the statistical summary of the total ND5 sequences and ND5 sequences infected with Spiroplasma and non-infected.

Tajima's D neutrality test was performed on the total number of sequences *(n*=21), *Spiroplasma* positive samples (n=13) and *Spiroplasma* negative samples (*n*=9) and the Tajima's D estimates were all positive values and not statistically significant, as shown in table 3. This indicates that the *Spiroplasma* is evolving randomly with no external force affecting the evolution. In addition, these suggests that *Spiroplasma* has no effect of mosquito evolution. There was no difference in the nucleotide diversity (π) values for the infected and non-infected groups with 0.00475 (*P* > 0.10) and 0.00490 (*P* > 0.10), respectively.

4.5 Microsporidia infection prevalence

Amplification of the small sub-unit ribosomal *RNA* of microsporidia in the samples collected from the two sites indicated that the prevalence of microsporidia was ~35% and ~9% in Mwea and Mbita, respectively (Figure 11).

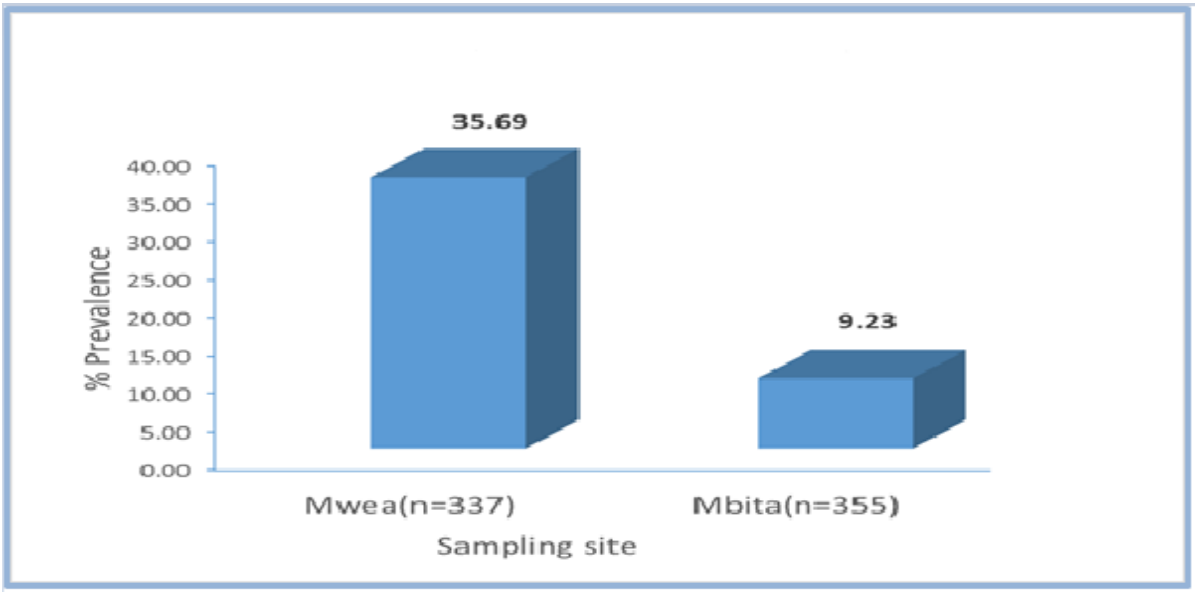


Figure 11: Microsporidia prevalence; Bar graph showing microsporidia prevalence in Mwea and Mbita. Mwea has a higher prevalence (35.69%) compared to Mbita (9.23%)

Sequencing and phylogenetic analysis showed that our samples had microsporidia strains related to genera *Crispospora, Hazardia, Parathelohania* and *Takaospora* (Figure 12). In addition, *Crispospora* (90%) was the most abundant species in Mwea, while in Mbita *Parathelohania* (73%) was the most abundant (Figure 12). This suggests adaptation of a specific microsporidia to specific geographical area.

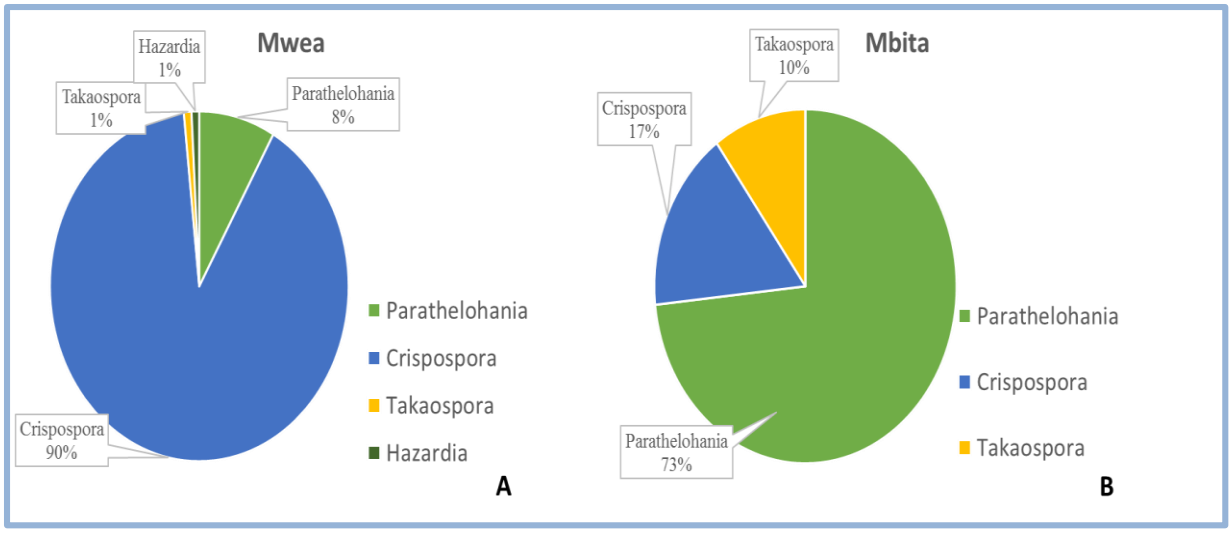


Figure 12: Microsporidia Species Composition Chart: Pie-chart showing the percentage composition of the microsporidia species in two sampling site (A) Mwea (B) Mbita. *Crispospora* is more dominant in mosquitoes collected in Mwea while *Parathelohania* was dominant in the mosquitoes collected in Mbita

