DOCTOR OF VETERINARY SCIENCES

East Coast fever in the African Great Lakes region

Population genetics and epidemiology of Theileria parva and Rhipicephalus appendiculatus

Amzati Sefu, Gaston

Award date:
2020

Awarding institution:
University of Namur

University of Namur
Faculty of Sciences
Department of Veterinary Medicine
Namur Research Institute for Life Sciences (NARILIS)
Integrated Veterinary Research Unit (URVI)

East Coast fever in the African Great Lakes region: Population genetics and epidemiology of *Theileria parva* and *Rhipicephalus appendiculatus*

Gaston AMZATI SEFU

A dissertation submitted in fulfilment
of the requirements for the degree of
Doctor (PhD) in Veterinary Sciences

Namur, March 2020

Supervisors:

Prof. Nathalie Kirschvink (Promotor), University of Namur
Dr. Tanguy Marcotty (Co-promotor), University of Namur

Jury:

Dr. Francesco Renzi (President), University of Namur
Prof. Benoît Muylkens, University of Namur
Dr. Jonathan Marescaux, University of Namur
Prof. Jean-Paul Dehoux, Université Catholique de Louvain
Prof. Luc Duchateau, Ghent University
Prof. Maxime Madder, University of Pretoria
Prof. Dirk Geysen, Institute of Tropical Medicine, Antwerp

Dedication

« Sois fort mon papa Gato, je suis venue en Belgique pour te remonter le moral »

(**Nathalie Aziza Amzati**)

To my lovely wife Esther Sakina Masaki, a woman of strong faith

To my wonderful daughters Nathalie Aziza and Amida Abasi

To my beloved parents Sabiti Amzati and Medi Sophie

To my young brother and best friend Ngton Saidi Amzati

I dedicate this dissertation

You are so fabulous

Acknowledgements

The most important lesson I have learned during this long trip was that it is not possible to complete a PhD work without the assistance and support of many people.

In the first instance, I would like to express my sincere gratitude to my supervisors Professor Nathalie Kirschvink and Dr Tanguy Marcotty for their expertise, understanding and patience giving me outstanding guidance on conducting good scientific research. Thanks for giving me intellectual freedom in my work and demanding high-quality work. You have allowed me to achieve one of my biggest dreams.

I wish to thank my committee and jury members who were more than generous with their expertise. Thank you for your precious time providing me with useful suggestions to improve the manuscript and your patience throughout the entire process.

May I also express my sincere appreciations to all colleagues and members of the department of Veterinary Medicine of the University of Namur for their support in many ways, especially Eric Kazadi, Damien Coupeau, Florian Poulain, Fatima Amarir, Nicolas Burton and Srđan Pejaković for valuable discussions and encouragements. Special thanks to Chantal Ippersiel, Marianne Raes and Laetitia Wiggers for generously helping me solve administrative issues along the PhD period.

A very special thanks goes out to Professor Gustave Mushagalusa, the Rector of the Université Evangélique en Afrique (UEA). Thanks for believing in me, encouraging me and supporting me from the very beginning of my career as an assistant lecturer and always pushing me for a culture of excellence.

I greatly acknowledge the donors of the project: The University of Namur through the UNamur-CERUNA institutional Grant, the Namur Research Institute for Life Sciences (NARILIS), the BecA-ILRI hub through the African Bioscience Challenge Fund (ABCF), the International Foundation for Sciences (IFS), the Agence Universitaire de la Francophonie (AUF) and the Communauté Economique des Pays des Grands Lacs (CEPGL).

Thanks to the co-authors of the publications included in this thesis, notably Esther Kanduma, Appolinaire Djikeng, David Odongo, Jean-Berckmans Muhigwa, Roger Pelle, Herman Nimpaye and Kgomotso Sibeko. Your contribution was valuable and essential. Very special

thanks to Bisusa Muhimuzi, Akilimali Itongwa and colleagues from the CRNS-Lwiro and UEA for technical assistance during data collection, especially a fastidious work looking for ticks in the vegetation.

I would like to thank my friends for being with me and giving me motivation throughout my PhD journey, especially Josué Aruna, Luc Borauzima, Muheshimiwa Mbozi, Silatsa Assongo, Mariette Anoumaa, Justin Buhendwa, Christian Mugisho, Kisse Kisekelwa, Laurent Kibungo, Doux Baraka and many others.

My thanks to many friends and church family for your prayers and kind support throughout the process, especially Pastor Zadi Florent, Pastor Musafiri, Pastor Igunzi and Mama Moise. I will always appreciate all you have done for me.

I extend my thanks to our generous families living in Belgium for always sharing valuable time with me: Mac Mugumaoderha and Lune Maruba, Jean Hamuli and Marie-Jeanne, Jenny Fimpadio and David Baruti and Ariane and Joshua.

Last but not least, I wish to thank all members of my family for their support, patience and love: François Abasi, Marie Amida, Jean de la Fontaine, Saidi Ngton, Espérance Aziza, Rogatien and Abedi Risasi.

To God be the Glory

Table of contents

Abstract

East Coast fever (ECF) is one of the most severe and economically important tick-borne diseases of cattle in the Great Lakes region but also occurs in other regions of the sub-Saharan Africa. The disease is caused by the apicomplexan pathogen *Theileria parva* and transmitted by the ixodid tick *Rhipicephalus appendiculatus.* The Great Lakes region of central Africa is characterised by diversified agro-ecological conditions and extensive cattle movements, where the epidemiology of ECF was previously reported to be complex and unstable: high mortality and high incidence in highlands, but clinical cases confined to calves in lowlands and midlands with reported occasional epidemics in lowlands. The available immunisation approach is based on the use of the Infection and Treatment method (ITM) which involves inoculation of live *T. parva* sporozoites and simultaneous treatment with oxytetracycline. The live vaccination approach provides strong immunity against homologous strains, but variable cross-protection. The efficacy and safety of live vaccine are limited due to the genetic diversity of *T. parva* in field populations and the risk of spreading "foreign" parasite strains in new areas. A better understanding of factors affecting the instability of ECF in the region and epidemiological differences between agro-ecological zones (AEZs) could contribute to the formulation of effective control strategies. In addition, knowledge of the genetic composition of *T. parva* is crucial prior to the use of live vaccine to prevent the risk of spreading exotic parasite variants in new areas.

The aim of this thesis was to assess the population genetics of *R. appendiculatus* and *T. parva* and the transmission and infection dynamics of *T. parva*, as so to further understand the epidemic instability of ECF reported in different agro-ecological zones of the Great Lakes region (based on altitudes: lowlands <1200 m, midlands: 1200-1800 m and highlands: 1600-2800 m). To achieve this objective, we examined the genetic variation and the phylogeography of *R. appendiculatus* populations using mitochondrial genes (*cox1* and 12S rRNA), the genetic and antigenic variation of *T. parva* using antigen-coding genes (*Tp1* and *Tp2*) and the transmission intensity and dynamics of *T. parva* to define the epidemiological states of ECF in contrasting AEZs by means of vectorial inoculation rate (VIR). *Rhipicephalus appendiculatus* ticks and cattle blood samples were collected during cross-sectional surveys in different AEZs of DRC, Rwanda and Burundi. Previously published sequences of *R. appendiculatus* and *T. parva* in their distribution range in Africa were also included in the analyses. The findings have contributed substantially to improve our understanding of the relationships between the genetic

and biogeographical distribution of *R. appendiculatus* and *T. parva* and the transmission dynamics of *T. parva* to understand the complex epidemiology of ECF in the Great Lakes region:

Two major *R. appendiculatus* genetic groups were identified in the Great Lakes region: lineage A (east African group) and lineage B (south African group). Their distribution in Africa strongly correlated with differences in phenotypic features including diapause behaviour, vector competence and body size. The east African lineage (the more diverse and ubiquitous) is widely distributed and has been longer established in the Great Lakes region whereas the south African lineage (less abundant) has settled a founder population from recent colonisation events and its distribution decreases with altitude. The two genetic groups have been subjected to sudden demographic and spatial population expansion associated with cattle movements in Africa. *Rhipicephalus appendiculatus* ticks are more diverse in lowlands than highlands and its biogeographical distribution suggests a sympatric coexistence in central and eastern Africa and allopatric distribution in south Africa.

Population genetic and phylogenetic analyses provided strong evidence of genetic similarity among *T. parva* genetic variants circulating in the Great Lakes region and the components of the Muguga trivalent vaccine. There was high degree of genetic variation within populations and limited agro-ecological structure due to the widespread major genotypes identical or closely similar to vaccine strains. *Theileria parva* populations from lowlands and midlands were more diverse than those from highlands areas. The genetic structure and biogeographical distribution of *T. parva* genetic variants were found to be driven by gene flow facilitated by cattle movement and ecological conditions affecting tick population dynamics. Importantly, the fact that ubiquitous *T. parva* alleles were genetically identical or closely-related to the components of Muguga trivalent vaccine, together with the admixture of *T. parva* populations, provides the evidence for safe use of existing trivalent live vaccine cross-protection field trials in the Great Lakes region.

The agro-ecological and seasonal variations in the transmission intensity of *T. parva* were primarily predicted by the abundance of tick vector rather than the differences in the proportion of infected ticks. *Rhipicephalus appendiculatus* ticks were present on cattle throughout the year and experience at least two generations in the Great Lakes region. The exposure of cattle to *T. parva* infection was found to be significantly different among agro-ecological zones. The prevalence of *T. parva* infection in cattle, the tick challenge on cattle and the transmission

intensity were significantly higher in lowlands and midlands while the highlands areas exhibited lower tick challenge and lower and constant transmission intensity.

Based on these three studies, different epidemiological situations were described according to contrasting AEZs: (i) In the highlands area, where high cattle mortality was previously reported, only *T. parva* isolates identical to the trivalent Muguga vaccine strains occur, associated with the presence of the east African tick lineage and low and constant transmission intensity of *T. parva*. In this area, the situation is suggested to be epidemic; (ii) In midlands and lowlands areas, where previous data reported clinical disease confined to calves, there was high genetic diversity of *T. parva* and *R. appendiculatus*. The two tick lineages were sympatric in these areas and there was high tick challenge and high transmission intensity of *T. parva*. This was an indication of endemic situation in these areas. Particularly, the coexistence of *R. appendiculatus* lineages well adapted to the climatic conditions of the region, together with continuous introduction of ticks through the extensive cattle movement may disrupt the endemicity and lead to occasional epidemics reported in the region. In view of these results, different control strategies should be implemented to be adapted in each epidemiological situation. In endemic areas (lowlands and midlands), a control method using the ITM vaccine on the susceptible animals (calves and exotic) is more indicated whereas strategic tick control during epidemics, treatment of sick animals and cost-effective vaccination of calves should be recommended in the epidemic area (highlands).

Abbreviations and symbols

ABI	Applied Biosystems
AEZ	Agro-ecological zone
AMOVA	Analysis of molecular variance
BecA	Biosciences Eastern and Central Africa
BIC	Bayesian information criterion
BLAST	Basic local alignment search tool
CEPGL	Communauté Economique des Pays des Grands Lacs
CI	Confidence interval
Cox1	Cytochrome c oxidase subunit I
CTL	Cytotoxic T lymphocytes
D	Tajima's neutrality test
DEFT	Root design effects
DNA	Deoxyribonucleic acid
DRC	Democratic Republic of the Congo
ECF	East Coast Fever
EDTA	Ethylenediamine tetra-acetic acid
EMBOSS	European Molecular Biology Open Software Suite
FAO	Food and Agriculture Organization
Fs	Fu's neutrality test
F_{ST}	Wright's fixation index
GALVmed	Global Alliance for Livestock Veterinary Medicine
gDNA	Genomic deoxyribonucleic acid
GLM	Generalised linear model
h	Number of haplotypes
Hd	Haplotype diversity
ILRAD	International Laboratory for Research on Animal Diseases
ILRI	International Livestock Research Institute
ITM	Infection and treatment method
ITS2	Internal transcribed spacer 2
K	Mean number of pairwise nucleotide differences within population

K_{XY}	Mean number of nucleotide differences between populations
MHC	Major histocompatibility complex
MJ	Median joining network
ML	Maximum likelihood
mtDNA	Mitochondrial deoxyribonucleic acid
NJ	Neighbour joining
nPCR	Nested polymerase chain reaction
PCR	Polymerase chain reaction
PIM	Polymorphic immunodominant molecule
PIS	Parsimony informative sites
RFLP	Restriction fragment length polymorphism
RI	Harpending's raggedness index
RNA	Ribonucleic acid
rRNA	ribosomal RNA
S	Number of segregating sites
SD	Standard deviation
SNP	Single nucleotide polymorphisms
SSD	Sum of squared deviations
SSR	Simple sequence repeat
TBD	Tick-borne disease
TCR	T cell receptor
VIR	Vectorial inoculation rate
VNTR	Variable number tandem repeats
π	Nucleotide diversity

List of definitions

1. **Allele (or gene allele)**: One of two or more alternative versions of a gene, having a unique nucleotide sequence.
2. **Allopatric speciation (or geographic speciation)**: Speciation following geographical isolation of two or more populations of a species, restricting the gene flow between the separated subpopulations (reproductive isolation).
3. **Allopatry**: Complete geographic separation of populations within a species
4. **Bottleneck effect**: A form of genetic drift that occurs when a population is rapidly and drastically reduced in size from an ecological crisis or an expansion to new habitat that wipes out most of members within a population, reducing the population genetic variation. The population bottleneck may also be caused by a *founder effect*.
5. **Founder effect (founder event)**: The loss of genetic variation and the consequent change in genotype frequencies that occurs when a few individuals split away and initiate a new population (founder population), with only a small proportion of the genetic diversity from a larger population. The founder effect potentially results in speciation.
6. **Gene flow:** The movement of genes (or individuals) between two or more populations following genetic admixture, usually as a result of migration.
7. **Genetic drift**: The changes in allele frequencies that occur in a population over time due to stochastic processes produced by sampling in finite populations from generation to generation.
8. **Genotype**: The genetic constitution of an organism (allele composition), typically with respect to a set of DNA versions (alleles) found at one or a few genes of interest.
9. **Haplogroup**: A subset of related haplotype sequences that form a phylogenetic clade (also referred to as *lineage*).
10. **Haplotype:** A sequence variant of a gene, used as alternative to “allele”.
11. **Lineage** (see *haplogroup*)
12. **Population**: All the individuals of a particular species that live in a specific geographic area or agro-ecological zone.
13. **Speciation**: Divergent evolution of related individuals or populations of a species, resulting into separate species.
14. **Stock**: A group of related organisms isolated from a common geographical origin within a species.

15. **Strain**: A genetic type or subtype within a species (see also genotype).
16. **Sympatric speciation**: Speciation occurring within populations with overlapping geographic ranges.
17. **Sympatry**: Occurring in the same geographic region (compare with *allopatry)*.
18. **Variant (or antigen variant)**: One of two or more alternative versions of an antigen (protein), having a unique amino acid sequence.

Chapter 1. General introduction and literature review

1.1 Introduction

The Great Lakes region of central Africa consists of the three countries members of the CEPGL (*Communauté économique des Pays des Grands Lacs*) that are closely interconnected: The Democratic Republic of Congo, Rwanda and Burundi. This region of Africa is close to the equator and is characterised by diversified agro-ecological conditions strongly regulated by a wide range of altitudes (700-3000 m), the diverse landscape relief, the protected areas and the water bodies including Lakes Tanganyika and Kivu and the Ruzizi River flowing between the two Lakes. Overall, the climate is tropical with a bimodal rainfall pattern: an early wet season from September to December and a late wet season from February to May. The dry season lasts from June to August with an intervening short dry period of approximately 15 days in January or February. The annual rainfall ranges between 700 and 2000 mm. Livestock are highly valuable and play multiple roles in the agricultural and rural livelihoods and economies in the Great Lakes region. They provide animal products for food, income and employment and contribute to mixed crop-livestock production as valuable source of manure fertilizer (Desiere et al., 2015; Klapwijk et al., 2019). Cattle represent the main livestock species and are considered as capital reserve for smallholder farmers in the Great Lakes region and in the sub-Saharan Africa in general (Herrero et al., 2013). However, cattle population have declined due to high human population demographic pressure, recent violent conflicts and animal diseases (Cox, 2012; Mazimpaka et al., 2017; Wurzinger et al., 2006). Ticks and tick-borne diseases are among the major biotic constraints that affect animal production and hamper its development in this part of the word, causing significant economic losses (Bazarusanga et al., 2007a; Kaiser et al., 1988; Kalume et al., 2013; Randolph and Rogers, 1997). In the context of political unrest during the past three decades, there were increased movements of live animals from neighbouring countries and within countries for trade, breeding and pasture, with very limited veterinary control. In addition, during the pre-colonial period, people immigrated from Rwanda and Burundi and established with their cattle in DRC (Verweijen and Brabant, 2017; Vlassenroot and Huggins, 2005). These unrestricted cross-border movements of cattle together with local transhumance system are thought to have important implications in the epidemiological complexity of tick-borne diseases (TBD) through the spatial spread of non-endemic ticks and pathogens strains, as ticks have potential to establish in new environments (Barre and Uilenberg, 2010; Estrada-Peña and Salman, 2013).

