CHARACTERISATION OF MICRO- AND MINISATELLITE DNA MARKERS FOR GENETIC DIVERSITY ANALYSIS OF THE TICK VECTOR *Rhipicephalus appendiculatus* (ACARI: IXODIDA)

By

ESTHER GATHONI KANDUMA

H80/8028/2006

A thesis submitted in fulfillment of the requirements for the award of Doctor of Philosophy in Biochemistry, University of Nairobi, Kenya.

September, 2014

DECLARATION

This thesis is my own, original work and to my knowledge it has not been presented for a degree or diploma in any other university.

Signed: ……. Date ……..

Esther G. Kanduma

This thesis has been submitted for examination with our approval as University supervisors

Signed: ……. Date ……..

Prof. P.W. Kinyanjui

Department of Biochemistry, University of Nairobi

Signed: ……. Date ……..

Dr. Richrad P. Bishop

International Livestock Research Institute (ILRI)

Signed: ……. Date ……..

Dr. R. A. Skilton

Biosciences eastern central Africa-International Livestock Research Institute (BecA-ILRI) Hub

DEDICATION

To my husband Steve and sons Jeremy and Adrian for their great love, moral support and understanding through out my PhD program

ACKNOWLEDMENTS

I wish to sincerely thank my supervisors, Prof. P.W. Kinyanjui, Dr. R. P. Bishop, Dr. R.A Skilton and the late Dr. P. P. Wambua and Dr. M. Kibe for their great technical advice and support, countless hours of reading, encouraging, and most of all patience throughout the entire process.

I am grateful to collegues at the International Livestock Research Institute (ILRI) and the Biosciences eastern and central Africa (BecA) Hub at ILRI for their kind assistance and moral support through out my PhD work. Special thanks go to Salma Hassan, Gaby Mbanjo, Naftaly Wangombe, Sonal Patel, Moses Njahira, Lucy Muthui, Inosters Inzuki, Timothy Kingori, Alice Njeri, Valerian Aloo, Rachel Mwangi and Berine Ada. Their encouragement and friendship made my stay at BecA memorable and the completion of my research an enjoyable experience. I am also grateful for the great encouragement I received from all my friends and collegues at the Department of Biochemistry, University of Nairobi; thank you Prof Mbaya, Lucy Kariuki, Virginiah Wangondu, Lucy Kamau, Faustus Mbaya, Eva Gakii and Evans Nyamboga for your prayers and kind support.

I am indebted to Joram Mwacharo (Centre for Genetics and Genomics, University of Nottingham) for his assistance with data analysis and the Kenya Wildlife Service (KWS) for granting permission to sample ticks from wildlife parks under their jurisdiction. I appreciate the technical assistance with tick identification I received from Edward Kariuki (Kenya Wildlife Service (KWS)), Abdalla Latif and Heloise Heyne (Agricultural Research Centre-Onderstepoort Veterinary Institute (ARC-OVI)).

I am greatly indebted to many donors who supported this project financially. This project was supported through a fellowship granted by the Biosciences eastern and central Africa Network (BecANet) through the New Partnership for Africa's Development (NEPAD), which was funded by the Canadian International Development Agency (CIDA). The project was also supported by the German Academic Exchange Service (DAAD) and the African Women in Agricultural Research and Development (AWARD). I gratefully acknowledge the financial support provided to the BecA-ILRI Hub by the Australian Agency for International Development (AusAID) through a partnership between Australia's Commonwealth Scientific and Industrial Research Organisation (CSIRO) and the BecA-ILRI Hub; and by the Syngenta Foundation for Sustainable Agriculture (SFSA), which made this work possible. **To God be the Glory.**

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF PUBLICATIONS, DRAFT MANUSCRIPTS AND CONFERENCE PRESENTATIONS

Publication

1. **Esther G. Kanduma**, Joram M. Mwacharo, Jack D. Sunter, Inosters Nzuki, Stephen Mwaura, Peter W. Kinyanjui, Michael Kibe, Heloise Heyne, Olivier Hanotte, Robert A. Skilton & Richard P. Bishop. (2012). Micro- and minisatellite-expressed sequence tag (EST) markers discriminate between populations of *Rhipicephalus appendiculatus*. *Ticks and Tick-borne Diseases*. Volume 3, Issue 3, Pages 128–13.

Draft Manuscripts

1. Multi-locus genotypes reveal the population genetic structure and dynamics of the brown ear tick (*Rhipicephalus appendiculatus*) in Kenya
2. Phylogenetic analysis of mitochondrial genes reveal that two genetically divergent lineages of *Rhipicephalus appendiculatus* exists in Sub-saharan Africa

Conference presentations

1. **12th Kari Biennial Scientific Conference: 8 – 12 November 2010. KARI Headquarters, Nairobi, Kenya**- Application of EST-SSR markers in the study of genetic diversity of *R. appendiculatus* (Acari: Ixodidae).
2. **7th Tick and Tick-Borne Pathogens International Conference (TTP 7): 28th August-2nd September, 2011. City Auditorium, Zaragoza, Spain**- Polymorphic EST-SSR markers and their application in assessing the genetic diversity of laboratory and Kenyan field populations of the tick *Rhipicephalus appendiculatus* (Acari: Ixodidae).
3. **9th University of Nairobi, Faculty of Veterinary Medicine Biennial Scientific conference and Exhibition**- Genetic diversity and population structure of *Rhipicephalus appendiculatus* in Kenya.

LIST OF ABBREVIATIONS AND ACRONYMS

ABI	Applied Biosystems
AFLP	Amplified fragment length polymorphism
AMOVA	Analysis of Molecular Variance
ARC-OVI	Agricultural Research centre- Onderstepoort Veterinary Institute
BecA	Biosciences eastern and central Africa
BIC	Bayesian Information Criterion
bp	base pair
COI	Cytochrome oxidase subunit I
Cytb	Cytochrome b
DAPC	Discriminant analysis of principal components
DNA	Deoxyribonucleic Acid
dNTP	deoxyribonucletide triphosphate
ECF	East Coast Fever
EDTA	Ethylenediaminetetra-acetic acid
gDNA	genomic Deoxyribonucleic Acid
GIS	Geographic Information System
H_e	Expected Heterozygosity
H_O	Observed Heterozygosity
HWE	Hardy-Weinberg equilibrium
ILRI	International Livestock Research Institute
ITM	Infection and treatment method
ITS2	Internal transcribed spacer 2
KWS	Kenya Wildlife Service
MCoA	Multiple Co-inertia Analysis
MJ	Median Joining Network
ML	Maximum Likelihood
MNA	Mean Number of Alleles
mtDNA	Mitochondrial Deoxyribonucleic Acid
mv	median vector
NJ	Neighbour joining
PA	Private allele

PCA	Principal Component Analysis
PCR	Polymerase Chain Reaction
PIC	Polymorphic Information content
RAPD	Random Amplified Polymorphic DNA
rDNA	Ribosomal Deoxyribonucleic Acid
RFLP	Restriction Fragment Length Polymorphism
rfu	Relative fluorescent units
RH	Relative Humidity
RI	Raggedness Index
RNA	Ribonucleic acid
SD	Standard Deviation
SNP	Single Nucleotide Polymorphisms
SSD	Sum of Squares Deviation
TBD	Tickborne disease
TC	Tentative consensus
TNA	Total number of alleles
Tv	Topological value
UV	Ultraviolet
VNTR	Variable number tandem repeats

GENERAL ABSTRACT

Rhipicephalus appendiculatus (Acari: Ixodidae) is the tick vector of *Theileria parva*, the causative agent of East Coast fever (ECF), considered to be the most important disease of cattle in 11 countries in Sub-Saharan Africa. ECF kills some one million cattle per year and is responsible for nearly half of calf deaths where the disease occurs. It is associated with high levels of mortality, especially in exotic and cross-bred cattle and is a major constraint to improvement of livestock production in East, Central and Southern Africa. Previous studies on variation within *R. appendiculatus* have focused on biological and behavioral differences between populations from different geographic regions. However, the genetic diversity and population structure of this important tick vector remains unknown due to absence of appropriate genetic markers. Thus molecular tools are required to characterize the genetic variation of *R. appendiculatus* and other rhipicephaline species in order to develop optimal control strategies for tick-borne diseases (TBDs). Expressed Sequence Tags (ESTs) and sequenced data such as bacterial artificial chromosome (BAC) are valuable resources for identifying micro- and minisatellite markers through data mining. This approach is fast, inexpensive and efficient. The availability of ESTs and BAC sequences for *R. appendiculatus* provided a unique opportunity to develop micro and minisatellite markers to accelerate research aimed at studying the population structure and genetic diversity of this important tick vector. This study focused on the development and evaluation of a panel of micro- and minisatellite markers which were used to characterize the genetic diversity within and between populations of *R. appendiculatus* and other rhipicephaline species. The sequences of the mitochondrial cytochrome c oxidase subunit I (COI), 12S rDNA and the complete nuclear second internal transcribed spacer (ITS2) were used to confirm the taxonomic status of *R. appendiculatus* and define the phylogenetic relationship between *R. appendiculatus* and three other closely related rhipicephaline species which are difficult to discriminate phenotypically due to their morphological similarities. Sixty six micro- and minisatellite markers were identified through analysis of the *R. appendiculatus* Gene Index (RaGI) EST database and selected BAC sequences. PCR primers were designed to flank these 66 markers and tested. Twenty nine markers all from the EST database were found to be polymorphic and therefore informative for genetic studies. These markers were used to genotype 979 individual ticks from 10 field populations, 10 laboratory-bred stocks and five additional rhipicephaline species. Based on the alleles generated, the percentage Principal Component Analysis (PCA) implemented using Multiple Co-inertia analysis (MCoA) clustered populations of *R. appendiculatus* into two groups. Twenty-three out of the 29 polymorphic EST

markers generated amplicons in *R. zambeziensis, R. praetextatus, R. pravus, R. pulchellus,* and *R. evertsi.* Genetic diversity analysis showed that the phylogeographic structure of *R. appendiculatus* may not be influenced by animal hosts and agroecological factors. STRUCTURE and discriminant analysis of principle components (DAPC) revealed a very weak genetic structure/substructure and differentiation among field ticks. The spatial genetic structure was characterized by a significant pattern of lack of isolation by distance with no evidence for the presence of distinct genetic clusters for the field tick populations. Phylogenetic analysis of COI identified two major haplogroups of *R. appendiculatus* in Kenya. Mismatch distribution analysis based on the COI gene revealed that one of the haplogroup has diverged from the other and has experienced recent sudden population expansion at a much faster rate than its ancestor.Analysis of the ITS2 gene in *R. appendiculatus* suggested high conservation of this gene within tick species while the 12S rRNA gene resulted in two major haplotypes that closely correlated with those of COI gene. COI and 12S rRNA genes were found to be good genetic markers for intra-species population genetic studies of *R. appendiculatus* compared to ITS2. Results of this study revealed that the genetic differentiation of *R. appendiculatus* may not be just driven by agro-ecological and climatic factors as previously thought and thus identifying other forces driving its differentiation may help in understanding the apparent sudden population expansion observed within the Sub-Saharan region. The findings of this study may have important taxonomical and distribution implications for *R. appendiculatus* and may point to an ongoing speciation in Sub-Saharan Africa where the tick occurs. Knowledge of the population structure of *R. appendiculatus* and its determining parameters is crucial in understanding the dispersal and transmission dynamics of the vector and its pathogens. Understanding the population structure of *R. appendiculatus* would be important in the design of sustainable control strategies as different tick populations may present differences in vector competence, rates of infectivity with *T. parva* and acaricide resistance. Thus, it is important to establish how the biology of the two major halogroups compares in regards to acquisition and transmission of ECF as this may be useful in the design and implementation of effective control strategies including the development of effective anti-*R. appendiculatus* tick vaccines.

CHAPTER ONE: GENERAL INTRODUCTION AND LITERATURE REVIEW

1.0 General introduction

Ticks are obligate haematophagous animal ectoparasites with a world-wide distribution. They are classified in the Phylum *Arthropoda*, Subphylum *Chelicerata*, Class *Arachnida*, Order *Acarina* and Superfamily *Ixodidoidea*. The Order *Acari* is divided into three families, *Ixodidae* (hard ticks), *Argasidae* (soft ticks) and *Nuttalliedae* (an intermediate group) (Hoogstraal & Aeschlimann, 1982; Mans & Neitz, 2004). Ticks of the family *Ixodidae* are vectors of infectious agents that cause life-threatening diseases of livestock and human. The most important genera of hard ticks are *Amblyomma, Dermacentor, Haemaphysalis, Hyalomma, Ixodes, Rhipicephalus and Boophilus.* The genus *Rhipicephalus* comprises 70 species which are found on many mammals in Africa (Walker *et al.*, 2000). *R. appendiculatus*, the brown ear tick, is the most important rhipicephalid tick in East and Southern Africa where it occurs on a wide variety of domestic and wild ruminants (Norval *et al.*, 1992). It is the main vector of *Theileria parva*, a protozoan parasite that causes East coast fever (ECF) in cattle. *R. appendiculatus* occurs in Eastern, Central and Southern regions of Africa (Norval *et al.*, 1992) and its distribution is correlated with wet humid climatic conditions where animal host availability and vegetation are favourable (Norval & Perry, 1990; Walker *et al.*, 2003).

The major economic importance of ticks is their ability to transmit a wide range of pathogenic micro-organisms: arboviruses, fungi, bacteria, spirochetes and parasitic protozoa of human, domestic and wild animal diseases. Tick-borne protozoan diseases (e.g. theileriosis and babesiosis) and rickettsial diseases (e.g. anaplasmosis and cowdriosis) are major health and management problems of livestock industry in many developing countries. These diseases result in morbidity and mortality of livestock and decreased productivity (Obenchain & Galun, 1982). Global economic

losses of livestock due to tick-borne diseases, together with costs of control measures have been estimated at 14-17 billion dollars annually (de Castro, 1997). It is estimated that 80% of the world's cattle are infested with ticks and as a result, vast regions are unable to sustain viable livestock production without active tick or tick-borne disease control measures.

East Coast fever is a fatal lymphoproliferative disease, considered to be the most economically important tick-borne disease of cattle in East, Central and Southern Africa. It is associated with high levels of mortality, especially in exotic and cross-bred cattle. This disease is a major constraint to improvement of livestock production across this part of Africa. The life cycle of *T. parva* in the bovine host involves two intracellular stages: the schizont, which transforms lymphocytes and is responsible for disease pathology, and the intra-erythrocytic piroplasm, which is infective to the tick vector, *Rhipicephalus appendiculatus* (Norval *et al.,* 1992). The mammalian life cycle stages are haploid, but there is a transient diploid phase in the tick gut after fusion of gametes (Gauer *et al.,* 1995) and sexual recombination has been observed between and within *T. parva* stocks during experimental transmission (Morzaria *et al.,* 1993; Bishop *et al.,* 2002a; Katzer *et al.,* 2006). The African buffalo (*Syncerus caffer*) is the natural reservoir of *T. parva*, but the parasite does not cause disease in this species. Transmission of buffalo-derived *T. parva* to cattle results in a rapidly lethal disease, but in many cases the parasites do not differentiate to the erythrocyte-infective stage and are not transmissible by ticks (Morrison *et al.,* 1989; Uilenberg, 1999). Hence, although *T. parva* parasites that originate from buffalo are genotypically and antigenically closely related to *T. parva* maintained in cattle, available evidence indicates that a significant proportion of the buffalo-derived population cannot be transmitted between cattle.

The main strategy for ECF control has been predominately based on controlling *R. appendiculatus* through intensive application of acaricides in an integrated approach. Acaricides such as amitraz are applied once or twice a week to prevent tick infestation and subsequently prevent transmission of *T. parva.* Control of ticks allows introduction of improved exotic breeds of cattle and upgrading of the indigenous cattle population in ECF affected areas. However, intensive tick control using acaricide prevents any contact of cattle with ticks which does not allow calves to develop immunity to tick borne diseases (TBDs). Resistance to existing classes of acaricides is widespread and its incidence is increasing. Newer classes of acaricides have tended to be significantly more expensive than their predecessors (Willadsen, 2004). In addition, there is increasing concern about the use of chemicals in all forms of agriculture, both for their potential environmental impact as well as for their presence in food products.

ECF has also been controlled through vaccination with live sporozoites prepared from laboratory stocks of *R. appendiculatus* ticks infected with *T. parva* (Di Giulio *et al.,* 2009*).* The live *T. parva* sporozoites, stored as cryopreserved stabilates, are inoculated simultaneously together with a long-acting oxytetracycline dose as an infection and treatment method (ITM). This vaccine has been used in several African countries, including Zambia (Uilenberg, 1999), parts of Kenya, Tanzania (Di Giulio *et al.,* 2009, Babo Martins *et al.,* 2010), Uganda and Malawi (Kasibule, 2013, McKeever *et al.,* 1995) as a cocktail of three strains of *T. parva* stabilates with an infection and treatment protocol to protect cattle against East Coast fever. Immunization of cattle with ITM is able to successfully protect 96% of immunized calves from ECF. Immunisation of cattle by infection with *T. parva* and simultaneous treatment with long-acting tetracycline results in long term immunity against the homologous parasite strain but variable protection against challenge with heterologous parasite strains. Hence, vaccination of cattle in the field by this method requires a mixture of

parasite strains. Development of vaccines that target both ticks and pathogen transmission may provide a sustainable method of controlling tick-borne diseases through immunization (Willadsen, 2004). Control of ticks by vaccination has the advantages of being cost-effective, reducing environmental contamination and preventing the selection of drug-resistant ticks that result from repeated acaricide application. Development of vaccines against ticks would also allow for inclusion of multiple antigens that could target a broad range of tick species and may also prevent transmission of pathogens (Bishop *et al*., 2004, Willadsen, 2004).

Most studies regarding genetic variation of *R. appendiculatus* have focused on biological and behavioral differences between populations of *R. appendiculatus* in different geographical regions (Norval *et al.,* 1992, Shaw & Young, 1994, Short & Noval, 1981, Rechav, 1982, Pegram & Banda, 1990, Berkvens *et al.,* 1995; Chaka *et al.,* 1999; Madder *et al.,* 2002; Speybroeck *et al.,* 2002, 2004). It is possible that in a given geographical region, a diverse genetic population of ticks exist. An assessment of a tick/pathogen population structure would be relevant to aspects of disease control such as live vaccine development, spread of strains and strain typing. The geographical distribution and separation of disease vector populations through physical and/or biological factors may create barriers to gene flow between vector populations and are important in population differentiation.

Molecular characterization of tick species has been undertaken by analyzing DNA variations in both coding and non-coding regions in the genome. Availability of a wide array of molecular genetic markers offers tools for quick detection and characterization of genetic variations. Simple sequence repeats (SSRs) also referred to as microsatellites and minisatellites have been used to study population structures of many organisms (Hamada *et al.,* 1982; Tauz & Renez, 1984; Gemayel,

2010). Microsatellite analysis of genetic diversity provides information on allele frequency differences within and between populations as well as providing information that allows estimation of levels of gene flow among populations. It provides information on the genetic relatedness and diversity in geographically separated vector populations. Understanding the extent and importance of genetic exchange in *R. appendiculatus* populations has practical implications, as sexual reproduction along with high levels of recombination would generate a population with a high degree of genetic diversity.

This study used several PCR-based DNA markers to analyze the phylogeny of *R. appendiculatus* and other closely related rhipicephaline ticks species. A new set of 29 polymorphic EST-micro- and misatellite DNA markers were identified, evaluated and characterized for the first time. The 29 informative polymorphic micro- and minisatellite markers are the first available tools for the analysis of the phylogeography and population genetics of *R. appendiculatus.* They were applied to estimate the genetic diversity within populations of *R. appendiculatus* from the field and laboratory stocks and tick populations from closely related rhipicephaline species. The markers can also be used for differentiating *R. appenduculatus* from closely related tick species in the field. In addition these markers could be applied to ensure standardisation and quality control of the live *T. parva* vaccine. The sequences of mitochondrial cytochrome c oxidase subunit I (COI), 12S rDNA and the complete nuclear second internal transcribed spacer 2 (ITS2) were used to confirm the taxonomic status of *R. appendiculatus* and define the phylogenetic relationship between *R. appendiculatus* and three other closely related rhipicephaline species which are difficult to discriminate phenotypically due to their morphological similarities. Knowledge of the population structure of *R. appendiculatus* and its determining parameters is crucial in understanding the dispersal and transmission dynamics

of the vector and its pathogens as well as in the design of optimal control strategies including the development of effective anti-*R. appendiculatus* tick vaccines.

1.1 Literature review

1.1.1 Importance of livestock and agriculture to economies of developing countries

Livestock play a vital role in the agricultural and rural economies of the developing world. Economic development of many farmers in many developing countries is contributed by livestock. In most African regions, livestock provides food, income, employment and contributes to improvement of agriculture through draft power, manure, fuel, and as a fertilizer. Livestock does not only produce food directly, it is also a major and sometimes the only income earner for many populations especially pastoralists in developing countries. Larger animals such as cattle are a capital reserve, built up in good times to be used when crops are poor. However, protozoan, rickettsial and viral diseases of livestock which are transmitted by tick vectors result in morbidity and mortality of livestock and lead to decreased productivity and huge losses in incomes (Obenchain & Galun, 1982)

1.1.2 Ticks

Ticks are haematophagus ectoparasites that belong to the Phylum *Arthropoda*. They make up the largest collection of creatures in the order *Acarina*. Ticks are the most important ecto-parasites of livestock in tropical and sub-tropical areas, and are responsible for severe economic losses in livestock. The medical and economic importance of ticks is due to their ability to transmit diseases to humans and animals. They are vectors of a variety of viral, prokaryotic and eukaryotic organisms which when transmitted to mammalian hosts may cause disease (Sonenshine, 1991; Rotz *et al.*, 2002).

1.1.3 Classification of ticks

Ticks belong to the suborder *Ixodida* in the order *Acari*, class *Arachnida*, phylum Arthropoda. The order *Acari* is divided into three families, *Ixodidae* (hard ticks), *Argasidae* (soft ticks) and *Nuttalliedae* (an intermediate group) (Hoogstraal & Aeschlimann, 1982; Mans & Neitz, 2004). There are 899 tick species that parasitize vertebrates including Argasidae (185 species), Ixodidae (713 species) and Nuttalliellidae (1 species) (Barker & Murrell, 2004). Argasid ticks include four genera, namely *Argas, Carios, Ornithodoros* and *Otobius*, whereas the ixodid ticks consists of over 240 species in the genus *Ixodes* and over 440 species in the remaining genera (Horak *et al.*, 2002). Globally, the most important genera of hard ticks are *Amblyomma, Dermacentor, Haemaphysalis, Hyalomma, Ixodes, Rhipicephalus and Boophilus.* The genus *Boophilus* was recently found to be phylogenetically related to genus *Rhipicephalus* by Horak *et al.,* (2002) and thus the five species of *Boophilus* are now considered as a subgenus of *Rhipicephalus* (Murrell *et al.,* 2000, 2001, Horak *et al.,* 2002).

1.1.4 Global importance of ticks

Ticks and tick-borne diseases affect animal and human health worldwide and are the cause of significant economic losses. Approximately 10% of the currently known 899 tick species act as vectors of a broad range of pathogens of domestic animals and humans and are also responsible for damage directly due to their feeding behaviour. The most important tick species and the diseases/effects they cause are listed in Tables 1. Tick-borne protozoan diseases (e.g. theilerioses and babesioses) and rickettsial diseases (e.g. anaplasmoses and heartwater or cowdriosis) are the major causes of mortality and morbidity of cattle and small ruminants, constraining livestock farming in communities in Africa, Asia and Latin America (Perry *et al.*, 2002; Minjauw &McLeod, 2003). The impact of ticks and tick-borne diseases on the global economy is considered to be high

and although some estimates are given, there is a lack of reliable data (Jongejan & Uilenberg, 2004). The genus *Rhipicephalus* comprises 70 species which are found on many mammals in Africa. *R. appendiculatus*, the brown ear tick, is the most important rhipicephalid tick in East and Southern Africa where it occurs on a wide variety of domestic and wild ruminants. The tick transmits *Theileria parva* which causes East Coast Fever, an important disease of cattle that occurs in 11 countries in sub-Saharan Africa. The tick also transmits Thogoto virus which causes Nairobi Sheep Disease. *Rhipicephalus zambeziensis* found in certain parts of Southern Africa including Zambia, Zibambwe and South Africa also transmits *T. parva* (Norval *et al.*, 1992).

1.1.5 Geographical distribution of ticks

Tick distributions are known to vary through space and time due to interactions of many factors, including climate, host diversity, levels of resistance of hosts, absence of tick control measures and management practices that affect host behavior (Randolph, 2000). Tick species parasitise domestic animals such as cattle, sheep goats, pigs, chicken, pets (such as dogs and cats), and virtually all wildlife mammals as well man. Geographic distribution of tick species, and indirectly of tick-borne diseases, is limited by climatic factors (Estrada-Peña, 2001). Ticks are well-adapted to survive in tropical, temperate, and even subarctic habitats. Both temperature and relative humidity are key factors governing the main aspects of the tick life-cycle. The “growth periods” (i.e. oviposition or moulting) are temperature mediated, in a very different way according to tick species.

Table 1: Important tick species in the family *Ixodidae*

Tick	**Disease caused**	**Pathogen**
Amblyomma variegatum (Fabricius, 1794)	Cowdriosis of ruminants	*Ehrlichia (Cowdria) ruminantium*
	Benign African theileriosis of cattle	*Theileria mutans;*
	African tick bite fever in man;	*Rickettsia africae*
Hyalomma anatolicum anatolicum (Koch, 1844)	Bovine tropical theileriosis	*Theileria annulata*
	Theileriosis of small ruminants	*Theileria lestoquardi*
Haemaphysalis leachi (Audouin, 1826)	Canine babesiosis	*Babesia rossi*
Haemaphysalis longicornis (Neumann, 1901)	Bovine babesiosis	*Babesia ovata*
	East Asian bovine theileriosis	*Theileria buffeli*
	Canine babesiosis	*Babesia gibsoni*
	Human rickettsiosis	*Rickettsia japonica*
Ixodes ricinus (Linnaeus, 1758)	Tick-borne encephalites in man and animals	*Borrelia burgdorferi sensu lato*
	Lyme borreliosis	*Borrelia burgdorferi sensu lato*
	Babesiosis in cattle and sporadically in man	*Babesia divergens*
	Human babesiosis	*Babesia microti*
Ixodes scapularis (Say, 1821)	Lyme borreliosis	*Anaplasma (Ehrlichia) phagocytophilum*
	human babesiosis	*Babesia microti*
Rhipicephalus evertsi evertsi (Neumann, 1897)	Equine piroplasmosis	*Babesia caballi* *Theileria equi*
	Bovine anaplasmosis	*Anaplasma marginale*
Rhipicephalus sanguineus (Latreille, 1806)	Canine ehrlichiosis	*Ehrlichia canis*
	Canine babesiosis	*Babesia vogeli*
	Tick bite fever in humans	*Rickettsia conorii*
Rhipicephalus appendiculatus (Neumann, 1901)	East Coast Fever in cattle	*Theileria parva*
	Benign bovine theileriosis	*Theileria taurotragi*
	Nairobi Sheep Disease virus	*Thogoto virus*
Rhipicephalus zambeziensis (Walker, Norval & Corwin, 1981)	East Coast fever in cattle	*Theileria parva*
	Benign bovine theileriosis	*Theileria taurotragi*
Rhipicephalus (Boophilus) decoloratus (Koch, 1844)	Bovine babesiosis	*Babesia bigemina*
	Bovine anaplasmosis	*Anaplasma marginale*
Rhipicephalus (Boophilus) microplus (Canestrini, 1888)	Bovine babesiosis	*Babesia bovis* and *Babesia bigemina*
	Equine piroplasmosis	*Theileria equi*
	Bovine anaplasmosis	*Anaplasma marginale*

1.1.6 Tick body structure

Ticks are ventro-dorsally compressed and usually have no definite division between the head, thorax and abdomen. The detailed tick structure is given by Sonenshine (1991). Sexes are separate. Larvae have 3 pairs of legs and the nymphs and adults 4 pairs. The nymph is sexually immature (has no genital aperture). There is no distinct head region as we normally define it. The synganglion (the primitive "brain"), the eyes and the salivary glands are all found in the body of the tick. Mouthparts are borne on a moveable part of the body called the basis capituli. Together, the basis capituli, the mouthparts and the palps form the capitulum ("little head"). The cuticle is thick, leathery and flexible with folds which allow for expansion during feeding. Mouthparts comprise 2 palps (sensory function), 2 chelicerae (cutting) and 1 hypostome (anchorage and feeding tube). The hypostome is armed with backward projecting teeth. The chelicerae are armed with movable denticles and the lateral stigmata are without sinuous peritremes. The "soft" ticks are characterised by the absence of a dorsal cuticular shield (scutum) and when viewed from above, no projecting mouth parts are visible (they project ventrally). The "hard" ticks have a dorsal shield (scutum) and their mouth parts (capitulum) project forward (anteriorly) when viewed from above. The scutum covers the entire dorsal surface in males but only part of the dorsal surface of females. Tick species identification is based on several structural features such as size and shape of mouth parts, palps and basis capitulum (capitulum), presence or absence of eyes and festoons, presence and shape of anal groove, shape, patterns as well as nature of punctations on scutum. Figure 1 shows the generalized structure of ixodid male and female ticks. The detailed structure of *R. appendiculatus* has been illustrated by Walker *et al.*, (2000, 2003).

1.1.7 Tick life cycle

Different tick species exhibit life cycle patterns involving feeding on one, two or three hosts during the life cycle, resulting in one-host, two-host or three-host ticks respectively (Walker *et al.*, 2000, 2003). The life cycle may take up to 3 weeks in total. Pathogens are usually acquired by immature ticks during blood feeding and transmitted by the next tick instar, although some organisms are vertically transmitted within tick populations via a transovarial route e.g. in *Boophilus* subgenus. According to the number of hosts they require during their life cycle, the hard ticks generally can be classed in three groups (one host tick, two host tick and three host tick). In one host ticks, all three instars engorge on the same animal, the two ecdyses also taking place on the host e.g. *Boophilus microplus, B. decoloratus & B. annulatus*. They stay on the body of an individual host throughout the feeding and molting phases of their life cycle. The larva is the only instar involved in host location and when engorged does not drop off the host, but molt into nymphs which engorge while on the host and in approximately three weeks engorged adults emerge. Two-host ticks such as some species of *Rhipicephalus* and *Hyalomma* initially locate hosts as larvae and again as adults. The larva engorges and moults on the host and the nymph drops after also having engorged; it moults on the ground and the imago seeks a new host- eg *Rhipicephalus evertsi, R. bursa*. Most species of *Ixodidae* e.g. *R. appendiculatus* are three-host ticks and each instar, larva, nymph and adult locates and then feed on a separate host. These require a different host for every instar; they drop off each time after having engorged and moult on the ground e.g. *Ixodes holocyclus, I. ricinus, R. appendiculatus* and most other ticks. In two- and three- host ticks, different individuals of a host species are parasitized during the life cycle. Each of the life cycle stages, larvae, nymphs and adult of ixodid ticks has a single instar. Ixodid ticks remain attached to the host for up to ten days while feeding. The period of blood ingestion in these ticks can last 2-3 days in the larvae and nymphs to more than 1 week in some adult females.

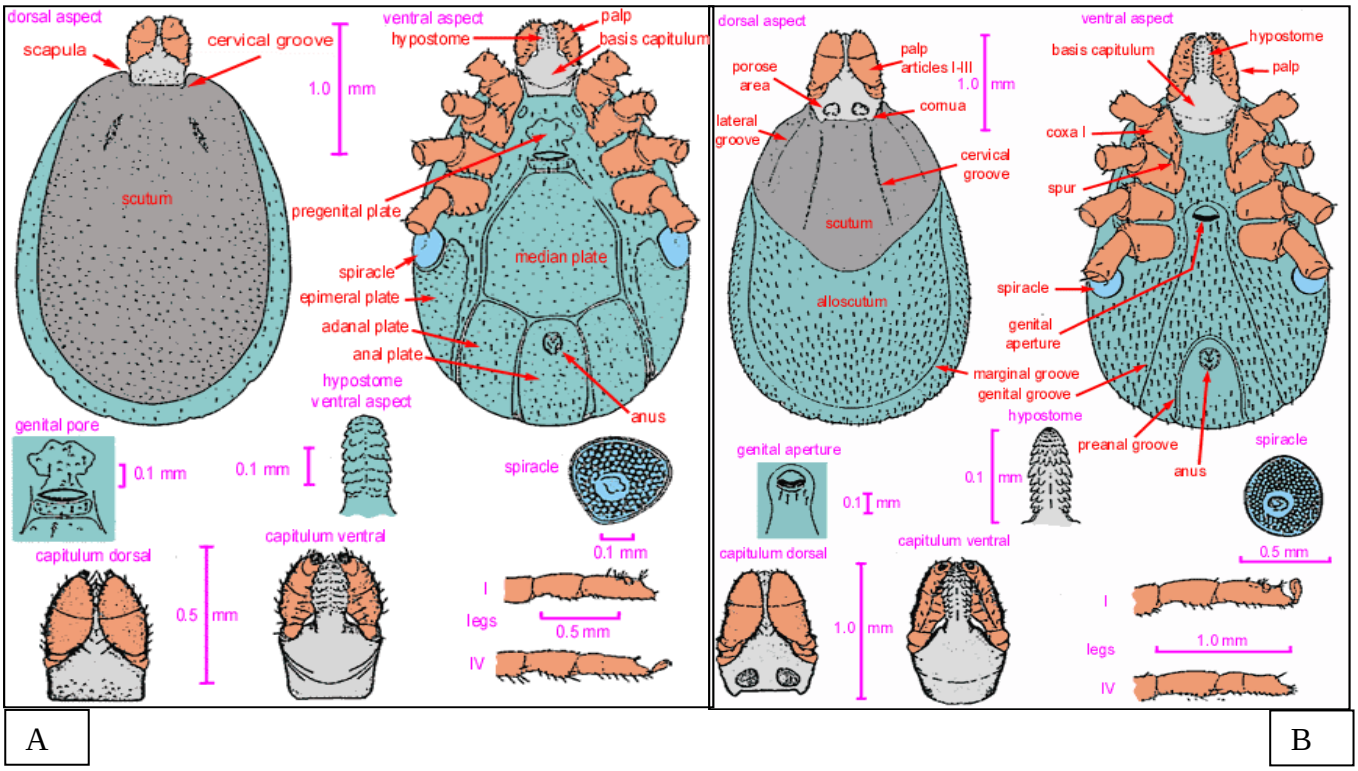


Figure 1: Generalised structure of an ixodid male tick (A) and female tick (B).
Adopted from http://www.tickalert.org.au/index.htm based on Sonenshine, 1991: Biology of Ticks, 2 volumes: Oxford University Press, New York, Oxford, 1991

1.1.8 Tick feeding /salivary glands/osmoregulation

For ixodid ticks, the salivary glands are vital organs, playing several key roles during the arthropod's life cycle. A tick bite should result in strong responses from the host defence systems (haemostatic, immune and inflammatory) but tick saliva appears to have evolved to counter these responses. The saliva components include enzymes, enzyme inhibitors, amine-binding proteins and cytokine homologues that act as anti-haemostatic, anti-inflammatory or immuno-modulatory agents. When attaching to a host, the salivary glands of most species produce cement that binds the tick's mouthparts to the host's skin (Moorhouse & Tatchell, 1966). There is usually pathogen replication in the tick and migration of organisms to the tick salivary gland which is responsible for osmoregulation via saliva (Sauer *et al.,* 2000). Subsequently, a cocktail of enzymes and other

bioactive molecules is secreted, the composition of which appears to change throughout the extensive feeding period (7–14 days). Release of pathogens coincides with the excretion of excess water during blood feeding and infectivity may be enhanced by components of tick saliva which modulate host immune, haemostatic and inflammatory responses (shown in Figure 2) (Nuttall *et al.*, 2000; Wikel & Alacron-Chaidez, 2001). The bioactive saliva creates a feeding lesion, maintains blood flow to the feeding site, antagonizes host haemostatic and inflammatory mediators, and helps the tick evade the host's rejection responses (Nuttall, 1998; Ribeiro, 1989; Wikel, 1996). Throughout the feeding period, the salivary glands also excrete excess fluid of the blood meal, which results in nutrient concentration and osmoregulation (Sauer *et al.*, 1995). The risk of acquiring tick-borne diseases after a tick bite depends on several factors, the length of exposure to the tick, the prevalence rate of a pathogen in the tick population, coinfection with some other pathogens in the tick, the pathogen strains in the infected tick, and the species of tick that thrives in a particular region. Among all these factors, it seems likely that the duration of tick attachment is one of the most important risk factors associated with successful transmission of *T. parva* by *R. appendiculatus* (Konnai *et al.*, 2007). Thus, the acquisition, development and transmission of pathogens involve complex multi-stage processes and the factors which influence vectorial capacity. Indigenous cattle breeds are known to resist tick infestation by acquiring resistance to ticks. It is thought that they induce anti-tick immune responses which limit the ability of ticks to successfully feed and reproduce (Willadsen, 2001).

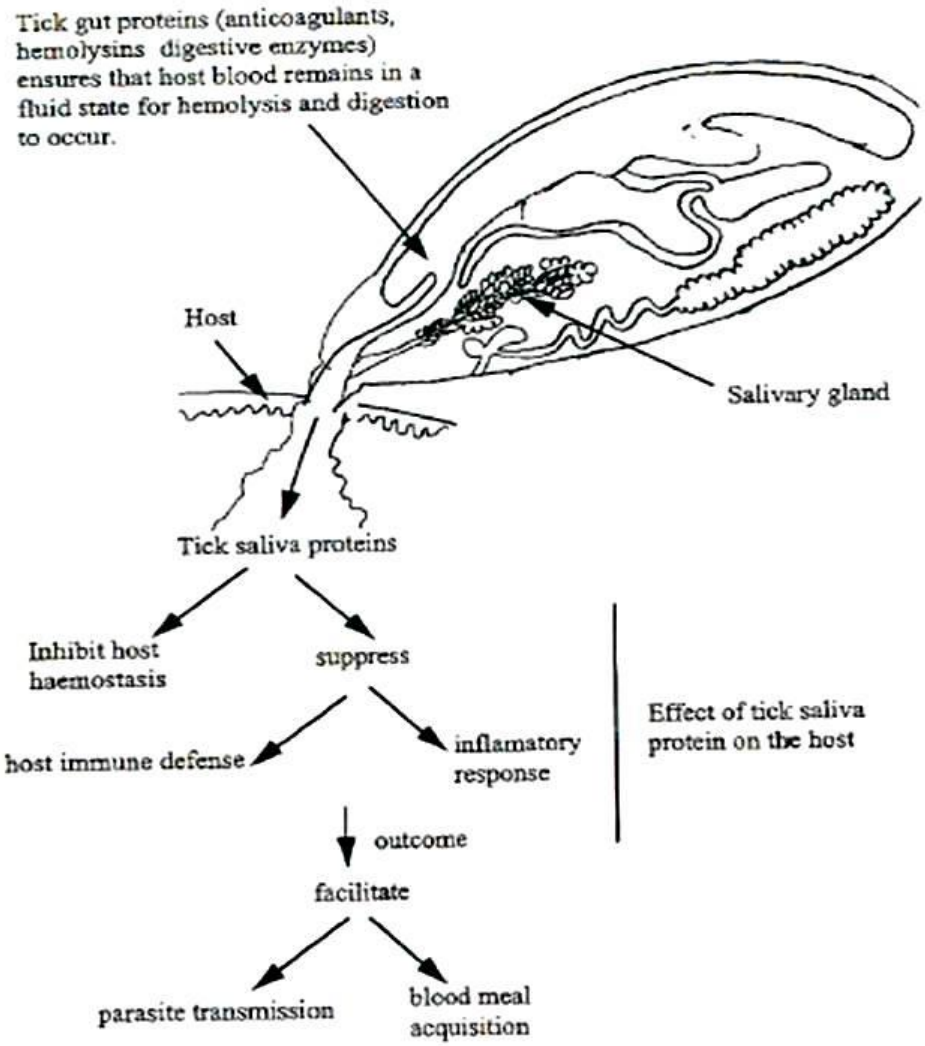


Figure 2: Schematic illustration of a feeding female tick showing the effects of tick saliva on host
The tick saliva components include enzymes, enzyme inhibitors, amine-binding proteins and cytokine homologues that act as anti-haemostatic, anti-inflammatory or immuno-modulatory agents. A cocktail of enzymes and other bioactive molecules is secreted, the composition of which appears to change throughout the extensive feeding period (7–14 days).

1.1.9 Control of ticks

Currently, tick control, where it occurs at all, is heavily dependent on two approaches: the use of chemical acaricides and the use of tick-resistant animals. Application of acaricides has been the main method of controlling ticks and tickborne diseases (TBDs). Arsenical compounds were among the first chemicals to be used as acaricides and they were successfully used to control ECF in South Africa (Lawrence, 1992). Ticks quickly developed resistance to arsenic compounds which were also implicated of being toxic to cattle and arsenic residues remaining in animal products (Shaw *et al.*, 1970). *Boophilus microplus* and *B. decoloratus* ticks developed resistance and arsenics were replaced with organochlorides in mid 1940s among them DDT and benzenehexachloride (BHC) (Cobbett, 1947; Maunder, 1949, Whitnall *et al.*, 1951). However, resistance to organochlorine chemicals was also reported and at the same time they have been shown to persist in the environment. Other compounds were investigated as acaricides among them carbamates but just like their predecessors, they were found to have cross-resistance with organophosphates (Roulston, 1968, Schuntner *et al.,* 1972). A group of chemicals effective against ticks include formamidines, chlordimeform and amitraz (Palmer *et al.,* 1971, George *et al.,* 2004). Amitraz is unstable in dipping vats though this can be overcome by adding sufficient calcium hydroxide to raise and maintain pH of the vat solution to 12 to ensure stability of the active ingredient (Stanford *et al.*, 1981; George *et al.*, 1998). Pyrethroids have also been used to control ticks (Ware, 2000, Davey & Ahrens, 1984) but just like other chemicals used before them, ticks were observed to develop resistance to this costly acaricide (Nolan *et al.,* 1979). Since the first report of the development of resistance of *B. microplus* to arsenic in Australia in 1937 (Newton, 1967) and *B. decoloratus* in South Africa in 1939 (Whitehead, 1958), the progressive evolution of resistance of ticks affecting cattle to almost all of the available acaricides has frustrated the efforts of cattle producers to manage

ticks and tick-borne diseases affecting their animals. The spectrum of chemical groups to which ticks have evolved resistance continues to broaden. Tick populations susceptible to a variety of acaricides exist and can be controlled, but it is more critical to use existing and improved diagnostic tools to determine where products are still useful and to employ tick control strategies that minimize the rate of selection for resistance (George *et al.*, 2004). The use of acaricides is thus greatly limited by development of resistant strains of ticks, environment pollution and residues in meat and other animal products (Willadsen, 1997, Nolan, 1990). In East Africa the cost of acaricide application by hand-spraying or dipping in both small-holder and traditional cattle systems amounts to US $6-36 per adult cattle per year (Minjauw & Mcleod, 2003). While there are efforts to improve on use of acaricides and mitigate resistance of ticks to acaricides, other sustainable control strategies are needed to tackle ticks and tick borne diseases. Biological control agents have been investigated for their potential in controlling ticks. These include nematodes, entomopathogenic fungi and baculoviruses. Strains of several nematode species including *Steinernema carpocapsae* and *Heterorhabditis bacteriophora* induce mortality in engorged female *R. appendiculatus, R. avertis, Amblyomma variegatum, A. gemma and Boophilus decoloratus* under experimental conditions (Kaaya *et al.,* 2000; Samish & Glazer, 2001). Unfed females and immature stages are however generally resistant (Kaaya *et al.,* 2000). The full potential of biological agents in controlling ticks needs to be explored further.

The use of disease resistant breeds is recognized as an economically realistic alternative to acaricide application and drug administration in tick and tick-borne disease control (Mattioli & Dempfle, 1995). Cattle resistance to tick attack and tick-borne micro-organisms (TBMs) varies among different breeds of cattle (Lemos *et al*., 1985; Scholtz *et al*., 1991), the species of infesting tick (Norval *et al*., 1997a, b) and the level of infestation (Sutherst *et al*., 1983). The influence of cattle

breed on tick infestation has been reported by Wambura and others who were able to show that Zebus (*Bos indicus*) had a relatively higher resistance to tick infestation compared to crossbreds (Wambura *et al.*, 1998). In a study to assess host resistance to ticks on cross-bred cattle in Burundi (Moran *et al.*, 1996), groups of cattle of four different cross-breeds (Ankole x Friesian, Ankole x Brown-Swiss, Ankole x Guernsey, Ankole x Sahiwal) plus a group of pure Ankole cattle were immunized against ECF, anaplasmosis, babesiosis and heartwater. The results indicated cattle resistance to the most abundant species of ticks, *R. appendiculatus*, as follows: Ankole > Ankole x Sahiwal > Ankole x Brown Swiss > Ankole x Friesian > Ankole x Guernsey. Another study on tick infestation on Zebu cattle in Western Kenya, revealed that highly resistance cattle showed little or no seasonal fluctuations in tick numbers compared to animals of low resistance which showed a sevenfold increase in magnitude of the tick burden when tick challenge was high (Latif *et al.*, 1991). There is evidence that the N'Dama breed found in West Africa is resistance against ticks (Murray *et al.*, 1991, Mattioli *et al.*, 1995) and this resistance is attributed to the host's skin cellular mechanisms (Mattioli *et al.*, 2000). Generally, zebu (*Bos indicus*) cattle possess a higher resistance to ticks and TBMs than European (*Bos taurus*) cattle breeds (Rechav & Kostrzewski, 1991). The host's immune system would appear to be the single most important factor that regulates this resistance such that breeds that resist tick infestation have more effective anti-tick skin reactivity (Mattioli *et al.*, 2000).

1.1.10 Control of ticks by vaccination

A novel approach designed to target tick vectors in such a way that they protect against pathogens they transmit is ant-tick vaccines. Thus this new strategy is being explored as means of controlling tick vectors and the pathogens they transmit. The feasibility of controlling tick infestations through immunization of hosts with selected tick antigens was demonstrated with the development of

vaccines that reduced *Boophilus spp*. infestations on cattle (de la Fuente & Kocan, 2003, Willadsen, 2004). Though development of an anti-*B. microplus* vaccine has largely been successful, development of a vaccine against other economically important ticks such as *Rhipicephalus* is just beginning. Anti-tick effects are expressed as tick and egg mortality, decreased engorgement and egg mass weights, prolonged tick feeding and inhibition of moulting.

Studies on tick vaccination has shown that partial to strong immunity to tick infestation can be induced by vaccination with a variety of antigenic materials, including whole tick homogenates, salivary glands and salivary gland extracts, tick internal organs, including tick gut material, cement material and so on (Willadsen, 2004). In a study by Rechav *et al.*, (1992) guinea-pigs inoculated with crude homogenate of unfed nymphs of the tick *R. appendiculatus* and with three semipurified fractions of the homogenate obtained by gel permeation chromatography, acquired a significant degree of immunity to infestation with adults of this tick. This immunity correlated with increased production of gamma-globulins in calves inoculated with the homogenate. In another study by Varma and others (Varma *et al.*, 1990), guinea-pigs immunized with homogenates of unfed larvae and nymphs of the tick *R. appendiculatus* developed significant levels of protective immunity to infestation with adults of this species. Willadsen & Kemp, (1988) have discussed tick antigens as belonging to one of two groups. The first of these includes the immunogenic materials exposed to a host by the normal processes of tick attachment and feeding. These have been referred to as 'exposed' antigens (Willadsen, 2004). These antigens are exposed naturally to the host immune system during tick infestation. Hosts immunized with these antigens are boosted by continuous tick exposure. The second group of antigens is referred to as 'concealed' antigens. These are antigens which are not part of the normal host–parasite interaction and which do not under normal circumstances stimulate an immunological response. It is possible to raise a response to these

immunogenic molecules by vaccination whose immunological response will subsequently damage the feeding tick. Typical examples of 'concealed' antigens are the immunogenic proteins located in the gut of the tick where, once an antibody to the protein is raised by vaccination, uptake of specific immunoglobulin during feeding leads to damage to the parasite (Willadsen, 2004). Exposed antigens include salivary gland proteins and cement constituents (Shapiro *et al.*, 1989, Brown & Askenase, 1986). These include a 94kDa antigen purified from *R. appendiculatus* attachment cement which was shown to elicit resistance to tick feeding by inducing a delayed-type hypersensitivity immune reaction at the site of feeding using laboratory guinea pig models (Shapiro *et al.,* 1989). Trimnell *et al.,* (2002) identified a putative *R. appendiculatus* protein, 64P whose recombinant constructs were able to stimulate significant protection to nymphal and adult infestations. Truncated constructs of 64P (64TRPs), were shown to provide cross-protection against *R. sanguineus* and *I. ricinus*, apparently by targeting antigens in the midgut and salivary glands of adults and nymphs, causing mortality (Trimnell *et al.,* 2005). A salivary gland and cement cone protein (RIM36) has been characterized from *R. appendiculatus* (Bishop *et al.,* 2002b). The protein induced a strong antibody response in tick-exposed cattle. Sugino *et al.,* (2003) have cloned, sequenced and characterized cDNA encoding *H. longicornis* serpin (HLS1). Vaccination of rabbits with recombinant HLS1 (rHLS1) expressed in *E. coli* resulted in 43.9 and 11.2% mortality of nymph and adult ticks which were fed on immunized rabbits. There has been a recent characterization of double-headed serine proteinase inhibitors from *B. microplus* (Tanaka *et al*., 1999). There is experimental evidence that these offer some immunoprotection against *B. microplus* larvae (Andreotti *et al.,* 1999, 2002). Three female-specific high-affinity histamine-binding proteins (HBPs) discovered in the saliva of *R. appendiculatus* ticks (Paesen *et al*., 1999) have been shown to supress the host inflammatory pathway. The first ever commercial anti-anthropod vaccine used against *Boophilus microplus* (Tellam *et al.*, 1992; Willadsen & Jongejan, 1999) has been developed

from a concealed antigen, a gut glycoprotein (Bm86). These commercial vaccines are TICKGARD and TickGARD^{TM} Plus which are used in Australia and GAVAC^{TM}, manufactured in Cuba (Rand *et al.*, 1989, Rodriguez *et al.*, 1994). In the field Bm86 vaccine mainly reduce the tick population in successive generations through reducing reproductive capacity. When the vaccine is used as a component of integrated control of ticks, a near 100% control of *B. microplus* infestations is achieved.

Control of ticks by vaccination has the advantages of being cost-effective, reducing environmental contamination and preventing the selection of drug-resistant ticks that result from repeated acaricide application. Development of vaccines against ticks would also allow for inclusion of multiple antigens that could target a broad range of tick species and may also prevent transmission of pathogens. A vaccine has the potential to be non-contaminating and sustainable. The development of resistance to a vaccine is less likely than for a pesticide. A single point mutation in the target molecule is sufficient to render a pesticide ineffective, while such point mutations are likely to be of little relevance to most vaccines (Willadsen, 2004). The scientific challenge in development of vaccines is producing truly efficacious vaccines able to provide tick control that is both practical and cost effective. It is considerably more difficult to develop an effective vaccine against highly diverse pathogens than against clonal pathogens and thus the population structure of a pathogen is important in relation to vaccine development. In both Australia and South America, isolate-to-isolate differences in Bm86 vaccine susceptibility was shown to occur (Garcia-Garcia *et al.*, 1999).

1.2 *Rhipicephalus appendiculatus*

R. appendiculatus (Neumann, 1901) *(Acari: Ixodidae*) is an important tick in the genus *Rhipicephalus*. It is also referred to as the brown ear tick because of its brown colour and preference

for feeding on the ears of cattle. The tick is a common ectoparasite of cattle and also feeds on sheep and goats and the larval stage is frequently found on small rodents. *R. appendiculatus* is endemic to most of Africa (Hoogstraal, 1956) and is of major veterinary and medical importance since it transmits the causative agents of East coast fever (*Theileria parva)* (Theiler, 1911) *(*Family: *Theileriidae)* and red water fever (*Babesia bigemina*) to cattle, Nairobi sheep disease virus (a nairovirus of the *Bunyaviridae* family) to sheep and goats, and boutonneuse fever (*Rickettsia conorii*) to man.

1.2.1 Animal hosts of *R. appendiculatus*

R. appendiculatus is an important ectoparasite of both domestic and wild ruminants. The tick feeds on a variety of hosts ranging from domestic cattle to buffalo which form reservoirs for parasites (Walker *et al*., 2003). All stages of the tick can feed on cattle. It also infests goats, sheep and dogs. The main wildlife reservoir for the tick in Kenya is the African Buffalo (*Syncerus caffer*). It is also an ectoparasite of elands, waterbucks, greater kudus and antelopes. In these animals, the tick transmits other species of *Theileria* which are transmissible to cattle. In domestic livestock, it is mainly infests cattle where it is a vector of several parasites. Buffalo-derived *T. parva* can be infective to cattle as well as other wild animals such as Theileria-free waterbuck (*Kobus defassa*) (Stagg *et al.,* 1994). In 1977, Young and others isolated a *Theileria* species from Eland (*Taurotragus oryx*) which was shown to be infective to cattle (Yound *et al.,* 1977).

1.2.2 Distribution of *R. appendiculatus*

R. appendiculatus shows a wide distribution in extensive parts of Africa. It occurs in Eastern, Central and Southern regions of Africa (Norval *et al.,* 1992). It also occurs in South Sudan, through to south eastern coast of South Africa. To the west, the occurrence of the tick is recorded in Western Zambia and Central Democratic Republic of Congo. The majority of *Rhipicephalus* species are

confined to Sub-Saharan Africa (Norval *et al.,* 1992; Walker *et al.,* 2003). Distribution of *R. appendiculatus* in these regions is characterized by short dry seasons and high ambient humidity throughout most of the year. The tick distribution is correlated with wet humid climatic conditions where host availability and vegetation are favourable (Norval and Perry, 1990; Walker *et al.,* 2003 Bazarusanga *et al.,* 2007). In parts of Burundi, Kenya, Tanzania and Uganda, it cycles through two or more generations per year (Yeoman, 1966; McCulloch *et al.,* 1968; Branagan, 1973; Newson, 1978; Kaiser *et al.,* 1982; Kaiser *et al.,* 1988; Tatchell & Easton, 1986). The potential distribution of *R. appendiculatus* in Africa is based on eco-climatic factors and is influenced by high temperatures and humidity (Olwoch *et al.,* 2008). Using a computer program, CLIMEX, areas of predicted climatic suitability for the tick were shown to extend through most of sub-Saharan Africa. The greatest suitability is seen in coastal zones of Cameroon and Equatorial Guinea, the periphery of the Zaire basin, and the coastal strip of Natal and Transkei in South Africa. Intermediate and low suitability is seen through much of central and southern Africa (Perry *et al.,* 1990). The most unsuitable areas are found in the deserts of the Horn of Africa and southern Africa (Figure 3). Recorded occurrence of *R. appendiculatus* in Kenya has been compiled by Lessard *et al.,* 1990 (Figure 4). Many arthropods including ticks synchronize their lifecycle by means of dormancy (Tauber *et al.,* 1986). In *R. appendiculatus*, this dormancy is expressed as a behavioural diapause which is thought to occur in response to photoperiod; newly molted ticks, exposed to short day length, postpone questing behavior until the next rainy season (Berkvens *et al.,* 1995; Rechav, 1981). The resulting phenology regulated by the occurrence of diapause, the number of generations per annum and whether or not the different instars feed simultaneously, determines the epidemiological character of East Coast Fever in any given region (Madder *et al.*, 1999). In areas with high tick burden throughout the year, endemic stability (a state where majority of calves become infected and immune by six months of age and little or no clinical disease occurs) can

develop (Norval *et al.,* 1992). Studies by Madder *et al.,* 1999 have indicated that considerable genetic variation for diapause induction exists in South African *R. appendiculatus* ticks as phenotypic while it is non-existent in East African ticks.

1.2.3 Life cycle of *R. appendiculatus*

R. appendiculatus is a 3-host tick (Sonenshine, 1993; Walker *et al.,* 2003). The engorged female adult drops off the host and starts ovipositing 5 to 6 days later. During the next 14 days, 3000-5000 eggs are laid. The eggs hatch into larvae after 20-90 days and the larvae climb onto vegetation in search of suitable hosts. Once a host is located, the larvae attach and feed to engorgement in 3-7 days. The engorged larvae then drop off the host and molt into nymphs in 18 to 21 days. The nymphs climb onto vegetation and attach to an available host and engorge within 3 to 7 days. Adults develop after approximately 21 days. Up to three generations of tick vector can occur per year in favourable areas of Eastern Africa such as the Lake Victoria Basin but only one generation occurs in Southern and Central Africa as a result of the behavioral diapause, controlled by photo period, in the adult tick. This behavioral diapauses result in a strict seasonal occurrence of the different tick stages on cattle. Behavioral diapause allows the tick to survive during the long dry seasons occurring in the Southern parts of Africa (Norval *et al.,* 1992; Walker *et al.,* 2003).

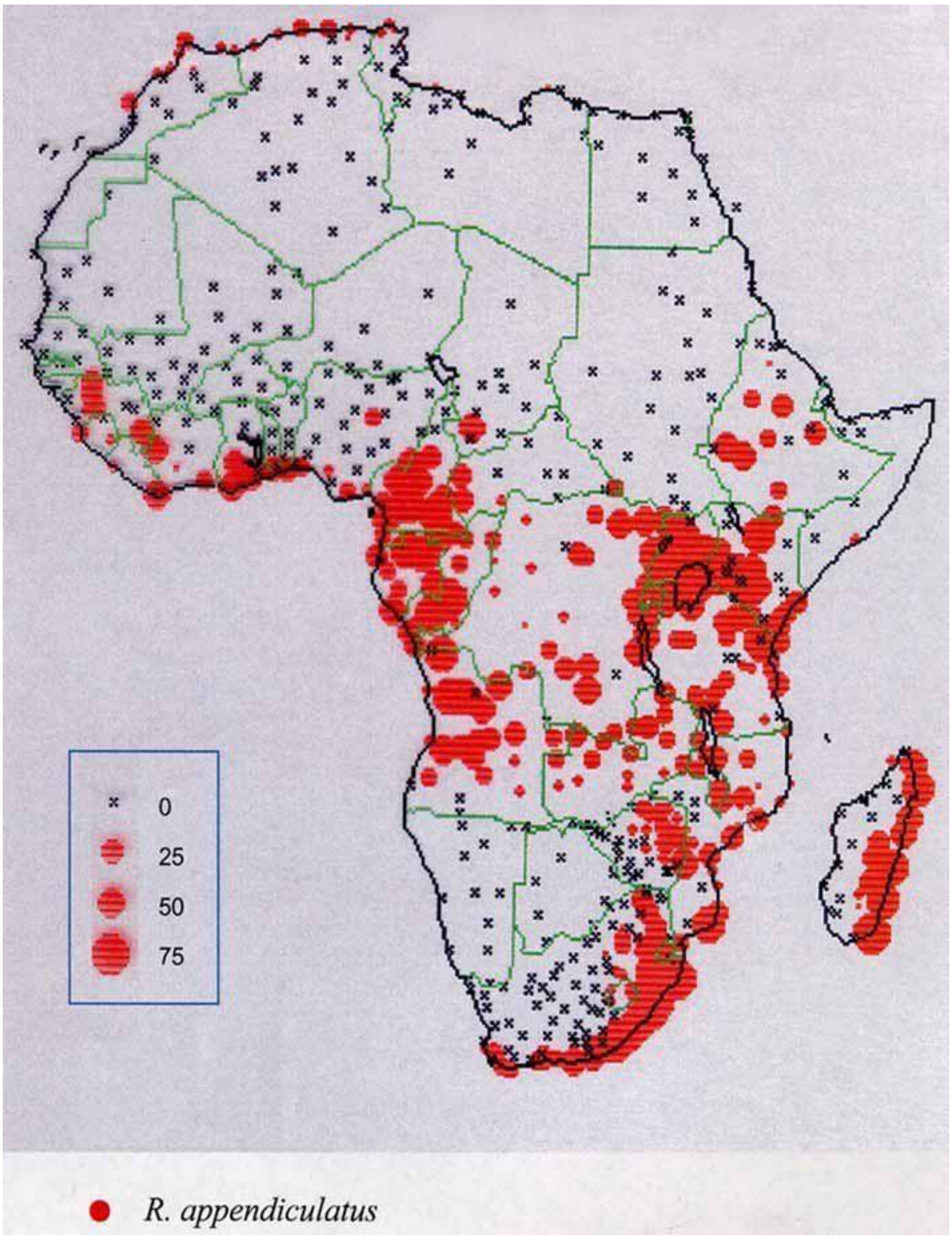


Figure 3: Potential climatic suitability for the distribution of *R. appendiculatus* in Africa. The map is adopted from Kibuka-Sebitosi *et al.*, (2006). Numbers indicate the ecoclimatic index on a scale of 0–75 where 0 is not suitable and 75 is very suitable, for *R. appendiculatus.*

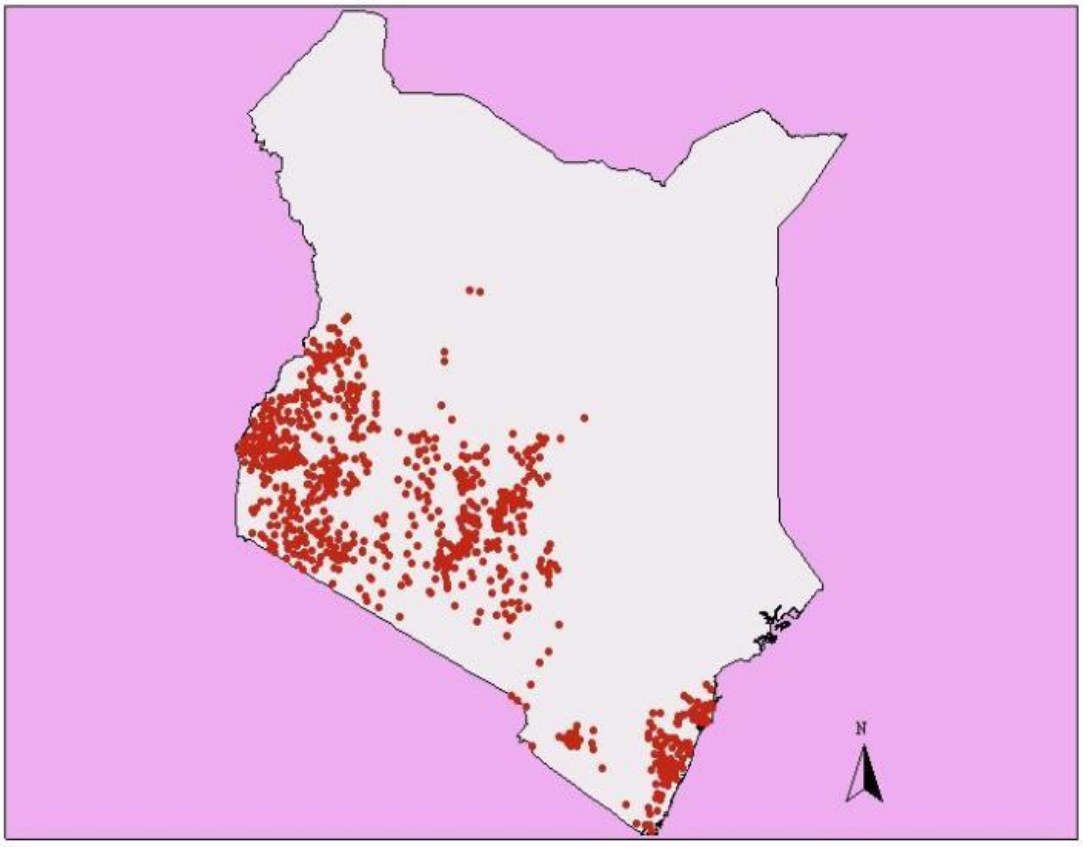


Figure 4: Distribution of the *R. appendiculatus* ticks in Kenya.
The map shows data based primarily on tick collections done by Walker, (1974).

1.2.4 Diseases transmitted by *R. appendiculatus*

The history of tick-borne diseases of cattle in Africa dates to the colonial settlement of the interior of eastern and southern Africa in early 1900s when a deadly disease, later determined to be East Coast Fever (ECF), was diagnosed in 1902 in cattle at several locations south of the Zambezi River. The origin of the disease was determined to be importations of cattle from Dar-es-Salaam in Tanzania after the cattle population of much of southern Africa had been destroyed by an epidemic of rinderpest (Lawrence, 1992). It is thought that ECF must have been present in E. Africa for many generations in areas populated by indigenous cattle, particularly in the Lake Victoria Basin and along the costal strip of eastern Africa (Perry, 1992) before it was diagnosed in early 1900s. *R.*

appendiculatus is the principal vector of *T. parva*, which causes East Coast fever in cattle, in East and Southern Africa. *Theileria lawrencei, T. taurotragi, T. velifera* and *T. mutans* have been isolated from African buffalo and are transmissible to cattle. Of these, *T. lawrencei* is the only species highly pathogenic for cattle. *T. taurotragi* a parasite of elands causes benign bovine theilerosis. Five species of *Theileria* have so far been described in Kenya (Table 2 below). *T. parva* cattle-derived and maintained between cattle, manifests itself in the form of classical ECF while *T. parva* buffalo-derived when transmitted to cattle causes Corridor Disease. *T. mutans*, *T. velifera* and *T. taurotragi* infections usually cause mild transient fever and anaemia and are not often reported in the field. *T. buffeli* was isolated in Kenya and also causes mild transient fever in cattle (Ngumi *et al.*, 1994).

Table 2: Known species of *Theileria* recognised in Kenyan cattle

Theileria species	**Disease**	**Tick vector**	**Animal host**
T. parva	East Coast fever	*R. appendiculatus*	Cattle, African buffalo *(Syncerus caffer)*, waterbuck *(Kobus spp.)*
T. taurotragi	Benign theileriosis (turning sickness), Tzaneen disease	*R. appendiculatus & R. pulchellus*	Cattle, eland *(Taurotragi oryx)*
T. mutans	Benign/mild theileriosis, African bont tick theileriosis	*Amblyomma spp*	Cattle, African buffalo
T. velifera	Benign theileriosis	*Amblyomma spp*	Cattle, African buffalo
T. buffeli	Benign theileriosis	*Haemaphysalis* spp	Cattle

T. parva is the causative agent of classical ECF is (Young *et al.*, 1973, Young *et al.*, 1981, Young & Purnell, 1973). The first clinical sign of ECF in cattle appears 7 to 15 days after attachment of infected ticks. This is seen as a swelling of the draining lymph node, usually the parotid, for the ear is the preferred feeding site of the vector. This is followed by a generalized lymphadenopathy in which superficial subcutaneous lymph nodes such as the parotid, prescapular, and prefemoral lymph nodes, can easily be seen and palpated. Death usually occurs 18 to 30 days after infestation of susceptible cattle by infected ticks. The severity and time course of the disease depend on, among other factors, the magnitude of the infected tick challenge, for ECF is a dose- dependent disease, and on the strain of parasites. Morbidity and mortality depend on, among other factors, the magnitude of the infected tick challenge and susceptibility of the host and strain of parasite. East Coast fever in susceptible cattle, which are not indigenous to the enzootic area, is very severe with a mortality approaching 100 percent. Some stocks of parasites cause a chronic wasting disease. In recovered cattle, chronic disease problems can occur that result in stunted growth in calves and lack of productivity in adult cattle. However, this syndrome tends to be in the minority of recovered clinical cases; in a majority of cases, asymptomatic carriers have no effect on their productivity (Moll *et al.*, 1986).

R. appendiculatus ticks have also been shown to transmit a virus of the species "*Thogoto virus*". Thogoto virus, a single stranded RNA virus of the family *Orthomyxoviridae* and genus *Thogotovirus*. The Thogoto viruses have been isolated from cattle, camels and humans and are known to infect a number of other species including goats where their antibodies have been detected. The virus can replicate in mosquitos and ticks. The virus can infect humans causing a febrile illness and encephalitis. In sheep, Thogoto viruses cause febrile infections and abortions. In a study by Jones and others (Jones *et al.*, 1990) the vector efficiency of *R. appendiculatus* and

Amblyomma variegatum, to transmit Thogoto virus was assessed using guinea pigs models which do not develop a detectable viraemia. Infected *R. appendiculatus* ticks transmitted the virus to uninfected ticks when co-feeding on uninfected guinea-pigs. Adult *R. appendiculatus* were shown to be more efficient than *A. variegatum* in mediating non-viraemic transmission of the Thogoto virus. *Anaplasma bovis* which causes ehrlichiosis and *Rickettsia conori* which causes tick typhus in humans are also transmitted by *R. appendiculatus*. The tick also transmits Nairobi sheep disease virus. Massive *R. appendiculatus* infestations lead to immuno-suppression and re-emergence of other diseases and may lead to anaemia due to blood ingested by ticks. It is shown that each female engorging fully results in 4g loss in potential for cattle growth (Walker *et al.*, 2003).

1.2.5 Life cycle of *T. parva*

Transmission of *T. parva* is strictly trans-stadial as the parasite is transmitted only by the nymphal and adult stages after acquiring infections as feeding larvae or nymphs. The piroplasm stage of *T. parva*, which resides in bovine erythrocytes, is infective to ticks. A zygote formed after gamete fusion enters tick gut epithelial cells, and differentiates into a kinete that is subsequently released into the haemolymph. The kinete invades the *e* cell of type III acini of the salivary glands and undergoes limited differentiation. Completion of sporogony only occurs after the tick molt into the next instar and is primarily stimulated by blood feeding, with production of mature mammalian infective sporozoites peaking at days 4–5 after initiation of feeding. The life cycle of *T. parva* is summarized in Figure 5. *T. parva* is acquired by immature ticks during blood feeding, and transmission of the parasite is strictly transstadial as the parasite is transmitted only by the nymphal and adult tick stages after acquiring infections during feeding as larvae and nymphs, respectively. The initiation of feeding of infected *R. appendiculatus* adults induces the rapid development of *T. parva* sporoblasts within the salivary gland acini leading to the production of numerous sporozoites

which are inoculated into the mammalian host initiating infection. The pattern of development, host cell specificity and emission of *T. parva* sporozoites within the salivary glands of heavily infected, 4-day fed adult *R. appendiculatus* ticks was examined by Shaw & Young (1995). Infected acini were randomly distributed throughout the salivary gland and sporozoite development within each gland was not synchronized. Wide variation in the rate of parasite development, which correlated with the secretory activity of the individual acinus, was also observed in all glands examined. *T. parva* have been shown to develop primarily in Type III *'e'* cells. However, in heavily infected salivary glands sporogony and the emission of mature sporozoites also occur in *'c'* cells of Type II acini (Shaw & Young, 1995).

Transmission of the *T. parva* parasite by both laboratory and field ixodid tick populations and between different instars of ticks has been studied extensively (Young *et al.,* 1983, Ochanda *et al.*, 1996, Piesman, 1991, Shih *et al.,* 1995). Early studies on *T. parva* transmission reported that experimental batches of adult *R. appendiculatus* ticks, used during the course of several years, showed considerable variation which could not always be correlated with the piroplasm parasitaemia in the cattle on which the ticks engorged as nymphs (Purnell, 1974). Transmission rate of *T. parva* from infected cattle to ticks and vice versa has been shown to be relatively high (Konnai *et al.,* 2006). Ochanda and others (1996) demonstrated that the abundance of *T. parva* infection differed between different instars of ticks when fed on cattle undergoing acute infection. When nymphs and adult ticks were compared in terms of prevalence of several *T. parva* strains namely; *T. parva* muguga, *T. parva* kiambu, marikebuni and zimbabweni *T. parva* Boleni, adult ticks were shown to consistently have higher prevalence of *T. parva* infection than nymphs feeding on the same cattle. Several factors could influence the ability of the vector tick to acquire and transmit *T. parva* from infected-ticks to cattle. High levels of *T. parva* prevalence, abundance and intensity

were associated with high levels of parasitaemia in ticks (Young *et al*., 1996). In a study by Konnai *et al.*, (2006), 70.8 % of ticks that acquired *T. parva* from an infected cow with low-level parasitemia successfully transmitted the parasite to naive cattle.

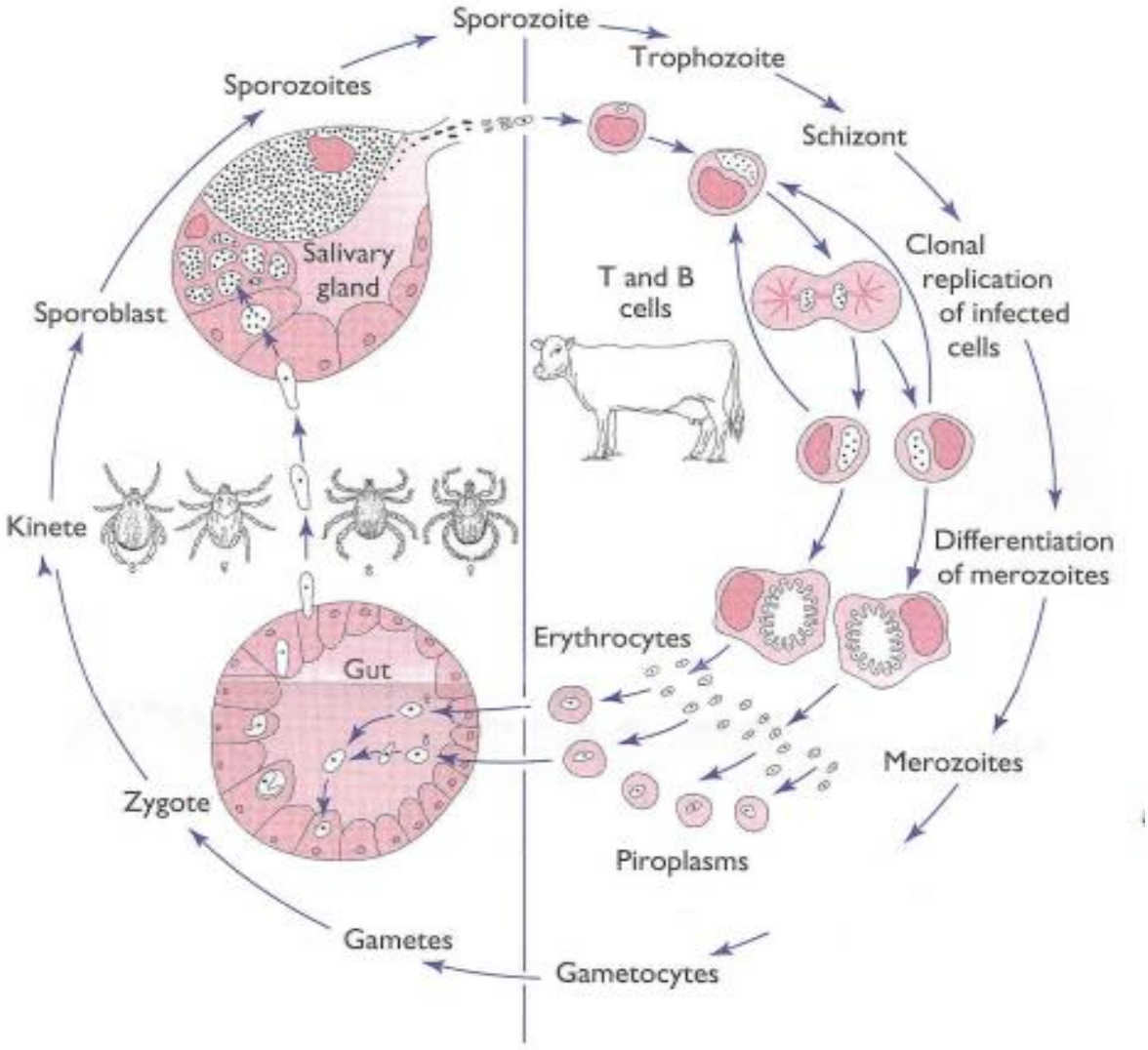


Figure 5: Life cycle of *T. parva* in cattle and *R. appendiculatus* tick vector.
Transmission in ticks is trans-stadial where larval or nymphal instars of the tick acquire infections from a blood-meal, which is then transmitted to a new cattle host, after moulting by nymphs or adults, respectively.

Some of the factors affecting the transmission rate of *T. parva* between ticks and cattle, both under experimental and field conditions transmission were investigated using a quantitative real-time PCR assay targeting the *T. parva* single-copy gene encoding the 104 kDa antigen (p104) (Konnai *et al.,* 2007). The study showed that *T. parva* transmission from tick to animal host begins 72 hrs after tick attachment and the duration of tick attachment was associated with increased risk of *T. parva* transmission. The *T. parva* infection rate in adult *R. appendiculatus* ticks feeding on low parasitaemic, *T. parva* carrier cattle has been observed to be lower than in ticks feeding on calves with acute, high parasitaemic infections (Marcotty *et al*., 2002; Ogden *et al*., 2003). These findings suggest that the level of parasitaemia in the infected animals contribute to the efficiency of acquisition of *T.parva* infection by ticks (Konnai *et al.,* 2006). In this study, *T. parva* transmission trials using *R. appendiculatus* as vector ticks indicated that the ticks had the ability to acquire *T. parva* infection from carrier animals with low levels of parasitaemia. In addition, the ticks that acquired *T. parva* from the calf during the low *T. parva* DNA load of infection successfully transmitted the parasite to naive calves. Although the number of infected adult or nymph ticks required to effectively transmit *T. parva* under natural conditions still remains unknown, this result indicated that one of the risks of transmission of *T. parva* from carrier animals with low levels of parasitaemia to naive calves was ticks having a high ability to acquire *T. parva*. *T. parva* carrier state has been observed in cattle in the field in endemic areas and the infectivity of different tick batches to susceptible cattle produced wide varied spectrum of theileriosis in an investigation on *T. parva* carrier state in cattle in commercial farms in Zimbabwe (Latif *et al.,* 2001).

At the same time differences in vector competence between different tick stocks has been observed (Ochanda *et al.,* 1998a). The competence of seven different stocks *of R. appendiculatus* and *R. zambeziensis* to transmit two *T. parva* stocks from Kenya and Zimbabwe in the laboratory was

investigated. The results showed that specific tick stocks transmitted certain parasites efficiently than others, with sporozoite infection increasing with time in days post-attachment. The effect was a combination of the tick and parasite genotype (Ochanda *et al.,* 1998b). Furthermore the heritability of susceptibility to *T. parva* susceptibility was shown to differ between off springs of individual ticks within families (Young *et al.,* 1995). Individual lab or field ticks were shown to vary greatly in their infection burdens of *T. parva*. The authors were able to select tick lines of high or low susceptibility for *T. parva* infection, suggesting that highly selected lines can be used to identify linked marker genes influencing such a trait (Young *et al.,* 1995).

Laboratory tick colonies are generally maintained by using limited numbers of breeding individuals and, as a result, are inevitably inbred. Extensive genotypic diversity in a recombining *T. parva* population from different infected laboratory *R. appendiculatus* populations has been reported (Katzer *et al.,* 2006), suggesting that different tick populations select particular genotypes of parasites. These observations suggest that different inbred tick populations may favor particular *T. parva* genotypes and repeated passage of the stock through individual tick lines is associated with increasing homogeneity with respect to the favored population (Katzer *et al.,* 2006). For the maintenance of stabilates for large-scale infection and treatment immunization, laboratory tick populations will show variation in the specificity of selection with various tick lines selecting for particular parasite genotypes (Radley *et al.,* 1975). However, variation among tick populations in the specificity of selection for *T. parva* parasites has major consequences for the population structure of the parasite in the field, since it identifies tick population heterogeneity as a major determinant of parasite diversity. In this regard, it is useful to consider the potential selection imposed by the tick (Katzer *et al.,* 2006).

1.2.6 Control of ECF by vaccination

Vaccination against ECF involves the simultaneous inoculation of a lethal dose of *T. parva* stabilates (cryopreserved sporozoites) and a long-acting oxytetracycline dose in an infection and treatment mode in several African countries, including Zambia (Uilenberg, 1999), parts of Kenya, Tanzania, Uganda and Malawi (McKeever *et al.,* 1994). After the development of the infection and treatment method (ITM) of immunisation in the mid-1970s, an immunising stabilate referred to as the "Muguga Cocktail" stabilate, was prepared from pooled ticks so that equal numbers of infected acini were present from each component stock consisting of sporozoites of three *T. parva* stocks: cattle-derived Muguga, cattle-derived Kiambu 5 and buffalo-derived Serengeti-transformed. The preparation of *T. parva* sporozoite stabilates has been described by Cunningham *et al* (1973). The main steps in the production of a live immunizing *T. parva* stabilite are summarized in Figure 6.

There is strong evidence that protection induced by the live vaccine is mediated by parasite-specific major histocompatibility complex (MHC) class I-restricted cytotoxic T lymphocytes (CTLs), which target the schizont-infected lymphoblast (McKeever *et al.,* 1994, Morrison *et al.*, 1987, 1989). Radley (1981) noted that a pool of these parasite stocks protected cattle against challenge with stocks from Kenya, Malawi, Tanzania and Uganda than the individual immunizing components (Radley *et al* 1975a; Radley *et al.,* 1975b). Inoculation of the trivalent stabilate and simultaneous administration of an appropriate tetracycline allows these three stocks to establish in the animal. A mild infection of *T. parva* occurs in the immunised cattle, which stimulates cell-mediated immunity to schizont-infected cells. This immunity can last up to 36 months in the absence of further challenge (Burridge *et al.,* 1972) but may be expected to become life long through boosting from the continuous challenge expected in areas where ITM is necessary. This vaccine has been used

successfully in East and Southern Africa to control ECF (Melewas *et al.*, 1999, Nsubuga-Mutaka, 1999). The live vaccine has successfully been used to control ECF in many areas where it is in use.

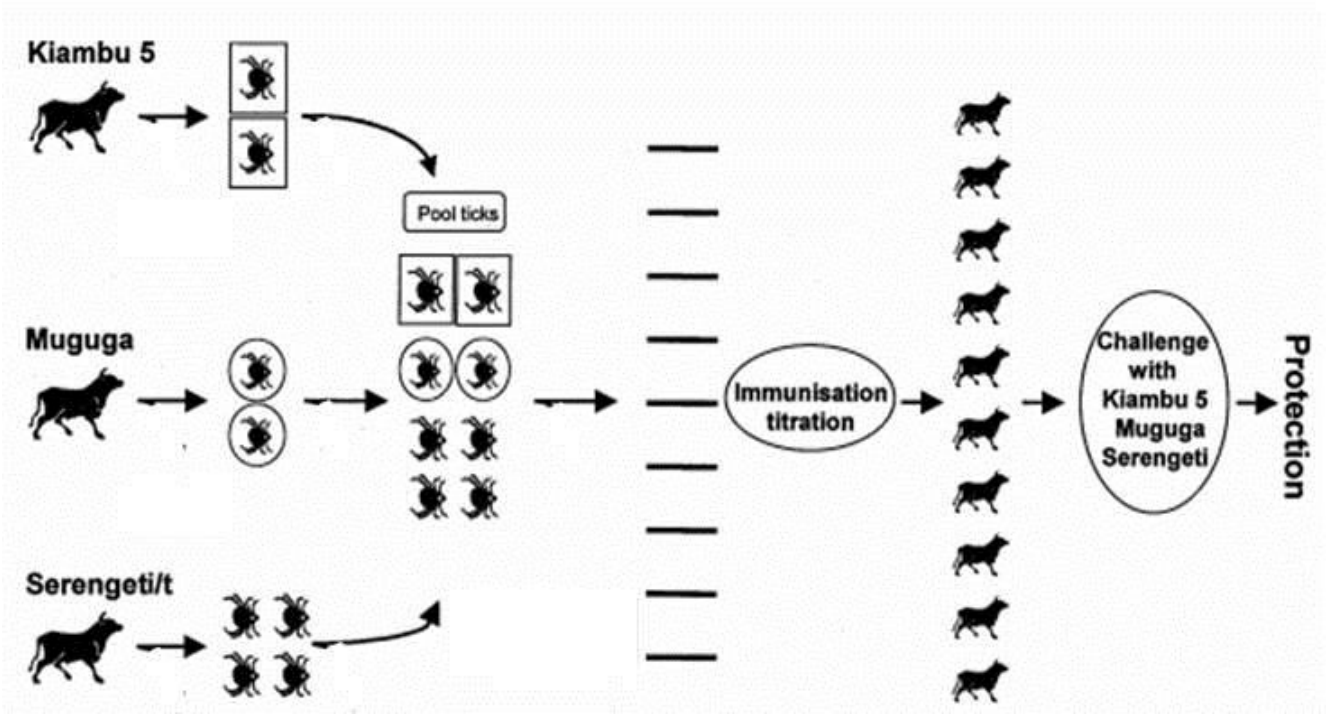


Figure 6: A schematic diagram showing the main steps in preparation of a composite vaccine stabilate containing components of the Muguga cocktail.
Three *T. parva* stocks are used: cattle-derived Muguga, cattle-derived Kiambu 5 and buffalo-derived Serengeti-transformed. Inoculation of the trivalent stabilate and simultaneous administration of an appropriate tetracycline allows these three stocks to establish in the animal. Adopted from Morzaria & Williamson (eds) (1999).