A phylogenetic tree constructed using sequences obtained from this study were compared to other sequences obtained from National Centre for Biotechnology Information (NCBI)This tree showed the four main strains of microsporidia found in the mosquitoes collected in this study These species include *Crispospora, Hazardia, Takaokaspora* and *Parathelohania*

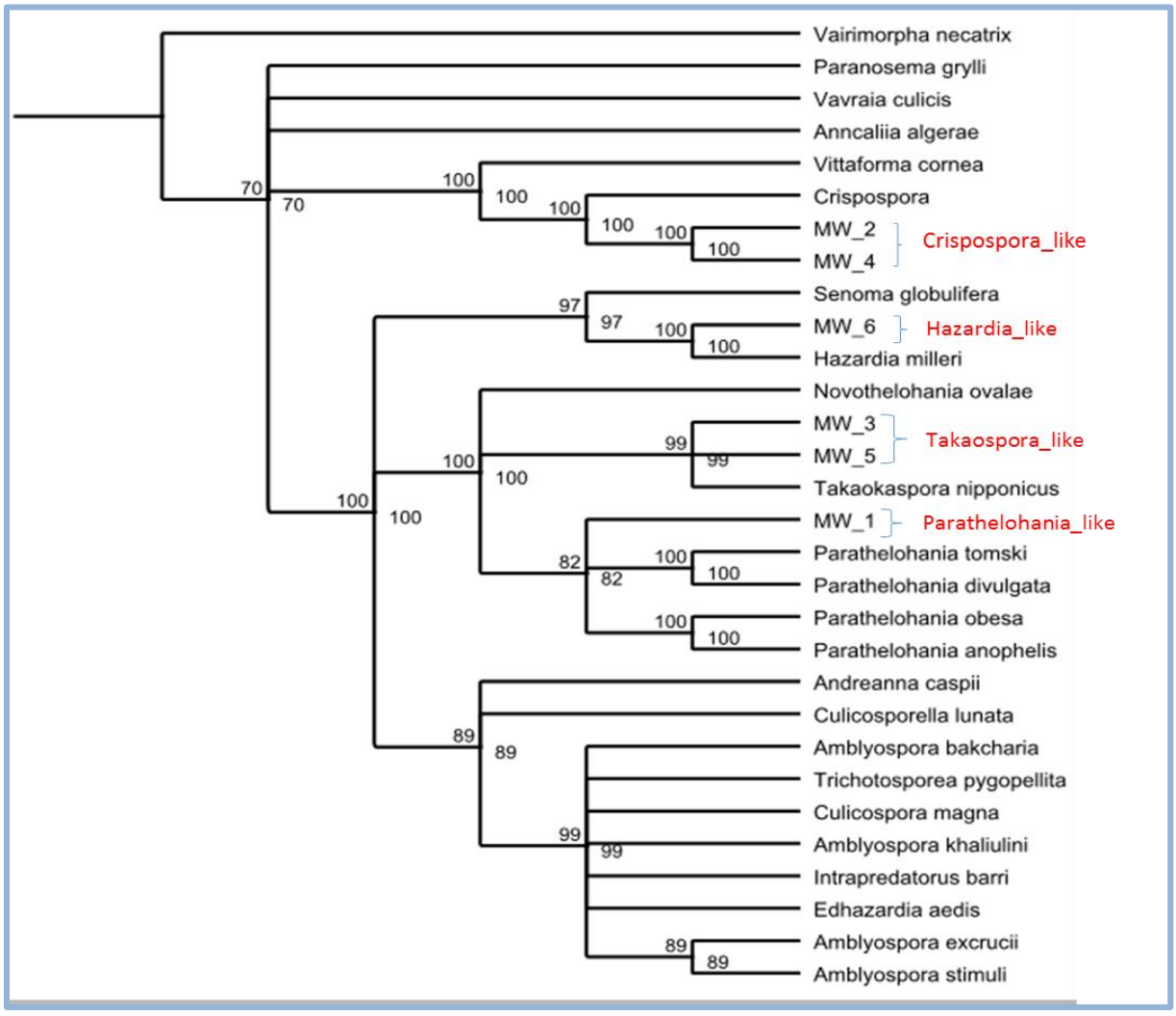


Figure 13: Microsporidia UPGMA phylogenetic tree analysis: Labelled in Red are representative microsporidia samples showing the strains circulating in the mosquito samples collected in this study

5.0 DISCUSSION

This study has developed a new pan-*Spiroplasma* screening method involving the use of universal primers targeting the different clades of *Spiroplasma* and a combination of specific primers to identify the *Spiroplasma* strain in the samples. This contributes to the knowledge about the biology and occurrence of *Spiroplasma* in mosquitoes and can also be applied to other organisms. This PCR-based pipeline is cheap and more convenient compared to other methods like high throughput screening and enzyme-linked immunosorbent assay (ELISA), especially when working with a large sample set. The use of high resolution melting PCR with the controls in our pipeline clearly shows the specific clade that the *Spiroplasma* belongs and representative samples can be sequenced and used for identification and profiling of the rest of the mosquito samples collected.