East Coast fever (ECF) is one of the most severe and economically important TBD of cattle in the Great Lakes region but also occurs in other regions of central, eastern and southern Africa (Perry, 2016). It induces large direct production losses and high cattle mortality, as well as indirect economic losses associated with control measures, hindering the development of livestock production in sub-Saharan Africa (Minjauw and McLeod, 2003). The disease is caused by the cell-transforming apicomplexan pathogen *Theileria parva* and transmitted by the three-host ixodid tick *Rhipicephalus appendiculatus* (Bishop et al., 2004). The geographic distribution and population dynamics of *R. appendiculatus* are driven by ecological conditions, host availability and vegetation (Leta et al., 2013; Perry et al., 1990). According to the wider suitable ecological conditions of the tick vector in sub-Saharan Africa, there is evidence that *T. parva* can spread through animal movement and establish in new areas (Boucher et al., 2019; De Deken et al., 2007; Yssouf et al., 2011). In the Great Lakes region, cattle mobility and agro-ecological variability have been suggested to be the main factors underlying the tick vector dispersal and consequently the epidemic instability of ECF (Amzati, 2011; Bazarusanga et al., 2007b; Kalume et al., 2012).

Theileria parva infection in susceptible cattle usually results in a rapid proliferation of schizont-infected lymphocytes, leading to death within two to three weeks in the absence of treatment. It has been demonstrated that cattle that spontaneously recover from an ECF infection develop a long-life immunity against homologous strains. However, these immune animals remain carrier of the pathogen, a source of persistent transmission and spreading of the disease to susceptible naïve cattle (Kariuki et al., 1995; Olds et al., 2018). In areas where agro-ecological conditions are favourable to continuous transmission of the parasite, this phenomenon may contribute to attain a state of endemic stability of ECF (Billiouw et al., 2002; Medley et al., 1993). The immune status observed in recovered cattle enabled the development of an immunisation approach using the Infection and Treatment Method (ITM), which involves inoculation of live sporozoites of *T. parva* and simultaneous treatment with oxytetracycline. A widely used formulation is a trivalent vaccine known as the "Muguga cocktail", which consists of a mixture of three *T. parva* stocks (Muguga, Kiambu-5 and Serengeti transformed) (Radley et al., 1975). This formulation has been successfully deployed on a large-scale in Tanzania, Kenya and Uganda, and on a limited scale in a few other countries (Di Giulio et al., 2009; McKeever, 2007). Immunisation using live vaccine (Muguga cocktail or other local strains) remains the only efficient method which provide a long-lasting and solid cell-mediated

immunity against homologous challenge, but not always full cross-protection against heterologous strains (Radley et al., 1975; Sitt et al., 2015; Taracha et al., 1995a).

However, despite the success of the Muguga cocktail vaccine, which have provided broad protection against ECF in some geographic areas, its widespread deployment is still limited due to the epidemiological complexity of ECF and several concerns which include the possibility of introducing new *T. parva* strains to unvaccinated cattle (Geysen et al., 1999; McKeever, 2007; Oura et al., 2007). It has been demonstrated that exotic *T. parva* strains introduced through vaccination can establish in local tick population and may change the *T. parva* population structure (Oura et al., 2004a; Oura et al., 2007; Skilton et al., 2002). These exotic strains may also undergo sexual recombination with local ones and potentially modify parasite genetic composition (Henson et al., 2012; Katzer et al., 2011). In addition, extensive cross-border movement of cattle may introduce new parasite genotypes and more competent tick stocks to a region (De Deken et al., 2007; Marcellino et al., 2017; Yssouf et al., 2011). On the other hand, the dynamics of population expansion and colonisation ability exhibited by *R. appendiculatus* is probably changing the genetic landscape of *T. parva* in the sub-Saharan Africa (Yssouf et al., 2011). These factors continuously change the transmission dynamics and the population structure of *T. parva* and *R. appendiculatus* which have the potential to significantly impact the epidemiology of ECF and the success of control measures (Katzer et al., 2006; Katzer et al., 2010; McKeever, 2007). The evolutionary dynamics and genetic diversity are main factors that facilitate the survival of *T. parva* and complicate the epidemiology and control of ECF (Katzer et al., 2010). Thus, understanding the complex interplay of genetic diversity and population structure of *T. parva* and its tick vector *R. appendiculatus* associated with agro-ecological conditions is useful to shed more light in the complexity of ECF epidemiology in the Great Lakes region.

In this study mitochondrial genes (*cox1* and 12S rRNA) were used to investigate the phylogeography of *R. appendiculatus* (chapter 3) and antigen-coding genes (*Tp1* and *Tp2*) to characterise the genetic and antigenic composition of *T. parva* (chapter 4). The transmission dynamics of *T. parva* were then assessed in different agro-ecological contexts (chapter 5) to further conciliate the genetic composition and biogeographical distribution of vector and pathogen with the epidemiology of *T. parva*. The results are crucial in understanding the transmission dynamics of *T. parva* as well as its impact on the epidemiological states and integrated management of ECF.

1.2 The tick vector: *Rhipicephalus appendiculatus*

1.2.1 Ixodid ticks: importance, systematic and morphological characteristics

Ticks are among the most important and obligate blood-feeding ectoparasites of livestock found throughout the world, and particularly abundant in the tropical and sub-tropical regions. They are responsible for considerable economic losses worldwide due to their important role as vectors of a range of infectious organisms, including protozoa, bacteria and virus, which are causative agents of tick-borne diseases affecting humans and animals (Jongejan and Uilenberg, 2004). In addition to their role as vectors of harmful infectious agents, tick bites may lead to severe toxic and allergic reactions that sometimes evolve to deadly paralytic symptoms. Heavy tick infestation on an animal can also lead to reduced weight gain and body condition, abortion and reduced milk production (Sonenshine and Roe, 2013).

Ticks belong to the suborder Ixodida (Metastigmata) in the order Parasitiformes, subclass Acari, class Arachnida, subphylum Chelicerata and phylum Arthropoda. The suborder Ixodida consists of three families: the Ixodidae (hard ticks), Argasidae (soft ticks) and Nuttalliedae (an intermediate monospecific family) (Mans and Neitz, 2004; Sonenshine and Roe, 2013). The ixodid ticks or hard ticks account for at least 700 species, among which approximately 200 species occur in the Afrotropical region (Barker and Murrell, 2004; Horak et al., 2002; Nava et al., 2009). The most important genera of hard ticks are *Rhipicephalus*, *Boophilus* (a subgenus of *Rhipicephalus*), *Dermacentor*, *Hyalomma*, *Ixodes*, *Amblyomma* and *Haemaphysalis* (Horak et al., 2002; Murrell et al., 2000, 2001). Their common characteristics are the presence of a hard sclerotised dorsal scutum, the mouthparts projecting forward and the scutum, which covers the entire dorsal body surface in adult males but limited to the anterior half to third dorsal body in females and immature stages (Guglielmone et al., 2014). *Rhipicephalu*s is one of the largest genera of veterinary importance widely distributed throughout the Word, with the majority of species found in Africa (Keirans, 1992; Walker et al., 2005). Representative species of the genus *Rhipicephalus* are recognised by short hypostome and palps, their basis capitulum is typically hexagonal in shape from the dorsal view. These ticks have eyes, festoons and the males have adanal plates. The most common feature is that most species of the genus *Rhipicephalus* do not have coloured pattern on their scutum, so they are conventionally referred to as “brown ticks”. *Rhipicephalus appendiculatus* (Neumann 1901) tick **Figure 1.1**, which is placed into the genus *Rhipicephalus* under the subfamily Rhipicephalinae, is widely distributed in relatively cool and humid ecosystems in sub-Saharan Africa. The tick has a uniform brown

colour, short mouthparts and a hexagonal basis capitulum particularly in female. *Rhipicephalus appendiculatus* transmits the protozoan pathogen *T. parva*, the causative agent of ECF in cattle. It also transmits strains of *T. parva* that cause Corridor disease (buffalo-derived *T. parva*) and Zimbabwean theileriosis (January disease).

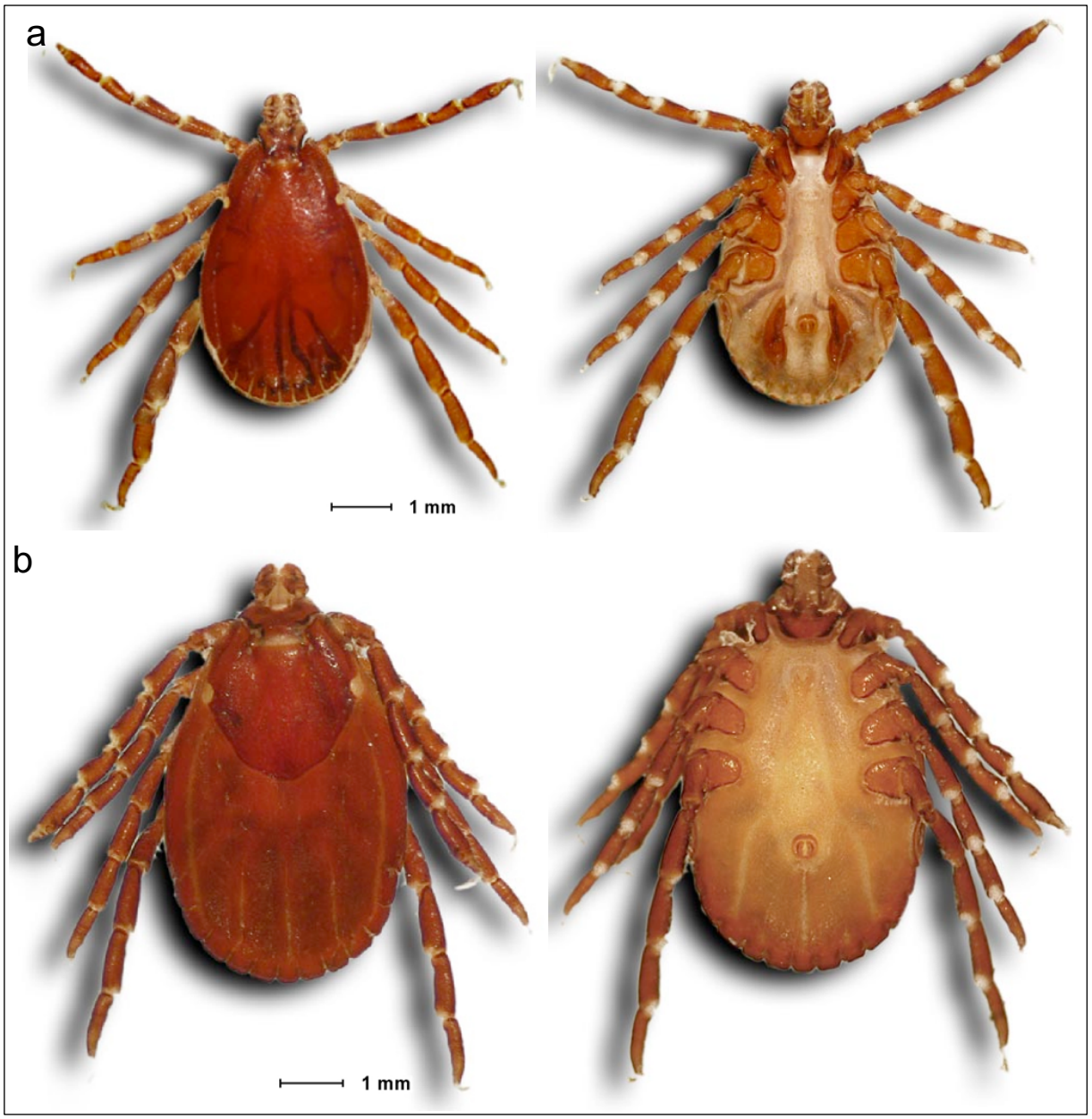

Figure 1.1 *Rhipicephalus appendiculatus* **a)** Male: dorsal (left) and ventral (right) view, and **b)** Female: dorsal (left) and ventral (right) view (Mtambo, 2008)

1.2.2 Host specificity and preference of *Rhipicephalus appendiculatus*

Rhipicephalus appendiculatus occurs in a variety of both domestic and wild ruminants in Africa (Walker, 2003; Walker et al., 2005). Both immature and adult stages of the tick can feed on cattle as the tick has become well adapted to the presence of domestic cattle. It also infests goats, sheep, dogs and wild ruminants including antelopes (*Hyppotragus niger*), waterbucks (*Kobus defassa*), greater kudus (*Tragelaphus strepsiceros*) and elands (*Taurotragus oryx*) (Githaka et al., 2014; Stagg et al., 1994). However, cattle remain the only preferred domestic hosts of all stages of *R. appendiculatus*. The tick is also an ectoparasite of African buffalo (*Syncerus caffer*), which is the main wildlife reservoir of various genotypes *T. parva* (Bishop et al., 2015; Pelle et al., 2011). In all these hosts, the predilection attachment sites of adult ticks are the pinna surfaces of ears and the head, although the tick can be found on other parts of the body during heavy infestation (Jongejan and Uilenberg, 2004). This tick is referred to as "brown ear" tick because of its brown colour and its preference for attachment on the ears of animals. The burdens of adult ticks can be very large exceeding 1,000 adult ticks on cattle and some wild domestic animals (Walker et al., 2005). Immatures are present with less predilection for their feeding sites.

1.2.3 Life cycle of *Rhipicephalus appendiculatus*

The development pattern of *Rhipicephalus appendiculatus* is typically a three-host cycle, during which the tick undergoes different phases according to their behaviour on or off the host: host seeking, feeding and off-host moulting (Sonenshine and Roe, 2013; Walker et al., 2005). Its life cycle comprises four developmental stages: eggs, larvae, nymphs and adults (male and female) (**Figure 1.2**). Larvae, nymph and adult go through an attachment and a free-living period. After repletion and detachment, ticks seek sheltered microenvironments either for moulting (larvae and nymphs) or for ovipositing (adult female). After moulting and in certain cases after quiescence or diapause, the unfed ticks go through a period of hardening, and thereafter they start seeking for suitable host, a process known as "questing" (Sonenshine and Roe, 2013). As soon as the tick has successfully attached on the host, it starts the feeding period, which takes approximatively 4 to 8 days. The feeding period varies with the stage and the temperature. Mating occurs after at least 4 days of the female feeding. Mating is required for the female continues to complete engorgement. Following repletion, the engorged female drops off the host, seek for favourable microhabitats for oviposition and starts ovipositing 5 to 10 days later. Ticks usually drop in host resting areas or drinking points, enhancing the chance of

survival of the ticks off their host (cooler and wetter areas) and facilitating host questing of the next instar (Minshull and Norval, 1982; Mwangi et al., 1991). Within the next 14 days, *R. appendiculatus* female lays a batch of around 3000-5000 eggs. However, the oviposition may be prolonged up to a month if the temperature is low. The egg hatching produces larvae after an incubation period of 20-90 days and the larvae ascend the vegetation in search of suitable hosts. Once on the host, larvae feed to repletion, after which they drop off from their host. The engorged larvae undergo metamorphosis to the nymphal stage in 18 to 21 days. Subsequently, the unfed emerged nymphs' quest for available host, feed and drop off the host. Engorged nymphs undergo metamorphosis to become adults after approximately 21 days. The adults (males and females) then find their suitable host, move to their predilection site on the host (i.e. on the ears) to feed and start a new cycle. The whole life cycle duration from freshly deposited eggs to adults can be completed in three months under most suitable conditions. However, much longer development periods may be expected under various environmental conditions. The main factor affecting the development rate of engorged ticks is the temperature, while their survival and that of unfed ticks is largely driven by the combined effect of temperature and humidity. The survival and development of the free-living stages of ticks are thus governed, to a large extent, by the microclimatic conditions (soil type, vegetation, temperature, humidity) prevailing in the habitats in which the ticks occur (Norval et al., 1992). The seasonal microclimatic variations and cattle mobility have also a strong impact on the drop-off rhythms and the distribution and questing activity of free-living instars of the tick in space and time (Norval et al., 1992; Speybroeck et al., 2002).