However, a few problems are associated with ITM (Mbogo, 1999; Di Giullio *et al.,* 2009) which include a small proportion (3 %) of the immunised animals developing the disease. This is because the MHC class I restricted cytotoxicity offered by the vaccine does not take effect until the schizont parasitosis is established (Morrison *et al.,* 1987). These animals are commonly referred to as reactors and are identified through follow-up monitoring between days 10 and 14 after immunisation. Another concern with vaccination using live sporozoites is the possibility of introduction of vaccine strains in areas where they were previously absent. *T. parva* vaccine parasites may establish a carrier state in immunised cattle as indicated by studies by Kariuki *et al.,* 1995 and Young *et al.,* 1986. Since many different strains exist, immunisation can result in the introduction of new strains of parasites into a new area which has been demonstrated by Oura *et al* (2004, 2007). However, Theilerial parasites isolated from carrier cattle after immunization with *T. parva* by the infection and treatment method indicated that the carrier state is not likely to produce new antigenic strains which would be dangerous to immunized cattle (Maritim *et al.,* 1989). Other major costs associated with this method of immunisation include those of the oxytetracycline, an expensive antibiotic, the need for a monitoring exercise to identify reactors, use of antitheilerial drugs to treat reactors and the possible loss of milk due to milk withdrawal after administering drugs to lactating cows.

Other studies have resulted in evaluation of *T. parva* antigens as possible vaccine candidates. Antisera from cattle in ECF-endemic areas are polyspecific and on immunoblots bind to several sporozoite molecules which include proteins of relative molecular weights (Mr) 105, 85, 67, 55, and 31, kilo-Daltons (kDa) (Musoke *et al.,* 1992, 1999). Candidate vaccine antigens were identified from this repertoire of sporozoite antigens by raising monoclonal antibodies (MAbs) that neutralized sporozoite infectivity (Musoke *et al* 1984; Dobbelaere *et al* 1984). The majority of the

MAbs bound to a sporozoite stage-specific 67 kDa molecule (p67) that is expressed on the surface membrane of the *T. parva* parasite. An *E. coli* recombinant fusion protein, NS1-p67, containing all 709 amino acid residues of p67 fused to the first 85 residues of NS1, a non-structural protein of influenza virus A, has been used in cattle immunisation experiments and a degree of protection against *T. parva* infection was demonstrated with recombinant p67 (Musoke *et al.,* 1992). A number of CTL target antigens have recently been identified. Graham and others (2006) have reported identification of five candidate vaccine antigens that are the targets of MHC class I-restricted $CD8^+$ CTL from immune cattle. $CD8^+$ T cell responses to these antigens were boosted in *T. parva*-immune cattle resolving a challenge infection and, when used to immunize naive cattle, induced CTL responses that significantly correlated with survival from a lethal parasite challenge. Cytotoxic T lymphocytes (CTLs) obtained from genetically diverse Zebu cattle immunized with a cocktail of *T. parva* stocks have been shown to recognize the CTL-target antigens (Akoolo *et al.,* 2008). *T. parva* Muguga specific polyclonal $CD8^{(+)}$ CTL lines were generated and confirmed to specifically lyse autologous infected cells. However, it has been shown that several parasite antigens are recognized by the CT lymphocytes and thus immunity is partially strain specific (Pelle *et al.,* 2011).

1.2.7 Biological and morphological differences in *R. appendiculatus*

Biological differences occur between populations of *R. appendiculatus* in different geographical regions (Norval *et al.,* 1992, Shaw & Young, 1994). Some populations of *R. appendiculatus* in Southern Africa are larger than those in Eastern Africa, a feature which contributes to a longer development interval of the tick instars in the former region (Norval *et al.,* 1992, Shaw & Young, 1994). Strains of tick may differ in their ability to transmit organisms that cause disease, for example *R. appendiculatus* and *R. Zambeziensis* have different levels of vector competence for *T. parva* (Ochanda *et al.,* 1998a) and strains of ticks differ in their responses to acaricides (George *et*

al., 2004). Norval *et al*. (1991) and Kubasu (1992) reported that brown ear ticks from geographically isolated areas differ in their *T. parva* vectorial capabilities by assessing the parasite load in the salivary glands. Genetic variations, isoenzyme polymorphism and protein variations have been shown to occur between tick species and populations (Bull *et al.*, 1984, Davey & Hilburn, 1991; Gomes & Wouters, 1991; Healy, 1979, Hilburn & Sattler, 1986).

Histological differences in salivary glands of *R. appendiculatus* ticks feeding on rabbits and cattle of varying tick resistance was demonstrated by Walker and Fletcher (1990). In this study, adult *R. appendiculatus* ticks were fed as three sequential infestations on both rabbits and cattle. The feedings at first infestation on naive hosts were optimum for the ticks, whereas at third infestation the hosts were resistant, as expressed by reduced tick feeding performance. Ticks from naive and resistant hosts were examined for histological differences of salivary glands. In ticks fed on resistant rabbits there was a large increase in the synthesis of glycoprotein secretory granules in the c1 cells compared with ticks fed on naive rabbits. In ticks fed on naive and resistant cattle, the activity of the c1 cells was less than in ticks fed on naive and resistant rabbits. This suggests that the salivary glands are able to respond selectively to conditions at the feeding site**.** Marked polymorphism of salivary gland proteins has been shown among individual ticks of the same species in *Amblyomma variegatum, Ixodes ricinus* and *R. appendiculatus* (Wang *et al.*, 1999). In this study, individual ticks of the same species showed variation in biochemical processes such as salivary gland protein profiles of the tick species studied. Potential intra-species diversity should be considered when pooled samples are used as is the case when preparing live *T. parva* vaccine from infected salivary glands of *R. appendiculatus* laboratory stocks.

Considerable genetic variation for diapause induction in ticks has been demonstrated. Southern and central African *R. appendiculatus* can undergo a behavioural diapause in the adult instar, which results only in one generation of ticks occurring per year in these areas (Short & Noval, 1981, Rechav, 1982, Pegram & Banda, 1990). Studies in the biology, ecology and behaviour of *R. appendiculatus* in Zambia have shown considerable geographic variation within and between populations (Berkvens *et al*., 1995; Chaka *et al*., 1999; Madder *et al*., 2002; Speybroeck *et al*., 2002, 2004). At different latitudes different seasonal activity patterns are observed. Near the equator, ticks usually feed throughout the year and numbers vary less. In southern Zambia there are marked annual cycles of abundance of each life-stage (larva, nymph and adult) punctuated by near total absence during the dry season of each year (Speybroeck *et al*., 2002). During this dry season (June- November), adult ticks enter into a diapause that they terminate at the beginning of the rainy season (December). Within Zambia, the presence of a transition zone with respect to activity patterns of *R. appendiculatus* can be noticed: in southern Zambia, *R. appendiculatus* adults have a single peak of host-seeking activity (during the rainy season in February), whereas in eastern Zambia, a second generation of adult *R. appendiculatus* is possible (at the beginning of the dry season in May; Chaka *et al*., 1999). It is also in Zambia where a difference in body size between southern and eastern *R. appendiculatus* (Speybroeck *et al*., 2002) has been reported. This is an indication of a link between the size and population dynamics of *R. appendiculatus* (Speybroeck *et al*., 2004).

Berkvens (1994) was first to recognize the importance of body size when attempting to understand *R. appendiculatus* activity patterns on the population level. Chiera *et al.*, (1985) reported lower survival rates for small *R. appendiculatus* individuals under unfavorable conditions. Madder *et al*., (1996) demonstrated low heritability of body size under laboratory conditions. Chaka *et al*., (1999)

showed that in eastern Zambia, *R. appendiculatus* adults were smallest during the dry season, but specimens became larger as the rainy season progressed. Second generation (non-diapausing) adults were on average smaller than those of the first generation. Eastern African ticks do not show diapausing behavior and they can undergo two or more generations per year (Kaiser *et al.,* 1982). This difference in tick seasonality contributes to the difference in the epidemiology of theileriosis between the two areas (Norval *et al.,* 1992). Instars of *R. appendiculatus* infesting hosts in southern and central Africa thus demonstrated a marked seasonality. In a study by Madder and others investigating behavioral diapause in ticks from different geoegraphical areas, newly-moulted adults of three *R. appendiculatus* stocks, originating from Kenya, Zambia and Zimbabwe, exhibited different phenotypic differences in behavioral diapause (Madder *et al.,* 1999). The authors suggest that the phenology of ticks in a given geographical region is determined by the occurrence of diapause, the number of generations per annum as well as the feeding behavior of different tick instars. These factors may then determine the epidemiological character of ECF in such a given region. In areas with high tick burdens throughout the year, endemic stability (a state whereby the majority of calves become infected and immune by six months of age and little or no clinical disease occurs (Norval *et al*., 1992) can develop. In areas with unimodal rainfall patterns, the disease challenge is more concentrated precluding development of full endemicity.

Characterisation of *R. appendiculatus* by use of protein profiles in Kenya has been reported by Baliraine *et al.,* 2000. In this study the protein profiles of two distinct populations of the brown ear tick (*R. appendiculatus*) in Kenya, which differed in their *T. parva* vector competence were compared by two-dimensional gel electrophoresis. A majority of the proteins were homologous to both populations. However, a few proteins were found to be population-specific. The findings of this study indicate the presence of population-specific proteins suggesting genetic differences

between the populations. As with other vector-borne diseases, an understanding of the vector phenology and its determining parameters and mechanisms is essential for the development of optimal control strategies (Madder *et al.,* 1999).

1.3 Molecular markers for studying tick phylogenetics

Variation at the DNA sequence level underlies the observed phenotypic differences between species (Chakravarti, 1999). These variations are referred to as polymorphisms (Storz, 2005). Some non-coding DNA variations also occur in the genome and are useful for molecular characterisation of species. Measuring genetic variation and patterns of population/species (intra-/inter-specific level) genetic diversity among biota helps us to obtain information on population relationships (Avise, 1994). Availability of a wide array of molecular genetic markers offers tools for quick detection and characterization of genetic variation. The correct identification of species is essential to the performance of ecological and evolutionary research (Graham *et al.,* 2004). Morphological identification of organisms has been used to accurately identify of many taxa (Gaston & O'Neill, 2004). However, for taxa that are closely related or not well studied, or for which distinguishing morphological characters have not been discerned, identification can be difficult. Accurate identification is especially problematic for closely related species of ticks and especially with different life stages of ticks. Walker *et al.,* (2000) and Warburton (1912) observed that *Rhipicephalus* species of ticks were extremely difficult to classify because of their high level of morphological intrageneric uniformity and intraspecific variability. Molecular markers may be able to circumvent some of the problems associated with morphological key based taxonomical methods (Lee, 2004; Dunn, 2003).

Molecular genetic markers can be divided into two classes: a) Biochemical markers which detect variation at the gene product level such as changes in proteins and amino acids b) molecular markers which detect variation at the DNA level such as nucleotide changes: deletion, duplication, inversion and/or insertion. These sites or loci in the genome at which DNA variations reside are referred to as genetic, molecular or DNA markers (Wright, 1995; Vignal *et al.*, 2002)

Biochemical markers detect variations as electrophoretic enzyme polymorphisms (isozymes). Isozymes are able to detect the diversity at the functional gene level and have simple inheritance, co-dominant expression, complete penetrance and no pleiotropic and epistatic interactions. Polymorphism in ticks has been studied by using isoenzymes. Genotypic variation in the enzyme α-Glycerophosphate dehydrogenase has been shown in Irish and Swedish *I. ricinus* ticks questing at different time of the day in different habitats (Healy *et al.*, 2004) while genotypic variation in the occurrence of malate dehydrogenase at different time intervals among nymphal stages of Danish *I. ricinus* tick populations has been reported (Jensen *et al.,* 2003). Two distinct populations of *R. appendiculatus* in Kenya which differed in their *T. parva* vector competence have been shown to occur (Baliraine *et al.,* 2000) suggesting the presence of population-specific proteins. Biochemical genetic markers are relatively inexpensive, rapid, and technically easy to apply as compared to DNA techniques. They show lower levels of diversity or polymorphism in a population as they assay relatively small numbers of genetic loci. They are unsuitable when a very high resolution of diversity is required. They may also be affected by the environmental factors such as age and animal part sampled. However, these markers may be the choice when the cost involved for the routine testing is lower than using molecular markers.

Molecular markers that detect variation at the nucleotide level can be classified into two major groups a) based on DNA-DNA hybridization (e.g. Restriction fragment length polymorphism (RFLP) and b) based on Polymerase Chain Reaction (PCR) amplification of ribosomal, mitochondrial and nuclear DNA fragments. Detection of RFLPs involves the fragmentation of genomic DNA by restriction enzymes that recognize specific DNA sequences/motifs (generally frequent cutters 4-10bp in length). The restricted DNA fragments are separated by gel electrophoresis and transferred onto a membrane by Southern blotting (Southern, 1979). Hybridization of the membrane to a labeled DNA probe (a piece of known DNA sequence) then determines the size of the fragments that are complementary to the probes. The polymorphism arises from sequence changes in the restriction sites as well as from the detection of insertion/deletion in the restriction fragments detected by the probe. Analysis of RFLP variation is an important tool in genome mapping, identification of disease resistance genes, paternity testing and genetic fingerprinting. PCR-based markers include nuclear ribosomal and mitochondrial protein coding genes such as nuclear ribosomal genes, mitochondrial ribosomal genes and mitochondrial protein coding genes and micro- and minisatellites.

1.3.1 Nuclear ribosomal genes

The non-coding region the second internal transcribed spacer region of the nuclear ribosomal gene cluster (ITS2) has been used to study genetic variation in ticks. The ITS2 cluster consists of three genes (18S rDNA, 5.8S rDNA and 28S rDNA) which are transcribed into RNA but not translated into protein. These RNA molecules form parts of the ribosome, a cellular structure involved in making proteins. The three rDNA genes are transcribed as a single transcript of RNA separated by two regions (ITS1 and ITS2) which are subsequently spliced out and serve no further purpose. Since they have no other function, these ITS sequences are under very little selection pressure and

can accumulate substitutions very quickly. This can be very useful for distinguishing between closely related species. By designing primers within the slowly evolving genes that span the rapidly evolving spacers, primers have been developed which work well in a wide variety of taxa. ITS1 and ITS2, 18S and 28S ribosomal genes have been used as genetic markers to distinguish species. A 274bp part of ITS2 sequence was used to show conspecificities among 6 closely related tick species which form the *R. sanguineus* complex. These species share a lot of phenotypic and morphological characteristics (Zahler *et al* 1997). Barker (1998) studied the entire ITS2 of rDNA in *B. microplus* and tick populations from Australia, Kenya, South Africa and Brazil and *R. appendiculatus* from Kenya, Zimbabwe and Zambia. Population specific primers were able to distinguish between populations of these species. ITS2 was also able to distinguish between *R. appendiculatus* and *R. zambeziensis* which are very closely related and very similar morphologically. A novel real-time PCR assay for speciation of medically important ticks using ITS2 has been described by Shone *et al.,* 2006. With this marker, four medically important ticks in the US; *I. scapularis, I. pacificus, Dermacentor variabilis* and *A. americium* were successfully differentiated. Mitochondrial 12S rDNA and ITS2 have been used to confirm the taxonomic status of *R. appendiculatus* and *R. zambeziensis* two closely related species of ticks from Eastern and Southern provinces of Zambia, whose morphological identification in the areas where they co-occur is difficult (Mtambo *et al.,* 2007c). The problems associated with internal transcribed spacers are those typical of rapidly evolving genes. Failure to amplify some species with universal ITS2 primers has been reported for mites (Yli-Mattila *et al*., 2000, Fenton *et al*., 1997). Heterogeneity of ITS2 sequences within individuals has been reported in ticks (Rich *et al*., 1997). In order to minimise these problems, internal transcribed spacers should be used only for studies of intraspecific variation or phylogenies of very closely related species. It may be necessary to clone PCR products and sequences should be treated as haplotypes rather than individuals. Phylogenetic relationships of 9 species of ticks were

inferred from nucleotide sequences of the D1 domain of the large subunit of rDNA 28S (Crampton *et al.,* 1996). Using 18S sequences of 13 hard ticks, 5 soft ticks and 2 mesostigmatid mites of the genus *Aponomma,* the family *Amblyomminae* was found to be paraphyletic (Dobson & Barker, 1999). The phylogenetic relationships among six European hard-tick genera based on the 18S rDNA gene revealed low genetic variation between more closely related species (Mangold *et al.*, 1998). These findings suggested that the 18S rRNA gene is not a suitable tool for studies of closely related genera of ticks but it seems to be a good marker for taxa grouping purposes. Ribosomal genes are most likely to be useful at the population level where highly variable sites have not yet experienced multiple substitutions and at deep levels of divergence where the more conserved sites which are free of homoplasy, supply useful phylogenetic information (Simon *et al.*, 2004).

1.3.2 Mitochondrial ribosomal genes

Mitochondrial ribosomal genes have been widely used in molecular systematics. Their high copy number makes them much easier to work with than single-copy nuclear genes and their strictly maternal inheritance has been particularly useful at the intraspecific level. Mitochondrial genes fall into two categories; ribosomal genes and protein-coding genes. There are two mitochondrial ribosomal genes; 12S and 16S rDNA. Norris *et al*. (1999) sequenced domain III of the 12S rDNA gene for 51 species of ticks and compared this with previously published (Black & Piesman, 1994) and newly sequenced 16S sequences. 12S and 16S gave poorly resolved trees and Norris *et al*. (1999) put this lack of resolution down to the high AT content of the mitochondrial genome causing a large amount of homoplasy and suggest that these genes may be more suited to resolving relationships at the intraspecific level or among closely related taxa. However, a portion of the mitochondrial 12S rDNA sequence has been used to resolve relationships among recently diverged ticks within the rhipicephaline genera (Murrell *et al.,* 1999, Norris *et al.,* 1999). Based on the 12S

rDNA ticks of the genus *Boophilus* clustered closely with species of the *R. evertsi* and *R. pravus* (Murrell *et al.,* 1999) and with *R. evertsi* in another study by Beati & Keirans (2001).

1.3.3 Mitochondrial protein coding genes

Cytochrome c oxidase subunit I (COI) is one of the most conserved mitochondrial protein coding genes. COI has been shown to be a suitable marker for global bioidentification system for animals (DNA barcoding) (Hebert *et al.,* 2002). COI does have two important advantages. First, the universal primers for this gene are very robust, enabling recovery of its 5′-end from representatives of most, if not all, animal phyla (Folmer *et al*., 1994; Zhang & Hewitt 1997). Second, COI appears to possess a greater range of phylogenetic signal than any other mitochondrial gene (Knowlton & Weigt, 1998). The evolution of this gene is rapid enough to allow the discrimination of not only closely allied species, but also phylogeographic groups within a single species (Cox & Hebert, 2001; Wares & Cunningham, 2001). The usefulness of COI as a phylogenetic marker over taxonomical or morphological means of species identification was demonstrated by Hebert *et al.*, (2004) who studied COI sequences of butterflies in Costa Rica region. DNA barcoding of butterflies from these region as well as museum specimens together with morphological data showed that the butterfly *Astraptes fulgerator* is a complex of 10 species and not the seven species previously thought. COI is more likely to provide deeper phylogenetic insights than alternatives such as cytochrome b (Simmons & Weller, 2001) because changes in its amino-acid sequence occur more slowly (Lynch & Jarrell, 1993). The phylogeny of 21 species and subspecies of ticks from the subfamilies *Rhipicephalinae* and *Hyalomminae* inferred using COI and 12S mitochondrial ribosomal RNA (rRNA) showed the genus *Rhipicephalus* to be paraphyletic with a subgenus *Rhipicephalus Boophilus* (Murrell *et al.,* 2000). COI has been used to show further evidence for geographical differentiation in *R. appendiculatus* vector ticks from eastern and southern Provinces

of Zambia (Mtambo *et al.,* 2000a, 2007b) where several closely related haplotypes existed in eastern province of Zambia while a single haplotype was reported in Southern Province. Kain *et al*. (1999) have used a 355bp fragment of the cytochrome oxidase III gene to investigate intraspecific variation in *I. pacificus* and found widespread haplotype sharing between populations. The cytochrome b gene has been used extensively in vertebrates because of the availability of highly conserved primers that work well to amplify a wide variety of species (Kocher *et al.,* 1989, Irwin *et al.,* 1991).

1.3.4 Other PCR-based DNA markers

Several different types of PCR-based DNA markers are currently available for genetic analysis and new higher resolution marker types are being developed. Markers differ from each other in the initial effort and costs of developing the system, running costs, ease of use, level of polymorphisms, dominance, and number of loci analysed per assay, reproducibility and distribution on the chromosomes. The choice of the best marker system to use depends on whether it will be used in evolutionary or population diversity studies or in genetic mapping. Locus specific genetic markers are powerful tools for investigating the population genetics and epidemiology of parasites. For both of these purposes, the most effective approach is to use multiple markers that are selectively neutral. PCR-based markers are less technically demanding than other DNA-based markers and require only a small amount of DNA. In addition, PCR provides flexibility in detecting genetic variation as a variety of primers can be designed and used to reveal particular types of polymorphism. Examples of other DNA PCR-based markers are randomly amplified polymorphic DNA (RAPD), amplified fragment length polymorphism (AFLP), variable number tandem repeats such (Simple Sequence Repeats (SSRs) which are referred to as microsatellites and minisatellites.

1.3.4.1 Random amplification of polymorphic DNA (RAPD) and amplified fragment length polymorphism (AFLP)

During random amplification of polymorphic DNA, primers of short length (~10 nucleotides) are used to amplify random locations in/across the genome (Williams *et al*., 1990). Due to their short length, they have the possibility of annealing at a number of locations in the genome. The number of PCR amplification products is directly related to the number and orientation of the sequences that are complementary to the primer in the genome. Lan *et al.*, (1996) successfully amplified fragments of *B. microplus* DNA by RAPD-PCR using a 10bp primer and suggested that random amplified polymorphic DNA may be used to detect DNA polymorphism in this species. RAPDs require no prior knowledge of sequence information for primer designing. It is simple, fast, relatively cheap and widely used for population diversity studies, construction of genetic maps, tagging desirable traits for marker assisted selection, etc. However, reproducibility can be affected by low stringency PCR amplifications, the source of the tissue used and the protocol used for DNA extraction. It being a dominant marker, the recessive allele expression may not be found.

AFLP combines restriction digestion and PCR amplification. Restriction fragments have adapters ligated to their ends allowing PCR amplification from primers derived from the adapters (Vos *et al*., 1995). Polymorphisms arise from sequence changes in the restriction sites or the selective bases used by the primers. AFLP is generally a dominant molecular marker and requires no prior sequence information. It is highly sensitive and reproducible. It is widely used in the identification of genetic variation in populations, hybrids, closely related species, criminal and paternity tests, to determine linkage studies, creating genetic maps for QTL analysis, etc. However, it involves many steps, is labor intensive and time consuming. Problems may also arise due to homoplasy (the co-migration of non-related PCR products).

1.3.4.2 Variable number tandem repeat (VNTR) markers

Multiple copies of simple repeated DNA sequences arranged in arrays of differing repeat unit size and copy number are an integral component of eukaryotic genomes and may comprise between 50-90% of total DNA in vertebrate and plant genomes (Armour *et al.,* 1999, Hancock, 1999). Repetitive DNA may be classified as either interspersed or tandemly repeated. In case of interspersed repeats, the repeated DNA motifs occur at multiple sites throughout the genome. Tandem repeats on the other hand consist of arrays of two to several thousand basic motifs which are arranged in a head-to-tail fashion. Although this kind of organization is also exhibited by protein coding genes (e.g. the transcription units for histone and ribosomal RNA) and some families unique to protozoa, most tandem repeats are non-coding DNA. A common feature of tandemly arranged repetitive sequences is their degree of polymorphism, mostly due to variation in the number of copies of the tandem repeat motif but sometimes also to the internal structure of individual repeats, particularly in the case of longer repeated units. The frequent variation in the copy number has led to these sequences being designated variable number tandem repeats (VNTRs) (Nakumura *et al.*, 1987).

1.3.4.2.1 Micro- and minisatellites markers

SSRs, also known as microsatellites, are a group of tandem repeated sequences of mono-, di-, tri-,tetra-, penta-, or hexa-nucleotide units such as (A)10, (GA)8, (CAC)6, (GATA)4 or (GATAG)4. Minisatellites are blocks of repeated core elements consisting of short motifs (usually 10-30 bp) in tandem arrays up to ~30 kb or more in length. In a range of organisms that have been studied, micro- and minisatellite sequences/markers have been used for genotyping. They have the advantage of being widely dispersed and often selectively neutral, are also frequently polymorphic and can be directly typed by PCR amplification. While some satellite sequences, such as those in

intergenic regions are likely to be selectively neutral, others are probably under selective pressure, particularly those exposed to the immune system. In addition to their neutrality, non-coding genetic markers have other useful properties including their generally high levels of polymorphism, distribution throughout the genome, and ease of assaying with high reproducibility using PCR amplification. Oligonucleotide primers are synthesized in sequences flanking these repeat motifs, and can be used in PCR reactions to amplify loci from genomic DNA (Holmes *et al.,* 1993). A unique microsatellite-containing locus is usually a few hundred base pairs long and is usually characterized by PCR amplification followed by resolution of the alleles by electrophoresis on high resolution denaturing acrylamide gels (Litt & Luty, 1989; Weber & May, 1989).

Microsatellites have been detected in the genomes of many organisms and many are polymorphic due to variation in the number of copies of the repeated motif present at the locus (Hamada *et al.*, 1982; Tauz & Renez, 1984; Gemayel, 2010). Microsatellites are highly informative and have become the most widely used genetic markers for linkage mapping projects, pedigree, forensic and population assessments because they exhibit a high degree of variability (highly polymorphic and therefore poly-allelic), are abundant throughout the genome, are relatively easy to genotype and yield reliable results (Weber, 1990; Bruford & Wayne, 1993). Isolation of useful genomic micro- and minisatellite loci can be a time consuming, laborious and expensive process. However, the advent of the genomics age has resulted in the production of vast amounts of publicly available DNAsequence data, including large collections of expressed sequence tags (ESTs) from a variety of different taxa (Ellis & Burke, 2007). Public databases containing a gene index of sequences expressed in the salivary gland of *R. appendiculatus* (Nene *et al*., 2004) and *A. variegatum* (Nene, *et al.,* 2002) are now available. Recent research has revealed that ESTs are a potentially rich source of micro- and minisatellite markers (Ellis & Burke, 2007). Development of EST-based markers is

comparatively less laborious, fast and cheap. Compared to other molecular markers, microsatellite markers are uniquely characterized by their simplicity, abundance, ubiquity, variation, co-dominance and multi-alleles among genomes (Powell *et al*., 1996).

Micro- and minisatellites however have certain disadvantages. The major drawback is that they need to be isolated *de novo* from most species being examined for the first time. This is due to the fact they are usually found in non-coding regions where the nucleotide substitution rate is higher than in coding regions (Zane *et al.*, 2002). Genetic data based on microsatellite loci underestimate genetic differentiation because (F_{ST}), the index of genetic differentiation cannot take values greater than the level of homozygosity (Hedrick, 1999, 2005). Allele size homoplasy (i.e difficult to tell whether two alleles identical in state are also identical by descent) is a common problem with microsatellites because the infinite allele mutation model is not appropriate for microsatellites (Estoup *et al.,* 1995). There are reports that for certain groups of organisms' microsatellites are difficult to isolate (Beaumont & Bruford, 1999) and the tendency of *Taq* polymerase to add a 3´ terminal thymidine to both ends of PCR products and *Taq* polymerase generated slippage products in some mono- and di-nucleotide microsatellite loci, causing base shifts and sizing problems (Schotterer & Tauz, 1992). The task of micro- and minisatellite isolation for the first time can be quite involving in terms of effort and time because it traditionally consists of screening genomic libraries with appropriate probes (Rassmann *et al.,* 1991). However, recent advances have simplified the isolation of microsatellites through enrichment with a hybridization capture protocol (Hamilton *et al.* 1999; Marcet *et al.* 2006) and through data mining from EST libraries. Despite these disadvantages, microsatellites remain widely used as powerful genetic markers because of their high variability and thus resolving power (Zane *et al.,* 2002). Over the last few years, they have become the method of choice for defining differences between individuals in many eukaryotic

genomes. Such markers have been used extensively in population genetic analyses, for example in *P. falciparum*, to define population structures, geographical sub-structuring and levels of inbreeding (Anderson *et al.,* 1999, 2000). In recent studies, microsatellites have been used to address questions concerning the genetic diversity and population structure of a wide range of parasitic organisms such as *Trypanosoma cruzi* (Oliveira *et al.,* 1998) and *T. parva* (Oura *et al.,* 2003). Eight polymorphic microsatellites loci in the economically important tick species *Boophilus microplus* have been isolated and characterized (Chigagure *et al.,* 2000).

Minisatellites ranging in size from a few hundred bp to several Kb can also be used as genetic markers. The total size of repeat units varies from 10-60 Kbp (Tourmente *et al.,* 1994). Because of their length polymorphism which results from variations in the number of repeats and the ability of some of these arrays to cross-hybridise with tens of other similar loci throughout the genome, these highly polymorphic sequences have contributed to fundamental insights into human genetics including providing polymorphic, multi-allelic markers for linkage analysis (Nakumura *et al.,* 1987). Minisatellites sequence repeats of 10-60 b with high degree of repetition at a given locus have been shown to occur at many loci in the genome of *T. parva* (Bishop *et al.,* 2002; Oura *et al.,* 2003).

1.3.4.2.2 Micro- and Minisatellite detection methods and data analysis

Agarose and denaturing/non-denaturing polyacrylamide gel electrophoresis (PAGE) are the two common methods for manual analysis of fragments generated using molecular marker systems. Seperated fragments are detected either by ethidium bromide or silver staining of the gels. However accurate sizing is difficult with both agarose and polyacrylamide gels and these matrices do not allow resolution to within a single base pair unit. Morever, the mobility is also affected by

sequence composition so that repeat unit confounds migration of complementary strands in a gel based system. For example CA strands move faster on denaturing acrylamide gels than GT strands and this can result in two bands instead of one for the same allele (Saitoh *et al.,* 1998). The use of semi-automated systems such as the ABI Prism 377 (Applied Biosystems) has made accurate sizing of micro- and minisatellite alleles even to a difference of 1 bp possible. Automated flourescent-based capillary detection systems such as ABI 310, 3100, 3130 and 3730 have been developed. These genetic analyzers have the capability of detecting and genotyping microsatellites (Weber and May, 1989) randomly amplified polymorphic DNA (RAPD) markers (Williams, 1990), variable number of tandem repeats (NVTRs) (Jeffreys *et al*., 1985) and single strand conformational polymorphism (SSCP) (Orita *et al*., 1987) and amplified fragment length polymorphism (AFLP) (Vos *et al*., 1995). This fluorescence based detection system involves the measurement of emission radiation when laser light strikes a flourophore. To detect microsatellites in this system, a flourophore must be attached to the PCE product. This is accomplished by either attaching a flourophore to one of the PCR primers or by using fluorescently labeled dUTP in the PCR. Both labeling strategies can accommodate multiple flourophores. The flourophore-labelled DNA is then detected by electrophoresis through a matrix and as the DNA passes through a window at a fixed migration distance, a laser excites the flourophore and the fluorescence is measured. On the applied Biosystems method, multiple microsatellite loci can be labeled with different flourophores and run in the same lane. Applied Biosystems manufactures 6-FAM, HEX, TET, NED, PET and VIC amides and Pharmacia manufactures carboxyflourescin amide, all of which can be added during normal oligonucleotide synthesis to the 5′ end of the oligonucleotide. Thus, high throughput genotyping with several markers is now possible because the system offer the possibility of PCR multiplexing and post-PCR co-loading since markers are labeled with different dyes and hence can be detected simultaneously. These systems offer automated data analysis and reproducibility in

accurate sizing of micro- and minisatellite alleles, which are important for largescale germplasm genotyping projects (Wenz *et al*., 1998). The use of these systems has removed the earlier limitations of conventional gel based systems such as exposure to toxic chemicals, speed and inaccurate sizing of alleles.

1.4 Justification of the study

Rhipicephalus appendiculatus is the tick vector of *Theileria parva*, the causative agent of East Coast fever (ECF), considered to be the most important disease of cattle in East, Central and southern Africa. East coast fever kills one million cattle a year and is responsible for nearly half of calf deaths where the disease occurs. It is associated with high levels of mortality, especially in exotic and cross-bred cattle and is a major constraint to improvement of livestock production in 11 countries in Sub-Saharan Africa. Biological differences between populations of *Rhipicephalus appendiculatus* (*Acari: Ixodidae*) from different geographic region have been shown. However, the genetic diversity and population structure of this important tick vector remains unknown due to absence of appropriate genetic markers. Thus the genetic diversity and population structure of this important tick vector has not been established and neither have the diversity of laboratory tick stocks and the way they relate to the genotypes in the field been examined. While the epidemiology of East Coast fever is correlated to the distribution of *R. appendiculatus*, the role of field *R. appendiculatus* genotypes and how they contribute to this epidemiology remains unclear. Molecular tools are required to characterize the genetic variation of this tick vector and other closely related species in order to develop optimal controls strategies. Moreover, the only effective vaccine currently available for ECF is a live *T. parva* sporozoite vaccine, produced using three different stocks of *T. parva* and hundreds of cattle and thousands of ticks to generate large doses of the vaccine. Ensuring the standardisation of the tick stocks used in the vaccine production is important.

Recent advances on control of ticks and tick-borne diseases are focused on developing improved control strategies through a better understanding of tick biology especially in the area of tick vaccine antigens**.** Vaccines can circumvent some of the problems associated with use of acaricides such as increasing incidences of resistance, pollution of environment and rising costs of applications. There is now abundant evidence that vaccination with defined protein antigen is able to induce significant immunity to tick infestations (Willadsen, 2004). With anticipated completion of the tick genome sequencing, many potential vaccine candidates may be identified for further development. However, the development of these vaccines will depend on the understanding of tick vector genetic diversity as different tick populations will respond differently when challenged with a particular vaccine. The identification of micro and minisatellites will allow detailed studies of the genetic complexity of *R. appendiculatus*. Estimates can be made of the allele frequencies within populations. This will provide information on the genetic relatedness of geographically separated parasite populations, and allow estimation of the gene flow between them. Given the advantages of vaccines over the use of acaricides, the challenge is to produce a truly efficacious vaccine able to provide tick control that is both practical and cost effective. In order to develop such a vaccine against *R. appendiculatus*, the population structure of this vector is important as it is considerably more difficult to develop an effective vaccine against highly diverse vector populations. Thus, these markers will provide useful molecular tools for understanding tick genotypes, their distribution and how they relate to the epidemiology of East Coast fever. The markers will be useful in assessing and monitoring quality of *R. appendiculatus* laboratory colonies. To ensure quality of laboratory colonies, it is important that genetic diversity is monitored temporally and corrective measures taken whenever there is evidence to suggest that genetic diversity is significantly reduced, a phenomenon that may cause reduced colony fitness. In addition these markers could be applied to

ensure standardisation and quality control of the live *T. parva* vaccine. They will also be used for differentiating *R. appendiculatus* from closely related tick species in the field which may also be involved in transmission of ECF.

The study attempted to answer the following research questions:-

1. Are there any micro- and minisatellite markers in the *R. appendiculatus* Gene Index database?
2. Are the markers in the *R. appendiculatus* Gene Index database polymorphic and can they be useful for genotyping the tick?
3. What molecular genotypes of *R. appendiculatus* exist in diffent localities in Kenya and what is their genetic relationship?
4. How much sequence variation occurs within the cytochrome c oxidase I (COI), internal transcriber 2 (ITS2) and 12S rRNA genes of different *R. appendiculatus* populations?

1.4.1 Hypothesis tested

1. The *R. appendiculatus* Gene Index database contains micro- and minisatellite markers.
2. There exist distinct genetic populations of *R. appendiculatus* within Kenya which contain polymorphic micro and minisatellite DNA markers.
3. Significant sequence variation occurs within the cytochrome c oxidase I (COI), internal transcriber 2 (ITS2) and 12S rRNA genes of different *R. appendiculatus* populations.

1.4.2 Broad Objective

To identify and characterize a panel of polymorphic micro- and minisatellite genetic markers for genetic diversity analysis of *R. appendiculatus* tick vector.

1.4.3 Specific objectives

I. To identify a panel of micro- and minisatellite expressed sequence tags (EST) DNA marker within the *R. appendiculataus* Gene Index

II. To validate and characterize the identified micro- and minisatellite DNA markers using populations of *R. appendiculatus* and other rhipicephaline species

III. To assess the genetic diversity and population structure within and between populations of *R. appendiculatus* and other rhipicephaline species

IV. To assess the phylogenetic relationship between populations of field and laboratory *R. appendiculatus* ticks and and other rhipicephaline species using COI, ITS2 and 12S genes

1.5 References

Akoolo, L., Pellé, R., Saya, R., Awino, E., Nyanjui, J., Taracha, E. L. Kanyari, P., Mwangi, D. M., & Graham, S. P. (2008). Evaluation of the recognition of *Theileria parva* vaccine candidate antigens by cytotoxic T lymphocytes from Zebu cattle. *Veterinary immunology and immunopathology*, **121(3),** 216-221.

Anderson, T. J. C., Su, X. Z., Bockarie, M., Lagog, M., Day, K. P., (1999). Twelve microsatellite markers for the characterisation of *Plasmodium falciparum* from finger-prick blood samples. *Parasitology* **119**, 113–125.

Anderson, T.J.C., Haubold, B., Williams, J.T., Estrada-Franco, J.G., Richardson, J., Mollinedo, R., Bockarie, M., Mokili, J., Mharakurwa, S., French, N., Whitworth, J., Velez, I. D., Brockman, A. H., Nosten , F., Ferreira, M.U., Day, K.P., (2000). Microsatellite markers reveal a spectrum of population structures in the malaria parasite *Plasmodium falciparum*. *Molecular Biology and Evolution* **17**, 1467–1482.

Andreotti, R., Sampaio, C. A.M., Gomes, A. & Tanaka, A. S. (1999). A serine proteinase inhibitor immunoprotection from B. microplus unfed larvae in calves. *IV Seminario Internacional de Parasitologia Animal*, 20–22 October, 1999, Puerto Vallarta, Jalisco, Mexico.

Andreotti, R., Gomes, A., Malavazi-Piza, K. C., Sasaki, S. D., Sampaio, C. A.M. & Tanaka, A. S. (2002). BmTI antigens induce a bovine protective immune response against B. microplus tick. *International Immunopharmacology* **2,** 557–563.

Armour, J.A.L., (1999). Microsatellites and mutation processes in tandemly repetitive DNA. In: Goldstein, D., Schlotterer, C. (Eds.), *Microsatellites: Evolution and Applications, Oxford University Press, Oxford*, pp. 24–29.

Avise, J. C. (1994). Molecular Markers, Natural History and Evolution*:* Eds. Chapman & Hall New York and London. pp 3-15

Babo Martins, S., Di Giulio, G., Lynen, G., Peters, A., & Rushton, J. (2010). Assessing the impact of East Coast Fever immunisation by the infection and treatment method in Tanzanian pastoralist systems. *Preventive veterinary medicine*, **97(3),** 175-182.

Baliraine, F. N., Osir, E. O., Obuya, S. B., & Mulaa, F. J. (2001). Protein polymorphism in two populations of the brown ear tick, *Rhipicephalus appendiculatus* Neumann (Acari: Ixodidae). *International Journal of Tropical Insect Science*, **20(3),** 227-231.

Barker, S. C. (1998). Distinguishing species and populations of rhipicephaline ticks with its 2 ribosomal RNA. *The Journal of parasitology*, 887-892.

Barker, S. C., & Murrell, A. (2004). Systematics and evolution of ticks with a list of valid genus and species names. *Parasitology*, **129(S1),** S15-S36.

Bazarusanga, T., Geysen, D., Vercruysse, J., & Madder, M. (2007). An update on the ecological distribution of Ixodid ticks infesting cattle in Rwanda: countrywide cross-sectional survey in the wet and the dry season. *Experimental and Applied Acarology*, **43(4),** 279-291.

Beati, L & Keirans, J, E. (2001). Analysis of the systematic relationships among ticks of the genera *Rhipicephalus* and *Boophilus* (acari: ixodidae) based on mitochondrial 12S ribosomal DNA gene sequences and morphological characters. *Journal of parasitology*, **87(1),** 32–48.

Beaumont, M. A., and Bruford, M.W., (1999). "Microsatellites in conservation genetics. In: Microsatellites: Evolution and Application" Edited by D. B. Goldstein and Schlötterer, Orford University Press, New York, pp 165-183.

Berkvens, D. L., Young, A. S., & Pegram, R. G. (1994). Collaborative research on behavioural diapause in adult *Rhipicephalus appendiculatus* populations. In Tick and tick-borne disease control in east, central and southern Africa, 1991-1994. *Proceedings of a joint OAU, FAO and ILRAD workshop held in Lilongwe, Malawi, 25-28 April 1994*. ILRI.

Berkvens, D.L., Pegram, R.G., and Brandt, J.R.A., (1995). A study of the diapausing behaviour of *Rhipicephalus appendiculatus* and *Rhipicephalus zambeziensis* under quasi-natural conditions in Zambia. *Medical and Veterinary Entomology***, 9**, 307–315.

Bishop, R., Geysen, D., Skilton, R., Odongo, D., Nene, V., Allsopp, B., Mbogo, S., Spooner, P., Morzaria, S., (2002a). Genomic polymorphism, sexual recombination and molecular epidemiology of *Theileria parva*. In: Mckeever, D., Dobbelaere, D. (Eds.), *Theileria*, Kluwer Academic, Dordrecht, pp. 23–40.

Bishop, R., Lambson, B., Wells, C., Pandit, P., Osaso, J., Nkonge, C., Morzaria, S., Musoke, A. & Nene, V. (2002b). A cement protein of the tick *Rhipicephalus appendiculatus*, located in the secretory e cell granules of the type III salivary gland acini, induces strong antibody responses in cattle. *International Journal for Parasitology* **32,** 833–842.

Bishop, R., Musoke, A., Morzaria, S., Gardner, M., and Nene, V. (2004). Theileria: intracellular protozoan parasites of wild and domestic ruminants transmitted by idodid ticks. *Parasitology* **129**, S271-S283

Black, W. C and Piesman, J. (1994). Phylogeny of hard- and soft-tick taxa (Acari: Ixodida) based on mitochondrial 16S rDNA sequences. *Proceedings of the National Academy of Sciences*, USA. **91,** 10034-10038.

Branagan, D., (1973). The development periods of the Ixodid tick *Rhipicephalus appendiculatus* Neum. under laboratory conditions. *Bulletin of Entomological Research,* **63**, 155–168.

Brown, S. J., & Askenase, P. W. (1986). *Amblyomma americanum*: Physiochemical isolation of a protein derived from the tick salivary gland that is capable of inducing immune resistance in guinea pigs. *Experimental parasitology*, **62(1),** 40-50.

Bruford, M. W., and Wayne, R. K., (1993). Microsatellites and their application to population genetic studies. *Current opinion in Genetics and Development* **3**, 939-943.

Bull, C. M., Andrews, R. H., and Adams, M., (1984). Patterns of genetic variation in a group of parasites, the Australian reptile ticks. *Heredity*, **53**, 509–526.

Burridge M. J., Morzaria S. P., Cunningham M. P. & Brown C.G.D. (1972) Duration of immunity to EastCoast fever (*Theileria parva*) infection in cattle. *Parasitology*, **64,** 511.

Chaka, G., Billiouw, M., Geysen, D. M and Berkvens, D, L. (1999). Spatial and temporal variation in *Rhipicephalus appendiculatus* size in eastern Zambia. *Tropical Medical International Health*. **4(9),** 43-8.

Chakravarti, A. (1999). Population genetics—making sense out of sequence. *Nature genetics*, **21,** 56-60.

Chiera, J. W., Newson, R. M., & Cunningham, M. P. (1985). The effect of size on feeding and breeding performance of *Rhipicephalus appendiculatus* Neumann. *International Journal of Tropical Insect Science*, ***6*(4),** 555-560.

Chigagure, N. N., Baxter, G. D., and Barker. S. C., (2000). Microsatellite loci of the cattle tick *Boophilus microplus* (Acari: Ixodidae). *Experimental and Applied Acarology* **24**, 951–956.

Cobbett, N. G. (1947). Preliminary tests in Mexico with DDT, cube hexachlorocyclohexane (Benzene hexachloride), and combinations thereof, for the control of the cattle fever tick, *Boophilus annulata*. *American Journal of Veterinary Research* **8,** 280–283.

Cox, A. J. and Hebert, P. D. N. (2001). Colonization, extinction and phylogeographic patterning in a freshwater crustacean. *Molecular Ecology.* **10**, 371–386.

Crampton, A., McKay, I., & Barker, S. C. (1996). Phylogeny of ticks (Ixodida) inferred from nuclear ribosomal DNA. *International journal for parasitology*, **26(5),** 511-517.

Cunningham, M. P., Brown, C. G. D., Burridge, M. J., & Purnell, R. E. (1973). Cryopreservation of infective particles of *Theileria parva*. *International Journal for Parasitology*, **3(5),** 583-587.

Davey, R. B. & Ahrens, E. H. (1984). Control of *Boophilus* ticks on heifers with two pyrethroids applied as sprays. *American Journal of Veterinary Research* **45,** 1008–1010.

Davey, R. B., and Hilburn, L. R., (1991). Reduction in egg viability resulting from infestations on cattle of hybridized *Boophilus* ticks and *B. microplus* (Acari: Ixodidae) at various ratios. *Journal of Medical Entomology* **28**, 763–769

De Castro J. J., (1997). Sustainable tick and tick-borne disease control in livestock improvement in developing countries. *Veterinary Parasitology*, **71,** 77-97

De la Fuente, J., & Kocan, K. M. (2003). Advances in the identification and characterization of protective antigens for recombinant vaccines against tick infestations. *Expert Review of Vaccines*, **2(4),** 583-593.

Di Giulio, G., Lynen, G., Morzaria, S., Oura, C., & Bishop, R. (2009). Live immunization against East Coast fever–current status. *Trends in parasitology*, **25(2),** 85-92.

Dobbelaere, D. A. E., Spooner, P. R., Barry, W. C., & Irvin, A. D. (1984). Monoclonal antibody neutralizes the sporozoite stage of different *Theileria parva* stocks. *Parasite immunology*, ***6*(4),** 361-370.

Dobson, S. J., & Barker, S. C. (1999). Phylogeny of the Hard Ticks (Ixodidae) Inferred from 18S rRNA Indicates That the Genus *Aponomma* Is Paraphyletic. *Molecular Phylogenetics and Evolution*, **11(2),** 288-295.

Dunn, C. P. (2003). Keeping taxonomy based in morphology. *Trends in Ecology & Evolution*, **18(6),** 270-271.

Ellis, J. R and Burke, J. M. (2007). EST-SSRs as a resource for population genetic analyses. *Heredity*. **99,** 125–132.

Estoup, A., Tailliez, C., Cornuet, J. M., & Solignac, M. (1995). Size homoplasy and mutational processes of interrupted microsatellites in two bee species, Apis mellifera and Bombus terrestris (Apidae). *Molecular Biology and Evolution*,***12*(6),** 1074-1084.

Estrada-Peña, A. (2001). Forecasting habitat suitability for ticks and prevention of tick-borne diseases. *Veterinary parasitology*, **98(1),** 111-132.

Fenton, B., Malloch, G., & Moxey, E. (1997). Analysis of eriophyid mite rDNA internal transcribed spacer sequences reveals variable simple sequence repeats. *Insect molecular biology*, **6(1),** 23-32.

Folmer, O., Black, M., Hoeh, W., Lutz, R. & Vrijenhoek, R. (1994). DNA primers for amplification of mitochondrial cytochrome *c* oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology* **3**, 294–299.

Garci´A-Garci´A, J. C., Gonzalez, I. L., Gonzalez, D.M., Valde´ S, M., Mendez, L., Lamberti, J., D'agostino, B.,Citroni, D., Fragoso, H., Ortiz, M., Rodriguez, M., and De La Fuente, J., (1999). Sequence variations in the *B. microplus* Bm86 locus and implications for immunoprotection in cattle vaccinated with this antigen. Experimental and *Applied Acarology* **11**, 883–895.

Gaston, K. J., & O'Neill, M. A. (2004). Automated species identification: why not?. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, **359(1444),** 655-667.

Gauer, M. U., Mackenstedt, H, Mehlhorn, E., Schein, F., Zapf E, Njenga, E., Young, A.,and Morzaria, S. (1995) DNA measurements and ploidy determination of developmental stages in the life cycles of *Theileria annulata* and *T. parva*. *Parasitology Research* **81,** 565–574.

Gemayel, R., Vinces, M. D., Legendre, M. and Verstrepen, K. J. (2010). Variable tandem repeats accelerate evolution of coding and regulatory sequences. *Annual Review of Genetics* **44,** 445–77.

George, J. E., Davey, R. B., Ahrens, E. H., Pound, J.M. & Drummond, R. O. (1998). Efficacy of amitraz (Taktic1 12.5% EC) as a dip for the control of *Boophilus microplus* (Canestrini) (Acari: Ixodidae) on cattle. Preventive Veterinary Medicine **37,** 55–67.

George, J. E., Pound, J. M., & Davey, R. B. (2004). Chemical control of ticks on cattle and the resistance of these parasites to acaricides. *Parasitology*, **129(S1),** S353-S366.

Gomes, A.F., and Wouters, G., (1991). Species identification of *Rhipicephalus duttoni* in relation to the *Rhipicephalus appendiculatus* group (Acarina: Ixodidae). *J. Med. Entomol.* **28,** 16–18

Graham, C. H., Ferrier, S., Huettman, F., Moritz, C., & Peterson, A. T. (2004). New developments in museum-based informatics and applications in biodiversity analysis. *Trends in ecology & evolution*, **19(9),** 497-503.

Graham, S. P., Pellé, R., Honda, Y., Mwangi, D. M., Tonukari, N. J., Yamage, M., ... & Taracha, E. L. (2006). *Theileria parva* candidate vaccine antigens recognized by immune

bovine cytotoxic T lymphocytes. *Proceedings of the National Academy of Sciences of the United States of America*, **103(9),** 3286-3291.

Hamada, H., Petrino, M.G., Kukunaga, T.A., (1982). Novel repeated element with Z-DNA forming potential is widely found in evolutionary diverse eucaryotic genomes. *Proc. Natl. Acad. Sci. USA* **79,** 6465–6469

Hamilton, M. B., Pincus, E. L., Di-Fiore, A., Fleischer, R. C. (1999). Universal linker and ligation procedures for construction of genomic DNA libraries enriched for microsatellites. *Biotechniques*, **27,** 500-507.

Hancock, J. M., (1999). Microsatellites and other simple sequences: genomic context and mutational mechanisms. In: *Microsatellites.Evolution and Applications* (eds Goldstein DB, Schl.tterer C), pp. 1-9. Oxford University Press, New York. TE ATION: A REVIEW

Healy, J. A., E., Cross, T. F., & Healy, A. (2004). The α-Gpdh polymorphism in the tick Ixodes ricinus: similar diurnal trends in genotypic composition in Irish and Swedish population samples. *Experimental & applied acarology*, **32(1-2),** 111-118.

Healy, J.A., (1979). Phosphoglucomutase polymorphism in the tick *Ixodes ricinus*. *Parasitology* **78**, 7–17.

Hebert, P. D. N., Ratnasingham, S. & deWaard, J. R. (2002). Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London. Series B*, **270,** S596–S599.

Hebert, P.D. N., Penton, E, H., Burns, J. M., Janzen, D. H., Hallwachs, W. (2004). Ten species in one: DNA barcoding reveals cryptic species in the Neotropical skipper butterfly *Astraptes fulgerator*. *Proceedings of the National Academy of Sciences,* USA. **101 (41),** 14812–14817

Hedrick, P. W. (1999). Perspective: highly variable loci and their interpretation in evolution and conservation. *Evolution*, 313-318.

Hedrick, P. W. (2005). A standardized genetic differentiation measure.*Evolution, **59*(8),** 1633-1638.

Hilburn, L. R., & Sattler, P. W. (1986). Are tick populations really less variable and should they be. *Heredity*, **57(1),** 113-117.

Holmes, N. G., Mellersh, C. S., Humphrey, S.J., Binns, M. M.,Holliman, A., Curtis, R., and Sampson, J. (1993). Isolation and characterization of microsatellites from the canine genome. *Animal genetics* **24**, 289-292

Hoogstraal, H. (1956). *African ixodoidea* (Vol. 1). US Government Printing Office.

Hoogstraal, H., & Aeschlimann, A. (1982). Tick-host specificity. In *Bulletin de la société entomologique suisse*.

Hoogstral, H., and Aeschlimann, A., (1992). Ticks host specificity. *Bulletin de la societe Entomologigue Suisse,* **55**, 5-32

Horak, I. G., Camicas, J. L., & Keirans, J. E. (2002). The *Argasidae, Ixodidae* and *Nuttalliellidae* (Acari: Ixodida): a world list of valid tick names. *Experimental & applied acarology*, **28(1-4),** 27-54.

Irwin, D. M., Kocher, T.D., Wilson, A.C. (1991). Evolution of the cytochrome b gene of mammals. *Journal of molecular evolution*, **32,** 128–144

Jeffreys, A.J, Wilson, V., Thein, S. L. (1985). Hypervariable "minisatellite" regions in human DNA. *Natur*e, **314,** 67–73

Jensen, P. M., & Kaufmann, U. (2003). Seasonal and diel activity of *Ixodes ricinus* (Acari: Ixodidae) subpopulations in Denmark. Aspects of size, physiological age, and malate dehydrogenase genotype in a forest site without any undergrowth. *Experimental & applied acarology*, **30(4),** 289-303.

Jones, L. D., Davies, C. R., Williams, T., Cory, J., & Nuttall, P. A. (1990). Non-viraemic transmission of Thogoto virus: vector efficiency of *Rhipicephalus appendiculatus* and Amblyomma variegatum. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, **84(6),** 846-848.

Jongejan, F., and Uilenberg, G., (1994). Ticks and control methods. *Review of Science and Technology*, **13**, 1201-1226

Kaaya, G. P., Samish, M., and Glazer, I.,, (2000). Laboratory evaluation of pathogenicity of entomogenous nematodes to African tick species. *Annual New York Academy of Sciences*, **916**, 303-308

Kain D, E., Sperling, F.A., Daly, H.V., Lane, R. S. (1999). Mitochondrial DNA sequence variation in *Ixodes pacificus* (Acari: Ixodidae). *Heredity* **83,** 378-386.

Kaiser, M. N., Sutherst, R.W., and Bourne, A.S., (1982). Relationship between ticks and Zebu cattle in southern Uganda. *Tropical Animal Health and Production*, **14**, 63–74.

Kaiser, M.N., Sutherst, R.W., Bourne, A.S., Gorissen, L. and Floyd, R.B. (1988). Population dynamics of ticks on Ankole cattle in five ecological zones in Burundi and strategies for their control. *Preventive Veterinary Medicine*, **6,** 199–222.

Kariuki, D. P., Young, A. S., Morzaria, S. P., Lesan, A. C., Mining, S. K., Omwoyo, P., ... & Molyneux, D. H. (1995). *Theileria parva* carrier state in naturally infected and artificially immunised cattle. *Tropical animal health and production*, **27(1),** 15-25.

Kasibule, D. (2013). Impact of the immunisation of cattle against East Coast fever in Bugabula County, Kamuli District, Uganda (Doctoral dissertation).

Katzer, F., Ngugi, D., Oura, C., Bishop, R. P., Taracha, E. L., Walker, A. R., & McKeever, D. J. (2006). Extensive genotypic diversity in a recombining population of the apicomplexan parasite *Theileria parva*. *Infection and immunity*, **74(10**), 5456-5464.

Kibuka-Sebitosi, E. (2006). Potential role of lectins in ticks: *Rhipicephalus appendiculatus* and *Rhipicephalus pulchellus* (Acari: Ixodidae). *Systematic and Acarology Acarology Special Publications*, **21(1),** 1-15.

Knowlton, N. & Weigt, L. A. (1998). New dates and new rates for divergence across the Isthmus of Panama. *Proceedings of the Royal Society of London. Series B*, **265**, 2257–2263.

Kocher, T. D., Thomas, W. K., Meyer, A., Edwards, S. V., Pääbo, S., Villablanca, F. X., & Wilson, A. C. (1989). Dynamics of mitochondrial DNA evolution in animals: amplification and sequencing with conserved primers. *Proceedings of the National Academy of Sciences*, **86(16),** 6196-6200.

Konnai, S., Imamura, S., Nakajima, C., Witola, W. H., Yamada, S., Simuunza, M., Nambota, A., Yasuda, J., Ohashi, K., and Onuma, M. (2006). Acquisition and transmission of *Theileria parva* by vector tick, *Rhipicephalus appendiculatus*. *Acta tropica*, **99(1),** 34-41.

Konnai, S., Yamada, S., Imamura, S., Simuunza, M., Chembensof, M., Chota, A., Nambota, A., Ohashi, K., and Onuma, M. (2007). Attachment duration required for *Rhipicephalus appendiculatus* to transmit *Theileria parva* to the host. *Vector-Borne and Zoonotic Diseases*, **7(2),** 241-248.

Kubasu, S. S. (1992). The ability of *Rhipicephalus appendiculatus* (Acarina: Ixodidae) stocks in Kenya to become infected with *Theileria parva*. *Bulletin of Entomological Research,* **82,** 349-353.

Lan, M., W. Sheng, C. Huang, X. Gu, Y. Jiang, U. Zhao, And W. Lu. (1996). RAPD-PCR method for detecting DNA polymorphism in *Boophilus microplus* (Acari: Ixodidae). *Systematic and Applied Acarology* **1,** 11-14.

Latif, A. A., Punyua, D. K., Nokoe, S., & Capstick, P. B. (1991). Tick infestations on Zebu cattle in western Kenya: individual host variation. *Journal of medical entomology*, **28(1),** 114-121.

Latif, A. A., Hove, T., Kanhai, G. K., Masaka, S., & Pegram, R. G. (2001). Epidemiological observations of Zimbabwean theileriosis: disease incidence and pathogenicity in susceptible cattle during *Rhipicephalus appendiculatus* nymphal and adult seasonal activity. *The Onderstepoort journal of veterinary research*, **68(3),** 187.

Lawrence, J. (1992). History of bovine theileriosis in southern Africa. In The Epidemiology of Theileriosis in Africa (ed. Norval, R. A. I., Perry, B. D. & Young, A. S.) 1–40. Academic Press, London.

Lee, M. S. (2004). The molecularisation of taxonomy. *Invertebrate Systematics*, **18(1),** 1-6.

Lemos, A.M., Teodoro, R.L., Oliveira, G.P., Madalena, F.E., (1985). Comparative performance of six Holstein-Friesian x Guzera grades in Brazil. 3. Burdens of *Boophilus microplus* under field conditions. *Animal Production* **41,** 187–191.

Lessard, P., L'Eplattenier, R., Norval, R.A.I., Perry, T. T. Dolan, A. Burrill, H. Croze, M. Sorensen, J. G. Grootenhuis, and A. D. Irvin (1990). Geographical information systems for studying the epidemiology of cattle diseases caused by *Theileria parva*. *The Veterinary Record*, **126**, 255–262.

Litt, M., Luty, J. A., (1989). A hypervariable microsatellite revealed by *in vitro* amplification of a dinucleotide repeat within the cardiac muscle actin gene. *American Journal of Human Genetics*, **44**, 397-401.

Lynch, M. & Jarrell, P. E. (1993). A method for calibrating molecular clocks and its application to animal mitochondrial DNA. *Genetics* **135,** 1197–1208.

Madder, M., Torreele, G., & Berkvens, D. (1996). Inheritance of weight in *Rhipicephalus appendiculatus* ticks (Acari: Ixodidae) in the laboratory. *Experimental & applied acarology*, **20(11),** 659-665.

Madder, M., Speybroeck, N., Brandt, J., & Berkvens, D. (1999). Diapause induction in adults of three *Rhipicephalus appendiculatus* stocks. *Experimental & applied acarology*, **23(12),** 961-968.

Madder, M., Speybroeck, N., Brandt, J., Tirry, L., Hodek, I., & Berkvens, D. (2002). Geographic variation in diapause response of adult *Rhipicephalus appendiculatus* ticks. *Experimental & applied acarology*, **27(3),** 209-221.

Mangold, A. J., Bargues, M. D. & Mas-Coma, S. (1998). Mitochondrial 16S rDNA sequences and phylogenetic relationships of species of Rhipicephalus and other tick genera among Metastriata (Acari: Ixodidae). *Parasitology Research*, **84,** 478 - 484.

Mans, B. J., & Neitz, A. W. (2004). Adaptation of ticks to a blood-feeding environment: evolution from a functional perspective. *Insect biochemistry and molecular biology*, ***34*(1),** 1-17.

Mans, B. J., and Neitz A. W. H., (2004). Adaptation of ticks to a blood feeding environment: evolution from a functional perspective. *Insect Biochemistry and Molecular Biology* **34**, 1-17. **Marcet, P. L., Lehmann, T., Groner, G., Gurtler, R. E., Kitron, U., Dotson, E. M.,** (2006) Identification and characterization of microsatellite markers in the Chagas disease vector *Triatoma infestans* (Heteroptera: Reduviidae). *Infect Genet Evol.* **6**, 32-7

Marcotty, T., Brandt, J., Billiouw, M., Chaka, G., Losson, B., Berkvens, D., (2002). Immunisation against *Theileria parva* in eastern Zambia: influence of maternal antibodies and demonstration of the carrier status. *Veterinary Parasitology* **110,** 45–56.

Maritim, A. C., Kariuki, D. P., Young, A. S., & Mutugi, J. J. (1989). The importance of the carrier state of *Theileria parva* in the epidemiology of theileriosis and its control by immunization. Theileriosis in Eastern, Central and Southern Africa. International Laboratory for Research on Animal Diseases, Nairobi, 121-128.

Mattioli, R. C., & Dempfle, L. (1995). Recent acquisitions on tick and tick-borne disease resistance in N'dama (*Bos taurus*) and Gobra zebu (*Bos indicus*) cattle. *Parassitologia*, **37(1),** 63-67.

Mattioli, R. C., Pandey, V. S., Murray, M., & Fitzpatrick, J. L. (2000). Immunogenetic influences on tick resistance in African cattle with particular reference to trypanotolerant N'Dama (*Bos taurus*) and trypanosusceptible Gobra zebu (*Bos indicus*) cattle. *Acta tropica*, ***75*(3),** 263-277.

Maunder, J. C. J. (1949). Cattle tick control: Results achieved in the field with DDT and BHC. *Queensland Agricultural Journal*, 1–8.

Mbogo, S. (1999). The infection and treatment method of immunisation: Practice and problems. In Live Vaccines for *Theileria parva*: Deployment in Eastern, Central and Southern Africa*: Proceedings of an FAO/OAU-IBAR/ILRI Workshop Held at ILRI, Nairobi, Kenya, 10-12 March 1997* (p. 21). ILRI . Eds (Subhash P. Morzaria, Susanna Williamson).

McCulloch, B., Kalaye, W.J., Tungaraza, R., Suda, B'Q.J., and Mbasha, E.M.S., (1968). A study of the life history of the tick *Rhipicephalus appendiculatus* – the main vector of East Coast Fever – with reference to its behaviour under field conditions and with regard to its control in Sukumaland,Tanzania. *Bulletin of Epizootic Diseases in Africa,* **16**, 477–500.

McKeever, D. J., Taracha, E. L., Innes, E. L., MacHugh, N. D., Awino, E., Goddeeris, B. M., & Morrison, W. I. (1994). Adoptive transfer of immunity to *Theileria parva* in the CD8+ fraction of responding efferent lymph. *Proceedings of the National Academy of Sciences*, **91(5),** 1959-1963.

McKeever D. J., ed. (1995). *Novel immunization strategies against protozoan parasites*. Proceedings of a Workshop held at ILRAD, Nairobi, Kenya, 1–4 November 1993. ILRAD, **Nairobi.**

Melewas, J., Majaliwa, M. and Lynen, L., 1999. ECF immunisation in Tanzania. In: Morzaria, S. and Williamson, S. (eds), Live vaccines for *Theileria parva*: Deployment in Eastern, Central and Southern Africa. Proceedings of an FAO, OAU-IBAR and ILRI Workshop, Nairobi, 1997. ILRI, 16–25

Minjauw, B., and McLeod, A., (2003). Tick-borne diseases and poverty. The impact of ticks and tick-borne diseases on the livelihood of small scale and marginal livestock owners in India, eastern and southern Africa. *Animal health program report, Centre for Tropical Veterinary Medicine, University of Edinburg, UK*

Moll, G., Lohding, A., Young, A. S., & Leitch, B. L. (1986). Epidemiology of theileriosis in calves in an endemic area of Kenya. *Veterinary parasitology*, **19(3),** 255-273.

Moorhouse, D.E., and Tatchell, R.J., (1966). The feeding processes of the cattle-tick *Boophilus microplus* (*Canestrini*): a study in host–parasitic relations. I. Attachment to the host. *Parasitology* **56**, 623–632.

Moran, M. C., Nigarura, G., & Pegram, R. G. (1996). An assessment of host resistance to ticks on cross-bred cattle in Burundi. *Medical and veterinary entomology*, **10(1),** 12-18.

Morrison, W. I., Goddeeris, B. M., Teale, A. J., Groocock, C. M., Kemp, S. J., & Stagg, D. A. (1987). Cytotoxic T-cells elicited in cattle challenged with ***Theileria parva*** (Muguga): evidence for restriction by class I MHC determinants and parasite strain specificity. *Parasite immunology*, **9(5),** 563-578.

Morrison, W. I., Goddeeris, B. M., Brown, W.C., Baldwin, C. L., Teale, A. J. (1989). *Theileria parva* in cattle: characterization of infected lymphocytes and the immune responses they provoke. *Veterinary Immunology and Immunopathology*, **20:** 213–237.

Morzaria, S. P, Young, J. R., Spooner, P. R., Dolan, T.T., Bishop, R. P. (1993). *Theileria parva*: a restriction map and genetic recombination. In: Morzaria SP, editor. Genome analysis of protozoan parasites. ILRAD: Nairobi. pp. 67–73.

Morzaria, S. P., & Williamson, S. (Eds.). (1999). Live Vaccines for *Theileria parva*: Deployment in Eastern, Central and Southern Africa: *Proceedings of an FAO/OAU-IBAR/ILRI Workshop Held at ILRI, Nairobi, Kenya, 10-12 March 1997*. ILRI (aka ILCA and ILRAD).

Mtambo, J., Madder, M., Van Bortel, W., Geysen, D., Berkvens, D., & Backeljau, T. (2007a). Genetic variation in *Rhipicephalus appendiculatus* (Acari: sIxodidae) from Zambia: correlating genetic and ecological variation with *Rhipicephalus appendiculatus* from eastern and southern Africa. *Journal of Vector Ecology*, **32(2),** 168-175.

Mtambo, J., Madder, M., Van Bortel, W., Chaka, G., Berkvens, D., & Backeljau, T. (2007b). Further evidence for geographic differentiation in *R. appendiculatus* (acari: ixodidae) from eastern and southern provinces of Zambia. *Experimental and Applied Acarology*, **41(1-2),** 129-138.

Mtambo, J., Madder, M., Van Bortel, W., Berkvens, D., & Backeljau, T. (2007c). *Rhipicephalus appendiculatus* and *R. zambeziensis* (Acari: Ixodidae) from Zambia: a molecular reassessment of their species status and identification. *Experimental and Applied Acarology*, **41(1-2**), 115-128.

Munstermann, L. E. (1994). Unexpected genetic consequences of colonization and inbreeding: allozyme tracking in Culicidae (Diptera). *Annals of Entomological Society of America,* **87,** 157-164.

Murray, M., Stear, M.J., Trail, J.C.M., d'Ieteren, G.D.M., Agyemang, K., Dwinger, R.H. (1991). Trypanosomiasis in cattle. Prospects for control. In: Owen, J.B., Axford, R.F.E. (Eds.), *Breeding for Disease Resistance in Farm Animals*. C.A.B. International, Oxon, UK, pp. 203–223.

Murrell, A., Campbell, N. J. H., and Barker, S. C. (1999). Mitochondrial 12S rDNA indicates that the *Rhipicephalinae* (Acari: Ixodida:Ixodidae) is paraphyletic. *Molecular Phylogenetics and Evolution,* **12,** 83–86

Murrell, A., Campbell, N. J. H., and Barker, S. C. (2000). Phylogenetic analyses of the rhipicephaline ticks indicate that the genus *Rhipicephalu*s is paraphyletic. *Molecular Phylogenetics and Evolution* **16,** 1–7.

Murrell, A.N., Campbell, J.H. and Barker, S.C. (2001). Recurrent gains and losses of large (84–109 bp) repeats in the rDNA internal transcribed spacer 2 (ITS2) of rhipicephaline ticks. *Insect Molecular Biology* 1**0,** 587–596.

Musoke, A. J., Nantulya, V. M., Rurangirwa, F. R., & Buscher, G. (1984). Evidence for a common protective antigenic determinant on sporozoites of several *Theileria parva* strains. *Immunology,* **52(2),** 231.

Musoke, A., Morzaria, S., Nkonge, C., Jones, E., & Nene, V. (1992). A recombinant sporozoite surface antigen of *Theileria parva* induces protection in cattle. *Proceedings of the National Academy of Sciences,* **89(2),** 514-518.

Nakumura, Y., Leppert, M., O'Connell, P., Wolff, R., Holm, T., Culver, M., Martin, C., Fujimoto, E., Hoff, M., Kumlin, E., and White, R., (1987). Variable number of tandem repeats (VNTR) markers for human gene mapping. *Science,* **235**, 1616-1622

Nene, V., Lee, D., Quackenbush, J., Skilton, R., Mwaura, S., Gardner, M. J. & Bishop, R., (2002). AvGI, an index of genes transcribed in the salivary glands of the ixodid tick *Amblyomma variegatum. International Journal for Parasitology* **32**, 1447–1456.

Nene, V., Lee, D., Kang'a, S., Skilton, R., Shah, T., de Villiers, E., Mwaura, S., Taylor, D., Quackenbush, J., and Bishop, R., (2004). Genes transcribed in the salivary glands of female *Rhipicephalus appendiculatus* ticks infected with *Theileria parva*. *Insect Biochemistry and Molecular Biology,* **34**, 1117-1128.

Neumann, L. G. (1901). Re´vision de la famille des Ixodide´s. Me´moires de la Socie´te´ Zoologique de France **14:** 249–372.

Newson, R. M., (1978). The life cycle of *Rhipicephalus appendiculatus* on the Kenyan coast. In: *Tick-borne diseases and their vectors, J.K.H. Wilde (ed.). Lewis Reprints Ltd., Tonbridge*.

Newson, R. M., and Punyua, D.K., (1978). The life cycle of *Rhipicephalus appendiculatus* and associated species in an ecologically marginal situation in central Kenya. *In: Tick-borne diseases and their vectors, J.K.H. Wilde (ed.). Lewis Reprints Ltd., Tonbridge.*

Newton, L. G. (1967). Acaricide resistance and cattle tick control. *Australian Veterinary Journal* **43,** 389–394.

Ngumi, P. N., Lesan, A. C., Williamson, S. M., Awich, J. R., Morzaria, S. P., Dolan, T. T., & Young, A. S. (1994). Isolation and preliminary characterisation of a previously unidentified *Theileria* parasite of cattle in Kenya. *Research in veterinary science,* **57(1),** 1-9.

Nolan, J., Roulston, W. J., & Schnitzerling, H. J. (1979).The potential value of some synthetic pyrethroids for control of the cattle tick (*Boophilus microplus*). *Australian Veterinary Journal* **55,** 463–466.

Nolan, J., (1990). Acaricide resistance in single and muilti-host ticks and strategies for control. *Parasitologia*, **32**, 145-153

Norris, D. E., Klompen, J. S., H. & Black, W. C. (1999). Comparison of the mitochondrial 12S and 16S ribosomal DNA genes in resolving phylogenetic relationships among hard ticks (Acari: Ixodidae). *Annals of the Entomological Society of America*. **92,** 117 - 129.

Norris, D. E., Shurtleff, A. C., Yeya, T. T and Lanzaro G. C. (2001). Microsatellite DNA polymorphism and heterozygosity among field and laboratory populations of *Anopheles gambiae s.s.* (*Diptera: Culicidae*). *Journal of Medical Entomology* **38(2),** 336-340.

Norval, R. A. F., and Perry, B. D., (1990). Introduction, spread and subsequent disappearance of the brown year tick, *Rhipicephalus appendiculatus*, from the southern lowveld of Zimbabwe. *Experimental Applied Acarology* **9**, 103-111

Norval, R.A. F., Lawrence, J.A., Young, A.S., Perry, B.D., Dolan, T.T., Scott, J., (1991). *Theileria parva*: influence of vector, parasite and host relationships on the epidemiology of theileriosis in southern Africa. *Parasitology* **102**, 347–356.

Norval, R. A. F., Perry, B. D., and Young, A. S., (1992). The Epidemiology of Theileriosis in Africa. Academic Press Ltd London & San Diego

Norval, R.A.I., Sutherst, R.W., Jorgensen, O.G., Kerr, J.D. (1997a). The effects of the bont tick, *Amblyomma hebraeum*, on milk production of Sanga and Sanga x Brahman cattle. *Medical Veterinary Entomology* **11,** 143–147.

Norval, R.A.I., Sutherst, R.W., Kurki, J., Kerr, J.D., Gibson, J.D. (1997b). The effects of the brown-ear tick, *Rhipicephalus appendiculatus*, on milk production of Sanga cattle. *Medical Veterinary Entomology* **11,** 148–154.

Nsubuga-Mutaka, R. (1999). ECF Immunisation in Uganda- In Live Vaccines for *Theileria parva:* Deployment in Eastern, Central and Southern Africa: Proceedings *of an FAO/OAU-IBAR/ILRI Workshop Held at ILRI, Nairobi, Kenya, 10-12 March 1997* (p. 21). ILRI . edited by Subhash P. Morzaria, Susanna Williamson.

Nuttall, P.A. (1998). Displaced tick-parasite interactions at the host interface. *Parasitology* **116**, S65–72

Nuttall, P.A., Paesen, G.C., Lawrie, C.H. and Wang, H., (2000). Vector–host interactions in disease transmission. *J. Mol. Microbiol. Biotechnol.* **2**, 381–386.

Obenchain, F.D., and Galun, R., (1982). Preface. *In: Physiology of Ticks, edited by F.D Obechain et al. pp vii-ix. Pergamon Press, Oxford, U.K*

Ochanda, H., Young, A. S., Wells, C., Medley, G. F. & Perry, B. D. (1996). Comparison of the transmission of *Theileria parva* between different instars of *Rhipicephalus appendiculatus*. *Parasitology* **113,** 243–253.

Ochanda, H., Young, A. S., Medley, G. F and Perry, B.D. (1998a). Vector competence of 7 rhipicephalid tick stocks in transmitting 2 *Theileria parva* parasite stocks from Kenya and Zimbabwe. *Parasitology* 116, 539-545.

Ochanda, H., Young, A.S., Mutugi, J. J., Mumo, J., and Omwoyo, P. L. (1998b). The effect of temperature on the rate of transmission of *Theileria parva* infection to cattle by its vector, *Rhipicephalus appendiculatus*. *Parasitology* **97**, 239-245

Ogden, N.H., Gwakisa, P., Swai, E., French, N.P., Fitzpatrick, J., Kambarage, D., Bryant, M. (2003). Evaluation of PCR to detect *Theileria parva* in field-collected tick and bovine samples in Tanzania. *Veterinary Parasitology* **112,** 177–1183.

Oliveira, R.P., Broude, N. E., Macedo, A. M., Cantor, C. R., Smith, C. L., Pena, S. D. J., (1998). Probing the genetic population structure of Trypanosoma cruzi with polymorphic microsatellites. *Proceedings of the National Academy of Sciences, USA.* **95**, 3776–3780

Olwoch, J. M., Reyers, B., Engelbrecht, F. A., & Erasmus, B. F. N. (2008). Climate change and the tick-borne disease Theileriosis (East Coast fever) in sub-Saharan Africa. *Journal of Arid Environments*, **72(2),** 108-120.

Orita, M., Suzuki, Y., Sekiya, T. and Hayashi, K. (1987). Rapid and sensitive detection of point mutations and DNA polymorphisms using the polymerase chain reaction. Genomics 5: 874–879.

Oura, C. A. L., Odongo, D. O., Lubega, G. W., Spooner, P. R., Tait, A., and Bishop, R. P., (2003). A panel of microsatellite and minisatellite markers for the characterization of field isolates of *Theileria parva*. *Int ernational Journal of Parasitology*, **33**, 1641–53.

Oura C.A., Bishop R., Asiimwe B.B., Spooner P., Lubega G.W. and Tait, A. (2004a). *Theileria parva* live vaccination: parasite transmission, persistence and heterologous challenge in the field. *Parasitology*, **134,** 1205–1213

Oura, C. A. L., Bishop, R., Wampande, E. M., Lubega, G. W., & Tait, A. (2004b). The persistence of component *Theileria parva* stocks in cattle immunized with the 'Muguga cocktail'live vaccine against East Coast fever in Uganda. *Parasitology*, **129(1),** 27-42.

Oura, C. A. L., Bishop, R., Asiimwe, B. B., Spooner, P., Lubega, G. W., & Tait, A. (2007). *Theileria parva* live vaccination: parasite transmission, persistence and heterologous challenge in the field. *Parasitology,* **134(9),** 1205-1213.

Paesen, G. C., Adams, P. L., Harlos, K., Nuttall, P. A., & Stuart, D. I. (1999). Tick histamine-binding proteins: isolation, cloning, and three-dimensional structure. *Molecular cell,* **3(5),** 661-671.

Palmer, B. H., McCarthy, J. F., Kozlik, A. & Harrison, I. R. (1971). A new chemical group of cattleacaricides. *Proceedings of the 3rd International Congress of Acarology, Prague* 1971, 687–691.

Pegram, R.G and Banda, D.S. (1990). Ecology and phenology of cattle ticks in Zambia. Development and survival of free-living stages. *Experimental Applied Acarology,* **8,** 291-301.

Pelle, R., Graham, S. P., Njahira, M. N., Osaso, J., Saya, R. M., Odongo, D. O & Bishop, R. P. (2011). Two *Theileria parva* CD8 T cell antigen genes are more variable in buffalo than cattle parasites, but differ in pattern of sequence diversity. *Public Library of Science (Plos) One,* ***6*(4),** e19015.

Perry, B.D., Lessard, P., Norval, R.A.I., Kundert, K. and Kruska, R., (1990). Climate, vegetation, and the distribution of *Rhipicephalus appendiculatus* in Africa. *Parasitology Today,* **6,** 100–194.

Perry, B. D. (1992). History of east coast fever in eastern and central Africa. In The Epidemiology of Theileriosis in Africa (ed. Norval, R. A. I., Perry, B. D. & Young, A. S.), pp. 41–62. Academic Press, London

Perry, B. D., Randolph, T. F., Mcdermott, J. J., Sones, K. R.& Thornton, P. K. (2002). Investing in animal health research to alleviate poverty. International Livestock Research Institute, Nairobi, Kenya.

Piesman, J. (1991). Experimental Acquisition of the Lyme Disease Spiroehete, Borrelia burgdorferi, by Larval Ixodes dammini (Acari: Ixodidae) During Partial Blood Meals. *Journal of medical entomology*, **28(2),** 259-262.

Powell, W., Machray, G.C., and Provan, J. (1996). **Polymorphism revealed by simple sequence repeats.** *Trends in plant science***. 1 (7),** 215-222

Purnell, R. E., Ledger, M. A., Omwoyo, P. L., Payne, R. C., & Peirce, M. A. (1974). *Theileria parva:* Variation in the infection rate of the vector tick, *Rhipicephalus appendiculatus*. *International journal for parasitology*, **4(5),** 513-517.

Radley, D. E., Brown, C. G. D., Burridge, M. J., Cunningham, M. P., Kirimi, I. M., Purnell, R. E., & Young, A. S. (1975a). East coast fever: 1. Chemoprophylactic immunization of cattle against *Theileria parva* (Muguga) and five theilerial strains. *Veterinary Parasitology*, **1(1),** 35-41.

Radley, D. E., Young, A. S., Brown, C. G. D., Burridge, M. J., Cunningham, M. P., Musisi, F. L., & Purnell, R. E. (1975b). East coast fever: 2. Cross-immunity trials with a Kenya strain of *Theileria lawrencei*. *Veterinary Parasitology*, **1(1),** 43-50.

Radley, D. E. (1981). Infection and treatment method of immunization against Theileriosis. In: Irvin AD, Cunningham MP, Young AS (eds) Advances in the control of Theileriosis: *Proceedings of an international conference held at ILRAD, Nairobi, 9–13 February, 1981*. Martinus NijhoV Publishers, The Hague, pp 227–236.

Rand, K. N., Moore, T., Sriskantha, A., Spring, K., Tellam, R., Willadsen, P., and Cobon, G. S., (1989). Cloning and expression of a protective antigen from the cattle tick Boophilus microplus. *Proceedings of the National Academy of Sciences, USA.* **86**, 9657-9661

Randolph, S. E., (2000). Ticks and tick-borne disease systems in space and from space. *Advances in parasitology*, **47,** 217-243.

Rassmann, K., Schl.tterer, C., Tautz, D., (1991) Isolation of simple sequence loci for use in polymerase chain reaction-based DNA fingerprinting. *Electrophoresis*, **12**, 113-118.

Rechav, Y,. (1981). Ecological factors affecting the seasonal activity of the brown ear tick *Rhipicephalus appendiculatus*. In: Tick Biology and Control, *Proceedings of an International Conference held from 27–29 January 1981, under the auspices of the Tick Research Unit, Rhodes University, G.B. Whitehead and J.D. Gibson (eds.). Rahamstown, South Africa*.

Rechav, Y. (1982). Dynamics of tick populations (Acari: Ixodidae) in the eastern Cape Province of South Africa. *Journal of Medical Entomology*, **19(6),** 679-700.

Rechav, Y and Kostrzewski, M.W. (1991). Relative resistance of six cattle breeds to the tick *Boophilus decoloratus* in South Africa. *Onderstepoort Jornal of Veterinary Research* **58,** 181–186.

Rechav, Y., Spickett, A. M., Dauth, J., Tembo, S. D., Clarke, F. C., Heller-Haupt, A., & Trinder, P. K. (1992). Immunization of guinea-pigs and cattle against adult *Rhipicephalus appendiculatus* ticks using semipurified nymphal homogenates and adult gut homogenate. *Immunology* **75(4),** 700-706.

Ribeiro, J.M., (1989). Role of saliva in tick/host interactions. *Experimrntal Applied Acarology* **7**, 15–20.

Rich, S. M., Hudson, R. P., Ayala, F. J., (1997). *Plasmodium falciparum* antigenic diversity: evidence of clonal population structure. *Proc Natl Acad Sci USA;* **94**, 13040–5.

Rodri´Guez, M., Rubiera, R., Penichet, M. L., Montesinos, R., Cremata, J., Falcon, V., Sa´ Nchez, P. A., Cordove´ S, C., Valde´ S, M., Lleonart, R., Herrera, L. & De La Fuente, J. (1994). High level expression of the B. microplus Bm86 antigen in the yeast Pichia pastoris forming highly immunogenic particles for cattle. *Journal of Biotechnology* **33**, 135–146.

Rotz, L.D., Khan, A.S., Lillibridge, S.R., Ostroff, S.M., Hughes, J.M., (2002). Public health assessment of potential biological terrorism agents. *Emerging Infectious Diseases* **8**, 225–230.

Roulston, W. J., Stone, B. F., Wilson, J. T. & White, L. I. (1968). Chemical control of an organophosphorus- and carbamate-resistant strain of Boophilus microplus (Can.) from Queensland. *Bulletin of Entomological Research* **58,** 379–392.

Samish, M., and Glazer, I., (2001). Entomopathogenic nematodes for the biocontrol of ticks. *Trends in Parasitology* **17**, 368-371

Sauer, J. R., McSwain, J. L., Bowman, A. S., and Essenberg, R. C., (1995). Tick Salivary gland Physiology. *Annua. Review of Entomology,* 40, 245-267

Sauer, J.R., Essenberg, R.C., Bowman, A.S., (2000). Salivary glands in ixodid ticks: control and mechanism of secretion. *Journal of Insect Physiology,* **46**, 1069–1078.