The PCR-based method developed was used for the identification and characterization of *Spiroplasma* in *Anopheles arabiensis* collected in two selected sites in Kenya (Mwea and Mbita) these sites were selected due to their geographical location and their high mosquito prevalence. Screening using this method recorded the presence of *Spiroplasma* infection in about 3% of mosquitoes collected from Mwea with no mosquitoes infected in Mbita. To validate the developed method Miseq illumina sequencing was performed on the PCR amplicons using universal bacterial 16S primers, and the pooling of samples from each of the two sites. Amplicon sequencing indicated the diverse range of bacteria harboured by the mosquito samples. More importantly, it indicated the presence of *Spiroplasma* in Mwea with no infections found in Mbita, thus the sequencing results were consistent with the results observed in our PCR-based method. In addition to *Spiroplasma,* the MiSeq Illumina sequencing data from the two sites showed that the mosquitoes were also hosts of other bacterial strains for example; *Enterobateriaceae, Protobacteria, Gammaproteobacteria, Actinobacteria* and *Cyanobacteria*. Studies have demonstrated that bacterial microbiota play an important role in the host. For instance members of *Enterobateriaceae* family contain haemolytic enzymes that help in digestion of blood.

The differences in *Spiroplasma* incidence between the two sites is notable. This could be attributed to several factors. First, there might be a component of agrochemicals and fertilizers that is preferred by *Spiroplasma* infected mosquitos that is used in Mwea rice fields unlike in Mbita[104]. Secondly, *Spiroplasma* infection incidence could be linked to rain conditions. This is perhaps suggested by the finding that most of the infected samples were collected during the

first week of April just before the onset of the long rains with the prevalence decreasing in the samples collected in May (during the long rains).

Previous studies have indicated that *Spiroplasma* strains isolated in mosquitoes were either in the *apis* clade or closely related to *ixodetis*[37,54]. Molecular phylogenetics in this study demonstrated that there are two strains of *Spiroplasma* harboured by mosquitoes in Mwea. One strain is closely related to *Spiroplasma insolitum* isolated from the *Bidens sp* flowering plant[105].The other strain is closely related to *Spiroplasma melliferum,* a pathogen of bees. Notably, these two strains were not evenly distributed across mosquito sexes. The insolitum-type strain was predominant in female mosquitoes (89% *n*=6), whereas the melliferum-type strain was found only in males (98% *n*=6). While the significance of this difference is not entirely clear, it suggests that the insolitum-type strain is more likely to be vertically transmitted endosymbiont (since these often are at high titres in ovaries) and is potentially playing a role in reproductive manipulation. Elsewhere, an insolitum-type strain has been identified in flower bugs and was shown to be a vertically transmitted endosymbiont[106]. The significance of this apparent male-specificity in the melliferum-type strain is also not entirely clear. Though it has been suggested that since the melliferum-type normally affects bees, the male mosquitoes tend to pick it up when sugar feeding on plants.

Phylogenetic data analyses of mosquito ND5 sequencing showed six haplotypes with *Spiroplasma* infected samples being evenly distributed among all the haplotypes. We did not observe a clear correlation between mitochondrial *DNA* haplotype and *Spiroplasma* infection. Suggesting two major possibilities, first that infection could be from a common ancestor that has been maintained in this species for a very long period of time (enabling diversification of mitochondrial *DNA* within the infected lineage). Another possibility is that there is significant horizontal transmission of *Spiroplasma* between the *Anopheles arabiensis* mosquitoes. However, to confirm these an experiment to determine the transmission of *Spiroplasma* in mosquitoes should be performed. In summary, there appears to be no apparent correlation between *Spiroplasma* infection and mitochondrial *DNA* haplotype, since the infected individuals are not restricted to one or more related haplotypes. While this does not rule out vertical transmission being the predominant mode of transmission, it does suggest that there is an appreciable level of horizontal transmission, which is not uncommon for facultative endosymbionts[57]. Another, less probable scenario is that the *Spiroplasma* infections are ancient (the species became infected prior to the diversification of these mitochondrial *DNA*

haplotypes) and strictly vertically transmitted, may have been lost in some mitochondrial lineages (e.g. Hap_44).

In addition to the prevalence study of *Spiroplasma*, we also compared its presence with that of microsporidian parasites, our results demonstrate that approximately 35% of the mosquitoes from Mwea were infected with microsporidia while in Mbita, there was a prevalence of less than 10%. This shows a striking difference between the two sites which suggests that either the microsporidia could be favoured by weather conditions or the nature of the larval habitat, Larval habitats are mainly in water and in Mwea that is provided by rice paddies while in Mbita that is provided for by the lake, Chemicals and fertilizers used in the rice fields have some components that are preferred by mosquitoes larvae, increase in larvae prevalence also increase microsporidia infection. This is because microsporidia are spread via its spores hence one major explanation is that the abundance of water during the rainy season favours dispersal of these spores leading to higher prevalence while absence of rain reduces the prevalence of microsporidia. Microsporidia species composition in the two sites also varied significantly with *Crispospora* being abundant in Mwea while *Parathelohania* shows dominance in Mbita. This suggests that specific microsporidia prefer certain environments. Mwea is a rice growing area and is characterized by frequent use of fertilizers and chemicals unlike Mbita. Therefore this suggests that *Crispospora* could be favoured by this chemicals and fertilizers. However more sampling should be done to confirm this finding

Interestingly, none of the samples infected with *Spiroplasma* were also infected with Microsporidia. This suggests that the *Spiroplasma* could have protective effects in the mosquitoes against microsporidia, which would be in line with the finding that *Spiroplasma* confers protection to its host against parasites and pathogens[36] . For instance *Spiroplasma* protects aphids from a fungal pathogen[36] . However, more sampling is needed to confirm this correlation. In addition, this study simply acts as a descriptive study of what is there in mosquitoes and was not set out to answer the protective effects of *Spiroplasma* against microsporidia.