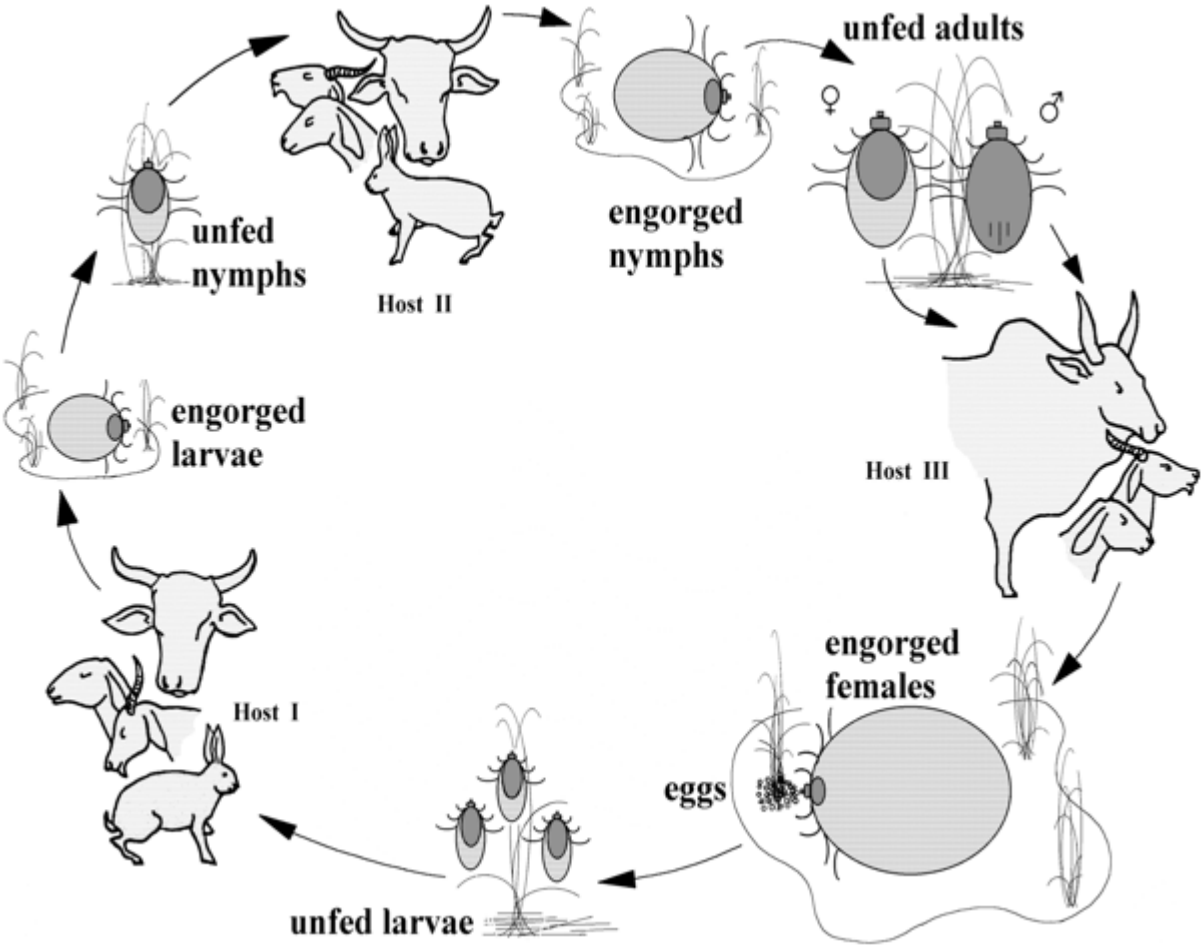

Figure 1.2 Life cycle of the three host ixodid tick *Rhipicephalus appendiculatus*

1.2.4 Geographic distribution, ecology and phenology of *Rhipicephalus appendiculatus*

Rhipicephalus appendiculatus has a wide distribution range confined to sub-Saharan Eastern and Central African countries, including the eastern DRC, Rwanda and Burundi (Bazarusanga et al., 2007a; Kaiser et al., 1988; Kalume et al., 2013; Perry et al., 1991). It has been reported from the equatorial region of east and central Africa to the more tropical region of South Africa and from hot coastal areas to cool highlands plateau where climate remains relatively cool and humid (Leta et al., 2013; Olwoch et al., 2008; Randolph, 2000). Climatic suitability is known to be the main factor that drive its distribution (Bett et al., 2019; Randolph, 2000). The predicted potential distribution of *R. appendiculatus* in Africa based on eco-climatic modelling **(Figure 1.3**), together with ecological studies in the field showed that the tick has a wide distribution range which is strongly influenced by bioclimatic conditions (temperature, humidity, vegetation) and the availability and density of suitable hosts (Berkvens et al., 1998; Cumming, 2000; Leta et al., 2013; Olwoch et al., 2008; Perry et al., 1990; Speybroeck et al., 2002). In actual fact, models indicate large potentially suitable areas in West Africa, where the tick has

not established yet but where severe theileriosis outbreaks could occur if infected ticks were introduced. In addition to climatic conditions, other factors are involved in the distribution and abundance of the tick. These include tick control practises (Vudriko et al., 2018), farming management system and host behaviour (Bazarusanga et al., 2007a; Bazarusanga et al., 2011) and host resistance to tick infestation (Latif et al., 1991). *Rhipicephalus appendiculatus* is found at altitudes ranging from just above the sea level to 2000 m, with annual rainfall varying between 500-2000 mm (Walker et al., 2005). It mainly occurs in areas which have substantial rainfall and moderate temperature (warm and humid) and requires the vegetation composed of both savannah grass, woodland and limited tree cover. Heavy forested areas are not suitable for this tick (Norval et al., 1992; Okello-Onen et al., 1999). Optimal vegetation composition provides favourable microclimatic conditions increasing the survival ability of the free-living stages of the ticks and their questing activity (Lessard et al., 1990). *Rhipicephalus appendiculatus* tends to disappear where the vegetation cover is reduced by overgrazing or human demographic pressure.

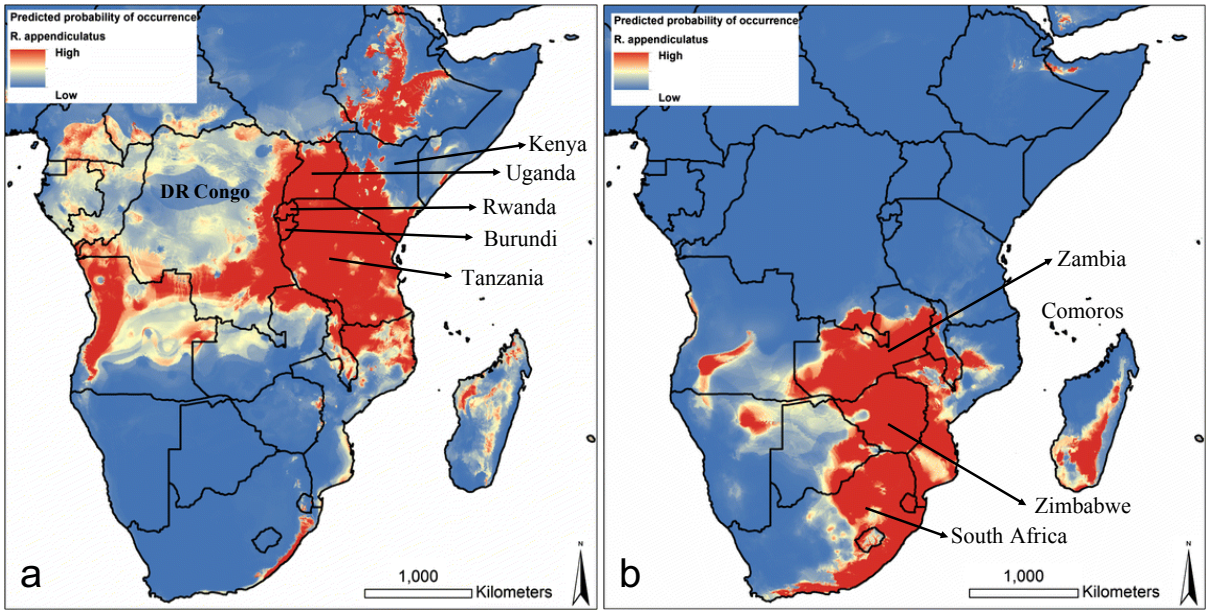

Figure 1.3 Predicted geographical distribution of *Rhipicephalus appendiculatus* based on occurrence data and bioclimatic variables. **a)** Equatorial distribution (eastern and central Africa), and **b)** Tropical distribution (southern Africa) **(**Leta et al., 2013**)**.

Rhipicephalus appendiculatus is expanding its colonisation to new zones in Africa. For instance, it has reached Grande Comore through cattle movements (Boucher et al., 2020; De Deken et al., 2007; Yssouf et al., 2011). Simulation models have explored the potential effect of global change on the dynamics and distribution of ticks (Dantas-Torres, 2015; Estrada-Pena, 2003, 2008). The findings show different rates of range alteration or suitability depending on the projected increased temperature or precipitation in specific areas. The projected increased temperature (from currently cooler to warmer future climate) in humid areas would become more suitable for *R. appendiculatus* to colonise new environments and shifts in its geographical distribution (Leta et al., 2013; Olwoch et al., 2008; Randolph, 2000). Highly suitable conditions occur in the highlands areas of the Horn of Africa (Ethiopia) which have moderate temperature and high precipitation where *R. appendiculatus* has not been reported yet (Leta et al., 2013; Olwoch et al., 2008). In contrast, given that some areas are already hot and dry, the scenario of further increases in temperature would make them unsuitable under the future climate scenarios (Bett et al., 2019).

Population dynamics of ticks are known to vary among agro-ecological zones due to variations in rainfall, temperature vegetation type and photoperiod (Norval et al., 1992; Okello-Onen et al., 1999). The phenology and population dynamics of *R. appendiculatus,* which are the main factors affecting the transmission intensity of *T. parva* (Gachohi et al., 2012), have been well documented throughout its distribution range: in the southern part of Africa (Berkvens et al., 1998; Pegram and Banda, 1990; Speybroeck et al., 2002) and in the eastern and central Africa (Bazarusanga et al., 2007a; Kaiser et al., 1982, 1991; Kaiser et al., 1988; Kalume et al., 2013; Laisser et al., 2017; Okello-Onen et al., 1999; Randolph, 2004). Findings from these studies revealed that the number of generations per year is strongly affected by seasonal variations of climatic conditions. The tick may undergo several overlapping generations per year in more suitable environment conditions. In most parts of the equatorial region of the eastern and central Africa (Burundi, Tanzania, Rwanda, Uganda, DRC, Kenya), *R. appendiculatus* cycles through two to more overlapping generations per year (from bimodal to multimodal phenology). Generally, these areas are characterised by regular rainfall throughout most of the year and no clear pattern of seasonal variation (or short dry seasons). In contrast, in the southern Africa (Zambia, Zimbabwe, South Africa, Grande Comore) and some unlike equatorial regions of central Africa where marked seasonal variations are reported, *R. appendiculatus* completes only a single adult phenology per year with almost no overlap in feeding periods of different instars, except for the fringe area of eastern province of Zambia where a bimodal phenology has been

described (Berkvens et al., 1998). The situation in the Southern Africa leads to behavioural diapause of the adult ticks, which allow the tick to survive harsh conditions during the prolonged dry seasons (Madder et al., 1999; Madder et al., 2002). This diapause behaviour results in seasonal synchronisation of different stages of the tick on cattle.

1.2.5 Intraspecific phenotypic variation in *Rhipicephalus appendiculatus*

There is evidence that environmental conditions affect substantially the phenotypic features of *R. appendiculatus* stocks and shape its geographical distribution. Studies on phenotypic, physiological and ecological characteristics of *R. appendiculatus* provided considerable differences between "east African" and "south African" tick stocks, representing the equatorial and tropical groups of *R. appendiculatus*, respectively. This indicated that the tick features two phenotypes in Africa, largely determined by ecological conditions. The "southern" stock found in South Africa, Southern Zambia and Zimbabwe while the "eastern" stock mainly distributed in eastern Zambia, Burundi, Kenya, Tanzania and Uganda. These tick stocks display substantial differences in their behavioural diapause (Berkvens et al., 1995; Madder et al., 1999; Madder et al., 2002), body size (Chaka et al., 1999; Speybroeck et al., 2004), engorgement weight (Kubasu et al., 2007), ecological plasticity and preferences (Berkvens et al., 1998; Speybroeck et al., 2002) and vector competence (Kubasu, 1992; Ochanda et al., 1998; Odongo et al., 2009). According to Madder et al. (2002), populations of *R. appendiculatus* from different latitudes and climatic conditions exhibit differences in diapause induction and intensity, varying from completely absent to obligate diapause with a significant geographic pattern (**Figure 1.4**). Variations were also reported in body size between geographically different populations of *R. appendiculatus* (Chaka et al., 1999; Speybroeck et al., 2004). Tick populations from the southern stock are larger, express obligatory and high intensity diapause, occur in more tropical areas characterised by prolonged dry seasons and experience univoltine adult phenology. The larger body size of these south African ticks contributes to slower developmental and reproductive cycles and longer period of time between tick instars, but provides strong ability to these ticks to survive harsher environmental conditions (Norval et al., 1992; Speybroeck et al., 2004). In contrast, the eastern stock populations are smaller, do not enter diapause and are present in equatorial areas where ticks cycle throughout the year and undergo bivoltine to multivoltine phenology. These smaller sized ticks cycle faster but have lower survival rates under unfavourable conditions. It is worth nothing that the two phenotypic groups coexist in fringes areas where ecological conditions seem to be intermediate (Madder et al., 2002). Studies

have also reported major variations in vector competence of *R. appendiculatus* stocks for different *T. parva* strains. A southern stock was demonstrated to be more competent than a eastern stock for *T. parva* Katete strains (Tempia, 1997) while the eastern stock was in turn more competent for *T. parva* Muguga strains than a southern stock in Zambia (Ochanda et al., 1998). Factors underlaying these specific interactions observed in the transmission system between tick stocks and *T. parva* strains are not known. However, it was suggested to be associated with geographic origin of involved ticks and *T. parva* stocks and the presence of population-specific proteins in different tick stocks (Baliraine et al., 2000; Ochanda et al., 1998; Wang et al., 2001). In addition, *Katzer* et al. (2006) showed that different tick colonies are also associated with particular genotypes of *T. parva* in the transmission system and impose non-random selective pressure on the parasite. This may have major consequences shifting the population structure of the pathogen in specific zones.