Schlötterer, C., and Tautz, D., (1992) Slippage synthesis of simple sequence DNA. *Nucleic Acids Research*, **20**, 211-215.

Scholtz, M.M., Spickett, A.M., Enslin, C.B. (1991). The effect of tick infestation on the productivity of cows of three breeds of cattle. *Onderstepoort Journal Veterinary Research* **58 (2),** 71–74.

Schuntner, C. A., Schnitzerling, H. J. & Roulston, W. J. (1972). Carbaryl metabolism in larvae of organophosphorus and carbamate-susceptible and -resistant strains of cattle tick *Boophilus microplus*. *Pesticide Biochemistry and Physiology* 1, 424–433.

Shapiro, S. Z., Voigt, W. P., & Ellis, J. A. (1989). Acquired resistance to ixodid ticks induced by tick cement antigen. *Experimental & applied acarology*, **7(1),** 33-41.

Shaw, M. K., and Young, A.S. (1994). The biology of Theileria species in ixodid ticks in relation to parasite transmission. *Advances in Disease Vector Reseach* **10,** 23-63.

Shaw, M. K., & Young, A. S. (1995). Differential development and emission of *Theileria parva* sporozoites from the salivary gland of *Rhipicephalus appendiculatus*. *Parasitology*, **111(2**), 153-160.

Shaw, R. D. (1970). Tick control on domestic animals (ii) The effect of modern treatment methods. *Tropical Science* **12,** 29–36

Shih, C. M., Telford, S. R., & Spielman, A. (1995). Effect of ambient temperature on competence of deer ticks as hosts for Lyme disease spirochetes. *Journal of clinical microbiology*, **33(4),** 958-961.

Shone, S. M., Dillon, H. J., Hom, S. S., & Delgado, N. (2006). A novel real-time PCR assay for the speciation of medically important ticks. *Vector-Borne & Zoonotic Diseases*, **6(2),** 152-160.

Short, N. J., and Norval, R.A., (1981). The seasonal activity of *Rhipicephalus appendiculatus* Neumann, 1901 (Acari: Ixodidae) in the highveld of Zimbabwe Rhodesia. *Journal of Parasitology*, **67,** 77-84.

Simmons, R. B., & Weller, S. J. (2001). Utility and evolution of cytochrome b in insects. *Molecular Phylogenetics and Evolution*, **20,** 196– 210.

Simon, C., Frati, F., Beckenbach, A., Crespi, B., Liu, H., and Flock, P. (1994). Evolution, weighting, and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Annals of Entomological Society of America*, **87,** 651–701.

Sonenshine, D.E., (1991). Biology of ticks (I). Oxford University Press, New York and Oxford.

Sonenshine, D.E., (1993). Biology of ticks (II). Oxford University Press, New York and Oxford.

Southern, E. M. (1979). Measurement of DNA length by gel electrophoresis. *Analytical biochemistry*, **100(2),** 319-323.

Speybroeck, N., Madder, M., Van Den Bossche, P., Mtambo, J., Berkvens, N., Chaka, G., Mulumba, M., Brandt, J., Tirry, L. and Berkvens, D. (2002). Distribution and phenology of ixodid ticks in southern Zambia. *Medical Veterinary Entomology,* **16(4),** 430-41.

Speybroeck, N., Madder, M., Thulke, H.H., Mtambo, J., Tirry, L., Chaka, G., Marcotty, T and Berkvens, D., (2004). Variation in body size in the tick complex *Rhipicephalus appendiculatus/Rhipicephalus zambeziensis*. *Journal of Vector Ecology,* **29(2),** 347-54.

Stagg, D. A., R. P. Bishop, S. P. Morzaria, M. K. Shaw, D. Wesonga, G. O. Orinda, J. G. Grootenhuis, D. H. Molyneux, and A. S. Young. (1994). "Characterization of *Theileria parva* which infects waterbuck (*Kobus defassa*)." *Parasitology* **108 (5),** 543-554.

Stanford, G. D., Baker, J. A. F., Ratley, C. V. & Taylor, R. J. (1981). The development of a stabilized amitraz cattle dip for control of single and multi-host ticks and their resistant strains in South Africa. In *Proceedings of a Conference on Tick Biology and Control* (ed. Whitehead, G. B. & Gibson, J. D.) pp 143–181. Rhodes University, Grahamstown, South Africa.

Storz, J. F. (2005). Using genome scans of DNA polymorphism to infer adaptive population divergence. *Molecular Ecology*, **14(3),** 671-688.

Sugino, M., Imamura, S., Mulenga, A., Nakajima, M., Tsuda, A., Ohashi, K., & Onuma, M. (2003). A serine proteinase inhibitor (serpin) from ixodid tick *Haemaphysalis longicornis* cloning and preliminary assessment of its suitability as a candidate for a tick vaccine. *Vaccine,* **21(21),** 2844-2851.

Sutherst, R.W., Maywald, G.F.M., Kerr, J.D., Stegeman, D.A. (1983). The effect of cattle tick (*Boophilus microplus*)on the growth of *Bos indicus* x *Bos taurus* steers. *Australian Journal of Agricultural Research* **34,** 317–327.

Tanaka, A. S., Andreotti, R., Gomes, A., Torquato, R. J., Sampaio, M. U., & Sampaio, C. A. (1999). A double headed serine proteinase inhibitor—human plasma kallikrein and elastase inhibitor—from< i> Boophilus microplus</i> larvae. *Immunopharmacology*, **45(1),** 171-177.

Tatchell, R.J., and Easton, E., (1986). (Tick) (Acari: Ixodidae) ecological studies in Tanzania. *Bulletin of Entomological Research*, **76**, 229–246.

Tauber, M.J., Tauber, C.A., and Masaki, S., (1986). Seasonal adaptations of insects. Oxford University Press, New York

Tautz, D., and Renz, M., (1984). Simple sequences are ubiquitous repetitive components of eukaryotic genomes. *Nucleic Acids Research,* **12**,4127–4138.

Tellam, R. L, Smith, D., Kemp, D. H., Willadsen, P., (1992). Vaccination against ticks. In: Wong WK, editor. Animal parasite control utilizing biotechnology. *Boca Raton: CRC Press;* pp 303–31.

Tourmente, S., Deragon, J. m., Lafleuriel, J., Tutois, S., Pelissier, T., Cuviller, C., Espagnol, M. C., AND Picard, G., (1994). Characterization of minisatellites in *Arabidopsis thaliana* with sequence similarity to the human microsatellite core sequence. *Nucleic Acids Research* **22,** 3317-3321

Trimnell, A. R., Hails, R. S. & Nuttall, P. A. (2002). Dual action ectoparasite vaccine targeting 'exposed' and 'concealed' antigens. *Vaccine,* **20**, 3360–3568.

Trimnell, A. R., Davies, G. M., Lissina, O., Hails, R. S., Nuttall, P. A. (2005). A crossreactive tick cement antigen is a candidate broad-spectrum tick vaccine. *Vaccine,* **23,** 4329–4341.

Uilenberg, G. (1999). Immunization against diseases caused by *Theileria parva*: a review. *Tropical Medicine & International Health,* **4(9),** A12-A20.

Varma M.G.R., Heller-Haupt A., Trinder P.K.E. & Langi A.O. (1990) Immunization of guinea pigs against *Rhipicephalus appendiculatus* adult ticks using homogenates from unfed immature ticks. *Immunology* **71,** 133-139.

Vignal, A., Milan, D., SanCristobal, M., & Eggen, A. (2002). A review on SNP and other types of molecular markers and their use in animal genetics. *Genetics Selection Evolution,* **34(3),** 275-306.

Vos, P., Hogers, R., Bleeker, M., Reijans, M., van de Lee, T., Hornes, M., Frijters, A., Pot, J., Peleman, J., Kuiper, M. (1995). AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research,* **23,** 4407–4414.

Walker, A. R., Fletcher, J. D., & Todd, L. (1990). Resistance between stages of the tick *Rhipicephalus appendiculatus* (Acari: Ixodidae). *Journal of medical entomology,* **27(6),** 955-961.

Walker, A. R., Bouattour, A., Camicas, J. L., Esatrada-Pena, A., Horak, I. G., Latif, A. A., Pergram, R. G., and Preston, P. M., (2003). Ticks of domestic animals in Africa. *Bioscience reports,* 42 Comiston Drive, Edinburg, UK

Walker, J. B. (1974). Ixodid ticks of Kenya.

Walker, J. B., Keirans, J. E., & Horak, I. G. (2000). The genus *Rhipicephalus* (Acari, Ixodidae): a guide to the brown ticks of the world. Cambridge University Press.

Wambura, P. N., Gwakisa, P. S., Silayo, R. S., & Rugaimukamu, E. A. (1998). Breed-associated resistance to tick infestation in *Bos indicus* and their crosses with *Bos taurus. Veterinary Parasitology*, **77(1),** 63-70.

Wang, H., Kaufman, W.R., and Nuttall, P. A., (1999). Molecular individuality: polymorphism of salivary gland proteins in three species of ixodid tick**.** *Experimental and Applied Acarology* **23,** 969–975

Warburton, C. (1912). Notes on the genus Rhipicephalus, with the description of new species, and the consideration of some species hitherto described. *Parasitology*, **5(1),** 1-20.

Ware, G.W. (2000). The Pesticide Book, 5th edn. Thomson Publications, Fresno, California. WHITEHEAD, G. B. (1958). A review of insecticide resistance in the blue tick, Boophilus decoloratus, in South Africa. *Indian Journal of Malariology* **12,** 427–432.

Wares, J. P. & Cunningham, C. W. (2001). Phylogeography and historical ecology of the North Atlantic intertidal. *Evolution* **12**, 2455–2469.

Weber, J. L., & May, P. E. (1989). Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *American journal of human genetics*, ***44*(3),** 388.

Weber, J. L., (1990). Informativeness of human (dC-dA)n(dG-dT)n polymorphisms. *Genomics,* **7,** 524-530

Weber, J. L., and May, P. E. (1989). Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *American Journal of Human Genetics*, **44**, 388-396.

Wenz, M.H., Robertson, J. M., Menchem, S., Oaks, F.,Demorest, D. M., Scheibler, D., Rosenblum, B, B., Wike, C., Gilbert, D. A., and Efcavitch, W. (1998). High precision genotyping using denaturing capillary electrophoresis. *Genome Research* **8,** 69-80

Wikel, S. K., (1996). Host Immunity to ticks. *Annu.Rev.Entomol.* **41,** 1-22

Wikel, S. K., and Alarcon-Chaidez F, J., (2001). Progress toward molecular characterization of ectoparasite modulation of host immunity. *Veterinary Parasitology* **101**, 275-287

Willadsen, P., & Kemp, D. H. (1988). Vaccination with 'concealed'antigens for tick control. *Parasitology Today*, **4(7),** 196-198.

Willadsen, P., (1997). Novel vaccines for ectoparasites. *Veterinary Parasitology* **71**, 209–222.

Willadsen, P., (2001). The molecular revolution in the development of vaccines against ectoparasites. Vet. *Parasitol.* **101**, 353–367.

Willadsen, P., (2004), Anti-tick vaccines. *Parasitology* **129**, S367–S387.

Willadsen, P., and Jongejan F., (1999). Immunology of the tick-host interaction and the control of ticks and tick-borne diseases. *Parasitology Today;* **15 (7),** 258–62.

Williams, J. G. K., Kubelik, A.R., Livak, K. J., Rafalski, J.A., Tingey, S.V. (1990). DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research*, **18,** 6531-6535

Wright, J. M., & Bentzen, P. (1995). Microsatellites: genetic markers for the future. In *Molecular genetics in fisheries* (pp. 117-121). Springer Netherlands.

Yeoman, G. H. (1966). Field vector studies of epizootic East Coast Fever. II. Seasonal studies of *R.appendiculatus* on bovine and non-bovine hosts in East Coast Fever enzootic, epizootic and free areas. *Bulletin of Epizootic Diseases in Africa*, **14,** 113–140.

Yli-Mattila, T., Paavanen-Huhtala, S., Fenton, B., & Tuovinen, T. (2000). Species and strain identification of the predatory mite *Euseius finlandicus* by RAPD-PCR and ITS sequences. *Experimental & applied acarology*, **24(10-11),** 863-880.

Young, A. S., & Purnell, R. E. (1973). Transmission ofTheileria lawrencei (Serengeti) by the ixodid tick, *Rhipicephalus appendiculatus*. *Tropical animal health and production*, **5(3),** 146-152.

Young, A. S., Brown, C. G. D., Burridge, M. J., Cunningham, M. P., Kirimi, I. M., & Irvin, A. D. (1973). Observations on the cross-immunity between *Theileria lawrencei* (Serengeti) and *Theileria parva* (Muguga) in cattle. *International journal for parasitology*, **3(6),** 723-728.

Young, A. S., Leitch, B. L., & Newson, R. M. (1981). The occurrence of a *Theileria parva* carrier state in cattle from an East Coast fever endemic area of Kenya. In *Advances in the Control of Theileriosis* (pp. 60-62). Springer Netherlands.

Young, A. S., Leitch, B. L., Dolan, T. T., Newson, R. M., Ngumi, P. N., & Omwoyo, P. L. (1983). Transmission of *Theileria parva* by a population of *Rhipicephalus appendiculatus* under simulated natural conditions. *Parasitology*, **86(2**), 255-267.

Young, A. S., Leitch, B. L., Newson, R. M., & Cunningham, M. P. (1986). Maintenance of *Theileria parva* infection in an endemic area of Kenya. *Parasitology*, ***93*(1),** 9-16.

Young, A. S., Dolan, T. T., Mwakima, F. N., Ochanda, H., Mwaura, S. N., Njihia, G. M., Dolan, R. B. (1995). Estimation of heritability of susceptibility to infection with *Theileria parva* in the tick *Rhipicephalus appendiculatus*. *Parasitology,* **111(1),** 31-38.

Zahler, M., Filippova, N. A., Morel, P. C., Gothe, R., & Rinder, H. (1997). Relationships between species of the Rhipicephalus sanguineus group: a molecular approach. *The Journal of parasitology*, 302-306.

Zane, L., Bargelloni, L., and Patarnello, T. (2002). Strategies for microsatellite isolation: a review. *Molecular Ecology* (2002). 1**1,** 1-16

Zhang, D.X. & Hewitt, G. M. (1997). Assessment of the universality and utility of a set of conserved mitochondrial primers in insects. *Insect Molecular Biology,* **6**, 143–150.

CHAPTER TWO: IDENTIFICATION OF A PANEL OF MICRO- AND MINISATELLITE MARKERS FOR *Rhipicephalus appendiculataus*

2.0 Abstract

Micro- and minisatellites are powerful tools for genetic research in genetic variation, population structure, mapping and analysis of quantitative traits. The isolation and characterization of such markers by screening DNA libraries is complex, expensive, time consuming and labor-intensive. However, satellite mining from nucleotide sequence databases, such as genomic and expressed sequence tag (EST) databases, is comparatively less laborious, fast and cheap. In this study, micro- and minisatellite markers of *R. appendiculatus* were mined from genomic sequences cloned in bacterial artificial chromosomes (BACs) and an EST database. PCR primer pairs for the satellite markers were designed and then optimised by PCR amplification. The EST database, comprising 6.26×10^6 bases from 7340 EST sequences, was screened using the Tandem Repeat Finder program, and 520 unique sequences were found to contain satellite repeat sequences. A total of 938 micro- and minisatellite markers were identified within the 520 unique sequences, representing a frequency of approximately 1 marker per 6.673 Kb. From a total of 3.608×10^5 bases from eleven BAC sequences, 150 micro- and minisatellite markers were identified, representing a frequency of 1 marker for every 2.405 Kb of genomic sequence indicating that repeats are three times more abundant in genomic sequences than in ESTs. From the total of 1088 micro- and mini-satellite markers identified, 66 markers were selected for further study based on (1) they had a repeat copy number of at least five, (2) they had flanking sequences that were suitable for designing unique PCR primers, and (3) the need to include both micro- and mini-satellite markers. Out of 66 PCR primer pairs, 42 pairs were able to amplify their respective markers from *R. appendiculatus* genomic DNA sample from the Muguga laboratory tick stock, and were selected for further study. The remaining 24 PCR primer pairs either gave a low yield of PCR product, or failed to give any product, and were thus discontinued from the study. The 42 EST-based micro and minisatellite markers identified through data mining from *R. appendiculatus* sequences are the first molecular markers available for studying the genetic diversity of this important tick vector. The markers represent the first step towards development of molecular tools to be applied to the genetic characterization of *R. appendiculatus*. Once validated, these newly developed markers can be useful for studying the population structure and understanding the genetic diversity of this tick as well as other closely related rhipicephaline tick species. Mining microsatellites from nucleotide databases was found to be an efficient way to identify molecular markers for diversity studies especially where reference genomes are unavailable.

2.1 Introduction

Variation at the DNA sequence level underlies the observed phenotypic differences between individuals and species. These variations are referred to as polymorphisms and they occur as point mutations, rearrangements, deletions, insertions or inversions involving blocks of several base pairs or as variations in nucleotide repeats. The sites or loci in the genome at which DNA variations reside are referred to as genetic, molecular or DNA markers. Several different types of DNA markers currently available for genetic analysis are found within multiple copies of simple repeated coding and non-coding DNA sequences arranged in arrays of differing repeat unit size and copy number. These repeated DNA sequences are an integral component of eukaryotic genomes and may comprise between 50-90% of total DNA in vertebrate and plant genomes (Armour *et al.*, 1999; Hancock, 1999).

Tandem DNA repeats consist of arrays of two to several thousand basic motifs that are arranged in a head-to-tail fashion. The frequent variation in the copy number of these repeats has led to these sequences being designated variable number tandem repeats (VNTRs) (Nakumura *et al.*, 1987). These hypervariable regions were first shown to be abundantly present in eukaryotic genomes (yeast and humans) (Hamada *et al*., 1982; Tautz and Renez, 1984). Minisatellite probes were then designed to detect repeats of 9-60 nucleotide units (Jeffreys *et al.*, 1985). Later, microsatellite markers consisting of 1-6 repeating nucleotide units were discovered (Lia & Luty, 1989; Tautz, 1989; Weber & May, 1989). Thus, VNTRs comprise of two related classes of loci, the minisatellites and microsatellites (Wright, 1994; Wright & Bentzen, 1995; Park and Moran, 1994).

Microsatellites (also known as simple sequence repeats (SSRs)) are defined as tandem repeat units ranging in length between 1-6 nucleotide bases (Weber, 1990; Tautz & Schlotterer, 1994; Goldstein and Pollock, 1997). Other authors, e.g. Armour *et al*., (1999) define microsatellites as 2–8 bp repeats, while Schlotterer (1998) refers to them as 1–5 bp repeat units. Microsatellites have been shown to be abundant in many eukaryotic genomes (Hamada *et al.*, 1982, Katti et al., 2001, Gemayel *et al*., 2010) as well as prokaryotic genomes (Coenye and Vandamme, 2005). Minisatellites are defined as blocks of repeat motifs, usually comprising10-60 bases, in tandem arrays of up to ~30 kb or more in length (Debrauwere *et al.*, 1997; Richard *et al.*, 2008). In this study, repeats of 1-9 nucleotide bases were regarded as microsatellites while those between 10-30 bases in length were regarded as minisatellites.

Microsatellites, which are commonly used to survey the distribution of genetic variation and gene flow among natural populations, have several advantages over other genetic markers. They are distributed throughout the genome (Toth *et al*., 2000), and occur in both coding and non-coding DNA sequences. In addition, they are highly polymorphic due to the variable number of copies of the repeated motifs (Hamada *et a.l*, 1982; Tauz & Renez, 1984; Hancock, 1999) and thus are highly poly-allelic. Variation in microsatellite length occurs primarily due to slipped-strand mis-pairing during DNA replication (Toth *et al.*, 2000; Katti *et al.*; 2001, Li *et al.*, 2002), and mutations of this type occur at a much higher frequency than do point mutations or insertions/deletions (Rossetto, 2001)). Another advantage is that microsatellite sequences that are found in non-coding intergenic regions are likely to be selectively neutral. Microsatellites are also preferred because they can be directly typed by PCR amplification, and yield reliable and reproducible results (Weber, 1990; Bruford & Wayne, 1993; Morgante & Olivieri, 1993). For these reasons, this class of markers has, over the last few years, become the method of choice for defining differences between individuals

in many eukaryotic genomes. They proved to be an extremely valuable tool for genome mapping in humans (Schuler *et al.,* 1996) and Zebrafish (Knapik *et al.,* 1998). They have been used extensively in population genetic analyses of *P. falciparum,* to define population structures, geographical sub-structuring and levels of inbreeding (Anderson *et al.,* 1999, 2000). They have been used to show the genetic diversity and population structure of *Trypanosoma cruzi* (Oliveira *et al.,* 1998) and *T. parva* (Oura *et al.,* 2003). Microsatellite markers have been applied show genetic polymorphism in *Rhipicephalu*s (*Boophilus) microplus* (Labruna *et al.,* 2009) and to dissect the genetic structure of amitraz-resistant and susceptible *R.* (*Boophilus) microplus* populations in Queensland, Australia (Cutulle *et al.,* 2009). Minisatellites have advantages similar to those of microsatellites. However, sometimes they are preferred because they can be easier to score due to their larger size.

Micro- and minisatellites however have certain disadvantages. The major drawback is that they need to be isolated *de novo* from most species being examined for the first time. The task of micro- and minisatellite isolation can be difficult in terms of effort and time because it traditionally consists of a long process of screening genomic libraries with appropriate probes (Rassmann *et al.,* 1991). Also, there are reports that for certain group of organisms microsatellites are difficult to isolate (Beaumont & Bruford, 1999), and the tendency of *Taq* polymerase to add a 3´ terminal thymidine to both ends of PCR products and to generate slippage products in some mono- and di-nucleotide microsatellite loci that can cause base shifts and sizing problems (Schotterer & Tauz, 1992).

Conventionally, micro- and minisatellite loci have been isolated from partial genomic libraries of the species of interest by screening several thousands of clones through colony hybridization with repeat containing probes (Rassmann *et al.,* 1991). A summary of this process is described in a review by Zane *et al.,* (2002). High quality genomic DNA is fragmented using restriction enzymes

and then size-selected to obtain small fragments (300-700 bp). DNA fragments are then ligated into a common plasmid vector. Transformation of bacterial cells with ligation product yields thousands of recombinant clones, which are subsequently screened for the presence of microsatellite sequences. Screening for positive clones is carried out by means of southern hybridization using repeat-containing probes. Repeat-containing probes can be synthesized *de novo*, or alternatively a genomic clone, which contains a microsatellite locus that has already been isolated, can be used. Hybridization probe(s) can be labeled by both radioactive (^{32}P, ^{33}P) or nonradioactive (e.g. digoxigenin) methods (Zane *et al.*, 2002). To avoid library construction and screening, modifications of the randomly amplified polymorphic DNA (RAPD, Williams *et al.*, 1990) approach has been used for the amplification of unknown microsatellites, by either using repeat-anchored random primers (Wu *et al.*, 1994) or using RAPD primers and subsequent Southern hybridization of polymerase chain reaction (PCR) bands with microsatellite probes (Cifarelli *et al.*, 1995; Richardson *et al.*, 1995). Other non-library PCR-based strategies rely on the use of repeat-anchored primers to isolate and then sequence one (Fisher *et al.*, 1996) or both regions (Lench *et al.*, 1996; Cooper *et al.*, 1997) flanking microsatellite repeats.

Although relatively simple, especially for microsatellite rich genomes, conventional isolation of satellites markers for the first time is complex, expensive and labor-intensive (Squirrel *et al*., 2003; Ellis and Burke, 2007; Zane *et al*., 2002). The process is extremely tedious and inefficient especially for species with low satellite frequencies (Zane *et al*., 2002). Therefore, several alternative strategies have been devised in order to simplify microsatellite isolation and to increase the efficiency of microsatellite isolation. Computer programs such as Tandem Repeat Finder (TRF) (Benson, 1999) have been developed to mine satellite markers from nucleotide database sequences

such as expressed sequence tag (EST) databases, whole genome databases, and also genomic sequences cloned in bacterial artificial chromosomes (BACs).

ESTs, which are fragments of mRNA sequences derived through single sequencing reactions performed on randomly selected clones from cDNA libraries (Parkinson & Baxter, 2009) have been shown to contain satellite markers (Ellis & Burke, 2007). By 2009, over 45 million ESTs had been generated from over 1400 different species of eukaryotes (Parkinson & Baxter, 2009). These EST projects were used to either complement existing genome projects or serve as low-cost alternatives for purposes of gene discovery. Since ESTs are derived from mRNA, they represent potential coding sequences. EST-based markers have been used in genetic studies of animals such as carps (Yue *et al*., 2004) and porcine (Rohrer *et al*., 2002) and plants such as bananas (Mbanjo *et al.*, 2013) maize, rice, sorghum and wheat (Kantety *et al*., 2002; Gupta *et al*., 2003), and shown to be powerful tools for genetic research in genetic variation, gene tagging and evolution, mapping and analysis of quantitative traits.

With high-throughput sequencing technology and an increased interest in genome-wide studies, the full genomes of more than 1000 organisms are now publicly available (NCBI; http://www.ncbi.nlm.nih.gov/guide/genomes-maps/). Screening genome databases for micro- and minisatellite repeats has been done for many organisms (Parkinson & Baxter, 2009, Kantety *et al*., 2002; Gupta *et al*., 2003). Mining satellites from nucleotide databases has several advantages. First, the process is less laborious and fast since there are several computer programs designed to quickly identify these markers. Secondly, the cost of data mining is very low because it avoids the expensive work associated with the initial steps of microsatellite development-namely, library construction and sequencing. Thirdly, markers obtained from coding sequences are derived directly

from gene expression and thus product identity and function can be identified by comparison with protein databases. This approach can integrate transcriptome studies and marker development into a single task and open avenues in linkage mapping, population genetics, and kinship analysis of species for which funding might be scarce. Fourth, EST-derived markers detect variation in the expressed portion of the genome and can therefore potentially detect marker-trait associations (Ellis & Burke 2007). Fifth, since these markers are from transcribed regions of the genome, they can be valuable for the investigation of population genetic phenomena in close relatives of species with existing genomic resources (Ellis & Burke, 2007). Indeed, these markers have been shown to be highly transferable across closely-related species (Perez *et al.*, 2005; Gupta *et al*., 2003; Thiel *et al*., 2003). Transferability means that the net cost per developed marker will be even lower if they are used for different species.

The primary limitation of using nucleotide databases such as ESTs and whole genomes as a source of molecular markers is that this approach relies on existing genomic resources, and suitable databases are often not available. Adding to this difficulty is the fact that in some organisms, only a fraction of all ESTs contain repeats (Perez *et al.*, 2005). An important concern with regard to use of EST markers is that, because they are derived from protein coding DNA, they may be subject to intense selection pressure. Selection on these loci may produce a distorted view of population structure, since selection rather than population history may determine the patterns of distribution of alleles within populations at these loci (Andersons *et al.*, 1999). However, a study by Woodhead and others (2005) showed that estimates of population differentiation (i.e F_{ST}) based on EST-markers were largely congruent with markers developed conventionally from non-coding regions.

For characterization of satellite markers, oligonucleotide primers designed from sequences flanking the repeat motifs are used in PCR reactions to amplify the repeat loci from genomic DNA (Holmes *et al.,* 1993). The amplified product is then resolved on gel electrophoresis on high resolution denaturing acrylamide gels (Litt & Luty, 1989; Weber & May, 1989). More recently, fluorescently-labeled PCR products have been analysed using capillary electrophoresis in automated genetic analyzers such as ABI 3730 (Applied Biosystems). The use of genetic analyzers has made accurate sizing of micro- and minisatellite alleles possible to a resolution of 1 bp. Currently, high-throughput genotyping with several markers is possible because the system offers the possibility of PCR multiplexing and post-PCR co-loading with markers that are labeled with different dyes that can be detected simultaneously. These systems offer automated data analysis and reproducibility in accurate sizing of micro- and minisatellite alleles, which are important for large-scale germplasm genotyping projects (Wenz *et al*., 1998). The use of these systems has removed the earlier limitations of conventional gel based systems such as exposure to toxic chemicals, low throughput and inaccurate sizing of alleles.

The creation of a gene index of sequences expressed in the salivary gland of *R. appendiculatus* (Nene *et al*., 2004; http://compbio.dfci.harvard.edu/cgi-bin/tgi/gimain.pl?gudb=r_appendiculatus), comprising 7340 non-redundant nucleotide sequences and the construction of *R. appendiculatus* BAC library (Sunter *et al*., 2008) provided a nucleotide sequence resource that allowed the development of a comprehensive panel of micro- and minisatellite markers for *R. appendiculatus*. This chapter describes the mining and identification of micro- and minisatellite markers from the RaGI EST database and some genomic sequences cloned in bacterial artificial chromosome (BACs). These molecular markers are the first genetic markers available to study population structure and genetic diversity of this important tick vector.

2.2. Materials and methods

2.2.1 Identification of *R. appendiculatus* micro- and minisatellite markers

The *R. appendiculatus* salivary gland gene index (RaGI EST database) (http://compbio.dfci.harvard.edu/cgi-bin/tgi/gimain.pl?gudb=r_appendiculatus (Nene *et al*., 2004) was screened for micro- and minisatellite tandem repeat sequences using the Tandem Repeat Finder program v4.0 (http://tandem.bu.edu/trf/trf.html; Benson, 1999). A total of 6.26 x 10^6 bases from 7340 non-redundant ESTs sequences, comprising of 2543 TC sequences and 4797 singletons, were screened. Eleven *R. appendiculatus* bacterial artificial chromosome (BAC) sequences with a total of 3.608 x 10^5 bases -constructed by the commercial company Amplicon Express (Pullman WA USA) (Sunter *et al*., 2008), were also screened for satellite markers (Table 3). Tandem Repeat Finder has two main components: detection and analysis components. The detection component uses a set of statistically-based criteria to find candidate tandem repeats. The outputs included the nucleotide sequence of each repeat unit, the size of the repeat unit, and the copy number (approximate number of tandem repeat units per locus). Several statistics derived from the alignment of each repeat sequence such as percent identity, percent of repeat units having insertions and deletions, nucleotide sequence composition and entropy measure were also obtained. Tandem repeat sequences ranging in size from 2-30 bp with copy numbers of five or greater were selected. Statistics about the alignment of the micro- and minisatellite repeat sequence such as percent identity, percent indels and the nucleotide sequence (composition, entropy measure) were extracted. Sequences containing poly (A) tails or tandem repeats with less than 30 bases of non-repeat sequence flanking either end of the EST sequence were excluded from further analysis.

2.2.2 Design of primer sequences flanking individual micro- and minisatellite markers

Repeat sequences of 2-30 bp in length with copy number of at least five were selected. Markers were selected for further study based on (1) they had a repeat copy number of at least five, (2) they had flanking sequences that were suitable for designing unique PCR primers, and (3) the need to include both micro- and mini-satellite markers. PCR primers were designed from the unique non-repeat flanking sequences using Primer3 software (v4.0) (http://frodo.wi.mit.edu/primer3/). In the software, the length of the required amplicons was set to 100-500 bp. Oligonucleotide parameters were set to a length of 18-24 bp with an optimum of 20 bp, a GC content of 45%-80% with an optimum of 50%, a melting temperature (Tm) of 50-60°C with an optimum of 55°C, and a primer Tm maximum difference of 1°C between the forward and the reverse primers. From a list of nearly 500 potential EST-SSR markers, 62 primer pairs were designed from the first 62 ESTs that had a flanking region long enough to generate suitable primer sequences and the expected product size of ≤400bp. Another four primer pairs were designed from four BAC sequences. Primers were synthesized by Bioneer Corporation (South Korea). They were reconstituted to 100 μM (100 pmol/μl) in TE (pH 7.6). The primers were further diluted to a working concentration of 25 μM. Primer stocks were stored at -20°C.

2.2.3 Extraction of whole tick genomic DNA

Whole genomic DNA was extracted from the Muguga laboratory ticks preserved in 70% ethanol using QIAGEN® DNeasy® Blood & Tissue Kit (QIAGEN GmbH-Germany) with a slight modification of the manufacturer's instructions. A shaking water bath was preheated at 56°C. Ticks stored in 70% ethanol had the ethanol drained, individual ticks removed, placed in a 15 ml falcon tube and washed twice in triple distilled water by vigorous shaking. Fresh live laboratory ticks were washed in 70% ethanol first and then twice in triple distilled water by vigorous shaking. The ticks

were left on a paper towel to dry for 5 min at room temperature. By use of a pair of forceps, individual ticks were placed into a 1.5 ml microcentrifuge tube pre-cooled in liquid nitrogen. Using another microcentrifuge tube, a few drops of liquid nitrogen were added and the tick ground using a sterile plastic pestle (Bel-Art products, USA). To the tick powder, 180 µl Buffer ATL were added followed by 20 µl Proteinase K. The contents were mixed thoroughly by vortexing and incubated at 56°C overnight in a shaking water bath. The solution was vortexed occasionally during the incubation period. After an overnight incubation, the tubes were removed from the shaking water bath, vortexed thoroughly and 200 µl Buffer AL added. The mixture was mixed thoroughly by vortexing and then incubated in a shaking waterbath at 70°C for 10 min, after which, 200 µl of absolute ethanol (96–100%) was added and the contents mixed thoroughly again by vortexing. All the mixture was then transferred into a DNeasy Mini spin column in a 2 ml collection tube and centrifuged at 8000 rpm for 1 min. The flow-through was discarded. The spin column was placed in a new 2 ml collection tube and 500 µl Buffer AW1 added and the contents centrifuged for 1 min at 8000 rpm. The flow-through was discarded. The spin column was placed in a new 2 ml collection tube and 500 µl Buffer AW2 added, then centrifuge for 3 min at 14,000 rpm. The spin column was then removed carefully so that it did not come into contact with the flow-through. The flow-through and collection tube were discarded. The spin column was placed in a new collection tube, then centrifuged for 1 min at 14,000 rpm to remove excess residual buffer AW2. The spin column was then placed in a new1.5 ml microcentrifuge tube, 60 µl Buffer AE added directly onto the DNeasy membrane for elution and then incubated at room temperature for 10 min. The microcentrifuge tube holding the spin column was then centrifuged at 8000 rpm for 2 min to obtain the DNA eluate. DNA samples were then aliquoted and stored at -20 °C.

2.2.3.1 Determination of nucleic acid concentration and purity

To determine concentration and yield of tick genomic DNA, 1.5 µl of gDNA sample from an individual tick was pipetted and used for spectrophotometric measurement of absorbances at 260 nm and 280 nm using the Thermo Scientific NanoDrop™ 2000 UV Spectrophotometer (Thermo Fisher Scientific Inc., USA). For analysis of the yield and quality of gDNA, 2 µl aliquot of genomic DNA from each sample was resolved on 1% (w/v) agarose gel in 1X TAE buffer, alongside Lambda DNA size marker. Electrophoresis was carried out in 1X TAE buffer at 50 V for 40 min. The gel was then stained with 200 ml of Ethidium Bromide (0.05µg/ml) solution on a rotating shaker for 25 min before being destained using distilled water on a rotating shaker for a further 25 min. The gel was visualized on a UV light box and images taken using a Kodak Polaroid Camera.

2.2.4 PCR optimization of primers flanking individual micro- and minisatellite loci

Primer pairs of the micro- and minisatellite markers were optimized by amplification of purified *R. appendiculatus* genomic DNA sample from the Muguga laboratory tick stock and subsequent visualization of the PCR product on a 2% agarose gel stained with 10 mg/ml ethidium bromide. Standard PCR conditions were used to optimize annealing temperature. The amplifications were performed in 25 µl reaction volumes consisting of 1x Promega colourless Taq DNA polymerase buffer, 1.5 mM $MgCl_2$, 0.2 mM dNTPs, 0.01 U Taq polymerase (Promega), 0.25 µM of each primer and 25 ng of template DNA. Reaction conditions comprised an initial denaturation at 95°C for 5 min followed by 35 cycles of denaturation (94°C) for 30 s, primer annealing (55-61°C) for 30 s, and primer extension (72°C) for 60 s. The final primer extension step was carried out at 72°C for 10 min. After PCR amplification, 2µl of PCR product was resolved by electrophoresis through 2 % (w/v) agarose gel in 1x TAE buffer (pH 8.0) for 30 min at 100 V. DNA was visualised under UV light following ethidium bromide staining of the gel, and the PCR product sizes were estimated by

comparison with a 100 bp DNA ladder (Promega). Further optimization was achieved by repeating the PCR reaction to determine the optimum annealing temperature.

2.3 Results

2.3.1 Identification of informative micro- and minisatellite markers

A total of 6.26×10^6 bases derived from 7340 non-redundant ESTs sequences (from salivary gland mRNA) comprising of 2543 Tentative consensus (TC) sequences and 4797 singletons were screened for satellite repeats. Five hundred and twenty unique sequences, comprising of 222 TC and 298 singleton sequences, were found to contain repeat sequences. This represented 7.1 % of the total non-redundant sequences. A total of 938 repeats were identified resulting in a frequency of at least 1 repeat marker per 6.673 Kb of the ESTs. The mean number of repeats in the 222 TC sequences was 2.2 while it was 1.48 in the 298 singletons. The highest number of repeats found in a single TC sequence (length = 1178) was 17, while it was 6 in a singleton sequence (length = 895). In the EST sequences it was observed that mononucleotide repeats represented the shortest repeat unit while the longest was 456 bases. The average repeat motif size was 37 bases and some of the repeated sequences occurred as compound repeats. Markers with 18 nucleotide repeats were most abundant while the average copy number of the repeats was six.

From a total of 3.608×10^5 bases from the eleven BAC sequences, 150 tandem repeats were identified representing a frequency of 1 repeat for every 2.405 Kb of genomic sequence. The highest number of repeats identified from a single BAC sequence was 31 (Table 3). The mean copy number of repeats in the BAC sequences was 19 while the mean repeat motif size was 17 bases. The longest repeat motif was 172 bases with the smallest motif being 1 bp.

Sixty two EST markers and four markers from RAHD_48 BAC clone were considered for design of primer sequences. These consisted of 42 microsatellites and 24 minisatellites. Table 4 summarises the statistics of the 7340 EST and 11 BAC sequences used to screen for presence of satellite repeats.

Table 3: List of BACs sequences, their length and the number of identified repeats

BAC sequence number	**Sequence identification number**	**Length (bp)**	**Number of Repeats**
1	RAHD_48	65414	31
2	RAHD_50	2259	1
3	RAHD_87	84879	35
4	RAHE_5	6177	1
5	RAHE_71	90515	31
6	RAHF_73	23722	7
7	RAHF_74	12314	11
8	RAHF_77	1844	1
9	RAHF_78	9300	1
10	RAHF_80	57084	28
11	RAHF_81	7295	3
Total		**360803**	**150**

Table 4: Summary statistics of nucleotide sequences screened for presence of micro- and minisatellite repeats

Source of nucleotide sequence	**ESTs**		**BACs**
	No. of TCs	No. of singletons	
	2543	4797	
Total number of sequences analysed	7340		11
Total number of bases	6.26×10^6		3.608×10^5
Number of sequences with repeats	TCs	Singletons	11
	221	298	
Total number of identified repeats	938		150
No. of sequences input into Primer 3	85		8
Number of sequences with suitable primers*	62		4
Number of primers pairs synthesized	62		4
Number of primer pairs that gave a PCR product	42		0

*Primer3 generated at least five primer pairs for each of these sequences but only one pair was optimized

2.3.2 Analysis of extracted whole tick gDNA

High quality tick gDNA was required for optimization of the designed primers pairs flanking micro- and minisatellite markers. Thus, the quality of gDNA extracted from 46 *R. appendiculatus* Muguga laboratory ticks preserved in 70% ethanol and kept at 4°C was analysed on agarose gel. Preserving ticks in alcohol and storing them in the refrigerator has been shown to be an effective method of preservation of gDNA which yields high quality DNA upon extraction (Mtambo *et al.*, 2006). Figure 7 below shows the resolution of the obtained gDNA. Some of the samples had sheared DNA shown by appearance of a strong smear at low molecular weight sizes.

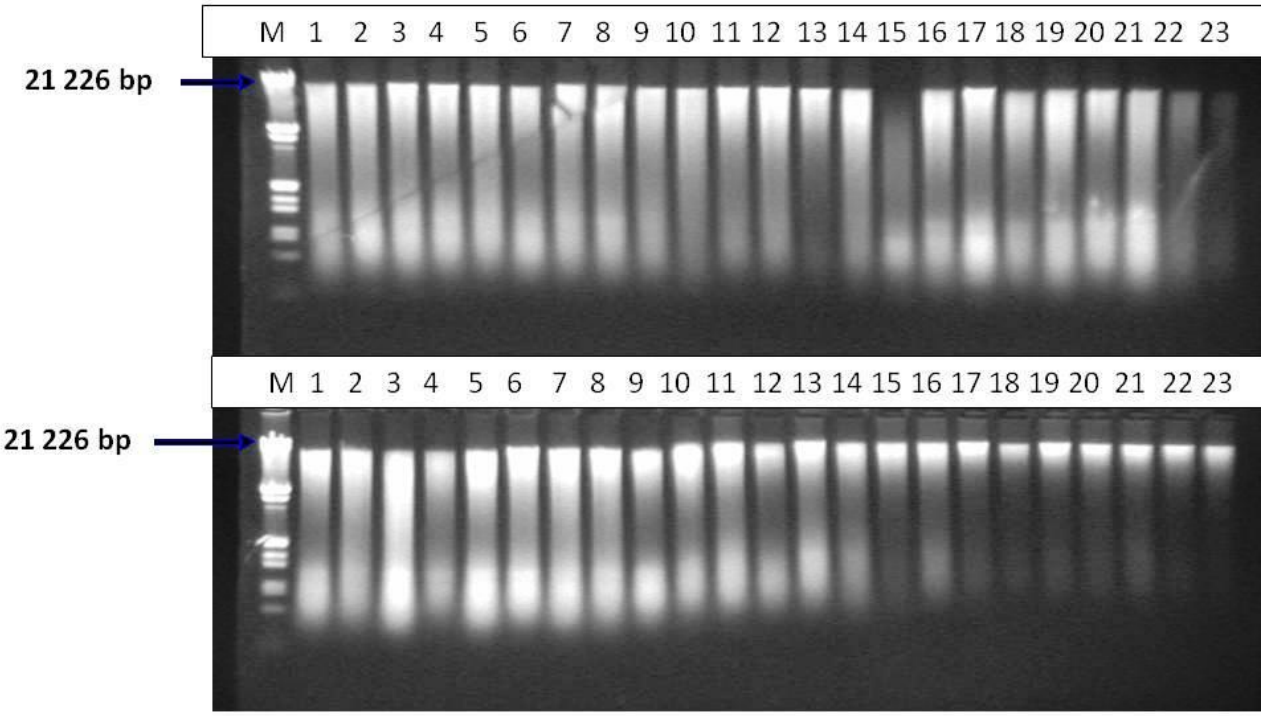


Figure 7: Agarose gel analysis of whole tick genomic DNA extracted from 46 *Rhipicephalus appendiculatus* Mugaga ticks.
Five µl of each sample were resolved on 1% TAE agarose gel. The size of the gDNA was compared to high molecular weight Lambda DNA/EcoRI + HindIII Markers DNA markers (Promega).

2.3.3 Primary primer screening and PCR conditions optimizations

Of the sixty six primer pairs designed, 42 pairs were able to amplify their respective markers from *R. appendiculatus* genomic DNA sample from the Muguga laboratory tick stock, and were selected for further study. These consisted of 28 microsatellites and 14 minisatellites. Twenty four PCR primer pairs either gave a low yield of PCR product, or failed to give any product, and were thus discontinued from the study. These consisted of 14 microsatellites and 10 minisatellites. Table 5 shows a summary of the characteristics of the 66 micro- and minisatellite markers identified, their primers sequences and the predicted PCR product size in base pairs. The last 24 markers in the list which failed to amplify were removed from the study. Figure 8 shows the annealing temperature (from 54-59°C) optimizations for four markers. Further validation of the optimization was done to confirm the right annealing temperature of each of the forty-two primer pairs that showed good amplification (Figure 9).

Table 5: Sixty-six micro- and minisatellite markers identified in this study and the PCR primer sequences for each marker

	Satellite name[a]	Repeat pattern	Sequence Length[b]	Position of repeat	Repeat size (bases)	Repeat Copy no.	Forward Primer 5'-3 (bases)	Reverse Primer 5'-3' (bases)	Predicted PCR product size (bp)
1	TC2392	TA	918	578-615	2	19.5	CTGCACTGTTGCCATGTA (18)	CAAGGAGGAAAAGGAACG (18)	160
2	TC2448	AT	1086	45-72	2	14	TGTGTGCCACGCAATAAC (18)	CATAAAGGAGGCACAGGAC (19)	126
3	TC705	CA	806	208-237	2	15	GGGAGCACTTTGGAACAA (18)	GGCCTTCTACTGTTCAAGC (19)	164
4	CD788726	AT	784	187-263	2	39	AAATGTCCACGGAACACC (18)	ATTGTGTGCTGACCTCTGTG (20)	129
5	CD790649	AC	903	397-421	2	12.5	CAGGATGAAACCTCAGTG (18)	CAACTGGGTGTACTTGCTACTG (22)	172
6	TC1002	GCA	905	504-552	3	16.3	GCAGTCTGGACCCTGGAAG (19)	GTCTTGGGTGGTGGTGCTT (19)	138
7	TC1227	AGC	940	445-478	3	11.3	CAGTTCCCTGGGCAGTTAT (19)	GTGTACTGGCTCCTTGACAG (20)	172
8	TC1307	ACA	1695	1494-1524	3	10.3	GATCAGCAGACCGCAAAG (18)	GCAACGACATTCAGTGGTG (19)	112
9	TC1312	ACA	803	621-651	3	10.3	CAATCAGAAGACCGGAAAGG (20)	GACATTGAGTGGCTTCGTC (19)	103
10	TC1716	TTC	1222	423-453	3	10.7	GTCCCAAGGTGAACACAGTA (20)	GACCTCTCTGACCTGCTATTTC (22)	117
11	TC2092	CTG	1544	266-349	3	28	AGTCAATATCGGGATCTCGG (20)	ACATTCGCAACCTTCTGG (18)	197
12	TC2543	CAG	846	202-256	3	18.3	GCTCCTTGAGAAGATGAAGC3 (20)	AGCAGCGCTTTCTGGTAT (18)	111
13	TC440a	TGC	2167	655-707	3	17.7	CTGCAGTGTCCTCGCATT (18)	GGACACTCAGCAACAAGC (18)	152
14	TC875	GCC	820	272-304	3	10.7	CTAGGAACCGCTGCTGTG (18)	CTGATGCCATTCTTCGAC (18)	171
15	CD788386	CAG	856	236-289	3	18	ACCTCAGGGTGGTATGAAC (19)	CATCATGCCTGTTGGATG (18)	160
16	TC2482	TGGA	670	285-327	4	10.8	CATGTCCGTGTCCGTATTAG (20)	TCCAAGTTACGGCTAGGTG (19)	164
17	CD788461	GTCT	814	443-501	4	14.8	TCGGTCTCTCTCTCTCTCTCT (21)	AGCTAGCTGAATCGTGTGAC (20)	162
18	TC1287	AAAGA	622	279-318	5	7.8	CGCGATTCATACCGCAAT (18)	CGCTGGGTTGCTGTTTCTC (19)	154
19	TC1353	CACCGC	899	301-334	6	5.7	GAGAGACGTCACGGAAACC (19)	GAAGACTCCGTAGTTCTAGCC (21)	102
20	TC1301	CTCCTGGAC	900	524-606	9	9.2	CAGGCTCACCGAACATAATC (20)	GAGTTGGGCCAGGTCTTG (18)	189
21	TC1302	CCAGGAGGC	967	299-404	9	11.8	CAGGTGGCGTAGGAATAGG (19)	CAGGCTCACCGAACATAATC (20)	159
22	TC1303	CCAGGAGGC	1537	305-392	9	9.8	CAGGTGGCATCGGAATAG (18)	CAGGCTCACCGAACATAATC (20)	141
23	TC1304b	CTCCTGGGC	1523	1129-1225	9	10.8	CAGGCTCACCGAACATAATC (20)	CAGGTGGCGTCGGAATAG (18)	149
24	TC1356	ACCCATGCC	1274	617-668	9	5.8	GTGGTTGCTGTTCTCTTCAC (20)	GCTGTTGTTGTGGAGCTT (18)	115
25	TC1398a	TGGTCTCGG	1163	313-363	9	5.7	CTCTCTGGTCTCTCTGGACTC (21)	GTATCCACCGAGCAGACC (18)	174
26	TC2077	GCTGCTGCA	1535	556-632	9	8.9	CTCAGCTCCTCCCATTTG (18)	TCCATTGATGTGGTCGAG (18)	176
27	TC2381	GTATCCACC	803	320-375	9	6.2	GATACCCATAGGTGCCACT (19)	GTACGGTGCGTACAATCC (18)	110

	Satellite name[a]	Repeat pattern	Sequence Length[b]	Position of repeat	Repeat size (bases)	Repeat Copy no.	Forward Primer 5'-3 (bases)	Reverse Primer 5'-3' (bases)	Predicted PCR product size (bp)
28	TC728b	CAGCAGCAG	785	399-546	9	16.4	GATGGATATGGGGGTAGTTC (20)	GAATGGGGTTGGTAGTATCC (20)	191
29	TC1968	TTGGATTCAGCC	1447	310-391	12	6.8	GAGAGGAGCTCGAAGTTTG (19)	AACTGCTGCTACGGAAGAG (19)	180
30	TC639a	CCGCTACCACCA	1494	338-431	12	7.8	GCTACCTTCACTCCTGCTT (19)	AGCGGTACAGGAGGACAT (18)	168
31	TC919	AATGCCAGCGCC	1313	805-872	12	5.7	AGCAGCCTAACGTGGAAG (18)	ATTGCTCCTGGTCTTTGCT (18)	151
32	TC1179	GCCAGCAATGCAGAC	775	112-264	15	10.2	CACAGACACCTCAGCCAAC (19)	CAGCACGGTGGTAGTTTG (18)	272
33	TC1284a	GGCAGCAGTGGCTAC	852	132-202	15	4.7	GGCTTTGGTGGAGACTTT (18)	GCTGCTACCTCCTCCAAG (18)	141
34	CD788063	TGATGAAGACCTAGA	442	265-348	15	5.6	CTTGGATGAGGAAGACGAC (19)	CATCCTCATCTTCGGACTC (19)	170
35	CD788164	GCCCAACAGCAAATG	895	280-365	15	5.7	ACAGCAGTATGCAGCACAG (19)	CCAAGTTCGGTTGCTGAG (18)	146
36	TC1026	GGACCGCGAGCGCATGCT	908	455-556	18	5.7	GAACCGAGGTCGGAAACTAC (20)	CATAGAAGTCGCGATCCAAC (20)	182
37	TC856	GAGGATCTCATGCCGCCTCT	1010	859-985	20	6.3	ATCTCTGTGGTCATGCAGTC (20)	TGTCTTTCTCGGCACTTG (18)	195
38	TC1398b	GGATTCGGCAGCCCACTCAGC	1163	419-685	21	12.7	GGCTACCCTGGTCTCTTTG (19)	CATGGACAGGGGACCGTAG (19)	338
39	TC1399b	GGATTCGGCAGCCCACTCAGC	918	388-633	21	11.7	GCTCTCGCTACGGAGGTC (18)	ATGGACAGGGGACCGTAG (18)	298
40	TC1400b	GGATTCGGCAGCCCACTCAGC	1129	397-684	21	13.7	5CCTGGTCTCTTTGGCTCTC (19)	ATGGACAGGGGACCGTAG (18)	353
41	TC1284b	TTCGGTGGCCTTGGAGGAAGCAGC	852	204-302	24	4	GCTTTGGTGGAGACTTTG (18)	CACTGAGGCCACTGCTTC (18)	213
42	TC23	TCCTCCGGACGGGTAGCTAGACAA	1650	190-453	24	11.1	CTTCCTGAACTCCAAGAGC (19)	GACTAACAAGCGGAGGTTAC (20)	319
Markers removed from the study									
43	TC2322	TA	954	692-719	2	14	GACAGCATGTAGTGGTTGC (19)	ACAGAGACACTTGTCAGACCTC (22)	207
44	TC267	TA	795	70-124	2	29	ACGCGCAAGAGATAGGTC (18)	TATCTCTCTCTCTCACCGACAC (22)	132
45	TC962	AT	928	770-806	2	18.5	CCCTACTGCCTGTCTAAATG (20)	GGGAGCCAGTAAAGAAGC (18)	113
46	RAHD48a	AC	65414	4334-4437	2	52	GCCCTAATGTGTTGGAGC (18)	CGAGCAAGTCCTCTGATTTC (20)	270
47	TC2121	CAG	1209	564-630	3	22.7	GTCTATGAGGGCACTCTCTGT (21)	CGAAAATCTCCTTGGAGCAC (20)	178
48	TC340	CTT	1445	38-74	3	12.3	CAGTCTCTCAGACTCGTCACT (21)	CCAGTCAAGCAAAGGAGAG (19)	116
49	TC2281a	ATGT	1134	157-223	4	16.8	GCATTCCACGGTTCTTAC (18)	CGGTTGCCCTTCTGTAAT (18)	151
50	RAHD48b	ATCT	65414	37751-37994	4	61	GTCGTTACAGAGAGAGAGTGTG (22)	CCGTTCGTAGCACAAATC (18)	443
51	RAHD48c	ACAC	65414	48713-49022	4	77.5	GTTGTAAGGAGGACACAGAGAG (22)	CCTCAGAACGGCTTTGAT (18)	393
52	RAHD48d	ATAC	65414	57495-58034	4	107.3	GCTGCTTGAACAGGCTTCT (19)	GATGCCAGAGTCTCCTGTTTC (21)	455
53	TC580	AAGGAA	1215	307-369	6	10.5	AGAGCATAGGGATGACCAC (19)	AGAGGAAGACGAGGCACT (18)	154
54	CD788906a	CACACA	852	26-192	6	27.8	CGGTTCCAACAGAAGTGC (18)	GGTACCTGGTTGACATCG (18)	235

	Satellite name[a]	Repeat pattern	Sequence Length[b]	Position of repeat	Repeat size (bases)	Repeat Copy no.	Forward Primer 5'-3 (bases)	Reverse Primer 5'-3' (bases)	Predicted PCR product size (bp)
55	TC2003	TATACATA	1126	780-826	8	6	TCGTATGCCTGTTGTGTG (18b)	ACGCTCACGAGTTCACAT (18)	162
56	TC307	TCTTCCTG	1667	1190-1248	9	6.9	CCTTCTTGGCCTGCTTTT (18)	AGGACGAAGAGGATGAGG (18)	150
57	TC2265	CTCGGTGGTGGC	676	344-435	12	7.7	CTACGGTAGTGGAAGCTACG3 (20)	CTTCCGAAACCAGATGAG (18)	234
58	CD787470	CCCCGACAACTA	710	214-283	12	5.8	GGCGTCGACTACCACAAC (18)	CTTTTTGGTCACGGACACT (18)	151
59	TC2298a	CAGCGGGGCCAGGAC	756	274-510	15	15.8	CTAGGAACAAACCCAGTGTC (20)	CCTCGTTTGTCCTTGACC (18)	354
60	TC1268	GGAGGCTTTGGTGGCCTA	2424	157-253	18	5.4	CAGGGCTGGTTAGTTATGG (19)	CTCCGAGACCGCTGTAAC (18)	156
61	TC1285b	GGCTTCGGCGGATCTAGC	1455	1158-1243	18	4.8	GGTGGAGGTTATGGTGGA (18)	GGATCCACCATAACCTCCAC (20)	153
62	TC22	GGATGGTCTAGCTACCCGTCTGGA	1698	1264-1472	24	8.7	AGGTAGCTATGGTGGAAGC (19)	CCAGAGCTTCCTCTGGAT (18)	276
63	TC110a	AGGCCACTGCCATAGCTGCCGTAGCCTCCA	1778	551-763	30	7	GTAACCACCGTAGCCGAGAC (20)	GCTACGGCAGTTATGGCAG (19)	320
64	TC110b	CCATAGCCAAGACCACCAAGCCCACTGCTG	1778	1046-1305	30	8.5	CGAAGACTGCCGTAGCTG (18)	GTCTCGGCTACGGTAGCT (18)	319
65	TC110c	AGACCACCAAGCCCACTACTGCCATAGCCG	1778	1055-1361	30	9.9	CGAAGACTGCCGTAGCTG (18)	GTGGACTTGGCAGCTACG (18)	406
66	TC111	GGAGGCTACGGCAGCTATGGCAGTGGCCTC	1613	1195-1407	30	7	GTAGCGGTCTTGGAGGTT (18)	CCAAATCCACTGCCATAG (18)	406

[a] Satellite names were derived from the names of the sequences from which they were derived (TC: Tentative consensus sequence; CD: Singletons; RAHD: BAC sequence)

[b] Length of TC, singleton or BAC sequence in bases

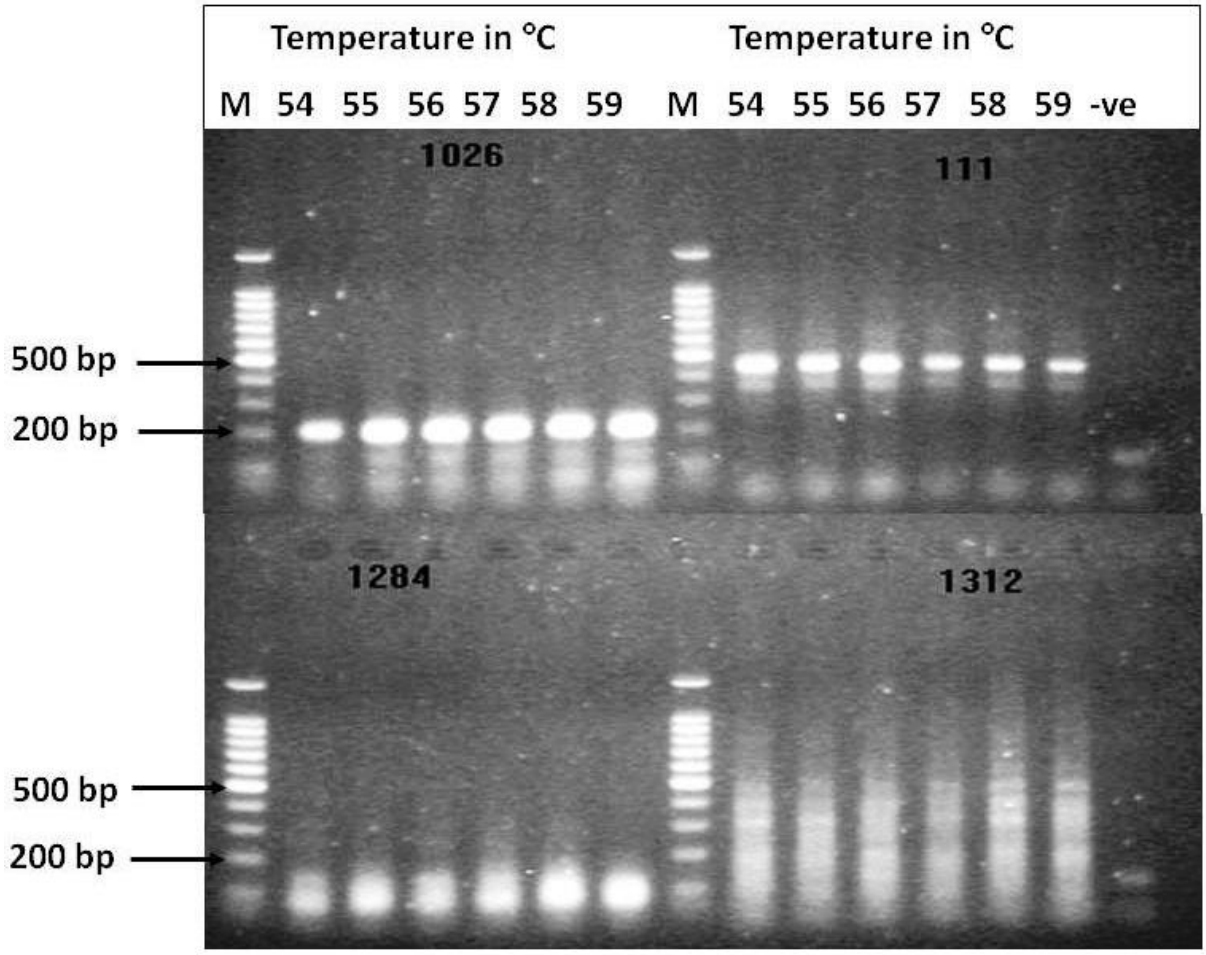


Figure 8. Agarose gel showing PCR products of four *R. appendiculatus* micro- and minisatellite markers in a range of annealing temperatures

The markers are TC1026, TC111, TC1284 and TC1312. A negative control sample denoted as (-) was included and contained water instead of the template gDNA. Optimization was done using a temperature gradient from 54 to 59°C. In the PCRs the concentration of $MgCl_2$ was 1.5 mM and the amount of template gDNA was 25 ng. Markers TC1284 failed to amplify while TC1312 showed non-specific amplification.

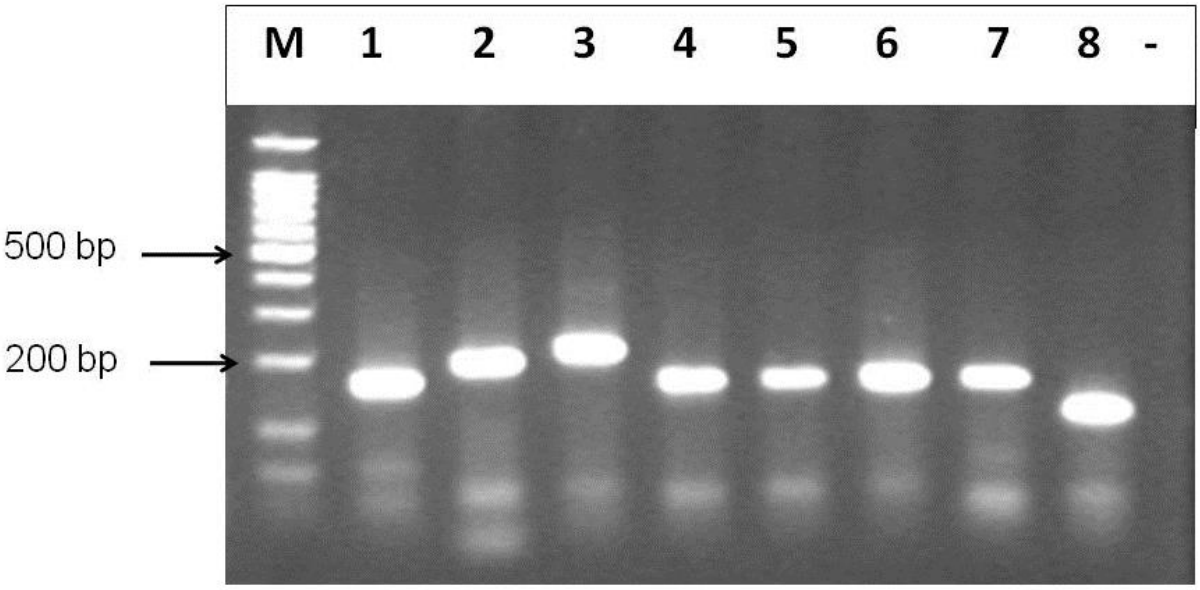


Figure 9: PCR products of eight micro- and minisatellite markers loci in *R. appendiculatus* following annealing temperature optimizations
The markers are 1: TC1002, 2: TC1227, 3: TC1301, 4: TC1302, 5: TC1303, 6: TC1304, 7: TC1287 and 8: TC1353. A negative control sample denoted as (-) was included and contained water instead of template gDNA. The amplification was carried out at an optimum annealing temperature of 58°C for all the markers. In the PCRs the concentration of $MgCl_2$ was 1.5 mM and the amount of template gDNA was 25 ng.

2.4. Discussion

Bioinformatics approaches are increasingly being used for molecular marker development since the sequences from many EST databases and genomes are freely available in public databases (Kantety *et al*., 2002; Varshney *et al*., 2002). These sources are mined for markers using computational tools, thereby eliminating the need for costly, laborious and time-consuming marker development from genomic libraries. The availability of the ESTs and BAC sequences for *R. appendiculatus* provided a unique opportunity to develop micro and minisatellite markers to accelerate research aimed at studying the population genetics and genetic diversity of this important tick vector. The computer program used in this study was able to detect and identify over 1000 micro- and minisatellite markers present in these *R. appendiculatus* nucleotide databases. Thus, the *R. appendiculatus* EST

database and BAC library were found to be valuable resources for rapid discovery of micro- and minisatellite repeats.

Large eukaryotic genomes are known to comprise significant amounts of non-coding, repetitive DNA (Elder and Turner, 1995) and, indeed, repetitive DNA has been reported in tick species. In *R. microplus*, highly repetitive DNA comprising of tandem and dispersed repeats was found to account for approximately 40% of the total tick genome (Ullmann *et al.*, 2005). Using reassociation kinetics, the *I. scapularis* genome was shown to contain 27% highly repetitive (HR) and 39% moderately repetitive (MR), while the *R. microplus* genome contained 31% HR and 38% MR (Ullmann *et al.*, 2005). A significant percentage (~35-40) of the *I. scapularis* and *A. americanum* genomes are comprised of highly repetitive sequences (Ullmann *et al.*, 2005; Palmer *et al.*, 1994). Although repetitive sequences have been shown to be abundant in these tick species, few micro- and minisatellite markers for ticks species have been characterized. Markers for *R. microplus* (Chigagure *et al.* 2000, Koffi *et al*. 2006) and *Ixodes ricinus* (Røed *et al.*, 2006) have been isolated from genomic libraries and characterized. They have been used to show paternity in *I. ricinus* (Gunnar et al, 2008) and *Dermacentor variabilis* (Ruiz-López *et al.*, 2012). The markers have been applied to show genetic polymorphism in *R. microplus* (Labruna *et al.*, 2009) and to dissect the genetic structure of amitraz-resistant and susceptible *R. microplus* populations in Queensland, Australia (Cutulle *et al.*, 2009).

The current study is the first to identify and characterize genetic markers of *R. appendiculatus.* The 7340 non-redundant salivary gland EST sequences with a total of 6.26×10^6 bases represented 0.6 % of the total *R. appendiculatus* genome, which is presumed to be 1.0×10^9 bases (Sunter *et al*., 2008). Repeats were present in 7.1 % of the EST sequences at a frequency of 1 repeat for every

6.673 Kb. Tentative consensus sequences were found to have more repeats compared to singletons. All the eleven BAC sequences with a total of 3.608×10^5 bases representing 0.036 % of the entire *R. appendiculatus* genome were found to contain repeat markers at a frequency of 1 repeat for every 2.405 Kb of genomic sequence. This implies that in *R. appendiculatus* relative abundance of repeats is three times more in genomic sequences compared to ESTs. The relatively low number of repeat sequences in ESTs may reflect the lower frequency of repeats in coding regions of *R. appendiculatus* genome. The 42 amplifiable markers obtained in this study confirm that EST and genomic sequences are useful sources of good micro- and minisatellite markers (Ellis & Burke, 2007) particularly if a reference genome is unavailable.

Optimal PCR conditions are essential for specificity, efficiency, and sensitivity of a PCR reaction. PCR optimization reduces commonly encountered problems such as undetectable products, low yield of desired products, or the presence of non-specific PCR products (Wang *et al*., 2010). Forty-two of the sixty-six primers designed and optimized in this study amplified a single fragment of the anticipated size suggesting that they exist as single copies. All the primers designed from the BAC sequences yielded poor and sometimes non-specific products. This is one of the most common problems encountered during micro and minisatellite analysis. This problem could possibly be because of poor source sequence or the possibility of a primer site spanning an exon/intron junction. The non-specific amplification observed for some primers may also be the result of multiple priming sites along the genome (Rallo *et al*., 2000), for example, because of tandem repeats or larger gene duplications. In post-PCR analysis, satellite markers that give strong clean signal with a minimum of stray bands are desirable (Gastier *et al*., 1995) and thus the 24 markers that did not generate PCR amplicons were dropped from the study.

In this study, a new set of 42 EST-based micro and minisatellite markers were identified through data mining from the *R. appendiculatus* EST database. This approach was found to be fast, affordable and an efficient way to identify molecular markers for diversity studies of *R. appendiculatus* because its reference genome is unavailable. This represents the first step towards development of molecular tools to be applied in genetic characterization of *R. appendiculatus*. These newly developed markers can be useful for studying the population structure and understanding the genetic diversity of this important tick vector. They will be used to confirm tick genotype identity and detect possible genotype mixtures among the field populations. The markers can also be applied in other molecular studies of *R. appendiculatus* such as standardization of the production of the live *T. parva* vaccine and in defining parameters of relevance to the epidemiology of ECF in the field. They will also enrich the molecular resources for *R. appendiculatus*. Since EST-based markers are more transferable between species than genomic markers (Gong *et al*., 2010; Mishra *et al*., 2011) these markers can be used to perform inter-species population genetics analysis as has been done in other species such as birds (Primmer *et al.*, 1996) and mosquitoes (Kamau *et al.,* 1999). Hence, the identified markers may be transferable to other tick species and could be applied in discriminating closely related rhipicephaline tick species in the field. Moreover, EST-based markers represent the transcribed part of the genome and thus could be present in genes that are translated into proteins, they can help establish gene traits and assist in marker assisted selection.

Results from this chapter have been presented in the following conferences;

1. **12th Kari Biennial Scientific Conference: 8 – 12 November 2010. KARI Headquarters, Nairobi, Kenya**- Application of EST-SSR markers in the study of genetic diversity of *R. appendiculatus* (Acari: Ixodidae).

2. **7th Tick and Tick-Borne Pathogens International Conference (TTP 7): 28th August-2nd September, 2011. City Auditorium, Zaragoza, Spain**-Polymorphic EST-SSR markers and their application in assessing the genetic diversity of laboratory and Kenyan field populations of the tick *Rhipicephalus appendiculatus* (Acari: Ixodidae).

The abstracts are given in Appendix II.

2.5 References

Anderson, T. J. C., Su, X. Z., Bockarie, M., Lagog, M., Day, K. P., (1999). Twelve microsatellite markers for the characterisation of *Plasmodium falciparum* from finger-prick blood samples. *Parasitology* **119**, 113–125.

Anderson, T.J.C., Haubold, B., Williams, J.T., Estrada-Franco, J.G., Richardson, J., Mollinedo, R., Bockarie, M., Mokili, J., Mharakurwa, S., French, N., Whitworth, J., Velez, I. D., Brockman, A. H., Nosten , F., Ferreira, M.U., Day, K.P., (2000). Microsatellite markers reveal a spectrum of population structures in the malaria parasite *Plasmodium falciparum*. *Molecular Biology and Evoution,***. 17**, 1467–1482.

Armour, J. A., Neumann, R., Gobert, S. and Jeffreys, A. J. (1994). Isolation of human simple repeat loci by hybridization selection. *Human Molecular Genetics,* 3, 599-565.

Armour, J.A.L., (1999). Microsatellites and mutation processes in tandemly repetitive DNA. In: Goldstein, D., Schlotterer, C. (Eds.), *Microsatellites: Evolution and Applications, Oxford University Press, Oxford,* pp. 24–29.

Beaumont, M. A., and Bruford, M.W., (1999). "Microsatellites in conservation genetics. In: Microsatellites: Evolution and Application" *Edited by D. B. Goldstein and Schlötterer, Orford University Press, New York,* pp 165-183.

Benson, G. (1999). Tandem repeats finder (TRF): a program to analyze DNA sequences. *Nucleic acid Research,* **27(2),** 573-580.

Chigagure, N. N., Baxter, G. D., Barker, S. C. (2000). Microsatellite loci of the cattle tick *Boophilus microplus* (Acari: Ixodidae). *Experimental Applied Acarology*, **24 (12**), 951-956.

Cifarelli, R.A, Gallitelli, M. and Cellini, F. (1995). Random amplified hybridization microsatellites (RAHM): isolation of a new class of microsatellite-containing DNA clones. *Nucleic Acids Research,* **23,** 3802-3803.

Coenye, T and Vandamme, P. (2005). Characterization of mononucleotide repeats in sequenced prokaryotic genomes. *DNA Research***, 12 (4**), 221-233.

Cooper, S., Bull, C.M., Gardner, M. G. (1997). Characterization of microsatellite loci from the socially monogamous lizard *Tiliqua rugosa* using a PCR-based isolation technique. *Molecular Ecology*, **6**, 793-795.

Cutulle, C., Jonsson, N. N., Seddon, J. (2009). Population structure of Australian isolates of the cattle tick *Rhipicephalus (Boophilus) microplus*. *Veterinary Parasitology* **161,** 283–291.

Debrauwere, H. C. G., Gendrel, S., and Lechat, M. D. (1997). Differences and similarities between various tandem repeat sequences: Minisatellites and microsatellites, *Biochimie*, **79(9–10),** 577–586.

Elder Jr, J. F., & Turner, B. J. (1995). Concerted evolution of repetitive DNA sequences in eukaryotes. *The Quarterly Review of Biology*, **70(3),** 297-320.

Ellis, J. R. and Burke, J. M. (2007). EST-SSRs as a resource for population genetic analyses. *Heredity* **99**, 125–132.

Fisher, P. J. Gardner, R.C., Richardson, T. E. (1996). Single locus microsatellites isolated using 5¢-anchored PCR. *Nucleic Acids Research,* **24,** 4369-4371.

Gastier, J. M., Pulido, J. C., Sunden, S., Brody, T., Buetow, K. H., Murray, J. C & Duyk, G. M. (1995). Survey of trinucleotide repeats in the human genome: assessment of their utility as genetic markers. *Human molecular genetics*, **4(10),** 1829-1836.

Gemayel, R., Vinces, M. D., Legendre, M. and Verstrepen, K. J. (2010). Variable tandem repeats accelerate evolution of coding and regulatory sequences. *Annual Review of Genetics*, **44,** 445–77.

Goldstein, D. B., and Pollock, D.D., (1997). Launching microsatellites: a review of mutation processes and methods of phylogenetic inference. *Journal of Heredity*, **88**, 335-342.

Gong, Y., Xu, S. C, Mao, W., Hu, Q., Zhang, G., Ding, J., Li, Y. (2010). Developing new SSR markers from ESTs of pea (*Pisum sativum* L). *Journal of Biomedical Biotechnology,* **11(9),** 02-707.

Gupta, P. K., Rustgi, S., Sharma, S., Singh, R., Kumar, N., & Balyan, H. S. (2003). Transferable EST-SSR markers for the study of polymorphism and genetic diversity in bread wheat. *Molecular Genetics and Genomics*, **270(4),** 315-323.

Hamada, H., Petrino, M. G. and Kakunaga, T. (1982). A novel repeated element with Z-DNA-forming potential is widely found in evolutionarily diverse eukaryotic genomes, *Proceedings of the National Academy of Sciences of USA*, **79,** 6465–6469.

Hancock, J. M. (1999). Microsatellites and other simple sequences: genomic context and mutational mechanisms. *Microsatellites: Evolution and applications*, **1,** 3-9.

Jeffreys, A. J., Wilson, V. and Thein, S. L. (1985). Hypervariable "minisatellite" regions in human DNA. *Nature* **314,** 67–73.

Kamau, L., Mukabana, W. R., Hawley, W. A., Lehmann, T., Irungu, L. W., Orago, A. A. S., Collins, F. H. (1999). Analysis of genetic variability in *Anopheles arabiensis* and *Anopheles gambiae* using microsatellite loci. *Insect Molecular Biology* **8(2),** 287-298.

Kantety, R. V., La Rota, M., Matthews, D. E., & Sorrells, M. E. (2002). Data mining for simple sequence repeats in expressed sequence tags from barley, maize, rice, sorghum and wheat. *Plant molecular biology,* **48(5-6),** 501-510.

Katti, M. V., Ranjekar, P. K. and Gupta, V. S. (2001). Differential distribution of simple sequence repeats in eukaryotic genome sequences. *Molecular Biology of Evolution,* **18(7),** 1161-1167.

Knapik, E. W., Goodman, A., Ekker, M., Chevrette, M., Delgado, J., Neuhauss, S.and Jacob, H. J. (1998). A microsatellite genetic linkage map for zebrafish *(Danio rerio*). *Nature genetics,* **18(4),** 338-343.

Koffi, B. B., Risterucci, A. M., Joulia, D., Durand, P., Barre, N., de Meeuˆ s, Chevillon, C. (2006). Characterization of polymorphic microsatellite loci within a young *Boophilus microplus* metapopulation. *Molecular Ecology Notes* 6, 502–504.

Labruna, M., Naranjo, V., Mangold, A., Thompson, C., Estrada-Peña, A., Guglielmone, A & De la Fuente, J. (2009). Allopatric speciation in ticks: genetic and reproductive divergence between geographic strains of *Rhipicephalus (Boophilus) microplus. BMC evolutionary Biology,* **9(1),** 46.

Lench, N. J., Norris, A., Bailey, A., Booth, A. and Markham, A. F. (1996). Vectorette PCR isolation of microsatellites repeat sequences using anchored dinucleotide repeat primers. *Nucleic Acids Research,* **24,** 2190-2191.

Li, Y. C., Korol, A. B., Fahima, T., Beiles, A. and Nevo, E. (2002). Microsatellites: genomic distribution, putative functions and mutational mechanisms: a review. *Molecular Ecology,* **11,** 2453-2465.

Litt, M. and Luty, J. A. (1989). A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. American *Journal of Human Genetics,* **44,** 397-401.

Mbanjo, E. G. N., Tchoumbougnang, F., Mouelle, A. S., Oben, J. E., Nyine, M., Dochez, C., Ferguson, M. E., & Lorenzen, J (2012). Development of expressed sequence tags-simple sequence repeats (EST-SSRs) for Musa and their applicability in authentication of a Musa breeding population. *African Journal of Biotechnology,* **11(71);** 13546-13559.

Mishra, R. K., Gangadhar, B. H., Yu, J. W., Kim, D. H., Park, S. W. (2011). Development and characterization of EST based SSR markers in Madagascar periwinkle (*Catharanthus roseus*) and their transferability in other medicinal plants. *Plant Omics Journal* **4(3):**154-162.

Morgante, M., & Olivieri, A. M. (1993). PCR-amplified microsatellites as markers in plant genetics. *The plant journal*, **3(1),** 175-182.

Nakumura, Y., Leppert, M., O'Connell, P., Wolff, R., Holm, T., Culver, M., Martin, C., Fujimoto, E., Hoff, M., Kumlin, E., and White, R., (1987). Variable number of tandem repeats (VNTR) markers for human gene mapping. *Science,* **235**, 1616-1622

Nene, V., Lee, D., Kang'a, S., Skilton, R., Shah, T., de Villiers, E., Mwaura, S., Taylor, D., Quackenbush, J. and Bishop, R. (2004). Genes transcribed in the salivary glands of female *Rhipicephalus appendiculatus* ticks infected with *Theileria parva*. *Insect Biochemstry and Molecular Biology*, **34,** 1117-1128.

Oliveira, R.P., Broude, N. E., Macedo, A. M., Cantor, C. R., Smith, C. L., Pena, S. D. J. (1998). Probing the genetic population structure of *Trypanosoma cruzi* with polymorphic microsatellites. *Proceedings of the National Academy of Sciences, USA.* **95**, 3776–3780.

Oura, C. A. L., Odongo, D. O., Lubega, G. W., Spooner, P. R., Tait, A., and Bishop, R. P., (2003). A panel of microsatellite and minisatellite markers for the characterization of field isolates of *Theileria parva*. *International Journal of Parasitology;* **33**, 1641–53.

Palmer, M. J., Bantle, J. A., Guo, X. and Fargoxy, W. S. (1994). Genome size and organization in the ixodid tick *Amblyomma americanum* (L.). *Insect molecular biology*, **3(1),** 57-62.

Park, L. K., & Moran, P. (1995). Developments in molecular genetic techniques in fisheries. *Molecular Genetics in Fisheries* (pp. 1-28). Springer Netherlands.

Parkinson, J. and Blaxter, M. (2009). Expressed sequence tags: an overview (2009). *Methods in Molecular Biology*, **533,** 1-12.

Pérez, F., Ortiz, J., Zhinaula, M., Gonzabay, C., Calderón, J., & Volckaert, F. A. (2005). Development of EST-SSR markers by data mining in three species of shrimp: *Litopenaeus vannamei, Litopenaeus stylirostris,* and *Trachypenaeus birdy*. *Marine Biotechnology*, ***7*(5),** 554-569.

Primmer, C. R., Muller, A. P. and Ellegren, H. (1996). A wide-ranging survey of cross-species amplification in birds. *Molecular Ecology,* **5,** 365-378.

Putney, S. D., Herlihy, W. C., and Schimmel, P. (1983). A new troponin T and cDNA clones for 13 different muscle proteins, found by shotgun sequencing. *Nature,* **302,** 718–721

Rassmann, K., Schlštterer, C., and Tautz, D. (1991). Isolation of simple sequence loci for use in polymerase chain reaction-based DNA fingerprinting. *Electrophoresis,* **12,** 113-118.

Richard GF, Kerrest A, Dujon B. (2008). Comparative genomics and molecular dynamics of DNA repeats in eukaryotes. *Microbiology and Molecular Biology Review,* **72,** 686–727.

Richardson, T., Cato, S., Ramser, J., Kahl, G. and Weising, K. (1995). Hybridization of microsatellites to RAPD: a new source of polymorphic markers. *Nucleic Acids Research,* **23,** 3798-3799.

Rohrer, G. A., Fahrenkrug, S. C., Nonneman, D., Tao, N., & Warren, W. C. (2002). Mapping microsatellite markers identified in porcine EST sequences. *Animal genetics*, **33(5),** 372-376

Rossetto, M. (2001). Sourcing of SSR markers from related plant species. In: Henry RJ, editor. Plant genotyping: the DNA fingerprinting of plants. Wallingford, CT: CABI Publishing; pp. 211-224.

Ruiz-López, M. J., Chaskelson, S., Gompper, M. E., & Eggert, L. S. (2012). Multiple paternities in the American dog tick, *Dermacentor variabilis* (Acari: Ixodidae). *Journal of Parasitology*, **98(3),** 498-501.

Røed, K. H., Hasle, G., Midthjell, V., Skretting, G. and Leinaas, H. P. (2006 Identification and characterization of 17 microsatellite primers for the tick, *Ixodes ricinus*, using enriched genomic libraries. *Molecular Ecology Notes,* **6(4),** 1165–1167.

Røed, K. H., Hasle, G., Midthjell, V., Skretting, G., Leinaas, H. P. (2006). Identification and characterization of 17 microsatellite primers for the tick, *Ixodes ricinus*, using enriched genomic libraries. *Molecular Ecology Notes* **6(4),** 1165–1167.

Schlotterer, C. (2000). Evolutionary dynamics of microsatellite DNA. *Chromosoma* **109,** 365-371.

Schlotterer, C. and Tautz, D. (1992). Slippage synthesis of simple sequence DNA. *Nucleic Acids Research,* **20,** 211-215.

Schlötterer, C. (1998). Microsatellites. *Molecular genetic analysis of populations: a practical approach*, ***2*,** 237-61.

Schuler, G. D., Boguski, S., Stewart, A., Gyapay, G., Rice, K., Page, D. C. and White, E. (1996). A Gene Map of the Human Genome. *Science*, **274 (5287),** 540-546.

Sunter, J.D., Patel, S.P., Skilton, R.A., Githaka, N., Knowles, D.P., Scoles, G.A., Nene, V., de Villiers, E. and Bishop, R.P. (2008). A novel SINE family occurs frequently in both genomic DNA and transcribed sequences in ixodid ticks of the arthropod sub-phylum *Chelicerata*. *Gene*, **415,** 13-22.

Tautz, D. (1989). Hypervariability of simple sequences as a general source for polymorphic DNA markers, *Nucleic Acids Research,* **17,** 6463–6471.

Tautz, D. and Schlotterer, C. (1994). Simple sequences, Current *Opinion in Genetics and Development,* **4,** 832–837.

Tautz, D., and Renz, M., (1984). Simple sequences are ubiquitous repetitive components of eukaryotic genomes. *Nucleic Acids Reserach*. **12**, 4127–4138.

Thiel, T., Michalek, W., Varshney R.K., Graner, A. (2003). Exploiting EST databases for the development and characterization of gene-derived SSR-markers in barley (*Hordeum vulgare L.*). *Theoretical Applied Genetics* **106,** 411–422.

Toth, G., Gaspari, Z. and Jurka, J. (2000). Microsatellites in different eukaryotic genomes: survey and analysis. *Genome Research* **10,** 967-81.

Ullmann, A. J., Lima, C. M. R., Guerrero, F. D., Piesman, J. and Black, W. C. (2005). Genome size and organization in the blacklegged tick, *Ixodes scapularis* and the Southern cattle tick, *Boophilus microplus*. *Insect molecular biology,* **14(2),** 217-222.