5.0 CONCLUSION

This study develops a PCR-based strategy for screening *Spiroplasma* in mosquitoes, this method is not only cost-effective but also less time consuming compared to other methods such as. High throughput screening and ELISA. This method can be used in screening for *Spiroplasma* in other mosquitoes and can also be applied to other insects. Secondly, it reports for the first time the presence of *Spiroplasma* in *Anopheles arabiensis* mosquitoes collected in Mwea, Kenya. This finding is an important discovery since more *Spiroplasma* studies in mosquitoes can be done. Additional study reveals that the *Spiroplasma* found in mosquitoes has no coinfection with microsporidia parasite. This suggests that *Spiroplasma* protects the mosquito against parasites and therefore sets base for more studies to be performed to determine the direct effect of *Spiroplasma* to microsporidia and other parasites and pathogens affecting mosquitoes (such as fungi and *Plasmodium*).

7.0 RECOMMENDATIONS

1. This study reports the presence of *Spiroplasma* in *Anopheles arabiensis*. To better understand how to apply this *Spiroplasma* in the control of malaria, experiments to determine their mode of transmission is required.
2. Results from this study shows no coinfection of *Spiroplasma* and microsporidia on the same mosquito, this suggests a possibility that *Spiroplasma* confers resistance against microsporidia, however, this needs to be examined further.
3. To further understand the protective characteristic of *Spiroplasma*, addition, studies on the relationship between *Spiroplasma* and other mosquito parasites, including *Plasmodium* and fungi should be done.

8.0 REFERENCES

1. World Health Organization World malaria report 2013 shows major progress in fight against malaria, calls for sustained financing. (2013).

2. World Health Organization. Implementation of the global malaria control strategy. Report of a WHO Study Group on the Implementation of the Global Plan of Action for Malaria Control 1993-2000. *World Health Organ. Tech. Rep. Ser.* **839,** 1–57 (1993).

3. World Health Organization | Zika virus. *WHO* (2016).

4. Nauen, R. Insecticide resistance in disease vectors of public health importance. **633,** 628–633 (2007).

5. Bourtzis, K. *et al.* Harnessing mosquito-Wolbachia symbiosis for vector and disease control. *Acta Trop.* **132 Suppl,** S150–63 (2015).

6. World Health Organization. World malaria report 2015. **Volume WHO,** 238 (2015).

7. Lindsay, S. W., Parson, L. & Thomas, C. J. Mapping the ranges and relative abundance of the two principal African malaria vectors, Anopheles gambiae sensu stricto and An. arabiensis, using climate data. *Proc. R. Soc. B Biol. Sci.* **265,** 847–854 (1998).

8. Gillies, M. T. & de Meillon, B. The Anophelinae of Africa south of the Sahara (Ethiopian Zoogeographical Region). (1968).

9. Ramírez, A. *et al.* Evolution of dengue virus type 3 genotype III in Venezuela: diversification, rates and population dynamics. *Virol. J.* **7,** 329 (2010).

10. Hill, A. V. S. Vaccines against malaria. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **366,** 2806–14 (2011).

11. Gubler, D. Vector-Borne Disease Emergence and Resurgence. (2008). at

12. Poopathi, S. & Tyagi, B. K. The Challenge of Mosquito Control Strategies : from Primordial to Molecular Approaches. **1,** 51–65 (2006).

13. Mcgraw, E. A. & Neill, S. L. O. Beyond insecticides : new thinking on an ancient problem. **11,** (2013).

14. Dale, C. & Moran, N. a. Molecular interactions between bacterial symbionts and their hosts. *Cell* **126,** 453–65 (2006).

15. Eleftherianos, I., Atri, J., Accetta, J. & Castillo, J. C. Endosymbiotic bacteria in insects: guardians of the immune system? *Front. Physiol.* **4,** 46 (2013).

16. Su, Q., Zhou, X. & Zhang, Y. Symbiont-mediated functions in insect hosts. 1–7 (2013).

17. Ratzka, C., Gross, R. & Feldhaar, H. Endosymbiont Tolerance and Control within Insect Hosts. *Insects* **3,** 553–572 (2012).

18. Ricci, I. *et al.* Symbiotic control of mosquito borne disease. *Pathog. Glob. Health* **106,** 380–5 (2012).

19. Werren, J. H. Biology of Wolbachia. *Annu. Rev. Entomol.* **42,** 587–609 (1997).

20. Hughes, H. & Britton, N. F. Opus : University of Bath Online Publication Store Modelling the use of Wolbachia to control dengue fever transmission. **75,** 796–818 (2013).

21. Kambris, Z., Cook, P. E., Phuc, H. K. & Sinkins, S. P. Immune activation by life-shortening Wolbachia and reduced filarial competence in mosquitoes. *Science* **326,** 134–6 (2009).

22. Sun, X.-J., Xiao, J.-H., Cook, J. M., Feng, G. & Huang, D.-W. Comparisons of host mitochondrial, nuclear and endosymbiont bacterial genes reveal cryptic fig wasp species and the effects of Wolbachia on host mtDNA evolution and diversity. *BMC Evol. Biol.* **11,** 86 (2011).