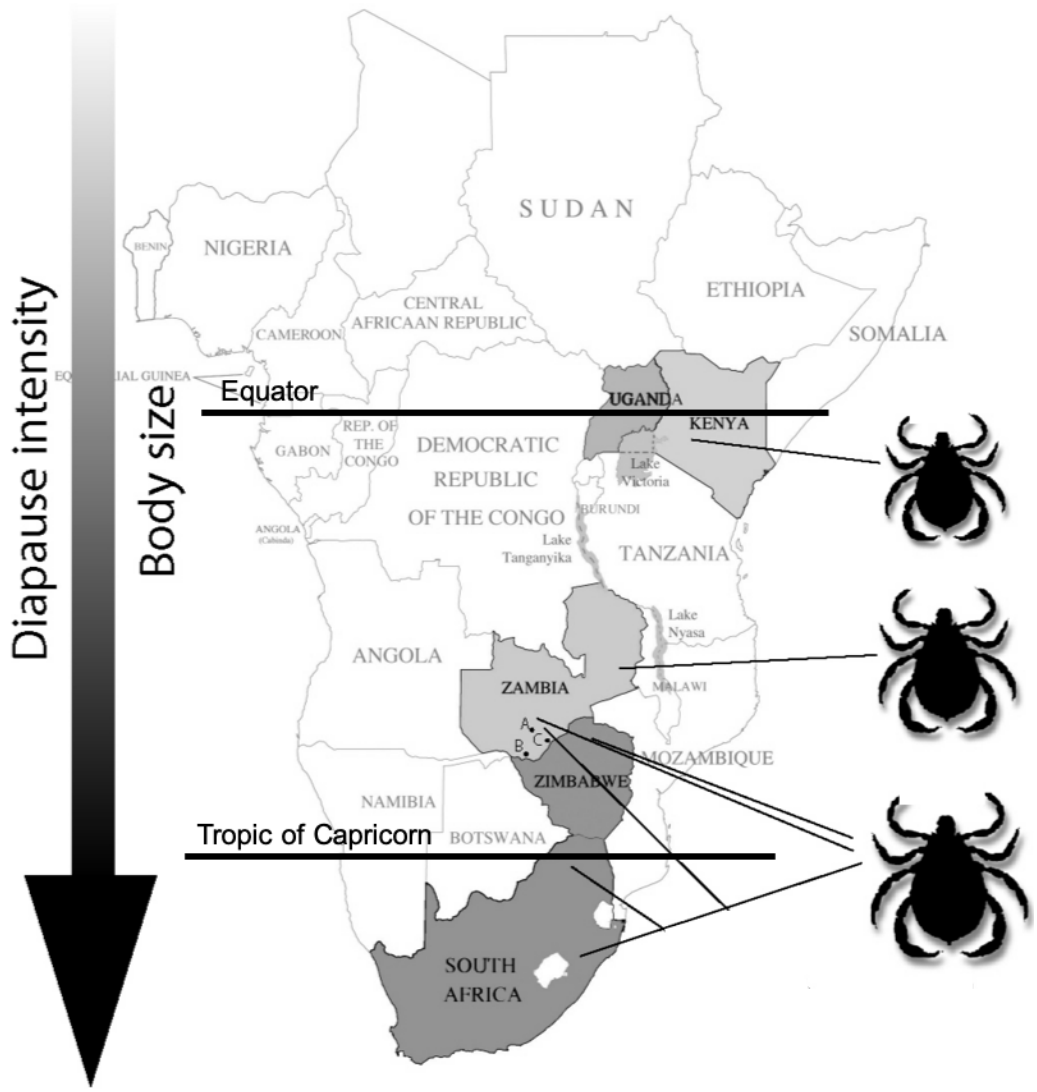

Figure 1.4 Geographical pattern of body size and diapause intensity of *Rhipicephalus appendiculatus* from different sub-Saharan African countries: equatorial and tropical distributions (Madder et al., 2002; Speybroeck et al., 2004).

1.2.6 Intraspecific genetic variation and population structure of *Rhipicephalus appendiculatus*

Intraspecific genetic variations are thought to occur within and between natural populations of tick species. The presence of variation at the DNA sequence level usually underlies phenotypic differences between and within species (Cruickshank, 2002). These variations can have major impact in the population structure of ticks and are thought to affect the epidemiology of the disease they transmit (Araya-Anchetta et al., 2015; Gooding, 1996; McCoy, 2008; Tabachnick and Black, 1995). Tick vectors are subjected to evolutionary forces such as mutation, migration

(promoting gene flow), natural selection, genetic drift and founder events leading to significant changes in the vector population structure and composition (Gandon and Michalakis, 2002; Gooding, 1996). Thus, investigating the genetic structure and diversity is essential in understanding microevolutionary changes in the vector population. Such analyses provide important clues to understand the epidemiology of arthropod-borne diseases for developing integrated control strategies (Gooding, 1996; McCoy, 2008).

Genetic variations can be detected in tick populations using a wide array of molecular markers that can be divided into two major groups: (i) biochemical markers which detect variations at protein (enzyme) level by electrophoresis (Baliraine et al., 2000), and (ii) PCR-based molecular markers using restriction fragment length polymorphism (RFLP) techniques (Lempereur et al., 2010) or direct DNA sequencing methods (Coimbra-Dores et al., 2018; Kanduma et al., 2016a). PCR-based method includes ribosomal, nuclear and mitochondrial genes and micro and minisatellites (SSR) markers. These markers allow to detect variation at the nucleotide and/or amino acid levels (Araya-Anchetta et al., 2015; Avise, 2012; Cruickshank, 2002; Kanduma et al., 2016b). Some of molecular markers widely used to characterise the genetic variability of *R. appendiculatus* comprise the second internal transcribed spacer region of the nuclear ribosomal gene cluster (ITS2) and mitochondrial genes: the 12S ribosomal rRNA (12S rRNA) and the cytochrome C oxidase subunit I (*cox1*). The choice of molecular markers is crucial and requires a good background of the nature of evolution of the gene and its usefulness for intraspecific or interspecific study. The combination of different markers is strongly recommended to provide deep resolution in the intraspecific phylogeny of ticks (Abouelhassan et al., 2019; Cruickshank, 2002; Dabert, 2006; Lv et al., 2014a). Two mitochondrial genes (12S rRNA and *cox1*) are the most widely used to assess the phylogenetic relationships among *R. appendiculatus* populations and their population structure. The fact that mitochondrial DNA (mtDNA) is maternally inherited, together with its faster evolution compared to nuclear genome make mtDNA a particularly useful tool for studying intraspecific population genetics (Simon et al., 1994). In addition, their high copy number and the small genome size of mitochondria make them easier to use in PCR and direct sequencing. Mitochondrial genes evolve at widely different rates. For instance, the protein coding gene *cox1* evolve faster than the non-coding gene 12S rRNA (Avise, 2000). The cox*1* gene is one of the most conserved mitochondrial protein coding genes and more suitable marker for DNA barcoding (Lv et al., 2014b). The evolution of this gene is rapid enough to provide deeper phylogeographic signal than any other mitochondrial gene (Latrofa et al., 2013; Wang et al., 2019).

The ITS2 nuclear marker have been used previously to resolve taxonomic problems of a wide variety of closely-related taxa and distinguish between populations of Rhipicephaline ticks, including *R. appendiculatus* versus *R. zambeziensis* and *Rhipicephalus (Boophilus) microplus* versus *Rhipicephalus (Boophilus) decoloratus* (Barker, 1998; Baron et al., 2018; Mtambo et al., 2007a). The marker was able to distinguish between populations of these tick species originating from diverse geographic regions (Barker, 1998). It also confirmed the taxonomic status of *R. appendiculatus* and *R. zambeziensis* which are morphologically very similar species of ticks from Zambia (Mtambo et al., 2007a). Therefore, the nuclear ITS2 marker is shown to be very informative for interspecific variation but lacks the resolution for intraspecific variation. Thus, it should be useful for interspecific phylogeny of closely similar species rather than intraspecific analysis.

Mitochondrial genes, 12S rRNA and *cox1* have been widely used in molecular studies of intraspecific variation in tick species as well as to resolve relationships among recently diverged ticks lineages (Murrell et al., 1999). The two markers provided evidence of geographical and genetic differentiation among *R. appendiculatus* stocks from eastern and southern provinces of Zambia (Mtambo et al., 2007b; Mtambo et al., 2007c). The geographic substructuring of *R. appendiculatus* genetic groups appeared to correlate with differences in phenotypic characteristics (body size and diapause) (Madder et al., 2002), adult phenologies (Berkvens et al., 1998; Speybroeck et al., 2002) and epidemiology of ECF between the two Zambian provinces (Billiouw et al., 1999; Billiouw et al., 2002; Mulumba et al., 2000). Extended analyses of *R. appendiculatus* populations from east Africa showed that the two lineages previously characterised where sympatric in Kenya, with the eastern lineage being more abundant than the southern lineage (Kanduma et al., 2016a). Another study reported that the tick recently introduced in Comoros island from Tanzania belonged to the southern mitochondrial lineage (Yssouf et al., 2011). These phylogenetic findings strongly correlate with phenotypic, ecological and physiological characteristics of the ticks and allow to confirm the occurrence of two diverged populations of *R. appendiculatus* in Africa. Surprisingly, the southern lineage appears to be shifting its ecological zone and settling in east Africa (Kanduma et al., 2016a; Yssouf et al., 2011).

1.3 The pathogen: *Theileria parva*

1.3.1 *Theileria parva*: classification, hosts and transmission

Theileria parva is an obligate intracellular protozoan pathogen, the causative agent of East Coast fever (ECF) in cattle, which was first identified by Sir Arnold Theiler in 1900 (Theiler, 1912). It is the most pathogenic and economically disastrous tick-borne pathogen of cattle in eastern, central and southern Africa (Mukhebi et al., 1992). The parasite infects both cattle and wild ruminant such as African buffalo (*Syncerus caffer*) and waterbuck (*Kobus ellipsiprymnus*) (Bishop et al., 2004; Pienaar et al., 2011; Stagg et al., 1994). African buffalo is the natural reservoir of *T. parva*, though it does not develop the disease and remains healthy carrier of multiple *T. parva* genotypes (Pelle et al., 2011; Sitt et al., 2018). *Theileria parva* is transstadially transmitted by its main vector, the ixodid tick *R. appendiculatus*, but it can also be transmitted by *Rhipicephalus zambeziensis* in southern Africa and *Rhipicephalus duttoni* in Angola (Norval et al., 1992). Based on clinical signs, severity and lethality of the infection and epidemiological characteristics, the disease caused by *T. parva* is usually described in three difference forms: (i) East coast fever and (ii) January diseases or Zimbabwe theileriosis caused by cattle-derived *T. parva* and (iii)corridor disease caused by buffalo-derived *T. parva.*

Theileria parva belongs to the kingdom Protista, subkingdom protozoa, phylum Apicomplexa, class Sporozoa, subclass Piroplasma, order Piroplasmida, family Theileriidae and genus *Theileria*. The genus *Theileria* comprises a number of cell-transforming pathogens that affect domesticated ruminants, such as *T. lestoquardi* and *T. taurotragi,* which cause diseases in sheep and *T. annulata* which is responsible for tropical theileriosis occurring in northern Africa, southern Europe and Asia. Other species members of the genus *Theileria* also affect domesticated animals and cause benign theileriosis in cattle. These include *Theileria velifera*, *Theileria orientalis*/*buffeli* and *Theileria mutans*. The species of the *Theileria* genus are phylogenetically closely related to species of the *Babesia* genus, and both fall in the order Piroplasmida and the phylum Apicomplexa (Lack et al., 2012; Morrison, 2009).

1.3.2 The life cycle of *Theileria parva*

The life cycle of *T. parva* is illustrated in **Figure 1.5**. Understanding the life cycle of *T. parva* provides useful information on the interface between population genetic dynamics, recombination processes and other mechanisms underlaying its evolution and diversity. The three-host tick *R. appendiculatus* transmits the parasite transstadially. There is no transovarian

transmission of *T. parva*. The parasite undergoes a complex life cycle involving different developmental stages, which include an obligate sexual phase in the tick and asexual multiplications occurring both in the bovine and the tick hosts (Jalovecka et al., 2018; Nene et al., 2016). The cycle involves two intra-cellular stages in the bovine host, the schizont which transforms lymphocytes and the intra-erythrocytic merozoite (piroplasm) stage, which is infective for the tick.

Infected nymph and adult ticks transmit the parasite as infective sporozoites to cattle during their blood-feeding. The sporozoites, once injected into the cattle, rapidly invade bovine T and B lymphocytes where they differentiate into intracellular schizonts and induce infected cells to undergo major changes (Dobbelaere and Heussler, 1999; Jalovecka et al., 2018). The schizonts become associated with the mitosis apparatus of the lymphocytes and divide in synchrony with the host cells, giving rises to an uncontrolled clonal expansion of infected cells. This enables the parasite to remain intracellular while proliferating (Shaw, 2003). This stage of the parasite is associated with cancer-like host cells transformation and is responsible for the pathology of the disease. A certain proportion of schizonts later undergo merogony after which they differentiate to produce merozoites, that upon released by host cell rupture, invade the erythrocytes (Gauer et al., 1995). Thereafter, the merozoites mature into intra-erythrocytic gametocytes, which are infective to ticks (Dobbelaere and Heussler, 1999). Tick larvae or nymphs become infected by ingesting bovine erythrocytes containing gametocytes when feeding on infected cattle or buffalo and transmit the infection to susceptible animal hosts during their subsequent post-moult instar (nymph or adult, respectively). The gametocytes differentiate into gametic stage inside the erythrocytes after being ingested by the tick and are released in the tick gut, where male and female gametes fuse to generate a zygote. The resulting diploid zygotes formed by syngamy invade the gut epithelial cells of the tick which then develop into motile kinetes, the final stage of the sexual reproduction. During the metamorphosis of the tick host, kinetes escape the gut cells and migrate to the salivary glands where they form sporoblasts, resting until the tick host starts feeding. During subsequent sporogony, meiotic division occurs and gives rise to haploid cells that result in the production of sporozoites, the infective stage for the bovine host (Gauer et al., 1995). The meiotic reduction division may involve genetic crossover events in the tick (Gauer et al., 1995; Katzer et al., 2006). The release of infective sporozoites from salivary glands occurs within a period of 4-8 days after the beginning of the tick meal (Shaw and Young, 1995).

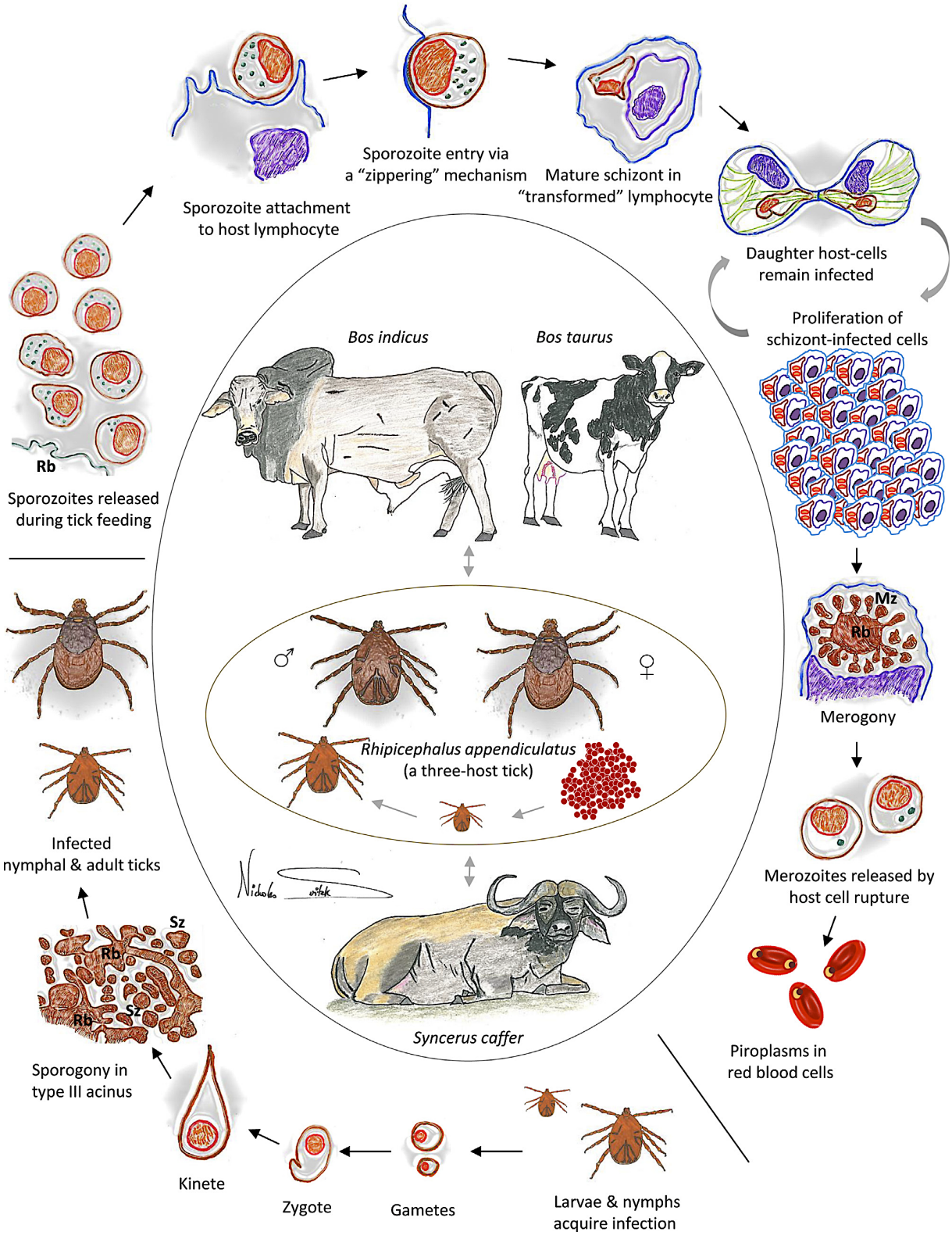

Figure 1.5 *Theileria parva* life cycle in cattle and the tick *Rhipicephalus appendiculatus* (Nene et al., 2016). *Developmental stages are not drawn to scale.*