Varshney, R. K., Thiel, T., Stein, N., Langridge, P., & Graner, A. (2002). In silico analysis on frequency and distribution of microsatellites in ESTs of some cereal species. *Cellular and Molecular Biology Letters*, **7(2A),** 537-546.

Weber, J. L. (1990). Informativeness of Human (Dc-Da)N.(Dg-Dt)N Polymorphisms. *Genomics*, **7,** 524-530.

Weber, J. L. and May, P. E. (1989). Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *American Journal of Human Genetics,* **44,** 388-396.

Williams, J. G. K., Kubelik, A.R., Livak, K. J., Rafalski, J.A., Tingey, S.V. (1990). DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Research,* **18,** 6531-6535

Wright, J. M. (1994). Mutation at VNTRs: Are minisatellites the evolutionary progeny of microsatellites? *Genome,* **37(2),** 345-347.

Wright, J. M., & Bentzen, P. (1995). Microsatellites: genetic markers for the future. In *Molecular genetics in fisheries* (pp. 117-121). Springer Netherlands.

Wu, K., Jones, R., Danneberger, L. and Scolnik, P. A. (1994). Detection of microsatellite polymorphisms without cloning. *Nucleic Acids Research,* **22**, 3257-3258

Yue, G. H., Ho, M. Y., Orban, L., & Komen, J. (2004). Microsatellites within genes and ESTs of common carp and their applicability in silver crucian carp. *Aquaculture*, **234(1),** 85-98.

Zane, L., Bargelloni, L., and Patarnello, T., (2002). Strategies for microsatellite isolation: a review. *Molecular Ecology* **11**, 1-16.

CHAPTER THREE: VALIDATION AND CHARACTERISATION OF MICRO- AND MINISATELLITE DNA MARKERS OF *Rhipicephalus appendiculatus*

3.0 Abstract

Biological differences, including vector competence for the protozoan parasite *Theileria parva* have been reported between populations of *Rhipicephalus appendiculatus* (Acari: Ixodidae) from different geographic regions. However, the genetic diversity and population structure of this important tick vector remains unknown due to absence of appropriate genetic markers. This chapter describes the validation of the 42 EST micro- and minisatellite markers described in Chapter Two. The potential of these markers to characterize the genetic diversity within and between populations of *R. appendiculatus* and other rhipicephaline species was also evaluated. The markers were optimized and validated using fluorescently labeled primers with subsequent detection of amplified products on fluorescence-based capillary electrophoresis. They were used to genotype 979 individual ticks from 11 field populations, 11 laboratory-bred stocks and five additional *Rhipicephalus* species. Twenty nine markers were polymorphic and therefore informative for genetic studies while six were monomorphic. The 29 polymorphic markers discriminated populations of *R. appendiculatus* and also four other *Rhipicephalus* species but not *R. zambeziensis*. The percentage Principal Component Analysis (PCA) implemented using Multiple Co-inertia analysis (MCoA) clustered populations of *R. appendiculatus* into two groups. Individual markers however differed in their ability to generate the reference typology using the MCoA approach. This indicates that different panels of markers may be required for different applications. The 29 informative polymorphic micro- and minisatellite markers are the first available tools for the analysis of the phylogeography and population genetics of *R. appendiculatus.*

3.1 Introduction

Previous studies on the variation within *R. appendiculatus* have focused on biological and behavioral differences between populations from different geographic regions (Norval *et al.*, 1992; Shaw and Young 1994; Berkvens *et al*., 1995; Chaka *et al*., 1999; Madder *et al*., 2002; Speybroeck *et al*., 2002, 2004). Different stocks of *R. appendiculatus* and species within the genus differ in their ability to transmit pathogens. For example *R. appendiculatus* and *R. zambeziensis* have been shown to have different levels of vector competence for *T. parva* (Ochanda *et al.,* 1998). By assessing the parasite load in salivary glands, Norval *et al*., (1991) and Kubasu (1992) demonstrated that brown ear ticks from geographically separated areas differ in their ability as vectors of *T. parva.* These studies indicate that there is considerable phenotypic variation between populations of *R. appendiculatus*.

Molecular characterization of tick species has been undertaken by analyzing DNA variations in both coding and non-coding regions of the genome and several PCR-based markers have been evaluated. The sequences of mitochondrial 12S rDNA and the complete nuclear second internal transcribed spacer (ITS2) have been used to confirm that *R. appendiculatus* and *R. zambeziensis* represent distinct taxa (Mtambo *et al.,* 2007a). These two closely related species of ticks are difficult to discriminate phenotypically due to their morphological similarities. Using the entire ITS2 region, Barker (1998) was able to distinguish populations of *Rhipicephalus* (*Boophilus) microplus* from Australia, Kenya, South Africa and Brazil as well as populations of *R. appendiculatus* from Kenya, Zimbabwe and Zambia. In another study, a 274 bp part of the ITS2 sequence was used to demonstrate close genetic relationship between 6 tick species within the *R. sanguineus* complex (Zahler *et al.,* 1997). A novel real-time ITS2 PCR assay was used to differentiate four medically

important ticks in the US; *Ixodes scapularis, I. pacificus, Dermacentor variabilis and Amblyomma americium* (Shona *et al.*, 2006). Phylogenetic relationship between several species within the *Rhipicephalinae* subfamily has been shown using the 12S rDNA (Beati & Keirans, 2001; Murrell *et al*, 1999).

Cytochrome c oxidase subunit I (COI) is one of the most conserved mitochondrial protein coding genes. It has been evaluated as a taxonomical tool and has served as the core of a global bioidentification (barcoding) system for many animal species (Hebert *et al.*, 2002). The phylogeny of 21 species and subspecies of ticks from the subfamilies *Rhipicephalinae* and *Hyalomminae* inferred using COI and 12S ribosomal RNA (rRNA) showed the genus *Rhipicephalus* to be paraphyletic with a subgenus *Rhipicephalus Boophilus* (Murrell *et al.*, 2000). COI has also been used to provide preliminary evidence of geographic differentiation between *R. appendiculatus* ticks from eastern and southern provinces of Zambia (Mtambo *et al*., 2007b).

Other PCR-based molecular genetic markers which offer quick detection and characterization of genetic variations include micro- and minisatellite markers. Microsatellites have been used to study the population structure and genetic diversity of human disease vectors such as tsetse flies and mosquitoes. Such markers have been used extensively in population genetic analyses, for example in *P. falciparum*, to define population structures, geographical sub-structuring and levels of inbreeding (Anderson *et al.*, 1999, 2000). They have been used to show gene flow of *Glossina pallidipes* in south-western Kenya (Ouma *et al*., 2006) as well as genetic differentiation and diversities of tsetse fly populations from different geographic regions in Africa (Gooding and Krafsur, 2004; Norris *et al.*, 2001). Using microsatellites, DNA polymorphism and heterozygosity have been shown to be drastically reduced among laboratory colonies of *Aedes egypti* compared to

field populations (Norris *et al,.* 2001) while gene flow between populations of *Anopheles arabiensis* and *A. gambiae* from East and West Africa has been inferred from the analysis of Variable Number Tandem Repeats (VNTRs) (Kamau *et al*., 1999). Microsatellites have been isolated and characterized from ixodid tick species, including *Rhipicephalu*s (*Boophilus) microplus* (Chigagure *et al.*, 2000, Koffi *et al*., 2006) and *Ixodes ricinus* (Røed *et al.*, 2006). They have been used to show paternity in *I. ricinus* (Hasle *et al.*, 2008) and *Dermacentor variabilis* (Ruiz-López *et al.*, 2012). These markers have been applied to show genetic polymorphism in *Rhipicephalu*s (*Boophilus) microplus* (Labruna *et al*., 2009) and to dissect the genetic structure of amitraz-resistant and susceptible *R. microplus* populations in Queensland, Australia (Cutulle *et al.*, 2009).

In this study, the identification of a new set of 42 amplifiable EST-based micro- and minisatellite markers from the *R. appendiculatus* EST database as described in Chapter two represented the first step towards development of molecular tools to be applied in genetic characterization of *R. appendiculatus*. The markers were characterized and evaluated for their usefulness in assessing genetic diversity within and between different field and laboratory-bred tick populations from the genus *Rhipicephalus*.

3.2 Materials and methods

3.2.1 Experimental tick populations

A total of 979 ticks were collected from different populations (Table 6). These included 460 *R. appendiculatus* ticks from 10 populations that have been maintained and expanded as closed laboratory colonies (Table 6). They also included 392 individual ticks collected from 10 localities in Kenya characterized by different agroecological conditions such as different mean annual rainfall and mean annual maximum temperatures (Figure 10 and 11). These included 180 individuals from

five populations sampled from pastures and herds of domestic cattle, 80 individuals from two populations from areas co-inhabited by wildlife and cattle, and 132 individuals from three areas exclusively inhabited by wild animals. The experimental samples also included 10 individuals from South Africa and 117 from five additional *Rhipicephalus* species, including a laboratory stock of *R. zambeziensis* which occurs in southern Africa and four other species from Kenya. The latter were included in the study to evaluate the species-specificity of the markers. Table 6 gives the origin detailed description of the agroecological characteristics of the 27 tick populations studied.

Table 6: Field and laboratory tick populations used to evaluate the micro- and minisatellite markers used in this study

Species	Category	Code	Area of origin	Population and sampling site description	Sample size
R. appendiculatus	Field tick populations	BU	Busia County	Ticks collected from cattle herds and pastures. Farming system: Small scale open grazing of indigenous breeds with minimal tick control. Area has dense *R. appendiculatus* populations	48
		KF	Kilifi County	Ticks collected from cattle herds and pastures. Farming system: Open grazing of indigenous breeds with minimal tick control. Area has dense *R. appendiculatus* populations	25
		MK	Makuyu, Muranga County	Ticks collected from cattle herds and pastures. Farming system: Small scale farming of exotic/indigenous crosses. Area has dense *R. appendiculatus* populations and tick control is minimal	46
		RU	Rusinga Island	Ticks collected from cattle herds and pastures. Farming system. Open small scale grazing of indigenous cattle breeds with minimal tick control. Area has dense *R. appendiculatus* populations	36
		KT	Kitale, Transzoia	Ticks collected from cattle herds and pastures. Farming system. Open grazing of exotic cattle breeds in large-scale with high tick control. Area has dense *R. appendiculatus* populations	25
		BO	Bomet County	Ticks collected from cattle herds and pastures. This a wildlife area with open grazing of indigenous and mixed cattle breeds in small scale with low tick control. Area has dense *R. appendiculatus* populations	34
		FP	Ol Pejeta Conservancy, Laikipia	Ticks collected from cattle and pastures. This is a wildlife area with open grazing of indigenous and mixed breeds in large-scale with high tick control. Area has dense *R. appendiculatus* populations	46
		BF	Buffaloes in Maasai Mara N. Reserve	Ticks collected from buffaloes. This is a wildlife area that has no contact with cattle. Area has dense *R. appendiculatus* populations	48
		MA	Maasai Mara N. Reserve	Ticks collected from pastures. This is a wildlife area that has no contact with cattle. Area has dense *R. appendiculatus* populations	48
		NB	Nairobi National Park	Ticks collected from pastures. This is a wildlife area that has no contact with cattle. Area has dense *R. appendiculatus* populations	36
		SN	South Africa Natal Province	Ticks were originally collected from wildlife pastures in Natal Province, South Africa and subsequently maintained at Onderstepoort Veterinary Institute (OVI), Pretoria, Republic of South Africa.	5
R. appendiculatus	Laboratory stocks originating from Kenya	KH	Kiambu Highline (shows high infectivity)	A selected stock derived from a family line of *R. appendiculatus* Kiambu that had high susceptibility to *T. parva* infection (Young et al. 1995). It has been maintained at ILRI	48
		KU	Kiambu unselected line	This stock was originally isolated from Kiambu County, Central Province, Kenya in 1981 and subsequently maintained at ILRI.	48
		LP	Laboratory Ol Pejeta	Stock was established in 1987 from a collection made in Ol Pejeta Ranch, Laikipia County. This ranch is inhabited by cattle and a large population of wildlife including African Buffalo. The stock is maintained at ILRI	48
		MF	Muguga Infected ticks	These were *T. parva*-infected ticks derived from the original unselected Muguga stock. The stock has been maintained at ILRI	48

		ML	Muguga Low line (shows low infectivity)	A Selected stock was derived from a family line of *R. appendiculatus* Muguga that had low susceptibility to *T. parva* infection (Young et al. 1995). The stock has been maintained at ILRI	48
		MU	Muguga unselected line	This is the laboratory stock used in Kenya for *T. parva* transmission studies. It was originally collected from the Central Highlands of Kenya in the 1950s and has subsequently been maintained at the East African Veterinary Research Organisation (EAVRO) (now Kenya Agricultural Research Institute (KARI)) and ILRI (Bailey 1960).	48
R. appendiculatus	Laboratory stocks from other African countries	ZE	Zambia (Eastern province)	The colony was established from ticks collected cattle from eastern province of Zambia in 1992. It is maintained at ILRI	48
		ZS	Zambia (Southern province)	The colony was established from ticks collected from pastures at Monze, southern province of Zambia in 1992. It is maintained at ILRI	48
		UG	Uganda stock	The colony was established from a tick collection made in Entebbe Uganda in 1984. The stock is maintained at ILRI	48
		ZM	Zimbabwe (West Mashonaland)	The colony was established from ticks collected from cattle at Aryshire farm, West Masonaland Zimbabwe in 1992. The stock is maintained at ILRI	28
		SL	South Africa stock	These were preserved ticks obtained in 2011 from Onderstepoort Veterinary Institute (OVI), Pretoria, Republic of South Africa.	5
Other *Rhipicephalus* species					
R. pravus	Field	RV	Bomet and Kitale districts	Cattle herds and pastures	12
R. praetextatus	Field	RP	Machakos District (Kapiti) and Ruma National Park	Cattle-wildlife pastures	24
R. zambeziensis	Laboratory	RZ	Zimbabwe (Kilkering)	The colony was established from ticks obtained from Onderstepoort Veterinary Institute (OVI), Republic of South Africa in 1992. The ticks were originally from Kilkering, Zimbabwe	48
R. evertsi	Field	RE	Kitale, Bomet and Makuyu Counties	Cattle herds and pastures	15
R. pulchellus	Field	RL	Machakos County (Kapiti) and Nairobi National Park	Cattle-wildlife pastures	18
				Total	**979**

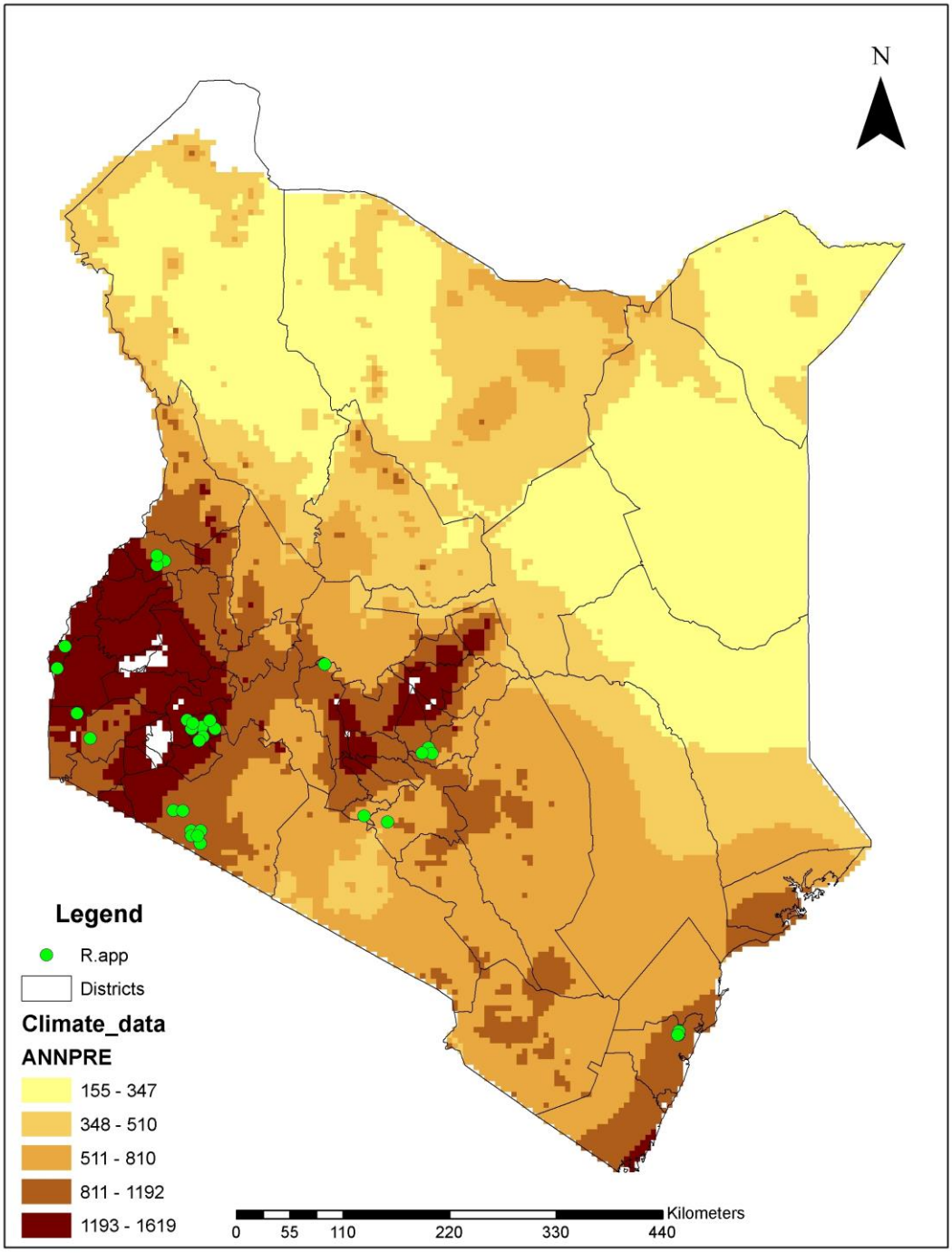


Figure 10: Map of Kenya showing the mean annual rainfall and the sampling sites from where the *R. appendiculatus* ticks were collected.
The map also shows the mean annual rainfall distribution in Kenya given by annual precipitation values (ANNPRE) in mm. The tick samples used in the study occurred in areas with a mean annual precipitation above 511 mm.

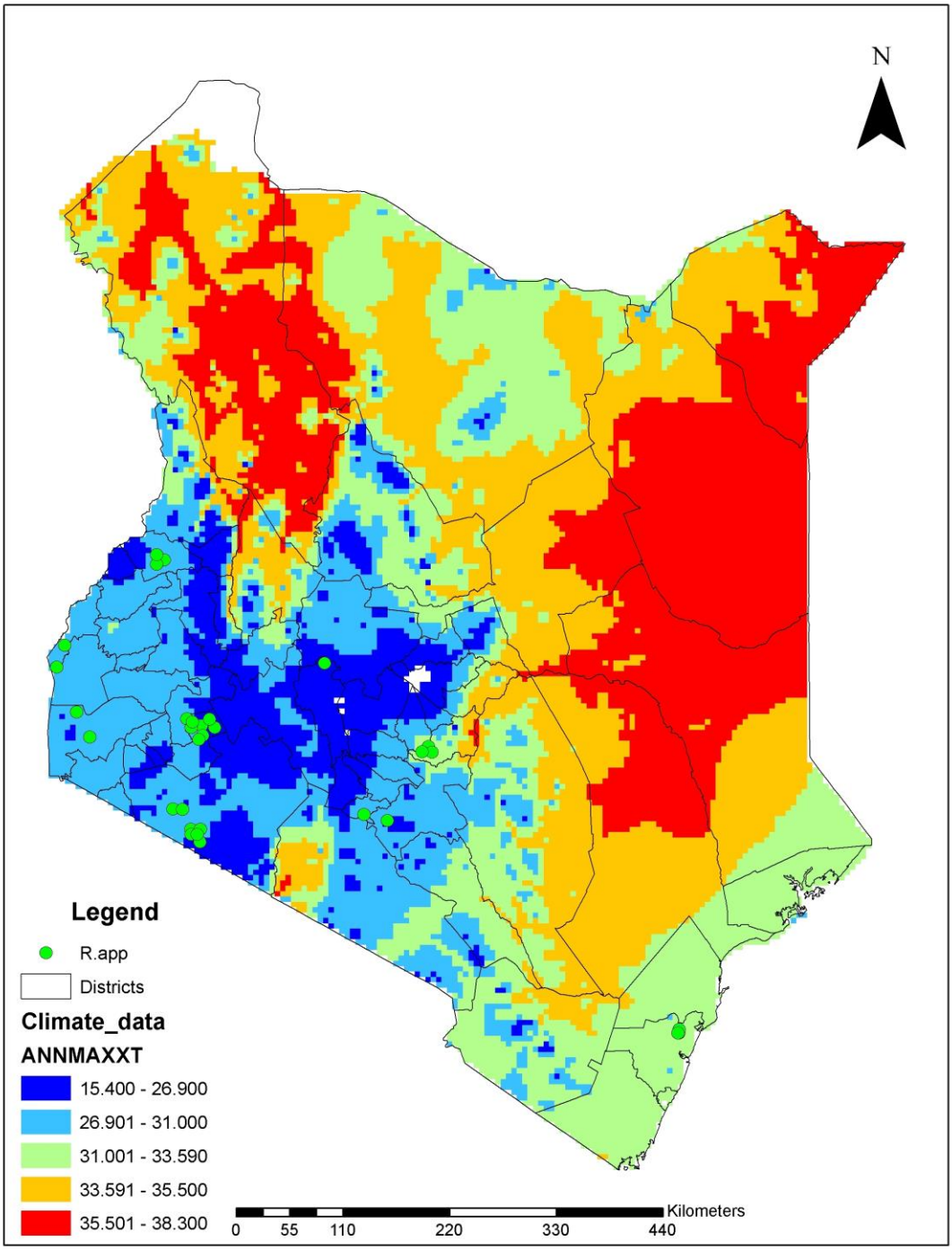


Figure 11: Map of Kenya showing the mean annual maximum temperature (ANNMAXXT) and the sampling sites from where of the *R. appendiculatus* ticks were collected
The ANNMAXXT values are given in °C. The tick samples used in the study occurred in areas with a mean annual maximum temperature below 33.6°C.

3.2.2 Maintenance and colony expansion of laboratory tick stocks

All the laboratory stocks used in this study were obtained from colonies maintained at the International Livestock Research Institute (ILRI) Tick Unit. These stocks had been reared, maintained and expanded as closed laboratory colonies. The description and origin of these stocks is shown in Table 6. Rearing and management conditions were as previously described and involved feeding ticks on either cattle or rabbits (Bailey, 1960; Irvin and Brocklesby, 1970). This method was used to rear and manage all the laboratory stocks used in this study.

3.2.2.1 Feeding laboratory tick stocks on cattle

Bos taurus or *Bos indicus* cattle from a closed herd aged between 6-9 months were screened to confirm absence of antibodies to tick borne diseases and then held under tick-free conditions. They were treated with acaricides until two weeks before use. The cattle were prepared for application of ticks by shaving an area of the back and/or base of the ears using animal clippers. Cloth ear bags or cloth back patches were then attached in place using zinc oxide plaster (BSN Medical GMbh & Co Hamburg, Germany) or rubber glue (Henkel chemicals, Nairobi, Kenya). To propagate different tick instars, 100 unfed adults' pairs or 800 unfed nymphs or 20,000 unfed larvae were applied to each back patch and up to 5 patches were attached to each animal. Adult ticks took 6-10 days to engorge, while larvae and nymphs took 4-7 days.

3.2.2.2 Feeding ticks on rabbits

New Zealand rabbits obtained from a closed colony maintained at ILRI were used. The base of each rabbit ear was shaved with small animal clippers and cloth cotton ear bags attached on combined ears using zinc oxide plaster (BSN Medical GmbH & Co., Hamburg, Germany). About 200 pairs of adult ticks were applied to combined ears and the top of each ear bag was sealed with tape to allow

the ticks to engorge and lay eggs. For producing other tick instars 500 unfed nymphs or unfed 10,000 larvae were applied to separate or combined ears. The back legs of rabbits were taped together loosely to prevent animals from scratching and detaching the bags. Adult ticks engorged between 6-10 days and larvae or nymphs engorged within 4-7 days post attachment.

3.2.2.3 Tick harvesting, egg laying, moulting and hardening

Engorged female adults were collected from the ear bags or back patches and debris was removed from the ticks by sieving using a mesh of appropriate size. The females were then dispensed into plastic Petri dishes at about 40 ticks per dish and then incubated for 4 weeks at 27^{o}C, with a relative humidity (RH) of 85%. Eggs were dispensed at 2g/tube (soda glass flat bottomed tube (Achelis, 28077 Bremen, Germany) and incubated at 27^{o}C, RH 85%. Hatching of eggs into larvae took 5 -7 weeks after females dropping from host. A sample of larvae (0.1g) was weighed and counted to ascertain numbers of larvae in the sample. The larvae were then dispensed as 2,000 (as gram equivalent) in flat bottomed tubes and incubated at 27-30^{o}C RH 85% for three weeks for hardening of the cuticle. Larvae were then stored at 20^{o}C, with RH 90% until they were required for feeding on rabbits or cattle to produce nymphs. Nymphs emerging from larvae hardened after two weeks and were stored at 20^{o}C, with RH 90%. They were maintained in this condition for up to 5 months before feeding them on cattle or rabbits for maturing and moulting into adults.

3.2.3 Collection and preservation of field tick populations

Adult *R. appendiculatus*, *R. praetextatus*, *R. pravus*, *R. evertsi* and *R. pulchellus* ticks were collected from 10 field sites in Kenya (Table 6) with different agro-ecological zones (Figure 10 and 11), different farming systems, tick control practices and different breeds of cattle. The ticks were either collected from vegetation by dragging or from animal hosts (cattle and buffaloes).

3.2.3.1 Collection of ticks from vegetation

Adult *Rhipicephalus* ticks were picked by hand from pastures where they normally quest on vegetation. More often the ticks were collected by dragging. A 1m piece of white cotton cloth or cotton towelling fitted with a bar at the front and a cord was used. The cloth was pulled slowly across vegetation for 5m-10m i.e. for approximately 30secs of walking. Adult ticks gripping onto the dragging cloth were collected with a pair of forceps. The process was repeated after removing the ticks.

3.2.3.2 Collection of ticks from animal hosts

R. appendiculatus adults prefer to feed on ears of cattle and buffaloes. The buffaloes were first sedated before the ticks could be removed. Ticks were plucked from their host by use of good quality steel forceps of medium size with blunt points and serrated inner surfaces to ensure that they maintained a good condition for identification. The forceps were used to grip the tick firmly over its scutum and mouthparts as closely to the host skin as possible, before pulling out the tick strongly and directly out from the skin. For purposes of nucleic acid analysis, unfed ticks were collected as some blood components present in the tick blood meal may inhibit DNA amplification.

3.2.3.3 Storage, labelling and transportation of field ticks

Live ticks were placed in strong tubes having perforated screw caps containing a small piece of damp paper or cotton wool. A piece of cotton cloth was placed around the mouth of the tube before replacing the cap to prevent any escape of the tick. The tubes were labelled, sealed tightly around the rim of the mouth with leucoplast tape and then kept in a sealed container containing wet cotton wool or paper towel to maintain high humidity. When it was not possible to keep the ticks live in the field, the ticks were preserved at the collection site by placing them directly into 1.5 ml microcentrifuge tubes containing 1ml of 70% ethanol. Preserved or live ticks in collection tubes

were packed in a nalgene container, placed in a carton box, sealed completely with masking tape and transported to the laboratory. Dead ticks in 70% ethanol were then stored at 4°C immediately for later species identification. Tubes containing live ticks had 70% ethanol added, then stored at 4°C until tick species identification was done.

3.2.3.4 Tick identification

The ticks were separated into pools based on collection site. The adult ticks were separated according sex, and species. Species identification was done by examining features of the dry dead ticks under a dissecting light microscope. The main features studied were the colour of the tick and ornamentation of the scutum/consoscutum, shape of the tick, size and distribution of punctuations, grooves and colour of legs. Preliminary identification of ticks was performed at ILRI's Tick Unit and species identification was confirmed at the Onderstepoort Veterinary Institute (OVI), Pretoria, South Africa.

3.2.4 Extraction of whole tick genomic DNA

Genomic DNA was extracted using the DNeasy® Blood and Tissue Kit (Qiagen GmbH, Germany) with minor modifications (see Chapter 2 Section 2.2.3).

3.2.4.1 Determination of nucleic acid concentration and purity

Determination of gDNA concentration and purity was done as stated in Chapter 2 section 2.2.3.1.

3.2.5 PCR amplification of individual mini- and microsatellite loci

Amplification of the forty two micro- and minisatellite markers was done using a PCR strategy that uses tailed primers (Schuelke, 2000). In this strategy, a 24 bp universal sequence (gctacagagcatctggctcactgg) was fused to the 5′-end of the forward primer. The same universal sequence was also used to synthesize four oligonucleotides (tails) that were directly-labeled with

FAM, VIC, NED and PET fluorescent dyes (Applied Biosystems). PCR amplifications were performed on 96-well plates using purified *R. appendiculatus* genomic DNA and PCR reagents from Promega. PCR amplifications were performed in 10 µl reaction volumes consisting of 1X Colourless GoTaq® Flex Buffer, 1.5 mM $MgCl_2$, 0.2 mM of each dNTP, 0.01 U GoTaq® DNA polymerase (Promega), 0.25 µM of each micro-/minisatellite primer, 0.175 µM of fluorescently-labeled tail primer and 25 ng of template gDNA. Reaction conditions comprised of an initial denaturation at 95°C for 5 min followed by 40 cycles of denaturation at 94°C for 30 s, primers annealing temperature (Table 7) for 30 s, and extension at 72°C for 60 s. The final extension was carried out at 72°C for 10 min. This assay reduced the expense associated with direct fluorescent labeling of primers and it allowed for multiplex electrophoresis and subsequent laser detection in an ABI 3730 Genetic Analyzer (Applied Biosystems). After PCR amplification 3 µl of PCR product was resolved by electrophoresis through 2% (w/v) agarose gel in 1x TAE buffer (pH 8.0) for 30 min at 100 V in a Mini Gel migration trough. DNA was visualised and photographed under UV light following ethidium bromide staining of the gel, and precise estimation of the PCR product sizes involved direct comparison with a 100 bp DNA ladder (Promega).

3.2.6 Preparation of samples for capillary electrophoresis analysis

To prepare the PCR products for capillary electrophoresis analysis using the ABI 3730 genetic analyzer, a mixture of HIDI formadide and LIZ500 size standard was prepared by adding 12 µl of LIZ500 to 1 ml of HIDI-formadide. Nine microlitres of the HIDI formadide-LIZ 500 mixtures was added to each well of a 96 PCR plate. The individual PCR products were centrifuged at pulse and 1.0 µl added to each well containing 9 µl HIDI-LIZ500 mixture. The contents were mixed by vortexing on a plate vortexer and the plate centrifuged for 1 min. The DNA fragments were then

denatured at 95°C for five minutes using a thermocycler. The mixture was then quickly chilled on ice before loading on ABI 3730 capillary electrophoresis instrument for fragment analysis.

3.2.6.1 Detection of fluorescently labeled markers on the ABI 3730 Genetic Analyzer

Samples were introduced into the 48 capillaries by electrokinetic injection where they migrated according to size by electrophoresis through a matrix containing urea in a long 36cm capillary array. A run module which consisted of 60°C run temperature, 10 seconds injection time, 15 KV run volatage and 20 min run time was used for all the samples. As the labeled PCR fragments passed through the detection cell, the fluorescent dyes were excited by an argon laser and emission spectra of the various dyes were detected by charge couple device (CCD) camera and the captured electrons were represented as relative fluorescent units (rfu). After the fragments were separated, the resulting electropherograms were evaluated and the peaks sized and alleles called using the GeneMapper Software v4.1 (Applied Biosystems). The reaction conditions that gave a clear single peak with relative fluorescent units (rfu) of $\geq$ 500 and minimum stutter peaks were selected as the optimal PCR assay conditions for that particular micro- or minisatellite.

Table 7: The 29 polymorphic micro- and minisatellite markers identified in this study and the primer sequences used in PCR amplification

Locus	**Repeat Motif**	**Repeat size**	**Forward Primer (**5'-3')	**Reverse Primer (**5'-3')	**Annealing Temperature (°C)**	**Size Range (bp)**
TC705	CA	2	GGGAGCACTTTGGAACAA (18bp)	GGCCTTCTACTGTTCAAGC (19bp)	58	161-196
CD790649	AC	2	CAGGATGAAACCTCAGTG (18bp)	CAACTGGGTGTACTTGCTACTG (22bp)	58	173-201
TC1002	GCA	3	GCAGTCTGGACCCTGGAAG (19bp)	GTCTTGGGTGGTGGTGCTT (19bp)	58	138-161
TC1227	AGC	3	CAGTTCCCTGGGCAGTTAT (19bp)	GTGTACTGGCTCCTTGACAG (20bp)	58	179-200
TC1307	ACA	3	GATCAGCAGACCGCAAAG (18bp)	GCAACGACATTCAGTGGTG (19bp)	58	109-138
TC1716	TTC	3	TCCCAAGGTGAACACAGTA (20bp)	GACCTCTCTGACCTGCTATTTC (22bp)	58	121-155
TC2092	CTG	3	AGTCAATATCGGGATCTCGG (20bp)	ACATTCGCAACCTTCTGG (18bp)	58	207-240
TC2543	CAG	3	GCTCCTTGAGAAGATGAAGC (20bp)	AGCAGCGCTTTCTGGTAT (18bp)	58	124-161
TC440a	TGC	3	CTGCAGTGTCCTCGCATT (18bp)	GGACACTCAGCAACAAGC (18bp)	58	155-227
TC875	GCC	3	CTAGGAACCGCTGCTGTG (18bp)	CTGATGCCATTCTTCGAC (18bp)	58	187-204
CD788386	CAG	3	ACCTCAGGGTGGTATGAAC (19bp)	CATCATGCCTGTTGGATG (18bp)	58	139-199
TC1287	AAAGA	5	CGCGATTCATACCGCAAT (18bp)	CGCTGGGTTGCTGTTTCTC (19bp)	58	152-181
TC1353	CACCGC	6	GAGAGACGTCACGGAAACC (19bp)	GAAGACTCCGTAGTTCTAGCC (21bp)	58	102-138
TC1301	CTCCTGGAC	9	AGGCTCACCGAACATAATC (20bp)	GAGTTGGGCCAGGTCTTG (18bp)	59	168-247
TC1302	CCAGGAGGC	9	CAGGTGGCGTAGGAATAGG (19bp)	CAGGCTCACCGAACATAATC (20bp)	61	146-215
TC1303	CCAGGAGGC	9	CAGGTGGCATCGGAATAG (18bp)	CAGGCTCACCGAACATAATC (20bp)	60	147-191
TC1304b	CTCCTGGGC	9	AGGCTCACCGAACATAATC (20bp)	CAGGTGGCGTCGGAATAG (18bp)	58	148-200
TC1398a	TGGTCTCGG	9	CTCTCTGGTCTCTCTGGACTC (21bp)	GTATCCACCGAGCAGACC (18bp)	59	182-218
TC2381	GTATCCACC	9	GATACCCATAGGTGCCACT (19bp)	GTACGGTGCGTACAATCC (18bp)	58	105-210
TC728B	CAGCAGCAG	9	GATGGATATGGGGGTAGTTC (20bp)	GAATGGGGTTGGTAGTATCC (20bp)	58	164-228
TC1968	TTGGATTCAGCC	12	GAGAGGAGCTCGAAGTTTG (19bp)	AACTGCTGCTACGGAAGAG (19bp)	58	172-244
TC639a	CCGCTACCACCA	12	GCTACCTTCACTCCTGCTT (19bp)	AGCGGTACAGGAGGACAT (18bp)	58	156-214
TC919	AATGCCAGCGCC	12	AGCAGCCTAACGTGGAAG (18bp)	ATTGCTCCTGGTCTTTGC (18bp)	58	160-211
TC1179	GCCAGCAATGCAGAC	15	CACAGACACCTCAGCCAAC (19bp)	CAGCACGGTGGTAGTTTG (18bp)	61	220-364
TC1284a	GGCAGCAGTGGCTAC	15	GGCTTTGGTGGAGACTTT (18bp)	GCTGCTACCTCCTCCAAG (18bp)	58	122-169
TC1398b	GGATTCGGCAGCCCACTCAGC	21	GGCTACCCTGGTCTCTTTG (19bp)	CATGGACAGGGGACCGTAG (19bp)	59	303-402
TC1399b	GGATTCGGCAGCCCACTCAGC	21	GCTCTCGCTACGGAGGTC (18bp)	ATGGACAGGGGACCGTAG (18bp)	59	321-400
TC1400b	GGATTCGGCAGCCCACTCAGC	21	CCTGGTCTCTTTGGCTCTC (19bp)	ATGGACAGGGGACCGTAG (18bp)	58	334-396
TC23	TCCTCCGGACGGGTAGCTAGACAA	24	CTTCCTGAACTCCAAGAGC (19bp)	GACTAACAAGCGGAGGTTAC (20bp)	58	104-342

To allow for accurate and robust mass screening with several markers while reducing the unit cost of high throughput genotyping, PCR products of micro- and minisatellite markers which gave clear sizeable peaks were put into groups of four based on their dye label, fragment size and relative fluorescent unit (rfu). Such a pool of PCR products from four different markers was referred to as a co-loading group and contained products from each of the four markers each labeled with one of the fluorescent dyes. In each co-loading group, empirical proportions of PCR products for each marker to be pooled were worked out based on the rfus and the relative intensities of each individual PCR product when resolved on a 2% TBE agarose gel. The pooled PCR products for each co-load group were centrifuged at pulse and 1.0 µl added to the 9.0 µl of HIDI-LIZ500 size standard mixture. The contents were mixed by vortexing on a plate vortexer and the plate centrifuged for 1 min. The DNA fragments were then denatured at 95 °C for five minutes using a thermocycler. The mixture was then quickly chilled on ice before loading in the ABI 3730 capillary electrophoresis instrument for fragment analysis. Based on the quality of the peaks and the rfus, the proportions of each marker were adjusted and analysed again so that each marker in the co-loading set had relative rfus ranging from 500-15000.

3.2.7 Selection of micro- and minisatellite markers for genetic diversity analysis of *R. appendiculatus*

Forty two micro- and minisatellite markers that gave reliable amplicons with strong clear single or multiple peaks on initial optimization were validated by initially amplifying a set of 12 gDNA samples. The samples were from four different tick populations (one field *R. appendiculatus* population, one non-*R. appendiculatus* species and two laboratory tick stocks). They were then validated further by genotyping a set of 72 tick samples belonging to six different populations (12 samples from each population) which included three non *R. appendiculatus* populations. Thirty five

out of forty two micro- and minisatellite markers were successfully scored while seven did not have clear sizeable peaks and thus were dropped from the study. The thirty five markers were subsequently subjected to polymorphism test using a set of 96 tick samples from eight different populations. Of these, six markers were monomorphic while the remaining 29 markers exhibited size polymorphisms between populations. Table 7 summarizes the nucleotide sequence, the primer sequences, primer annealing temperatures and the predicted size in base pairs of the 29 polymorphic micro- and minisatellite markers used in this study. The discrimination abilities of these 29 polymorphic markers were investigated by further genotyping another set of 288 tick samples (12 samples from each of the 24 population studied) and which included five non *R. appendiculatus* populations. These markers showed considerable size variation and were used to genotype the remaining 595 tick samples. In total, 979 tick samples belonging to 27 different populations (Table 6) were analyzed.

3.2.8 Data analysis

Assessment of allelic diversity: total number of alleles (TNA), polymorphic information content (PIC; Botstein *et al*. 1980; Anderson *et al*. 1993) and observed (Ho) and expected (He) heterozygosity for each marker were estimated from allele frequencies with FSTAT v2.9.4 (Goudet, 2001) and Microsatellite Toolkit (Park, 2001). PIC is an index that provides information on the power of a marker to detect polymorphisms in populations. For instance, a marker that reveals five alleles, with only a single allele occurring at very high frequency (e.g. 0.9) is less discriminating than one, which also has five alleles but all of which occur at equal frequencies. The power of markers to differentiate populations was assessed by calculating F_{ST} (θ) values for each marker individually and also across all markers following Weir & Cockerham's (1984) approach.

Multiple Co-inertia Analysis (MCoA) was performed on locus-population frequency data with the ADE-4 package (Chessel *et al.*, 2004) implemented in R (R Core Development Team, 2006). The MCoA approach is used to find successive axes from two or more tables of allele frequencies that maximize a covariance function. The approach makes it possible to extract common information from separate analyses, through establishing a reference typology and subsequently comparing separate typologies for the same markers. The efficiency of a marker for discriminating between populations is given by its typological value (Tv). Tv represents the contribution of a marker to the construction of the reference typology (Laloë *et al.*, 2007). In performing the MCoA, a reference percent Principal Component Analysis (%PCA) was generated using all the micro- and minisatellite markers that were genotyped successfully. Individual %PCA plots were also constructed for each marker. PCA represents a way of identifying patterns in data, and expressing the data in such a way as to highlight their similarities and differences. PCA creates graphics by measuring different dimensions in datasets, which are then ordered in such a way that PC1 accounts for the highest amount of genetic variation, and each corresponding PC explains a lower proportion of the total variation. In cases where variables are highly correlated, the first three PCs may explain the highest variation. In most cases, only the first two or three PCs are normally shown. Typological values were estimated for each marker to evaluate their percentage contributions in generating the reference %PCA.

To assess whether any of the markers used for genotyping were under selection pressure, the FDIST2 outlier test of Beaumont & Richard (1996) was used. The test identifies loci with low or high genetic differentiation (F_{ST}) with respect to the expected heterozygosity. Coalescent simulations (100 000) based on the infinite allele model were used to generate datasets with the mean F_{ST} equal to the neutral mean F_{ST} of the empirical distribution. The observed F_{ST} values were

compared with the simulated distribution of F_{ST} to identify potential outlier loci. In the first step, the neutral expectation was based on the overall mean value of F_{ST} calculated using all the markers and setting a probability value equal to 0.99. Markers with F_{ST} values whose probabilities fell outside the 0.99 threshold were excluded in subsequent analyses and the F_{ST} recalculated. The procedure was repeated until no probability value of any F_{ST} for any marker fell outside the 0.99 limit. Markers with F_{ST} values whose probabilities fell outside the 0.99 limit from the last analysis were considered as outliers and hence potential candidates for selection.

To investigate whether some of the markers were within protein coding genes, sequences of polymorphic micro- and minisatellites were searched using the BLAST algorithm (Altschul *et al.*, 1990) against the non-redundant nucleotide (BLASTN) and amino acid (BLASTX) genbank sequence databases. An E value of less than $1e^{-10}$ and an alignment length longer than 300 bp for BLASTN and 100 amino acids for BLASTX were used as threshold values and the best matches were selected.

3.3 Results

3.3.1 Diversity indices of the 29 polymorphic micro- and minisatellite markers

Allele sizes in bases for all the markers were determined from graph peaks obtained from electropherograms generated using ABI 3730 genetic analyzer. Figure 12 shows the allele sizes of four different markers (TC1002, TC1227, TC1301 and TC1302) as determined by the genetic analyzer. Locus variability of the 29 polymorphic markers was measured as TNA, Ho, He and PIC (Table 8). TNA per locus ranged from 4 (TC1400b) to 19 (TC440a and TC728b), while He ranged between 0.133 (for TC1002) and 0.719 (for TC705). PIC values ranged from 0.193 (for TC1002) to 0.822 (for TC23). The values of F_{ST}/θ for each locus ranged between 0.103 (TC23) and 0.431

(TC1002). There was a significant and positive correlation between He and TNA (Spearman's *rho* = 0.657, calculated *P* = 0.000) and between He and PIC (Spearmans *rho* = 0.940, calculated *P* = 0.000). TNA was also positively and significantly correlated with PIC (Spearmans *rho* = 0.413, calculated *P* = 0.002). F_{ST} on the other hand was significantly and negatively correlated with He (Spearmans *rho* = -0.434, calculated *P* = 0.019) and TNA (Spearmans *rho* = -0.569, calculated *P* = 0.001) but its correlation with PIC was negative but not significant (Spearmans *rho* = -0.289, calculated *P* = 0.128).

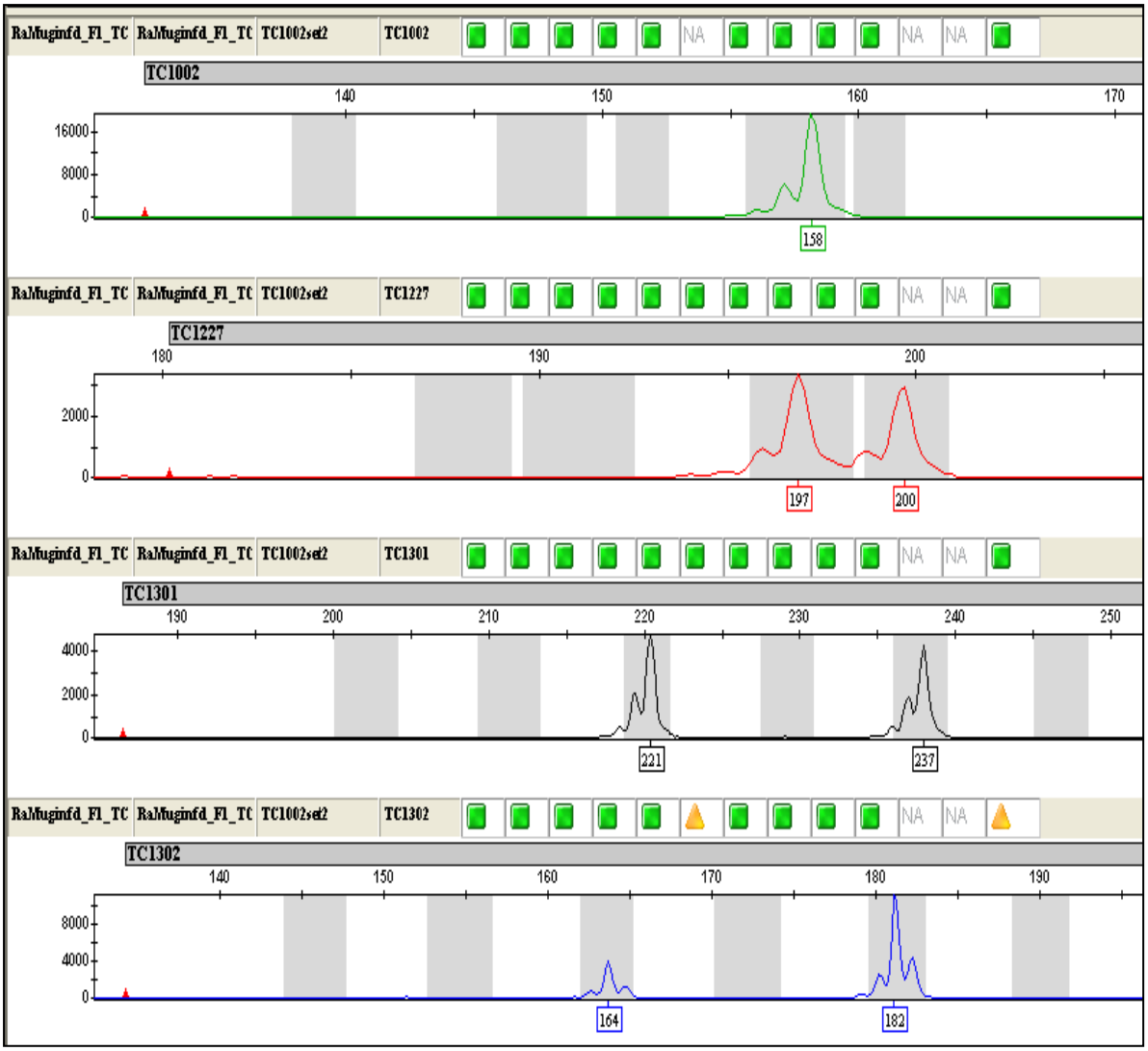


Figure 12: Electropherogram showing marker sizes (in base pairs) of 4 different loci in *R. appendiculatus*

The sizes were determined by the ABI 3730 Genetic Analyser. Muguga *T. parva* infected tick gDNA was used as a template to amplify the loci represented by markers TC1002, TC1227, TC1301 and TC1302. Each pane represents a graph showing the relative concentration of the labeled product in relative fluorescent units (RFUs) and a peak(s) of the predicted size in base pairs. A particular locus was identified by the florescent dye used to label the marker where green represents VIC, red PET, black NED and blue FAM.

Table 8: Diversity indices for the 29 micro- and minisatellite markers generated from 979 tick individuals belonging to 27 populations

Marker	**Expected Heterozygosity (He)**	**Observed Heterozygosity (Ho)**	**Total number of alleles (TNA)**	**Polymorphic information content (PIC)**	**Genetic differentiation (F_{ST})**
TC1002	0.133	0.097	10	0.193	0.431
TC1227	0.255	0.266	7	0.275	0.265
TC1301	0.615	0.436	13	0.764	0.229
TC1302	0.664	0.552	8	0.751	0.164
CD790649	0.553	0.451	13	0.628	0.133
TC1179	0.549	0.464	11	0.664	0.162
TC1303	0.638	0.395	7	0.764	0.247
TC1304b	0.616	0.516	7	0.754	0.231
TC1353	0.328	0.141	7	0.335	0.113
TC1400b	0.436	0.36	4	0.416	0.203
TC1716	0.639	0.634	13	0.737	0.17
TC1968	0.284	0.257	6	0.358	0.312
TC2092	0.629	0.618	12	0.709	0.189
TC1284a	0.281	0.065	6	0.33	0.225
TC1287	0.643	0.307	11	0.678	0.154
TC1307	0.664	0.237	10	0.75	0.174
TC23	0.618	0.802	12	0.656	0.103
TC1398a	0.453	0.455	8	0.458	0.133
TC1398b	0.396	0.313	5	0.421	0.293
TC1399b	0.463	0.433	5	0.459	0.204
TC705	0.719	0.588	16	0.822	0.169
TC2543	0.655	0.616	13	0.748	0.152
TC440a	0.715	0.662	19	0.782	0.121
TC639a	0.497	0.469	6	0.525	0.183
TC875	0.344	0.317	6	0.41	0.32
CD788386	0.692	0.628	18	0.771	0.14
TC2381	0.606	0.322	10	0.656	0.134
TC728b	0.583	0.542	19	0.638	0.134
TC919	0.612	0.37	9	0.689	0.165
Mean	**0.527**	**0.425**	**10**	**0.59**	**0.19**

3.3.2 Suitability of markers for discriminating between tick populations

To examine the usefulness of the 29 micro- and minisatellite markers for population genetic studies, a %PCA was constructed using all the study populations (Figure 13). The decrease in eigenvalues revealed three clusters in the reference typology where the first two axes only (PC1 and PC2) are shown. The first axis separated *R. appendiculatus* populations into two groups, the first one consisted of seven out of eleven laboratory stocks and one field population from South Africa (SN) and the other group comprises of all the field ticks sampled in Kenya and five laboratory stocks including *R. zambeziensis,* which is a different species, but also a vector of *T. parva*. This first axis also separated the four other *rhipicephaline* species into two groups, the first one made up of *R. evertsi* and *R. praetextatus* while the second comprised of *R. pulchellus* and *R. pravus* respectively. The second axis discriminated *R. appendiculatus* and *R. zambeziensis* from the other four *rhipicephaline* species sampled in Kenya (*R. evertsi*, *R. praetextatus*, *R. pulchellus* and *R. pravus*). The third axis was excluded because it did not reveal any clear clusters between the populations analysed and is therefore not informative. The Tv values for each marker are shown in Figure 14. High Tv values indicate a perfect match between the population clusters generated by each marker and the reference typology constructed using all the markers. In other words, high Tv values indicate that structuring among the loci is homogenous and each marker generates a carbon copy of the reference typology when analyzed independently. From Figure 14 it can be seen that the Tv values differ for all the markers and for the different axes of the reference topology. This indicates that some markers are more efficient than others in generating a typology that is close to the overall reference classification and makes it possible to select the most informative markers for further studies. For instance, TC728b had the highest Tv value for the first axis and therefore contributed most to the structuring observed for this axis. The next most informative loci for this axis were

TC919, TC1002, TC23 and TC1398a each of which contributed more than 5% of the total variance. For the second axis, six markers, TC1227, TC875, TC1304b, TC1303, TC1307 and TC1301 were the most informative each contributing more than 5% of the total variance. Markers contributing significantly to the separation of populations for the third axis were different from those that contributed most to the first and second axes. Loci TC1179, TC1301, TC1302, TC1398a, and TC2092 each contributed more than 5% of the variance in the third axis. Clearly the efficiency of markers in structuring populations differs, with markers contributing differently to different axes. Several loci do not contribute significantly to structuring populations on any of the three principal axes while in some cases, the Tv values for each locus sometimes exhibited inconsistencies. For instance TC728b, the most important marker for the first axes was among the least important for the second and third axes respectively. Similarly, TC1227 and TC875 were the most important markers for the second axis, but not for the first and third axes while TC1179 had the highest contribution to the third axis but not to the first and second axes. Thus the most efficient markers for discriminating between populations of *R. appendiculatus* are not necessarily the same as those that are most useful in distinguishing between the closely related *R. appendiculatus*-*R. zambeziensis* complex and other *rhipicephaline* species.

In this study, coordinated %PCA plots, which were drawn on the same scale for each marker and used to define how individual markers separate populations with respect to the common/reference structure, were generated for each marker. Figure 15 shows representative %PCA plots for four markers, TC728b, TC919, TC1227 and TC875 respectively. The first two (TC728b and TC919) were among those that contributed most to the structuring observed for PC1 of the reference typology (Figure 13) and thus contributed significantly in separating *R. appendiculatus* into two major groups while the other two (TC1227 and TC875) were among the most informative for PC2

and thus contributed significantly in discriminating *R. appendiculatus and R. zambeziensis* from the four other *Rhipicephalus* species sampled in Kenya. The first two PC axes of the %PCA for each of the 29 markers are shown in Supplementary Figures 1a, b, and c (Appendix I). In these plots, arrows indicate alleles, with those contributing most to discrimination being linked together by lines. A confidence ellipse ($P = 0.95$) representing the number of individuals sampled is drawn around each population point. From all the plots it can be seen that each marker generates a unique typology. With the exception of TC23, which shows a clear separation between *R. appendiculatus* and other *rhipicephaline* species, no other individual marker provides clear discrimination between *R. appendiculatus* populations or between *R. appendiculatus* and other species in the genus *Rhipicephalus.* In most cases either only one or two populations of *R. appendiculatus* are separated or, one or two non-*appendiculatus* populations are separated. Therefore, when used individually, the markers do not separate the populations as efficiently as they do when used in combination. Most of the populations are structured by a combination of 3 or 4 alleles. The remaining alleles, which are rare, are located close to the center of the %PCA plots and are not useful for discrimination. From this analysis it can be concluded that the key alleles that are useful for population discrimination tend to have high frequencies and therefore represent the most common alleles.

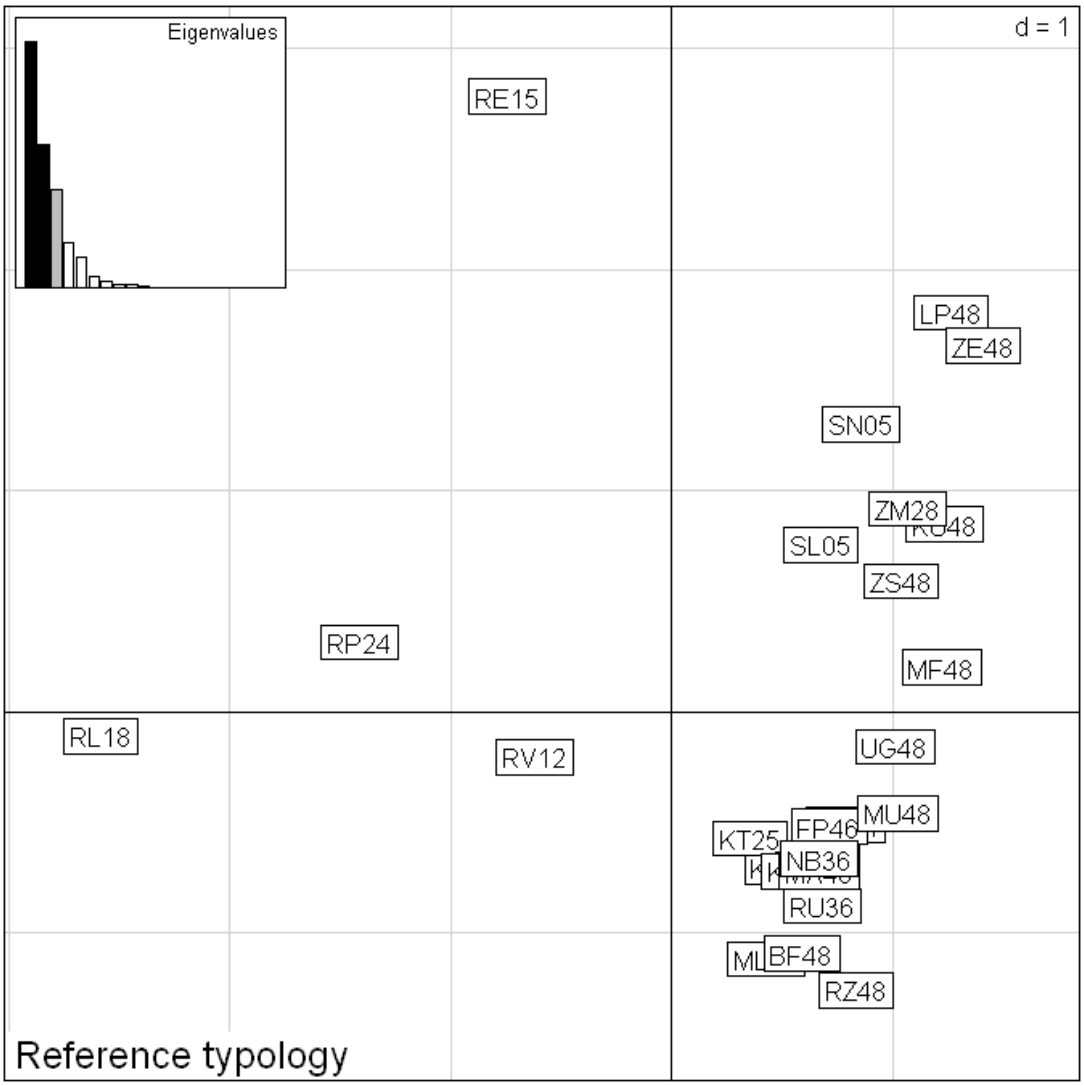


Figure 13: Reference percent (%) PCA plots showing the separation of the various tick populations The plots were generated using MCoA using all the tick species analysed. The first axis separates *R. appendiculatus* populations into two groups. This first axis also separates the four other *rhipicephaline* species into two groups, the first one is made up of *R. evertsi* and *R. praetextatus* while the second comprises *R. pulchellus* and *R. pravus* respectively. For the second axis, six markers, TC1227, TC875, TC1304b, TC1303, TC1307 and TC1301 were the most informative. Loci TC1179, TC1301, TC1302, TC1398a, and TC2092 each contributed more than 5% of the variance in the third axis.

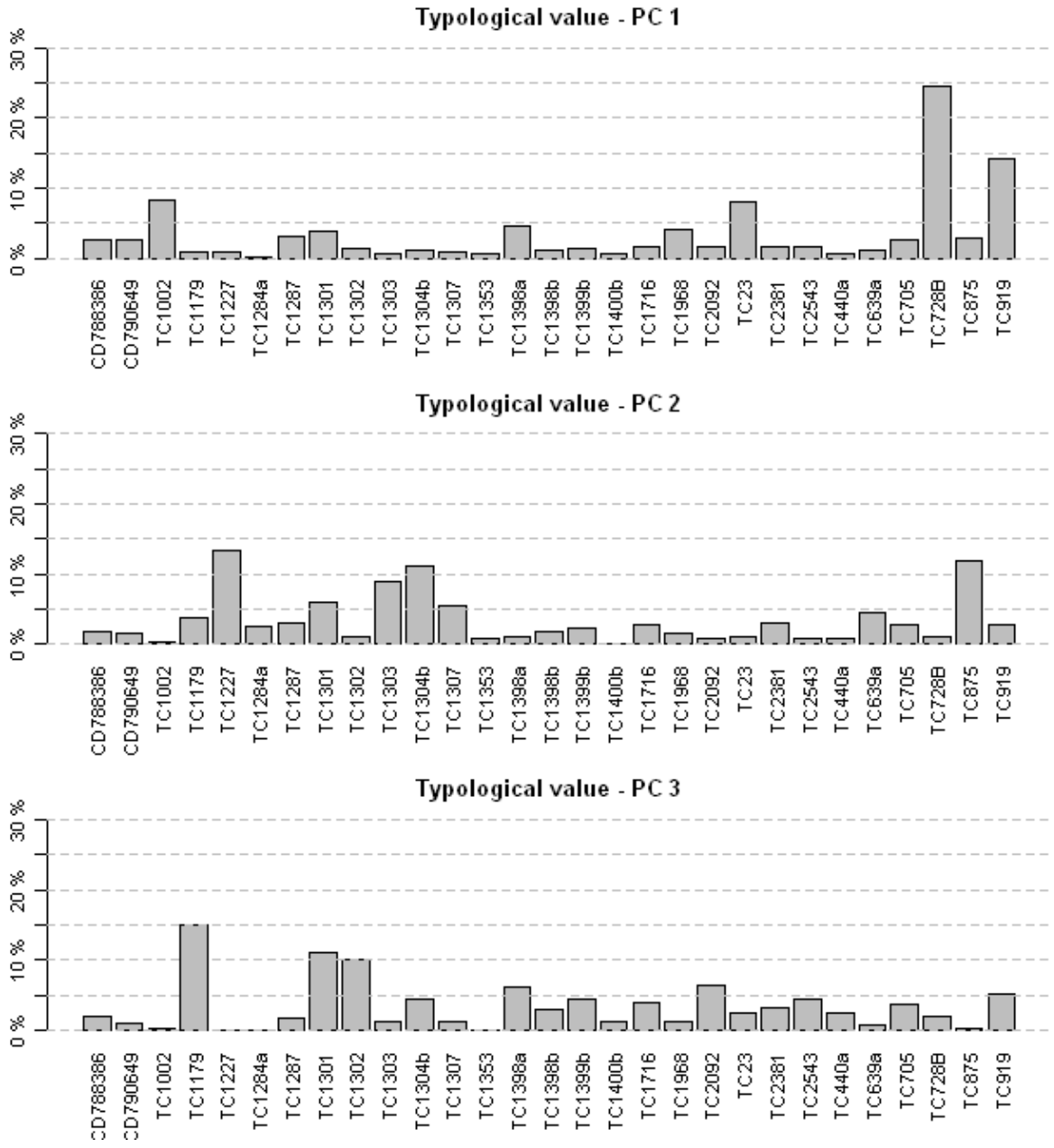


Figure 14: Derived typological values (Tv) for each marker for the first three principal components (PC) of the reference typology/structure

High Tv values indicate a perfect match between the population clusters generated by each marker and the reference typology constructed using all the markers. TC728b had the highest Tv value for the first axis and therefore contributed most to the structuring observed for this axis. The next most informative loci for this axis were TC919, TC1002, TC23 and TC1398a each of which contributed more than 5% of the total variance.

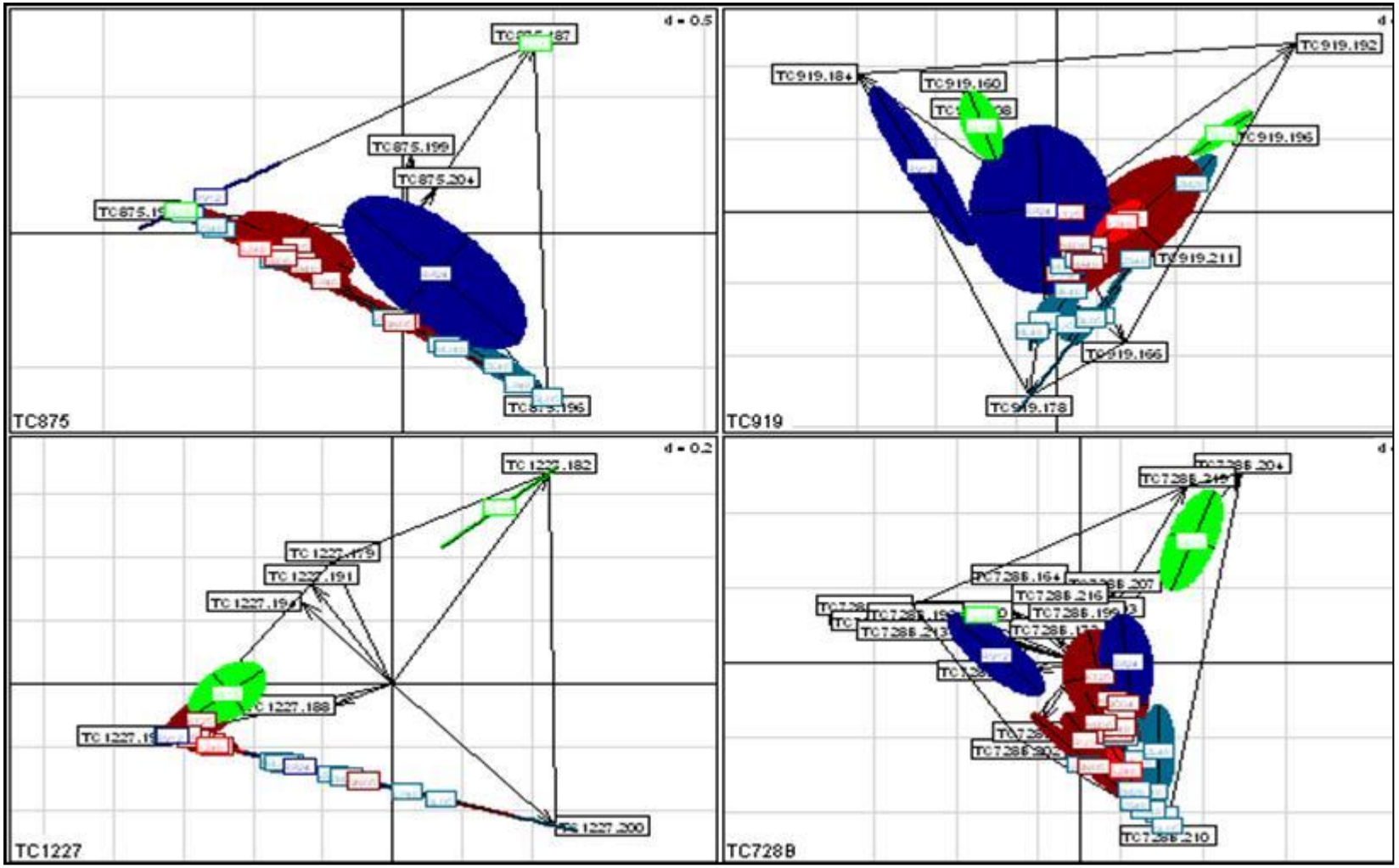


Figure 15: Coordinated %PCA plots for the 4 most discriminating markers

The markers are TC728b, TC919, TC1227 and TC875. The plots show the first two principal components (PC) revealing how well these markers separate the study populations with respect to the common/reference typology. The populations are labeled in their confidence ellipse (P = 0.95), within an envelope formed by alleles (arrows), the most discriminating alleles being joined by lines. The colored ellipses represent the different populations sampled as follows: Red = Lab stock other species (RZ); Dark red = Field ticks (BF, BU, KF, MK, RU, KT, BO, FP, MA, NB, SN); Deep Sky Blue = Lab stocks - *R. appendiculatus* (KH, KU, LP, MF, ML, MU, ZE, ZS, UG, ZM, SL); Dark blue = Field ticks - other species (RP, RV) and Green = Mixed stocks – other species (RE, RL).

3.3.3 Detecting markers under selection

EST-SSRs detect variation in the expressed regions of the genome. There is a possibility of such markers being under selective pressure, either directly or indirectly through linkage to a locus that encodes a trait under selection. Using the FDIST2 program, initial simulations revealed two (TC1002 and TC875) loci that appear to be subject to positive selection and one (TC440a) that may be under balancing selection. Subsequent simulations after excluding the three identified initially revealed one additional locus (TC1284a) as a potential candidate for positive selection (Figure 16). Performing the analysis again following the exclusion of these four loci did not reveal other genes that exhibited signatures of selection.

3.3.4 Protein coding potential of EST containing polymorphic micro- and minisatellite sequences

BLAST analysis revealed that of the 29 EST micro- and minisatellite sequences, seven were present within genes that encode proteins in ixodid tick species. These included; four putative proteins within *Ixodes scapularis* genome (hexosaminidase domain-containing protein, serine palmitoyltransferase, zinc finger protein and gastric mucin), one within the *Rhipicephalus (Boophilus) microplus* EST database (salivary gland metalloprotease) and two from *Rhipicephalus sanguineus* (putative cement protein and salivary gland-associated glycine rich protein). ESTs exhibiting homology to proteins in other species were; Proteasome 26S subunit complex (PSMD1) in *Homo sapiens*, 26S proteasome complex in *Heterocephalus glaber* and two proteins (Actin-dependent regulator chromatin and C-terminal domain of RNA polymerase II polypeptide A) in *Acromyrmex echinatior*. Five of the EST sequences encode genes of hypothetical functions in ticks (four in *Amblyomma maculatum* and one in *Ixodes scapularis*). The remaining 13 sequences did not have any significant protein matches in the databases.

3.3.5 Species specificity of the 29 polymorphic markers

The distribution of alleles across the study populations revealed that the 29 polymorphic markers had alleles that were unique to individual populations of *R. appendiculatus* while nine loci failed to amplify in some populations of *R. appendiculatus* (Table 9). Primers derived from three loci (TC1353, TC1284a and TC440a) did not generate amplicons with most *R. praetextatus* samples, while primers derived from two loci (TC1400b and TC1398b) failed to generate amplicons from *R. pulchellus*. Primers for TC1284a and TC705 failed to amplify in *R. pravus* and *R. evertsi* respectively. Thirteen loci had several alleles that were specific to the four additional *rhipicephaline* species included in the study.

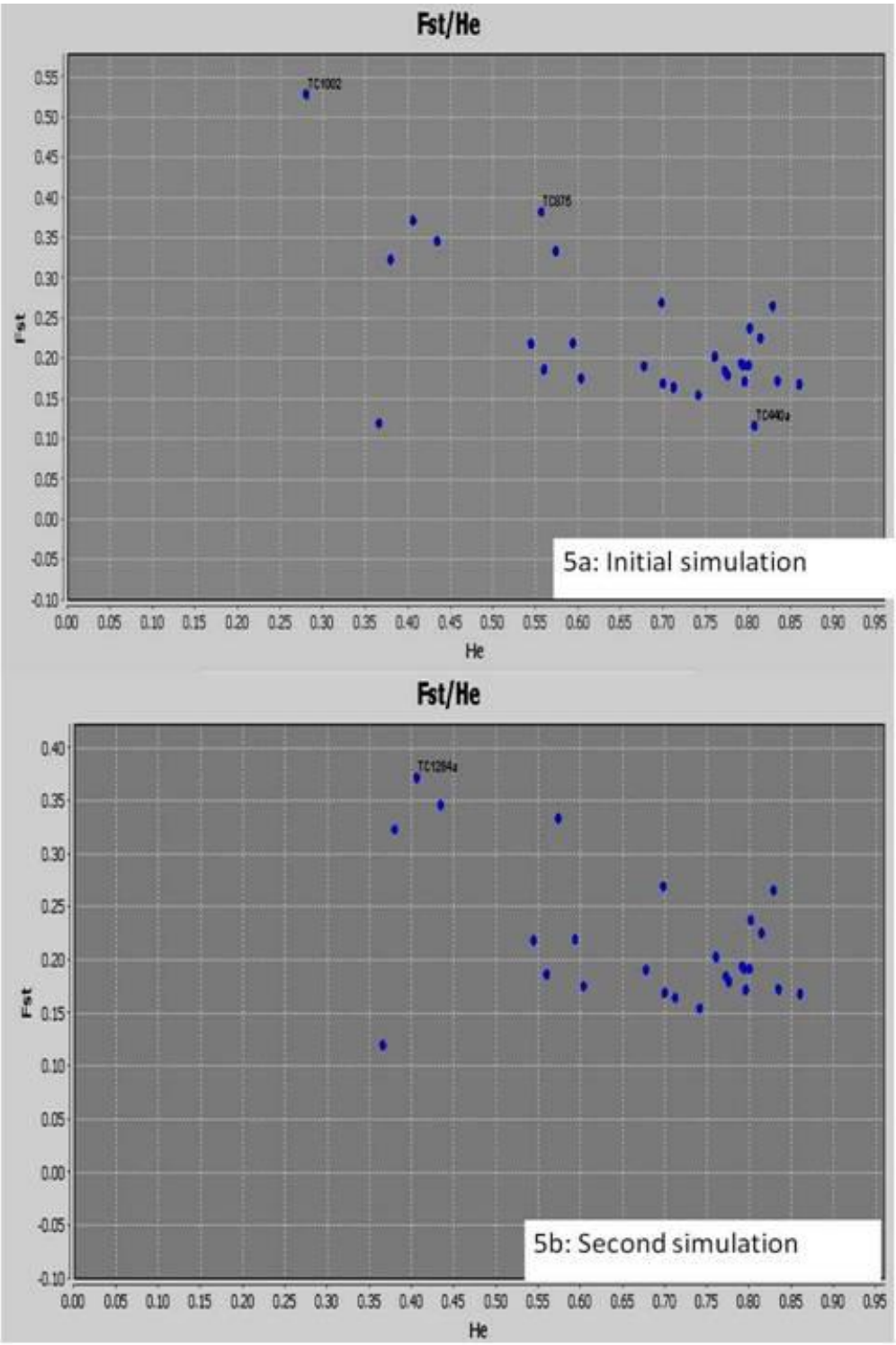


Figure 16: Graphical distributions of the Fst/He values of each marker in the 1st and 2nd simulation runs

The initial simulations revealed two (TC1002 and TC875) loci that appear to be subject to positive selection and one (TC440a) that may be under balancing selection. The 2nd simulation revealed one additional locus (TC1284a) as a potential candidate for positive selection.

Table 9: Number of micro- and minisatellite markers which showed species-specific/unique alleles different species of ticks

Species of tick	SSR with species-specific alleles	Loci that gave no PCR product
R. appendiculatus	All 29 SSRs in different populations	9 (TC1284a, TC1398a, TC1400b, TC1398b, TC1399b, TC705, TC728B, TC2381 and TC1179)
R. praetextatus	7 (TC1002, TC1301, TC1179, TC639a, TC1398a, TC1398b, CD788386)	3 (TC1284a, TC1353, TC440a)
R. pravus	2 (TC440a, CD788386)	1 (TC1284a)
R. evertsi	2 (TC1301, TC1968)	1 (TC705)
R. pulchellus	2 (TC1301, TC705)	2 (TC1398b, TC1400b)

3.4 Discussion

Micro- and minisatellites have become markers of choice for genotyping because of their high level of polymorphism and ease of assaying with high reproducibility using PCR amplification. These markers can be directly typed by PCR amplification (Holmes *et al.*, 1993) followed by resolution of the alleles by electrophoresis on high resolution denaturing acrylamide gels or capillary electrophoresis genetic analyzers (Litt & Luty, 1989; Weber & May, 1989). To determine the polymorphism and reproducibility of the 42 micro- and minisatellite markers that were found amplifiable, further optimization and validation using fluorescently labeled primers with subsequent detection of the amplified products on fluorescence-based capillary electrophoresis was done. Of the forty two VNTR loci evaluated in this study twenty micro and nine minisatellites were polymorphic while six were monomorphic. These six monomorphic markers suggest that they are highly conserved within and between *rhipicephaline* species. Though microsatellites from certain

ixodid tick species including *Rhipicephalu*s (*Boophilus) microplus* (Chigagure *et al*., 2000, Koffi *et al.*, 2006) and *Ixodes ricinus* (Røed *et al.*, 2006) have previously been isolated and characterized, this study is the first to report polymorphic micro- and minisatellite markers that have been evaluated for estimating genetic diversity within and between a significant number of ixodid tick populations.

The loci analysed showed considerable variability across the study populations with 18 and 19 loci having a He and PIC of above 0.50 respectively. The fact that some markers had high and others low levels of variability implies that using all the 29 markers characterized in this study for population genetic analyses should result in a balanced assessment of the levels of genetic diversity and divergence within and between populations. The positive correlations between He, TNA and PIC implies that any of these indices can be used to select markers for population genetic studies independently and that the most polymorphic markers are also high in diversity and therefore most likely to be useful in discriminating between populations. On the other hand, the negative correlations between He, TNA and PIC with F_{ST} suggest that the most diverse markers do not necessarily give the best estimate of genetic differentiation between populations.

Using the MCoA approach, Laloë *et al.* (2007) determined the efficiency of 30 microsatellite markers in structuring different breeds of cattle from France and West Africa. This study used the same approach to determine the efficiency of 29 markers to discriminate between populations of *R. appendiculatus* and between different species of *Rhipicephalus*. The first axis of the %PCA separated populations of *R. appendiculatus, R. zambeziensis* and four other *rhipicephaline* species from Kenya into two groups (Figure 13). For this axis, four markers (TC1002, TC23, TC728b and TC919) each with a Tv value of above 5 % appear to be the most efficient in separating populations

(Figure 13). However, when the efficiency of individual markers was assessed using the coordinated %PCA, TC23 was the only marker that clearly separated *R. appendiculatus* from the other *rhipicephaline* populations. Despite the power of TC23 for between-species differentiation, it typically identified multiple alleles which could not be scored easily and had a high heterozygosity. In post PCR, analysis, micro- and minisatellite markers that give strong clean signal with a minimum of stray bands are desirable (Gastier *et al.*, 1995). This marker might therefore be difficult to use in routine genotyping of field populations. With the exception of *R. zambeziensis* which clusters together with populations of *R. appendiculatus*, the second axis separated all populations of *R. appendiculatus* from the other four *Rhipicephalus* species sampled in Kenya. *R. zambeziensis* is morphologically similar to *R. appendiculatus* and the markers used here could not discriminate between them. This is because they had many alleles in common for most of the markers.

In this study, samples of *R. appendiculatus* were obtained from field and laboratory stocks. While the %PCA separated *R. appendiculatus* into two groups, these did not correspond to either the field or laboratory stocks. For instance population SN, a field stock, clusters together with laboratory bred stocks while UG and MU, two laboratory stocks, cluster with the field stocks. Therefore, though PC1 clusters *R. appendiculatus* into two groups, it fails to discriminate between laboratory and field stocks and thus it would be interesting to determine what, other than the origin of the samples contributes to the discrimination of the groups of *R. appendiculatus* populations by PC1. For each axis, the markers contributing significantly to structuring of populations were different. Therefore markers that are efficient in separating *R. appendiculatus* populations are not necessarily efficient in separating *R. appendiculatus* from other *rhipicephaline* species and especially *R. zambeziensis*. Therefore the choice of markers will be best determined by the species of interest and the research questions being addressed.

Each of the 29 micro- and minisatellite loci had certain alleles that were specific to populations of *R. appendiculatus* while others were species-specific. On the other hand, nine markers did not yield amplicons with some populations of *R. appendiculatus*. This suggests that either the priming sites for these loci have mutated leading to a change in the nucleotide sequence on the primer binding site or the loci are population specific. If this is the case, these markers can be used to discriminate specific populations of *R. appendiculatus* from the others. Twenty-three out of the 29 polymorphic EST markers also generated amplicons in *R. zambeziensis, R. praetextatus, R. pravus, R. pulchellus, and R. evertsi* and certain alleles in some of these markers were specific to these other species in the genus *Rhipicephalus*. These 23 markers together with the six that did not yield any amplicons in other *Rhipicephalus* species can potentially be used to discriminate and clarify genetic differentiation between populations in cases where morphological characters are not informative. They also have potential to be included in multiplex assays to discriminate between species. Therefore these EST markers though developed from *R. appendiculatus*, can be used to perform inter-species population genetics analysis as has been done in other species such as birds (Primmer *et al.,* 1996), mosquitoes (Kamau *et al.,* 1999), bovines and caprines (Pepin *et al.,* 1995) including humans and non-human primates (Deka *et al.*, 1994).

Three markers (TC1002, TC1284a and TC875) are likely to be under positive selection and TC440a is likely to be under stabilizing selection. There is a need to further evaluate these four markers to establish if they are linked to any trait or region of the genome under selection. If further information supports the hypothesis that they are under selection, then these markers may be excluded from future population genetics studies. However, Ellis & Burke (2007) suggest that although selection may be operating on a small number of EST-SSRs, this drawback can be overcome by using a large number of markers to minimize the potential bias of non-neutral loci on

estimates of population genetic parameters. Given that 11 EST-SSRs were in genes with putative functions based on homology with other organisms, the expectation is that some of them may be subject to selection. However it was not possible to investigate this possibility in detail in the absence of full length gene and protein sequences, which would have allowed the, potentially more definitive, analysis of the relative frequency of synonymous and non-synonymous substitutions.

This study provided an optimum marker system that can be applied in molecular studies of *R. appendiculatus*. The micro- and minisatellite markers described here are a valuable resource for studying the genetics of *R. appendiculatus* and related species of the genus *Rhipicephalus* because of their ability to discriminate both within *R. appendiculatus* and between other *rhipicephaline* species. It is recommended that the following panel of markers: TC728b, TC919, TC1002, TC23 and TC1398a, which contributed more than 5% of the total variance in the first PC, could be useful for intra-species discrimination while a different panel of markers: TC1227, TC875, TC1304b, TC1303, TC1307 and TC1301, are more useful for inter-species discrimination within the *rhipicephaline* complex. These markers represent the first molecular tools for defining *R. appendiculatus* genotypes, their distribution and degree of geographic differentiation. They will also be useful for defining parameters of relevance to the epidemiology of ECF in the field. This includes possible interactions between tick and pathogen genotypes, specifically whether differences exist in levels of *T. parva* challenge at different geographic locations that may relate to tick genotypes, as has been demonstrated in laboratory experiments (Ochanda *et al*., 1998), which might be mediated directly by vector competence factors at the tick/parasite interface or indirectly via differences in adaptation to different environments affecting tick densities. Future researchable areas include attempts to understand the biology underlying the variable contribution of markers to population and species differentiation. The ability to discriminate at higher resolution within the

rhipicephaline complex could also ultimately provide new insights into the distribution of different tick species in Africa and perhaps even result in the identification of novel tick vectors of *T. parva.*

It is interesting to note that the %PCA discriminates some stocks of laboratory and field ticks. Such differences have been noted previously using the single copy Ra86 gene (Kamau *et al.,* 2011). This is of more than theoretical importance since the Muguga laboratory stock which is used to produce the infection and treatment (ITM) live vaccine for immunization against ECF (Radley *et al.*, 1981, reviewed by Di giulio *et al*., 2009) has very likely diverged from the original field material with unknown implications for the vaccine.

Results from this chapter have been published in the article;

Esther G. Kanduma, Joram M. Mwacharo, Jack D. Sunter, Inosters Nzuki, Stephen Mwaura, Peter W. Kinyanjui, Michael Kibe, Heloise Heyne, Olivier Hanotte, Robert A. Skilton, Richard P. Bishop. (2012). Micro- and minisatellite-expressed sequence tag (EST) markers discriminate between populations of *Rhipicephalus appendiculatus*. *Ticks and Tick-borne Diseases*. Volume 3, Issue 3, Pages 128–13

3.5 References

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology* **215,** 403–410.

Anderson, J. A., Churchil, G. A., Autrique, J.E., Tanksley, S. D., Sorrells, M. E. (1993). Optimizing parental selection for genetic linkage maps. **A., Churchil, G. A., Autrique, J. E., Tanksley, S. D. Sorrells, M. E.** (1993). Optimizing parental selection for genetic linkage maps. *Genome* **36**, 181-186.

Bailey, K.P. (1960). Notes on the rearing of *Rhipicephalus appendiculatus* and their infection with *Theileria parva* for experimental transmission. **P.** (1960).*Bulletin of Epizoonotic Diseases of Africa,* **8,** 33–43.

Beati, L & Keirans, J, E. (2001). Analysis of the systematic relationships among ticks of the genera *Rhipicephalus* and *Boophilus* (acari: ixodidae) based on mitochondrial 12S ribosomal DNA gene sequences and morphological characters. *Journal of parasitology*, **87(1),** 32–48.

Beaumont, M. A and Richard, N. A. (1996). Evaluating loci for use in the genetic analysis of population structure. *Proceedings of the Royal Society Biological Sciences* **263,** 1619-1626.