23. Dyer, K. a. & Jaenike, J. Evolutionarily stable infection by a male-killing endosymbiont in Drosophila innubila: molecular evidence from the host and parasite genomes. *Genetics* **168,** 1443–55 (2004).

24. Baudry, E., Bartos, J., Emerson, K., Whitworth, T. & Werren, J. H. Wolbachia and genetic variability in the birdnest blowfly Protocalliphora sialia. *Mol. Ecol.* **12,** 1843–1854 (2003).

25. Johnstone, R. A. & Hurst, G. D. D. Maternally inherited male-killing microorganisms may confound interpretation of mitochondrial DNA variability. *Biol. J. Linn. Soc.* **58,** 453–470 (1996).

26. Poulson, D. F. & Sakaguchi, B. Nature ofsex-ratio agent in Drosophila. *Science* **133,** 1489–90 (1961).

27. Xie, J., Vilchez, I. & Mateos, M. Spiroplasma bacteria enhance survival of Drosophila hydei attacked by the parasitic wasp Leptopilina heterotoma. *PLoS One* **5,** e12149 (2010).

28. Keeling, P. J. & Fast, N. M. Microsporidia: Biology and Evolution of Highly Reduced Intracellular Parasites. *Annu. Rev. Microbiol.* **56,** 93–116 (2002).

29. Roberta, D. W. & Strand, M. A. Pathogens of medically important arthropods. *Bull. World Health Organ.* **55,** (1977).

30. Milbrath, M. O. *et al.* Comparative virulence and competition between Nosema apis and Nosema ceranae in honey bees (Apis mellifera). *J. Invertebr. Pathol.* **125,** 9–15 (2015).

31. Parr, N. Studies of Microsporidia in mosquitoes. **121,** 70–77 (1993).

32. Gubler, D. J. Resurgent Vector-Borne Diseases as a Global Health Problem. **4,** 442–450 (1998).

33. Bian, G., Xu, Y., Lu, P., Xie, Y. & Xi, Z. The endosymbiotic bacterium Wolbachia induces resistance to dengue virus in Aedes aegypti. *PLoS Pathog.* **6,** e1000833 (2010).

34. Hedges, L. M., Brownlie, J. C., O'Neill, S. L. & Johnson, K. N. Wolbachia and virus protection in insects. *Science* **322,** 702 (2008).

35. Teixeira, L., Ferreira, Á. & Ashburner, M. The Bacterial Symbiont Wolbachia Induces Resistance to RNA Viral Infections in Drosophila melanogaster. *PLoS Biol.* **6,** e1000002 (2008).

36. Lukasik, P., Guo, H., Van Asch, M., Ferrari, J. & Godfray, H. C. J. Protection against a fungal pathogen conferred by the aphid facultative endosymbionts Rickettsia and Spiroplasma is expressed in multiple host genotypes and species and is not influenced by co-infection with another symbiont. *J. Evol. Biol.* **26,** 2654–2661 (2013).

37. Lindh, J. M. *et al.* 16S rRNA Gene-Based Identi cation of Midgut Bacteria from Field-Caught. *Society* **71,** 7217–7223 (2005).

38. World Health Organization | Vector-borne disease. *WHO* (2010).

39. Keating, G. M. RTS,S/AS01 malaria vaccine (Mosquirix^{TM}): a guide to its use. *Drugs Ther. Perspect.* (2016). doi:10.1007/s40267-016-0290-9

40. Mateos, M. *et al.* Heritable endosymbionts of Drosophila. *Genetics* **174,** 363–376 (2006).

41. Hilgenboecker, K., Hammerstein, P., Schlattmann, P., Telschow, A. & Werren, J. H. How many species are infected with Wolbachia?--A statistical analysis of current data. *FEMS Microbiol. Lett.* **281,** 215–20 (2008).

42. Duron, O. *et al.* The diversity of reproductive parasites among arthropods: Wolbachia do not walk alone. *BMC Biol.* **6,** 27 (2008).

43. Werren, J. H., Skinner, S. W. & Huger, A. M. Male-killing bacteria in a parasitic wasp. *Science* **231,** 990–2 (1986).

44. Werren, J. H. *et al.* Rickettsial relative associated with male killing in the ladybird beetle (Adalia bipunctata). *J. Bacteriol.* **176,** 388–94 (1994).

45. Sanada-Morimura, S., Matsumura, M. & Noda, H. Male killing caused by a Spiroplasma symbiont in the small brown planthopper, Laodelphax striatellus. *J. Hered.* **104,** 821–9 (2013).

46. Stouthamer, R., Breeuwer, J. a & Hurst, G. D. Wolbachia pipientis: microbial manipulator of arthropod reproduction. *Annu. Rev. Microbiol.* **53,** 71–102 (1999).

47. Oliver, K. M., Campos, J., Moran, N. A. & Hunter, M. S. Population dynamics of defensive symbionts in aphids. *Proc. Biol. Sci.* **275,** 293–9 (2008).

48. Oliver, K. M., Smith, A. H. & Russell, J. a. Defensive symbiosis in the real world - advancing ecological studies of heritable, protective bacteria in aphids and beyond. *Funct. Ecol.* **28,** 341–355 (2014).

49. John Jaenike, Julie Stahlhut, Lisa Boelio, R. U. Association between Wolbachia and Spiroplasma within Drosophila neotestacea : an emerging symbiotic. *Mol. Ecol.* 414–425 (2010). doi:10.1111/j.1365-294X.2009.04448.x

50. Gasparich, G. E. Spiroplasmas and phytoplasmas: microbes associated with plant hosts. *Biologicals* **38,** 193–203 (2010).