1.3.3 Epidemiology of East Coast fever

1.3.3.1 Epidemiological states of East Coast fever

The epidemiological states of ECF usually vary among agro-ecological zones and are a result of interaction between ecological variables and management practices, tick vector, host and pathogen variables (Gachohi et al., 2012). The vector variables include seasonality of tick instars, tick burden, vector competence and the proportion of infected ticks (Norval et al., 1991). The epidemiological indicators related to host-pathogen interaction include disease incidence, host immunity, antibody prevalence (seroprevalence or seroconversion), case-fatality and mortality (Berkvens, 1991; Norval et al., 1992). These variables influence the epidemiology and control of ECF. Thus, choosing control measures in the context of integrated ECF control depends on the epidemiological state of ECF for the targeted area (Uilenberg, 1999). The epidemiological states of ECF are considered to be: the absence of ECF, the endemic situation and the epidemic situation based on a combination of epidemiological indicators. *Theileria parva* endemicity is defined as the epidemiological state in which a large majority of calves are infected and either die or become immune (Billiouw et al., 2002). This implies that all adult animals are immune. The majority of cattle develop a carrier state in endemic areas, which is defined as the asymptomatic persistence of the parasite in primary infected and recovered animals (Medley et al., 1993). Suitable ecological conditions for the tick that allow continuous exposure of cattle to all instars of *R. appendiculatus* (continuous transmission of the parasite) are thought to contribute to the endemicity. The endemically unstable state describes a state in which the majority of animals of all age classes are susceptible as they were never exposed before in their life. Clinical disease is experienced regardless the age groups with high case-fatality and mortality (Norval et al., 1992). Endemic situations may become epidemic for ecological disruption (unfavourable climatic conditions for tick), constant application of acaricides at high frequency, grazing system that reduce cattle exposure (low infestation rates) and introduction of susceptible animal from disease-free area or exposure of cattle to antigenically different *T. parva* and more competent tick stocks. The epidemic situation normally occurs where animals are kept under low levels of tick challenge for a certain period of time restricting cattle/tick contact (Gachohi et al., 2012; Norval et al., 1992).

1.3.3.2 Factors affecting the transmission dynamics of *Theileria parva*

Research have focused on the study of transmission of *T. parva* in field and laboratory stocks of ticks (Ochanda et al., 1998; Ochanda et al., 1988; Ochanda et al., 1996; Olds et al., 2018;

Young et al., 1983). These studies showed considerable variation in the transmission of *T. parva* among stocks and among instars of the tick (Konnai et al., 2006). According to Ochanda et al. (1996), adults ticks have consistently higher prevalence of *T. parva* than nymphs. Several other factors are believed to influence the ability of the vector to acquire and transmit the parasite. These include agro-ecological and climatic conditions, tick stocks behaviour, the duration of tick attachment, *T. parva* prevalence and level of parasitaemia in cattle. The successful attachment and its duration are associated with increased risk of *T. parva* transmission (Konnai et al., 2007). *Theileria parva* infection rate in adult *R. appendiculatus* ticks feeding on carrier cattle (low parasitaemia) has been proven to be lower than in ticks feeding on cattle with acute infection (high parasitaemia) (Marcotty et al., 2002; Ogden et al., 2003). Agro-ecological conditions affect not only the number of generations per year, but also the tick abundance, diapause behaviour and consequently the transmission potential of *T. parva*. The number of generations and whether or not different instars of the tick occur simultaneously, determine the epidemiological profile of East Coast Fever in a given region (Billiouw et al., 1999; Billiouw et al., 2002). In the equatorial conditions where ticks do not express diapause and have high burden throughout the year, cattle are continuously challenged by infected ticks (high transmission intensity) and calves are infected early in their life while all adult animals are immune (no clinical disease occur). In tropical conditions characterised by well-defined seasons and unimodal rainfall pattern, obligatory diapause of adult ticks results in tick phenology synchronised with the rains and loss of *T. parva* infection in adult ticks (reduced transmission intensity) (Bishop et al., 2008). In such areas and unlike equatorial regions, carrier animals play a crucial role of reservoir whereas the nymphal tick stage may transmit *T. parva* more efficiently than adult ticks (Mulumba et al., 2000). In fringes areas, there could be endemic instability and risk of occasional epidemics due to the introduction and temporal subsistence of infected ticks

1.3.4 Control of East Coast fever

1.3.4.1 Chemotherapy and control of the vector

East Coast fever can be controlled by direct killing *T. parva* using chemotherapeutic drugs and indirectly by using acaricides to reduce infestation by the tick vectors responsible for transmission of the pathogen. Acaricides are usually applied by spraying and/or dipping of cattle and have effectively been used in preventing ECF in most affected areas. However, although effective, there have been raised issues regarding the use of acaricides. The acaricides are toxic, raising concerns in terms of residues in the environmental and possible contamination of the food chain (residues in milk and meat products). Acaricides are expensive and need to

be applied regularly throughout the life of animals; poor farmers in most affected countries cannot afford the chemical for regular spraying (George et al., 2004). Furthermore, intensive use of acaricide creates a susceptible cattle population and favours the development of acaricide-resistance Rhipicephalus ticks (Jongejan and Uilenberg, 2004; Vudriko et al., 2016).

Affected animals can be treated using theilericidal drugs such as parvaquone or buparvaquone. Buparvaquone (traded as Butalex®), a derivative of parvaquone remains the drug of choice for treating *T. parva* infections in cattle (McHardy et al., 1985). Buparvaquone most likely inhibits mitochondrial electron transport in the parasite. This drug is effective against schizont and sporozoite stages. However, there are some limitations to the widespread use of this drug: (i) this molecule requires to be administrated during the early stage of infection before widespread invasion of the immune system by intracellular schizont stage; so, a rapid and accurate diagnostic is required, and (ii) buparvaquone is also relatively expensive and may not always be available to smallholder farmers (Morrison and McKeever, 2006). The findings of buparvaquone resistance in *T. annulata* also raises concerns for possible resistance in *T. parva* as well (Nyagwange et al., 2019).

1.3.4.2 Vaccination against *Theileria parva*

A live vaccination of cattle against ECF is available. It involves the inoculation of *T. parva* sporozoites and simultaneous treatment with a long-acting oxytetracycline. The antibiotic controls but does not kill the parasite and allow cattle to develop a protective immune response. Following live vaccine administration, a mild episode stimulates the cell-mediated immunity based on the MHC-I restricted CD8+ T cells (McKeever et al., 1994; Morrison and Goddeeris, 1990). Vaccinated animals are then protected against lethal challenge with homologous parasite strains. The immunity is expected to be long life especially if boosted by continuous tick challenge. This approach is known as infection and treatment method (ITM) of vaccination (Uilenberg, 1999). The ITM vaccination have been developed in the mid 1970s (Radley et al., 1975). A mixture of three *T. parva* stocks called the "Muguga Cocktail" is the widely used formulation. The three stocks incorporated in the vaccine are Muguga, Kiambu-5 and Serengeti-transformed. The use of the trivalent vaccine has been reported to confer broad-spectrum protection with variable degree of cross-protection in countries where the vaccine was introduced. However, the it does not protect against buffalo-derived *T. parva* in most cases (Sitt et al., 2015). The vaccine has been registered and used in several affected African countries, including Malawi, Kenya, Uganda, Tanzania and Zambia (McKeever, 2007; Nene and

Morrison, 2016). Meanwhile, other *T. parva* stocks a part from the initial Muguga cocktail have been used to immunise cattle in different African countries: Marikebuni stock in Kenya (Wanjohi et al., 2001), Katete and Chitongo stocks in Zambia (Geysen, 1999; Marcotty et al., 2002) and Boleni stock in Zimbabwe (Latif and Hove, 2011). Although the ITM vaccine has provided solid immunity, its large-scale adoption has been facing major concerns, which include the risk of introducing new strains in areas where they were previously absent (Di Giulio et al., 2009; McKeever, 2007; Oura et al., 2004b). It is known that vaccine strains may establish a tick-transmissible carrier state in immunised cattle which constitute the risk for spread of the disease (Kariuki et al., 1995). Other major concerns of this method of immunisation include the fact ITM vaccination provides immunity against homologous strains, while vaccinated animals remain susceptible to infection by heterologous strains (Radley et al., 1975; Sitt et al., 2015; Taracha et al., 1995a). Breakthrough infections in some areas have been attributed to the existence of genetically and immunologically more diverse strains in field population that probably do not cross-protect with vaccine strains. Thus, live parasite immunisation approach requires genetic and antigenic characterisation of the parasite population present in an area prior to safety and efficacy trials and deployment (Geysen, 1999). Efforts are still directed at scaling up its wide deployment in more affected countries. To overcome the concerns raised above, other vaccination methods have been evaluated and are currently under evaluation, using *T. parva* antigen vaccine candidates that are target of neutralization antibodies and CD8+ T cells (Lacasta et al., 2018; Musoke et al., 1992; Musoke et al., 2005; Svitek et al., 2018). Neutralizing antibodies are directed against sporozoite surface proteins to prevent sporozoite entry into host cells while CD8+ cytotoxic T cells (CTLs) are directed against antigens of the schizont stage to interrupt the infection by killing the schizont-infected host cells. A number of antigens that are targets of neutralizing antibodies have been identified: p67, p32, p104, PIM and p150 (Dobbelaere et al., 1985; Iams et al., 1990; Shapiro et al., 1987; Skilton et al., 1998; Skilton et al., 2000). The most tested antigen is the surface protein p67, which neutralized sporozoite infection in vitro. This antigen is conserved among cattle-derived *T. parva* strains making it a suitable candidate vaccine to allow cross-protection between cattle-derived parasite strains (Nene et al., 1996). There have been several vaccine trials using the p67 protein and protection from *T. parva* challenge following p67 immunisation induces some degree of protection ranging from 13 to 70% (Musoke et al., 1992). The major limitation with p67 vaccination trials is the lack of correlation between various antibody titres and immunity and immunity was relatively lower in field conditions (Musoke et al., 1992; Musoke et al., 2005). Research are still undertaken trying to improve the immunological effect

of the recombinant p67 surface protein. Recently, Lacasta et al. (2018) reported a preliminary improvement of the correlation between humoral immune responses (antibody titers) and protection against ECF. A number of candidate vaccine antigens that are target of the MHC-I restricted CD8+ T cells have been identified and are under evaluation attempting to produce next generation vaccine against *T. parva* (Akoolo et al., 2008; Graham et al., 2007; Graham et al., 2008; Svitek et al., 2018).

1.3.5 Genetic recombination during the sexual reproduction phase of *Theileria parva* in the tick

The life cycle scenario is consistent with *T. parva* being predominantly haploid and undergoing asexual proliferation during the developmental stages in the bovine, while a transient diploid zygote formed after fusion of gametes develops in the tick midgut. Tick vectors are likely to carry a mixture of parasite genotypes during their blood meal, as cattle and buffalo are usually infected by multiple *T. parva* genotypes in the field (Hemmink et al., 2018; Muleya et al., 2012; Salih et al., 2018). Studies have provided evidence for significant genetic exchange (recombination events) that can occur through sexual recombination when two or more *T. parva* genotypes infect *R. appendiculatus* (Katzer et al., 2006; Nene et al., 1998). This can generate allelic diversity and is presumed to play important role in the distribution of polymorphism and shifts in the population genetic structure of *T. parva* (Henson et al., 2012; Katzer et al., 2011; Katzer et al., 2006). Furthermore, the recombination processes may have substantial implication in shifting of immunological determinants and their recognition which may occur from reassortment between antigen-encoding loci recognised by the host cellular immune response (Katzer et al., 2006).

1.3.6 Pathogenesis and clinical signs of East coast fever

The pathogenic effect of *T. parva* is induced by the schizont-infected cells. During the intracellular phase of *T. parva* in cattle, schizont-infected cells acquire a cancer-like phenotype with massive and uncontrolled proliferation. This rapid multiplication of infected cells is responsible for most of the pathogenicity of *T. parva* and is associated with the main clinical signs of ECF (Shaw, 2003). The length of the incubation period and the severity of ECF depend on the parasite load, with initial clinical signs appearing approximately 8-12 days post infection. Schizont-infected cells appear in the lymph nodes draining the inoculation site (the predilection feeding site of the vector). In the case of *T. parva* often transmitted by *R. appendiculatus* ticks,

this results in unilateral hypertrophy of the preparotidian lymph node. Two to three days later, parasitized cells become detectable in other lymphoid tissues (spleen, kidneys, liver and lungs) and the animals suffer from generalised lymphadenopathy. The main clinical signs include enlargement of lymph nodes (containing lager number of parasitized cells), high fever (40-42°C), dyspnoea, diarrhoea, lacrimation, corneal opacity and nasal discharge. Throughout the course of the disease, rumination stops, the animals become anorexic, they lose weight and their body condition quickly deteriorates which leads to reduced milk yield (Morrison et al., 1996) Pregnant animals may abort. The disease causes high mortality in fully susceptible cattle, which usually die within 14 to 24 days after infection (Morrison, 2015). Cattle death is usually caused by severe pulmonary oedema leading to respiratory distress at the terminal stage (Ndungu et al., 2005). The acute phase is associated with the presence of multinucleated intracytoplasmic and free schizonts in lymph node biopsy smears. Some infected animal can also develop into a fatal condition referred to as "turning sickness" (neurologic signs). Neurological symptoms are caused by blockage of the capillaries of the central nervous system by infected cells (Mbassa et al., 2006). A wasting disease is observed in more chronic conditions with prolonged poor body state and ultimately death after a few weeks. Long-standing cases show foci of lymphocytic infiltrations in kidneys that appear as white infarcts in animals that have recovered. The is a marked variation in the susceptibility of cattle to infection and the severity of the disease can be altered by cattle breed, the strains of *T. parva* involved and magnitude of the infected tick challenge (Mbogo et al., 1996; Morrison et al., 2015). A proportion of animals may recover and develop strong immunity against homologous *T. parva* strain (carrier status) (Skilton et al., 2002). However, they may remain emaciated and unproductive and, in most cases, sustain a chronic infection transmissible to ticks over prolonged periods of time (Kariuki et al., 1995; Oura et al., 2007).