Berkvens, D. L., Peg**ram, R. G. and Brandt, J. R. A.** (1995). A study of the diapausing behaviour of *Rhipicephalus appendiculatus* and *Rhipicephalus zambeziensis* under quasi-natural conditions in Zambia. *Medical Veterinary Entomology* **9,** 307–315.

Botstein, D., White, L.W., Skolnick, M.,Botstein, D., White, L. W., Skolnick, M. and Davis, R. W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. American *Journal of Human Genetics* **32,** 314-331.

Chaka, G., Billiouw, M., Geysen, D. M. and Berkvens, D, L. (1999). Spatial and temporal variation in *Rhipicephalus appendiculatus* size in eastern Zambia. *Tropical Medicine and International Health* **4(9),** 43-48.

Chessel, D., Dufour, A. B. and Thioulouse, J. (2004). The ade4 package-I: one-table methods. *R News* **4,** 5-10.

Chigagure, N. N., Baxter, G. D. and Barker, S. C. (2000). Microsatellite loci of the cattle tick *Boophilus microplus* (Acari: Ixodidae). *Experimental Applied Acarology* **24 (12),** 951-956.

Cutulle, C., Jonsson, N. N., Seddon, J. (2009). Population structure of Australian isolates of the cattle tick *Rhipicephalus (Boophilus) microplus*. *Veterinary Parasitology*, **161,** 283–291.

Deka, R., Shriver, M. D., Yu, L. M. and Jin, L. (1994 Conservation of human chromosome 13 polymorphic microsatellite (CA)n repeats in Chimpanzees. *Genomics.* **22(1),** 226-230.).

Di giulio, G., Godelieve, L., Morzaria, S., Oura, C. and Bishop, R. P. (2009). Live immunization against East Coast fever-current status. *Trends in parasitology* **25(2),** 85-92.

Ellis, J. R and Burke, J. M. (2007). EST-SSRs as a resource for population genetic analyses. *Heredity* **99,** 125–132.

Gastier, J. M., Pulido, J. C., Sunden, S., Brody, T., Buetow, K. H., Murray, J. C. & Duyk, G. M. (1995). Survey of trinucleotide repeats in the human genome: assessment of their utility as genetic markers. *Human molecular genetics*, **4(10),** 1829-1836.

Gooding, R. H and Krafsur, E. S. (2005). Tsetse genetics: contributions to biology, systematics, and control of tsetse flies. *Ann. Rev. Entomol*. 50, 101–123. **S.** (2005). Tsetse genetics: contributions to biology, systematics, and control of tsetse flies. *Annual Reveiw of Entomology* **50,** 101–123.

Goudet, J. (2001). FSTAT, a program to estimate and test gene diversities and fixation indices version 2.9.3. **Irvin, A. D and Brocklesby, D. W.** (1970). Rearing and maintaining *Rhipicephalus appendiculatus* in the laboratory. *J. Inst. Anim Technol.* **21,** 106–112.

Hasle, G., Røed, K. H., & Leinaas, H. P. (2008). Multiple paternity in Ixodes ricinus (Acari: Ixodidae), assessed by microsatellite markers. *Journal of Parasitology*, **94(2),** 345-347.

Kamau, L., Mukabana, W. R., Hawley, W. A., Lehmann, T., Irungu, L. W., Orago, A. A. S., Collins, F. H. (1999). Analysis of genetic variability in *Anopheles arabiensis* and *Anopheles gambiae* using microsatellite loci. *Insect Mol. Biol*. **8(2),** 287-298.

Kamau, L., Skilton, R. A., Odongo, D. O., Mwaura, S., Githaka, N., Kanduma, E., Obura, M., Kabiru, E., Orago, A., Musoke, A. and Bishop, R. P. (2011**).** Differential transcription of two highly divergent gut-expressed Bm86 antigen gene homologues in the tick *Rhipicephalus appendiculatus* (Acari: Ixodida). *Insect Molecular Biology* **20(1),** 105–114.

Koffi, B. B., Risterucci, A. M., Joulia, D., Durand, P., Barre, N., de Meeuˆ s, Chevillon, C. (2006). Characterization of polymorphic microsatellite loci within a young *Boophilus microplus* metapopulation. *Molecular Ecology Notes* 6, 502–504.

Kubasu, S. S. (1992). The ability of *Rhipicephalus appendiculatus* (Acarina: Ixodidae) stocks in Kenya to become infected with *Theileria parva*. *Bulletin of Entomology Research* 82, 349-353.

Labruna, M., Naranjo, V., Mangold, A., Thompson, C., Estrada-Peña, A., Guglielmone, A & De la Fuente, J. (2009). Allopatric speciation in ticks: genetic and reproductive divergence between geographic strains of *Rhipicephalus (Boophilus) microplus*. *BMC evolutionary biology*, **9(1),** 46.

Laloë, D., Jombart, T., Dufour A.B. and Moazami-Goudarzi, K. (2007). Consensus genetic structuring and typological value of markers using multiple co-inertia analysis. *Genetics Selection Evolution,* **39(5),** 545-567.

Litt, M., & Luty, J. A. (1989). A hypervariable microsatellite revealed by in vitro amplification of a dinucleotide repeat within the cardiac muscle actin gene. *American journal of human genetics*, ***44*(3),** 397.

Mtambo, J., Madder, M., Van Bortel, W., Geysen, D., Berkvens, D. and Backeljau, T. (2007a). *Rhipicephalus appendiculatus* and *R. zambeziensis* (Acari: Ixodidae) from Zambia: a molecular reassessment of their species status and identification.*Exp. Experimental Applied Acarolology* **41,** 115-128.

Mtambo, J., Madder, M., Van Bortel, W., Geysen, D., Berkvens, D. and Backeljau, T. (2007b). Genetic variation in *Rhipicephalus appendiculatus* (Acari: Ixodidae) from Zambia: correlating genetic and ecological variation with *Rhipicephalus appendiculatus* from eastern and southern Africa. *Journal of vector Ecology* **32(2),** 168-75.

Norris, D. E., Shurtleff, A. C., Yeya, T. T. and Lanzaro G. C. (2001). Microsatellite DNA polymorphism and heterozygosity among field and laboratory populations of *Anopheles gambiae s.s.* (*Diptera: Culicidae*). *Journal of Medical Entomolog,* **38(2),** 336-340.

Norval, R. A. F., Lawrence, J. A., Young, A. S., Perry, B. D., Dolan, T. T., Scott, J. (1991). *Theileria parva*: influence of vector, parasite and host relationships on the epidemiology of theileriosis in southern Africa. *Parasitology* **102,** 347–356.

Norval, R. A. F., Perry, B. D. and Young, A. S. (1992). The epidemiology of Theileriosis in Africa. Academic Press Ltd London & San Diego.

Ochanda, H., Young, A. S., Medley, G. F. and Perry, B. D. (1998). Vector competence of 7 rhipicephalid tick stocks in transmitting 2 *Theileria parva* parasite stocks from Kenya and Zimbabwe. *Parasitology,* **116,** 539-545.

Ouma, J. O., Marquez, G. M. and Krafsur, E. S. (2006). Microgeographical breeding structure of the tsetse fly *Glossina pallidipes* in south-western Kenya. *Medical and Veterinary Entomology,* **22,** 222–227.

Park S. D. E. (2001). The Excel Microsatellite Toolkit. http://oscar.gen.tcd.ie/~sdepark/mstoolkit/index.php

Pepin, L., Amigues, Y., Lepingle, A., Berthier, J. L., Bensaid, A and Vaiman, D. Pepín, (1995). Sequence conservation of microsatellites between *Bos taurus* (cattle), *Capra hircus* (goat) and related species. Examples of use in parentage testing and phylogeny analysis. *Heredity*. **74 (1),** 53-61.

Primmer, C. R., Moeller, A. P. and Ellegren, H. (1996). A wide-ranging survey of cross-species microsatellite amplification in birds. *Molecular Ecology* **5(3),** 365-378.

Radley, D. E. (1981). Infection and treatment method of immunization. In Advances in the Control of Theileriosis. A.D. Irvin, M.P. Cunningham & A.S. Young, Eds.: 227-237. Martinus Nijhoff, The Hague.

Røed, K. H., Hasle, G., Midthjell, V., Skretting, G. and Leinaas, H. P. (2006 Identification and characterization of 17 microsatellite primers for the tick, *Ixodes ricinus*, using enriched genomic libraries. *Molecular Ecology Notes,* **6(4),** 1165–1167.

Ruiz-López, M. J., Chaskelson, S., Gompper, M. E., & Eggert, L. S. (2012). Multiple paternity in the American Dog Tick, *Dermacentor variabilis* (*Acari: Ixodidae*). *Journal of Parasitology*, **98(3),** 498-501.

Schuelke, M. (2000). An economic method for the fluorescent labeling of PCR fragments. *Nature biotechnology*, *18*(2), 233-234.**Shaw, M. K., and Young, A.S.** (1994). The biology of Theileria species in ixodid ticks in relation to parasite transmission. *Advances in Disease Vector Reseach* **10,** 23-63.

Speybroeck, N., Madder, M., Thulke, H.H., Mtambo, J., Tirry, L., Chaka, G., Marcotty, T. and Berkvens, D. (2004). Variation in body size in the tick complex *Rhipicephalus appendiculatus/Rhipicephalus zambeziensis*. *Journal of Vecor Ecology,* **29(2),** 347-54.

Speybroeck, N., Madder, M., Van Den Bossche, P., Mtambo, J., Berkvens, N., Chaka, G., Mulumba, M., Brandt, J., Tirry, L. and Berkvens, D. (2002). Distribution and

phenology of ixodid ticks in southern Zambia. *Medical and Veterinary Entomology* 16(4), 430-41.

Weber, J. L., & May, P. E. (1989). Abundant class of human DNA polymorphisms which can be typed using the polymerase chain reaction. *American journal of human genetics*, ***44*****(3),** 388.

Weir, B. S. and Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution,* **38,** 1358–1370.

CHAPTER FOUR: ASSESSMENT OF THE GENETIC DIVERSITY AND POPULATION STRUCTURE WITHIN AND BETWEEN POPULATIONS OF *Rhipicephalus appendiculatus*

4.0 Abstract

The interaction between parasites and their hosts may select for unique mutations with major micro/macro-evolutionary adaptive consequences with implications on effective parasite and pathogen control. *Rhipicephalus appendiculatus* is the most economically important tick vector in eastern and southern Africa and is a vector of several disease causing pathogens in domestic and wild animals, yet its underlying genetic variation and structure has not been explicitly studied. To investigate whether host species and geographic location shape the genetic and phylogeographic structure of *R. appendiculatus*, micro- and minisatellite genotypes from 392 individual ticks from 10 populations sampled from animal hosts (cattle and wildlife) and vegetation across different geographic locations in Kenya were analyzed. Allele data of 460 individual ticks from 10 populations that have been bred and maintained as closed laboratory stocks for over 20-30 years and 5 other species of *rhipicephaline* were also analyzed. Estimates of genetic and allelic diversity were high for field ticks though a low level of genetic variation among the 10 populations of field ticks (F_{ST} = 0.014 ± 0.002) was observed. Their spatial genetic structure was characterized by a significant pattern of lack of isolation by distance and no distinct genetic clusters corresponding to their geographic location and/or host species was observed. This contrasted markedly with *R. appendiculatus* ticks maintained as laboratory stocks (F_{ST} = 0.248 ± 0.015) and the five other *rhipicephaline* species (F_{ST} = 0.368 ± 0.032), which were strongly differentiated into distinct genetic groups by STRUCTURE and discriminant analysis of principle components (DAPC) analyses. Of the 10 laboratory bred *R. appendiculatus* populations, 8 diverged significantly from the field populations as they were assigned to different clusters by both STRUCTURE and DAPC indicating that they are genetically different. The results suggest that for most parasites, their genetic structure is linked to the vagility of their hosts, and that parasite-host systems provide a rich source of variability in population genetic structure. These results have direct implications for the coevolutionary interactions of *R. appendiculatus* and its hosts with subsequent consequences for the epidemiology of tick-borne diseases. The divergence between field and laboratory bred *R. appendiculatus* stocks has implications on the design and development of effective control strategies for *R. appendiculatus* and its different disease causing pathogens.

4.1 Introduction

R. appendiculatus is the most economically important tick vector in eastern and southern Africa and is a vector of several disease-causing pathogens in domestic and wild animals, yet its underlying genetic variation and structure has not been explicitly studied. The twenty nine polymorphic micro- and minisatellite markers (Kanduma *et al.*, 2012) were applied to elucidate the population structure and phylogeography of *R. appendiculatus* populations in Kenya as well as the genetic diversity of laboratory tick colonies. *R. appendiculatus* is an important ectoparasite of both domestic and wild ruminants (Norval & Perry, 1992). The tick mainly infests cattle, goats, sheep and dogs. The main wildlife reservoir for the tick is the African Buffalo (*Syncerus caffer*) which form reservoirs for *T. parva* parasites (Walker *et al*., 2005). It is also infests elands, waterbucks, greater kudus and antelopes. Understanding the genetic diversity of *R. appendiculatus* is important because interactions between hosts and parasites may favor unique mutations resulting in genetic adaptation and divergence between parasites (De Meeûs *et al.*, 1998; Gandon *et al.*, 1998). Host-parasite interactions may select for unique mutations with major micro/macro-evolutionary adaptive consequences with implications on effective parasite and pathogen control. Ecological divergence and niche specialization are common among parasites that are reproductively isolated in time and space. Parasites may therefore be spatially structured between individuals or populations of the same host and between different hosts (De Meeûs *et al.*, 1998; Clayton *et al.*, 2003). The host's physiological, behavioral and demographic variability may influence significantly the genetic landscape of parasites with limited dispersal ability such as ticks (Price, 1980; Poulin & George-Nascimento, 2007). For such parasites, their short and long-range dispersal depends to a large extent on the hosts' dispersal capacity and this may shape their genetic diversity and structure.

The local dynamics of *R. appendiculatus* can therefore play an important role in host population dynamics and in the co-evolutionary interactive outcomes between host, parasite and pathogens. The geographic distribution and population dynamics of *R. appendiculatus* are influenced to a great extent by biophysical barriers, ecological and climatic factors (Lessard *et al.*, 1988, 1990; Perry *et al.*, 1990; Norval *et al*., 1991; Estrada-Peña, 2001). Differences in diapause induction (Madder *et al.*, 1999), body size (Norval *et al.*, 1992; Shaw & Young 1994), histology of salivary glands (Walker & Fletcher, 1990), vector competence (Ochanda *et al.*, 1998), response to acaricides (Chigagure *et al.*, 2000) and in polymorphism of salivary gland proteins (Wang *et al*., 1999) suggest the existence of diverse tick populations.

Several stocks of *R. appendiculatus* have been bred and maintained for laboratory experiments, as stable standardized stocks for comparison purposes and as representatives of field populations. For instance, the standard laboratory Muguga strain (*R. appendiculatus*) has been used to produce the only available vaccine for ECF (Di Giulio, 2009). This is a live *T. parva* sporozoite vaccine that is produced using a pool of three different *T. parva* strains obtained from thousands of infected *R. appendiculatus* (Radley, 1981). Previously, assessment of the biology of laboratory stocks of *R. appendiculatus* has revealed differences in infection rates (Young *et al.*, 1995), susceptibility and efficiency to host *T. Parva* pathogens (Odongo *et al.*, 2009; Ochanda *et al.*, 1998). The genetic diversity and structure of these laboratory bred tick stocks *vis a vis* their field counterparts has never been ascertained. An understanding of the genetic make up of laboratory stock that is used in the generation of the live *T. parva* vaccine is important for vaccine standardisation and quality control. However, during vaccine development, colony genetics are rarely considered. In laboratory colonies of *Aedes egypti*, significant differences in genetic diversity and variation within and between strains has been reported (Norris *et al* 2001).

To develop effective control strategies for vector-borne diseases either through vector control and/or anti-vector/anti-pathogen vaccination strategies (Dai *et al.,* 2009; Gillet *et al.,* 2009), understanding the genetics of vector populations is important. Factors such as high host vagility, low host specificity, and frequent metapopulation extinction and recolonization events, could all act to reduce genetic structure in parasite populations while isolating factors such as sedentary hosts, host patchiness in space and time all promote structuration (Nadler, 1995). Population genetics will provides information on the taxonomic status of the vector species, their spatial limits and the extent of gene flow among host populations (McCoy, 2008; Noureddine *et al.,* 2011).

Microsatellite analysis has been used to show changes in population structure and development of resistance to amitraz in generations of *R. microplus* ticks created by the admixture and hybridization of amitraz and synthetic pyrethroids (SP) resistant and susceptible ticks (Corley *et al.,* 2013). Identification of local tick populations on domestic or maintenance hosts enhanced integrated management of tick borne diseases and led to the eradication of *R. appendiculatus* from South Africa (Gutsche, 1979; Cranefield, 1991) as well as *R. annulatus* and *R. microplus* from USA (George, 1989). Thus improvement of tick control can be enhanced by better understanding of tick population dynamics influenced by drivers of tick distribution, aggregation and stability (Walker, 2011). Whilst several studies have investigated the population dynamics, ecology and biology of *R. appendiculatus* (Norval & Perry, 1990, Norval *et al.*, 1992 and Randolph, 1994, 2004), there is a paucity of studies that have attempted to look at the genetic variability and describe the phylogeographic structure of *R. appendiculatus* populations.

Current efforts at controlling ticks attempt to identify tick antigens that can be exploited as possible vaccine candidates (reviewed in Willadsen, 2004 and Parizi *et al.*, 2012). A vaccine incorporating Bm86, a *Boophilus microplus* gut antigen, has been used successfully to control ticks of the *Boophilus* species (Willadsen *et al.,* 1995). However, some strains of *B. microplus* from different geographic regions have shown variation in the efficacy of protection against the Bm86 vaccine (Garcia-Garcia *et al.,* 1999). Therefore, an understanding of population structure of ticks is essential for the development of effective anti-tick vaccines.

Though microsatellites have been extensively been used to study human disease vectors such as tsetse flies and mosquitoes (de Rosas *et al.*, 2008; Sharakhov *et al.*, 2003, Ouma *et al.,* 2007; Krafsur, 2009, Loaiza *et al.*, 2012, Bouyer *et al.*, 2009 and Solano, 2010), few studies have reported the use microsatellites in population studies of tick vectors. These markers have been applied to show paternity in *I. ricinus* (Gunnar *et al.,* 2008) and *Dermacentor variabilis* (Ruiz-López *et al.*, 2012), genetic polymorphism in *Rhipicephalu*s (*Boophilus) microplus* (Labruna *et al.*, 2009) and genetic structure of amitraz-resistant and susceptible *R. microplus* populations in Queensland, Australia (Cutulle *et al.,* 2009). Though *R. appendiculatus* transmits the most economically important cattle disease in 11 countries in sub-saharan Africa, its genetic diversity and population structure has not been elucidated due lack of molecular markers. Knowledge of the population genetic structure and variability of *R. appendiculatus* is important in understanding its dispersal and transmission dynamics and those of its pathogens. It will improve the design of optimal control strategies and for the development of effective anti-*R. appendiculatus* vaccines. Allele data generated from 979 individual ticks using the twenty nine polymorphic EST based markers (Table 7, Chapter 3, Section 3.2.5; Kanduma *et al.*, 2012) was analysed to evaluate the genetic diversity and population structure of *R. appendiculatus* from different host species and geographic areas in

Kenya. Data from 10 laboratory stocks bred and maintained as closed stock for between 20-50 years was included. Some of these laboratory stocks have been used as representatives of field ticks in laboratory experiments and in the development of vaccines against ECF (Young, 1995, Radley *et al.*, 1981). This analysis was done to investigate whether *R. appendiculatus* diversity is spatially structured across its geographic range and between different host species and whether the genetic landscape of laboratory bred stocks is still representative of their field counterparts.

4.2. Materials and methods

4.2.1 Experimental samples and genomic DNA extraction

A total of 979 individuals comprising of 862 *R. appendiculatus* ticks and 117 of other rhipicephaline species were genotyped (Table 6, Chapter 3 Section 3.2.1). These included 392 individuals sampled in Kenya from ten field populations (Figure 10 and 11) and five individuals from Natal province, South Africa (total n = 397) and ten laboratory colonies which included five individuals from South Africa (n = 465). The former included ticks collected directly from i) areas grazed exclusively by cattle (five populations, n = 180), ii) areas co-grazed by wildlife and cattle (two populations, n = 80) and iii) areas grazed exclusively by wildlife (three populations, n = 132). The ten laboratory colonies had been bred and maintained as closed genetic stocks at the Tick Unit of the International Livestock Research Institute (ILRI) (Table 6, Chapter 3 Section 3.2.1). They were reared and managed as described in Chapter 3 Section 3.2.2.1 to 3.2.2.3 (Bailey, 1960; Irvin & Brocklesby, 1970). Of the ten laboratory bred populations, one was originally sampled in Uganda (n = 48) and Zimbabwe (West Mashonaland, n = 28) and two from Zambia (eastern province, n = 48; southern province, n = 48) while the rest were obtained in Kenya. One hundred and seventeen individuals from five other rhipicephaline species: *R. pravus* (n = 12), *R. praetextatus* (n = 24), *R.*

zambeziensis (n = 48), *R. evertsi* (n = 15) and *R. pulchellus* (n = 18) sampled in Kenya were included.

Genomic DNA was extracted using the DNeasy® Blood and Tissue Kit (Qiagen GmbH, Germany) with minor modifications (see Chapter 3 Section 3.2.4). Genotyping was done using twenty-nine micro- and mini-satellite markers as described in Chapter 3 Section 3.2.5.

4.2.2 Statistical analysis

4.2.2.1 Genetic diversity analysis

For each population, allelic [total number of alleles (TNA), mean number of alleles (MNA), private alleles (PA)] and genetic diversity [(expected (H_e) and observed (H_o) heterozygosity] from allele frequencies were estimated with Microsatellite toolkit (Park, 2001). For each locus-population combination Fisher's exact test with Bonferroni correction was used to test possible deviations from Hardy-Weinberg equilibrium (HWE) using GENEPOP 3.4 (Raymond & Rousset, 1995). Exact *P*-values were calculated using the Markov chain algorithm following 10 000 dememorizations, 500 batches and 5000 iterations per batch.

4.2.2.2 Population genetic differentiation analysis

When a population comprises of isolated subpopulations, there is less heterozygosity than there would be if it was undivided. The decline in heterozygosity due to subdivision within a population has usually been quantified using an index known as Wright's *F*-statistic, also known as the fixation index. The fixation index ranges from 0 (indicating no differentiation between the overall population and its subpopulations) to a theoretical maximum of 1. Several parameters of population differentiation (*F* statistics) can be used to estimate the degree of population differentiation. In this

study, the total genetic variation of the populations (F_{IT}) was partitioned into within (F_{IS}) and between populations (F_{ST}) components following Weir & Cockerham (1984). Mean estimates of *F*-statistics were obtained using the program FSTAT 2.9.3.2 (Goudet, 2001).

4.2.2.3 Mantel test for isolation-by-distance analysis

To investigate whether the ten field populations of *R. appendiculatus* were at migration-drift equilibrium, the presence of a pattern of isolation by distance (Slatkin, 1993) was determined by regressing geographic distance between populations against pairwise estimates of F_{ST} following 10 000 pseudorandom permutations performed in IBDWS 3.05 (http://ibdws.sdsu.edu). Geographic distances between populations were estimated using the MapCrow Travel Distance Calculator (http://www.mapcrow.info/) and they represented the distances between the central most towns within each sampling location.

4.2.2.4 Population structure and admixture analysis

The underlying population genetic architecture was investigated using Bayesian clustering implemented in STRUCTURE (Pritchard *et al.,* 2000), with the admixture and correlated allele frequency model (Falush *et al.,* 2003). Assuming that the data set could be partitioned into *K* separate genetic clusters, ten runs for each *K* were performed with 200 000 iterations following a burn-in of 100 000 without prior population information for four datasets: i) Global dataset ($1 \leq K \leq 27$); ii) field and laboratory stocks ($1 \leq K \leq 25$); iii) field ticks only ($1 \leq K \leq 10$) and iv) laboratory stocks only ($1 \leq K \leq 20$). The most optimal number of genetic clusters was inferred from the mean estimated log probability of the data (Pritchard *et al.,* 2000) and its second-order rate of change (ΔK) (Evanno *et al.,* 2005) as calculated in STRUCTUREHARVESTER (Earl, 2012). Cluster matching and permutation was done in CLUMPP (Jakobsson & Rosenberg, 2007) while DISTRUCT (Rosenberg, 2004) was used to generate a graphical display of the permuted results.

4.2.2.5 Discriminant analysis of principle components (DAPC) and Bayesian Information Criterion (BIC) analysis

The Bayesian approach implemented in STRUCTURE (Falush *et al.*, 2003; Pritchard, 2000) relies on a genetic model and requires that populations meet Hardy-Weinberg equilibrium expectations and that linkage equilibrium exists between loci. Such assumptions are often violated in natural and inbred populations. Various multivariate analyses that circumvent such assumptions and are not based on any model have been developed (Fisher, 1936; Lachenbruch, 1979). Discriminant analysis of principal components (DAPC) (Jombart *et al.*, 2010) was done to further investigate the genetic structure of the four groups of *R. appendiculatus* analysed with STRUCTURE. DAPC describes global diversity by optimizing the variance between groups while minimizing the variance within groups. In this way the approach seeks synthetic variables or discriminant functions, which reveal differences between groups in the best way possible while minimizing variation within clusters. This is a better approach than either Principal Component Analysis (PCA), Principal Coordinates Analysis (PCoA) or Factorial Correspondence Analysis (FCA), which focus on the total variance and thus describe the global diversity while overlooking differences between groups. To identify the optimal number of clusters that are useful to explain the data, *K*-means was run clustering sequentially with increasing values of *K*, and compared different clustering solutions using the Bayesian Information Criterion (BIC). The lowest BIC value in the regression curve of BIC values as a function of *K* gives the most optimal clustering solution (*K* value) of genetic diversity. DAPC was performed using ADEGENET V1.3-9.2 package (Jombart, 2008) in the R environment v2.15.3 (R Core Team, 2013).

4.2.2.6 Analysis of molecular variance (AMOVA)

To assess the magnitude of genetic variation and hierarchical portioning of variation within and between populations and groups of populations, an analysis of molecular variance (AMOVA) was performed as a weighted average over loci in Arlequin v3.5 (Excoffier & Lischer, 2010) at eight levels, three of which had two hierarchical clusters/groups and one had three hierarchical clusters/groups: (i) the global dataset (All ticks as one group); (ii) *R. appendiculatus* verses other rhipicephaline species (two groups); (iii) all ticks sampled in the field verses laboratory bred stocks (two groups) (iv) Field *R. appendiculatus* only (one group); (v) all laboratory stocks of *R. appendiculatus* (one group); (vi) field ticks analysed as cattle stocks, cattle-wildlife stocks and wildlife stocks (three groups); (vii) other species of rhipicephaline (one group). Levels of significance for the covariance components associated with the different hierarchical clusters/groups were determined with 1000 non-parametric permutations.

4.2.2.7 Demographic stability and expansion tests (Kurtosis (*k*-test) and variance test (*g*-test))

To infer population demographic history and dynamics, the tick populations were assessed if they were at mutation-drift equilibrium by performing the intra-locus kurtosis test (*k*-test) and the inter-locus variance test (*g*-test) (Reich & Goldstein 1998; Reich *et al*., 1999). The *k*-test is based on the premise that allele distribution patterns in expanding populations differ from those that are demographically stable. In expanding populations, the kurtosis (*k*), or the combination of the variance and kurtosis (Reich *et al.*, 1999) of allele size distributions is negative. The method uses a binomial test of the number of positive *k*-values based on the expectation of an almost equal probability ($P = 0.515$) of negative and positive *k*-values. The *g*-test on the other hand, compares the observed and estimated values of the variance in allele sizes across loci. In stable populations the variance is highly variable among loci, whereas in expanding populations it is much more even. For

this test, low variances in allele sizes were taken as evidence of expansion. The the cut-off values given in Table 1 (page 455) of Reich *et al*. (1999) were used for inference purposes. Both *k*- and *g*-tests were performed using the macro program "kgtests" (Bilgin, 2007) implemented in Microsoft Excel®.

4.3.0 Results

4.3.1 Genetic variability among and within tick populations

The overall genetic diversity indices were calculated by pooling all the genotypic data from both field and laboratory ticks. A total of 291 alleles were observed at the twenty-nine loci genotyped in 979 individuals from twenty-seven populations giving an average of 10.03 ± 4.24 alleles per population (Table 10). The overall average expected and observed heterozygosity were 0.633 ± 0.032 and 0.442 ± 0.003 respectively (Table 10). Laboratory bred stocks had the lowest levels of genetic and allelic diversity. Although the other five species of rhipicephaline had a higher level of expected heterozygosity, populations of *R. appendiculatus* that were sampled from the field had a higher level of observed heterozygosity and allelic diversity (MNA and TNA). The total number of private alleles ranged from zero in KF, SN, MU, ZE, ZM, SL and RZ to eleven in RL. The highest number of private alleles was observed in field populations of *R. appendiculatus* (37) and in the five other species of rhipicephaline (34). Except for KU that had a negative F_{IS} value indicating an excess of heterozygotes and MF that had a zero F_{IS} value indicating an equal number of homo- and heterozygotes, the rest had positive F_{IS} values indicating deficiencies in heterozygotes. KU is the only population where the observed heterozygosity exceeded the expected while MF is the only population that had the same values of observed and expected heterozygosity.

4.3.2 Genetic differentiation and relationship between tick populations

PCA was done to visualize pairwise differentiation (F_{ST}) between the 25 Kenyan tick populations and between samples (Supplementary Figure 2 and 3, Appendix I) using the genetic distance. Most field *R. appendiculatus* tick populations cluster within a central group, with field ticks (NB, FP, MA, MK, BU) clustering closely with KH laboratory stock. *R. zambeziensis* (RZ) clusters with most *R. appendiculatus* populations. However RE, RP, RV, RL which are populations within the other rhipicephaline species are separated from *R. appendiculatus*. The PCA also discriminates laboratory populations ZE, ML, LP. The separation of individual ticks is shown in Supplementary Figure 3 (Appendix I). Most *R. appendiculatus* ticks cluster together. Individuals of the standard tick stock (MU) are dispersed throughout the PCA. Individual ticks from populations BF, MF, RP, KF, ZM, and FP appear clustered in distict groups.

Pairwise F_{ST} estimates ranged between 0.0003 (between FP and NB populations) and 0.507 (between SL and RL populations) (Table 10 and Supplementary Figure 4 (Appendix I)). Laboratory bred stocks returned the highest values of F_{ST} between themselves and also between themselves and the field populations. Field populations on the other hand returned the lowest F_{ST} values between themselves with the most closely genetically related populations being NB and FP (Pairwise F_{ST} =-0.0003). Mantel test revealed a negative but non-significant correlation between geographic (km) and genetic distances measured as pairwise F_{ST} (r = -0.1636; P = 0.7396) (Figure 17).

Table 10: Genetic and allelic diversity for each of the 27 tick populations analysed in the study

			Genetic diversity		Allelic diversity				
Population	**Sample size**	**Loci typed**	**He (SD)**	**Ho (SD)**	**MNA (SD)**	**TNA**	**F_{IS}**	**Private alleles**	**HWE (P<0.01)**
Field ticks									
KF	25	29	0.588 (0.040)	0.514 (0.021)	4.38 (1.88)	127	0.137	-	-
MK	46	29	0.613 (0.037)	0.501 (0.014)	5.93 (2.88)	172	0.183	1	7
KT	25	29	0.642 (0.029)	0.523 (0.020)	5.59 (2.85)	162	0.189	5	3
RU	36	29	0.606 (0.035)	0.518 (0.017)	5.31 (2.02)	154	0.148	1	4
BU	48	29	0.616 (0.034)	0.526 (0.014)	5.83 (2.11)	169	0.148	2	3
BO	34	29	0.624 (0.035)	0.488 (0.017)	5.10 (2.40)	148	0.221	1	4
FP	46	29	0.629 (0.037)	0.504 (0.014)	5.76 (2.37)	167	0.200	3	6
NB	36	29	0.608 (0.040)	0.506 (0.016)	5.90 (2.47)	171	0.170	3	2
MA	48	29	0.614 (0.037)	0.473 (0.014)	5.86 (2.37)	170	0.230	1	8
BF	48	26	0.541 (0.046)	0.454 (0.016)	5.38 (2.42)	140	0.168	1	2
SN	5	29	0.527 (0.054)	0.316 (0.040)	2.86 (1.30)	83	0.446	-	-
Overall	**397**	**29**	**0.623 (0.035)**	**0.505 (0.005)**	**8.28 (3.88)**	**240**	**0.189**	**37**	**14**
Laboratory stocks									
LP	48	29	0.441 (0.046)	0.427 (0.014)	3.48 (1.72)	101	0.030	4	6
KH	48	29	0.472 (0.043)	0.388 (0.014)	3.69 (1.75)	107	0.186	2	6
KU	48	29	0.449 (0.042)	0.470 (0.014)	3.07 (1.33)	89	-0.044	1	7
MU	48	29	0.579 (0.032)	0.475 (0.014)	4.62 (1.95)	134	0.181	-	7
MF	48	29	0.566 (0.030)	0.430 (0.013)	4.38 (2.03)	127	0.243	1	8
ML	48	29	0.384 (0.048)	0.384 (0.013)	2.62 (1.15)	76	0.000	2	4
UG	48	29	0.442 (0.037)	0.387 (0.014)	2.97 (1.15)	86	0.126	1	4
ZS	48	29	0.364 (0.043)	0.334 (0.012)	2.83 (1.17)	82	0.082	1	5
ZE	48	29	0.409 (0.041)	0.391 (0.014)	3.10 (1.32)	90	0.042	-	2
ZM	28	29	0.491 (0.043)	0.474 (0.018)	3.45 (1.45)	100	0.035	-	2
SL	5	29	0.344 (0.051)	0.291 (0.040)	1.79 (0.56)	52	0.205	-	-
Overall	**465**	**29**	**0.592 (0.034)**	**0.411 (0.004)**	**6.59 (3.03)**	**191**	**0.305**	**14**	**26**
Other Rhipicephaline									

RE	15	28	0.474 (0.057)	0.391 (0.031)	3.11 (1.93)	87	0.206	3	-
RP	24	26	0.515 (0.063)	0.261 (0.030)	3.12 (1.61)	81	0.611	10	3
RV	12	28	0.587 (0.042)	0.370 (0.038)	3.36 (1.70)	94	0.410	2	-
RL	18	27	0.454 (0.045)	0.310 (0.024)	3.22 (1.31)	87	0.339	11	2
RZ	48	29	0.490 (0.040)	0.350 (0.013)	3.31 (1.17)	96	0.288	-	6
Overall	**117**	**29**	**0.680 (0.022)**	**0.343 (0.010)**	**7.10 (2.93)**	**206**	**0.497**	**34**	**28**
Total	**979**	**29**	**0.633 (0.032)**	**0.442 (0.003)**	**10.03 (4.24)**	**291**	**0.301**		

KF = Kilifi District, MK = Makuyu, KT = Kitale, RU = Rusinga Island, BU = Busia District, BO = Bomet, FP = Ol Pejeta Conservancy, Nanyuki, NB = Nairobi National Park, MA = Maasai Mara National Park, BF = Buffaloes in Maasai Mara National Park , SN = Natal Province, South Africa. LP = Laboratory Ol Pejeta, KH = Kiambu Highline (shows high infectivity), KU = Kiambu unselected line, MU = Muguga unselected line, MF = Muguga Infected line, ML = Muguga Low line (shows low infectivity), UG = Uganda stock, ZS = Zambia (Southern province), ZE = Zambia (Eastern province), ZM = Zimbabwe (West Mashonaland), SL = South Africa Laboratory stock. RE = *Rhipicephalus evertsi*, RP = *Rhipicephalus praetextatus*, RV = *Rhipicephalus pravus*, RL = *Rhipicephalus pulchellus*, RZ = *Rhipicephalus zambeziensis*.

Table 11: Population differentiation based on pairwise F_{ST} values between each of the 27 populations of ticks

	KF	**MK**	**BO**	**KT**	**BU**	**RU**	**FP**	LP	KH	KU	MF	ML	MU	UG	ZS	ZE	ZM	SL	SN	**BF**	**MA**	**NB**	RE	RP	RV	RL	RZ
KF	0.00	0.019	0.064	0.018	0.021	0.009	0.024	0.244	0.142	0.197	0.105	0.205	0.079	0.220	0.248	0.246	0.139	0.263	0.095	0.085	0.024	0.028	0.270	0.271	0.099	0.317	0.201
MK		0.00	0.024	0.009	-0.001	0.006	0.0003	0.183	0.151	0.189	0.075	0.159	0.049	0.162	0.247	0.191	0.121	0.234	0.074	0.071	0.001	0.001	0.241	0.252	0.089	0.302	0.139
BO			0.00	0.029	0.0278	0.037	0.026	0.1795	0.164	0.160	0.041	0.219	0.018	0.169	0.246	0.191	0.122	0.250	0.097	0.108	0.019	0.029	0.238	0.241	0.112	0.306	0.137
KT				0.00	0.0103	0.018	0.011	0.204	0.141	0.212	0.085	0.196	0.056	0.163	0.252	0.205	0.137	0.231	0.071	0.060	0.012	0.016	0.206	0.232	0.051	0.283	0.144
BU					0.00	0.012	0.005	0.174	0.122	0.195	0.078	0.159	0.049	0.159	0.251	0.185	0.127	0.235	0.085	0.072	0.001	0.004	0.242	0.243	0.079	0.303	0.145
RU						0.000	0.014	0.209	0.145	0.201	0.079	0.188	0.056	0.171	0.256	0.208	0.146	0.263	0.101	0.072	0.014	0.011	0.257	0.252	0.106	0.309	0.147
FP							0.000	0.17	0.150	0.176	0.072	0.165	0.049	0.156	0.240	0.172	0.114	0.215	0.078	0.073	0.003	-0.0003	0.223	0.243	0.077	0.299	0.141
LP								0.00	0.293	0.283	0.179	0.386	0.183	0.280	0.396	0.057	0.223	0.393	0.174	0.298	0.191	0.177	0.341	0.336	0.269	0.448	0.308
KH									0.000	0.302	0.185	0.267	0.188	0.201	0.333	0.294	0.271	0.386	0.227	0.209	0.146	0.135	0.376	0.369	0.217	0.397	0.265
KU										0.000	0.136	0.334	0.195	0.329	0.232	0.300	0.219	0.369	0.259	0.232	0.188	0.197	0.357	0.384	0.290	0.454	0.363
MF											0.000	0.265	0.057	0.181	0.149	0.170	0.145	0.294	0.134	0.132	0.070	0.079	0.279	0.278	0.167	0.356	0.187
ML												0.000	0.248	0.318	0.393	0.392	0.305	0.402	0.341	0.231	0.169	0.163	0.449	0.449	0.262	0.479	0.348
MU													0.000	0.200	0.276	0.186	0.135	0.274	0.112	0.126	0.034	0.054	0.280	0.274	0.164	0.351	0.155
UG														0.000	0.342	0.249	0.267	0.434	0.224	0.249	0.159	0.168	0.368	0.379	0.280	0.399	0.216
ZS															0.000	0.373	0.286	0.478	0.388	0.262	0.238	0.253	0.429	0.471	0.372	0.480	0.404
ZE																0.000	0.234	0.426	0.195	0.299	0.191	0.182	0.369	0.369	0.296	0.462	0.299
ZM																	0.000	0.313	0.129	0.194	0.117	0.125	0.311	0.333	0.241	0.412	0.279
SL																		0.000	0.306	0.300	0.228	0.240	0.445	0.374	0.337	0.507	0.414
SN																			0.000	0.203	0.089	0.089	0.257	0.253	0.129	0.373	0.231
BF																				0.000	0.074	0.080	0.308	0.230	0.119	0.349	0.184
MA																					0.000	0.002	0.233	0.246	0.094	0.301	0.137
NB																						0.000	0.245	0.249	0.080	0.307	0.143
RE																							0.000	0.379	0.271	0.368	0.380
RP																								0.000	0.260	0.334	0.408
RV																									0.000	0.313	0.262
RL																										0.000	0.414
RZ																											0.000

Pairwise F_{ST} estimates ranged between 0.0003 (between FP and NB populations) and 0.507 (between SL and RL populations). Laboratory bred stocks returned the highest values of F_{ST} between themselves and also between themselves and the field populations. Field populations on the other hand returned the lowest F_{ST} values between themselves with the most closely genetically related populations being NB and FP (Pairwise F_{ST} =-0.0003).

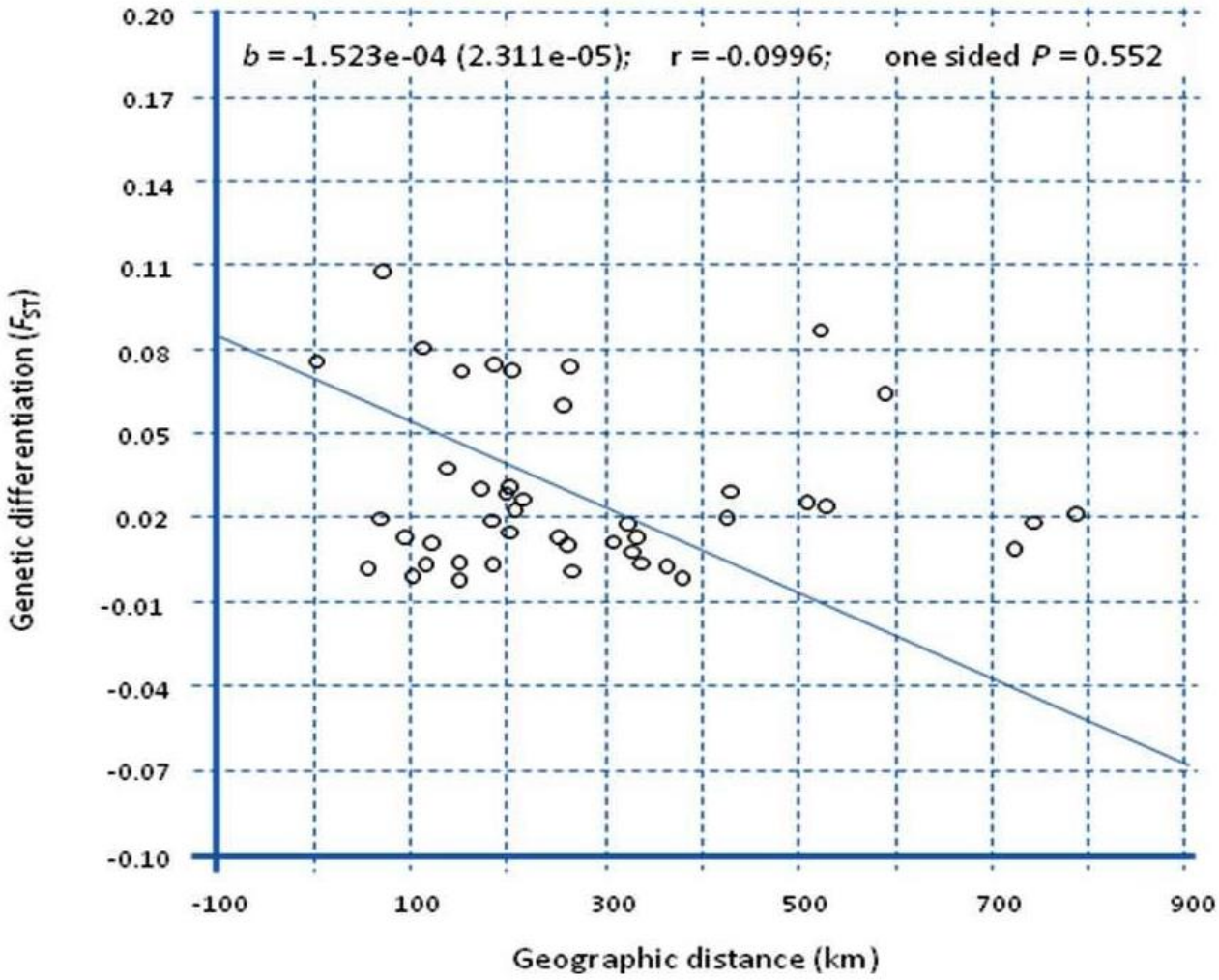


Figure 17: Regression of pairwise Fst against geographic distance for field ticks only
The graph shows correlation between geographic (km) and genetic distances in the Kenyan *R. appendiculatus* tick population. A negative but non-significant correlation between geographic (km) and genetic distances measured as pairwise F_{ST} (r = -0.0996; P = 0.552) was observed.

4.3.3 Genetic variance within and between tick populations

The level of genetic variation present between different groups of ticks was assessed using AMOVA (Table 12). For all the hierarchical clusters tested, the highest level of genetic variation was within individuals and ranged from 45.07% (in the five other species of rhipicephaline to 80.56% in the ten field stocks of *R. appendiculatus*). The lowest level of genetic variation (0.125%) was observed among the populations of *R. appendiculatus* grouped on the basis of their host (cattle, cattle-wildlife, and wildlife). However when these were analyzed as a single group, the level of genetic variation between populations rose to 1.77%. This value was lower than that between laboratory and field ticks (2.33%). The level of genetic variation between populations of laboratory bred *R. appendiculatus* was the second highest at 25.139% while that between the five other species of rhipicephaline was the highest at 38.62%. The variation among tick populations of *R. appendiculatus* and the five other species of rhipicephaline was 6.05% while that present among all the 27 tick populations was 18.63%.

4.3.4. Demographic expansion of tick populations

The "kg-tests" was performed to evaluate whether the study populations had experienced any changes in their effective population sizes (Table 13). When the twenty-seven populations were analyzed as a single group, the *k*-test showed twenty-six loci with a significant negative kurtosis. Apart from the ten laboratory bred populations of *R. appendiculatus* and the five other species of rhipicephaline which had seventeen and eighteen loci with non-significant negative kurtosis, the other groups of ticks had more than twenty loci with significant negative values of kurtosis (Table 13). The twenty populations of *R. appendiculatus* as a single group had twenty-two loci with a significant negative kurtosis ($P \leq 0.05$) supporting population expansion. However, when the analysis was performed separately for field and laboratory stocks, the former had twenty-seven loci

with significant negative values of kurtosis while the latter had 17 loci with no significant kurtosis values.

Table 12: Global AMOVA results as a weighted average over 29 loci in 27 tick populations

Source of variation	**Hierarchy**	**Variance components**	**Percent of variation**
1. All ticks	1 group	Within individuals	69.628
		Among individuals within populations	11.744
		Among populations	18.628
2. *R. appendiculatus* vs Other rhipicephaline species	2 groups	Within individuals	66.149
		Among individuals within populations	11.157
		Among populations within groups	16.637
		Among groups	6.057
3. Field vs Lab stocks	2 groups	Within individuals	71.494
		Among individuals within populations	11.831
		Among populations within groups	14.344
		Among groups	2.330
4. Field ticks only	1 group	Within individuals	80.146
		Among individuals within populations	18.076
		Among populations	1.778
5. Laboratory stocks only	1 group	Within individuals	67.755
		Among individuals within populations	7.106
		Among populations	25.139
6. Field ticks (cattle, cattle-wildlife, wildlife)	3 groups	Within individuals	80.561
		Among individuals within populations	17.846
		Among populations within groups	1.467
		Among groups	0.125
7. Other *Rhipicephalinae*	1 group	Within individuals	45.07937
		Among individuals within populations	16.29935
		Among populations	38.62129

Table 13: Demographic expansion parameters for different stocks of ticks as determined from the "kg-tests"

Type of tick stock	Number of loci with negative K value	k-test's p-value	g-test value
All tick populations	26	3.75E-06	6.256344
All *Rhipicephalus appendiculatus* (Field and laboratory stocks)	22	0.002504	5.255007
Other rhipicephaline species	18	0.100746	9.168623
Field tick stocks only	27	3.77E-07	6.2668
Laboratory stocks only	17	0.182863	4.729614
Wildlife ticks	27	3.77E-07	7.083718
Cattle ticks	28	2.45E-08	5.83376

4.3.5. Population structuring and divergence

STRUCTURE and DAPC reveal the partitioning of genetic diversity at the individual and population level. The Δ*K* (Evanno *et al.,* 2005) approach, the most optimal number of clusters revealed was *K* = 20 for the overall dataset of twenty-seven populations, *K*= 2 for field ticks only, *K* = 12 for the combined field and laboratory ticks and *K* = 3 for laboratory stocks only (Figures 18 and 19). At *K* = 20, all the ten field populations of *R. appendiculatus* (cattle, cattle-wildlife, wildlife) show no clear clustering pattern (Figure 20). The figure showed that the laboratory bred populations (KU, MU and MF) had high genetic admixture while seven (LP, KH, ML, UG, ZS, ZE, ZM) were genetically distinct (Figure 20 and Supplementary Figure 5). For the five other rhipicephaline species, four (RE, RP, RV, RL) were genetically admixed while one (RZ) was genetically distinct. Of the two *R. appendiculatus* populations sampled from South Africa, one (SN) was distinct and the other (SL) was admixed (Figure 20).

The regression of BIC values against *K* following DAPC reveal 23 clusters in the overall dataset (Figure 21a), six clusters in field *R. appendiculatus* ticks (Figure 21b), 18 clusters in the combined group of field and laboratory *R. appendiculatus* ticks (Figure 22a) and 13 clusters in laboratory *R. appendiculatus* ticks (Figure 22b) to be the most optimal to explain the variation observed. As with STRUCTURE, the ten field populations of *R. appendiculatus* are assigned to multiple clusters as are two populations of laboratory bred stocks (MU and MF). LP and ZE are assigned to three clusters, KU and ZM to two clusters one of which they share with very few individuals of MF and ZS (Cluster 21). KH, ML and UG are assigned to individual clusters (Figure 23). For the five other rhipicephaline species, RE and RZ are assigned to their own clusters (Cluster 5 and 6 respectively) while majority of the individuals of RP are assigned to cluster 19 which they share with less than 15 individuals of RV and RL (Figure 23). Worth noting is that STRUCTURE and DAPC showed that most laboratory bred stocks were assigned to distinct clusters from those of the field stocks. For instance, individuals of LP, a laboratory bred population originally sampled in Ol Pajeta wildlife ranch, did not share any cluster with those of FP, a field population sampled recently from the same ranch (Figure 20 and 23).

To visualize in greater detail the divergence of the field stocks from the laboratory bred stocks, STRUCTURE and DAPC analysis was performed for these two groups of ticks. The ΔK approach revealed the optimal number of clusters to be 12 (Figure 19a) while BIC gave a value of 18 (Figure Figure 22a). For STRUCTURE, apart from MU and MF, which were assigned to the same cluster with some degree of admixture, and KU, which was assigned to two clusters, the other populations were predominantly assigned to individual clusters. The ten field populations of *R. appendiculatus* were highly admixed with no clear observable genetic clusters (Figure 24). The clustering pattern revealed by DAPC (Figure 25) confirmed the admixed nature of the ten field populations of *R.*

appendiculatus; most of their individuals were assigned to multiple clusters. MU and MF were the only laboratory bred populations that were assigned to multiple clusters similar to those of field populations, while the rest were assigned to individual clusters. The other exceptions were LP and ZE, which were assigned to two clusters (1 and 13) at almost equal proportions and KU, which was also assigned to two clusters (2 and 5). A large number of ZS individuals together with about 15 of MF were assigned to one cluster (7) (Figure 25).

STRUCTURE and DAPC analyses were further performed for field and laboratory populations of *R. appendiculatus* separately to portray in finer detail the genetic structure in these two groups of ticks. The ΔK approach returned two (Figure 18b) while BIC returned six (Figure 21b) as the most optimal numbers of clusters for field populations. At $K = 2$, all field populations shared two genetic backgrounds disproportionately; there was no distinction between ticks sampled from areas grazed by cattle, cattle-wildlife and wildlife respectively (Figure 26). All field populations were assigned to the six clusters revealed by BIC disproportionately (Figure 27). In spite the high K value, there was no clarity of assignment for individuals sampled from areas grazed by cattle, cattle-wildlife and wildlife respectively. For laboratory bred stocks, the ΔK (Figure 22b) and BIC (Figure 21b) approaches returned 3 and 13 as the most optimal number of clusters in the dataset. At $K = 3$, STRUCTURE revealed LP, KU, ML, ZS, and ZE to be the least admixed populations and MU, MF and ZM to be the most admixed (Figure 28). KH and UG had intermediate levels of admixture. DAPC showed KH, ML and UG to be the most distinct. Although ZS was distinct it shared its cluster (11) with some individuals of MU and MF. LP and ZE were assigned to three clusters (2, 3 and 8) while KU and ZM were assigned to two clusters, 1 and 10, and 10 and 13 respectively (Figure 29).

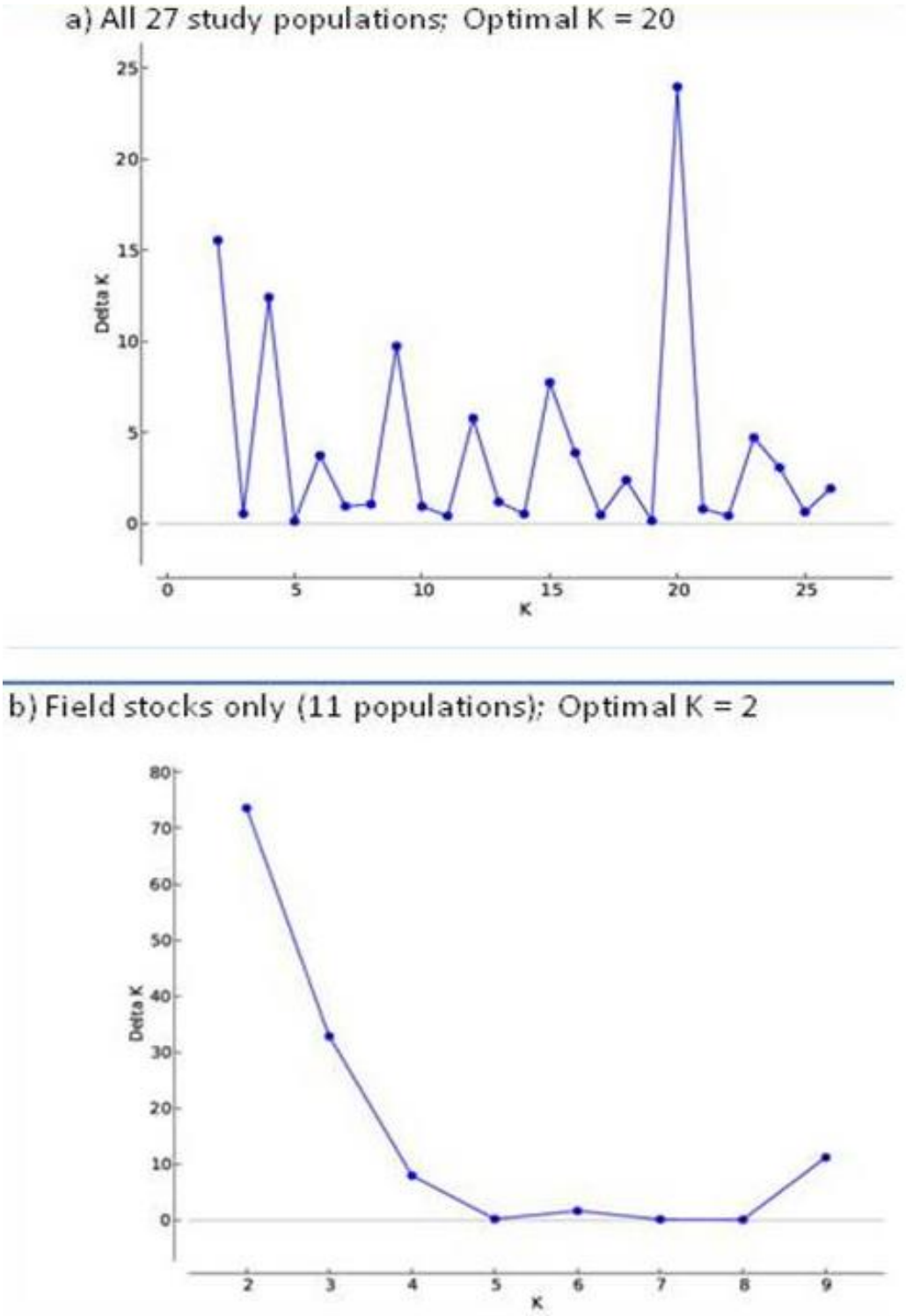
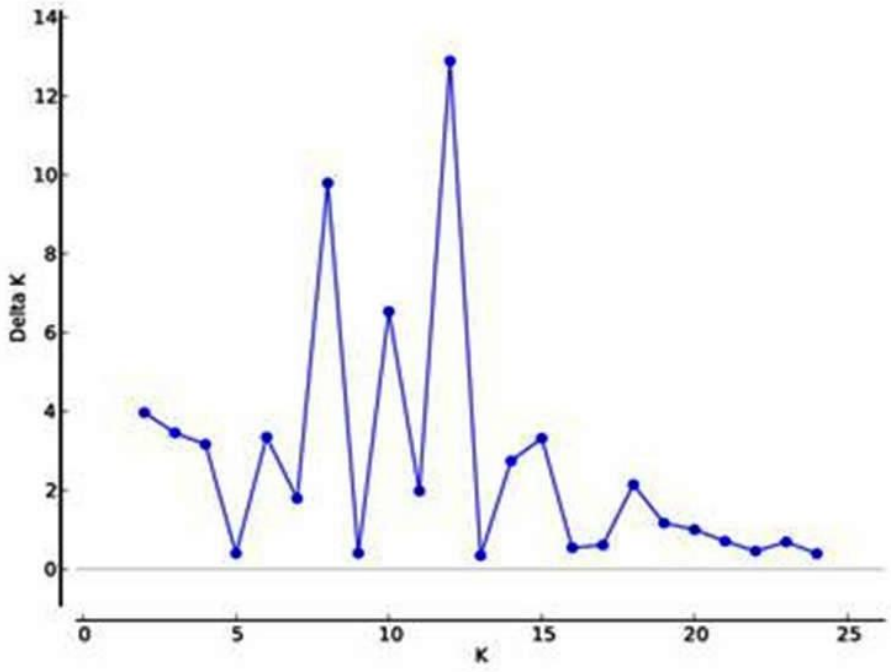


Figure 18a and b: Delta K (Evanno *et al*, 2005) plotted against assumed genetic clusters (K) for all the 27 tick populations and 11 field populations respectively

The most optimal number of clusters revealed was $K = 20$ for the overall dataset of twenty-seven populations and $K = 2$ for field ticks.

a) Field and Laboratory stocks (22 populations); Optimal K = 12



b) Laboratory stocks only (10 populations); Optimal K = 3

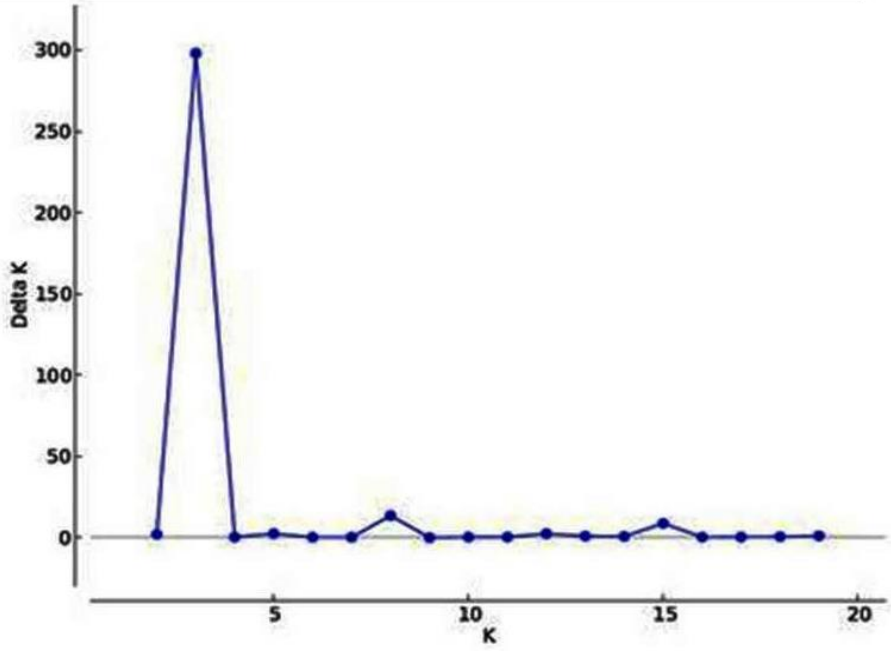


Figure 19a and b: Delta K (Evanno *et al*, 2005) plotted against assumed genetic clusters (K) for the combined 22 populations of field and laboratory ticks and 10 laboratory populations respectively

The most optimal number of clusters revealed was $K = 12$ for the combined field and laboratory ticks and $K = 3$ for laboratory stocks only.

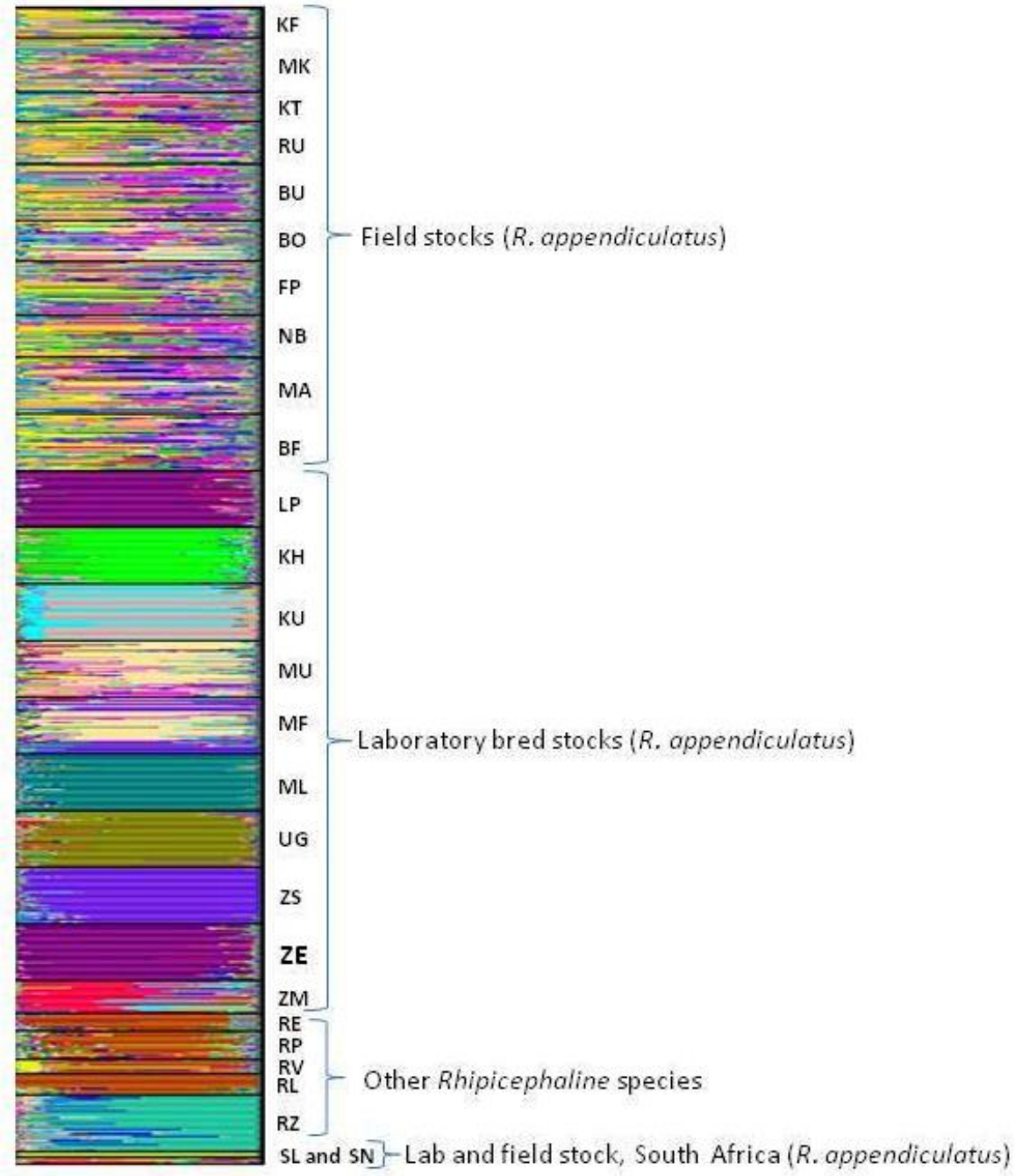
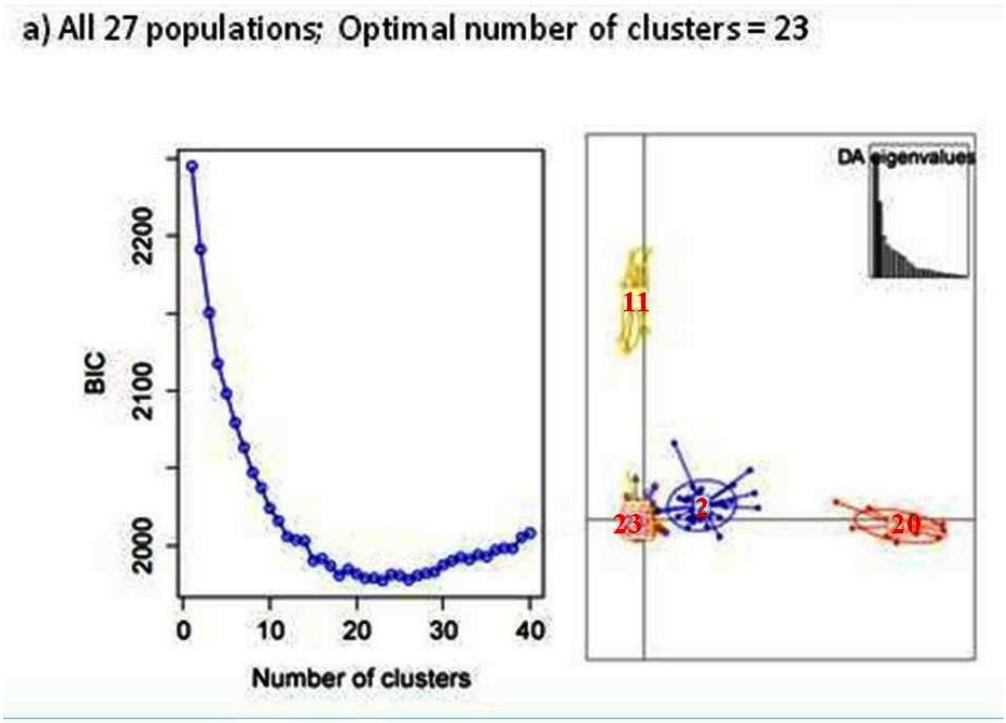


Figure 20: STRUCTURE bar plots at *K*=20 showing the population structure of the 27 tick populations

The ancestry of each population is represented by different colors. Laboratory bred populations (KU, MU and MF) had high genetic admixture while seven populations (LP, KH, ML, UG, ZS, ZE, ZM) were genetically distinct. For the five other rhipicephaline species, four (RE, RP, RV, RL) were genetically admixed while one (RZ) was genetically distinct.



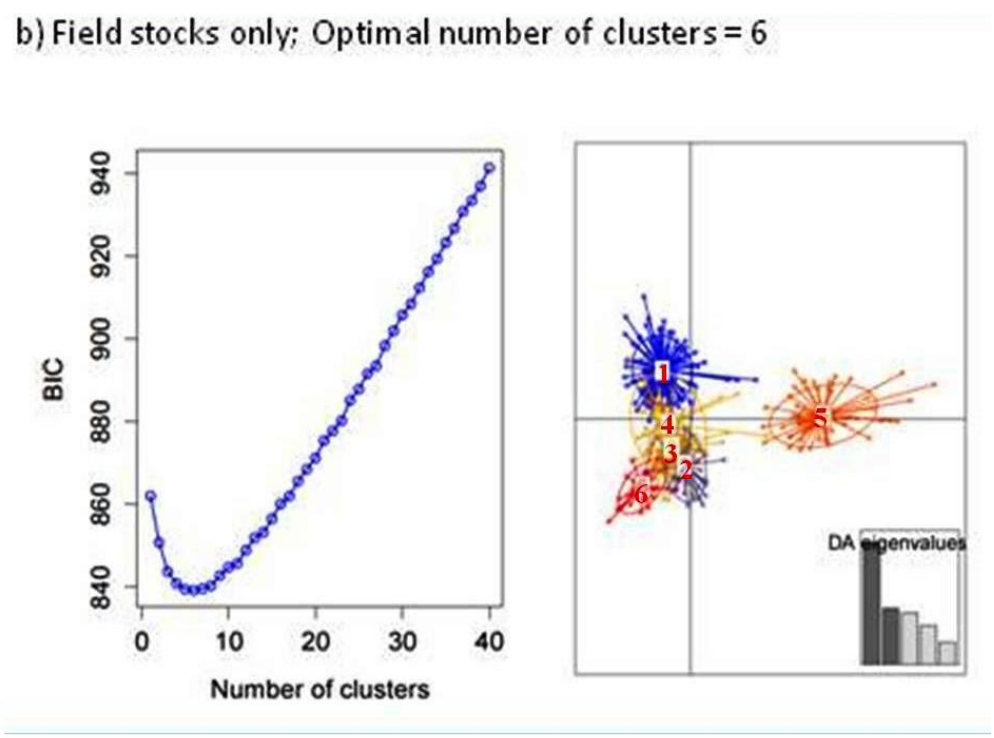
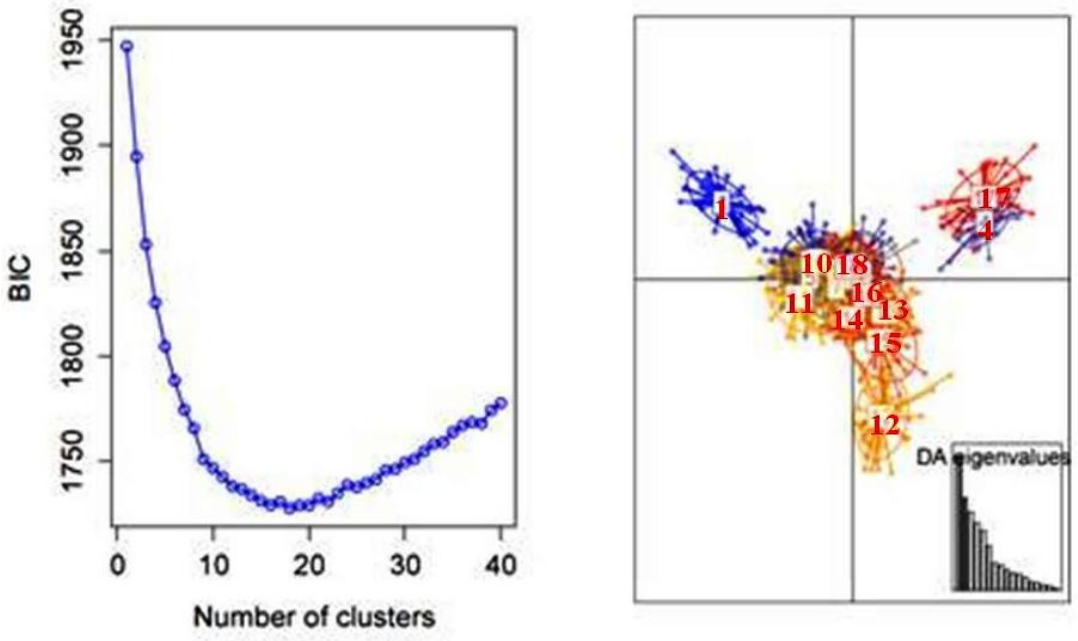


Figure 21a and b: DAPC and BIC plots for all the 27 tick populations and 11 field tick populations respectively

The regression of BIC values against *K* following DAPC revealed 23 clusters in the overall dataset (Figure 19a) and six clusters in field *R. appendiculatus* ticks (Figure 19b) to be the optimal number of clusters. These are represented by the number of clusters with the lowest BIC values.

a) Field and laboratory stocks; Optimal number of clusters = 18



b) Laboratory stocks only; Optimal number of clusters = 13

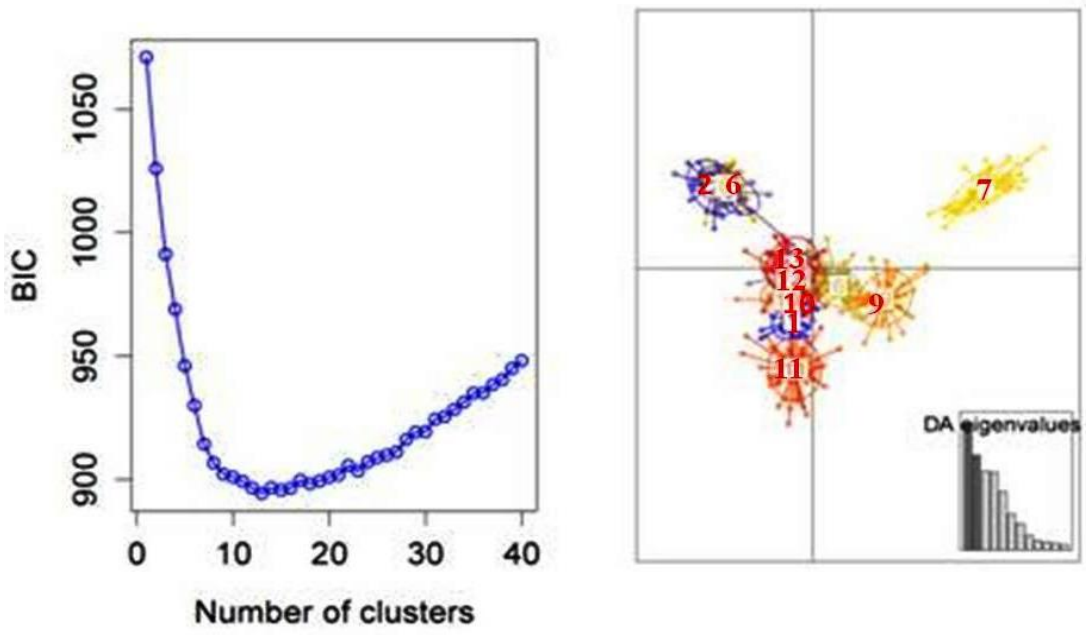


Figure 22a and b: DAPC and BIC plots for 22 combined field and laboratory tick populations and 10 laboratory tick populations respectively

The regression of BIC values against *K* following DAPC shows 18 clusters in the combined group of field and laboratory *R. appendiculatus* ticks (Figure 20a) and 13 clusters (Figure 20b) in laboratory *R. appendiculatus* ticks to be the most optimal number of clusters.

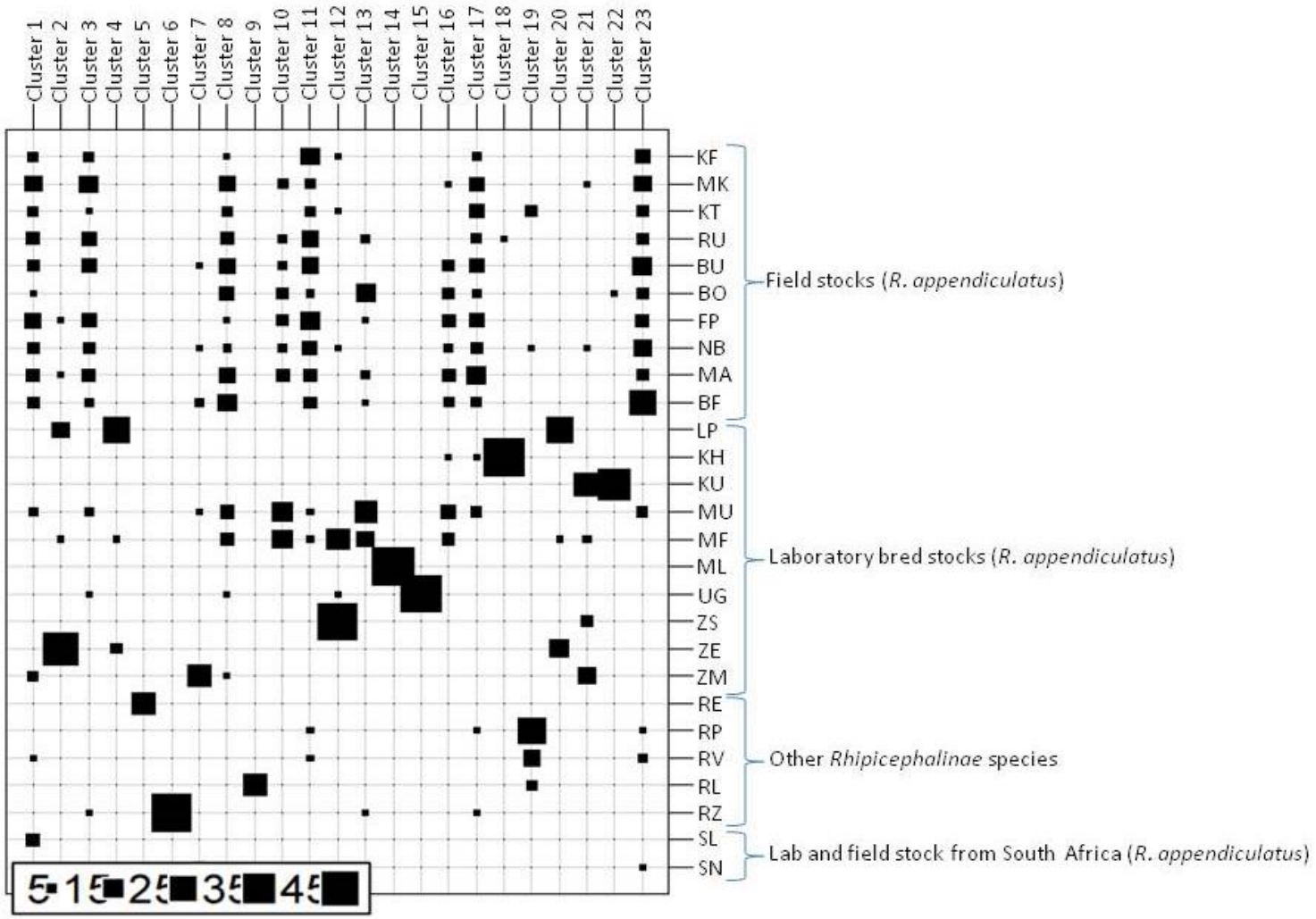


Figure 23: DAPC assignment of individuals from 27 field and laboratory populations of *R. appendiculatus*
The optimal clusters were found to be 23. The ten field populations of *R. appendiculatus* are assigned to multiple clusters as are two populations of laboratory bred stocks (MU and MF). LP and ZE are assigned to three clusters, KU and ZM to two clusters one of which they share with very few individuals of MF and ZS (Cluster 21). KH, ML and UG are assigned to individual clusters.

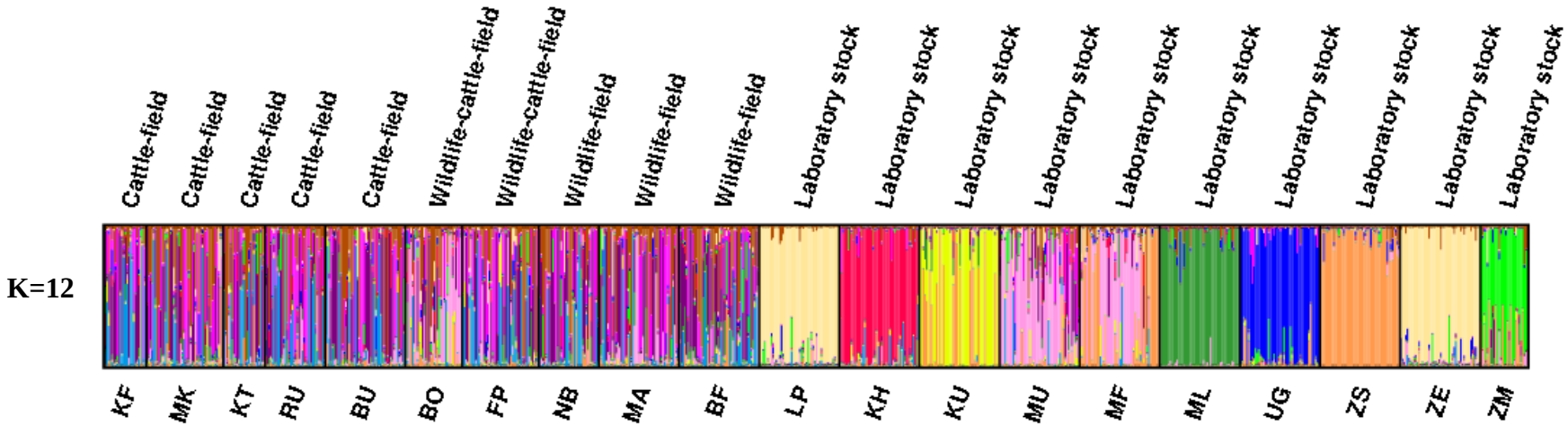


Figure 24: STRUCTURE bar plots at *K*=12 showing the population structure of the 20 combined field and laboratory populations The ancestry of each population is represented by different colors. The ten field populations of *R. appendiculatus* were highly admixed with no clear observable genetic clusters.

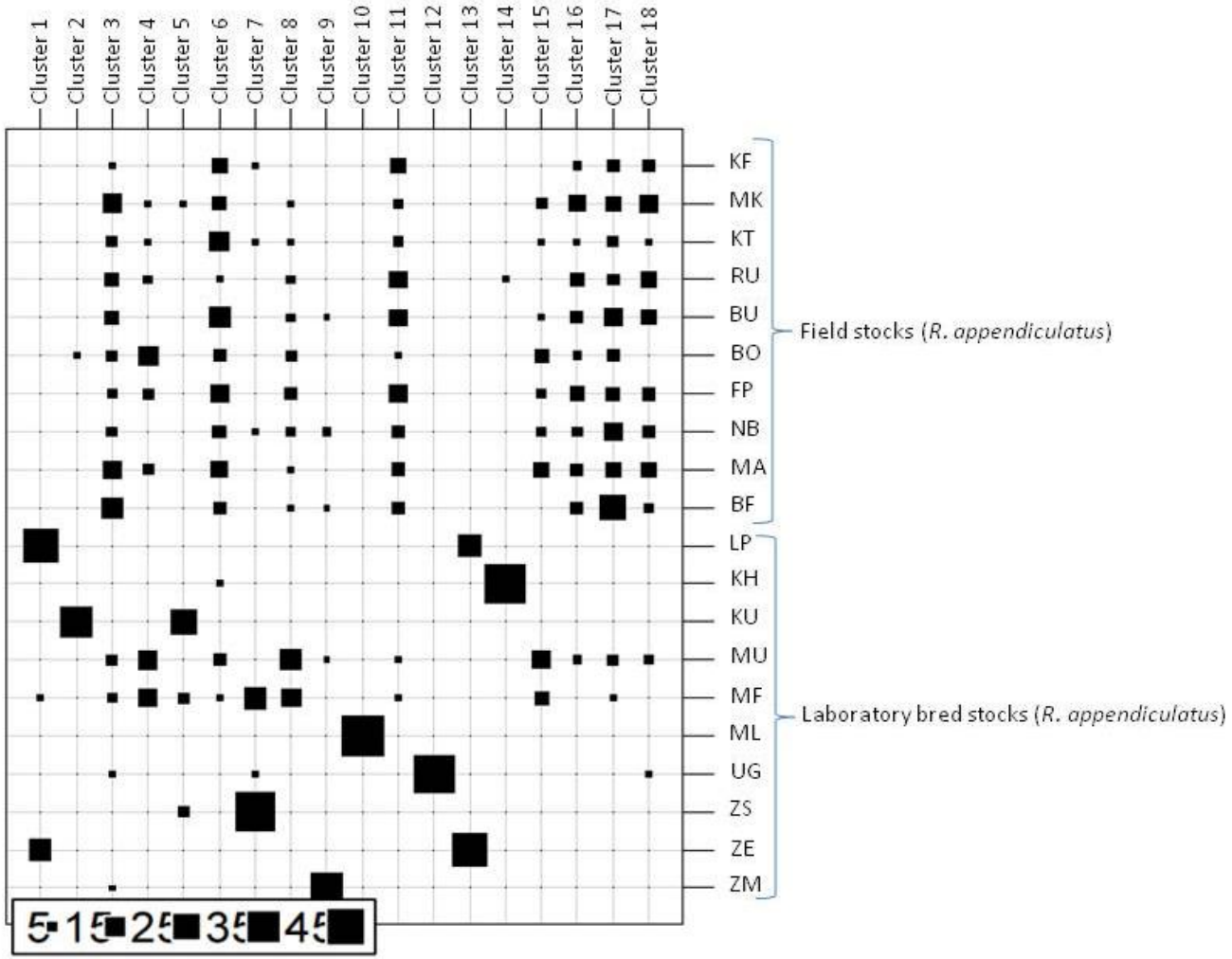


Figure 25: DAPC assignment of individuals from 20 field and laboratory populations of *R. appendiculatus*
The optimal clusters were 18. The ten field populations of *R. appendiculatus* belonged to multiple clusters. Laboratory stocks MU and MF were assigned to the same cluster with some degree of admixture, and KU, which was assigned to two clusters, the other populations were predominantly assigned to individual clusters.