51. Regassa, L. B. & Gasparich, G. E. Spiroplasmas: evolutionary relationships and biodiversity. *Front. Biosci.* **11,** 2983–3002 (2006).

52. Haselkorn, T. S. The Spiroplasma heritable bacterial endosymbiont of Drosophila. *Fly (Austin).* **4,** 80–7 (2010).

53. Saglio, P. *et al.* Spiroplasma citri gen. and sp. n.: A Mycoplasma-Like Organism Associated with 'Stubborn' Disease of Citrus. *Int. J. Syst. Bacteriol.* **23,** 191–204 (1973).

54. Lo, W.-S. S., Ku, C., Chen, L.-L. L., Chang, T.-H. H. & Kuo, C.-H. H. Comparison of metabolic capacities and inference of gene content evolution in mosquito-associated Spiroplasma diminutum and S. taiwanense. *Genome Biol. Evol.* **5,** 1512–23 (2013).

55. Majerus, T. The butterfly Danaus chrysippus is infected by a male- killing Spiroplasma bacterium. (2000).

56. Tinsley, M. C. & Majerus, M. E. N. A new male-killing parasitism: Spiroplasma bacteria infect the ladybird beetle Anisosticta novemdecimpunctata (Coleoptera: Coccinellidae). *Parasitology* **132,** 757–65 (2006).

57. Watts, T., Haselkorn, T. S., Moran, N. a & Markow, T. a. Variable incidence of Spiroplasma infections in natural populations of Drosophila species. *PLoS One* **4,** e5703 (2009).

58. Pool, J. E., Wong, a & Aquadro, C. F. Finding of male-killing Spiroplasma infecting Drosophila melanogaster in Africa implies transatlantic migration of this endosymbiont. *Heredity (Edinb).* **97,** 27–32 (2006).

59. Hackett, K. J., Lynn, D. E., Williamson, D. L., Ginsberg, A. S. & Whitcomb, R. F. Cultivation of the Drosophila sex ratio spiroplasma. *Science (80-.).* **232,** 1253–1255 (1986).

60. Herren, J. K., Paredes, J. C., Schüpfer, F. & Lemaitre, B. Vertical Transmission of a Drosophila Endosymbiont Via Cooption of the Yolk Transport and Internalization Machinery. **4,** 1–8 (2013).

61. Montenegro, H., Solferini, V. N., Klaczko, L. B. & Hurst, G. D. D. Male-killing Spiroplasma naturally infecting Drosophila melanogaster. *Insect Mol. Biol.* **14,** 281–7 (2005).

62. Hurst, G. D. D. & Majerus, M. E. N. Why do maternally inherited microorganisms kill males? *Heredity (Edinb).* **71,** 81–95 (1993).

63. Cheng, B., Kuppanda, N., Aldrich, J. C., Akbari, O. S. & Ferree, P. M. Male-Killing Spiroplasma Alters Behavior of the Dosage Compensation Complex during Drosophila melanogaster Embryogenesis. *Curr. Biol.* (2016). doi:10.1016/j.cub.2016.03.050

64. Kageyama, D. *et al.* Prevalence of a non-male-killing spiroplasma in natural populations of Drosophila hydei. *Appl. Environ. Microbiol.* **72,** 6667–73 (2006).

65. Hurst, G. D. D. & Hutchence, K. J. Host defence: getting by with a little help from our friends. *Curr. Biol.* **20,** R806–8 (2010).

66. Oliver, K. M., Degnan, P. H., Hunter, M. S. & Moran, N. A. Bacteriophages Encode Factors Required for Protection in a Symbiotic Mutualism. *Science (80-.).* **325,** (2009).

67. Xie, J., Tiner, B., Vilchez, I. & Mateos, M. Effect of the Drosophila endosymbiont Spiroplasma on parasitoid wasp development and on the reproductive fitness of wasp-attacked fly survivors. *Evol. Ecol.* **53,** 1065–1079 (2011).

68. Humphery-Smith, I., Grulet, O. & Chastel, C. Pathogenicity of Spiroplasma taiwanense for larval Aedes aegypti mosquitoes. *Med. Vet. Entomol.* **5,** 229–32 (1991).

69. Haine, E. R. Symbiont-mediated protection. *Proc. Biol. Sci.* **275,** 353–61 (2008).

70. Kambhampati, S. & Verleye, D. M. Frequencies of Mitochondrial DNA Haplotypes in Laboratory Cage Populations of the Mosquito, nAedes albopictus. *Genetics* (1992).

71. Hung, S. H. Y., Chen, T. A., Whitcomb, R. F., Tully, J. G. & Chen, A. N. D. Y. X. Spiroplasma culicicola sp . nov . from the Salt Marsh Mosquito Aedes sollicitanst. 365–370 (1987).

72. Chastel, C., Tully, J. G., Bove, J. M. & Gilot, R. F. W. B. Spiroplasma sabaudiense sp. nov. from Mosquitoes Collected in France. 260–265 (1987).

73. Rosen, L., Tully, J. G., Bove, J. M., Chastel, C. & Williamson, D. L. Mosquitoes Collected in Taiwan. **12,** 103–107 (1988).

74. Whitcomb, R. F. The Biology of Spiroplasmas. *Annu. Rev. Entomol.* **26,** 397–425 (1981).

75. Abalain-Colloc, M. L. *et al.* Spiroplasma sabaudiense sp. nov. from Mosquitoes Collected in France. *Int. J. Syst. Bacteriol.* **37,** 260–265 (1987).