1.3.7 Immune response to *Theileria parva* in the bovine host

There is evidence that immunity to *T. parva* can be activated in two ways: killing the schizont-infected host cells by cytotoxic T cells and blocking the sporozoite entry by neutralization antibodies specific for sporozoite surface antigen. However, maternally derived antibodies do not provide such protection, and a cell-mediated response to infection is always necessary for effective immune protection (McKeever et al., 1999; Morrison, 2007). The recovery from infection has been demonstrated to be correlated with the appearance of CD8+ T cells specific for parasitised lymphocytes following immunisation or natural challenge with *T. parva*. This is

a strong evidence that the solid and long-lasting protective mechanisms against *T. parva* is cell-mediated and directed against the schizont-infected lymphocytes (McKeever et al., 1994). The immune cells responsible for killing infected lymphoblasts belong to CD8+ T-cell subpopulation, mediated by their T cell receptor (TCR) which bind to the major histocompatibility complex class-I (MHC-I) at the surface of target cells, to produce a cytotoxic immune response (**Figure 1.6**) (Goddeeris et al., 1990; McKeever et al., 1994). The CD8+ T cell responses in different animals is mainly directed against different immunodominant antigens which exhibit significant allelic variation among *T. parva* strains, and make the protection strain specific (MacHugh et al., 2009; Taracha et al., 1995a; Taracha et al., 1995b). The strain specificity of induced CTL responses is consistent with the reported antigenic heterogeneity in *T. parva* populations, which is maintained to a greater extent in buffalo-derived *T. parva* (McKeever, 2001; Pelle et al., 2011; Sitt et al., 2015). The strain specific nature of protection has also been associated with the differences in MHC-I haplotypes among cattle, which tend to respond to different epitopes of *T. parva* (Goddeeris et al., 1990; Steinaa et al., 2018b) and the polymorphism in the T cell receptor (TCR) repertoires of the host might play a role in the affinity differences towards different epitopes of *T. parva* strains (Morrison et al., 1995). Therefore, the cell-mediated immune response to *T. parva* is the result of interaction between bovine MHC-I haplotypes, TCR recognition and the strain of *T. parva* (McKeever, 2006; Morrison et al., 1987; Morrison et al., 1996). Although the cell-mediated immunity has been reported to be the main mechanisms underlying the protective immunity against *T. parva*, sporozoite neutralising antibodies can also be induced after repeated challenge or vaccination using recombinant p67 sporozoite surface protein and reduce severity of infection. However, the lack of correlation between the antibody titers and immunity suggests weak immunogenicity of the protein (Musoke et al., 2005; Nene et al., 2016).

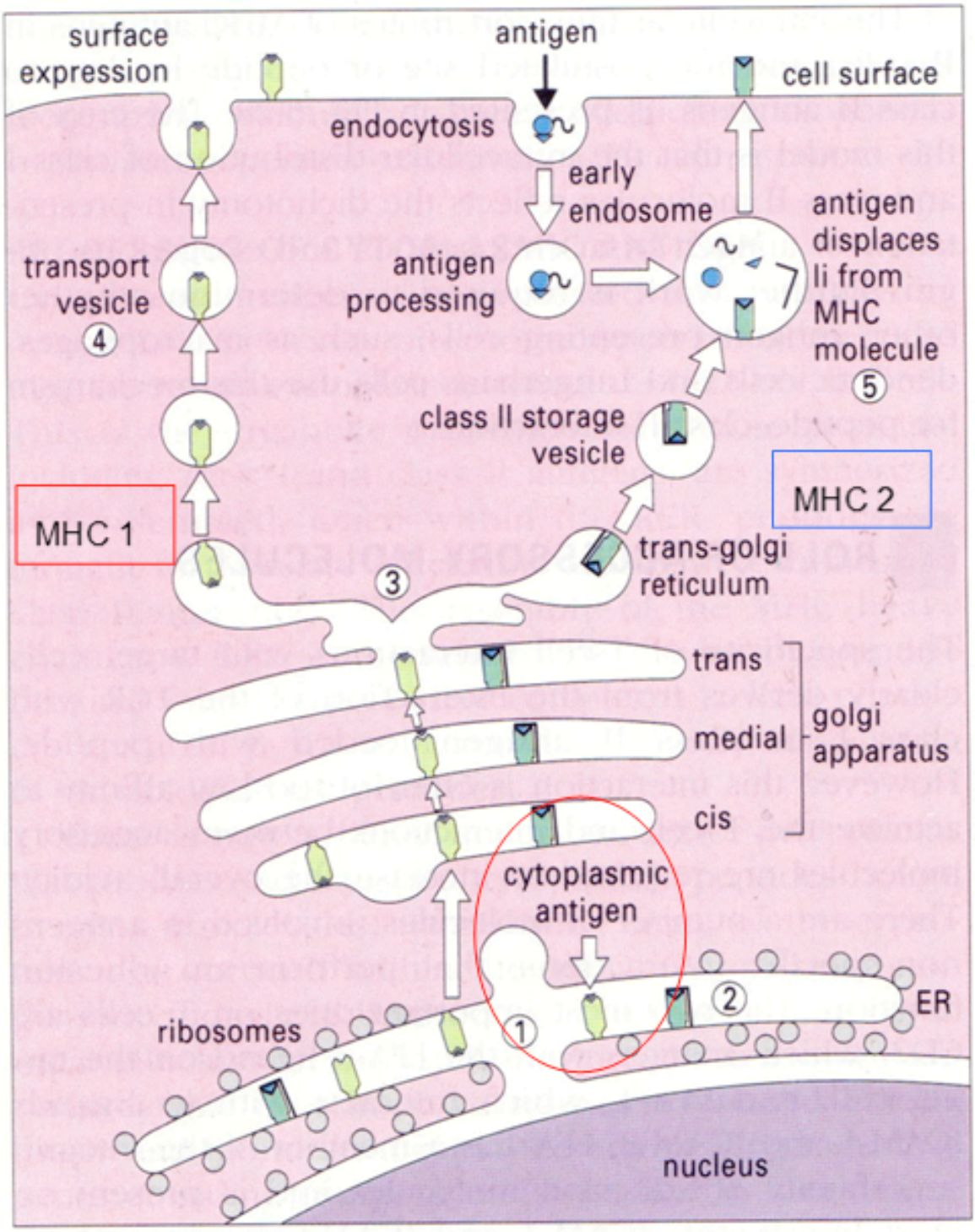

Figure 1.6 Proposed intracellular transport pathway for MHC I and II molecules involved in antigen presentation (Male and Brostoff, 2007)

1.3.8 Evolutionary history of *Theileria parva*: from African buffalo to cattle?

Theileria parva infects both cattle and African buffalo (*Syncerus caffer*). The parasite is non-pathogenic in buffalo and does not induces any clinical sign of the disease, while it gives rise to a clinical disease in cattle when infected for the first time. African buffalo, which is considered as native bovine host, remains long-term healthy carrier and natural reservoir of more diverse genotypes of *T. parva* (Bishop et al., 2015; Minami et al., 1983; Morrison, 2015; Sitt et al., 2015). Strikingly, cattle grazing in the same area with buffalo can acquire infection with buffalo-derived *T. parva* variants. Such infection results in an acute and usually highly fatal clinical disease called "corridor disease", associated with an acute death of infected cattle (Bishop et al., 2004; Latif et al., 2002; Sitt et al., 2015). However, given the poor differentiation of buffalo-derived infections to the erythrocytic tick-infective stage (piroplasm stage) and the fatal death of cattle during the corridor disease, transmission of the infection from cattle to ticks is highly inefficient (Bishop et al., 2015; Maritim et al., 1992; Uilenberg, 1999). Cattle-derived *T. parva* (causing East Coast fever and January disease) is considered to have originated from African buffalo before it adapted to cattle. From the evolutionary point of view, the initial evolution and genetic diversity of *T. parva* may have likely arisen in African buffalo prior to its establishment in cattle population (Hayashida et al., 2013; McKeever, 2009). It has been argued that *T. parva* co-evolved a long time with African buffalo before the introduction of cattle in Africa. This implies that, after introduction of domestic cattle in Africa, only a limited subset of the buffalo-derived *T. parva* gene pool may have developed the ability to cycle in cattle (i.e. keeping cattle alive long enough to be transmitted to ticks) and established a stable population that is currently maintained and circulating among cattle (Freeman et al., 2006; Gifford-Gonzalez and Hanotte, 2011; Hayashida et al., 2013; Pelle et al., 2011). This hypothesis reflects that the African buffalo may have strongly influenced the evolution of *T. parva* through their immune responses, which allowed the evolutionary adaptation to the parasite and the control of the infection (Bishop et al., 2004; McKeever, 2009). As cattle are still sharing grazing lands with African buffalo in some areas and the fact that cattle are continuously infected by *T. parva* from buffalo-derived in the presence of the tick vector *R. appendiculatus*, it is possible that the adaptation process has been continuous. However, there is no data that predict the evolutionary time scales of such co-evolution and the adaptation process is still unknown. In addition, researchers have argued that there is a genetic barrier to recombination events between cattle and buffalo-derived *T. parva* strains. They hypothesised that genetic recombination may have occurred in the very ancient ancestral bovine *Theileria* populations before their

geographical segregation (Hayashida et al., 2013). Then the coevolution between *T. parva* and cattle led likely to the divergence between buffalo and cattle parasite strains that evolved independently after geographic and host separation (Hayashida et al., 2013).

Parasite adaptation to cattle can be seen in two steps. First, the survival of the cattle should be long enough for the parasite to produce gametocytes and infect ticks shortly after infection. The disease caused by such infection is referred to as East Coast fever and is generally observed in Central and Eastern Africa where tick abundance is relatively constant throughout the year. In areas with uni or bimodal annual tick abundance as observed in Southern Africa, cattle must remain carrier of the infection for several months to assure the transmission from infective ticks (nymphs or adults) to receptive ticks (larva or nymphs), which are not found at the same time on animals. In this instance, the disease may present a peak of incidence in January, during the rains (hence the name January disease) and often appear less severe than ECF and Corridor disease.

1.3.9 Generation and maintenance of *Theileria parva* genetic diversity: evolutionary survival strategies

The evolutionary changes observed in *T. parva* parasites have significantly modified the phenotype characteristics, including virulence and evasion ability from host cells immunity (Sivakumar et al., 2014). There are several mechanisms and factors that explain the generation and maintenance of genetic variation in *T. parva* parasites. A better understanding of these mechanisms and their patterns in different ecological and epidemiological zones is crucial to predict theileriosis dynamics and vaccine outcomes. Genetic recombination during the sexual stage in the tick vector is considered to be the major mechanism generating genetic diversity in *T. parva* as discussed above (Henson et al., 2012; Katzer et al., 2011). Nevertheless, several other mechanisms and factors including bovine host immunity, mutation, gene conversion, gene isolation, gene flow among populations and genetic drift are thought to underline and maintain the genetic diversity among *T. parva* populations (Connelley et al., 2011; Geysen et al., 2004; Katzer et al., 2011; Katzer et al., 2006; McKeever, 2009; Sivakumar et al., 2014; Toye et al., 1991). *Theileria parva* is believed to have undergone evolutionary changes to survive immunological responses of its hosts (Deitsch et al., 2009). Thus, the influence of bovine immunity plays a crucial role generating genetic and antigenic diversity of *T. parva* (McKeever, 2009; McKeever et al., 1999). This survival strategy of *T. parva* often complicates the epidemiology of ECF (McKeever, 2009). In addition, evolutionary acquisition of mutations in

the immunodominant CD8+ T-cell epitopes play an important role in the survival strategies facilitating the parasite to escape bovine immune recognition as immunity is strain specific (Connelley et al., 2011). Furthermore, the diversity is very limited in populations characterised by geographic separation (restricted gene flow) and genetic drift, whereas gene flow is thought to increase the chance of sexual recombination between *T. parva* stocks and therefore increase the genetic diversity within parasite populations. Besides, it is important to consider factors such as host movements (mostly attributed to bovines since ticks themselves rely on their hosts for movements), which may be more crucial in the genetic composition and substructuring of *T. parva* (Bishop et al., 2004).

1.3.10 Molecular tools to study the population genetics of *Theileria parva*

Genetic composition and evolutionary processes of pathogens vary among populations within species, and has significant implications for the epidemiology and control of the disease (Ellegren and Galtier, 2016). Over the past three decades, molecular techniques and gene markers were developed for *T. parva* and other protozoan pathogens. These tools have been increasingly applied to understand the population genetics and the mechanisms underlying the genetic diversity in protozoa pathogens (Tibayrenc et al., 1990). Specifically, genetic diversity and molecular evolution of *T. parva* studies have been widely carried out to genotype *T. parva* populations with a major goal to support live *T. parva* vaccine development and use in sub-Saharan Africa. For that purpose, a range of genetic markers and molecular techniques were developed and widely used to analyse the diversity and structure of *T. parva* stocks isolated from cattle and buffalo in a wide range of geographic areas as well as laboratory stocks (**Figure 1.7**). These tools include anti-schizont monoclonal antibodies (MAbs) directed against a single immunodominant polymorphic antigen (Conrad et al., 1989; Minami et al., 1983; Toye et al., 1995b), southern blotting and restriction fragment length polymorphisms (RFLP) techniques using mainly DNA probes derived from multicopy gene families (Bishop et al., 2001; Bishop et al., 1993; Bishop et al., 1994; Conrad et al., 1987), PCR based approaches that target genes encoding antigens recognised by bovine antibodies by applying PCR-RFLP essay (Bishop et al., 2001; Bishop et al., 1992; De Deken et al., 2007; Elisa et al., 2014; Geysen, 1999; Pienaar et al., 2011) and sequencing of these antigen genes (Nene et al., 1992; Sibeko et al., 2011; Sibeko et al., 2010; Sitt et al., 2019). The polymorphic single copy antigen genes (PIM, p67, p104 and p150) used to discriminate between *T. parva* strains using the RFLP method enabled better characterisation of *T. parva* stocks before and after immunisation programs (Elisa et al.,

2014; Geysen et al., 1999). Later, the identification of a panel of micro- and mini-minisatellite DNA markers (SSR) greatly enhanced molecular characterisation of the population structure and diversity of *T. parva* (Asiimwe et al., 2013; Elisa et al., 2015; Muwanika et al., 2016; Odongo et al., 2006; Oura et al., 2005; Oura et al., 2007; Oura et al., 2003; Oura et al., 2011b; Patel et al., 2011; Rukambile et al., 2016). Most recently, the published whole genome sequence for *T. parva* gave hope to further understand the population genetics and evolutionary changes of the pathogen (Gardner et al., 2005; Hayashida et al., 2013; Hayashida et al., 2012; Henson et al., 2012; Norling et al., 2015; Pain et al., 2005). It provided more powerful polymorphic genetic tools, including antigens that are targets of CD8+ T cell responses induced by the live vaccine and additional SSR markers for fine characterisation of the genotypic and antigenic diversity and the population structure of the pathogen (Elisa et al., 2015; Graham et al., 2007; Graham et al., 2006; Graham et al., 2008; Hemmink et al., 2018; Hemmink et al., 2016; Katzer et al., 2010; Kerario et al., 2019; Muleya et al., 2012; Oura et al., 2004a; Pelle et al., 2011; Salih et al., 2018; Salih et al., 2017; Sitt et al., 2018).

Molecular characterisation based on monoclonal antibodies and DNA probing revealed an extensive polymorphism among *T. parva* isolates from various geographic regions, including Tanzania, Kenya, Zambia, Zimbabwe and Uganda (Bishop et al., 1993; Bishop et al., 1994). However, these studies did not provide overall picture of the diversity on a genome-wide scale (Bishop et al., 2002; Bishop et al., 1994; Toye et al., 1995b). Other research focused on the analysis of single copy antigen coding genes under the PCR-RFLP and genotyping with a panel of Variable Number Tandem Repeat (VNTR) markers have demonstrated that highly diverse *T. parva* populations occur in Uganda (Muwanika et al., 2016; Oura et al., 2005), Zambia (Geysen et al., 1999; Muleya et al., 2012), Kenya (Nene et al., 1992; Odongo et al., 2006; Oura et al., 2003) and South Sudan (Salih et al., 2018) and in the Grande Comore (De Deken et al., 2007). However, low level of diversity was found in cattle-derived *T. parva* from Tanzania (Elisa et al., 2015), while buffalo-derived parasites where more diverse in the same country (Rukambile et al., 2016) and in Uganda (Oura et al., 2011a). The majority of Zambian *T. parva* field samples had homogenous hybridisation patterns, but differ from the components of the Muguga trivalent vaccine (Geysen et al., 1999). Infection of a single animal with multiple *T. parva* genotypes was reported to be frequent and evidence of linkage disequilibrium was observed in most of these studies. In addition, sequencing of the polymorphic immunodominant molecule (PIM) and the p67 antigen revealed polymorphism between *T. parva,* especially in buffalo-derived *T. parva* (Bishop et al., 1997; Toye et al., 1995a). Different PIM patterns and

an extensive polymorphism where observed in buffalo-derived *T. parva* from south Africa (Sibeko et al., 2011; Sibeko et al., 2010). The levels of diversity in the gene encoding the p67 sporozoite surface protein were similar in *T. parva* from buffalo and cattle naturally infected from buffalo (Sitt et al., 2019). However, the sporozoite antigen p67 shows a considerable sequence conservation within cattle-derived *T. parva*, making it useless to characterise cattle-derived *T. parva* populations, but could serve to distinguish these parasites with buffalo-derived *T. parva (*Nene et al., 2016; Nene et al., 1996). Profiles obtained from buffalo stocks are also heterogeneous whereas cattle-derived *T. parva* stocks are often homogeneous.