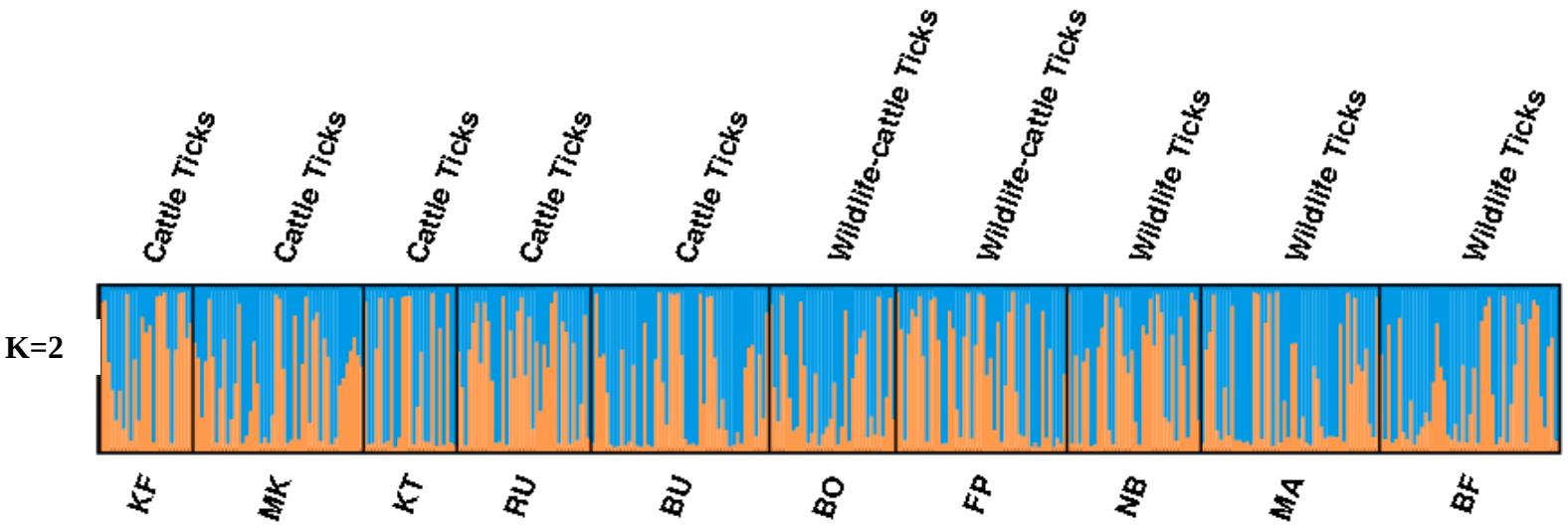


Figure 26: STRUCTURE bar plots at *K*=2 showing the population structure of the 10 field tick populations
The ancestry of each population is represented by different colors. The field populations shared two genetic backgrounds disproportionately; there was no distinction between ticks sampled from areas grazed by cattle, cattle-wildlife and wildlife respectively.

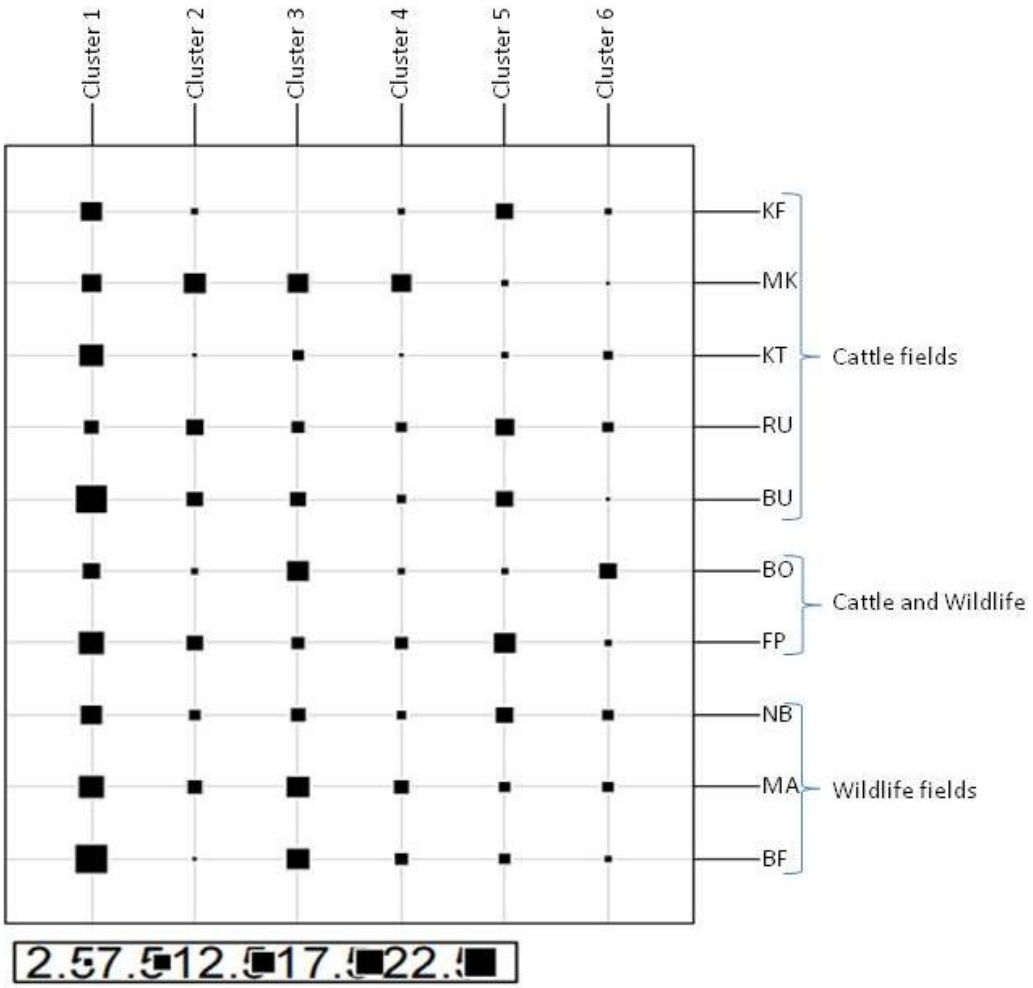


Figure 27: DAPC assignment of individuals from 10 field populations of *R. appendiculatus*
All field populations were assigned to the six clusters shareddisproportionately among all the populations.

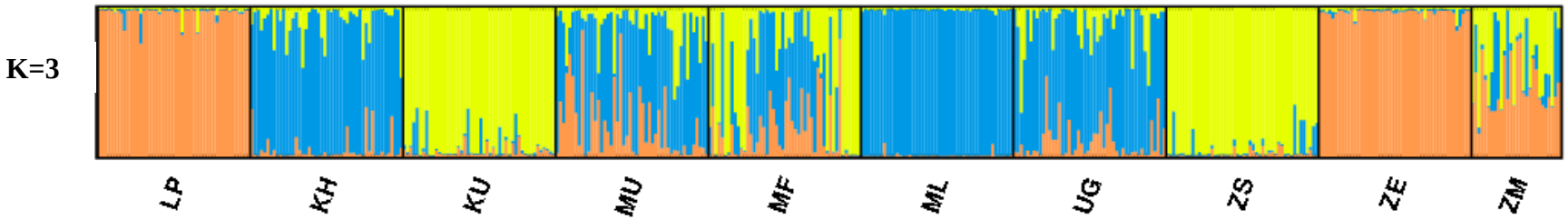


Figure 28: STRUCTURE bar plots at *K*=3 showing the population structure of the 10 laboratory tick populations
The ancestry of each population is represented by different colours. LP, KU, ML, ZS, and ZE appear to be the least admixed populations and MU, MF and ZM to be the most admixed. KH and UG had intermediate levels of admixture.

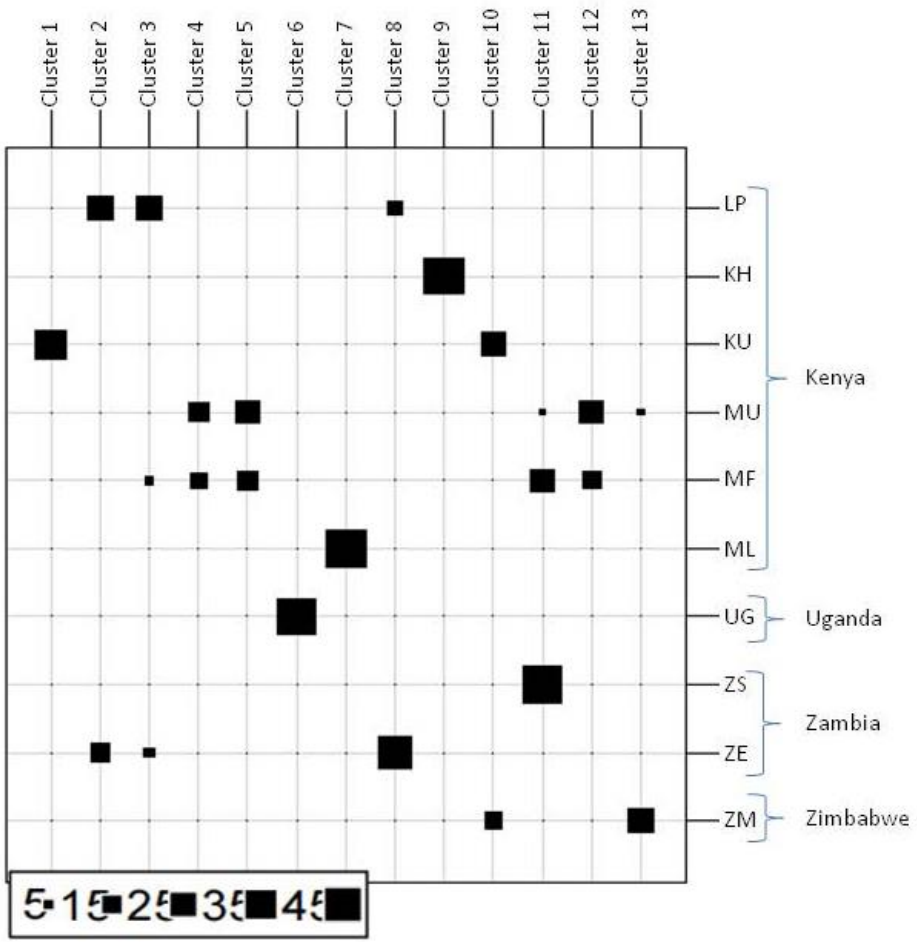


Figure 29: DAPC assignment of individuals from 10 populations of laboratory bred *R. appendiculatus*
The optimal number of clusters was 13. KH, ML and UG appeared to be the most distinct. ZS shared its cluster (11) with some individuals of MU and MF. LP and ZE were assigned to three clusters (2, 3 and 8) while KU and ZM were assigned to two clusters, 1 and 10, and 10 and 13 respectively.

4.4 Discussion

To develop effective methods against tick-borne diseases via integrated vector control and/or anti-tick vaccination strategies (Dai *et al.,* 2009; Gillet *et al.,* 2009), it is a pre-requisite to have a deeper understanding of the genetic variability of vector populations. *R. appendiculatus,* which is widely, distributed across eastern, central and southern Africa acts as a vector of several pathogens of veterinary and economic importance (De Vos 1981; Perry *et al.,* 1990; Walker *et al.,* 2003). Knowledge on the genetic structure and evolution of this species is of great epidemiological importance. In this study, the genetic and phylogeographic structure of *R. appendiculatus* in Kenya was examined to determine whether the genetic diversity in *R. appendiculatus* is spatially structured or is homogeneously distributed at different geographic scales and across a range of host species. Tick colonies which have been maintained as laboratory stocks for more than two decades were investigated to determine if they still represent the variation observed among field populations. Estimates of genetic and allelic diversity were high for field ticks (Table 10). As several ticks are often found in a single animal or herd, these findings suggest that field ticks likely maintain equilibrium genotype frequencies within and across individual animals, herds and populations implying that the effect of genetic drift and inbreeding is weak. This is supported by the mantel test which revealed field populations to be at migration-drift equilibrium (Figure 17). On the contrary, laboratory bred stocks had lower levels of allelic and genetic diversity due possibly to a higher level of inbreeding (overall F_{IS} = 0.305) and genetic drift due to their small effective population sizes.

The genetic data revealed positive F_{IS} coefficients for all the populations analyzed with the exception of KU with a negative value, and ML with a zero value (Table 10). The significant deficiency of heterozygote genotypes, revealed by positive F_{IS} values, suggests that inbreeding is

common among both field and lab stocks but it is weaker in the former. It could also have been caused, at least partly, by Wahlund effect (Wahlund, 1928), that is, sampling together of differentiated genetic groups. If each group exhibits different allele frequencies, then the overall heterozygosity within the pooled sample will be reduced. The excess of heterozygote genotypes in KU may possibly be due to a mating strategy that is designed to deliberately avoid inbreeding. ML on the other hand appears to be a perfectly outbred population where mating is taking place between completely unrelated individuals.

In this study, no evidence of host-related population genetic structure and specificity was found. Ticks that were collected from areas grazed by cattle, cattle-wildlife and wildlife herds shared the same genetic background (Figure 25 and 27). Furthermore, the level of genetic variation between these three groups was the lowest among all the clusters tested in AMOVA. This contrasted with the pattern observed among the five other species of rhipicephaline and the laboratory bred stocks, which aggregated in distinct clusters and had the highest and second highest levels of genetic variation (Table 12). The host is an essential component of a parasite's environment and is important in its evolution (Combes, 2001). In host-parasite associations with strong co-evolutionary signals, co-speciation events may result in host specificity. For parasites with free-living stages, mobile hosts and capacity for population intermixing, local adaptation may direct specificity (Gandon & Michalakis, 2002; Greischar & Koskella, 2007). The evolution of host-specificity has been demonstrated in *Ixodes uriae* (McCoy *et al.,* 2001, 2005), with experiments revealing an adaptive basis for these patterns (Dietrich *et al.*, 2011). The results obtained therefore suggest a lack of host-associated adaptation/specialization by *R. appendiculatus*. This could be a survival strategy by *R. appendiculatus* as it provides micro-niches that allow the ectoparasite to flourish in different ecological environments.

The two approaches used to evaluate genetic structure revealed a very weak genetic structure/substructure and differentiation among field ticks. STRUCTURE (Figure 26) and DAPC (Figure 27) returned $K = 2$ and $K = 6$ respectively, as the most optimal number of genetic clusters among field ticks. Interestingly, at both values of "*K*" all individuals shared the observed genetic backgrounds disproportionately with no correlation between the generated clusters and the type of host species (cattle, cattle-wildlife and wildlife). Therefore the results of STRUCTURE revealed the global genetic structure while those of DAPC revealed fine scale genetic structure among field stocks. Both approaches showed that irrespective of the host species, all field ticks were genetically admixed; they shared a proportion of their genetic makeup and have a low level of population subdivision over a large geographic range. This was supported by the lack of isolation by distance as revealed by the mantel test, low level of genetic differentiation between field ticks and the signal of expansion (demographic and spatial) observed among field ticks (Table 13), suggesting that tick mobility at large spatial scales may be frequent despite the low inherent vagility of *R. appendiculatus*. The high vagility in *R appendiculatus* may be related to the extensive transportation of cattle across Kenya for trade and breeding purposes. Indeed the extensive translocation of cattle has been used to explain the weak genetic structure observed amongst the Indigenous East African zebu cattle, the predominant breed of cattle across Kenya (Okomo, 1997). Previous studies that have examined the population structure of various species of *Ixodidae* found relatively high within-population genetic variation and weak genetic structure over a large spatial range. Delaye *et al.* (1997) suggested that *Ixodes ricinus* populations in Switzerland were panmictic due to a large number of host species at their disposal. McCoy *et al.* (2003) found that *Ixodes uriae* parasitizing black-legged kittiwakes were highly substructured than those parasitizing Atlantic puffins with estimates of population differentiation in kittiwake tick populations being almost twice as large as those of puffin tick populations. The former also showed evidence of isolation by

distance suggesting that opportunities for dispersal for puffin ticks is much greater than for kittiwake ticks due to differences in the movement patterns and social behaviour of host species. In the host-parasite system involving *Geomydoecus actuosi* and its host *Thomomys bottae*, low genetic variation within parasite populations and strong population structure has been directly linked to the movement and social behaviour of the host species (Nadler *et al.,* 1990). In a study that examined the genetic structure of five different nematode parasites from three host species, Blouin *et al.* (1995) observed a clear relation between parasite population structure and the host species exploited. Nematodes parasitizing livestock transported between distant locations showed weak genetic structure compared to nematodes that parasitized white-tailed deer. It is evident from these studies that what drives the genetic structure of parasites with low vagility such as ticks is the dynamics of transmission by their host species. Where there is relatively high dispersal rate among hosts, the parasites become mixed and panmixia is more likely.

Both STRUCTURE and DAPC analyses showed that the laboratory bred ticks appear to be diverging from the field ticks. Only two laboratory bred ticks (MU and MF) shared the same genetic cluster with the field ticks (Figure 24 and 25). On the other hand the laboratory bred stocks seem also to have diverged from each other, each being clearly represented by a distinct genetic cluster. Two selected lineages ML which is significantly less susceptible to *T. parva* infection and KH which is highly susceptible (Young *et al.*, 1995; Odongo *et al.*, 2009) appear to have completely diverged from their original unselected stocks (MU and KU) respectively. Effects of intense selection, several years of isolation and inbreeding have probably resulted in differentiation of these ticks into new distinct populations given that the original unselected stocks were from same geographical region in Central Kenya. Though both strains of ticks have been shown to be able to transmit *T. parva* infection experimentally, it remains to be determined if similar biological and

genetic differences are present in field populations and how they affect transmission of *T. parva* in the field. Two laboratory stocks from Zambia (ZE from eastern province and ZS from southern province) share no genetic similarity. Biological, morphological and genetic differences between *R. appendiculatus* from southern and eastern Zambia have been reported previously (Speybroeck *et al*. 2002; Mtambo *et al.*, 2007a, 2007b, Ochanda *et al.*, 1998) and hence our results confirm that these two populations are genetically differentiated. The LP population shared a similar genetic cluster with ZE (eastern Zambia) population but has completely diverged from FP, a field population sampled recently from the same ranch, while FP has no genetic similarity with ZE (Figure 25 and 26). It is possible that there has been accidental mixing of the two populations, during the long period of maintenance and expansion in the lab. Though few ticks from UG population shared a cluster with some field ticks (Figure 24 and 25) it remained genetically distinct from other laboratory stocks by DAPC (Figure 28 and 29) which could also be attributed to selection, isolation and inbreeding. The ZM population shares a cluster with KU, suggesting a genetic similarity. These two tick stocks have been ranked closely in regards to vector competence represented by abundance of *T. parva* infection in the tick salivary glands (Ochanda *et al.*, 1998).

In conclusion, it has been shown that tick populations sampled from different geographic locations and across three different host species were characterised by a clear lack of genetic structure. Genotypes of individual ticks were highly admixed and indicate extensive gene flow across the geographic range of the species. Of importance was the clear genetic divergence between field tick populations and most of laboratory bred stocks. These results have direct implications for the coevolutionary interactions of *R. appendiculatus* and its hosts with subsequent consequences for the epidemiology of tick-borne diseases. More broadly, the results suggest that for most parasites, their genetic structure is linked to the vagility of their hosts, and that parasite-host systems provides a

rich source of variability in population genetic structure. Host-dependent rates of dispersal between colonies will alter infestation probabilities and local dynamics in domestic or maintenance hosts. Genetic diversity analysis of *R. appendiculatus* can establish if agro-ecological factors such as type of cattle breed, livestock farming systems practiced and nature of tick control method influence its population structure. This information can be used to identify tick population changes associated with acaricide resistance as well as areas where transmission cycles between domestic and maintenance hosts overlap, which helps to measure the risk of tick reinvasion and thus allows more targeted integrated tick control management. Designing effective control strategies for ECF will depend on understanding whether recurrent infestations are due to residual domestic tick populations that survive spraying with acaricides or to reinvasion of ticks from alternative hosts. Successful eradication of *R. appendiculatus* in South Africa (Gutsche, 1979; Cranefield, 1991) and *R. annulatus* and *R. microplus* from USA (George, 1989) relied on identification and isolation of surviving population foci which were targeted for tick control.

4.5 References

Bailey, K.P. (1960). Notes on the rearing of *Rhipicephalus appendiculatus* and their infection with *Theileria parva* for experimental transmission. *Bulletin of Epizoonotic Diseases of Africa,* **8,** 33–43.

Bilgin, R. (2007). Kgtests: a simple Excel Macro program to detect signatures of population expansion using microsatellites. *Molecular Ecology Notes,* ***7*****(3),** 416-417.

Blouin, M. S., Yowell, C. A., Courtney, C. H. and Dame, J. B. (1995). Host movement and the genetic structure of populations of parasitic nematodes. *Genetics* **141,** 1007 - 1014.

Bouyer, J., Ravel, S., Dujardin, J. P., De Meeüs, T., Vial, L., Thévenon, S. and Solano, P. (2007). Population structuring of Glossina palpalis gambiensis (Diptera: Glossinidae) according to landscape fragmentation in the Mouhoun river, Burkina Faso. *Journal of medical Entomology,* **44(5),** 788-795.

Chigagure, N. N., Baxter, G. D. and Barker. S. C. (2000). Microsatellite loci of the cattle tick *Boophilus microplus* (Acari: Ixodidae). *Experimental and Applied Acarology***, 24,** 951–956

Clayton, D. H., Bush, S. E., Goates, B. M. and Johnson, K. P. (2003). Host defense reinforces host–parasite cospeciation. *Proceedings of the National Academy of Sciences,* **100(26),** 15694-15699.

Combes, C. (2001). *Parasitism: the ecology and evolution of intimate interactions*. University of Chicago Press.

Corley, S. W., Jonsson, N. N., Piper, E. K., Cutullé, C., Stear, M. J., & Seddon, J. M. (2013). Mutation in the RmβAOR gene is associated with amitraz resistance in the cattle tick *Rhipicephalus microplus*. *Proceedings of the National Academy of Sciences*, **110(42),** 16772-16777.

Cranefield, P. F. (1991). Science and empire: East Coast fever in Rhodesia and the Transvaal. Cambridge University Press, Cambridge, UK.

Cutulle, C., Jonsson, N. N., Seddon, J. (2009). Population structure of Australian isolates of the cattle tick *Rhipicephalus (Boophilus) microplus*. *Veterinary Parasitology,* **161,** 283–291.

Dai, J., Wang, P., Adusumilli, S., Booth, C. J., Narasimhan, S., Anguita, J., and Fikrig, E. (2009). Antibodies against a tick protein, Salp15, protect mice from the Lyme disease agent. *Cell host and microbe,* **6(5),** 482-492.

De Meeûs, T., Michalakis, Y. and Renaud, F. (1998). Santa Rosalia revisited: "or why are there so many kinds of parasites in the garden of early delights?". *Parasitology Today*, **14(1),** 10-13.

De Vos, A. J. (1981). *Rhipicephalus appendiculatus*: cause and vector of diseases in Africa. *Journal of the South African Veterinary Association*, **52(4),** 315-322.

Delaye, C. Beati, L., Aeschlimann, A. A., Renaud, F. and De Meeuis, T.. (1997). Population genetic structure of Ixodes ricinus in Switzerland from allozymic data: no evidence of divergence between nearby sites. *International Journal of Parasitology*, **27,** 769-773.

Di Giulio, G., Lynen, G., Morzaria, S., Oura, C. and Bishop, R. (2009). Live immunization against East Coast fever--current status. Trends in parasitology **25(2),** 85-92.

Dietrich, M., Gómez-Díaz, E., and McCoy, K. D. (2011). Worldwide distribution and diversity of seabird ticks: implications for the ecology and epidemiology of tick-borne pathogens. *Vector-borne and Zoonotic Diseases*, **11(5),** 453-470.

Earl, D. A. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation genetics resources*, **4(2),** 359-361.

Estrada-Peña, A. (2001). Forecasting habitat suitability for ticks and prevention of tick-borne diseases. *Veterinary parasitology*, **98(1),** 111-132.

Evanno, G., Regnaut, S., and Goudet, J. (2005). Detecting the number of clusters of individuals using the software structure: a simulation study. *Molecular Ecology*, **14,** 2611–2620

Excoffier, L., and Lischer, H. E. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular ecology resources*, **10(3),** 564-567.

Falush, D., Stephens, M., Pritchard, J. (2003). Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics* **164,** 1567-1587.

Fisher, R. A. 1936). The Use of Multiple Measurements in Taxonomic Problems. *Annals of Eugenics*, **7,** 179-188.

Gandon, S., Ebert, D., Olivieri, I. and Michalakis, Y. (1998). Differential adaptation in spatially heterogeneous environments and Host-Parasite coevolution. In: Genetic Structure and Local Adaptation in Natural Insect Populations (S. Mopper and S. Strauss, eds), pp. 325–340. Chapman and Hall, London.

Gandon, S., and Michalakis, Y. (2002). Local adaptation, evolutionary potential and host–parasite coevolution: interactions between migration, mutation, population size and generation time. *Journal of Evolutionary Biology*, **15(3),** 451-462.

Garci´A-Garci´A, J. C., Gonzalez, I. L., Gonzalez, D.M., Valde´ S, M., Mendez, L., Lamberti, J., D'agostino, B.,Citroni, D., Fragoso, H., Ortiz, M., Rodriguez, M., and De La Fuente, J., (1999). Sequence variations in the *B. microplus* Bm86 locus and implications for immunoprotection in cattle vaccinated with this antigen. Experimental and *Applied Acarology* **11**, 883–895.

George, J. E. (1989). Wildlife as a constraint to tick eradication of *Boophilus* spp. In The Eradication of Ticks: Animal Production and Health Paper, Number 75, (ed. Anonymous), pp. 146–153. *Food and Agriculture Organisation, Rome, Italy*.

Gillet, P., Mori, M., Van Esbroeck, M., Van den Ende, J. and Jacobs, J. (2009). Assessment of the prozone effect in malaria rapid diagnostic tests. *Malaria Journal*, **8,** 271.

Goudet, J. (2001). FSTAT, a program to estimate and test gene diversities and fixation indices version 2.9.3. **Irvin, A. D and Brocklesby, D. W.** (1970). Rearing and maintaining *Rhipicephalus appendiculatus* in the laboratory. *J. Inst. Anim Technol.* **21,** 106–112.

Greischar, M. A., and Koskella, B. (2007). A synthesis of experimental work on parasite local adaptation. *Ecology Letters*, **10(5),** 418-434.

Gunnar, H., Knut, H., Røed, and Hans, P. L. (2008). Multiple Paternity in Ixodes ricinus (Acari: Ixodidae), Assessed by Microsatellite Markers. Journal of Parasitology, **94(2)**, 345-347.

Gutsche, T. (1979). There was a Man: the Life and Times of Sir Arnold Theiler K.C.M.G. of Onderstepoort. Howard Timmins, Cape Town, South Africa.

Jakobsson, M., and Rosenberg, N. A. (2007). CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, **23(14),** 1801-1806.

Jombart, T. (2008). Adegenet: a R package for the multivariate analysis of genetic markers *Bioinformatics*, **24,** 1403-1405.

Jombart, T., Devillard, S. and Balloux, F. (2010). Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics,* **11(1),** 94.

Kanduma, E. G., Mwacharo, J. M., Sunter, J. D., Nzuki, I., Mwaura, S., Kinyanjui, P. W., Kibe, M., Heyne, H., Hanotte, O., Skilton, R. A., Bishop, R. P. (2012). Micro- and

minisatellite-expressed sequence tag (EST) markers discriminate between populations of *Rhipicephalus appendiculatus*. *Ticks and Tick-borne Diseases,* **3(3):** 128–136

Krafsur, E. S. (2009). Tsetse flies: genetics, evolution, and role as vectors. *Infection, Genetics and Evolution,* **9(1),** 124-141.

Labruna, M., Naranjo, V., Mangold, A., Thompson, C., Estrada-Peña, A., Guglielmone, A and De la Fuente, J. (2009). Allopatric speciation in ticks: genetic and reproductive divergence between geographic strains of *Rhipicephalus (Boophilus) microplus*. BMC evolutionary biology, **9(1),** 46.

Lachenbruch, P. A. and Goldstein, M. (1979). Discriminant analysis. *Biometrics,* **35,** 69-85.

Lessard, P., L'Eplattenier, R., Norval, R. A. I., Perry, B. D., Dolan, T. T., Burrill, A., Croze, H., Sorensen, M. Grootenhuis, J. G. and Irvin, A. D. (1988). The use of geographical information systems in estimating east coast fever risk to African livestock. *Acta Veterinaria Scandinavia Supplementum,* **(84),** 234-236

Lessard, P., L'Eplattenier, R. and Norval, R. A. I. (1990). Geographical information systems for studying the epidemiology of cattle diseases caused by *Theileria parva*. *The Veterinary Record*, **126,** 255–262.

Loaiza, J. R., Bermingham, E., Sanjur, O. I., Scott, M. E., Bickersmith, S. A. and Conn, J. E. (2012). Review of genetic diversity in malaria vectors (*Culicidae: Anophelinae*). *Infection, Genetics and Evolution,* 12(1): 1-12.

Madder, M., Speybroeck, N., Brandt, J. and Berkvens, D. (1999) Diapause induction in adults of three *Rhipicephalus appendiculatus* stocks. *Experimental and Applied Acarology* **23,** 961–968

McCoy, K. D., Boulinier, T., Chardine, J. W., Danchin, E. and Michalakis, Y. (1999). Dispersal and Distribution of the Tick *Ixodes uriae* within and among Seabird Host Populations: The Need for a Population Genetic Approach. *The Journal ofParasitology* **85(2),** 196-202.

McCoy, K. D., Boulinier, T., Tirard, C., and Michalakis, Y. (2001). Host specificity of a generalist parasite: genetic evidence of sympatric host races in the seabird tick Ixodes uriae. *Journal of Evolutionary Biology*, **14(3),** 395-405.

McCoy, K. D., Boulinier, T., Tirard, C. and Michalakis, Y. (2003). Host-dependent genetic structure of parasite populations: differential dispersal of seabird tick host races. *Evolution,* **57(2),** 288-296.

McCoy, K. D., Chapuis, E., Tirard, C., Boulinier, T., Michalakis, Y., Le Bohec, C., Yvon Le Maho, and Gauthier-Clerc, M. (2005). Recurrent evolution of host-specialized races in a globally distributed parasite. *Proceedings of the Royal Society B: Biological Sciences*, **272(1579),** 2389-2395.

McCoy, K. D. (2008). The population genetic structure of vectors and our understanding of disease epidemiology. *Parasite*, **15(3),** 444-448.

Mtambo, J., Madder, M., Van Bortel, W., Geysen, D., Berkvens, D. and Backeljau, T. (2007a). *Rhipicephalus appendiculatus* and *R. zambeziensis* (Acari: Ixodidae) from Zambia: a molecular reassessment of their species status and identification. *Experimental and Applied Acarology*, **41,** 115-128.

Mtambo, J., Madder, M., Van Bortel, W., Geysen, D., Berkvens, D and Backeljau, T. (2007b). Genetic variation in *Rhipicephalus appendiculatus* (Acari: Ixodidae) from Zambia: correlating genetic and ecological variation with *Rhipicephalus appendiculatus* from eastern and southern Africa. *Journal of Vector Ecology*, **32(2),** 168-75.

Nadler, S. A., Hafner, M. S., Hafner, J. C., and Hafner, D. J. (1990). Genetic differentiation among chewing louse populations (*Mallophaga: Trichodectidae*) in a pocket gopher contact zone (*Rodentia: Geomyidae*). *Evolution*, 942-951.

Norris, D. E., Shurtleff, A. C., Touré, Y. T. and Lanzaro, G. C. (2001). Microsatellite DNA polymorphism and heterozygosity among field and laboratory populations of Anopheles gambiae ss (Diptera: Culicidae). *Journal of medical entomology*, **38(2),** 336-340.

Norval, R. A. I. and Perry, B. D. (1990). The introduction and subsequent disappearance of the brown ear-tick *Rhipicephalus appendiculatus*, from the southern lowveld of Zimbabwe. *Experimental and Applied Acarology*, **9,** 103–111.

Norval, R. A. F., Lawrence, J. A., Young, A. S., Perry, B. D., Dolan, T. T. and Scott, J. (1991). *Theileria parva*: influence of vector, parasite and host relationships on the epidemiology of theileriosis in southern Africa. *Parasitology* **102,** 347–356.

Norval, R. A. F., Perry, B. D. and Young, A. S. (1992). The epidemiology of Theileriosis in Africa. Academic Press Ltd London and San Diego.

Noureddine, R., Chauvin, A. and Plantard, O. (2011). Lack of genetic structure among Eurasian populations of the tick Ixodes ricinus contrasts with marked divergence from north-African populations. *International Journal of Parasitology*, **41,** 183-192.

Ochanda, H., Young, A. S., Medley, G. F. and Perry, B. D. (1998). Vector competence of7 rhipicephalid tick stocks in transmitting 2 *Theileria parva* parasite stocks from Kenya and Zimbabwe. *Parasitology*, **116,** 539-545.

Odongo, D. O., Ueti, M. W., Mwaura, S. N., Knowles, D. P., Bishop, R. P. and Scoles, G. A. (2009). Quantification of *Theileria parva* in *Rhipicephalus appendiculatus* (Acari: Ixodidae) confirms differences in infection between selected tick strains. *Journal of medical entomology*, **46(4),** 888-894.

Okomo, M. A. (1997). Characterisation of the genetic diversity of East African cattle breeds by microsatellite DNA analysis, *(MSc thesis, University of Nairobi)*

Ouma, J. O., Marquez, J. G. and Krafsur, E. S. (2007). Patterns of genetic diversity and differentiation in the tsetse fly *Glossina morsitans morsitans* Westwood populations in East and southern Africa. *Genetica*, **130(2),** 139-151.

Parizi, L, F., Githaka, N. W., Logullo, C., Konnai, S., Masuda, A., Ohashi, K. and da Silva, Vaz Jr I. (2012). The quest for a universal vaccine against ticks: Cross-immunity insights. *The Veterinary Journal*, **194,** 158–165

Park S. D. E. (2001). The Excel Microsatellite Toolkit. *http://oscar.gen.tcd.ie/~sdepark/mstoolkit/index.php*

Perry, B. D., Lessard, P., Norval, R. A. I., Kundert, K. and Kruska, R. (1990). Climate, vegetation, and the distribution of *Rhipicephalus appendiculatus* in Africa. *Parasitology Today*, **6,** 100–194.

Perry, B. D., Lessard, P., Norval, R. A. I., Kundert, K., and Kruska, R. (1990). Climate, vegetation and the distribution of *Rhipicephalus appendiculatus* in Africa. *Parasitology Today*, **6(4),** 100-104.

Poulin, R. and George-Nascimento, M. (2007). The scaling of total parasite biomass with host body mass. *International journal for parasitology* **37(3),** 359-364.

Price, P. W. (1980). Evolutionary biology of parasites (Vol. 15). Princeton University Press.

Pritchard, J. K., Stephens, M. and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* **155**, 945-959.

R Core Team. (2013). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria.

Radley, D. E. (1981). Infection and treatment method of immunization. In Advances in the Control of Theileriosis. A.D. Irvin, M.P. Cunningham and A.S. Young, Eds.: 227-237. Martinus Nijhoff, The Hague.

Randolph. S E. (1994). Population dynamics and density-dependent seasonal mortality indices of the tick *Rhipicephalus appendiculatus* in eastern and southern Africa . *Medical and Veterinary Entomology,* **8(4),** 351–368

Randolph, S. E. (2004). Tick ecology: processes and patterns behind the epidemiological risk posed by ixodid ticks as vectors. *Parasitology*-Cambridge, **129,** S37.

Reich, D. E., & Goldstein, D. B. (1998). Genetic evidence for a Paleolithic human population expansion in Africa. *Proceedings of the National Academy of Sciences*, **95(14),** 8119-8123.

Reich, D. E., Wayne, R. K., & Goldstein, D. B. (1999). Genetic evidence for a recent origin by hybridization of red wolves. *Molecular Ecology*, **8(1),** 139-144.

Rosenberg, N. A. (2004). DISTRUCT: a program for the graphical display of population structure. *Molecular Ecology Notes*, *4*(1), 137-138.

Ruiz-López, M. J., Chaskelson, S., Gompper, M. E., & Eggert, L. S. (2012). Multiple Paternity in the American Dog Tick, *Dermacentor variabilis* (*Acari: Ixodidae*). *Journal of Parasitology*, **98(3),** 498-501.

Sharakhov, O., Braginets, O., Grushko, A., Cohuet, W. M., Guelbeogo, D., Boccolini, M., Weill, C., Costantini, N'. F., Sagnon, D., Fontenille, G., Yan, N. J. and Besansky. A. (2004). Microsatellite Map of the African Human Malaria Vector *Anopheles funestus*. *Journal of Heredity*, **95(1),** 29-34.

Shaw, M. K. and Young A. S. (1994). The biology of *Theileria* species in ixodid ticks in relation to parasite transmission, pp. 23-63. In K. F. Harris [ed.], Advances in disease vector research, volume 10. Springer, New York.

Short, N. J. and Norval, R. A. I. (1981). The seasonal activity of *Rhipicephalus appendiculatus*, Neumann, 1901(*Acarina: Ixodidae*) in the highveld of Zimbabwe. *Journal of Parasitology*, **67,** 77-84.

Solano, P., Ravel, S. and De Meeûs, T. (2010). How can tsetse population genetics contribute to African trypanosomiasis control? *Trends in Parasitology*, **26(5),** 255-263.

Speybroeck, N., Madder, M., Van Den Bossche, P., Mtambo, J., Berkvens, N., Chaka, G., Mulumba, M., Brandt, J., Tirry, L. and Berkvens, D. (2002). Distribution and phenology of ixodid ticks in southern Zambia. *Vet.Medical and Veterinary Entomology* **16(4),** 430-41.

Wahlund, S. (1928). The combination of populations and the appearance of correlation examined from the standpoint of the study of heredity. *Hereditas*, ***11*,** 65-106.

Walker, A. R. and Fletcher, J. D. (1990). *Rhipicephalus appendiculatus* feeding on rabbits and cattle: salivary-gland responses to varying host resistance. Experimental *and Applied Acarology* **8(4),** 285-290.

Walker, A. R., Bouatour, A., Camicas, J. L., Estrada-Peña, A., Horak, I. G., Latif, A. A., Pegram, R. G. and Preston, P. M. (2003). Ticks of domestic animals in Africa: a guide to identification of species. Bioscience Reports, Scotland *Bioscience reports,* 42 Comiston Drive, Edinburg, UK

Walker, A. R. (2011). Eradication and control of livestock ticks: biological, economic and social perspectives. *Parasitology*, **138(08),** 945-959.

Walker, J. C., Keirans, J. E. and Horak, I. G. (2005). The genus *Rhipicephalus (Acari, Ixodidae)*. A guide to the brown ticks of the world. Cambridge University Press. 655 p.

Wang, H., Kaufman, W. R. and Nuttall, P. Wang, H., Kaufman, W.R., and Nuttall, P. A. (1999). Molecular individuality: polymorphism of salivary gland proteins in three species of ixodid tick**.** *Experimental and Applied Acarology* **23,** 969–975

Willadsen, P., Bird, P., Cobon, G. S. and Hungerford, J. (1995). Commercialisation of a recombinant vaccine against *Boophilus microplus. Parasitology;* **110**, S43–50.

Willadsen, P. (2004). Anti-tick vaccines. *Parasitology* **129,** S367–S387.

Young, A. S., Dolan, T. T., Mwakima, F. N., Ochanda, H., Mwaura, S. N., Njihia, G. M., Muthoni, M. W. and Dolan, R. B. (1995). Estimation of heritability of susceptibility to infection with *Theileria parva* in the tick *Rhipicephalus appendiculatus*. *Parasitology* **111**, 31-38.

CHAPTER FIVE: ASSESSMENT OF THE PHYLOGENETIC RELATIONSHIP BETWEEN POPULATIONS OF *Rhipicephalus appendiculatus* USING CYTOCHROME C OXIDASE SUBUNIT I (COI), 12S RIBOSOMAL DNA AND INTERNAL TRANSCRIBED SPACER 2 (ITS2) REGION

5.0 Abstract

Rhipicephalus appendiculatus transmits *Theileria parva*, the causative agent of East Coast Fever (ECF), the most economically important cattle disease in eastern, central and southern Africa. *R. appendiculatus* is a three-host generalist tick that mostly depends on the host for transport between localities and on climatic conditions to survive in particular geographical regions. Thus, the phylogeography and population genetic structure of this species may be influenced by both intra- and inter-population processes such as gene flow, mutations, genetic drift and natural selection. To further assess the phylogenetic relationship between field and laboratory isolates of *R. appendiculatus* and other rhipicephaline tick species from Kenya, this study examined the DNA sequence variation in a 558 bp fragment of the mitochondrial COI gene, a 345 bp fragment of the mitochondrial 12S rDNA and an 1149 bp region of ribosomal nuclear ITS2. Analysis of 332 *R. appendiculatus* COI sequences revealed 28 haplotypes differentiated by a total of 30 polymorphic sites. The COI gene showed high intra- and inter-species diversity. Genetic diversity indicators based on the COI gene were highest in ticks collected from areas grazed by wildlife compared to those collected from areas grazed by cattle while nine laboratory tick stocks were found to have no gene diversity. Phylogenetic and network relationship inferred from the COI gene sequences showed that field *R. appendiculatus* ticks occurring in Kenya are characterized by the existence of two distinct and highly differentiated genetic haplogroups separated by at least 12 mutation points. These two groupings were not correlated to geographical origin of the samples. Tick populations showed high within-population genetic variation based on the COI gene and relatively low inter-population differentiation. Analysis of the ITS2 gene in *R. appendiculatus* revealed one major haplotype suggesting high conservation of this gene within the tick species. The 12S gene resulted in two major haplotypes that closely correlated with the two major haplotypes differentiated by COI analysis. Thus, COI and 12S genes were found to be superior genetic markers for intra-species population genetic studies in *R. appendiculatus* over the ITS2. Mismatch distribution analysis based on the COI gene further revealed that one haplogroup of *R. appendiculatus* may have diverged from the other and has experienced recent sudden population expansion at a much faster rate than its ancestor. These findings may have important taxonomical and distribution implications for *R. appendiculatus* and may point to an ongoing speciation of the tick in sub-Saharan Africa. Moreover, these results suggest that genetic differentiation of *R. appendiculatus* may not be driven by agro-ecological and climatic factors alone as previously thought, and thus identifying other factors may help explain the apparent sudden population expansion observed within the sub-Saharan region. The diversity of the ticks as well as the possible ongoing speciation could have implications on the epidemiology of ECF and therefore may call for consideration when designing effective control strategies.

5.1 Introduction

The polymorphic ribosomal and mitochondrial DNA markers are frequently utilized for deducing genetic diversity within and between natural populations. Mitochondrial DNA (mtDNA) has been extensively used to analyze phylogeny, phylogeography and population structure of many organisms. It is an ideal molecular marker for evolutionary and population-wide genetic studies (Avise *et al.*, 1987). It has been used to study the phylogeny and population structure of arthropods of the subphylum Chelicerata (ticks and mites) (Norris *et al.*,1999; Anderson & Trueman, 2000; Roehrdanz *et al*., 2002; Hassanin, 2006; Stevens & Hogg, 2006**;** Nava *et al.*, 2008; Trout *et al.*, 2010). The mtDNA genome has several important characteristics that make it suitable for genetic studies within and among populations and that influence both analysis and interpretation of data (Avise, 2000). Firstly, because of the absence of a relatively efficient mutation repair mechanism, it evolves rapidly compared to nuclear genes (Wilson *et al.*, 1985), thus making it sufficiently variable to address evolutionary questions at low taxonomic levels such as variation within species, genera or families (Harrison, 1989; Saccone *et al.*, 2000; Shao & Barker, 2007). Secondly, mtDNA was thought to have evolved almost in a neutral way, allowing a differentiation between historical and current processes (Holderegger *et al*., 2006) but recent evidence suggests that mtDNA may be under some selection (Dowling *et al*, 2008). Thirdly, the mitochondrial genome is usually maternally inherited, enabling genealogical researchers to trace maternal lineage far back in time. Because of a lack of recombination, a mitochondrial gene represents a single ancestry lineage (Avise *et al.*, 1987; Avise, 1991; Avise, 2009). Fourthly, the effective population size of mtDNA is smaller than that of nuclear DNA because of its haploid nature and maternal inheritance. Thus, within a population the frequency of a single mtDNA haplotype should fluctuate more rapidly than nuclear DNA alleles, and consequently mtDNA should be more sensitive to demographic incidents such as founder events and bottlenecks than nuclear DNA (Avise, 1991; Birky *et al*., 1989; DeSalle & Giddings,

1986). As a consequence, the possible loss or gain of an mtDNA haplotype would be considerable for a small population (Neigel & Avise, 1985). Notably, amongst the 37 mitochondrial genes, cytochrome oxidase subunit I (COI), cytochrome b (Cytb), 12S and 16S (small and large rRNA subunits) have been the most used in studies of population genetics and evolution of parasitic arthropods (Navajas & Fentom, 2000; Shao & Barker, 2007).

Although possessing almost the ideal characteristics and having contributed significantly towards an understanding of the evolution and natural history of organisms, mtDNA markers have some limitations. In some cases, recombination and heteroplasmy (Rokas *et al*., 2003; Tsaousis *et al.*, 2005; Galtier *et al.*, 2009), absence of neutrality (Ballard & Kreitman, 1995; Dowling *et al.*, 2008), non-constant mutation rate (Martin, 1995; Nabholz *et al.*, 2007; Galtier *et al*., 2009), mitochondrial pseudogenes in the nuclear genome (Zhang & Hewitt, 1996; Bensasson *et al.*, 2001; Gaziev & Shaikhaev, 2010), and duplicated genes (Shao *et al.*, 2005; Morris-Pocock *et al.*, 2010) can discourage the use of mtDNA markers. Conflict between gene and species trees is often reported and mtDNA only provide a single window on the evolutionary history of species (Pollard *et al*., 2006). The use of multiple genes drawn from different sources can overcome the drawbacks or increase the accuracy in the construction of a phylogenetic history of species (Gadagkar *et al*., 2005; Balmer *et al*., 2011).

Cytochrome c oxidase subunit I (COI) is one of the most conserved mitochondrial protein coding genes. COI has been shown to be a suitable marker for global bioidentification system for animals (DNA barcoding) (Hebert *et al.*, 2002). The universal PCR primers for COI gene are very robust, enabling recovery of its 5′ end from representatives of most, if not all, animal phyla (Folmer *et al*. 1994; Zhang & Hewitt 1997). COI appears to possess a greater range of phylogenetic signal than

any other mitochondrial gene (Knowlton & Weigt 1998). The evolution of this gene is rapid enough to allow the discrimination of not only closely-related species, but also phylogeographic groups within a single species (Cox & Hebert 2001; Wares & Cunningham 2001). Although COI may be matched by other mitochondrial genes in resolving such cases of recent divergence, this gene is more likely to provide deeper phylogenetic insights than alternatives such as cytochrome b (Simmons & Weller 2001) because changes in the COI amino-acid sequence occur more slowly than those in cytochrome b or any other mitochondrial gene (Lynch & Jarrell 1993). The usefulness of COI as a phylogenetic marker over taxonomical or morphological means of species identification was demonstrated by Hebert *et al.,* (2004) who showed that the butterfly *Astraptes fulgerator* is a complex of 10 species and not seven as previously thought. The phylogeny of 21 species and subspecies of ticks from the subfamilies *Rhipicephalinae* and *Hyalomminae* inferred using COI and 12S showed the genus *Rhipicephalus* to be paraphyletic with a subgenus *Boophilus* (Murrell *et al*, 2000). Variation in the COI gene has provided further evidence of geographical differentiation in *R. appendiculatus* vector ticks from Eastern and Southern Provinces of Zambia (Mtambo *et al,* 2007) with several closely related haplotypes in Eastern Province of Zambia and a single haplotype in Southern Province being reported. Recently, using the COI gene, the phylogeography of two ixodid ticks, *H. rufipes* and *A. habraeum* was shown to be influenced by interspecific competition and host preference (Cangi *et al,* 2013).

Nuclear ribosomal genes such as internal transcribed spacers have been used to resolve phylogenetic relationships of ticks (Rich *et al*., 1997; Rees *et al.*, 2003; Marreli *et al.*, 2007) and mites (Zahler *et al*., 1995; Navajas *et al*., 1999) both at the species and population levels. Although the mtDNA genes provide data with greater resolution for phylogeographic studies, nuclear DNA can present a more complete genealogy, resulting from historical and environmental processes

experienced by populations (Hare, 2001). The nuclear genome is usually biparentally inherited with regular recombination (Freeland, 2005). The large nuclear ribosomal locus in eukaryote cells comprises three genes 18S, 5.8S and 28S, that encode the ribosomal subunits. These regions are typically tandemly formed by several copies that are homogenized by concerted evolution (Roderick *et al*., 1996; Zhang & Hewitt, 2003). Between these genes are the non-coding regions, namely internal transcribed spacers 1 (ITS1) and the internal transcribed spacers 2 (ITS2) (Dixon, 1991; Hwang & Kim, 1999). In eukaryotes internal transcribed spacers (ITS) separate the 18S (small subunit), 5.8S and 28S (large subunit) rRNA genes. ITS1 separates 18S and 5.8S, whereas ITS2 separates the 5.8S and 28S genes. The transcribed spacers contain signals for processing rRNA transcripts and apparently form stable secondary structures (Hillis & Dixon, 1991). In general, the ITS tend to evolve more rapidly than the coding regions because of the near absence of selection in this region (Hwang & Kim, 1999). The high polymorphism, paternal and matrilineal history, neutrality, large effective population size (four times larger than the mtDNA), high copy number and easy amplification (Hlinka *et al*., 2002) make these nuclear regions almost the perfect markers to combine with mtDNA when studying the genetic structure of species (Zhang & Hewitt, 2003; Moore, 1995). However, some limitations that can distort the evolutionary history must be considered when using the ITS region and other nuclear markers. These include false information caused by recombination (Jiggins *et al*., 2001), concerted evolution, absence of neutrality (Elder & Turner, 1995; McLain *et al.*, 1995), difficulties in scoring alleles due to heterozygosity (two different haplotypes), in some cases low divergence and polytomy, and PCR and sequencing difficulty (Zhang & Hewitt, 2003).

The ITS regions are useful for defining species boundaries and inferring phylogenies because there tends to be little intraspecific variation but considerable interspecific variation (Murrell *et al*, 2001).

Highly conserved rRNA genes flank both ITS regions so it is relatively simple to amplify them by PCR. This region has been used to study genetic variation in ticks. A 274bp part of ITS2 sequence was used to differentiate between six tick species in the *R. sanguineus* complex which share a lot of common phenotypic and morphological characteristics (Zahler *et al*., 1997). Barker (1998) investigated the intraspecific variation in the ITS2 of *Rhipicephalus* and *Boophilus* spp. and found evidence of an 82 base pair (bp) repeat common to *R. appendiculatus, R. zambeziensis, R. evertsi, B. microplus* and *B. decoloratus*. He was able to distinguish between Australian populations of *B. microplus* and those from Kenya, South Africa and Brazil as well as *R. appendiculatus* from Kenya, Zimbabwe and Zambia using the ITS2 rDNA. Using a novel real-time ITS2 PCR assay, *I. scapularis, I. pacificus, D. variabilis* and *A. americium* which are medically important ticks in the US were successfully differentiated (Shona *et al,* 2006). ITS2 has also been sequenced extensively in many other ticks to study closely related species (McClain *et al*., 1995; Zahler *et al*., 1995; Rich *et al*., 1997; Fukunaga *et al*., 2000).

Mitochondrial ribosomal genes have been widely used in molecular systematics. Their high copy number makes them much easier to work with than single-copy nuclear genes and their strictly maternal inheritance has been particularly useful at the intraspecific level. Mitochondrial 12S rDNA together with ITS2 have been used to confirm the taxonomic status of *R. appendiculatus* and *R. zambeziensis,* two closely related species of ticks from Eastern and Southern provinces of Zambia, whose morphological identification in the areas where they co-occur is difficult (Mtambo *et al.*, 2007b). The phylogeny of rhipicephaline ticks has been studied with mitochondrial 12S rDNA (Murrell *et al*., 1999, 2000).

The intimate relationship between the host and the parasite can lead to complex differentiation and speciation patterns in the parasite that may be affected by the evolutionary history and genetic structure of the host (Wickström *et al.*, 2003; Reed *et al.*, 2004). Population genetic divergences are mainly determined by the amount of gene flow (Hilburn & Sattler, 1986), which is influenced by dispersal ability, barriers to dispersal, species reproduction biology, fragmented habitats and metapopulations dynamics (Slatkin, 1987; Freeland, 2005). In ticks, host availability, host migration, ecological requirements of juvenile and adult stages and tick population sizes should be taken into account in order to understand the genetic variability (Lampo *et al.*, 1998; Kain *et al.*, 1999). *Rhipicephalus* tick species lay eggs off the host and use between one and three host species to complete their life cycles. In addition to feeding, the ticks depend on the vertebrate host for transport between localities and on climate conditions to survive in an area. Thus, the phylogeography and genetic structure of these species may be influenced by intra-population and inter-population processes such as gene flow, mutations, genetic drift and natural selection factors that are all specific to a particular species (Amos & Harwood, 1998). To further assess the phylogenetic relationship between field and laboratory rhipicephaline tick species, this study examined the DNA sequence variation in the mitochondrial COI, a fragment of the mitochondrial 12S rDNA and ribosomal nuclear ITS2 region.

5.2 Materials and methods

5.2.1 Tick samples

A total of 419 tick samples were used for the COI gene amplification and sequence analysis. This included 332 individuals of *R. appendiculatus* from 10 field populations and 12 laboratory stocks. In the 10 Kenyan field populations (Figure 10, Chapter 3), 118 individuals were collected directly from areas grazed exclusively by cattle (six populations) while 89 individuals were from areas co-

grazed by wildlife and cattle (four populations). A total of 125 individuals belonged to the 12 laboratory colonies which had been bred and maintained as closed genetic stocks at the Tick Unit of the International Livestock Research Institute (ILRI). Of the 12 laboratory stocks, one originated from ticks collected in UG (Uganda) (n = 12), one from ZM (Zimbabwe, West Mashonaland, n = 12) and two from ZE (Zambia eastern province, n = 12; ZS, southern province, n = 8) while the rest were obtained from Kenya. These were reared and managed as described in Chapter 3 Section 3.2.2 to 3.2.2.3 (Bailey, 1960; Irvin and Brocklesby, 1970). The detailed description of the tick populations used in this study is given in Chapter 3 Table 6 while the full names of the populations are listed in Table 14 below.

Eighty one field ticks from five other rhipicephaline species: *R. pravus* (n = 16), *R. praetextatus* (n = 5), *R. zambeziensis* (n = 12), *R. evertsi* (n = 12), *R. pulchellus* (n = 11), and an unidentified species from Ruma National Park in Kenya (n = 25), were included (Table 14). For the 12S gene amplification and sequence analysis, a total of 147 tick samples were used. These comprised of 93 *R. appendiculatus* samples consisting of 64 field and 29 laboratory samples, 31 field *R. preatextatus* individuals and 23 field *R. pravus* samples. A total of 152 tick samples were used for the ITS2 gene amplification and sequence analysis. These comprised of 96 *R. appendiculatus* samples, 30 field *R. praetextaus* samples and 26 field *R. pravus* samples.

5.2.2 Extraction of whole tick genomic DNA

Genomic DNA was extracted using the DNeasy® Blood and Tissue Kit (Qiagen GmbH, Germany) with minor modifications as described in Chapter 3 Section 3.2.4.

5.2.3 PCR amplification of COI gene

Cytochrome oxidase (CO1) gene sequences were obtained by PCR amplification of individual *R. appendiculatus* lab and field ticks using the forward primer LCO1490 (5′-GGTCAACAAATCATAAAGATATTGG-3′) and reverse primer HC02198 (5′-TAAACTTCAGGGTGACCAAAAAATCA-3′) (Folmer *et al*, 1994). The 50 µl PCR reaction consisted of 10 µl Taq polymerase 5x colourless buffer (Promega), 5 µl 25mM $MgCl_2$, 0.5 µl 20 mM dNTP mix, 1.25 U Taq polymerase (Promega), 1 µl 0.25 pmol of each of the respective primers and 2 µl of individual *R. appendiculatus* gDNA as a template. The final volume of the reaction mixture was made up to 50 µl with nuclease free water (Promega). The PCR conditions consisted of an initial denaturation at 95°C for 5 min followed by 35 cycles of 94°C for 1 min, annealing at 40°C for 1 min and extension at 72°C for 1.5 min. The final extension was carried out for 10 min at 72°C.

5.2.4 PCR amplification of 12S ribosomal RNA gene

The primers pairs SR-J-1499 (5′-TACTATGTTACGACTTAT-3′) and SR-N-14594 (5′-AAACTAGGATTAGATACCC-3′) (Simon *et al*., 1994) were used to amplify the 12S rDNA. A 50 µl PCR reaction consisted of 10 µl 5x colourless buffer (Promega), 5 µl 25mM $MgCl_2$, 0.5 µl 20 mM dNTP mix, 1.25 U Taq DNA polymerase (Promega), 1 µl 0.25 pmol of each of the respective primers and 3 µl of individual tick sample gDNA as a template. The final volume of the reaction mixture was made up to 50 µl with nuclease free water (Promega). The PCR conditions consisted of an initial denaturation at 95°C for 5 min followed by 35 cycles of 94°C for 1 min (denaturation), 50°C for 1 min (annealing) and 72°C for 2 min (primer extension). A final primer extension was carried out at 72°C for 10 min.

5.2.5 PCR amplification of the ITS2 region

The ITS2 region (1-1.25 kb) was PCR amplified as two fragments: a full-length fragment, plus an internal 721 bp fragment to ensure good sequence coverage. The full-length fragment was amplified with the forward primer 3SAF (5′-CTAAGCGGTGGATCACTCGG-3′) (Barker, 1998) and reverse primer ITS2R (5′-ATATGCTTAAATTCAGCGGG-3′) (Domanico *et al*. 1997). The inner ITS2 fragment was amplified with primers ITS2_Int_F1 (5′-AAGAGCCTGCAGGGAAAG-3′) and ITS2_Int_R1 (5′-CACGTTCGTAAACCCATC-3′). The 50 µl PCR reaction consisted of 10 µl 5x colourless buffer (Promega), 5 µl 25 mM $MgCl_2$, 0.5 µl 20 mM dNTP mix, 1.25 U Taq DNA polymerase (Promega), 1 µl of 0.25 pmol of each of the respective primers and 3 µl of individual tick gDNA as the template. The final volume of the reaction mixture was made up to 50 µl with nuclease-free water (Promega). The PCR conditions consisted of an initial denaturation at 95°C for 5 min followed by 35 cycles of 94°C for 1 min (denaturation), 55°C for 1 min (for primer pair 3SAF and ITS2R) or 52°C (for primer pair ITS2_Int_F1 and ITS2_Int_R1) (annealing) and 72°C for 2 min (primer extension). A final primer extension was carried out at 72°C for 10 min.

5.2.6. Purification and sequencing of amplified PCR products

Amplified COI, 12S and ITS2 PCR products were purified using the QIAquick PCR Purification Kit (Qiagen GmbH, Germany) following the manufacturer's protocol as follows: Five volumes of Buffer PB were added to 1 volume of the PCR sample, mixed thoroughly using a pipette and the mixture transferred into a QIAquick spin column placed in a 2 ml collection tube. The column with the mixture was then centrifuged at 14,000 rpm for 60 s to bind the DNA. The flow-through was discarded and the QIAquick column with bound DNA placed back into the same tube, washed twice with 0.75 ml Buffer PE. The QIAquick column was then placed back into the same tube and centrifuged for an additional 1 min at 14, 000 rpm to remove residual ethanol. The column was then

placed in a sterile 1.5 ml microcentrifuge tube and 30 µl nuclease free water added to the center of the QIAquick membrane. The column was then incubated at room temperature for 5 min and centrifuged at 14, 000 rpm for 2 min to recover the purified DNA sample. The quality of the purified DNA was analyzed by running 5 µl on a 1.2% agarose gel. DNA bands were visualised using a UV lamp after post-staining with about 200 ml 1x TAE containing 10 µl ethidium bromide (10 mg/ml) for 30 min and destaining in deioinised water for 25 min. Purified COI, 12S and ITS2 PCR products were sequenced directly using the BigDye terminator v3.1 kit (Applied Biosystems) at the Segolip unit of the BecA-ILRI Hub. The COI and 12S genes were sequenced using the gene specific forward and reverse primer pair used for PCR amplification while the ITS2 region was sequenced with the two primer pairs used for amplification. Sequencing reactions were separated by electrophoresis on an AB1 3730 96-capillary automated sequencer (Applied Biosystems, UK).

5.2.7 Data analysis

5.2.7.1 Sequence editing and multiple alignments

Sequences trace files obtained from the ABI sequence software were visually inspected and manually edited using CLC Main Workbench version 6.8.3 software (CLC bio). The sequences were trimmed to remove poor quality 3′ and 5′ sequences. Consensus sequences were generated from a forward and reverse sequence for each of the COI and 12S PCR products (i.e. two individual sequences). The ITS2 consensus sequences were generated using forward and reverse sequences from both the full-length and inner PCR products (i.e. four individual sequences). Species identity was confirmed by BLASTN searches (http://blast.ncbi.nlm.nih.gov/Blast.cgi). The *R. appendiculatus* sequences were trimmed to uniform sizes (558 bp for COI, 345 bp for 12S and 1149 bp for ITS2). Multiple sequence alignments were constructed using ClustalW2 in CLC Main

Workbench. For alignment of *R. appendiculatus* ITS2 with *R. pravus*, the sequences were trimmed to cover a uniform region of 878 bp while that of 12S covered a uniform region of 347 bp.

5.2.7.2 Population genetic variability and structure

The multiple alignments generated were used for haplotype reconstruction in DnaSP v5.10.01 (Librado & Rozas, 2009). Percent identities between haplotypes were generated using Clustal Omega (clustalo) v1.2.1 program at EMBL-EBI (http://www.ebi.ac.uk/Tools/msa/clustalo/). Genetic variation (nucleotide and haplotype diversity and mean nucleotide differences) based on the COI gene for each population was calculated using DnaSP v5.10.01. Population genetic structure was assessed by nested analysis of molecular variance (AMOVA) using Arlequin v3.5. The groupings used for AMOVA were as follows: (i) the overall dataset assuming no groups; (ii) between *R. appendiculatus* and other rhipicephaline species (iii) between *R. appendiculatus* tick populations from cattle and from wildlife areas; (iv) between *R. appendiculatus* from field and laboratory populations; (v) between three groups of *R. appendiculatus* populations (cattle, cattle-wildlife, wildlife).

5.2.7.3 Population demographic structure

Population dynamics based on the COI gene were inferred on the basis of mismatch distribution patterns (Rogers & Harpending, 1992) for all *R. appendiculatus* populations. The frequency distribution of the number of mutational differences between haplotypes (mismatch analyses) was calculated using Arlequin (v3.5) (Excoffier & Lischer, 2010) to determine whether the number of pairwise differences among all DNA sequences reflected expanding or stable populations. The goodness-of-fit (discrepancy between observed and expected values) for the expansion model was performed using the sum of squares deviation (SSD) and Harpending's raggedness index ''RI'' (Harpending, 1994) following 1000 coalescent simulations. The two tests were augmented with the

Fu's FS statistic (Fu, 1997) and Tajima's D (Tajima, 1989, 1996) coalescent-based estimators of selective neutrality, whose significance was also tested with 1000 coalescent simulations in Arlequin v3.5.

5.2.7.4 Phylogenetic and median-joining network analysis

Phylogeny reconstruction and evolutionary analyses were performed separately on the COI, 12S rDNA and ITS2 haplotype sequences. The evolutionary models that best suited these gene sequence data were selected using Molecular Evolutionary Genetics Analysis (MEGA) v6.0 software (Tamura *et al.,* 2013). Maximum likelihood (ML) statistical method based on the best-fit Tamura 3-parameter model of substitution (Tamura, 1992), assuming a discrete Gamma distribution (5 categories (+G, parameter = 0.3476)) to model evolutionary rate differences among sites, and an evolutionarily invariable rate of variation of sites ([+I], 0.0000% sites) was used to construct the COI distance tree in MEGA 6. Stability of the ML tree so obtained was assessed via 1000 bootstrap replicates. The trees were drawn to scale, with branch lengths measured in the number of substitutions per site. Haplotype sequences from other rhipicephaline ticks were included in the analysis. The COI analysis involved 33 nucleotide sequences (28 *R. appendiculatus* haplotypes, one reference *R. appendiculatus* sequence from Genbank (Accession number AF132833) and four sequences from other rhipicephaline species) each with a length of 558 bases. The 12S evolutionary history was also inferred by using the ML method based on the Tamura 3-parameter model. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+G, parameter = 0.0943)). The analysis involved 12 haplotype sequences (five *R. appendiculatus* haplotypes, four *R. pravus* haplotypes, and three *R. praetextatus* haplotypes) and 9 Genbank sequences each with a length of 323 bases. The ITS2 analysis was also done similarly using the ML method. A discrete Gamma distribution was used to model evolutionary rate

differences among sites (5 categories (+G, parameter = 0.4598)). It involved eight haplotype sequences (two *R. appendiculatus* haplotypes, three *R. pravus* haplotypes, and three *R. preatextatus* haplotypes) and five sequences from Genbank each with a length of 861 bases. In the COI analysis, gaps were treated as missing data, while for ITS2 and 12S rDNA analysis gaps were treated as complete deletion sites. The tree with the highest log likelihood was drawn for each gene. Initial tree(s) for the heuristic search were obtained by applying the Neighbor-Joining method to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach. The trees were drawn to scale, with branch lengths measured in the number of substitutions per site. Median-joining networks (Bandelt *et al*., 1999) were constructed for the COI gene using the program NETWORK 4.6 with default settings (fluxus-engineering.com).

5.3 Results

5.3.1 Amplification of COI, 12S and ITS region

PCR amplification of COI generated a 710 bp product, 12S resulted to a product of 345 bp while the internal ITS2 fragment resulted to a product of 721 bp and the outer fragment generated a 1.1kb DNA fragment. Figure 30 below shows purified PCR products of COI, 12S and ITS2 internal fragment. A total of 419 COI, 147 12S and 152 ITS2 good quality consensus sequences were obtained.

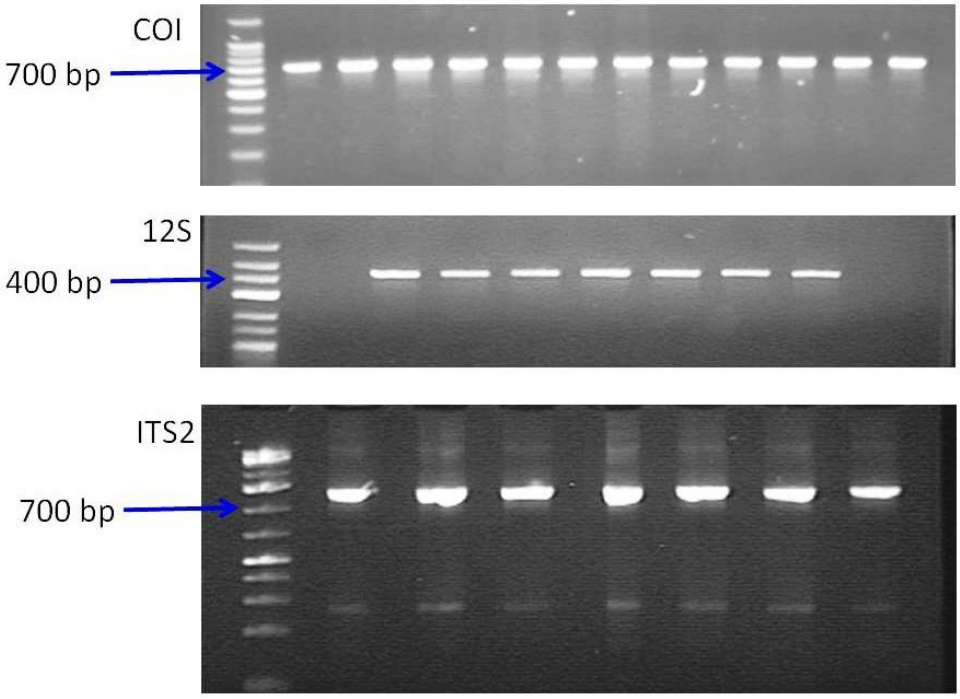


Figure 30: Purified COI, 12S and ITS2 PCR products
Upper gel: Purified COI PCR products of 12 individual *R. appendiculatus* samples ran in a 2% agarose gel against DNA markers (100 bp DNA ladder; Promega). **Middle gel**: Purified 12S PCR products of 7 individual *R. appendiculatus* ran in a 2% agarose gel against DNA markers (GeneRuler Low Range DNA Ladder; Thermoscientific) One sample (well 1) failed to amplify. **Lower gel:** Purified ITS2 (internal fragment) of 7 individual *R. appendiculatus* samples ran in a 2% agarose gel against DNA markers (GeneRuler 100 bp DNA Ladder; Thermoscientific). The expected product sizes were 710 bp for COI, 380 bp for 12S and 721 bp for the internal ITS2 fragment.

5.3.2 Morphological and molecular identification of tick species

Morphological species identification of ticks was based on taxonomic descriptions and illustrations whereas molecular identity was confirmed by searching the COI sequences against the non-redundant nucleotide (BLASTN) genbank sequence databases using the BLAST algorithm (Altschul *et al.*, 1990). Of the five hundred and seventy six tick samples subjected to COI PCR, 419 tick samples generated a high quality consensus sequence of 558 bases while the remaining 157 generated poor quality sequences and were not studied further.

Based on morphologic characteristics and molecular identification, 332 COI sequences were confirmed to belong to *R. appendiculatus* (97-100% identity). Four out of 12 *R. zambeziensis* matched to *R. zambeziensis* with an identity of 94% while the remaining 8 *R. zambeziensis* sequences matched to *R. appendiculatus* (97-98% identity). Twelve ticks were confirmed to belong to *R. evertsi* (99-100% identity). Eleven sequences belonged to *R. pulchellus* (99% identity). It was noted that for some species, morphological and molecular identity were not congruent. Five sequences of ticks morphologically identified as *R. praetextatus* matched to *R. simus* COI sequences (96-98% identity). Sixteen sequences of ticks morphologically identified as *R. pravus* matched with an identity of 90% to *Rhipicephalus* (*Boophilus) geigyi* and 88% to *R. pravus* (Murell *et al.,* 2001). The sequences of 25 ticks from the Ruma population morphologically identified as *R. praetextatus* matched to *R. simus* COI sequences (90-92% identity). This group of ticks was treated as unknown *Rhipicephalus* species. Seven sequences of ticks whose morphological characterization was not done were excluded from further analysis.

5.3.3 COI sequence variability and haplotype diversity

Comparison of sequence homology revealed an average percent identity of 94% between *R. appendiculatus* and *R. zambeziensis* sequences was, 87% between *R. appendiculatus* and *R. praetextatus* sequences, 85% between *R. appendiculatus* and *R. pravus* sequences, 86% between *R. appendiculatus* and sequences of the unidentified species, and an average of 86% between *R. appendiculatus* and *R. pulchellus* sequences.

Analysis of the 558 bp COI region in the 332 *R. appendiculatus* sequences revealed a total of 28 haplotypes (Table 14). A total of 30 polymorphic sites were found in this region (Table 15). COI sequence identities ranged from 97-100% in the 28 *R. appendiculatus* haplotypes. The two major

haplotypes (Hap_4 and Hap_1), which had 107 (32.2%) and 94 (28.3%) sequences, respectively, accounted for 60.5% of all the sequences analysed. Hap_4 represented 100% of the sequences from populations ML, LP, ZE and eight out of the 12 sequences in KH, while Hap_1 represented all the sequences from populations SAL, SAN, KU, ZM and eight out of the 12 sequences in MF and ZS and 19 out of the 20 sequences from KF (Table 16). All the sequences (n = 12) from UG were in a single haplotype (Hap_8). The overall mean number of haplotypes in all populations of *R. appendiculatus* was 3.77. The highest number of haplotypes found in the tick populations was 10 (KT and FP) while the lowest number of haplotypes found was one in nine laboratory stocks. The mean number of haplotypes was 6 in tick populations from areas grazed by cattle, 2 in laboratory stocks of *R. appendiculatus* and 7 in populations from areas grazed by wildlife.

The COI sequence variability and haplotype diversity measures of the 22 *R. appendiculatus* populations and six other rhipicephaline populations are summarized in Table 14. The overall COI haplotype (gene) diversity in the 22 *R. appendiculatus* populations was 0.802±0.014 with a range of 0.900±0.161 (RUM2) to 0.000 in nine laboratory stocks with an overall mean of 0.402±0.004. The mean of haplotype diversity was 0.694 ±0.083 in ticks from cattle, 0.143±0.029 in laboratory tick stocks and 0.767±0.0064 in populations from wildlife areas. Overall nucleotide diversity was 0.0123±0.00019 in the 332 sequences with a range of 0.010±0.06 (KT) to 0.000 in nine laboratory stocks with an overall mean of 0.0048 ±0.0016. The overall mean number of nucleotide differences was 6.865 with a range from 8.100±4.534 (RUM2) to 0.000 in nine laboratory stocks and an overall mean of 2.679±1.402. The mean nucleotide differences were 4.132 ±2.212 in tick populations from areas grazed by cattle, 1.45±0.747 in laboratory populations and 4.179±2.155 in tick populations from areas grazed by wildlife.

Other rhipicephaline species showed a total of 25 haplotypes: *R. pravus* (n = 6), *R. praetextatus* (n = 3), *R. zambeziensis* (n = 2), *R. evertsi* (n = 2), *R. pulchellus* (n = 1), and an unidentified *Rhipicephalus* species (Ruma) (n = 11). The mean number of haplotypes in these species was four (Table 14). The haplotype diversity ranged from 0.857±0.054 (Ruma) to 0.000 (*R. pulchellus*) with an overall mean of 0.595±0.074. The nucleotide diversity ranged from 0.01564 ±0.0017 (*R. zambeziensis)* to 0.000 in *R. pulchellus* with an overall mean of 0.012±0.002. The mean nucleotide differences in these species ranged from 8.727±4.335 (*R. zambeziensis*) to 0.000 (*R. pulchellus*).

The phylogenetic relationship between the 53 COI haplotypes obtained from the 413 sequences analysed is shown in Figure 31. It was observed that the *R. appendiculatus* haplotypes clustered in one clade. There was a strong support for the close genetic relationship between *R. appendiculatus* and *R. zambeziensis*. However, one *R. zambeziensis* haplotype clustered with the *R. appendiculatus* group suggesting a possible mix up of the two species during maintenance since the *R. zambezienis* was from a laboratory stock. There was a strong support for the close relationship between *R. praetextatus* and *R. simus* which fell in one clade. These two species appeared closely related to the unknown species from Ruma. *R. pravus* haplotypes obtained from this study were strongly separated from the *R. pravus* reference sequence (RpvAF132837) from the databases suggesting that they possibly represent two different species. However, the haplotypes appeared closely related to *R. evertsi*, though the support was weak. There is also strong support for the closer genetic relationship between the unknown *Rhipicephalus* species from Ruma and *R. praetextatus* and *R. simus. R.evertsi* and *R. pravus* appear closely related the when compared to other species.

Table 14: Summary of COI sequence variability and genetic diversity measures of the 22 *R. appendiculatus* populations and six other rhipicephaline species

Population[a]	Sample size	No. of Haplotypes	Haplotype Diversity (SD)	Nucleotide Diversity±(SD)	Mean number of nucleotide differences (SD)	Fu's FS (P-value)	Tajima's D (P-value)	Sum of squared deviation (SSD) (P-value)	Harpending's Raggedness Index (RI)
Field (Cattle)									
Kilifi (KF)	20	2	0.100 (0.088)	0.00018 (0.00016)	0.100 (0.1775)	-0.8793 (0.080)	-1.1644 (0.123)	0.0000 (0.288)	0.650 (0.810)
Makuyu (MK)	25	9	0.817 (0.055)	0.00505 (0.0014)	2.820 (1.539)	-1.340 (0.264)	-1.015 (0.175)	0.0285 (0.140)	0.0801 (0.228)
Kitale (KT)	29	10	0.865 (0.037)	0.0100 (0.0056)	5.584 (2.760)	0.747 (0.678)	0.999 (0.860)	0.0390 (0.489)	0.0302 (0.747)
Busia (BU)	18	7	0.784 (0.085)	0.01108 (0.0062)	6.1830 (3.0824)	2.062 (0.8440)	0.4631 (0.715)	0.0887 (0.210)	0.0985 (0.421)
Rusinga (RU)	21	5	0.700 (0.073)	0.00360 (0.0018)	2.0095 (1.178)	0.781 (0.690)	**-1.749 (0.026)**	0.1279 (0.180)	0.0956 (0.280)
Ruma (RUM2)	5	4	0.900 (0.161)	0.01452 (0.0040)	8.100 (4.534)	1.261 (0.657)	1.497 (0.932)	0.1199 (0.131)	0.130 (0.839)
Mean	**20**	**6**	**0.694 (0.083)**	**0.0075 (0.003)**	**4.132 (2.212)**	**0.439 (0.536)**	**-0.163 (0.472)**	**0.0673 (0.240)**	**0.181 (0.554)**
Laboratory stocks									
Lab Ol pejeta (LP)	14	1	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.0000	0.000 (1.000)	0.0000	0.00 (0.000)
Kiambu Highline (KH)	12	2	0.485 (0.106)	0.01043 (0.0023)	5.8182 (2.9926)	9.2418 (0.998)	1.9499 (0.984)	**0.4701 (0.0000)**	0.7355 (0.924)
Kiambu unselected line (KU)	10	1	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.0000	0.000 (1.000)	0.0000	0.000 (0.000)
Muguga infected (MF)[i]	12	3	0.545 (0.144)	0.01162 (0.00268)	6.4849 (3.300)	6.887 (0.995)	1.701 (0.978)	**0.4418 (0.000)**	0.449 (0.957)
Muguga unselected (MU)[i]	12	4	0.682 (0.102)	0.0092 (0.00312)	5.1364 (2.677)	3.752 (0.953)	0.1463 (0.570)	0.1545 (0.103)	0.3693 (0.700)
Muguga lowline (ML)[i]	11	1	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.0000	0.000 (1.000)	0.0000	0.000 (0.000)
Uganda (UG)	12	1	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.0000	0.000 (1.000)	0.0000	0.000 (0.000)
South Africa Natal province (SAN)	5	1	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.0000	0.000 (1.000)	0.0000	0.000 (0.000)
South Africa Lab stock (SAL)	5	1	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.0000	0.000 (1.000)	0.0000	0.000 (0.000)
Zambia Sothern province (ZS)	8	1	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.0000	0.000 (1.000)	0.0000	0.000 (0.000)
Zambia Eastern province (ZE)	12	1	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.0000	0.000 (1.000)	0.0000	0.000 (0.000)
Zimbabwe West Mashonaland (ZM)	12	1	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)	0.0000	0.000 (1.000)	0.0000	0.000 (0.000)
Mean	**10**	**2**	**0.143 (0.029)**	**0.0026 (0.001)**	**1.45 (0.747)**	**6.6269 (0.982)**	**0.316 (0.961)**	**0.0889 (0.034)**	**0.129 (0.215)**

Field (Wildlife)									
Field Ol Pejeta (FP)	23	10	0.830 (0.067)	0.0054 (0.0017)	2.9881 (1.620)	-2.2583 (0.109)	**-1.5363 (0.048)**	0.0039 (0.792)	0.0154 (0.983)
Maasai Mara (MA)	22	6	0.788 (0.050)	0.01022 (0.00218)	5.7013 (2.839)	3.618 (0.925)	1.086 (0.889)	0.0691 (0.304)	0.0761 (0.556)
Nairobi National Park (NB)	21	6	0.729 (0.065)	0.00450 (0.00188)	2.5095 (1.408)	0.468 (0.640)	**-1.600 (0.038)**	0.0246 (0.279)	0.079 (0.521)
Bomet (BO)	23	6	0.719 (0.074)	0.00989 (0.00244)	5.5178 (2.7532)	3.6284 (0.936)	0.4785 (0.7160)	0.1003 (0.168)	0.2270 (0.090)
Mean	**22**	**7**	**0.767 (0.064)**	**0.0075 (0.002)**	**4.179 (2.155)**	**1.364(0.652)**	**-0.3930 (0.423)**	**0.0495 (0.38575)**	**0.099 (0.537)**
Other rhipicephaline spp.									
R. evertsi (RE)	12	2	0.409 (0.133)	0.00147 (0.00048)	0.8182 (0.629)	1.961 (0.809)	0.687 (0.784)	0.3347 (0.0000)	0.6834 (0.924)
R. praetextatus 1(RP)	5	3	0.700 (0.054)	0.00358 (0.002)	2.000 (1.343)	0.6435 (0.595)	-1.1239 (0.080)	0.0939 (0.380)	0.2900 (0.420)
R. pravus (RV)	16	6	0.833 (0.127)	0.00687 (0.00260)	3.833 (2.125)	-0.596 (0.324)	-1.473 (0.09)	0.0759 (0.366)	0.1466 (0.475)
R. pulchellus (RL)	11	1	0.000 (0.000)	0.0000 (0.0000)	0.000 (0.000)	0.0000	0.000 (0.000)	0.0000	0.000 (0.000)
R. zambezienis (RZ)	12	2	0.545 (0.062)	0.01564 (0.00176)	8.727 (4.335)	11.820 (1.00)	2.794 (1.000)	0.5950 (0.000)	0.802 (0.922)
Unidentified species from Ruma (RUM)	25	11	0.857 (0.059)	0.00467 (0.00113)	2.606 (1.610)	-1.807 (0.160)	-1.0846 (0.145)	0.0142 (0.675)	0.043 (0.815)
Mean	**12**	**4**	**0.595 (0.074)**	**0.012 (0.002)**	**2.58 (3.32)**	**3.072 (0.650)**	**0.496 (0.501)**	**0.1795 (0.307)**	**0.284 (0.675)**

[a]Refers to the source population or origin of the tick stock. Tick populations were grouped on the basis of the source of the sequences as field ticks (collected from areas grazed by cattle), laboratory stocks maintained at ILRI Tick unit, wildlife ticks (collected from areas grazed by wildlife) and other rhipicephaline species. The number of individual sequences analysed (sample size) and the corresponding number of haplotypes for each population is listed. The detailed description of these populations is given in Chapter 3 Table 6. Tajima's D was negative and statistically significant in RU, FP and NB but positive and not significant for all the other tick populations except the nine laboratory stocks that did not show diversity. KH had a high Fu's FS (9.2418). SSD was statistically significant in KH, MF, RE and RZ.