76. Lo, W. S., Ku, C., Chen, L. L., Chang, T. H. & Kuo, C. H. Comparison of metabolic capacities and inference of gene content evolution in mosquito-associated Spiroplasma diminutum and S. taiwanense. *Genome Biol. Evol.* **5,** 1512–1523 (2013).

77. Phillips, R. N. & Humphery-Smith, I. The Histopathology of Experimentally Induced Infections of Spiroplasma taiwanense (Class: Mollicutes) in Anopheles stephensi Mosquitoes. *J. Invertebr. Pathol.* **66,** 185–195 (1995).

78. Becnel, J. J. & Takvorian, P. M. Checklist of Available Generic Names for Microsporidia with Type Species and Type Hosts. 671–686 (2014).

79. Andreadis, T. G. Microsporidian parasites of mosquitoes. *J. Am. Mosq. Control Assoc.* **23,** 3–29 (2007).

80. Dunn, A. M. & Smith, J. E. Microsporidian life cycles and diversity: The relationship between virulence and transmission. *Microbes Infect.* **3,** 381–388 (2001).

81. Dunn, A. M. & Smith, J. E. Microsporidian life cycles and diversity: the relationship between virulence and transmission. *Microbes Infect.* **3,** 381–8 (2001).

82. Terry, R. S., Smith, J. E. & Dunn, A. M. Impact of a Novel, Feminising Microsporidium on its Crustacean Host. *J. Eukaryot. Microbiol.* **45,** 497–501 (1998).

83. Vossbrinck, C. R., Andreadis, T. G., Vavra, J. & Becnel, J. J. Molecular Phylogeny and Evolution of Mosquito Parasitic Microsporidia (Microsporidia: Amblyosporidae). *J. Eukaryot. Microbiol.* **51,** 88–95 (2004).

84. Andreadis, T. G. & Hall, D. W. Development, Ultrastructure, and Mode of Transmission of Amblyospora sp. (Microspora) in the Mosquito*. *J. Protozool.* **26,** 444–452 (1979).

85. Becnel, J. J., White, S. E. & Shapiro, A. M. Review of microsporidia-mosquito relationships: from the simple to the complex. *Folia Parasitol. (Praha).* **52,** 41–50 (2005).

86. Andreadis, T. G., Takaoka, H., Otsuka, Y. & Vossbrinck, C. R. Morphological and molecular characterization of a microsporidian parasite, Takaokaspora nipponicus n. gen., n. sp. from the invasive rock pool mosquito, Ochlerotatus japonicus japonicus. *J. Invertebr. Pathol.* **114,** 161–172 (2013).

87. QGIS Development Team. QGIS Geographic Information System. Open Source Geospatial Foundation Project. http://qgis.osgeo.org. *Qgisorg* (2014).

88. Muturi, E. J. *et al.* Mosquito species diversity and abundance in relation to land use in a riceland agroecosystem in Mwea, Kenya. *J. Vector Ecol.* **31,** 129–37 (2006).

89. Ijumba, J. N. & Lindsay, S. W. Impact of irrigation on malaria in Africa: paddies paradox. *Med. Vet. Entomol.* **15,** 1–11 (2001).

90. Muturi, E. J. *et al.* Diversity of riceland mosquitoes and factors affecting their occurrence and distribution in Mwea, Kenya. *J. Am. Mosq. Control Assoc.* **24,** 349–58 (2008).

91. Olanga, E. A., Okombo, L., Irungu, L. W. & Mukabana, W. R. Parasites and vectors of malaria on Rusinga Island, Western Kenya. *Parasit. Vectors* **8,** 250 (2015).

92. Minakawa, N., Dida, G. O., Sonye, G. O., Futami, K. & Njenga, S. M. Malaria vectors in Lake Victoria and adjacent habitats in Western Kenya. *PLoS One* **7,** (2012).

93. M.T.Gillies;Maureen Coetze. A Supplement to the Anophelinae of Africa South of the Sahara. (1987).

94. Wood, E. Molecular Cloning. A Laboratory Manual. *Biochem. Educ.* **11,** 82 (1983).

95. Broderick, N. A., Raffa, K. F., Goodman, R. M. & Handelsman, J. Census of the bacterial community of the gypsy moth larval midgut by using culturing and culture-independent methods. *Appl. Environ. Microbiol.* **70,** 293–300 (2004).

96. Vos, M., Quince, C., Pijl, A. S., de Hollander, M. & Kowalchuk, G. A. A Comparison of rpoB and 16S rRNA as Markers in Pyrosequencing Studies of Bacterial Diversity. *PLoS One* **7,** e30600 (2012).

97. Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Mentjies, P., & Drummond, A. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data.Bioinformatics, 28(12), 1647-1649. (2012).

98. Lane, D. In Nucleic Acid Techniques in bacterial systematics. 115 – 175 (1991).

99. Besansky, N. J. *et al.* Patterns of mitochondrial variation within and between African malaria vectors, Anopheles gambiae and An. arabiensis, suggest extensive gene flow. *Genetics* **147,** 1817–1828 (1997).

100. Ghosh, K. & Weiss, L. M. Molecular diagnostic tests for microsporidia. *Interdiscip. Perspect. Infect. Dis.* **2009,** 926521 (2009).

101. Zianni, M. R., Nikbakhtzadeh, M. R., Jackson, B. T., Panescu, J. & Foster, W. a. Rapid discrimination between Anopheles gambiae s.s. and Anopheles arabiensis by High-Resolution Melt (HRM) analysis. *J. Biomol. Tech.* **24,** 1–7 (2013).

102. Rozas, J. & Rozas, R. DnaSP version 2.0: a novel software package for extensive molecular population genetics analysis. *Comput. Appl. Biosci.* **13,** 307–11 (1997).