The recent identification of genes encoding *T. parva* antigens that are targets of the CD8 + T cells allowed to further investigate sequence polymorphism and epitope diversity in *T. parva* populations from cattle and buffalo in different geographic regions of Africa: in Kenya and selected laboratory isolates from different African countries (Pelle et al., 2011), Tanzania (Elisa et al., 2015; Kerario et al., 2019), South Sudan (Salih et al., 2017), laboratory isolates component of the trivalent vaccine from Kenya (Hemmink et al., 2016) and buffalo-derived isolates from Kenya and South Africa (Hemmink et al., 2018; Pelle et al., 2011; Sitt et al., 2018). These studies revealed the presence of large number of allelic variants both at nucleotide and antigenic level and within the CTL epitopes in buffalo-derived *T. parva* (Hemmink et al., 2018; Pelle et al., 2011; Sitt et al., 2018). However, a limited polymorphism (nucleotide and antigenic) was reported within cattle-derived and in the component of the trivalent vaccine, further suggesting that cattle contain a subset of *T. parva* population that have originally evolved in buffalo and that the vaccine does not cover the overall diversity of *T. parva* in the field (Hemmink et al., 2016). The application of these markers is crucial in determining the population structure of *T. parva* in different eco-epidemiological contexts, predicting vaccine outcomes prior to vaccination and follow up the vaccination (Graham et al., 2006; Graham et al., 2008; Nene et al., 2016). No such studies have been undertaken in the Great Lakes region, where field vaccination has not been deployed yet, except in Burundi and Rwanda where a small-scale live immunisation programme was implemented between 1981 and 1987, to test the protective effect of local *T. parva* isolates (Tama, 1989).

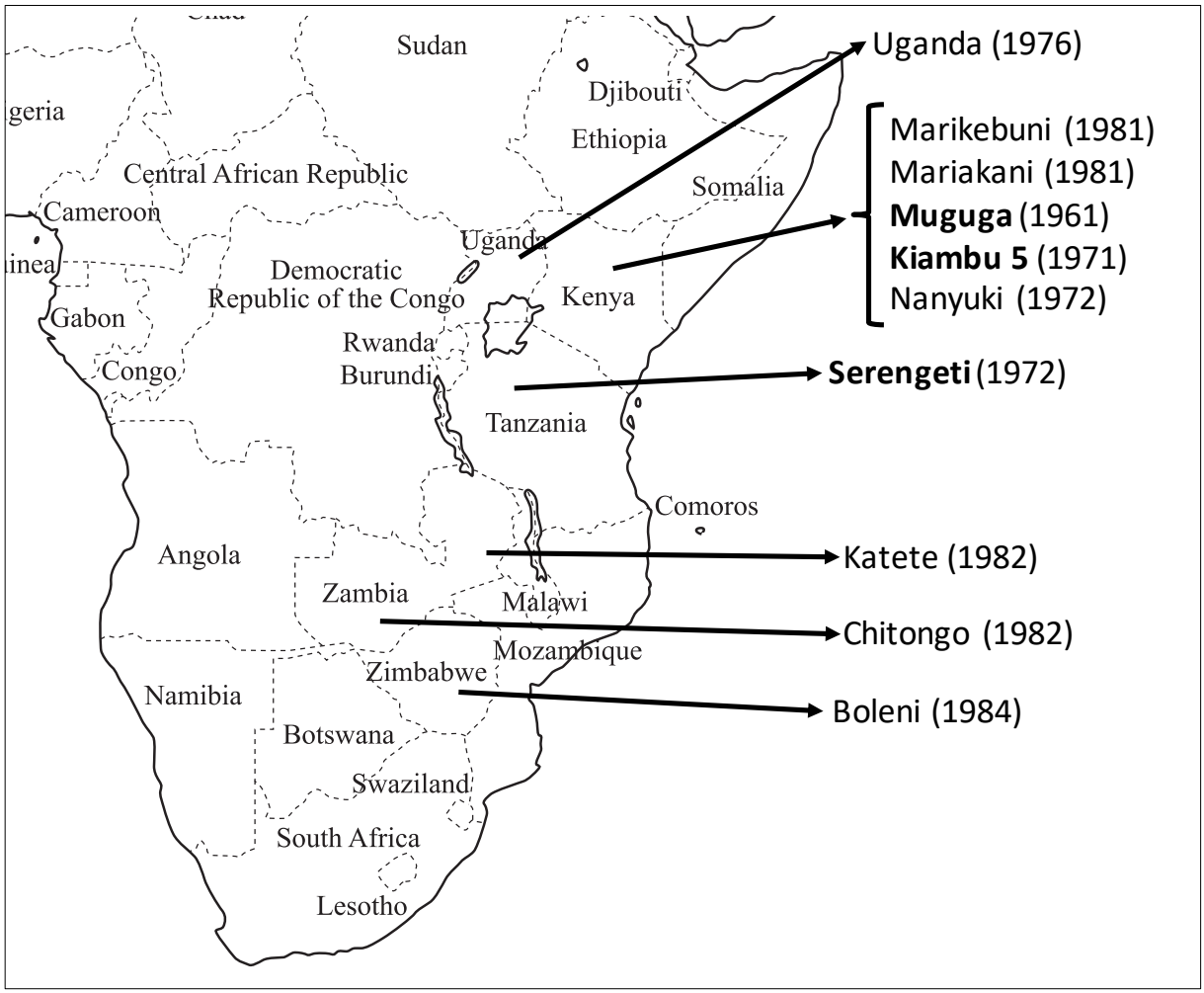

Figure 1.7 Geographical distribution of *Theileria parva* laboratory stocks described in sub-Saharan Africa. Bold indicates the component of the trivalent Muguga vaccine. The numbers in brackets indicate the year of isolation.

1.3.11 *Theileria parva* whole genome: questioning the protective value of Muguga and Serengeti strains in the vaccine

The *T. parva* genome, which has a small size (8.5-12 Mbp) compared to other apicomplexan parasites, is divided into 4 chromosomes (Bishop et al., 2004; Nene et al., 1998). Whole genome sequences from cattle and buffalo-derived *T. parva* revealed the genetic complexity of *T. parva* and identified higher number of single nucleotide polymorphisms (SNPs) in *T. parva* from buffalo than in those from cattle (Hayashida et al., 2013; Henson et al., 2012). In addition, Buffalo-derived isolates were clustered in separate phylogenetic group diverged from cattle-derived parasites (Norling et al., 2015). Single nucleotide polymorphisms are more prevalent in protein coding than non-coding regions, under different rate of evolutionary pressure

(Hayashida et al., 2013). However, the genes that are involved in mediating immunity to ECF, virulence and biological differences between strains remain to be determined. The whole genome sequencing has also identified high rate of recombination events and a considerable number of crossover events, suggesting gene conversion that might be leading to divergence and diversity between *T. parva* strains (Gardner et al., 2005; Henson et al., 2012; Katzer et al., 2011). Besides, the sequence differences between the components of the trivalent vaccine (Muguga, Serengeti-transformed and Kiambu-5) are becoming problematic. For instance, there is high degree of genome similarity between Muguga and Serengeti-transformed, questioning the protective value of the two strains and their importance to be used together in the vaccine. However, the differences Muguga and Kiambu-5 are consistently different across the genome, particularly within the protein-coding genes (Norling et al., 2015). A more recent experiment demonstrated that immunisation with the Muguga strain alone provides similar level of protection compared to trivalent Muguga cocktail (Steinaa et al., 2018a), supporting the evidence of limited antigenic diversity in the three components of the Muguga cocktail vaccine (Hemmink et al., 2016). In view of these data, the protection value of the three components strains of the Muguga cocktail vaccine should be re-assessed under fine-scale biological and vaccination trials.

1.4 East Coast fever in the Great Lakes region

1.4.1 Physical characteristics of the study area: DRC, Rwanda and Burundi

The ecological diversity in the Great Lakes region (Figure 1.8) is largely defined by the landscape topography and altitude which subsequently regulate other bioclimatic attributes, including rainfall amount and pattern, temperature and vegetation. In addition, protected areas and shared water bodies including Lakes Kivu (DRC and Rwanda) and Tanganyika (DRC and Burundi) and the Ruzizi river also influence the bioclimatic conditions in the region. Generally, the climate is essentially tropical and strongly tempered by altitude (cooler and more humid with increasing altitude). It comprises three major seasons with a long and bimodal rainfall pattern: (i) the early rainy season (wet1: September to December), (ii) the late rainy season (wet2: February to May) and (iii) the dry season (dry: June-August). There is an intervening short dry period between the two rainy seasons of approximately 15 days in January and/or February. These contrasting agroecological and climatic conditions are thought to influence the socio-economic characteristics, cattle management systems, tick vector population dynamics

and epidemiological features of ECF among AEZs (Bazarusanga et al., 2007a). A summary description of the characteristics of the six AEZs is presented in **Table 1.1.**

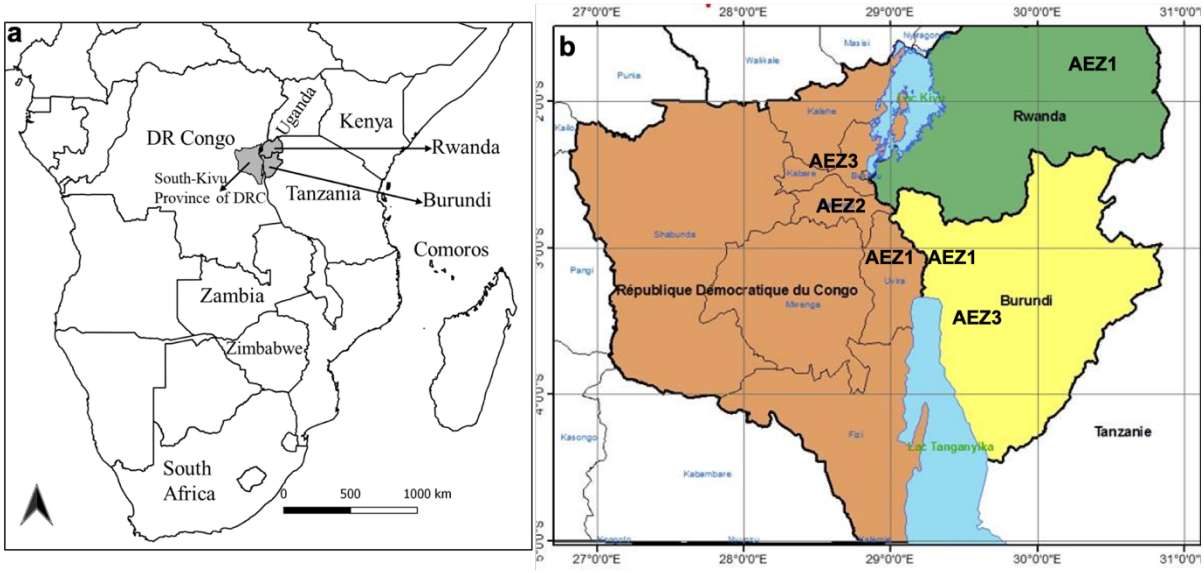

Figure 1.8 Geographic map of the study area: a) Map of Africa showing the study area (in grey) and other countries where *R. appendiculatus* and *T. parva* were previously described; b) The Great Lakes region. Altitudes: lowlands (AEZ1 < 1200 m); midlands (AEZ2: 1200–1800 m) and highlands (AEZ3: 1600-2800 m).

Table 1.1 Geographical and climatic characteristics of the agro-ecological zones (AEZ) included in the study

Country	Agro-ecological zone (AEZ)	Altitude (m)	Temperature (°C)	Rainfall (mm/year)	Rainy season
DRC	AEZ1	780–1100	23–25	800–1000	October-April
	AEZ2	1200–1800	17–21	1000–1500	September-May
	AEZ3	1800–2800	12–19	1350–2000	September-May
Burundi	AEZ1	774–1100	23–24	800–1100	November-May
	AEZ3	1700–2800	14–15	1300–2000	September-May
Rwanda [a]	AEZ1	1000–1500	21–24	800–950	November-May

Abbreviations: DRC (AEZ1: Lowlands, AEZ2: Midlands, AEZ3: Highlands); Burundi (AEZ1: Lowlands, AEZ3: Highlands); Rwanda (AEZ1: eastern low plateau which is the lowlands of Rwanda)

The Democratic Republic of Congo

Sampling was conducted in the South-Kivu province in the eastern Democratic Republic of Congo. The South-Kivu province lies between longitudes 26- 29°E and latitudes 1°-5°S and covers 64,851 km². Cattle blood samples and ticks (attached and free-living) were collected in three AEZs along an altitudinal transect, defined as lowlands (DRC AEZ1) located in the Ruzizi Valley, midlands (DRC AEZ2) in the administrative district of Walungu and highlands (DRC AEZ3) located in the district of Kabare and the highlands part of Walungu (Mulumemunene). The lowlands area is the warmest semi-arid AEZ, characterised by a tropical warm and dry climate, with a cool dry season of 4–5 months and warm rainy season. The vegetation is strongly dominated by savannah grasslands and small patches forest. The "upland Kivu" (midlands and highlands) falls within a montane humid tropical climate which consists of a warm and humid tropical climate in midlands and much cooler conditions in highlands (lower to mild temperature). The highlands area experiences abundant and frequent rainfalls with a warm dry season receiving occasional and poorly distributed rainfall. The dominant vegetation is more variable in the upland Kivu, composed of montane and savannah grasslands, scrubs woodland and degraded forests in the midlands and highlands while specifically in some places of highlands, it is marked by relict species of seasonal deciduous forests (Klapwijk et al., 2019).

Burundi

Burundi is located between latitude 2-4°S and longitude 28-30°E and covers 27.834 km^2. The country is bordered by Rwanda, Tanzania, Lake Tanganyika and the Democratic Republic of Congo. Burundi is divided into five AEZs based on their altitudes and climatic conditions: the Imbo valley (780 – 1100 m), the western slope of the Congo-Nile ridge (1000 – 1700 m), the Congo-Nile ridge (1700 – 2500 m), the central plateau (1350 – 2000 m) and the northeast depressions (1100 – 1400 m). Cattle blood samples and ticks were collected from two AEZs: The administrative districts of Rugombo and Gihanga in the Imbo lowlands (Burundi AEZ1) and the districts of Muramvya, Mwaro and Mugamba in the Congo-Nile highlands (Burundi AEZ3). The Imbo region extends along the Ruzizi River and the North of Lake Tanganyika. Like the Ruzizi valley in RDC, the vegetation is composed of savannah and small patches of forest. The Imbo valley is characterised by a warm tropical climate with low rainfall and a long dry season while the Congo-Nile ridge (highlands) is characterised by a tropical montane climate with low temperature (cooler conditions) and high rainfall (comparable with the highlands of DRC) (Hatungumukama et al., 2007).

Rwanda

Rwanda lies between latitudes 1–3°S and longitudes 29–31°E and covers 26,300 km^2. It borders the Democratic Republic of Congo, Burundi, Tanzania and Uganda. The country is just below the equator (south of the equator). Based on altitudes, climatic conditions and the diverse relief, Rwanda is divided into four major AEZs: (i) The highlands zone (altitude:1,950 m) found in the northern cold and high rainfall region, (ii) the medium tempered zone in the central high plateau (1,650–1,950) where the climate (cooler conditions) is moderated by the Lakes Kivu, (iii) the medium continental zone located in the western high plateau (1,550–1,900) which is characterised by cooler climate, and (iv) the lowlands in the eastern plateau (1,000 and 1,500 m) which is the semi-arid region of Rwanda and is ecologically comparable with the lowlands of DRC and Burundi (Ruzizi valley and Imbo valley, respectively). Ticks samples from Rwanda were obtained in the districts of Nyagatare, Gatsibo and Kayonza of the eastern low plateau (Rwanda AEZ1) (**Table 1.1**). This AEZ includes the savannah area of eastern province of Rwanda and is the most important pastoral region in Rwanda, holding 40% of the national herd raised under a communal grazing system. The lowlands area of Rwanda is located in the north-eastern part of Rwanda close to the border with Tanzania and Uganda. The vegetation type is largely savannah and some river bank woods (Bazarusanga et al., 2007a).