Table 15: *R. appendiculatus* COI haplotypes showing variable sites and the number sequences in each haplotype

POSITION[a]	17	44	71	77	101	113	149	155	170	176	235	236	251	257	269	281	308	311	322	323	343	347	359	398	404	428	467	521	530	550	N*
Hap_1	T	A	C	A	A	T	G	A	G	C	G	C	A	G	A	T	A	A	G	T	G	T	A	T	A	T	A	T	A	A	94
Hap_2	.	C	T	.	G	C	.	.	A	.	.	.	.	A	G	.	.	.	A	.	.	.	.	.	G	C	G	.	C	G	8
Hap_3	.	C	T	.	G	C	.	.	A	T	.	.	G	A	G	.	.	.	A	.	.	.	.	.	G	.	G	.	T	G	6
Hap_4	.	C	T	.	G	C	.	.	A	.	.	.	.	A	G	.	.	.	A	.	.	.	.	.	G	.	G	.	C	G	107
Hap_5	.	C	T	.	G	C	.	.	A	T	.	.	.	A	G	.	.	.	A	.	.	.	.	.	G	.	G	.	T	G	28
Hap_6	.	C	T	.	G	C	.	.	A	.	.	.	.	A	G	.	.	.	A	.	.	.	.	.	G	.	G	.	T	G	5
Hap_7	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	21
Hap_8	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	12
Hap_9	.	C	T	.	G	C	.	.	A	.	.	.	.	A	G	.	.	.	A	.	.	.	.	.	.	.	G	.	C	G	6
Hap_10	.	C	T	.	G	C	.	.	A	T	.	.	.	A	G	.	.	.	A	C	.	.	.	.	G	.	G	.	T	G	14
Hap_11	.	T	T	.	G	C	.	.	A	.	.	.	.	A	G	.	.	.	A	.	.	.	.	.	G	.	G	.	C	G	4
Hap_12	.	T	T	.	G	C	.	.	A	.	.	.	.	A	G	.	.	.	A	.	A	.	.	.	T	.	G	.	C	G	1
Hap_13	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	C	.	.	1
Hap_14	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	4
Hap_15	.	C	T	.	G	C	.	.	A	T	.	.	.	A	G	.	.	.	A	.	.	.	.	.	.	.	G	.	T	G	1
Hap_16	.	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	2
Hap_17	.	C	T	.	G	C	.	.	A	.	.	.	.	A	G	.	.	.	A	.	.	.	.	C	.	.	G	.	C	G	1
Hap_18	.	T	T	.	G	C	.	.	A	.	.	.	.	A	G	.	G	.	A	.	.	.	.	.	G	.	G	.	C	G	1
Hap_19	.	T	T	.	G	C	.	.	A	.	.	T	.	A	G	.	.	.	A	.	.	.	.	.	G	.	G	.	T	G	4
Hap_20	.	C	T	.	G	C	.	.	A	T	.	.	G	A	G	.	.	G	A	.	.	.	.	.	G	.	G	.	T	G	1
Hap_21	.	.	.	G	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	4
Hap_22	.	C	T	.	G	C	.	.	A	.	.	.	.	A	G	C	.	.	A	.	.	.	.	.	G	.	G	.	C	G	1
Hap_23	.	C	T	.	G	C	T	.	A	.	.	.	.	A	G	.	.	.	A	.	.	.	.	.	G	.	G	.	T	G	1
Hap_24	.	C	T	.	G	C	.	.	A	.	.	T	.	A	G	.	.	.	A	.	.	.	.	.	.	.	G	.	C	G	1
Hap_25	.	C	T	.	G	C	.	.	A	T	.	T	.	A	G	.	.	.	A	.	.	.	.	.	G	.	G	.	T	G	1
Hap_26	.	T	T	.	G	C	.	.	A	T	.	T	.	A	G	.	.	.	A	.	.	.	.	.	G	.	G	.	T	G	1
Hap_27	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	1
Hap_28	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.	.	.	.	.	1

Sequence variation of 28 COI haplotypes derived from 332 *R. appendiculatus* ticks. There were a total of 30 polymorphic sites. [a]Represents the nucleotide position where variation occurred. *N; Represents the number of individuals sharing the same haplotype. Dots (.) denote identity between sequences.

Table 16: Distribution of tick samples from different populations in the four major haplotypes

Population Name	**Number of sequences in Hap_4**	**Number of sequences in Hap_1**	**Number of sequences in Hap_5**	**Number of sequences in Hap_7**	**Total Sample size***
NB	8	-	8	1	21
KT	8	4	6	1	29
ML	11	-	-	-	11
MU	6	1	-	-	12
MF	2	8	-	-	12
MA	7	1	2	4	22
RUM2	1	2	1	1	5
LP	14	-	-	-	14
FP	9	-	3	1	23
BO	11	-	-	-	23
KH	8	4	-	-	12
KU	-	10	-	-	10
MK	9	1	6	-	25
ZE	12	-	-	-	12
RU	1	6	-	10	21
ZS	-	8	-	-	8
SAL	-	5	-	-	5
SAN	-	5	-	-	5
ZM	-	12	-	-	12
BU	-	8	2	3	18
KF	-	19	-	-	20
Total	**107**	**94**	**28**	**21**	

*Refers to the total number sequences from each of the populations that were analysed. (–) means no samples from that particular population. Only haplotyes having more than 20 sequences are represented.

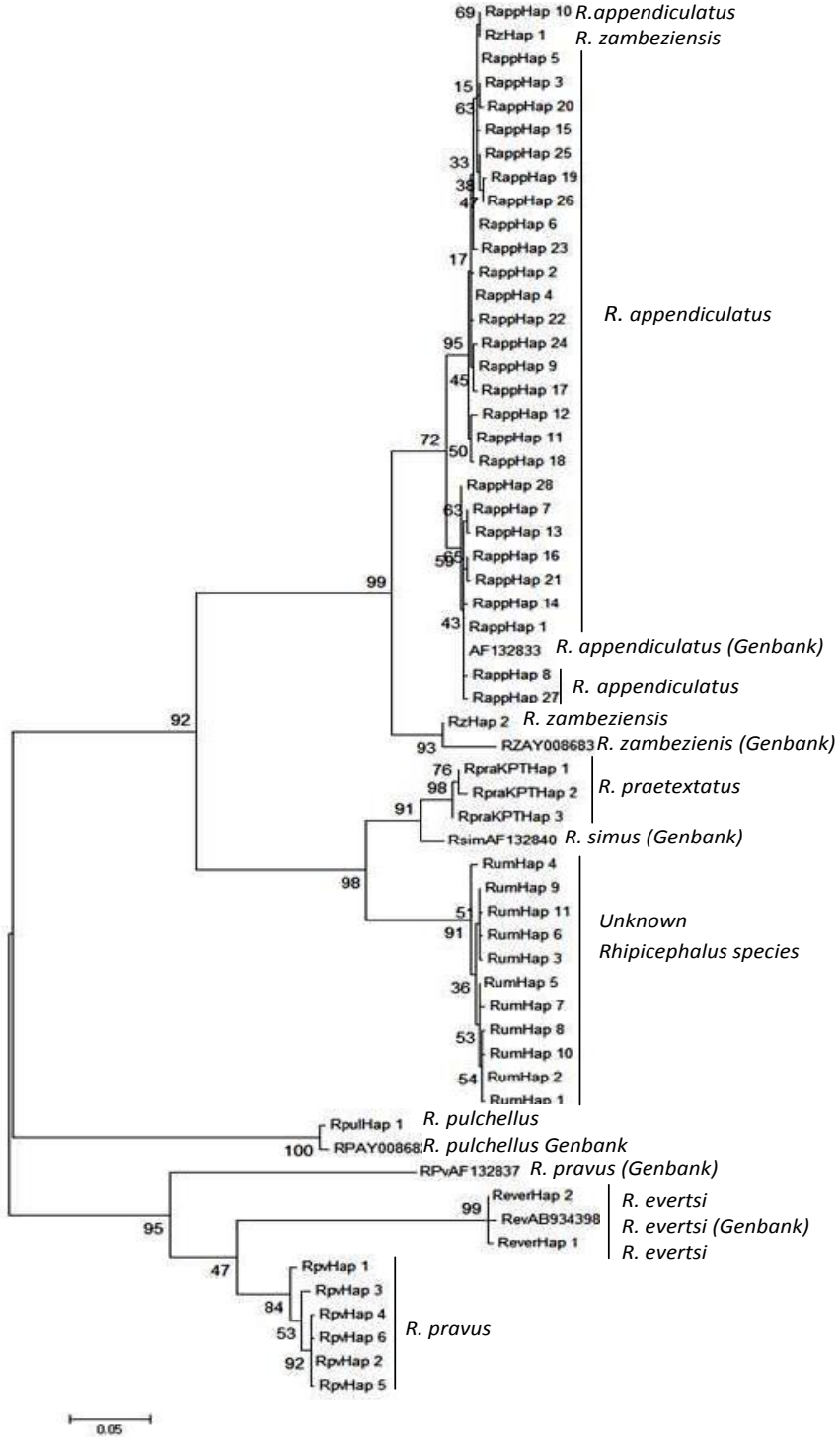


Figure 31: Phylogenetic relationship between all the COI haplotypes belonging to all the rhipicephaline species

The scale represents 0.05 nucleotide substitutions per site. Bootstrap values (1000 replications) are shown. Sequences from Genbank were included in the analysis. *R. appendiculatus* haplotypes clustered in one clade. *R. praetextatus* and *R. simus* appeared closely related to the unknown species from Ruma. *R. pravus* haplotypes obtained from this study were strongly separated from the *R. pravus* reference sequence (RpvAF132837). *R. appendiculatus* haplotypes are represented as RappHap 1-28, *R. zambeziensis* as Rz1-2, *R. praetextatus* as Rpra1-3, *R. pulchellus* as Rpul1, *R. evertsi as* Rev1-2, *R. pravus* as Rpv1-6 and the unknown *Rhipicephalus* from Ruma as Rum1-11.

5.3.4 Phylogeny and median-joining network profiles of *R. appendiculatus* COI haplotypes

To gain insight into the phylogenetic relationship between the 28 COI haplotypes, a phylogenetic tree (NJ) and a median-joining network (MJ) were constructed for the COI gene (Figures 32 and 33 respectively). The COI NJ tree revealed two well-supported groupings of *R. appendiculatus* with a bootstrap value of 100%. Since the sequences in these groups clustered closely, the two groups were treated as different haplogroups and designated haplogroup A and haplogroup B. Haplogroup A with a total of 19 haplotypes including one of the major haplotype (containing 107 individuals) further separated into two groups with one group being designated as sub-group A (8 haplotypes) and the other having 11 haplotypes. Haplogroup B had 9 haplotypes including one of the other major haplotype (containing 94 individuals). From the tree, haplogroup B was observed to be the ancestor of haplogroup A. The unidentified *Rhipicephalus* species clustered closely with *R. praetextatus* with a bootstrap value of 98%.

MJ also revealed two major groups also referred to as haplogroup A and B represented by the same number of haplotypes as in the NJ tree. Nineteen haplotypes (Hap_2, 3, 4, 5, 6, 9, 10, 11, 12, 15, 17, 18, 19, 20, 22, 23, 24, 25, 26) were grouped in haplogroup A. Nine haplotypes (Hap_1, 7, 8, 13, 14, 16, 21, 27 and 28) were grouped in haplogroup B. The links between the two haplogroups were well resolved. Haplogroup A was separated from subgroup A by two mutations while it was separated from B by twelve mutations. Haplogroup A was connected to B by two median vectors (mv1 and mv3) while median mv1 connects the minor subgroup to A (Figure 33). The median vectors may represent either un-sampled haplotypes, haplotypes never introduced to the region or haplotypes that were introduced but became extinct. A star like pattern radiating from haplotypes H_1, H_4 and H_5, which are central to haplogroup A, subgroup A and B respectively could represent common

ancenstral haplotype. From this analysis, haplogroup A was considered to have diverged from haplogroup B supporting the phylogeny observed with the NJ tree. The network represented a population structure that is not related to the geographical origin of the tick samples.

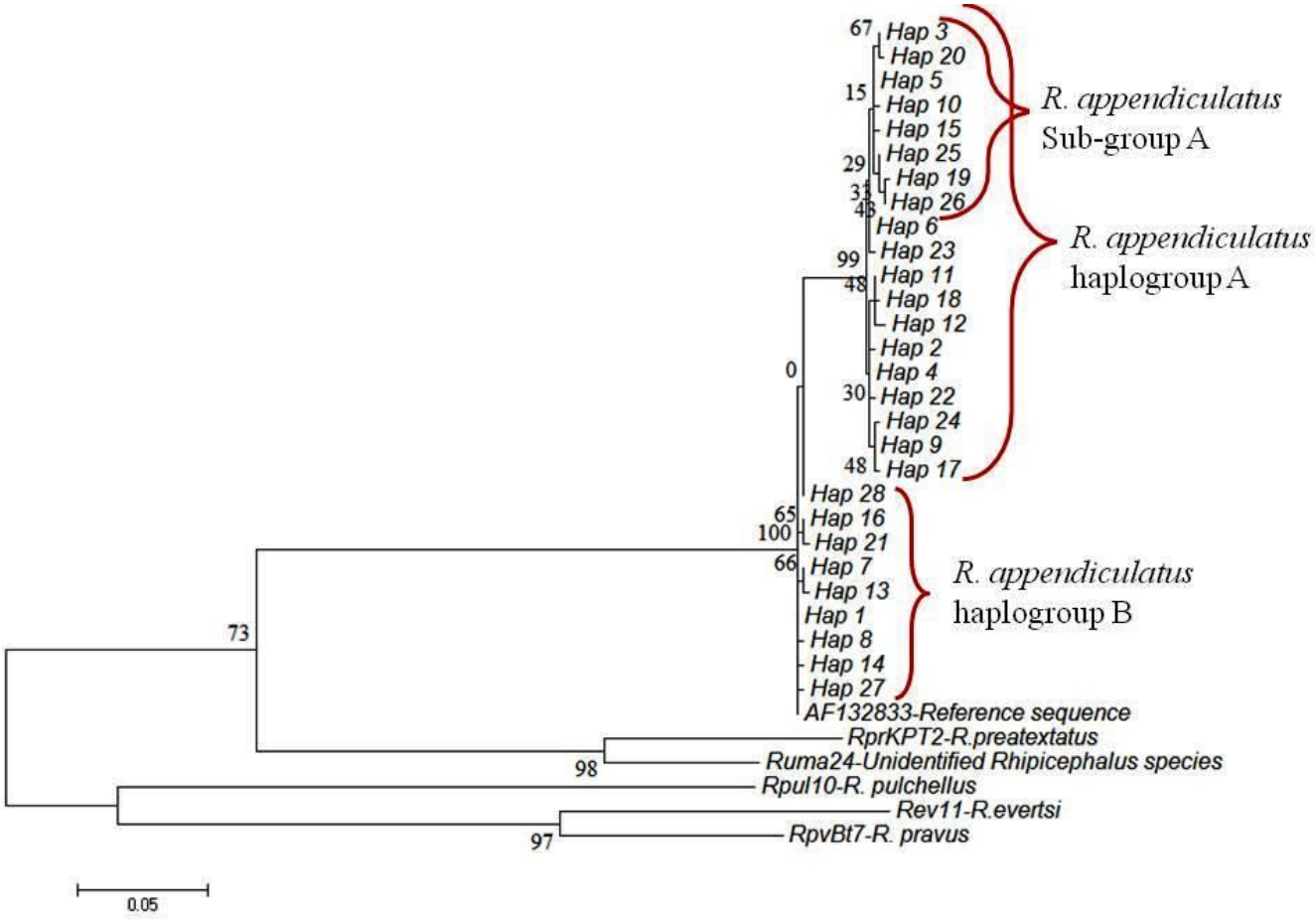


Figure 32: Molecular phylogenetic analysis of *R. appendiculatus* COI haplotype sequences by Maximum Likelihood (ML) method

The 28 haplotypes of *R. appendiculatus* are represented by the names Hap 1-28. The scale represents 0.05 nucleotide substitutions per site. Bootstrap values (1000 replications) are shown. Closely related *R. appendiculatus* haplotypes that clustered in a group supported by high bootstrap values were regarded as a haplogroup. One sequence, AF132833 from Genbank was included in the analysis. Sequences of other rhipicephaline species used in this study were included to help root the tree. These included one sequence from ticks identified as *R. praetextatus* (RprKPT2), *R. pulchellus (Rpul10), R. evertsi (Rev11), R. pravus (RpvBt7)* and one sequence from the unidentified *Rhipicephalus* species (Ruma24).

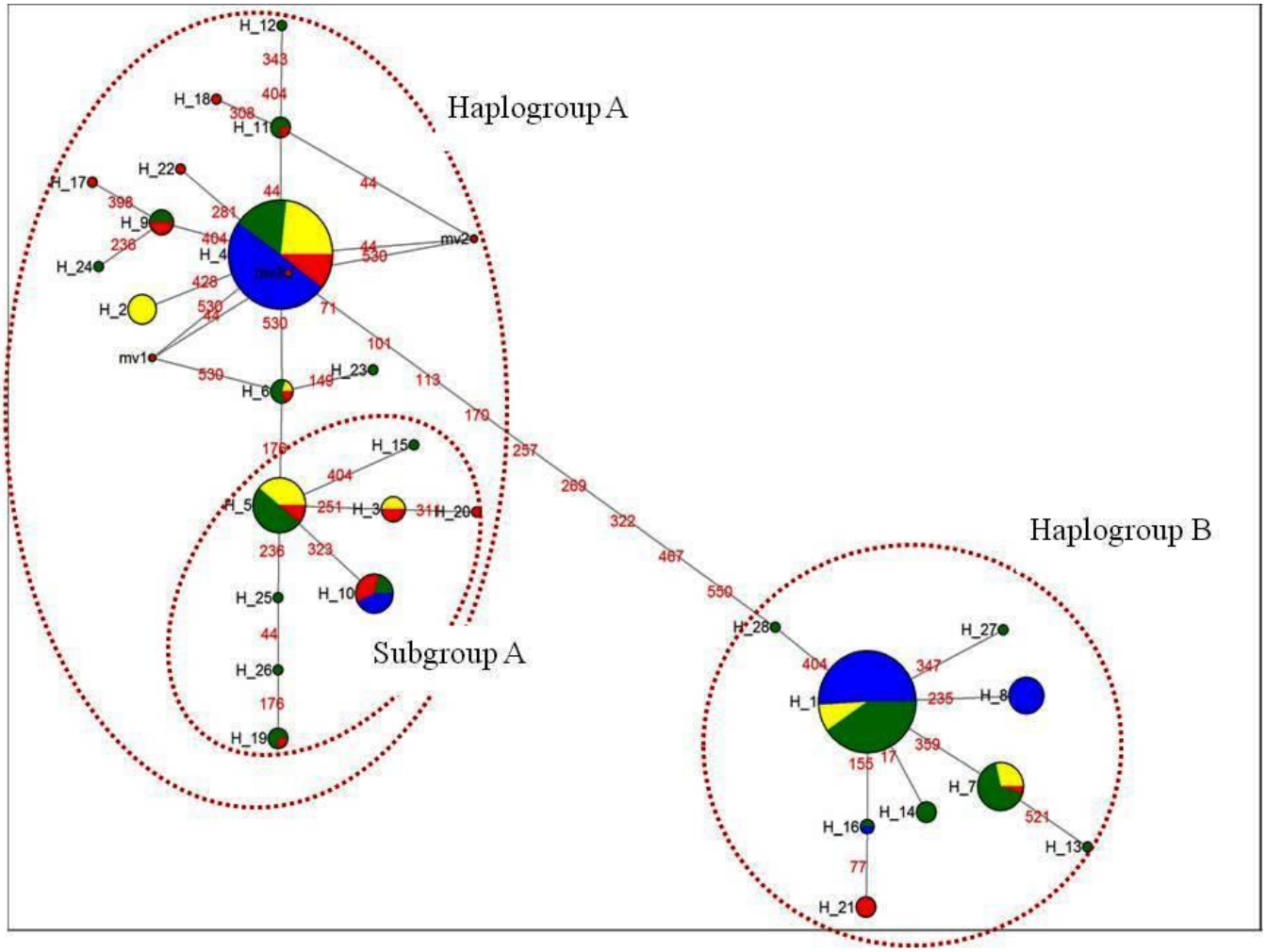


Figure 33: Median-Joining network of 28 haplotypes observed in 332 *R. appendiculatus* ticks
The network was based on the polymorphic sites in the 558 bp COI gene segment. Each circle represents one haplotype and the area of the circle is proportional to the haplotype frequencies. Numbers represent nucleotide position. Colours represent a group of tick populations classified on the basis of the source of the sequences. Blue represents laboratory populations, yellow represents populations sampled from pastures grazed by wildlife, red represents populations sampled from pastures grazed by both cattle and wildlife while green represents populations sampled from cattle pastures. Median vectors are represented by ‘‘mv’’.

5.3.5 Population genetic variability and structure based on the COI gene

5.3.5.1 Analysis of molecular variance (AMOVA)

Population genetic structure was estimated from the hierarchical analysis of molecular variance (AMOVA) based on the frequencies of the haplotypes observed in each group. The level of genetic variation present was assessed in five different clusters of ticks as follows (i) the overall dataset assuming no groups; (ii) between *R. appendiculatus* vs. other rhipicephaline species (iii) between tick populations from cattle vs. wildlife ticks; (iv) between field vs. laboratory populations; (v) between three groups of tick populations (cattle vs. cattle-wildlife vs. wildlife populations) (Table 17). For the overall dataset (22 *R. appendiculatus* populations), level of genetic variation among populations was 85.96% and 14.0% within population while 62.5% variation was observed between *R. appendiculatus* and other rhipicephaline species. Genetic variation between tick populations from cattle and those from wildlife was 14.9% among the two groups and 35.38% among populations within the two groups. The variation was higher at 49.7% within tick populations in these two groups when analysed as one group. The lowest level of genetic variation (-1.43%) occurred between field and laboratory *R. appendiculatus* groups. However, the genetic variation was 57.3% among different populations within these two groups of ticks compared to 44.15% within populations when analysed as one group. Field populations of *R. appendiculatus* grouped on the basis of area of collection i.e. cattle, cattle-wildlife, and wildlife showed a variation of 4.94% among the three groups while the level of genetic variation within the three populations rose to 42.69%. When tick populations in three groups were analysed as a single group, the genetic variation rose to 52.37%. This demonstrates low genetic differences between populations of *R. appendiculatus* originating from different animal host species. Analysis of COI revealed two major haplogroups of *R. appendiculatus*. Analysis of genetic variation between these two haplogroups was

found to be significantly high (90.8%) (Table 17) in addition to a high genetic differentiation index (F_{ST}) of 0.907 (P= 0.000) meaning that the two groups were genetically differentiated.

Table 17: Global analysis of molecular variance (AMOVA) for different groups of ticks at different hierarchical levels

Clusters	**Hierarchy**	**Variance components**	**Percentage variation**
Overall	1	Among populations	85.96
		Within populations	14.04
R. appendiculatus vs. other rhipicephaline species	2	Among groups	62.50
		Among populations within groups	29.50
		Within populations	8.00
Cattle vs. wildlife *R. appendiculatus* populations	2	Among groups	14.91
		Among populations within groups	35.38
		Within populations	49.71
Field vs. laboratory *R. appendiculatus* populations	2	Among groups	-1.43
		Among populations within groups	57.28
		Within populations	44.15
Cattle vs. Cattle-Wildlife vs. Wildlife *R. appendiculatus* populations	3	Among groups	4.94
		Among populations within groups	42.69
		Within populations	52.37
Haplogroup A vs. haplogroup B (between the two major *R. appendiculatus* haplogroups)	2	Among populations	90.8
		Within populations	9.3

Clusters are based on a priori groupings of sampling localities. Cattle vs. wildlife *R. appendiculatus* populations refer to ticks collected from areas grazed by cattle versus those collected areas grazed by wildlife. Field vs. laboratory *R. appendiculatus* populations refers to all *R. appendiculatus* ticks collected from field localities versus laboratory *R. appendiculatus* while Haplogroup A vs. haplogroup B was between the two major *R. appendiculatus* haplogroups identified by MJ network.

5.3.5.2 Mismatch distribution analysis

To understand the historical dynamics of the study populations and processes that could have caused the observed genetic variation, the frequency distribution of the number of mutational differences between haplotypes (mismatch distribution) was evaluated across the 28 *R. appendiculatus* haplotypes to test for population expansion or contraction. In an expanding population, mismatch distribution is expected to be a bell shaped smooth curve, whereas neutrally evolving, stable populations tend to have multimodal, ragged mismatch distributions. Populations undergoing sudden population expansions or exponential growth are expected to produce smooth and unimodal mismatch distributions (Slatkin & Hudson 1991; Rogers & Harpending 1992; Harpending, 1994; Rogers, 1995). Appearance of multiple peaks suggests a population subdivision. In this study, mismatch distribution analysis was also done separately for the two major haplogroups designated as A and B (Figure 34). The overall mismatch distribution pattern of the 28 haplotypes showed two distinct bimodal peaks and a smooth curve (Figure 34) suggesting two distinct populations of *R. appendiculatus* that seem to be undergoing sudden population expansion independently. These two peaks correspond to the two major haplogroups A and B as defined by the phylogenetic tree and the median-joining network. Mismatch distribution patterns for the two haplogroups analysed independently were smooth and unimodal suggesting that sudden population expansion had occurred.

Arlequin also reports two statistics based on mismatch distributions that can be used to assess the goodness-of-fit sudden (demographic) or spatial population expansion models for the data. These are the sum of squared deviations (SSD) and Harpending's raggedness index (RI) (Harpending 1994). A significant P value of simulated (expected) SSD value ≥ observed SSD value under any

expansion model indicates a good-fit model and thus supports population expansion. A significant RI value ($P \leq 0.05$) indicates deviation from the expansion model and is taken as evidence for rejecting that population expansion model (Schneider & Excoffier, 1999). In this study, SSD was estimated to support the sudden expansion model while RI was estimated to test whether the sequence data from each population deviated from what is expected under a sudden expansion model. Mean SSD was 0.0673 (P= 0.240) in ticks collected from areas grazed by cattle while it was 0.0889 (P (Simulated SSD ≥ Observed SSD) = 0.034) in laboratory stocks (Table 14). The mean SSD was 0.0495 (P (Simulated SSD ≥ Observed SSD) = 0.3857) in ticks collected from areas grazed by wildlife while it was 0.1795 (P (Simulated SSD ≥ Observed SSD) (Simulated SSD ≥ Observed SSD) = 0.307) in other rhipicephaline species. SSD was statistically significant in KH and MF tick populations. In the 332 *R. appendiculatus* COI sequences, the overall SSD was 0.0761 (P (Simulated SSD ≥ Observed SSD) = 0.0340) (Table 18). The overall RI was 0.0610 (P (Simulated RI ≥ Observed RI = 0.06) (Table 18). Mean RI was 0.181 (P (Simulated RI ≥ Observed RI) = 0.554) in ticks collected from areas grazed by cattle while it was 0.129 (P (Simulated RI ≥ Observed RI) = 0.215) in laboratory stocks (Table 14). The mean RI was 0.099 (P (Simulated RI ≥ Observed RI) = 0.537) in ticks collected from areas grazed by wildlife (Table 14). In haplogroup A, the SSD was 0.01263 (P = 0.55) while the RI index was 0.05029 (P = 0.74). For haplogroup B, the SSD was 0.00281 (P = 0.17) and RI was 0.10201 (P = 0.21). Since the overall SSD under the sudden population expansion model was statistically significant (P = 0.034), it showed that there was demographic expansion in these populations supporting the observed mismatch distribution patterns. The SSD and RI for either haplogroup did not differ significantly from the observed (P > 0.05) and thus although the distribution patterns are consistent with sudden demographic expansion growth model, the SSD and RI do not support this model for any of these two haplogroups.

5.3.5.3 Selective neutrality tests

Tajima's D (Tajima, 1989) and Fu's Fs (Fu, 1997) tests were then conducted to determine whether patterns of mitochondrial sequence variation were consistent with predictions of the neutral model. Along with detecting the influence of selection on a gene, these tests can also be potentially informative about the demographic forces that have affected a population or species (Tajima, 1989; Fu, 1997). To test for neutrality or if any of the tick populations were undergoing selection, Tajima's D and FU's Fs statistic were estimated in Arlequin v3.5. The null hypothesis of the Tajima's D test is neutral evolution in an equilibrium population. This implies that no selection is acting at the locus and that the population has not experienced any recent growth or contraction (Tajima, 1989). A positive Tajima's D value indicates an excess of intermediate frequency (polymorphic) alleles signifying an excess of high frequency polymorphisms, indicating a decrease in population size and/or balancing selection while a negative value indicates an excess of rare alleles signifying an excess of low frequency polymorphisms relative to expectation, indicating population size expansion (e.g. after a bottleneck or a selective sweep) and/or purifying selection. Although the Tajima's D statistic is commonly employed as a test for natural selection, this is confounded by the influence of population history. For instance, balancing selection may yield a positive value for Tajima's D. However, demographic processes such as population reduction, population subdivision, a recent bottleneck (Maruyama & Fuerst 1985), or migration can also produce this result. Similarly, linkage to a selective sweep or purifying selection at the studied locus can generate a negative value for Tajima's D. Yet, this pattern can also be due to population growth and a less recent bottleneck (Maruyama & Fuerst, 1985), or migration.

The overall Tajima's D in the 28 *R. appendiculatus* haplotypes was 1.244 (P = 0.907) (Table 19). When the ticks were grouped into the two major haplogroups and compared, Tajima's D for

haplogroup A was -1.65063 (P = 0.017) while it was -1.08714 (P = 0.125) for haplogroup B (Table 19). Although Tajima's D values are negative for both haplogroups, it is only statistically significant in haplogroup A (P = 0.017) suggesting that Haplogroup A is experiencing population expansion more rapidly than haplogroup B. The magnitude of Tajimas D (-1.65063 for A) indicates the rate of population expansion is faster in haplogroup A than in haplogroup B. Though five field *R. appendiculatus* populations (KT, BU, RUM2, MA and BO) and three laboratory stocks (KH, MF and MU) had positive Tajima's D value (Table 14) suggesting a decrease in population size and/or balancing selection, the P values for these populations were not significant. Five field populations (KF, MK, RU, FP and NB) had negative Tajima's D values with significant P values except for KF and MK suggesting an excess of low frequency polymorphisms relative to expectation, indicating population size expansion for these populations. *R. pravus* and ticks from the unidentified *Rhipicephalus* species from Ruma also had negative Tajima's D values while *R. evertsi, R. zambeziensis* and *R. praetextatus* had high positive Tajima's D suggesting indicating a decrease in population size and/or balancing selection.

The overall Fu's FS in the 28 *R. appendiculatus* haplotypes was -0.122 (P = 0.559) (Table 19). When the two major haplogroups were compared, the Fu's FS for haplogroup A was -10.3479 (P =0) while it was -3.46287 (P = 0.057) for haplogroup B. A negative value of FS is evidence for an excess number of alleles, as would be expected from a recent population expansion or from genetic hitchhiking (genetic draft) which results in changes in the frequency of an allele because of linkage with a positively selected allele at another locus. Though the overall Fu's Fs indicated a negative value suggesting a recent population expansion or genetic hitchhiking of the COI gene, it was however not statistically significant (P = 0.559) and thus the tick populations cannot be said to be expanding.

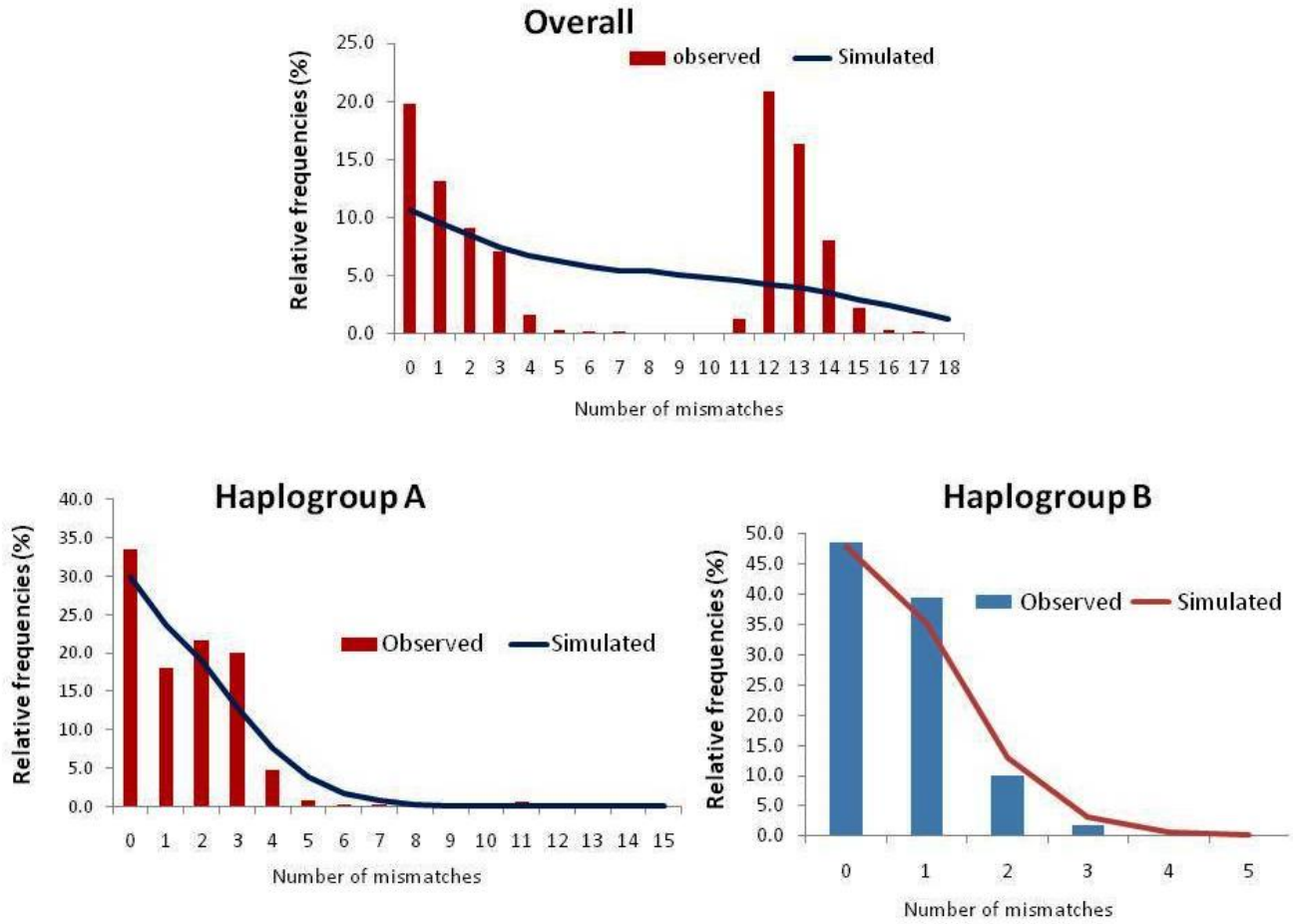


Figure 34: Mismatch distribution patterns for the overall 28 haplotypes, haplogroup A and haplogroup B, for COI gene region generated using 332 COI sequences

The graph shows frequency distributions of the number of pairwise nucleotide differences (mismatch) between the haplotypes. Overall distribution represents differences between the 28 haplotypes. The overall mismatch distribution showed two distinct bimodal populations. Haplogroup A shows nucleotide differences of haplotypes falling within haplogroup A while haplogroup B consists of haplotypes falling within haplogroup B as defined by the MJ network.

The negative value of F_S (-10.3479) (Table 19) for haplogroup A which is statistically significant (P = 0) is evidence for an excess number of alleles from a recent population expansion or from genetic hitchhiking. The magnitude of Fu Fs also indicates that haplogroup A is expanding at a much faster rate than haplogroup B. This result is also supported by the Tajima's D result.

Table 18: Overall *R. appendiculatus* mismatch distribution analysis

Parameters	**Overall**	**Haplogroup A**	**Haplogroup B**
Sum of Squared deviation (SSD)	0.0761	0.0126	0.0028
P(Simulated SSD ≥ Observed SSD)	0.034	0.550	0.17
Harpending's Raggedness index (RI)	0.061	0.05029	0.102
P(Simulated RI ≥ Observed RI)	0.060	0.740	0.210

Table 19: Selective neutrality tests statistics for the two major *R. appendiculatus* haplogroups

Neutrality test	**Statistics**	**Overall**	**Haplogroup A**	**Haplogroup B**
Tajima's D test	**Sample size**	**332**	**193**	**139**
	No. of substitution sites (S)	30	23	7
	Mean no. of pairwise difference (pi)	6.864	1.5693	0.65186
	Tajima's D	1.244	**-1.65063**	**-1.08714**
	Tajima's D p-value	0.907	**0.017**	**0.125**
Fu's FS test	**Sample size**	**28**	**20**	**8**
	Mean no. of pairwise difference (pi)	6.898	1.5693	0.65186
	FS	-0.1222	**-10.3479**	**-3.46287**
	FS p-value	0.559	**0**	**0.057**

5.3.6 Diversity and phylogeny relationships of ITS2 and 12S sequences

Sequencing of 12S and ITS2 was only done for a small number of *R. appendiculatus* from the two major haplogroups, *R. pravus* and *R. preatextatus* ticks to support the COI data. A total of 140 12S sequences were used in the phylogenetic analysis. Of the 140 sequences, 93 sequences were from *R. appendiculatus* distributed among 12 populations (Table 20). Diversity analysis of the 12S sequences revealed a total of five haplotypes consisting of 2 major haplotypes, one with 38 sequences and the other with 52 sequences and three minor haplotypes each containing a singleton. The grouping or distribution of the 12S sequences in the two major haplotypes was consistent with the grouping of the two major COI haplotypes thus supporting the grouping observed with COI gene. A phylogenetic tree constructed for all the 12S haplotype sequences revealed two groups of *R. appendiculatus* (Figure 35). There was a weak support (56%) for the close genetic relationship between *R. turanicus* and *R. appendiculatus*. The tree also showed a strong support for the close genetic relationship (91%) between *R. praetextatus,* the *Ruma* species and *R. simus* from genbank (Murell *et al.,* 1999). These species are also closely related to *R. compositus* and *R. muhsamae.* One *R. pravus* genbank sequence clustered with *R. geigyi* and appears to be distinct from other sequences of *R. pravus* haplotypes. Figure 36 shows some of the homologous and mutation points in the 12S gene across different tick species.

Of the 126 ITS2 sequences analysed, 87 sequences from 12 *R. appendiculatus* populations were confirmed to belong to *R. appendiculatus* (Table 20) while *R. pravus* had 13 sequences and *R. praextextatus* had 26 sequences. An 1149 bp region of the ITS2 was used to determine the number of haplotypes in the 87 *R. appendiculatus* sequences. This resulted in three haplotypes with one major haplotype containing 67 sequences while the other two had 9 and 11 sequences respectively. The distribution of the ITS2 haplotypes did not seem to support the grouping observed with COI

sequences. This finding suggests that the gene is conserved within species and thus can be a good genetic marker for interspecies population studies rather for intraspecies diversity. Phylogenetic analysis of the eight haplotypes (two haplotypes for *R. appendiculatus*, three for *R. pravus* and three for *R. praetextatus* together with five ITS2 sequences available from genbank clustered the *R. appendiculatus* haplotypes in one group. The resulting tree also strongly supported the close genetic relationship between *R. appendiculatus* and *R. zambeziensis* with a bootstrap value of 100% (Figure 37). Insertion/deletion (indels) mutations were found to be common in the gene. Some of the homologous and mutation points in the 861 bp ITS2 gene segment across different tick species are shown in Figure 38.

Table 20: List of the number of ITS2 and 12S sequences analysed from eight field and four laboratory populations of *R. appendiculatus*

***R. appendiculatus* population**	**No. of sequences**	
	12S	**ITS2**
BU	5	9
BO	10	13
FP	6	7
KF	4	8
MK	13	4
MF	8	7
KU	7	9
KT	12	7
ML	8	7
MU	6	5
RU	9	11
RUM2	5	0
Total	**93**	**87**

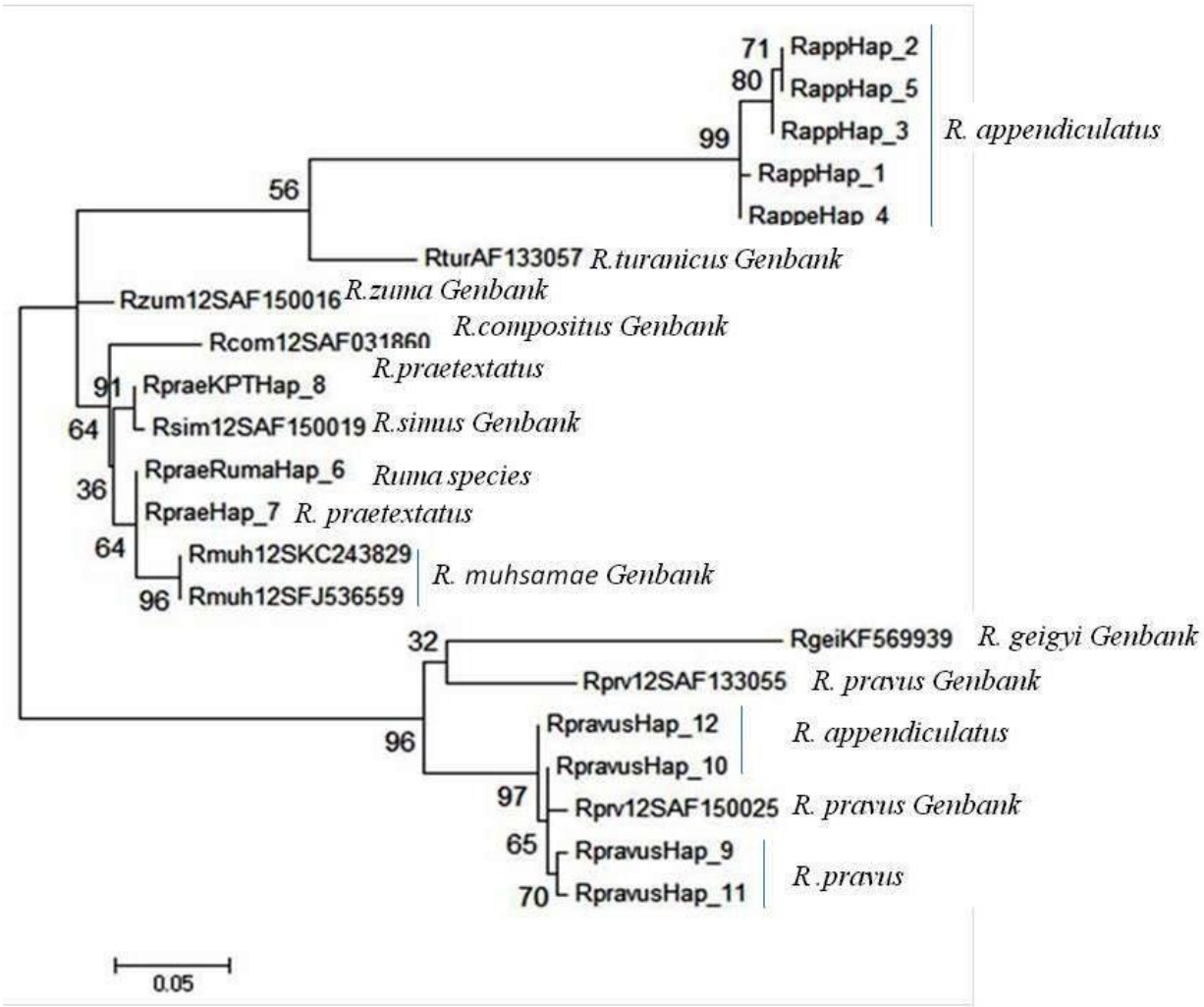


Figure 35: Phylogenetic analysis of all the 12S haplotypes

The analysis involved 21 sequences which included 9 sequences from genbank. *RpravusHap9-12* refers to haplotypes unique to *R. pravus* while *RappH*ap1-5 refers to haplotypes unique to *R. appendiculatus* while *RpraeRumaHap6, RpraeHap7* and *RpraeKPTHap8* refers to haplotypes observed with *R. praetextatus* ticks two of which were from Kapiti (KPT) and Ruma sites respectively. The tree revealed two clusters of *R. appendiculatus* and strongly supported the close genetic relationship (91%) between *R. praetextatus,* the *Ruma* species and *R. simus*. These species are also closely related to *R. compositus* and *R. muhsamae.* One *R. pravus* genbank sequence clusters with *R. geigyi* and appears to be distinct from other sequences of *R. pravus* haplotypes.

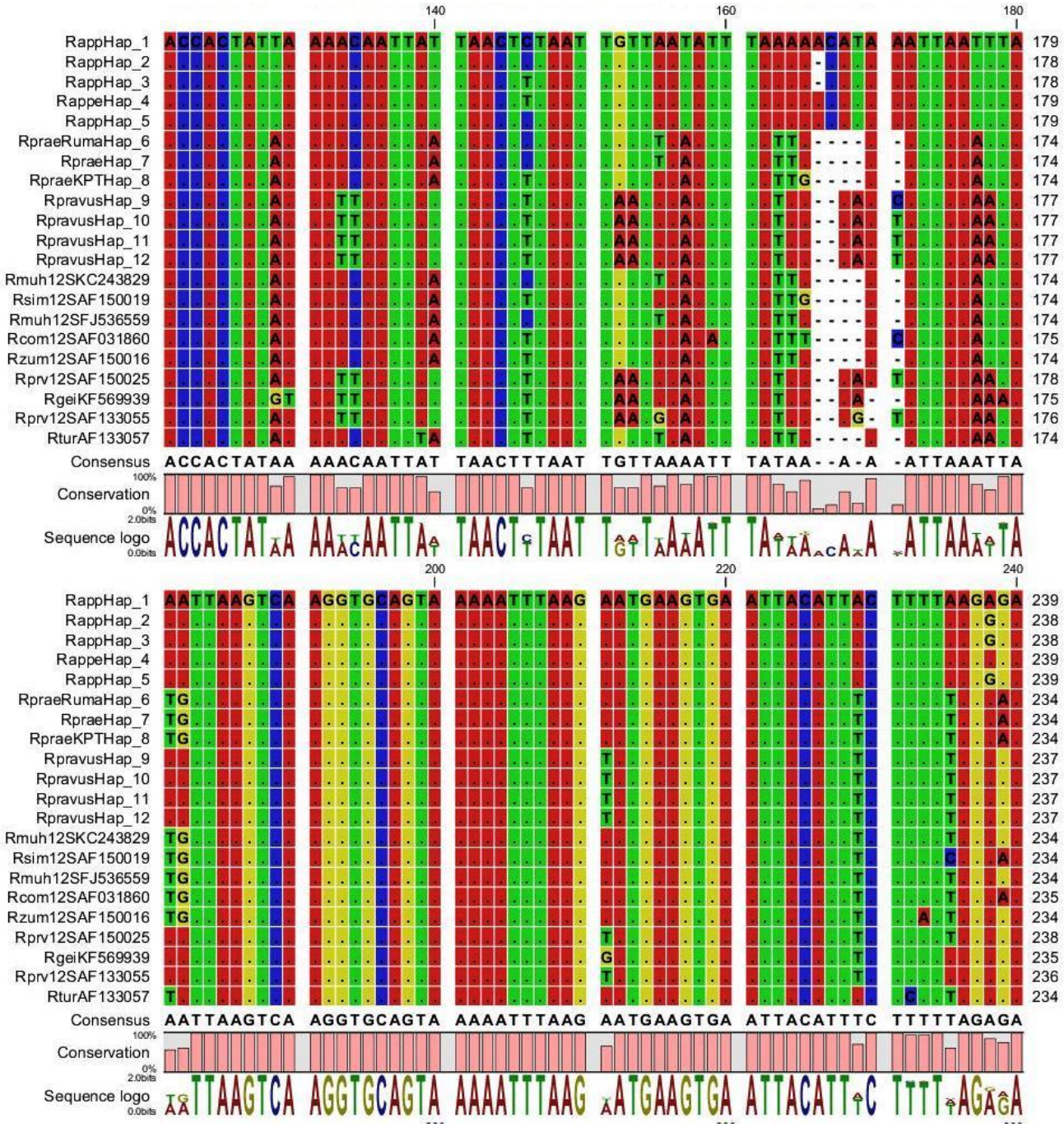


Figure 36: A section of a multiple sequence alignment of the 12S gene showing some variable sites within the gene

The alignment was generated using 5 *R. appendiculatus*, 3 *R. praetextatus*, 4 *R. pravus* haplotypes and 9 genbank sequences. Identical nucleotides representing an exact match of nucleotides in the aligned sequences are labeled in the same colour. (-) denotes a gap which represents a nucleotide inserted or deleted from one or the other sequence producing indel mutations. A consensus sequence generated by the software appears at the bottom of the alignment.

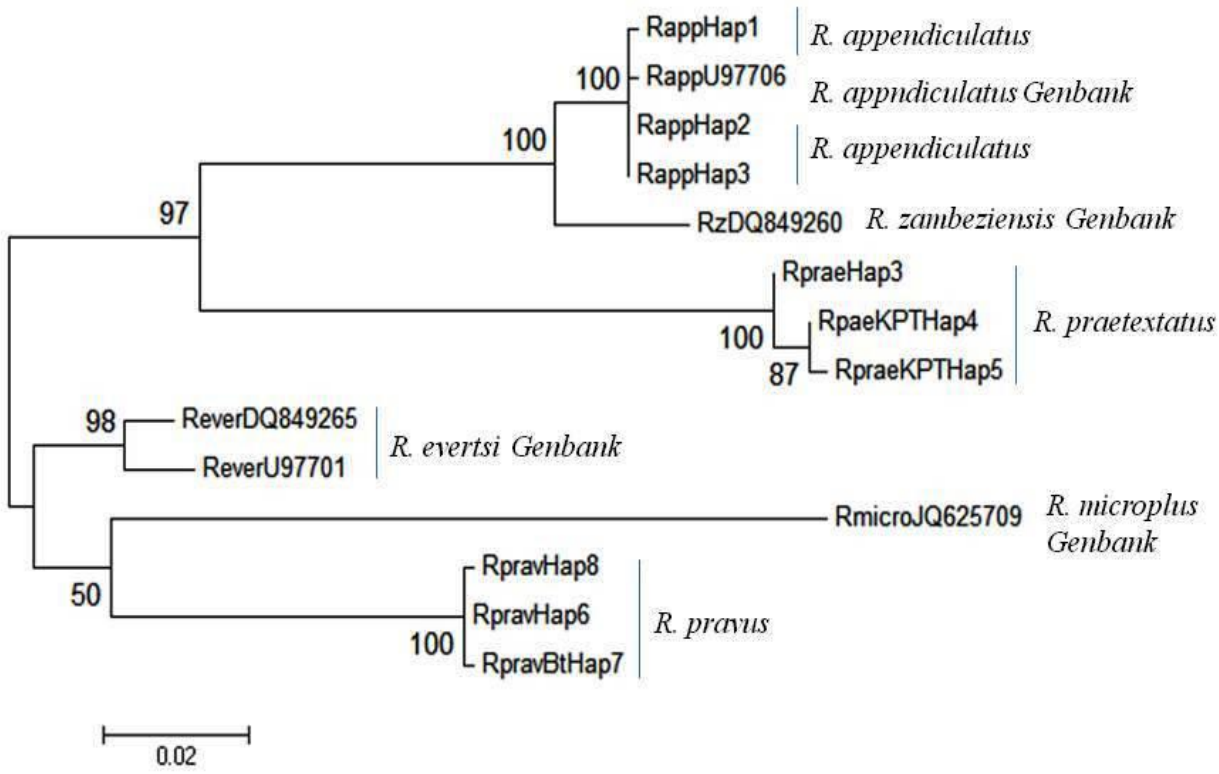


Figure 37: Phylogenetic analysis of all the ITS2 haplotypes sequences
The analysis involved 8 haplotype sequences and 5 sequences from genbank. *RpravHa6-8* refers to haplotypes unique to *R. pravus* while *RappHap 1-3* refers *to* haplotypes unique to *R. appendiculatus* while *RpraeKPTHap4* and *RpraeKPTHap5* refers to haplotypes unique to *R. praetextatus* ticks which were from Kapiti (KPT). The tree strongly supported the close genetic relationship (100 %) between *R. appendiculatus* and *R. zambeziensis* which are closely related to *R. praetextatus* ticks. *R. pravus* associated closely with *R. microplus*.

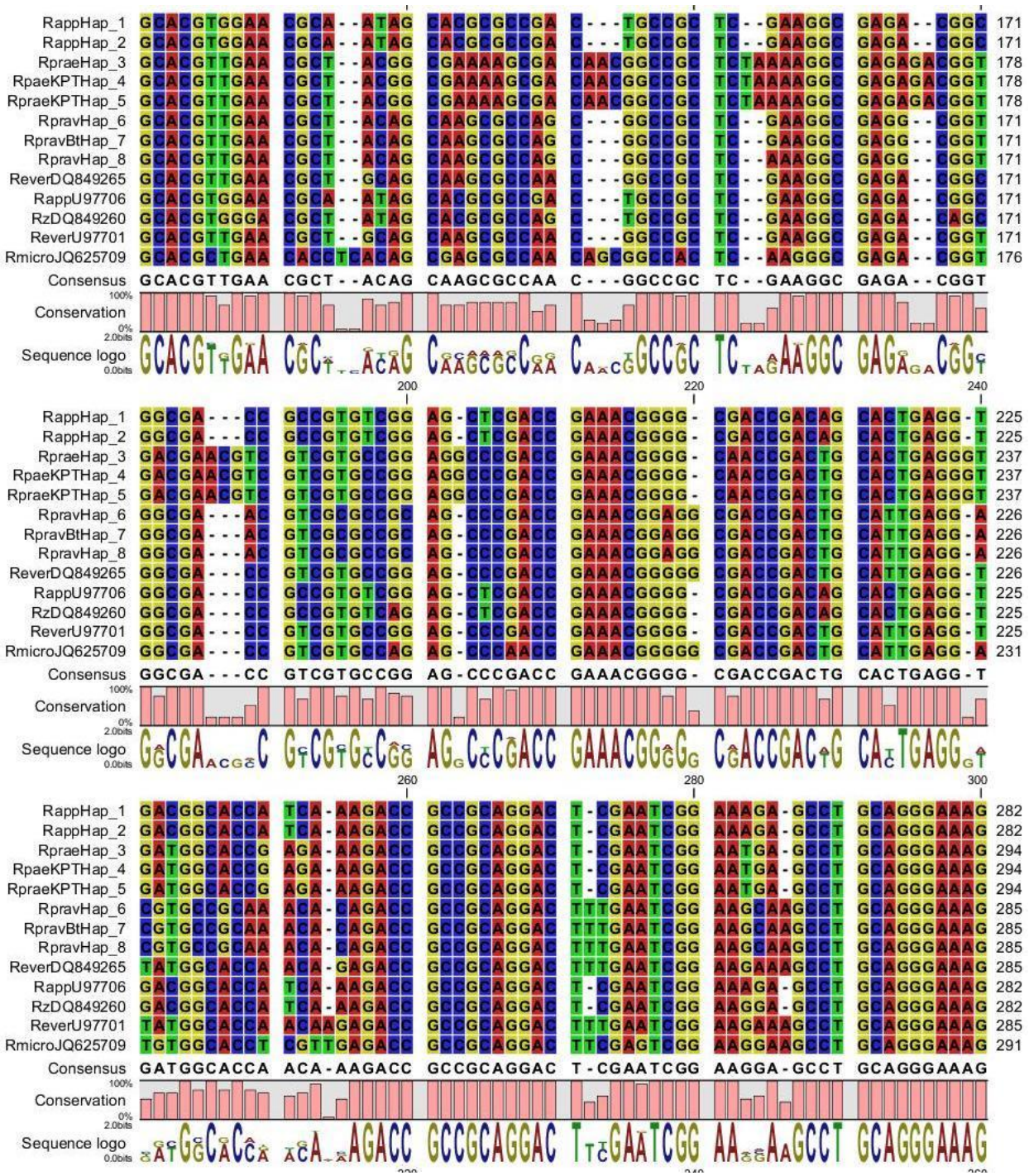


Figure 38: A section of a multiple sequence alignment of the ITS2 gene showing some variable sites within the gene.

The alignment was generated using 3 *R. appendiculatus*, 2 *R. praetextatus*, 3 *R. pravus* haplotypes and five genbank sequences. Identical nucleotides representing an exact match in the aligned sequences are labeled in the same colour. (-) denotes a gap which represents a nucleotide inserted or deleted from one or the other sequence producing indel mutations. A consensus sequence generated by the software appears at the bottom of the alignment.

5.4 Discussion

This study further assessed the phylogenetic relationship between Kenyan field and laboratory *R. appendiculatus* and other rhipicephaline tick species by examining the DNA sequence variation in the mitochondrial COI gene supported by data from ITS2 and 12S genes. COI is the most widely used marker for DNA barcoding. COI barcoding has been extensively used to confirm species identity and discriminate closely related taxa (Hebert *et al.,* 2004). This study used COI to confirm species identity and assess the phylogenetic relationship within *R. appendiculatus* and the rhipicephaline family. The evolution of COI gene is known to be rapid enough to allow the discrimination of not only closely-related species, but also phylogeographic groups within a single species (Cox & Hebert, 2001; Wares & Cunningham, 2001).

The usefulness of COI as a phylogenetic marker for *R. appendiculatus* had been demonstrated earlier (Murell *et al.,* 2000); Mtambo *et al.,* 2007a; 2007b). In the present study, COI gene conservation within *R. appendiculatus* was found to be high as revealed by a sequence identity that ranged between 97-100 %. Although highly conserved within species, a total of 28 haplotypes were obtained from the 22 *R. appendiculatus* tick populations analysed in this study (Table 14) indicating high intra-species diversity. The gene also showed high inter-species diversity and was found to be a powerful genetic marker for discrimination of closely-related *Rhipicephalus* tick species. It was able to clearly discriminate *R. appendiculatus* from other rhipicephaline species analysed (Figure 31). COI sequence identity ranged from an average of 87% between *R. appendiculatus* and *R. praetextatus,* 85% between *R. appendiculatus* and *R. pravus*, 84% between *R. appendiculatus* and *R. evertsi,* 86% between *R. appendiculatus* and the unidentified *Rhipicephalus* species and an average of 86% between *R. appendiculatus* and *R. pulchellus*. Thus, COI variation within the *R.*

appendiculatus species was observed to be 3% or less while among species, the variation was observed to be about 16%.

Though the COI gene was able to confirm the species identity of ticks belonging to *R. appendiculatus, R. pulchellus and R. evertsi* species, the identity of other species could not be resolved. For some tick species, morphological identity and COI characterization were not congruent. For example, 4 out 8 sequences of *R. zambezienis* clustered and matched with high identity to *R. appendiculatus* sequences (Figure 32). Since the *R. zambeziensis* sequences were from a laboratory stock, it is possible that there was a mix up of tick stocks during growth and maintenance. Thus, COI could be used to detect a possible mix up of samples. COI could not also confirm the species identity of ticks morphologically identified as *R. praetextatus, R. pravus* and the unknown species from Ruma. No sequences matching with a high identity were found in genbank or the Barcode of Life (BOLD) databases. For *R. praetextatus*, the sequence closely matched to *R. simus* in genbank database while for *R. pravus* one genbank sequence (AF132832; Murell *et al.*, 2001) was available but failed to match with high identity (Figure 32) which could indicate species mis-identification. Molecular identification of species is based on homology searches against sequences deposited in major databases such as genbank. High identities imply genetic similarities as levels of divergence among individuals are usually much lower within the same species than between closely related species (Tautz *et al.,* 2002). When highly identical sequences are missing from the databases, it is difficult to conclusively identify the species. It is possible that the COI sequences of some the ticks analysed are not available in the databases. Thus, some of the non-identical tick sequences could not be accurately resolved to the species level. Indeed, one of the concerns raised regarding molecular identifications is that non-identical sequences may remain unidentifiable or may be unambiguously wrongly placed (Will & Rubinoff, 2004). At the same

time, public nucleotide databases are known to sometimes contain information of misidentified species, inaccurate taxonomies and sequences showing errors or obtained from contaminated samples (Shen *et al.,* 2013). Although the intent of DNA barcoding is to assign unknown individuals to species (Moritiz & Cicero, 2004), for some of the tick species studied, COI could not conclusively confirm their identity. Thus, DNA-based identification system can only be useful after morphological species identification on the basis of the work of taxonomists and biologists. DNA sequences can then be used to augment taxonomy and species delimitation as either corroborating evidence for existing hypotheses or as starting points for further testing by other means (DeSalle *et al.,*. 2005).

Within the rhipicephaline family and based on the COI, *R. appendiculatus* was found to be closely related to *R. zambezienis* (Figure 32) as reported by Murell *et al.,* (2001). Morphological differentiation of the two species is usually difficult because they share a lot of similar traits (Norval *et al.,* 1992). *R. zambeziensis* occurs in southern Africa and is also a vector of *T. parva.* However, using ITS2 and 12S genes, the two have been shown to be distinct species (Mtambo *et al.,* 2007). *R. praetextatus* was found to be closely related to *R. simus.* Based on COI, there was also a strong support for the closer genetic relationship between the unknown *Rhipicephalus* species from Ruma and *R. praetextatus*, and *R. simus* (Figre 32). *R. pravus* haplotypes observed in this study were strongly separated from the *R. pravus* reference sequence (RpvAF132837) from the databases suggesting that they possibly represent two distinct species (Figure 32). However, the haplotypes were found to be closely related to *R. evertsi* compared to other species, an observation made earlier by Murell *et al.* (2001).

This study also analysed the phylogenetic relationship between the *R. appendiculatus* populations using the ITS2 and the 12S gene (Figure 35 and 37) to support the results obtained using the COI data. The use of multiple genetic markers targeting different regions of the genome has been recommended to overcome drawbacks associated with one marker as well as to increase the accuracy in the construction of a phylogenetic history of species (Gadagkar *et al*., 2005; Balmer *et al*., 2011). COI identified a total of 10 haplotypes having more than 5 sequences (Table 14 and 16) while 12S identified two major haplotyes and two singletons from 93 *R. appendiculatus* sequences and ITS2 identified one major haplotype with two minor haplotypes from 87 *R. appendiculatus* sequences. The two major haplotypes revealed by 12S (Figure 35) correlated well with those observed using COI indicating a high level of congruence between the two genes. COI and 12S had been successfully used previously to show that ticks from eastern and southern Zambia constitute two genetically differentiated groups while ITS2 failed to discriminate between these populations (Mtambo *et al.*, 2007). The failure of ITS2 to resolve geographic structuring in *A. habraeum* and *H. rufipes* has also been reported (Cangi *et al.*, 2013). These findings prove that COI and 12S regions are better markers for studying intra-species diversity of closely related tick species while ITS2 region may be useful for discriminating between species because there tends to be little intra-specific variation but considerable inter-specific variation with this gene (Murrell *et al.*, 2001). Within the rhipicephaline family, *R. praetextatus*, the *Ruma* species and *R. simus* were found to be closely related based on the 12S gene. These species are also closely related to *R. compositus* and *R. muhsamae.*

Morphological, physiological and phylogenetical data has previously supported the existence of at least two distinct *R. appendiculatus* groups thought to represent two geographic genetically differentiated lineages (Matambo *et al.,* 2007a, 2007b; Madder *et al.,* 1999, Speybroeck *et al.,*

2004). Ticks found in Southern African region (South Africa, Southern Zambia and Zimbabwe) and those distributed mainly in Eastern Africa region (parts of Kenya, Tanzania and Uganda, Burundi and Rwanda) are thought to constitute two geographic groups that display major morphological, ecological and epidemiological differences (Mtambo *et al.,* 2007b). In this study, phylogeny of *R. appendiculatus* based on COI and supported by 12S revealed a strong genetic partitioning of the species into two major lineages (Figure 32 and 33). These well separated monophyletic groups are herein referred to as haplogroup A and B. Hence, *R. appendiculatus* in Kenya is characterized by the existence of two distinct highly differentiated genetic haplotypes separated by a remarkable 12 mutation points which lack geographic or ecological structuring. The genetic variation between these two haplogroups was found to be significantly high (90.8%), (Table 4) supported by the high differentiation index (F_{ST}) of 0.907 (P= 0.000). These findings provide evidence for evolutionary divergent lineages for *R. appendiculatus*.

Median-Joining network analysis revealed haplogroup A to have diverged from haplogroup B and thus B can be said to be the ancestor of A. This is the first study to demonstrate the actual COI divergence within *R. appendiculatus* species. This data further confirms earlier findings that the existence of the two genetically differentiated groups of *R. appendiculatus* extends beyond Zambia (Mtambo *et al.*, 2007a) where *R. appendiculatus* from eastern and southern provinces represent two distinct geographic ecological groups. While the differentiation of these two groups was thought to be significantly driven by agro-ecological and climatic factors (Madder *et al.*, 2002, Speybroeck *et al.*, 2004; Mtambo *et al.*, 2007a, 2007b) the distribution of the Kenyan field *R. appendiculatus* as revealed by COI and supported by 12S phylogenetics does not appear to be influenced by these factors. In the Kenyan stocks, COI haplotypes consisting of ticks from two geographically separated regions (KF (Coastal Kenya) and RU (Rusinga Island, in Lake Victoria) as shown in Figure 10

(Chapter 3) were observed to fall within the same haplogroup (haplogroup B) based on the NJ and MJ analysis (Figure 32 and 33). These two tick populations were observed to fall within the ancestor group. Other populations with significant number of ticks in this group include BU and laboratory stocks MF and KU. Ticks from KF, RU and BU were collected mainly from indigenous breeds of cattle or pastures grazed by these cattle. Conversely, haplotypes consisting of ticks from laboratory stocks from Uganda (UG), Zimbabwe (ZM), Southern Province of Zambia (ZS), South Africa (SAL) and field ticks from Natal, Province, South Africa (SAN) were also found to belong to this ancestor haplogroup, B. The clustering of ticks from Southern Africa, Southern Province of Zambia and Zimbabwe in the same haplogroup is consistent with the separation observed in earlier studies by Mtambo *et al.* (2007a). Haplotypes consisting of ticks from the laboratory stocks LP, ML and eight out of the 12 samples from KH occurred in haplogroup A (Table 16) which appeared to have recently diverged from B. All the ticks from Eastern province of Zambia also clustered in haplogroup A. Most haplotypes consisting of field ticks collected from areas grazed by both cattle and wildlife or strictly wildlife (NB, FP, BO and MA) also clustered in haplogroup A in addition to populations from field ticks from areas grazed by cattle (MK and KT) (Table 16). Ticks from MK population were mainly collected from mixed cross-breeds while those from KT were mainly collected from exotic breeds. The standard *R. appendiculatus* laboratory stock (MU) and a related selected lineage that is known to show low infectivity with *T. parva* (ML) also occurred in this haplogroup. Findings from this study suggest that haplogroup A ticks have recently diverged from the ancestral haplogroup B. Thus the Kenyan tick populations represent two genetic haplogroups that lack geographic or ecological structuring, suggesting that some alternative factors have shaped the observed differentiation. The genotypic divergence displayed by the *R. appendiculatus* lineages could have probably arisen as a result of groups experiencing different selective pressures or haplogroup B has undergone genetic drift or mutations have arisen in the haplogroup B gene pool

overtime to give rise to haplogroup A. Other than suitable climate and habitat, inter-specific competition in immature stages of *H. rufipes* has been shown to influence its genetic structure (Cangi *et al.,* 2013). It would be important to establish the alternative factors that have influenced the genetic differentiation observed for *R. appendiculatus*.

The studied tick populations showed high within-population genetic variation based on the COI gene and relatively low population differentiation. The level of genetic variation among the 22 *R. appendiculatus* populations was very high (85.96%) while within the populations the molecular variance was 14.0 % (Table 17). Host-dependent rates of tick dispersal between colonies can alter infestation probabilities and local tick dynamics and may thus modify the adaptation potential of ticks to local hosts (McCoy *et al*., 2003). Domestic cattle in Kenya are frequently moved over large distances, which would facilitate tick dispersal over a large scale while the movement of the natural reservoirs of *R. appendiculatus* (wild bovidae) within the wildlife areas considered in this study is limited since these areas are fenced. While different tick hosts have been shown to influence the spatial genetic structure of *I. uriae* (McCoy *et al.,* 2003), the observed genetic variation of *R. appendiculatus* collected from areas harbouring different animal hosts was low (4.94 %) while within populations was 52.37 % (Table 17). This demonstrates low genetic differences between populations of *R. appendiculatus* originating from different animal host species suggesting an independent origin and history of the different populations in the three groups (cattle, cattle wildlife and wildlife). Although COI haplotype diversity and gene diversity was least in laboratory *R. appendiculatus* stocks compared to field populations, no genetic variation was observed between the two groups (-1.43%). The lack of haplotype diversity in most of the laboratory stocks could be attributed to closed colony expansion and high inbreeding.

Neutrality tests on the mitochondrial COI sequence variation revealed patterns that were not consistent with predictions of the neutral model. These tests showed that overall, *R. appendiculatus* in Kenya was experiencing a decrease in population size and/or balancing selection though the P values were not significant (Table 19). Five field *R. appendiculatus* populations (KT, BU, RUM2, MA and BO) and three laboratory stocks (KH, MF and MU) were also found to have COI patterns suggesting a decrease in population size and/or balancing selection though the P values were also not significant (Table 14). However, when the two major haplogroups were analysed separately they showed patterns consist with either a purifying selection or a demographic population expansion though the statistics were only significant for haplogroupp A. These findings suggest that the two major haplogroups of COI have experienced different processes and historical dynamics that have shaped the structure observed. Haplogroup A was found to be either undergoing a purifying selection or was experiencing a recent population expansion more rapidly or at a much faster rate than haplogroup B. Though five field populations (KF, MK, RU, FP and NB) showed patterns to suggest population size expansion, only RU, FP and NB were supported by a significant P values (Table 14). *R. pravus* and ticks from the unidentified *Rhipicephalus* species from Ruma showed patterns to suggest a decrease in population size and/or balancing selection. Patterns of distribution of pairwise differences supported the results obtained with neutrality tests (Figure 34). These results showed that the population structure of *R. appendiculatus* is defined by two distinct differentiated lineages that seem to be undergoing sudden population expansion independently. These lineages correspond to the two major haplogroups A and B identified by the phylogenetic and the median-joining network (Figure 32 and 33). These findings support an earlier observation that within tick population of the same species, COI appear to be more sensitive to demographic incidents such as founder events and bottlenecks than nuclear DNA (Avise, 1991; Birky *et al*., 1989; DeSalle & Giddings, 1986).

The findings of this study may have important taxonomical and distribution implications for *R. appendiculatus* and may point to an ongoing speciation in sub-Saharan Africa where the tick occurs. The observed genetic differentiation of *R. appendiculatus* may not be just driven by agro-ecological and climatic factors as previously thought and thus identifying other forces driving its differentiation may help in understanding the apparent sudden population expansion within the Sub-Saharan region. Although *R. appendiculatus* is a generalist tick, buffalos are the natural reservoirs and cattle the preferred domestic hosts of all stages of its development (Norval and Lightfoot, 1982; Okello-Onen *et al*, 1999). However, it is an ectoparasite of several wild and domestic animals. Such a generalist vector can easily spread through ecosystems and alter disease transmission cycles. Understanding the population structure of *R. appendiculatus* would be important in the design of sustainable control strategies as different tick populations may present differences in vector competence, acaricide resistance and rates of infectivity with *T. parva*. Thus, it would be important to establish how the biology of the two major halogroups compares in regards to acquisition and transmission of ECF.

Results from this chapter have been presented in the following conference;

1. **9th University of Nairobi, Faculty of Veterinary Medicine Biennial Scientific conference and Exhibition**- Genetic diversity and population structure of *Rhipicephalus appendiculatus* in Kenya

The abstract is given in Appendix II.

5.5 References

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of molecular biology*, ***215*****(3),** 403-410.

Amos, W. & Harwood, J. (1998). Factors affecting levels of genetic diversity in natural populations. *Philosophical Transactions of the Royal Society B: Biological Sciences of London*. **353(1366),** 177 – 186.

Anderson, D. L. & Trueman, J. W. H. (2000). *Varroa jacobsoni (Acari: Varroidae)* is more than one species. *Experimental and Applied Acarology*. **24,** 165 - 189.

Avise, J. C., Arnold, J., Martin Ball, R., Bermingham, E., Lamb, T., Neigel, J. E., Reeb, C. A. and Saunders, N. C. (1987). Intraspecific Phylogeography: the Mitochondrial DNA Bridge between Population Genetics and Systematics. *Annual Review of Ecology and Systematics* **18,** 489 –522.

Avise, J. C. (1991). Ten unorthodox perspectives on evolution prompted by comparative population genetic findings on mitochondrial DNA. *Annual Review of Genetics*. 25: 45 - 69.

Avise, J. C. (1994). Molecular markers, natural history and evolution. New York: Chapman and Hall.

Avise, J. C. (2000). Phylogeography: The history and formation of species. USA: Harvard University Press.

Avise, J. C. (2009). Phylogeography: retrospect and prospect. *Journal of Biogeography* **36(1),** 3 - 15.

Bailey, K.P. (1960). Notes on the rearing of *Rhipicephalus appendiculatus* and their infection with *Theileria parva* for experimental transmission. *Bulletin of Epizoonotic Diseases of Africa,* **8,** 33–43.

Ballard, J. W. and Kreitman, M. (1995). Is mitochondrial DNA a strictly neutral marker? *Trends in Ecology and Evolution* **10(12),** 485 - 488.

Balmer, O., Beadell, J. S., Gibson, W. and Adalgisa, C. (2011). Phylogeography and Taxonomy of Trypanosoma brucei. *PLoS Neglected Tropical Diseases* **5(2),** e961.

Bandelt, H.-J., Forster, P. and Röhl, A. (1999). Median-joining networks for inferring intraspecific phylogenies. *Molecular Biology and Evolution* **16,** 37–48.

Barker, S. C. and Murrell, A. (2002). Phylogeny, evolution and historical zoogeography of ticks: a review of recent progress. *Experimental and Applied Acarology*, **28,** 55-68.

Barker, S. C. and Murrell, A. (2008). Systematics and evolution of ticks with a list of valid genus and species names. In P. NuttallAS Bowman. Ticks: biology, disease and control. *Cambridge University Press.* pp. 1-39.

Bensasson, D., Zhang, D., Hartl, D. L. and Hewitt, G. M. (2001). Mitochondrial pseudogenes: evolution's misplaced witnesses. *Trends in Ecology and Evolution***, 16(6),** 314 - 321.

Berkvens, D. L., Geysen, D. M., Chaka, G., Madder, M. and Brandt, J. R. A. (1998). A survey of the ixodid ticks parasitising cattle in the Eastern Province of Zambia. *Medical and Veterinary Entomology*, **12,** 234-240

Birky Jr, C. W., Fuerst, P. and Maruyama, T. (1989). Organelle gene diversity under migration, mutation, and drift: equilibrium expectations, approach to equilibrium, effects of heteroplasmic cells, and comparison to nuclear genes. *Genetics,* **121,** 613 - 627.

Black, W. C. and Piesman, J. (1994). Phylogeny of hard and soft-tick taxa (Acari: Ixodida) based on mitochondrial 16S rDNA sequences. *Proceedings of the National Academy of Sciences,* **91,** 10034 - 10038.

Blouin, M. S., Yowell, C. A., Courtney, C. H. and Dame, J. B. (1995). Host movement and the genetic structure of populations of parasitic nematodes. *Genetics* 141: 1007 - 1014.

Cangi, N., Horak, I.G., Apanaskevich, A. D., Matthee, S., das Neves, L C. B. G., Estrada-Peña, A., Matthee, C.A. (2013). The Influence of Interspecific Competition and Host Preference on the Phylogeography of Two African Ixodid Tick Species. *Public Library of Science One.* **8(10),** e76930

Cox, A. J. and Hebert, P. D. N. (2001). Colonization, extinction and phylogeographic patterning in a freshwater crustacean. *Molecular Ecology.* **10**, 371–386.

Cruickshank, R. H. (2002). Molecular markers for the phylogenetics of mites and ticks. *Systematic and Applied Acarology,* **7,** 3 - 14.

Cumming, G. S. (1999). Host distributions do not limit the species ranges of most African ticks (*Acari: Ixodidae*). *Bulletin of Entomological Research,* **89,** 303 – 32.

De Meeûs, T. and Renaud, F. (2002). Parasites within the new phylogeny of eukaryotes. *Trends in Parasitology,* **18,** 247-251.

De Meeûs, T., Koff, B. B., Barré, N., Garine-Wichatitsky, M. and Chevillon, C. (2010). Swift sympatric adaptation of a species of cattle tick to a new deer host in New Caledonia. *Infectious Diseases and Evolution,* **10(7),** 976 - 983.

DeSalle, R. and Giddings, L. V. (1986). Discordance of nuclear and mitochondrial DNA phylogenies in Hawaiian Drosophila. *Proceedings of the National Academy of Science of the U SA,* **83(18),** 6902 - 6906.

Dobson, S. J. and Barker, S. C. (1999). Phylogeny of the Hard Ticks (Ixodidae) Inferred from 18S rRNA indicates that the genus *Aponomma* is paraphyletic. *Molecular Phylogenetics and Evolution,* **11(2),** 288 - 295.

Dowling, D. K., Friberg, U. and Lindell, J. (2008). Evolutionary implications of non neutral mitochondrial genetic variation. *Trends in Ecology and Evolution,* **23(10),** 546 - 554.

Elder, J. F. and Turner, B. J. (1995). Concerted evolution of repetitive DNA sequences in eukaryotes. *Quarterly Review of Biology,* **70,** 297 - 320.

Excoffier, L., and Lischer, H. E. L. (2010). Arlequin suite version 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Molecular and Ecological Resources* **10,** 564-567.

Folmer, O., Black, M., Hoeh, W., Lutz, R. & Vrijenhoek, R. (1994). DNA primers for amplification of mitochondrial cytochrome *c* oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology.* **3**, 294–299.

Freeland, J. (2005). Molecular Ecology. England: John Wiley & Sons.

Fu, Y. X. (1997). Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics 147, 915–925.

Gadagkar, S. R., Rosenberg, M. S. and Kumar, S. (2005). Inferring species phylogenies from multiple genes: concatenated sequence tree versus consensus gene tree. *Journal of Experimental Zoology B*: *Molecular and Developmental Evolution* **30(4B),** 64 -74.

Galtier, N., Nabholz, B., Glémin, S. and Hurst, G. D. (2009). Mitochondrial DNA as a marker of molecular diversity a reappraisal. *Molecular Ecology***, 18(22),** 4541 – 4550

Gaziev, A. I. and Shaikhaev, G. O. (2010). Nuclear Mitochondrial Pseudogenes. *Molecular Biology,* **44(3),** 358 - 368.

Hare, M. P. (2001). Prospects for nuclear gene phylogeography. *Trends in Ecology and Evolution* **16(12),** 700 - 706.

Harpending, H. (1994). Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. *Human Biology,* **66,** 591–600.

Harrison, R. G. (1989). Animal mitochondrial DNA as a genetic marker in population and evolutionary biology. *Trends in Ecology and Evolution,* **4,** 6 - 11.

Hassanin, A. (2006). Phylogeny of Arthropoda inferred from mitochondrial sequences: Strategies for limiting the misleading effects ofmultiple changes in pattern and rates of substitution. *Molecular Phylogenetics and Evolution***, 38,** 100 - 116.

Hilburn, L. R. and Sattler, P. W. (1986). Are tick populations really less variable and should they be? *Heredity,* **57,** 113 - 117.

Hillis, D. M. and Dixon, M. T. (1991). Ribosomal DNA: molecular evolution and phylogenetic inference. *Quarterly Review of Biology,* **66,** 411- 429.

Hlinka, O., Murrell, A. and Barker, S. C. (2002). Evolution of the secondary structure of the rRNA internal transcribed spacer 2 (ITS2) in hard ticks (*Ixodidae, Arthropoda*). *Heredity* **88(4),** 275 - 279.

Holderegger, R., Kamm, U. and Gugerli, F. (2006). Adaptative vs. neutral genetic diversity: implications for landscape genetics. *Landscape Ecology*, **21,** 797 – 807.

Hwang, U. W. and Kim, W. (1999). General properties and phylogenetic utilities of nuclear ribosomal DNA and mitochondrial DNA commonly used in molecular systematics. *The Korean Journal of Parasitology***, 37(4),** 215 –228.

Jiggins, F. M., Schulenburg, J. H., Hurst, G. D. and Majerus, M. E. N. (2001). Recombination confounds interpretations of Wolbachia evolution. *Proceedings of the Royal Society of London. Series B*, **268,** 1423 - 1427.

Jongejan, F. and Uilenberg, G. (2004). The global importance of ticks. *Parasitology,* **129,** 1 - 12.

Kain, D. E., Sperling F. A., Daly, H. V. and Lane, R. S. (1999). Mitochondrial DNA sequence variation in *Ixodes pacificus* (*Acari: Ixodidae*). *Heredity*, **83,** 378 - 386.

Klompen, J. S. H., Black, W. C., Kierans, J. E. and Oliver Jr, J. H. (1996). Evolution of ticks. *Annual Review of Entomology*, **41,** 141 - 161.

Lampo, M., Rangel, Y. and Mata, A. (1998). Population genetic structure of a three-host tick, *Amblyomma dissimile*, in eastern Venezuela. *Journal of Parasitology*, **84,** 1137 – 1142.

Librado, P. and Rozas, J. (2009). DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics,* **25,** 1451–1452.

Madder, M., Speybroeck, N., Brandt, J., Berkvens, D. (1999). Diapause induction in adults of three *Rhipicephalus appendiculatus* stocks. *Experimental Applied Acarology* **23,** 961-968.

Mangold, A. J., Bargues, M. D. and Mas-Coma, S. (1998). Mitochondrial 16S rDNA sequences and phylogenetic relationships of species of *Rhipicephalus* and other tick genera among Metastriata (*Acari: Ixodidae*). *Parasitology Research*, **84,** 478 - 484.

Marrelli, M. T., Souza, L. F., Marques, R. C.,Labruna, M. B., Matioli, S. R., Tonon, A. P.,Ribolla, P. E. M., Marinotti, O. and Schumaker, T. T. S. (2007). Taxonomic and phylogenetic relationships between Neotropical species of ticks from genus *Amblyomma* (*Acari: Ixodidae*) inferred from second internal transcribed spacer sequences of rDNA. *Journal of Medical Entomology,* 44: 222 - 228.

Martin, A. P. (1995). Metabolic rate and directional nucleotide substitution in animal mitochondrial DNA. *Molecular Biology and Evolution,* **12(6),** 1124 – 1131.

Mattioli, R. C., Pandey, V. S., Murray, M. and Fitzpatrick, J. L. (2000). Immunogenetic influences on tick resistance in African cattle with particular reference to trypanotolerant N'Dama (*Bos taurus*) and trypanosusceptible Gobra zebu (*Bos indicus*) cattle. *Acta Tropica* **75(3),** 263 - 277.

McCoy, K. D., Boulinier, T., Tirard, C. and Michalakis, Y. (2003). Host-dependent genetic structure of parasite populations: differential dispersal of seabird tick host races. *Evolution,* **57(2),** 288-296.

McLain, D. K., Daw, M. W., Collins, F. H. and Oliver Jr, J. H. (1995). Evolution of the rDNA spacer, ITS 2, in the ticks *Ixodes scapularis* and *I. pacificus* (*Acari: Ixodidae*). *Heredity,* **75,** 303- 319.

Morris-Pocock, J. A., Taylor, S. A., Birt, T. P. and Friesen, V. L. (2010). Concerted evolution of duplicated mitochondrial control regions in three related seabird species. *BMC Evolutionary Biology,* **10,** 14.

Mtambo, J., Madder, M., Van Bortel, W., Berkvens, D. and Backeljau, T. (2007a). *Rhipicephalus appendiculatus* and *R. zambeziensis* (*Acari: Ixodidae*) from Zambia: a molecular reassessment of their species status and identification. *Experimental and Applied Acarology,* **41,** 115-128.

Mtambo, J., Madder, M., Van Bortel, W., Berkvens, D., Chaka, G. and Backeljau, T. (2007b). Further evidence for geographic differentiation in *R. appendiculatus* (Acari: Ixodidae) from Eastern and Southern provinces of Zambia. *Experimental and Applied Acarology,* **41,** 129-138

Murrell, A., Campbell, N. J. H. and Barker, S. C. (2001). A total-evidence phylogeny of ticks provides insights into the evolution of life cycles and biogeography. *Molecular Phylogenetics and Evolution,* **21(2),** 244 – 258.

Nabholz, B., Glémin, S. and Galtier, N. (2008). Strong variations of mitochondrial mutation rate across mammals- the longevity hypothesis. *Molecular Biology and Evolution* **25(1),** 120-130.

Nava, S., Szabo, M. P. J., Mangold, A. J. and Guglielmone, A. A. (2008). Distribution, hosts, 16SrDNA sequences and phylogenetic position of the Neotropical tick *Amblyomma parvum* (Acari: Ixodidae). *Annals of Tropical Medicine and Parasitology* **102(5),** 409 - 425.

Navajas, A., Lagnel, J., Fauvel, G. and De Moraes, G. (1999). Sequence variation of Ribosomal Internal Transcribed Spacers (ITS) in commercially important Phytoseiidae mites. *Experimental and Applied Acarology,* **23(11),** 851 - 859.

Navajas, M. (1998). Host plant associations in the spider mite *Tetranychus urticae* (*Acari: Tetranychidae*): insights from molecular phylogeography. *Experimental and Applied Acarology,* **22,** 201 - 214.

Navajas, M. and Fenton, B. (2001). The application of molecular markers in the study of diversity in acarology: a review. *Experimental and Applied Acarology,* **24,** 751 - 774.

Neigel, J. E. and Avise, J. C. (1985). The precision of histocompatibility response in clonal recognition in tropical marine sponges. *Evolution,* **39,** 729 - 732.

Norris, D. E., Klompen, J. S. H. and Black, W. C. (1999). Comparison of the mitochondrial 12S and 16S ribosomal DNA genes in resolving phylogenetic relationships among hard ticks (*Acari: Ixodidae*). *Annals of the Entomological Society of America,* **92,** 117 - 129.