103. Omondi, D. *et al.* Unraveling Host-Vector-Arbovirus Interactions by Two-Gene High Resolution Melting Mosquito Bloodmeal Analysis in a Kenyan Wildlife-Livestock Interface. *PLoS One* **10,** e0134375 (2015).

104. Kibuthu, T. W. *et al.* Agricultural chemicals: life changer for mosquito vectors in agricultural landscapes? *Parasit. Vectors* **9,** 500 (2016).

105. Hackett, K. J. *et al.* Spiroplasma insolitum sp. nov., a New Species of Group I Spiroplasma with an Unusual DNA Base Composition. *Int. J. Syst. Bacteriol.* **43,** 272–277 (1993).

106. Watanabe, M., Yukuhiro, F. & Maeda, T. Novel Strain of Spiroplasma Found in Flower Bugs of the Genus Orius (Hemiptera : Anthocoridae): Transovarial Transmission , Coexistence with Wolbachia and Varied Population Density AS. 219–228 (2014). doi:10.1007/s00248-013-0335-8

107. Haselkorn, T. S. T. S. T. S. The spiroplasma heritable bacterial endosymbiont of Drosophila. *Fly (Austin).* **4,** 80–87 (2010).

9.0 Appendix 1: *Spiroplasma insolitum* BLAST on MiSeq data

BLASTN 2.2.10 [Oct-19-2004]
Reference: Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schaffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", Nucleic Acids Res. 25:3389-3402.
Query= (1515 letters)
Database: C:\BioEdit\database\Galaxy8-[FASTQ_to_FASTA_on_data_5].fasta

225,347 sequences; 76,912,062 total letters

Sequences producing significant alignments:	Score (bits)	E Value
M00532:128:000000000-AK1TL:1:2117:9891:21261 1:N:0:116	662	0
M00532:128:000000000-AK1TL:1:1115:11308:2211 1:N:0:116	662	0
M00532:128:000000000-AK1TL:1:1108:21400:14891 1:N:0:116	662	0
M00532:128:000000000-AK1TL:1:2119:8579:18916 1:N:0:116	654	0
M00532:128:000000000-AK1TL:1:2117:20488:11148 1:N:0:116	654	0
M00532:128:000000000-AK1TL:1:2114:12118:19724 1:N:0:116	654	0
M00532:128:000000000-AK1TL:1:2114:12134:19714 1:N:0:116	654	0
M00532:128:000000000-AK1TL:1:2114:4926:11177 1:N:0:116	654	0
M00532:128:000000000-AK1TL:1:2110:3716:12066 1:N:0:116	654	0
M00532:128:000000000-AK1TL:1:2110:17727:7670 1:N:0:116	654	0
M00532:128:000000000-AK1TL:1:2109:13820:17133 1:N:0:116	654	0
M00532:128:000000000-AK1TL:1:2109:13816:17115 1:N:0:116	654	0
M00532:128:000000000-AK1TL:1:2103:13944:4883 1:N:0:116	654	0
M00532:128:000000000-AK1TL:1:1118:13388:21819 1:N:0:116	654	0
M00532:128:000000000-AK1TL:1:1105:10918:9142 1:N:0:116	654	0
M00532:128:000000000-AK1TL:1:1102:17924:12066 1:N:0:116	654	0
M00532:128:000000000-AK1TL:1:2112:15776:9545 1:N:0:116	650	0
M00532:128:000000000-AK1TL:1:2117:14347:5574 1:N:0:116	646	0
M00532:128:000000000-AK1TL:1:2109:9326:10275 1:N:0:116	646	0
M00532:128:000000000-AK1TL:1:2109:9309:10267 1:N:0:116	646	0
M00532:128:000000000-AK1TL:1:2102:9806:2351 1:N:0:116	646	0
M00532:128:000000000-AK1TL:1:1109:8317:8965 1:N:0:116	646	0
M00532:128:000000000-AK1TL:1:1103:11599:6871 1:N:0:116	646	0
M00532:128:000000000-AK1TL:1:2117:25948:8905 1:N:0:116	642	0
M00532:128:000000000-AK1TL:1:2117:6608:3623 1:N:0:116	638	0
M00532:128:000000000-AK1TL:1:2107:3387:14191 1:N:0:116	638	0
M00532:128:000000000-AK1TL:1:2101:19298:4198 1:N:0:116	638	0
M00532:128:000000000-AK1TL:1:1107:21634:20470 1:N:0:116	638	0
M00532:128:000000000-AK1TL:1:1106:27231:13066 1:N:0:116	638	0
M00532:128:000000000-AK1TL:1:2114:19384:19846 1:N:0:116	638	e-179
M00532:128:000000000-AK1TL:1:2105:9305:16971 1:N:0:116	630	e-179
M00532:128:000000000-AK1TL:1:2107:15131:20474 1:N:0:116	622	e-177
M00532:128:000000000-AK1TL:1:2107:15124:20491 1:N:0:116	615	e-175
M00532:128:000000000-AK1TL:1:2101:5572:11361 1:N:0:116	615	e-175
M00532:128:000000000-AK1TL:1:1101:20898:11234 1:N:0:116	615	e-175
M00532:128:000000000-AK1TL:1:2102:10518:3695 1:N:0:116	575	e-163
M00532:128:000000000-AK1TL:1:2119:22513:2141 1:N:0:116	496	e-139
M00532:128:000000000-AK1TL:1:2111:20678:4880 1:N:0:116	480	e-134
M00532:128:000000000-AK1TL:1:2111:20691:4895 1:N:0:116	464	e-129
M00532:128:000000000-AK1TL:1:1107:17189:9714 1:N:0:116	428	e-118