Norval, R. A. I. and Lightfoot, C. J. (1982). Tick problems in wildlife in Zimbabwe. Factors influencing the occurrence and abundance of *Rhipicephalus appendiculatus*. Zimbabwe. *Veterinary Journal,* **13,** 11–20.

Norval, R. A. I., Lawrence, J. A., Young, A. S., Perry, B. D., Dolan, T. T. and Scott, J. (1991). *Theileria parva*: Influence of vector, parasite and host relationships on the epidemiology of theileriosis in southern Africa. *Parasitology,* **102,** 347-356.

Norval, R. A. I., Perry, B. D. and Young, A. S. (1992). The epidemiology of theileriosis in Africa. Academic Press, London.

Okello-Onena, J., Tukahirwa, E.M., Perry, B.D., Rowlands, G.J., Nagda, S.M., Musisi, G., Bode, E., Heinonen, R., Mwayi, W. and Opuda-Asibo, J. (1999). Population dynamics of ticks on indigenous cattle in a pastoral dry to semi-arid rangeland zone of Uganda. *Experimental & Applied Acarology,* **23,** 79–88

Palumbi, S. R., Prager, E. M., Sage, R. D. and Stoneking, M. (1985). Mitochondrial DNA and two perspectives on evolutionary genetics. *Biological Journal of the Linnean Society* **26,** 375 -400.

Pollard, D. A., Iyer, V. N., Moses, A. M. and Eisen, M. B. (2006). Widespread discordance of gene trees with species tree in Drosophila: evidence for incomplete lineage sorting. *PLoS Genetics* **2(10),** e173.

Rees, D. J., Dioli, M. & Kirkendall, L. R. (2003). Molecules and morphology: evidence for cryptic hybridization in African *Hyalomma* (*Acari: Ixodidae*). Molecular Phylogenetics and *Evolution*. **27,** 131 - 142.

Rich, S. M., Rosenthal, B. M., Telford, S. R., Spielman, A., Harti, D. L. and Ayala, F. J. (1997). Heterogeneity of the internal transcribed spacer (ITS-2) region within individual deer ticks. *Insect Molecular Biology*, **6,** 123 - 129.

Roderick, G. K. (1996). Geographic structure of insect populations: gene flow, phylogeography, and their uses. *Annual Review of Entomology*, **41,** 325 - 352.

Rogers, A. R. and Harpending, H. (1992). Population growth makes waves in the distribution of pairwise genetic differences. *Molecular Biology and Evolution* 9, 552–569.

Rokas, A., Ladoukakis, E. and Zouros, E. (2003). Animal mitochondrial DNA recombination revisited. *Trends in Ecology and Evolution*, **18(8),** 411 - 417.

Saccone, C., Gissi, C., Lanave, C., Larizza, A., Pesole, G. and Reyes, A. (2000). Evolution of the mitochondrial genetic system: an overview. *Gene* **261(1),** 153 - 159.

Shao, R., Barker, S. C., Mitani, H., Aoki, Y. and Fukunaga, M. (2005). Evolution of duplicate control regions in the mitochondrial genomes of metazoa: a case study with Australasian Ixodes ticks. *Molecular Biology and Evolution***, 22(3),** 620 – 629.

Shao, R. and Barker, S. C. (2007). Mitochondrial genomes of parasitic arthropods: implications for studies of population genetics and evolution. *Parasitology*, **134,** 153 - 167.

Simon, C., Frati, F., Beckenbach, A., Crespi, B., Liu, H. and Flook, P. (1994). Evolution, weighting and phylogenetic utility of mitochondrial gene sequences and a compilation of conserved polymerase chain reaction primers. *Annals of Entomological Society of America* **87,** 651- 701.

Slatkin, M. (1987). Gene flow and the geographic structure of natural populations. Science **236,** 787 -792.

Slatkin, M. and Hudson, R. R. (1991). Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. *Genetics*, **129,** 555-562.

Speybroeck, N., Madder, M., Thulke, H. H., Mtambo, J., Tirry, L., Chaka, G., Marcotty, T., Berkvens, D. (2004). Variation in body size in the tick complex *Rhipicephalus appendiculatus/Rhipicephalus zambeziensis*. *Journal of Vector Ecology* **29,** 347-354.

Swofford, D. L. (2003). PAUP*. Phylogenetic Analysis Using Parsimony (* and Other Methods), version 4. Sunderland, MA: Sinauer Associates

Tamura K., Stecher G., Peterson D., Filipski A. and Kumar S. (2013). MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular Biology and Evolution* 30, 2725-2729.

Tatchell, R. J. (1987). Interaction between ticks and their hosts. *International Journal of Parasitology* **17(2),** 597 - 606.

Tautz, D., Arctander, P., Minelli, A., Thomas, R. H., Vogler, A. P. (2002). DNA points the way ahead in taxonomy. *Nature*, **418**, 479–479.

Tsaousis, A. D., Martin, D. P., Ladoukakis, E. D., Posada, D. and Zouros, E. (2005). Widespread recombination in published animal mtDNA sequences. *Molecular Biology and Evolution*, **22,** 925 - 933.

Wares, J. P. & Cunningham, C. W. (2001). Phylogeography and historical ecology of the North Atlantic intertidal. *Evolution* **12**, 2455–2469.

Wilson, A. C., Cann, R. L., Carr, S. M., George, M., Gyllensten, U. B., Helmbychowski, K. M., Higuchi, R. G., Palumbi, S. R., Prager, E. M., Sage, R. D. and Stoneking, M. (1985). Mitochondrial DNA and two perspectives on evolutionary genetics. *Biological Journal of the Linnean Society*, 26: 375 -400.

Zahler, M., Gothe, R. and Rinder, H. (1995). Genetic evidence against a morphologically suggestive conspecificity of *Dermacentor reticulatus* and *D. marginatus*. *International Journal for Parasitology*, **25(12),** 1413 - 1419.

Zhang, D. X. and Hewitt, G. M. (2003). Nuclear DNA analyses in genetic studies of populations: practice, problems and prospects. *Molecular Ecology*, **12(3),** 563 - 584.

CHAPTER SIX: GENERAL DISCUSSION, CONCLUSION AND RECOMMENDATIONS

6.0 General discussion

Athough the economic and veterinary importance of *R. appendiculatus*, the tick vector for *T. parva* that causes East coast fever is well known, the genetic diversity and population structure of this tick has not been elucidated due to lack of appropriate molecular markers. The availability of a public EST database and some BAC sequences for *R. appendiculatus* (Nene *et al.,* 2004; Sunter *et al.,* 2008) provided a unique opportunity to accelerate the elucidation of the population structure and genetic diversity of this vector through development of micro- and minisatellite markers. The present study is the first to identify and evaluate genetic markers of *R. appendiculatus* and utilize them to estimate the genetic diversity between and within a significant number of important rhipicephaline tick species. To further assess the phylogenetic relationship between field and laboratory stocks of *R. appendiculatus* and other rhipicephaline tick species from Kenya, this study examined sequence variation in a 558 bp DNA fragment of tick mitochondrial COI gene, a 345 bp fragment of the mitochondrial 12S rDNA and an 1149 bp region of ribosomal nuclear ITS2. The use of multiple genetic markers targeting different genomic regions has been recommended to overcome drawbacks associated with one marker as well as to increase the accuracy in the construction of a phylogenetic history of species (Gadagkar *et al*., 2005; Balmer *et al*., 2011).

This study identified over 1000 micro- and minisatellite markers present in the two available *R. appendiculatus* nucleotide databases (Chapter TWO). The 7340 non-redundant salivary gland EST sequences which represented 0.6 % of the total *R. appendiculatus* genome were found to have repeats at a frequency of 1 repeat for every 6.673 Kb. The eleven BAC sequences that represent 0.036 % of the entire *R. appendiculatus* genome were found to contain repeat markers at a

frequency rate of 1 repeat for every 2.405 Kb of genomic sequence. These findings suggest that in *R. appendiculatus* the relative abundance of repeats is three times higher in genomic sequences compared to ESTs. The relatively low number of repeat sequences in ESTs may reflect the lower frequency of repeats in coding regions of *R. appendiculatus* genome since ESTs represent the transcribed regions of the genome. Thus, these databases were found to be valuable resources for rapid discovery of micro- and minisatellite repeats. A total of 29 informative markers characterized in this study confirmed that EST and genomic sequences are useful as a source of helpful micro- and minisatellite markers (Ellis & Burke, 2007) in the absence of a reference genome for *R. appendiculatus*. Repetitive DNA has been reported to occur in other tick species. In *R. microplus*, highly repetitive DNA comprising of tandem and dispersed repeats was found to account for approximately 40% of the total tick genome (Ullmann *et al.*, 2005) while a significant percentage (~35-40) of the *I. scapularis* and *A. americanum* genomes are comprised of highly repetitive sequences (Ullmann *et al.*, 2005; Palmer *et al.*, 1994).

The diversity indices used to characterize the 29 polymorphic loci described in this study displayed considerable genetic variability across the study populations. Analysis of allelic diversity showed 18 and 19 loci having a He and PIC of above 0.50 respectively. The fact that some markers had high and others low levels of variability suggests that the use of all the 29 markers developed in this study for population genetic analyses may result in a balanced assessment of the extent of genetic diversity and divergence within and between populations. The positive correlations between He, TNA and PIC observed in this study (Chapter Three) suggest that any of these indices can be used to select markers for population genetic studies independently. It was also noted that most polymorphic markers had high diversity and therefore they were most likely to be useful in discriminating between populations. On the other hand, the negative correlations observed between

He, TNA and PIC with F_{ST} suggest that the most diverse markers do not necessarily give the best estimate of genetic differentiation between populations.

The suitability and usefulness of the 29 micro- and minisatellite marker in discriminating the 27 tick populations studied was determined using the Multiple Co-inertia Analysis (MCoA) approach of Laloë *et al.,* (2007) and percent Principle Coordinate Analysis (%PCA) (Chapter 3). The MCoA approach had been used to determine the efficiency of 30 microsatellite markers to determine population structuring of different breeds of cattle from France and West Africa. On the basis of this analysis, the markers separated populations of *R. appendiculatus, R. zambeziensis* and four other rhipicephaline species from Kenya into two groups. Four markers (TC1002, TC23, TC728b and TC919) with a high Tv value seem to be the most efficient in separating these populations. With the exception of *R. zambeziensis* which clusters together with populations of *R. appendiculatus*, the other four rhipicephaline species sampled in Kenya segregated distinctively from *R. appendiculatus. R. zambeziensis* is morphologically similar to *R. appendiculatus* and the markers analysed in this study could not discriminate between the two species. In addition, though PCA analysis grouped *R. appendiculatus* into two clusters, it failed to discriminate between laboratory and field stocks. It would be of interest therefore to determine what other factors, other than geographical separation of the ticks contributes to the segregation of *R. appendiculatus* populations analysed. From MCoA and %PCA analyses, the markers with significant contribution to structuring of populations were found to be different. This suggests that markers that are efficient in separating *R. appendiculatus* populations are not necessarily efficient in separating *R. appendiculatus* from other rhipicephaline species and especially *R. zambeziensis*. This can be extrapolated to mean that the choice of markers is best determined by the species of interest and the research questions being addressed. Results from percent PCA reveal a distinct line of laboratory *R. appendiculatus* derived from the Muguga

stock exists. These genotypic differences have been reported previously using the single copy gene Ra86 (Kamau *et al.,* 2011). This finding is worthwhile since the Muguga laboratory stock is used to produce the infection and treatment (ITM) live vaccine for immunization against ECF (Radley *et al.,* 1981; reviewed by Di giulio *et al.,* 2009) and has most likely diverged from the original field collection with unknown implications for the vaccine.

The 29 polymorphic micro- and minisatellite loci reported in this study displayed varied species specificity (Chapter Three). Each of the 29 markers had certain alleles that were specific to populations of *R. appendiculatus* while others were species-specific. On the other hand, nine markers did not yield amplicons with some populations of *R. appendiculatus*. These markers could be used to discriminate specific populations of *R. appendiculatus* from the others. Twenty-three out of the 29 polymorphic EST markers also generated amplicons with *R. zambeziensis, R. praetextatus, R. pravus, R. pulchellus, and R. evertsi* and certain alleles in some of these markers were specific to these other species in the genus *Rhipicephalus*. These 23 markers together with the six that did not yield any amplicons in other *Rhipicephalus* species can potentially be used to discriminate and clarify genetic differentiation between populations in cases where morphological characters are not sufficiently descriptive. These too have the potential to be included in multiplex assays to discriminate between species. Consequently, though these EST markers have been derived from *R. appendiculatus,* they can be used to perform inter-species population genetics analysis as it has been done in some other tick species (Leo *et al.,* 2012) *Tephritidae* species (Stratikopoulos *et al.,*, 2009), birds (Primmer *et al.,* 1996), mosquitoes (Kamau *et al.,* 1999) and shrimps (Perez *et al.,* 2005).

BLAST analysis revealed that 7 out of the 29 informative markers found in this study are present within protein coding regions in ixodid tick genomes. These markers can therefore be of use in

determining gene traits and assisting in marker-assisted selection. However, there is a possibility of such markers being under selective pressure, either directly or indirectly through linkage to a locus that encodes a trait under selection. Analysis to detect if any of the markers were subject to selection revealed that three markers (TC1002, TC1284a and TC875) are likely to be under positive selection and TC440a is likely to be under stabilizing selection. Further evaluation of these four markers will establish if they are linked to any trait or region of the genome under selection. In the case that these markers are under selection, their use in future population genetics studies will be limited. Given that 11 EST-SSRs were in genes with putative functions based on homology with DNA sequences from other organisms, it is possible that some of them are subject to selection. It was however not possible to investigate this possibility in detail in the absence of full length gene and protein sequences, which would have allowed the, potentially more definitive, analysis of the relative frequency of synonymous and non-synonymous substitutions.

The identification of the 29 EST-based micro and minisatellite markers by data mining and validation by PCR provided an optimum marker system which was applied to elucidate the genetic diversity and population structure of *R. appendiculatus* (Chapter Four). The phylogeographic structure of *R. appendiculatus* in Kenya was examined to determine whether the genetic diversity in *R. appendiculatus* is spatially structured or is homogeneously distributed at different geographic scales and across a range of host species. Estimates of genetic and allelic diversity were found to be high for field-collected *R. appendiculatus* ticks in comparison with laboratory-maintained stocks. These findings suggest that field ticks likely maintain equilibrium genotype frequencies within and across individual animals, herds and populations implying that the effect of genetic drift and inbreeding is weak. This is supported by the mantel test which revealed field populations to be at migration-drift equilibrium. On the contrary, laboratory-bred stocks had lower levels of allelic and

genetic diversity possibly due to a higher level of inbreeding (overall F_{IS} = 0.305) and genetic drift due to their small effective population sizes.

The present study provided no evidence of host-related population genetic structure and specificity. Ticks that were collected from areas grazed by cattle, cattle-wildlife and wildlife herds shared the same genetic background. Furthermore, the level of genetic variation between these three groups was the lowest among all the clusters tested in AMOVA. These findings suggest a lack of host-associated adaptation/specialization by *R. appendiculatus*. This trait could be a survival strategy by *R. appendiculatus* as it provides micro-niches that allow the ectoparasite to flourish in different ecological environments. In host-tick associations with strong co-evolutionary signals, co-speciation events may result in host specificity (Combes, 2001). For parasites with free-living stages, mobile hosts and capacity for population intermixing, local adaptation may direct specificity (Gandon & Michalakis 2002; Greischar & Koskella 2007). The evolution of host-specificity has been demonstrated in *Ixodes uriae* (McCoy *et al.*, 2001, 2005), with experiments revealing an adaptive basis for these patterns (Dietrich, 2011).

The two approaches used to evaluate genetic structure revealed a very weak genetic structure/substructure and differentiation among field ticks (Chapter 4). STRUCTURE and discriminant analysis of principle components (DAPC) returned K = 2 and K = 6 respectively, as the most optimal number of genetic clusters among field ticks. Interestingly, at both "K" values, all individuals shared the observed genetic backgrounds disproportionately with no correlation between the generated clusters and the type of host species (cattle, cattle-wildlife and wildlife). Both approaches showed that irrespective of the host species, all field ticks were genetically admixed, that is, they share a relatively undifferentiated gene pool and a proportion of their genetic makeup.

They have a low level of population subdivision over a large geographic range. This was supported by the lack of isolation by distance as revealed by the mantel test, low level of genetic differentiation between field ticks and the signal of expansion (demographic and spatial) observed among field ticks, suggesting that tick mobility at large spatial scales may be frequent despite the low inherent vagility of *R. appendiculatus*. Previous studies that have examined the population structure of various species of *Ixodidae* found relatively high within-population genetic variation and weak genetic structure over a large spatial range. Delaye *et al.* (1997) suggested that *Ixodes ricinus* populations in Switzerland were panmictic due to a large number of host species at their disposal. McCoy *et al.* (2003) found that *Ixodes uriae* parasitizing black-legged kittiwakes were highly substructured than those parasitizing Atlantic puffins with estimates of population differentiation in kittiwake tick populations being almost twice as large as those of puffin tick populations. The former also showed evidence of isolation by distance suggesting that opportunities for dispersal for puffin ticks is much greater than for kittiwake ticks due to differences in the movement patterns and social behaviour of host species. The genetic structure of parasites with low vagility such as ticks is driven by the dynamics of transmission by their host species as it has been observed with *Geomydoecus actuosi* and its host *Thomomys bottae* (Nadler *et al.*, 1990), and nematodes and their livestock and deer hosts (Blouin *et al.*, 1995). Where there is relatively high dispersal rate among hosts, the parasites become mixed and panmixia is more likely.

Both STRUCTURE and DAPC analyses indicated that the laboratory bred ticks may be diverging from the field strains. Only two laboratory bred ticks (MU and MF) shared the same genetic cluster with the field ticks (Chapter 4). On the other hand the laboratory bred stocks seem also to have diverged from each other, each being clearly represented by a distinct genetic cluster. Two selected lineages, ML which is considerably less susceptible to *T. parva* infection and KH which is highly

susceptible (Young *et al.,* 1995; Odongo *et al.,* 2009) appear to have completely diverged from their original unselected stocks (MU and KU) respectively. Effects of intense selection, several years of isolation and inbreeding have probably resulted in differentiation of these ticks into new distinct populations given that the original unselected stocks were from the same geographical region in Central Kenya.

Genetic diversity analysis carried out in this study showed that tick populations sampled from different geographic locations and across three different host species were characterised by a clear lack of genetic sub-structuring (Chapter 4). Genotypes of individual ticks were highly admixed and indicate extensive gene flow across the geographic range of the species. Of importance was the clear genetic divergence between field tick populations and most of laboratory bred stocks examined. These results have direct implications for the coevolutionary interactions of *R. appendiculatus* and its hosts with subsequent consequences for the epidemiology of tick-borne diseases. More broadly, the results suggest that for most parasites, their genetic structure is linked to the vagility of their hosts, and that parasite-host systems provides a rich source of variability in population genetic structure. Host-dependent rates of dispersal between colonies will alter infestation probabilities and local dynamics in domestic or maintenance hosts.

Molecular identification of species based on homology searches against sequences deposited in major databases is used to augment taxonomy and species delimitation, either as corroborating evidence for existing hypotheses or as a starting point for further testing by other methods (DeSalle *et al,.* 2005). This study used COI to confirm species identity and assess the phylogenetic relationship within *R. appendiculatus* and the larger rhipicephaline family. The evolution of COI gene is known to be rapid enough to allow the discrimination of not only closely-related species, but

also phylogeographic groups within a single species (Cox & Hebert, 2001; Wares & Cunningham, 2001).

The COI divergence within the *R. appendiculatus* species was observed to be 3% or less while among species, the divergence was observed to be about 16%. The gene conservation within *R. appendiculatus* was found to be high as revealed by a sequence identity that ranged between 97-100% while inter-species diversity was high. COI was able to clearly discriminate *R. appendiculatus* from other rhipicephaline species analysed. Sequence identity ranged from an average of 87% between *R. appendiculatus* and *R. praetextatus*, 85% between *R. appendiculatus* and *R. pravus*, 84% between *R. appendiculatus* and *R. evertsi*, 86% between *R. appendiculatus* and an unidentified *Rhipicephalus* species from Ruma National park, Kenya, and an average of 86% between *R. appendiculatus* and *R. pulchellus* (Chapter 5). Thus, COI was found to be a powerful genetic marker for discriminating between closely-related *Rhipicephalus* tick species.

Though the COI gene was able to confirm the species identity of ticks belonging to *R. appendiculatus, R. pulchellus and R. evertsi* species, the identity of other species could not be resolved. For some tick species, morphological identity and COI characterization were not congruent. For example, 4 out 8 sequences of *R. zambezienis* clustered and matched with high identity to *R. appendiculatus* sequences. Since the *R. zambeziensis* sequences were from a laboratory stock, it is possible that there was a mix up of tick stocks during growth and maintenance. COI could not also confirm the species identity of ticks morphologically identified as *R. praetextatus, R. pravus* and an unknown species from Ruma National park.

Based on the COI analysis, *R. appendiculatus* was found to be closely related to *R. zambezienis* as observed with the micro- and minisatellite data. This finding has been reported before by Murell *et al.,* (2001). Morphological differentiation of the two species is usually difficult because they share a lot of similar traits (Norval *et al.,* 1992). *R. praetextatus* was found to be closely related to *R. simus.* Based on the COI gene, there was also a strong support for the closer genetic relationship between the unknown *Rhipicephalus* species from Ruma and *R. praetextatus,* and *R. simus.*

Morphological, physiological and phylogenetical data has previously supported the existence of at least two distinct *R. appendiculatus* groups thought to represent two geographic genetically differentiated lineages (Mtambo *et al.,* 2007a, 2007b; Madder *et al.,* 1999, Speybroeck *et al.,* 2004). Ticks found in Southern African region (South Africa, Southern Zambia and Zimbabwe) and those distributed mainly in Eastern Africa region (parts of Kenya, Tanzania and Uganda, Burundi and Rwanda) are thought to constitute two geographic groups that display major morphological, ecological and epidemiological differences (Mtambo *et al.,* 2007b). In this study, phylogeny of *R. appendiculatus* based on COI and supported by 12S revealed a strong genetic partitioning of the species into two major lineages (Chapter 5 Figure 31) here in identified as haplogroup A and B. Hence, *R. appendiculatus* in Kenya can be said to be characterized by the existence of two distinct highly differentiated genetic haplotypes which lack geographic or ecological structuring. In the Kenyan stocks, COI haplotypes consisting of ticks from two geographically separated regions (KF (Coastal Kenya) and RU (Rusinga Island, in Lake Victoria) were observed to fall within the same haplogroup (haplogroup B) based on the NJ and MJ analysis. These two tick populations were observed to fall within the same ancestor group. This result supports the observation made with the micro- and minisatellite data (Chapter 4 Figure 16). The genetic variation between the two haplogroups of *R. appendiculatus* was found to be significantly high (90.8%), (Chapter 5 Table 17)

supported by the high differentiation index (F_{ST}) of 0.907 (P = 0.000). These findings provide evidence for evolutionary divergent lineages for *R. appendiculatus*.

Median-Joining network analysis revealed haplogroup A to have diverged from B and thus B can be said to be the ancestor of A. This data further confirms earlier findings that the existence of the two genetically differentiated groups of *R. appendiculatus* extends beyond Zambia (Mtambo *et al.*, 2007a) where *R. appendiculatus* from eastern and southern provinces represent two distinct geographic ecological groups. The clustering of ticks from Southern Africa, Southern Province of Zambia and Zimbabwe in the same haplogroup is consistent with the separation observed in earlier studies by Mtambo *et al.*, (2007a). Findings from this study suggest that haplogroup A ticks have recently diverged from the ancestral haplogroup B. Thus the Kenyan tick populations represent two genetic haplogroups that lack geographic or ecological structuring, suggesting that some alternative factors have shaped the observed differentiation. The genotypic divergence displayed by the *R. appendiculatus* lineages could have probably arisen as a result of groups experiencing different selective pressures or haplogroup B has undergone genetic drift or mutations have arisen in the haplogroup B gene pool overtime to give rise to haplogroup A. Other than suitable climate and habitat, inter-specific competition in immature stages of *H. rufipes* has been shown to influence its genetic structure (Cangi *et al.*, 2013).

The studied tick populations showed high within-population genetic variation based on the COI gene and relatively low population differentiation. The level of genetic variation among the 22 *R. appendiculatus* populations was very high (85.96%) while within the populations the molecular variance was 14.0 %. Host-dependent rates of tick dispersal between colonies can alter infestation probabilities and local tick dynamics and may thus modify the adaptation potential of ticks to local

hosts (McCoy *et al*, 2003). Domestic cattle in Kenya are frequently moved over large distances, which would facilitate tick dispersal over a large scale while the movement of the natural reservoirs of *R. appendiculatus* (wild *bovidae*) within the wildlife areas considered in this study is limited since these areas are fenced. While different tick hosts have been shown to influence the spatial genetic structure of *I. uriae* (McCoy *et al.*, 2003), the observed genetic variation of *R. appendiculatus* collected from areas harbouring different animal hosts was low (4.94 %) while within populations was (52.37 %). This demonstrates low genetic differences between populations of *R. appendiculatus* originating from different animal host species suggesting an independent origin and history of the different populations in the three groups (cattle, cattle wildlife and wildlife). Although COI haplotype diversity and gene diversity was least in laboratory *R. appendiculatus* stocks compared to field populations, no genetic variation was observed between the two groups (-1.43%). The lack of haplotype diversity in most of the laboratory stocks could be attributed to closed colony expansion and high inbreeding.

Neutrality tests on the COI sequences showed that overall, *R. appendiculatus* in Kenya was experiencing a decrease in population size and/or balancing selection though the P values were not significant. Five field *R. appendiculatus* populations (KT, BU, RUM2, MA and BO) and three laboratory stocks (KH, MF and MU) were also found to have COI patterns suggesting a decrease in population size and/or balancing selection though the P values were also not significant. However, when the two major haplogroups were analysed separately they showed patterns consistent with either a purifying selection or a demographic population expansion though the statistics were only significant for haplogroup A. These findings suggest that the two major haplogroups of COI have experienced different processes and historical dynamics that have shaped the structure observed. Haplogroup A was found to be either undergoing a purifying selection or was experiencing a recent

rapid population expansion or at a much faster rate than haplogroup B. Though five field populations (KF, MK, RU, FP and NB) showed patterns to suggest population size expansion, only RU, FP and NB were supported by a significant P value. *R. pravus* and ticks from the unidentified *Rhipicephalus* species from Ruma National park showed patterns to suggest a decrease in population size and/or balancing selection. Patterns of distribution of pairwise differences supported the results obtained with neutrality tests. These results showed that the population structure of *R. appendiculatus* is defined by two distinct differentiated lineages that seem to be undergoing sudden population expansion independently. These lineages correspond to the two major haplogroups A and B identified by the phylogenetic and the median-joining network. These findings support an earlier observation that within tick population of the same species, COI appear to be more sensitive to demographic incidents such as founder events and bottlenecks than nuclear DNA (Avise, 1991; Birky *et al*., 1989; DeSalle & Giddings, 1986).

This study also analysed the phylogenetic relationship between the *R. appendiculatus* populations using the ITS2 and the 12S gene to corroborate the observations obtained with COI data. 12S identified two major haplotyes from 93 *R. appendiculatus* sequences while ITS2 identified one major haplotype from 87 *R. appendiculatus* sequences analysed. The two major haplotypes revealed by 12S correlated well with those observed using COI indicating a high level of congruence between the two genes. These findings prove that COI and 12S regions are better gene markers for studying intra-species diversity of closely related tick species while ITS2 region may be useful for discriminating between species because there tends to be little intra-specific variation but considerable inter-specific variation with this region (Murrell *et al.*, 2001). Within the rhipicephaline family, *R. praetextatus,* the *Ruma* species and *R. simus* were found to be closely

related based on the 12S gene. These species are also closely related to *R. compositus* and *R. muhsamae.*

Based on the findings of the satellite and gene markers analysis, this study showed that, depending on availability of public nucleotide databases, data mining can generate enough micro- and minisatellite markers for a variety of genetics tasks in many organisms. For new projects, a quick download of ESTs from the species of interest or closely related taxa, combined with the appropriate in silico analysis, might save money and months of bench work. The genetic markers described here were shown to be valuable resource for studying the genetics of *R. appendiculatus* and related species of the genus *Rhipicephalus* because of their ability to discriminate both within *R. appendiculatus* and between other rhipicephaline species. These markers have revealed that *R. appendiculatus* ticks occurring in Kenya are characterized by the existence of two distinct and well differentiated lineages. The satellite markers developed and validated in this study represent the first molecular tools for defining *R. appendiculatus* genotypes, their distribution and degree of geographic differentiation.

The findings of this study may have important implications on the taxonomy and distribution of *R. appendiculatus* and may point to an ongoing speciation in sub-Saharan Africa where the tick occurs. While the differentiation of the lineages in South Africa was thought to be significantly driven by agro-ecological and climatic factors (Madder *et al.,* 2002, Speybroeck *et al.,* 2004; Mtambo *et al.,* 2007a, 2007b) the distribution of the Kenyan field *R. appendiculatus* as revealed by COI and supported by 12S phylogenetics does not appear to be influenced by these factors. It would be important in future to establish the alternative factors that have influenced the genetic differentiation

observed for *R. appendiculatus*. This may help in understanding the apparent sudden population expansion within the Sub-Saharan region and associated ECF disease patterns.

Although *R. appendiculatus* is a generalist tick, African Cape buffalos are the natural reservoirs while cattle are the preferred domestic host for all of its development stages (Norval & Lightfoot, 1982; Okello-Onen *et al.*, 1999). However, it is an ectoparasite of several wild and domestic animals. Such a generalist vector can easily spread through ecosystems and alter disease transmission cycles. Genetic differences among tick populations in the specificity of selection for *T. parva* parasites may have major consequences for the population structure of the *T. parva* parasite in the field, since it has been shown that tick population heterogeneity is a major determinant of parasite diversity (Katzer *et al.,* 2006). In this regard, it may be useful to consider the potential selection imposed by the different tick populations when designing control strategies such as ITM vaccination. Understanding the population structure of *R. appendiculatus* would be important therefore in the design of sustainable control strategies as different tick populations may present differences in vector competence, acaricide resistance and rates of infectivity with *T. parva*. Thus, it would be important to establish how the biology of the two major halogroups compares in regards to acquisition and transmission of ECF. Further diversity analysis of *R. appendiculatus* using micro- and minisatellite markers can establish if agro-ecological factors such as type of cattle breed, livestock farming systems practiced and nature of tick control method influence its population. This information can be used to identify tick population changes associated with acaricide resistance as well as areas where transmission cycles between domestic and maintenance hosts overlap, which helps to measure the risk of tick reinvasion and thus allows more targeted integrated tick control management.

The markers described herein will be useful for defining parameters of relevance to the epidemiology of ECF in the field. This includes possible interactions between tick and pathogen genotypes, specifically whether differences exist in levels of *T. parva* challenge at different geographic locations that may relate to tick genotypes, as has been demonstrated in laboratory experiments (Ochanda *et al.,* 1998), which might be mediated directly by vector competence factors at the tick/parasite interface or indirectly via differences in adaptation to different environments affecting tick densities. The markers can also be applied in other molecular studies of *R. appendiculatus* such as standardization of the production of the live *T. parva* vaccine and in defining parameters of relevance to the epidemiology of ECF in the field. Since EST-based markers are more transferable between species than genomic markers (Gong *et al.,* 2010; Mishra *et al.,* 2011) they can be used to perform inter-species population genetics analysis as has been done in some tick species (Leo *et al.,* 2012), Tephritidae species (Stratikopoulos *et al.,*, 2009), birds (Primmer *et al.,* 1996), mosquitoes (Kamau *et al.,* 1999) and shrimps (Perez *et al.,* 2005).

Other future researchable areas include attempts to understand the biology underlying the variable contribution of markers to population and species differentiation. The ability of markers to discriminate at higher resolution within the rhipicephaline complex could also ultimately provide new insights into the distribution of different tick species in Africa and perhaps even result in the identification of novel tick vectors of *T. parva.* Designing effective control strategies for ECF will depend on understanding whether recurrent infestations are due to residual domestic tick populations that survive spraying with acaricides or to reinvasion of ticks from alternative hosts. Successful eradication of *R. appendiculatus* in South Africa (Gutsche, 1979; Cranefield, 1991) and *R. annulatus* and *R. microplus* from USA (George, 1989) relied on identification and isolation of surviving population foci which were targeted for tick control. This study has also provided

protocols and characterized molecular markers for any future work involving genetic diversity and population structure of *R. appendiculatus* and closely related rhipicephaline species.

6.1 Conclusions

In conclusion, the *R. appendiculatus* Gene Index (RaGI) EST database and BAC sequences contain polymorphic micro- and minisatellite markers. The markers are able to discriminate *R. apppendiculatus* from other rhipicephaline species. They can be used for inter-species population genetics analysis.

Genotypes of individual field *R. appendiculatus* ticks are highly admixed and the genetic and allelic diversity is high in field-collected ticks compared to laboratory-maintained stocks. There is a very weak genetic structure/substructure and differentiation among field ticks. Based on the 29 polymorphic micro- and minisatellite markers, there are two main clusters of *R. appendiculatus* in Kenya.

COI is a powerful genetic marker for discriminating between closely-related *Rhipicephalus* tick species. The COI divergence within the *R. appendiculatus* species is about 3% or less while the divergence in other rhipicephaline species is about 16%. Based on this gene, *R. appendiculatus* in Kenya is characterized by the existence of two distinct highly differentiated genetic haplogroups which lack geographic or ecological structuring. One of the haplogroups is experiencing recent population expansion. COI and 12S gene regions are better gene markers for studying intra-species diversity of closely related rhipicephaline tick species while ITS2 region is useful for inter-species analysis.

6.2 Reccommendations and future work

This study recommends that;

1. There is need to understand the biology underlying the variable contribution of different markers to population and species differentiation.
2. The four markers shown to be under selection should be evaluated to establish if they are linked to any trait or region of the *R. appendiculatus* genome under selection.
3. It would be important to establish whether particular tick genotypes select particular genotypes of *T. parva* and how the biology of the two major halogroups compares in regards to acquisition and transmission of ECF.
4. Factors contributing to the observed segregation of *R. appendiculatus* populations should be determined. Based on the COI analysis, it would be important to establish the alternative factors that have influenced the genetic differentiation observed.
5. Since the FST value between the two *R. appendiculatus* haplogroups is highly significant, it is important to determine if the tick exist as a cryptic species complex.
6. The true species identity of ticks morphologically identified as *R. praetextatus*, *R. pravus* and an unknown species from Ruma National Park needs to be confirmed.

6.3 References

Avise, J. C. (1991). Ten unorthodox perspectives on evolution prompted by comparative population genetic findings on mitochondrial DNA. *Annual Review of Genetics,* 25: 45 - 69.

Balmer, O., Beadell, J. S., Gibson, W. and Adalgisa, C. (2011). Phylogeography and Taxonomy of Trypanosoma brucei. *PLoS Neglected Tropical Diseases,* **5(2),** e961.

Birky Jr, C. W., Fuerst, P. and Maruyama, T. (1989). Organelle gene diversity under migration, mutation, and drift: equilibrium expectations, approach to equilibrium, effects of heteroplasmic cells, and comparison to nuclear genes. *Genetics,* **121,** 613 - 627.

Blouin, M. S., Yowell, C. A., Courtney, C. H. and Dame, J. B. (1995). Host movement and the genetic structure of populations of parasitic nematodes. *Genetics* **141,** 1007 - 1014.

Cangi, N., Horak, I.G., Apanaskevich, A. D., Matthee, S., das Neves, L C. B. G., Estrada-Peña, A., Matthee, C.A. (2013). The Influence of Interspecific Competition and Host Preference on the Phylogeography of Two African Ixodid Tick Species. *PLOS One.* **8(10),** e76930, 1-12.

Combes, C. (2001). *Parasitism: the ecology and evolution of intimate interactions*. University of Chicago Press.

Cox, A. J. and Hebert, P. D. N. (2001). Colonization, extinction and phylogeographic patterning in a freshwater crustacean. *Molecular Ecology,* **10**, 371–386.

Cranefield, P. F. (1991). Science and empire: East Coast fever in Rhodesia and the Transvaal. Cambridge University Press, Cambridge, UK.

Delaye, C. Beati, L., Aeschlimann, A. A., Renaud, F. and De Meeuis, T. (1997). Population genetic structure of Ixodes ricinus in Switzerland from allozymic data: no evidence of divergence between nearby sites. *International Journal of Parasitology,* **27,** 769-773.

DeSalle, R. and Giddings, L. V. (1986). Discordance of nuclear and mitochondrial DNA phylogenies in Hawaiian Drosophila. *Proceedings of the National Academy of Science of the USA,* **83(18),** 6902 - 6906.

Di giulio, G., Godelieve, L., Morzaria, S., Oura, C. and Bishop, R. P. (2009). Live immunization against East Coast fever-current status. *Trends in parasitology* **25(2),** 85-92.

Dietrich, M., Gómez-Díaz, E., and McCoy, K. D. (2011). Worldwide distribution and diversity of seabird ticks: implications for the ecology and epidemiology of tick-borne pathogens. *Vector-borne and Zoonotic Diseases,* **11(5),** 453-470.

Ellis, J. R and Burke, J. M. (2007). EST-SSRs as a resource for population genetic analyses. *Heredity* **99,** 125–132.

Gadagkar, S. R., Rosenberg, M. S. and Kumar, S. (2005). Inferring species phylogenies from multiple genes: concatenated sequence tree versus consensus gene tree. *Journal of Experimental Zoology B*: *Molecular and Developmental Evolution* **30(4B),** 64 -74.

Gandon, S., and Michalakis, Y. (2002). Local adaptation, evolutionary potential and host–parasite coevolution: interactions between migration, mutation, population size and generation time. *Journal of Evolutionary Biology*, **15(3),** 451-462.

Gong, Y., Xu, S. C, Mao, W., Hu, Q., Zhang, G., Ding, J., Li, Y. (2010). Developing new SSR markers from ESTs of pea (*Pisum sativum* L). *Journal of Biomedical Biotechnology*, 11(9), 02-707.

Greischar, M. A., and Koskella, B. (2007). A synthesis of experimental work on parasite local adaptation. *Ecology Letters*, **10(5),** 418-434.

Gutsche, T. (1979). There was a Man: the Life and Times of Sir Arnold Theiler K.C.M.G. of Onderstepoort. Howard Timmins, Cape Town, South Africa.

Kamau, L., Mukabana, W. R., Hawley, W. A., Lehmann, T., Irungu, L. W., Orago, A. A. S., Collins, F. H. (1999). Analysis of genetic variability in *Anopheles arabiensis* and *Anopheles gambiae* using microsatellite loci. *Insect Mol. Biol*. **8(2),** 287-298.

Kamau, L., Skilton, R. A., Odongo, D. O., Mwaura, S., Githaka, N., Kanduma, E., Obura, M., Kabiru, E., Orago, A., Musoke, A. and Bishop, R. P. (2011**).** Differential transcription of two highly divergent gut-expressed Bm86 antigen gene homologues in the tick *Rhipicephalus appendiculatus* (Acari: Ixodida). *Insect Molecular Biology* **20(1),** 105–114.

Katzer, F., Ngugi, D., Oura, C., Bishop, R. P., Taracha, E. L., Walker, A. R., & McKeever, D. J. (2006). Extensive genotypic diversity in a recombining population of the apicomplexan parasite *Theileria parva*. *Infection and immunity*, **74(10**), 5456-5464.

Laloë, D., Jombart, T., Dufour A.B. and Moazami-Goudarzi, K. (2007). Consensus genetic structuring and typological value of markers using multiple co-inertia analysis. *Genetics Selection Evolution*, **39(5),** 545-567.

Leo, S. S. T., Davis, C. S. and Sperling, F. A. H. (2012). Characterization of 14 microsatellite loci developed for *Dermacentor albipictus* and cross-species amplification in *D. andersoni* and *D. variabilis* (Acari: Ixodidae). *Conservation Genetics Resources*, 4(2), 379-382.

Madder, M., Speybroeck, N., Brandt, J., & Berkvens, D. (1999). Diapause induction in adults of three *Rhipicephalus appendiculatus* stocks. *Experimental & applied acarology*, **23(12),** 961-968.

Madder, M., Speybroeck, N., Brandt, J., Tirry, L., Hodek, I., & Berkvens, D. (2002). Geographic variation in diapause response of adult *Rhipicephalus appendiculatus* ticks. *Experimental & applied acarology*, **27(3),** 209-221.

McCoy, K. D., Boulinier, T., Tirard, C., and Michalakis, Y. (2001). Host specificity of a generalist parasite: genetic evidence of sympatric host races in the seabird tick *Ixodes uriae*. *Journal of Evolutionary Biology*, **14(3),** 395-405.

McCoy, K. D., Boulinier, T., Tirard, C. and Michalakis, Y. (2003). Host-dependent genetic structure of parasite populations: differential dispersal of seabird tick host races. *Evolution*, **57(2),** 288-296.

McCoy, K. D., Chapuis, E., Tirard, C., Boulinier, T., Michalakis, Y., Le Bohec, C., Yvon Le Maho, and Gauthier-Clerc, M. (2005). Recurrent evolution of host-specialized races in a globally distributed parasite. *Proceedings of the Royal Society B: Biological Sciences*, **272(1579),** 2389-2395.

Mishra, R. K., Gangadhar, B. H., Yu, J. W., Kim, D. H., Park, S. W. (2011). Development and characterization of EST based SSR markers in Madagascar periwinkle (*Catharanthus roseus*) and their transferability in other medicinal plants. *Plant Omics Journal* **4(3):**154-162.

Mtambo, J., Madder, M., Van Bortel, W., Geysen, D., Berkvens, D., & Backeljau, T. (2007a). Genetic variation in *Rhipicephalus appendiculatus* (*Acari: Ixodidae*) from Zambia: correlating genetic and ecological variation with *Rhipicephalus appendiculatus* from eastern and southern Africa. *Journal of Vector Ecology*, **32(2),** 168-175.

Mtambo, J., Madder, M., Van Bortel, W., Chaka, G., Berkvens, D., & Backeljau, T. (2007b). Further evidence for geographic differentiation in *R. appendiculatus* (*Acari*: *Ixodidae*) from eastern and southern provinces of Zambia. *Experimental and Applied Acarology*, **41(1-2),** 129-138.

Murrell, A.N., Campbell, J.H. and Barker, S.C. (2001). Recurrent gains and losses of large (84–109 bp) repeats in the rDNA internal transcribed spacer 2 (ITS2) of rhipicephaline ticks. *Insect Molecular Biology* 1**0,** 587–596.

Nadler, S. A., Hafner, M. S., Hafner, J. C., and Hafner, D. J. (1990). Genetic differentiation among chewing louse populations (*Mallophaga: Trichodectidae*) in a pocket gopher contact zone (*Rodentia: Geomyidae*). *Evolution*, 942-951.

Nene, V., Lee, D., Kang'a, S., Skilton, R., Shah, T., de Villiers, E., Mwaura, S., Taylor, D., Quackenbush, J., and Bishop, R., (2004). Genes transcribed in the salivary glands of female *Rhipicephalus appendiculatus* ticks infected with *Theileria parva*. *Insect Biochemistry and Molecular Biology*, **34**, 1117-1128

Norval, R.A.I. and Lightfoot, C.J. (1982). Tick problems in wildlife in Zimbabwe. Factors influencing the occurrence and abundance of *Rhipicephalus appendiculatus*. *Zimbabwe Veerinary Journal*, **13,** 11–20.

Norval, R. A. F., Perry, B. D., and Young, A. S., (1992). The Epidemiology of Theileriosis in Africa. Academic Press Ltd London & San Diego

Ochanda, H., Young, A. S., Medley, G. F and Perry, B.D. (1998). Vector competence of 7 rhipicephalid tick stocks in transmitting 2 *Theileria parva* parasite stocks from Kenya and Zimbabwe. *Parasitology* **116,** 539-545.

Odongo, D. O., Ueti, M. W., Mwaura, S. N., Knowles, D. P., Bishop, R. P. and Scoles, G. A. (2009). Quantification of Theileria parva in *Rhipicephalus appendiculatus* (*Acari: Ixodidae*) confirms differences in infection between selected tick strains. *Journal of medical entomology*, **46(4),** 888-894.

Okello-Onena, J., Tukahirwa, E.M., Perry, B.D., Rowlands, G.J., Nagda, S.M., Musisi, G., Bode, E., Heinonen, R., Mwayi, W. and Opuda-Asibo, J. (1999). Population dynamics of ticks on indigenous cattle in a pastoral dry to semi-arid rangeland zone of Uganda. *Experimental & Applied Acarology*, **23,** 79–88

Palmer, M. J., Bantle, J. A., Guo, X. and Fargoxy, W. S. (1994). Genome size and organization in the ixodid tick *Amblyomma americanum* (L.). *Insect molecular biology*, 3(1), 57-62.

Pérez, F., Ortiz, J., Zhinaula, M., Gonzabay, C., Calderón, J., & Volckaert, F. A. (2005). Development of EST-SSR markers by data mining in three species of shrimp: *Litopenaeus vannamei, Litopenaeus stylirostris,* and *Trachypenaeus birdy*. *Marine Biotechnology*, ***7*(5),** 554-569.

Primmer, C. R., Moeller, A. P. and Ellegren, H. (1996). A wide-ranging survey of cross-species microsatellite amplification in birds. *Molecular Ecology* **5(3),** 365-378.

Radley, D. E. (1981). Infection and treatment method of immunization. In Advances in the Control of Theileriosis. A.D. Irvin, M.P. Cunningham & A.S. Young, Eds.: 227-237. Martinus Nijhoff, The Hague.

Speybroeck, N., Madder, M., Thulke, H.H., Mtambo, J., Tirry, L., Chaka, G., Marcotty, T and Berkvens, D., (2004). Variation in body size in the tick complex *Rhipicephalus appendiculatus/Rhipicephalus zambeziensis*. *Journal of Vector Ecology*, **29(2),** 347-54.

Stratikopoulos, E. E., Augustinos, A. A., Pavlopoulos, I. D., Economou, K. P., Mintzas, A., Mathiopoulos, K. D. and Zacharopoulou, A. (2009). Isolation and characterization of microsatellite markers from the Mediterranean fruit fly, *Ceratitis capitata*: cross-species amplification in other *Tephritidae* species reveals a varying degree of transferability. Molecular Genetics and Genomics, 282(3), 283-306.

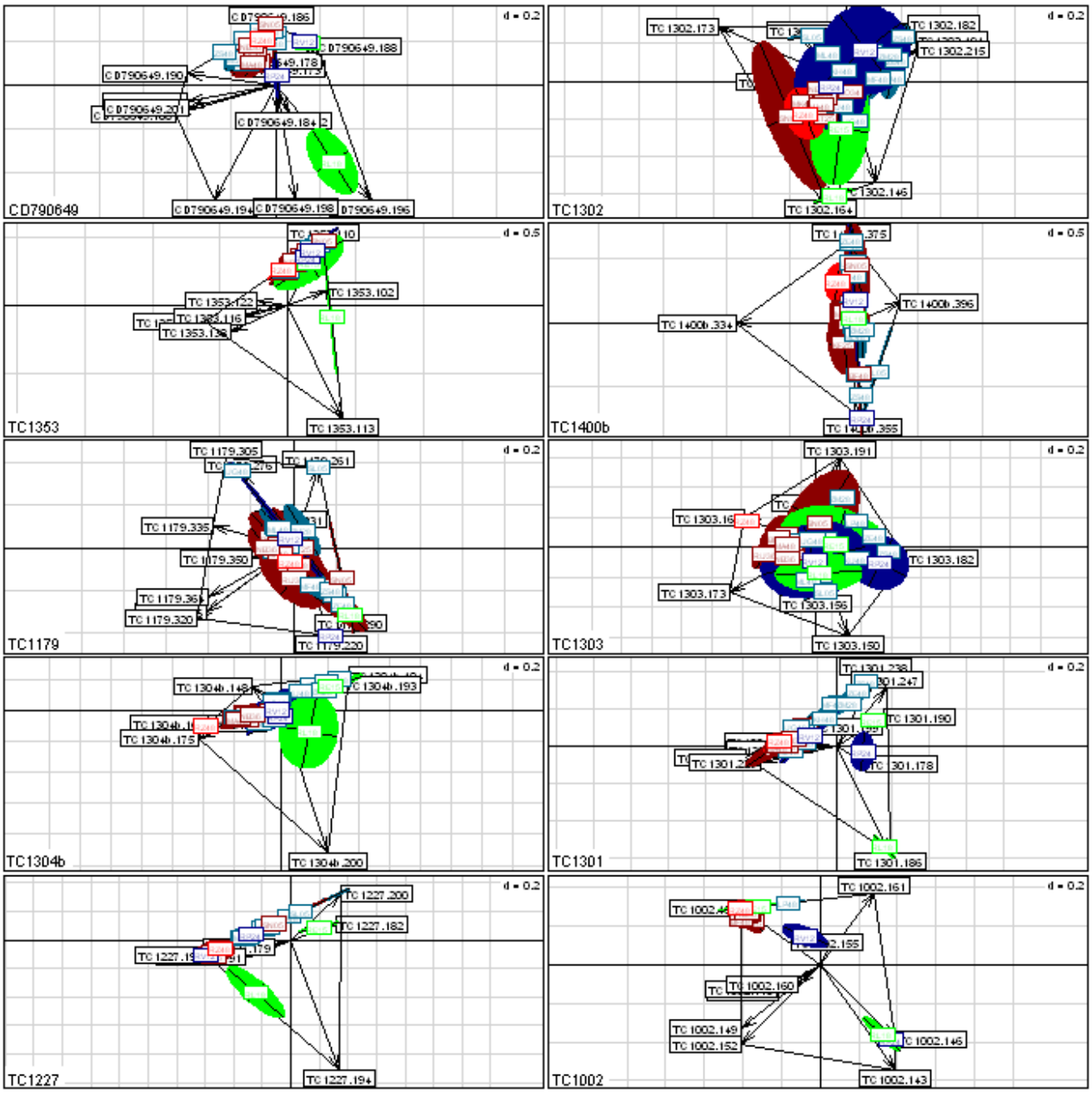
Sunter, J.D., Patel, S.P., Skilton, R.A., Githaka, N., Knowles, D.P., Scoles, G.A., Nene, V., de Villiers, E. and Bishop, R.P. (2008). A novel SINE family occurs frequently in both genomic DNA and transcribed sequences in ixodid ticks of the arthropod sub-phylum *Chelicerata*. *Gene*, **415,** 13-22.

Ullmann, A. J., Lima, C. M. R., Guerrero, F. D., Piesman, J. and Black, W. C. (2005). Genome size and organization in the blacklegged tick, *Ixodes scapularis* and the Southern cattle tick, *Boophilus microplus*. *Insect molecular biology*, **14(2),** 217-222.

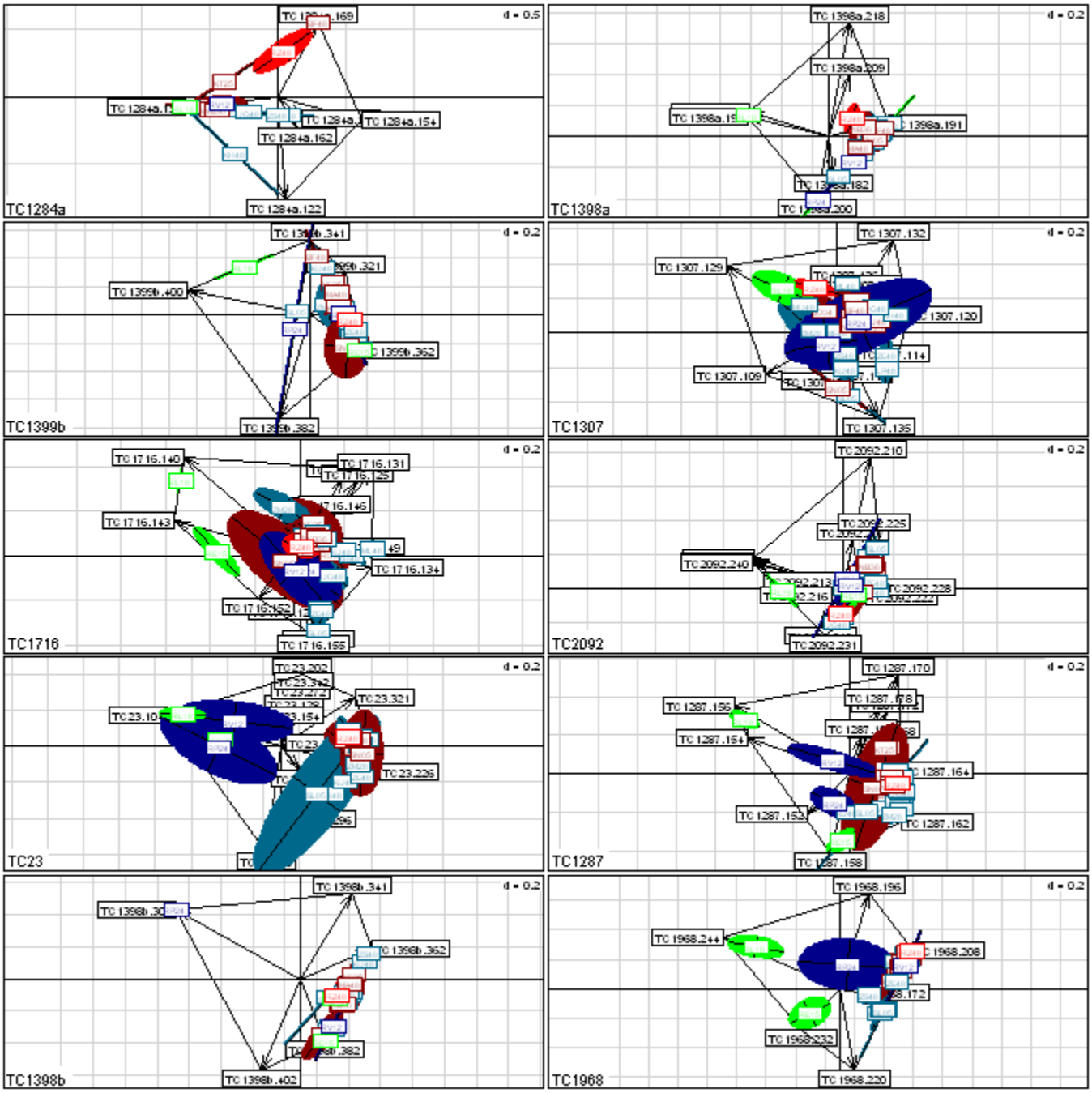
Wares, J. P. & Cunningham, C. W. (2001). Phylogeography and historical ecology of the North Atlantic intertidal. *Evolution* **12**, 2455–2469.

Young, A. S., Dolan, T. T., Mwakima, F. N., Ochanda, H., Mwaura, S. N., Njihia, G. M., Muthoni, M. W. and Dolan, R. B. (1995). Estimation of heritability of susceptibility to infection with *Theileria parva* in the tick *Rhipicephalus appendiculatus*. *Parasitology*, **111**, 31-38.

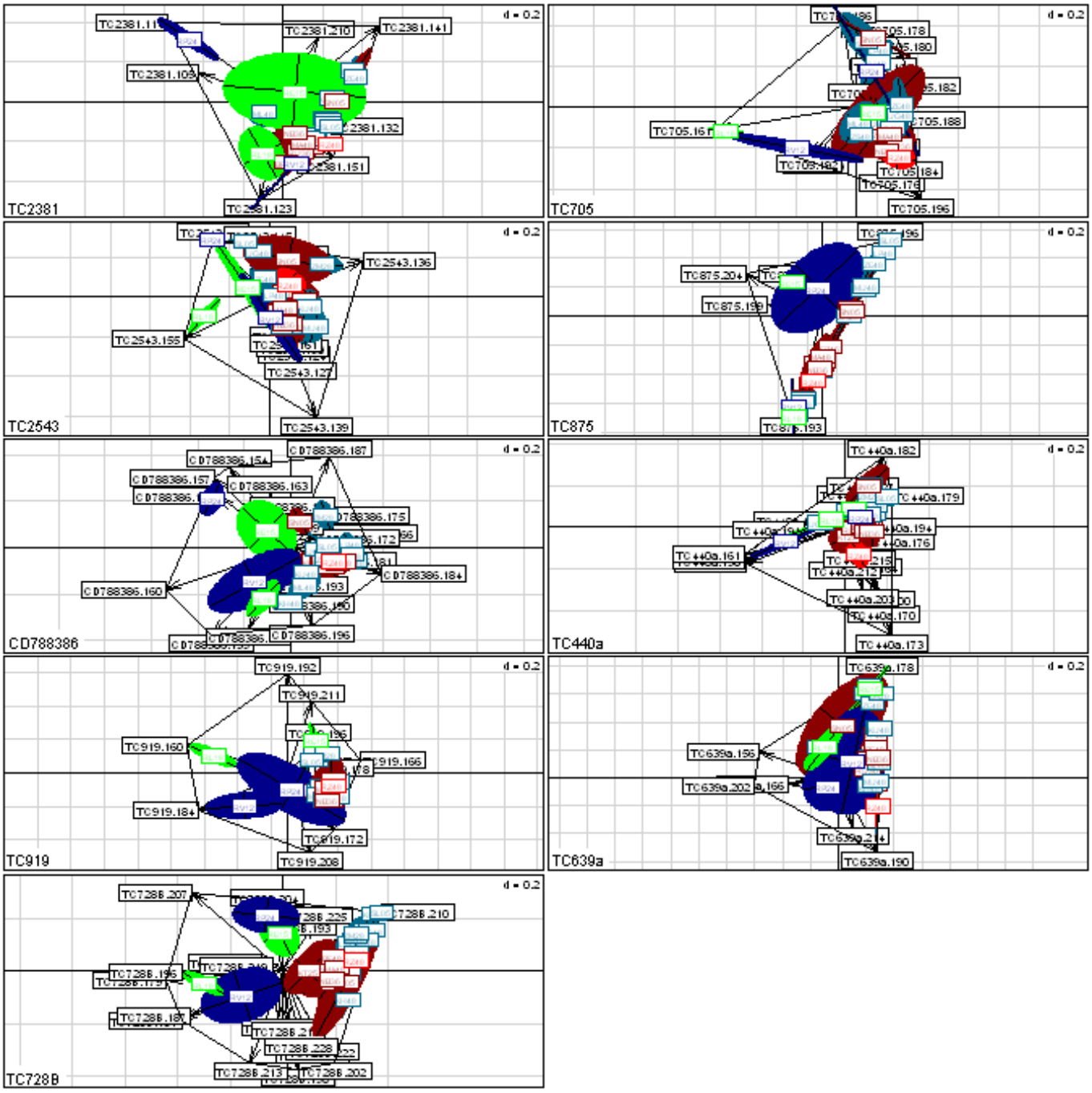
APPENDIX I: SUPPLEMENTARY FIGURES



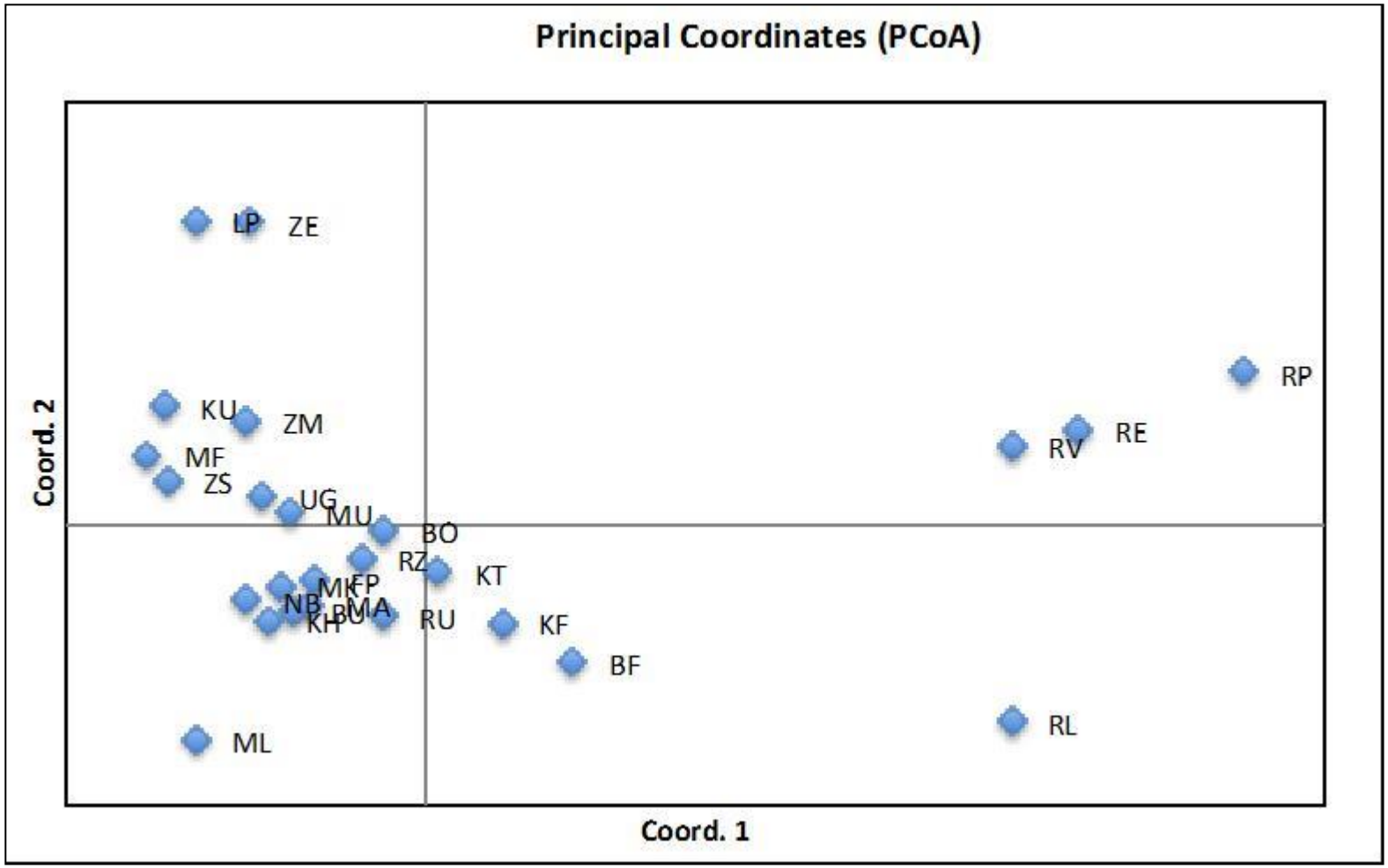
Supplementary figure 1a: Coordinated % PCA's for each of the 10 markers showing how well each marker separates the study population with respect to the common/reference structure
The populations are labeled in their confidence ellipse (P = 0.95), within an envelope formed by alleles (arrows), the most discriminating alleles being joined by lines. The colored ellipses represent the different populations sampled as follows: Red = Lab stock other species (RZ); Dark red = Field ticks (BF, BU, KF, MK, RU, KT, BO, FP, MA, NB, SN); Deep Sky Blue = Lab stocks - *R. appendiculatus* (KH, KU, LP, MF, ML, MU, ZE, ZS, UG, ZM, SL); Dark blue = Field ticks - other species (RP, RV) and Green = Mixed stocks - other species (RE, RL).



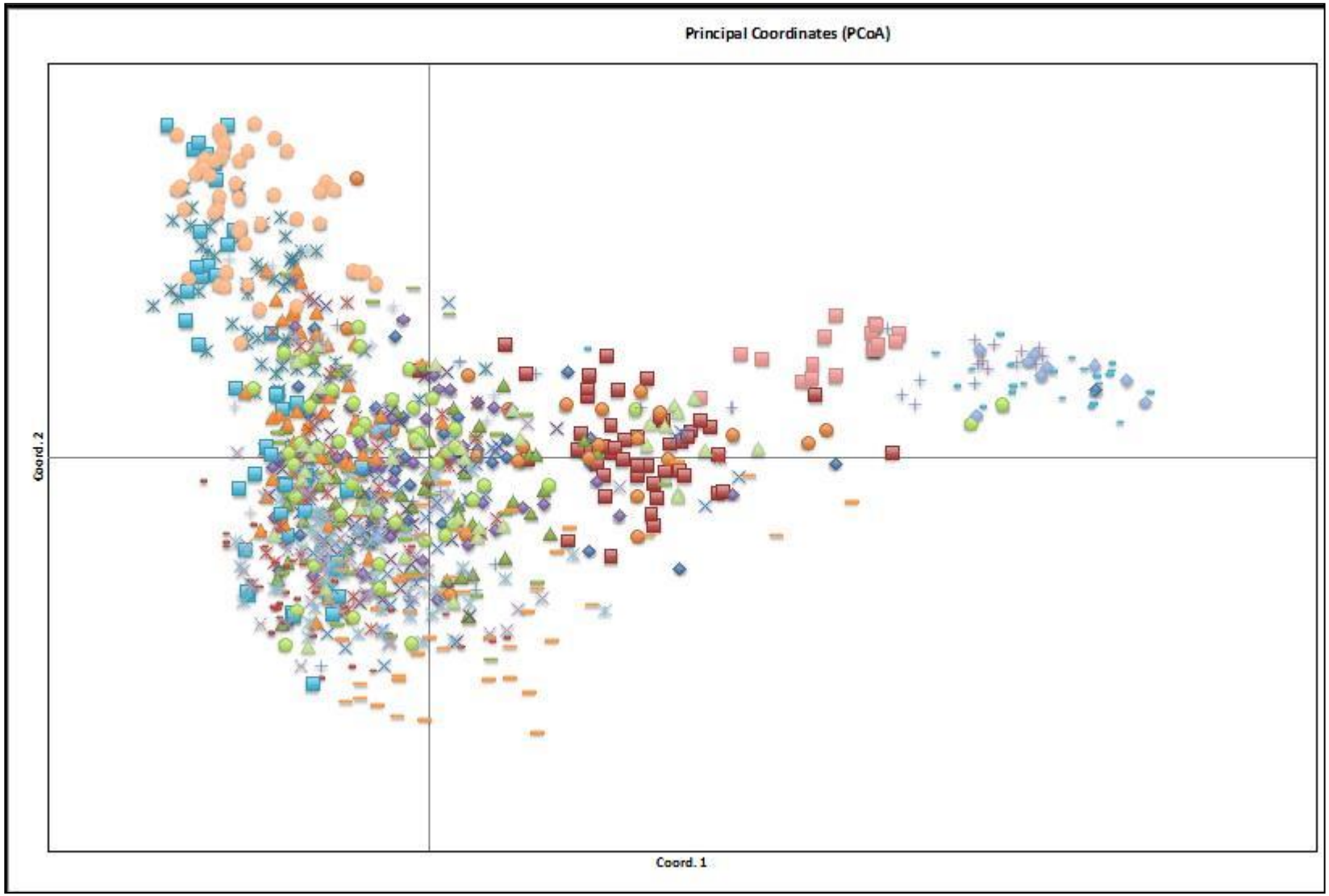
Supplementary figure 1b: Coordinated % PCA's for each of the 10 markers showing how well each marker separates the study population with respect to the common/reference structure
The populations are labeled in their confidence ellipse (*P* = 0.95), within an envelope formed by alleles (arrows), the most discriminating alleles being joined by lines. The colored ellipses represent the different populations sampled as follows: Red = Lab stock other species (RZ); Dark red = Field ticks (BF, BU, KF, MK, RU, KT, BO, FP, MA, NB, SN); Deep Sky Blue = Lab stocks - *R. appendiculatus* (KH, KU, LP, MF, ML, MU, ZE, ZS, UG, ZM, SL); Dark blue = Field ticks - other species (RP, RV) and Green = Mixed stocks - other species (RE, RL).



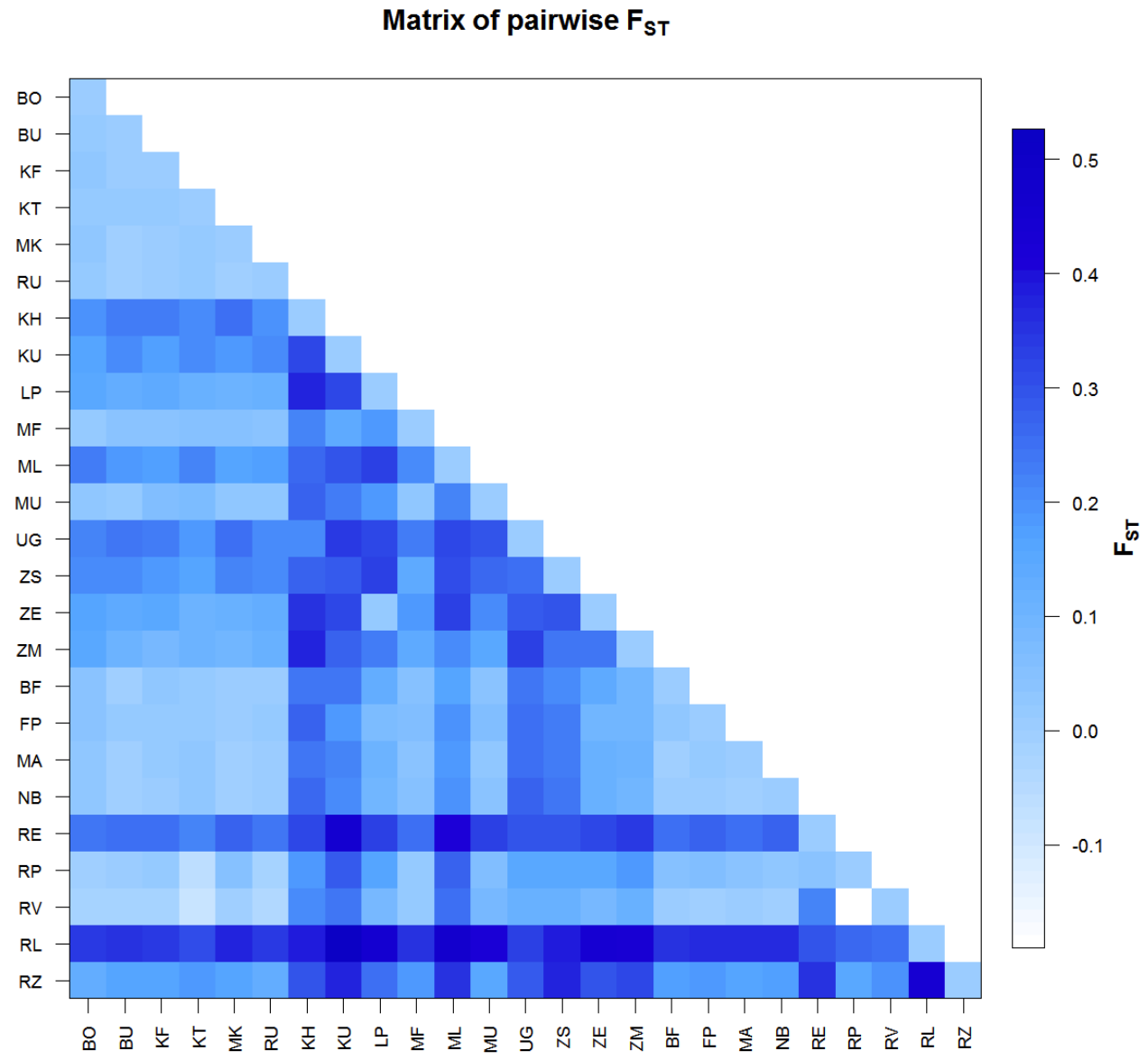
Supplementary figure 1c: Coordinated % PCA's for each of the 9 markers showing how well each marker separates the study population with respect to the common/reference structure
The populations are labeled in their confidence ellipse (*P* = 0.95), within an envelope formed by alleles (arrows), the most discriminating alleles being joined by lines. The colored ellipses represent the different populations sampled as follows: Red = Lab stock other species (RZ); Dark red = Field ticks (BF, BU, KF, MK, RU, KT, BO, FP, MA, NB, SN); Deep Sky Blue = Lab stocks - *R. appendiculatus* (KH, KU, LP, MF, ML, MU, ZE, ZS, UG, ZM, SL); Dark blue = Field ticks - other species (RP, RV) and Green = Mixed stocks - other species (RE, RL).



Supplementary figure 2: **Principle Coordinate Analysis (PCA) of the 25 tick populations analysed in this study**
Most field *R. appendiculatus* tick populations cluster within a central group, with field ticks (NB, FP, MA, MK, BU) clustering closely with KH laboratory stock. *R. zambeziensis* (RZ) clusters with most *R. appendiculatus* populations. However RE, RP, RV, RL which are populations within the other rhipicephaline species are separated from *R. appendiculatus*. The PCA also discriminates laboratory populations ZE, ML, LP.

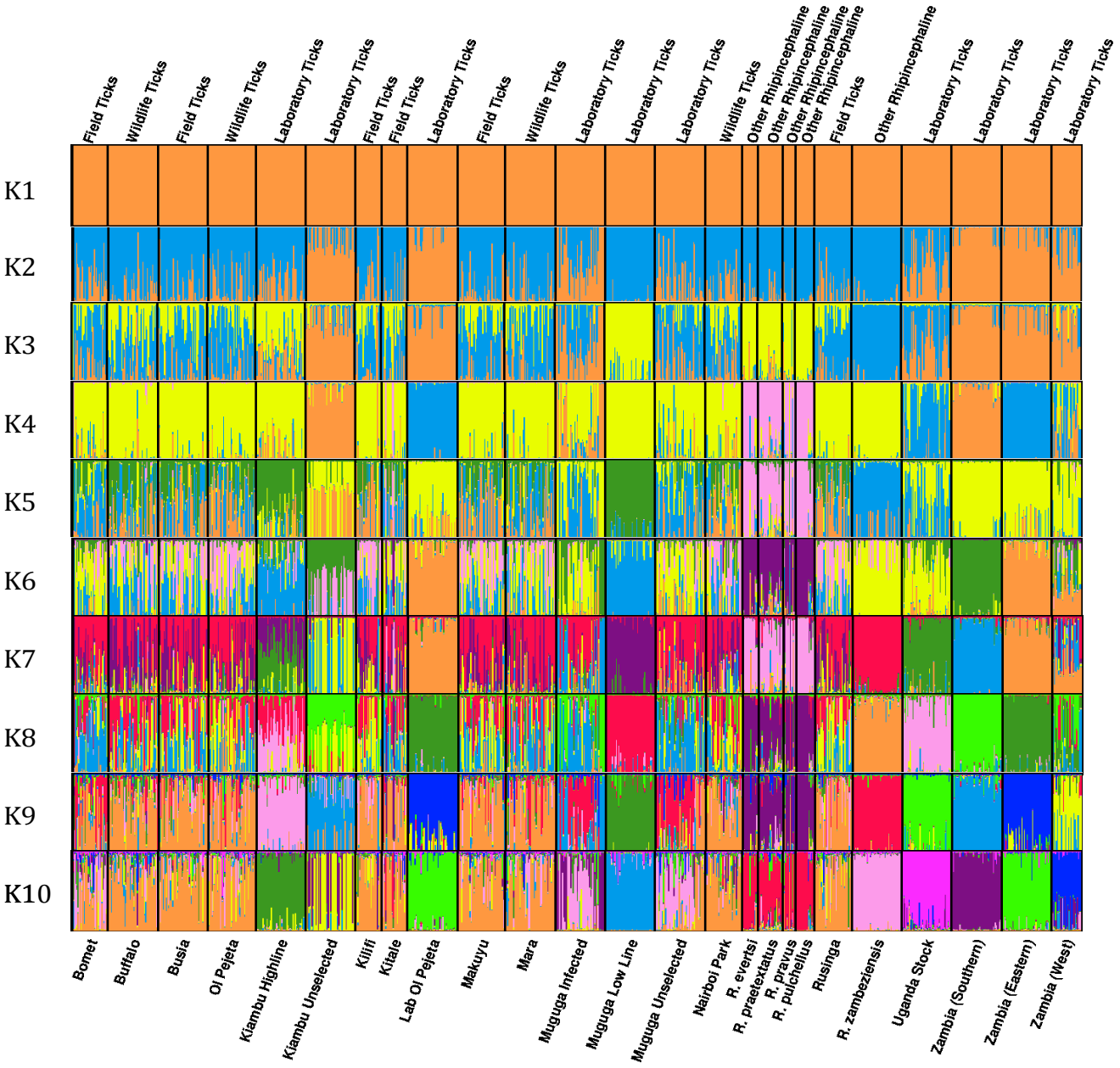


Supplementary figure 3: Principle Coordinate Analysis (PCA) of individual ticks in the 25 Kenyan populations analysed in this study Most *R. appendiculatus* ticks cluster together. Individuals of the standard tick stock (MU) are dispersed through out the PCA. Individual ticks from populations BF, MF, RP, KF, ZM, and FP appear clustered in distict groups.



Supplementary figure 4: A colorimetric graph to ease visualization of pairwise F_{ST} (Table 11, Chapter 4) between populations

Increasing colour intensity corresponds to increasing F_{ST} values. The graph was generated using the R- function in ArlequinV3.5. Pairwise F_{ST} estimates ranged between 0.0003 (between FP and NB populations) and 0.507 (between SL and RL populations). Laboratory bred stocks returned the highest values of F_{ST} between themselves and also between themselves and the field populations. Field populations on the other hand returned the lowest F_{ST} values between themselves with the most closely genetically related populations being NB and FP (Pairwise F_{ST} =-0.0003)



Supplementary figure 5: STRUCTURE bar plots at *K*= 1-10 showing the population structure of the 25 Kenyan tick populations

The ancestry of each population is represented by different colors. The size of the bar plot represents the population sample size. Field ticks high genetic admixture while seven laboratory stocks (LP, KH, ML, UG, ZS, ZE and ZM) were genetically distinct.

APPENDIX II: CONFERENCE ABSTRACTS

1. 12th Kari Biennial Scientific Conference: 8 – 12 November 2010. KARI Headquarters, Nairobi, Kenya

Title: Application of EST-SSR markers in the study of genetic diversity of *R. appendiculatus* (Acari: Ixodidae)

Authors: Esther Kanduma, Robert A. Skilton, Insoters Nzioki, Richard P. Bishop

Abstract

R. appendiculatus, the brown ear tick, transmits *Theileria parva* which causes East Coast Fever (ECF), a fatal lymphoproliferative disease, considered to be the most economically important tick-borne disease of cattle in East, Central and Southern Africa. It is associated with high levels of mortality, especially in exotic and cross-bred cattle. This disease is a major constraint to improvement of livestock production across these parts of Africa.

The objective of this study was to develop and apply molecular markers to estimate the genetic variation or polymorphism in *R. appendiculatus* populations and the genetic relationship between these populations. We evaluated the use of Expressed Sequence Tag –Simple Sequence Repeat (EST-SSR) markers to quantify the genetic diversity within *R. appendiculatus* tick populations from laboratory stocks and from several locations around Kenya. The data will help us understand the genetic diversity of laboratory tick stocks, which are used for the production of live vaccines against ECF and in laboratory experiments where ticks are used to challenge cattle with *T. parva*, and how these laboratory stocks compare with field populations. 938 micro- and minisatellite repeats were identified among 7328 *R. appendiculatus* ESTs. The SSRs were defined by their length which ranged from 1 to 454 bp with an average size of 37 bp. The mean copy number was 6. From about 600 ESTs sequences, 66 primer pairs were designed from flanking regions of SSRs with a

maximum length 30 bp and a copy number equal to or greater than five. As a representative sample, 42 EST-SSR primer pairs were synthesized and used to assess genetic diversity of eight field ticks populations collected from different agro-ecological sites in Kenya and six laboratory bred tick stocks. Out of 42 EST-SSR primer pairs examined, four primer pairs were unamplifiable and another six showed very weak amplification. The discrimination abilities of the remaining 32 EST-SSR primer pairs were investigated by initially genotyping 72 tick samples. Of these, 2 SSRs were monomorphic and another eight showed poor non-specific amplifications. Considerable variation was found at 19 out of 32 EST-SSRs loci analysed and these were validated further by genotyping another set of 96 tick samples. The number of alleles per locus ranged from 3 to 11 with a mean of 6 and a mean observed heterozygosity (*H*o) across all 19 loci of 0.424. The number of genotypes in the 168 samples ranged from 5 to 29 with a mean of 12. The Polymorphic Information Content (PIC) ranged from 0.24 to 0.765, with a mean of 0.529 and mean gene diversity of 0.578. Analysis of resulting data showed that laboratory tick stocks are genetically distinguishable, from other laboratory stocks and also from field collected ticks. However, genetic-relatedness of field collected ticks does not appear to be consistent with geographic origins.

2. 7th Tick and Tick-Borne Pathogens International Conference (TTP 7): 28th August-2nd September, 2011. City Auditorium, Zaragoza, Spain

Title: Polymorphic EST-SSR markers and their application in assessing the genetic diversity of laboratory and Kenyan field populations of the tick *Rhipicephalus appendiculatus* (Acari: Ixodidae)

Authors: Esther Kanduma, Robert A. Skilton, Joram Mwacharo, Inosters Nzuki, Stephen Mwaura, Peter Kinyanjui, Michael Kibe, Richard P. Bishop

Abstract

Rhipicephalus appendiculatus is the main tick vector of *Theileria parva*, a protozoan parasite that causes East Coast Fever, a lethal disease of cattle in East, Central and Southern Africa. It is unclear how genetic diversity of *R. appendiculatus* in the field affects the epidemiology of East coast fever. The genetic diversity and population structure of this important tick vector has not been established, since no useful markers are currently available. . Neither has the diversity of laboratory stocks of *R. appendiculatus* held at the International Livestock Research Institute (ILRI) and how they relate to the genotypes in the field been examined. Here we describe the evaluation and application of Expressed Sequence Tag-Simple Sequence Repeat (EST-SSR) markers to characterize the genetic diversity within both laboratory stocks and Kenyan field *R. appenduculatus* tick populations. Sixty six EST-SSRs were identified from the *R. appendiculatus* Gene Index (RaGI) EST database, primer pairs were designed and optimized for amplifications of these VNTRs s. Twenty nine SSR markers were found to be polymorphic and informative. These were further validated by genotyping 454 tick samples derived from nine field populations, ten laboratory-bred stocks and four additional rhipicephaline species. The number of alleles per locus ranged from 4 to 29 with a mean of 11 and a mean observed heterozygosity (*H*o) of 0.42. The Polymorphic Information Content (PIC) ranged from 0.13 to 0.66, with a mean of 0.45 and mean expected heterozygosity of 0.56. Field ticks showed high diversity and admixing of genotypes. There was a negative correlation between genetic differentiation and geographic distance (r= -0.164). These polymorphic EST-SSR markers will be useful molecular tools for discriminating tick genotypes, analyzing their distribution and determining how they relate to the epidemiology of East coast fever. They can also be used for differentiating *R. appenduculatus* from closely related tick species in the field. The results will be discussed in more detail.

3. 9th University of Nairobi, Faculty of Veterinary Medicine Biennial Scientific conference and Exhibition

Title: Genetic diversity and population structure of *R. appendiculatus* in Kenya.

Authors: E.G. Kanduma, J. M. Mwacharo, N.W. Githaka, P.W. Kinyanjui, J. N. Njuguna, L. M. Kamau, O. Hanotte, R. A. Skilton and R. P. Bishop

Abstract

Rhipicephalus appendiculatus transmits *Theileria parva,* the causative agent of East Coast fever (ECF), the most economically important cattle disease in eastern, central and southern Africa. The genetic diversity and population structure of this important vector remains unknown. To assess the population structure of *R. appendiculatus* in Kenya, this study examined the DNA sequence variation in a 558 bp fragment of the mitochondrial COI gene, a 345 bp fragment of the mitochondrial 12S rDNA and an 1149 bp region of ribosomal nuclear ITS2. Analysis of 332 *R. appendiculatus* COI sequences revealed a total of 28 haplotypes. Genetic diversity indicators based on the COI gene were highest in ticks collected from areas grazed by wildlife compared to those collected from areas grazed by cattle while nine laboratory tick stocks were found to have no gene diversity. Phylogenetic and network relationship inferred from the COI gene sequences showed that field *R. appendiculatus* ticks occurring in Kenya are characterized by the existence of two distinct and highly differentiated genetic haplogroups. These two groupings were not correlated to geographical origin of the samples. Analysis of the ITS2 gene in *R. appendiculatus* revealed one major haplotype while the 12S gene resulted in two major haplotypes that closely correlated with the two major haplotypes differentiated by COI analysis. Mismatch distribution analysis based on the COI gene further revealed that one haplogroup of *R. appendiculatus* may have diverged from the other and has experienced recent sudden population expansion at a much faster rate than its

ancestor. These findings suggest that genetic differentiation of *R. appendiculatus* may not be due to agro-ecological and climatic factors alone as previously thought, and thus identifying other factors may help explain the apparent sudden population expansion observed within the sub-Saharan region. The results may point to an ongoing speciation of the tick in sub-Saharan Africa.