1.4.2 Cattle production systems in the Great Lakes region

The livestock species plays important economic and socio-cultural roles in the Great Lakes region where cattle remain an important cultural symbol of power and prosperity (Hatungumukama et al., 2007; Kibwana et al., 2012; Mazimpaka et al., 2017). The contribution of livestock to the Gross Domestic Product (GDP) is estimated at 16% in Rwanda, 13% in Burundi and 4-6% in DRC. The cattle population consists of approximately 1,194,895 cattle in Rwanda, 467,000 in Burundi, 237,000 in Nord Kivu and 180,000 in South Kivu (eastern DRC). The majority of cattle populations are found in the eastern Province (lowlands) of Rwanda, in the Imbo valley (lowlands) and the Congo-Nile ridge (highlands) of Burundi and in the Ruzizi valley (lowlands) of DRC. The cattle populations are mainly dominated by pure indigenous Ankole Longhorned cattle breed (Sanga type), ecotypes that are found only in the Great Lakes region of East and Central Africa. They are estimated to represent more than 80-90% of the cattle population in both countries, but it remains difficult to define the degree of cross-breeding (Manirakiza et al., 2017; Ndumu et al., 2008). These cattle have the advantage to be highly adapted to the local environmental conditions. They have the potential to withstand periodic

feed shortage better than exotic breeds, they are tolerant to heat stress and against tropical diseases (most tick-borne diseases) as well as walking long distances for grazing lands (Paling et al., 1991; Wurzinger et al., 2006). However, local breeds are characterised by poor dairy and beef potential compared to exotic breeds.

In the Great Lakes region, cattle are kept under different production systems varying from traditional extensive systems to intensive or semi-intensive systems. Indigenous Ankole cattle are mostly kept under the traditional extensive system in open communal grazing lands characterised by little or absence of tick control. This system is mainly practised in semi-arid areas (lowlands) of both countries where communal natural grazing areas are the most available. It includes cattle mobility through seasonal transhumance in search of pasture and water. During the last three decades, the cattle population declined drastically and livestock production systems have undergone drastic changes often attributed to human population growth and land pressure which led to loss and degradation of natural pastures, shortage of communal grazing areas and disease epidemics that followed the influx of cattle with migrants of the violent conflicts (1990-2006) (Cox, 2012; Desiere et al., 2015; Hatungumukama et al., 2007; Jenicek and Grofova, 2015; Wurzinger et al., 2006). In most affected areas, especially in much of the midlands and highlands of both countries, smallholder farmers have adopted the integrated mixed crop-livestock farming system which also includes off-farm activities and micro-livestock (Klapwijk et al., 2019; Maass et al., 2012). In addition, there has also been a gradual shift from extensive systems to commercial intensive or semi-intensive management practices such as zero-grazing with feed supplementation, cultivated forage and use of agricultural residues to overcome the shortage of animal feed, mostly in highlands areas where grazing lands are becoming inexistent (replaced by habitations or crop production). These systems consist of integrating highly productive exotic breeds of cattle kept together with indigenous Ankole cattle (Manirakiza et al., 2017). They are characterised by frequent acaricide application. However, the intensive and semi-intensive systems are still emerging and are rare in the region.

After the tragic war period (1990-2006), livestock rehabilitation programmes have been initiated in Rwanda and Burundi and very recently in DRC trying to revitalise the sector by reversing the trends in cattle ownership among households. These projects were initiated to respond, in line with vision 2020, to the rapidly increasing demand for livestock products which is driven by current population growth and urbanisation. The major agenda was to modernise

agriculture and livestock production to contribute to both food security and poverty alleviation. Emphasis has been focused on the introduction and dissemination of exotic cattle breeds to rural farmers in Rwanda and Burundi; especially the Holstein and Friesian so as to improve dairy productivity in the region to meet the challenges of food security. In Rwanda, the government started the strategy of cattle distribution aiming at intensifying livestock production through the social framework "One cow per poor family". Since its introduction in 2006, the project has distributed around 300 thousand dairy cattle, with the purpose to reach more than 350 thousand by the end of the project (Nilsson et al., 2019). Similar livestock-oriented programs are gaining popularity in neighbouring countries. For instance, the Burundi government, with the support of technical and financial partners, promoted the importation of Holstein crossbred cows from Uganda to be given to rural poor smallholders. Some progress has already been made so far by these various projects in the region. Nevertheless, there are still major challenges of designing and developing appropriate livestock production system related to feed availability and animal diseases management, especially tick and tick-borne diseases (Klapwijk et al., 2014).

1.4.3 Cattle movement: local transhumance and cross-border movements of cattle

Socio-economic vulnerability in the Great Lakes Region is increasing by the recurrent violent conflicts and the effect of climate change, which increase poverty and internal migration of human population with their cattle. As rainfall become temporally and spatially more variable, seasonal movements of cattle, known as transhumance, represent a key resilience behaviour of cattle keepers in response to the environmental and socio-cultural constraints. The transhumance is a common practice characterised by regular movement of cattle during the dry season across all the main production zones in the region. It is motivated by the need to access grazing and water resources to overcome harsh environmental conditions towards regions of different climate and seasonal availability of vegetation. However, uncontrolled movements and contact patterns of cattle can contribute to the dispersal of vectors and multiple infectious animal diseases, as well as drive social conflicts between pastoralists and sedentary crop farmers (Brabant and Nzweve, 2013; Kerfua et al., 2018). Thus, the lack of integrated management of transboundary diseases between countries remains a big challenge in the region (Bouslikhane, 2015).

In general, three different types of cattle movements can be distinguished in the Great Lakes region: (i) short distance transhumance occurring within a production zone where cattle keepers

move with their animals in search of suitable grazing in the same AEZ, (ii) long-distance movements occurring to overcome seasonal variability of feed and water and to engage in livestock trade between AEZs (internal movements) and between countries (cross-border movements), and (iii) human population migrations with their cattle due to political unrest. The general trend of transhumance practice usually occurs from April to August. During this period, a large proportion of cattle located in high plateau migrate for long-distance transhumance to littoral regions of Lakes Kivu and Tanganyika and along the Ruzizi River. In particular, short-distance transhumance represents an integral component of the livestock production system in the Ruzizi Valley, the Imbo Valley and the Eastern Province of Rwanda (Nyagatare district), where cattle keepers undertake more local migration largely within the AEZs throughout the year. The main transhumance destination and corridor within countries are represented by savannah areas in lowlands near protected areas or Lakes and Rivers (mostly lowlands areas). Cattle move towards internal migration as well as cross-border seasonal movements from Burundi, Rwanda and Tanzania to DRC in search for grazing. Cross-border cattle trading routes within the Great Lakes region are from Rwanda, Burundi and Uganda (via Rwanda or Burundi) to the South Kivu (Ruzizi valley and Kabare district) and Nord Kivu Provinces. Cattle entering DRC for trade from neighbouring countries are oriented on provincial markets and towards main urban agglomerations of the country in other Provinces of DRC (Katanga, Maniema and Tanganyika). The particularity of these movements is that cattle do not move from Congo to other countries. The Ruzizi valley of DRC remains the main corridor and destination of cattle from neighbouring countries (Brabant and Nzweve, 2013).

The protracted violent conflicts which started in 1990 has immerged the region for much of the last three decades and also cleared the cattle population across the Great Lakes region. During the tragic conflict (1994-2006) and almost continuous and ongoing instability in the region, people migrated with their cattle to neighbouring countries (Cox, 2012). These migrations include several waves of exodus of cattle keepers originating from Rwanda and Burundi into DRC where they settled with their cattle in the Ruzizi valley (Uvira district) and the high plateau (Fizi and Uvira districts) of the South Kivu Province in the DRC (Vlassenroot and Huggins, 2005). Since the war of 1994 onwards, there has been an influx of returning refugees with a large number of cattle from neighbouring countries. In addition to recent migrations, there is broad consensus that the largest group of cattle-keeping population moved to DRC and Uganda during the pre-colonial period at the end of the nineteenth century in the context of migratory flux of population related to political unrest and other social and economic calamities in the

Rwandan Kingdom (Verweijen and Vlassenroot, 2015). The suitable grazing lands along the Ruzizi River and the Lakes Tanganyika also attracted migration of cattle keepers to DRC, where they settled with their cattle (Vlassenroot and Huggins, 2005). There was also forced migration of Rwandans who fled their country as independence approached (1959-1961) and the assisted migration of Rwandans to eastern Congo under Belgium rule (between 1937-1955) who moved with their cattle to DRC.

Both risk of spreading animal diseases and the potential generation of social conflicts related to livestock movements have led the three governments of the CEPGL to sign a regional agreement to address the regulation of cross-border cattle movements. However, there are still uncontrolled movements of cattle entering DRC through the Ruzizi valley (Brabant and Nzweve, 2013).

1.4.4 History and control challenges of East Coast fever in the Great Lakes region

Theileriosis was first reported in 1919 in the Great Lakes region while it was already described under its endemic form since 1910 in Uganda (Mortelmans and Kageruka, 1986). A veterinary laboratory was established in Butare (Rwanda) and used to serve Rwanda, Burundi, and the Belgian Congo. The disease was observed in the Gitega region (in Burundi) among tick-infested animals concentrated along the trade routes. Firstly, it was described as a relatively mild disease of calves and not of adult cattle; the case-fatality rate was estimated to range between 20-75% associated with a mortality of 30% among calves during their first year of life. The great majority of calves (80-90%) experienced a clinical form of the disease accompanied by hyperthermia and marked enlargement of lymphatic nodes. Since that period the disease was identified throughout the region and the severe disease was named by local communities as *Amakebe* in Uganda, *Ikibagalira* or *Kivagilira (Kiwagarira)* in Rwanda, Congo and Burundi, meaning “the killer” (Mortelmans and Kageruka, 1986; Van Saceghem, 1925). Later, the epidemiology and transmission of theileriosis were further described and the disease was declared as endemic in Rwanda-Burundi and Belgian Congo with mortalities confined to calves. In Rwanda, ECF cases were reported throughout the country whereas in Congo and Burundi it was found in localised areas. The high mortality and most of clinical disease usually occurred at the end of the wet seasons and calves were the most exposed to the disease, but adult cattle that move from different regions developed highly fatal disease. Researchers suggested the existence of antigenically different *T. parva* strains. *Rhipicephalus appendiculatus* was present in the whole region with the exception of the mountain regions at

high altitude ranging from 2000-2800m where it was absent or was represented by fewer number of ticks. It was reported that *R. appendiculatus* does not survive in highlands due to the cooler conditions that prevail in these areas. Control measures relied considerably on the restriction of cattle movement (between endemic and disease-free areas) and a number of dipping tanks and hand-spraying facilities have been installed to control the vector; however, only about 10% of cattle population was regularly dipped (Paling et al., 1991). At the time, farmers were using arsenic (dipping tank) to reduce tick infestation, but resistance was reported for arsenic and they started using organochlorides family (DDT, HCH, Toxaphen). Before the independence in 1960, ECF was supposed to be completely controlled among the local breeds, but exotic breeds were still at risk of developing the disease. However, only few dipping tanks remained operational and the disease was again declared as major constraint causing epidemics in the region. The disease further hinders the introduction of more productive exotic breeds of cattle, hampering the development of the livestock sector.

To cope with these challenges, live vaccination against theileriosis was initiated at the phase of laboratory testing and field validation between 1975-1991 in Rwanda and Burundi, but no such initiative was done in the Eastern Congo. A national research program was initiated in Burundi to develop ITM-based vaccination with the technical assistance from the German Agency for Technical Cooperation (GTZ) and the International Laboratory for Research on Animal Diseases (ILRAD) (Tama, 1989). The infection and treatment method of immunisation was then introduced Burundi in 1981 using a combination of stabilates from three local *T. parva* stocks (Gatumba, Gitega and Ngozi) and tetracycline treatment. The Ngozi isolate did not provide conclusive result and was later discarded due to its apparent association with ophthalmic problems. Immunisation programme have been conducted on a small scale with satisfactory results from 1981-1985 in the government farms and only 10% of calves own by private farmers. From 1986-1991, around 1781 calves were immunised, among which only 30% belonged to private farmers and 70% to the government farms. A cross-protection immunisation trial was also conducted in 1990 using Muguga cocktail stabilates, but this trial was not conclusive. The immunisation method has been used mainly on young cattle on government farms. Between 1975 and 1982, the Government of Rwanda together with the Food and Agriculture Organization (FAO) of the United Nations set up a project (1977-1982) to control TTBD is Rwanda. Immunisation against ECF began during that project and study was carried out to identify the species and strains of *Theileria*. The disease was diagnosed throughout the country, except in highlands areas were low temperature was believed to limit the distribution

of the vector. Two strains of *T. parva* were identified and used in the vaccination: *T. parva* Nyakizu and *T. parva* Satinski (Paling and Geysen, 1981). The development of a vaccine against theileriosis, based on local strains was considered to be the only practical solution at that moment. The *T. parva* Nyakizu isolate was chosen as the vaccine strain and in order to characterise this strain, various doses of the stabilates were tested in Ankole and crossbred animals. An average of 300 cattle was vaccinated each year on a voluntary basis. However, many farmers were reluctant to vaccinate their cattle because of the high cost of vaccination, the lack of qualified trained staff to monitor vaccinated animals in rural areas and the expense of drugs (tetracycline and parvaquone). During the civil war which started in 1994, the veterinary service has lost many of its facilities and qualified personnel. The local *T. parva* stocks were also unfortunately lost during the war due to lack of liquid nitrogen. Since Rwanda has emerged from civil war, there has been some progress in rehabilitation of the livestock population throughout the Great Lakes region. The current control strategy against ECF is achieved by acaricide application and treatment of sick animals using the therapeutic Buparvaquone drug. But these methods are not sustainable and are very expensive to smallholder farmers.

Recent studies conducted in Rwanda and in the Nord Kivu Province of DRC showed that the tick vector *R. appendiculatus* remains the most common Ixodid ticks and that the epidemiology of ECF is influenced by various ecological conditions and management practices which in turn affect the vector dynamics. The climate change, human population growth, change in land use and cattle movements are potential factors that affect the epidemiology of ECF in the Great Lakes region (Bazarusanga et al., 2007a; Bazarusanga et al., 2011; Bazarusanga et al., 2007b; Kalume et al., 2012; Kalume et al., 2013). Currently, the initiative of the Global Alliance of Veterinary Medicine (GALVmed) is playing the role of facilitator, accessing the feasibility of sustainable deployment of the available Muguga Cocktail vaccine in the region or to answer the question whether local isolates could serve as vaccination strains instead of the existing Muguga Cocktail. The ITM vaccine is not yet registered in the region due to unavailability of information on *T. parva* strains circulating in the field. The major challenge for successful immunisation against ECF using the ITM approach is the genetic diversity of *T. parva* genotypes in the field. As a progress, GALVmed is testing the Muguga trivalent vaccine under various controlled trials in Rwanda, DRC (Nord- Kivu) and Burundi prior to request the right to import the vaccine (Perry, 2016).

Chapter 2. Thesis objectives and overview

2.1 Introduction

As stated in the **Chapter 1**, ECF remains one of the major challenges to cattle production in the Great Lakes region, hampering the development of cattle production in the region. Although efforts have been done trying to control the disease, the epidemiology of ECF is still complex in the region: high mortality and high incidence in highlands, but clinical cases confined to calves in lowlands and midlands with reported occasional epidemics in lowlands (Amzati, 2011; Bazarusanga et al., 2007b; Kalume et al., 2012). A better understanding of factors affecting the instability of ECF in the region and epidemiological differences between AEZs could contribute to the improvement and formulation of effective and integrated control strategies. These factors may be related to the genetic composition and ecological dynamics of the tick vector and the pathogen it transmits. The biogeographical dynamics and population structure of *R. appendiculatus* and *T. parva* stocks are thought to be affected by short and long-distance movements of cattle and agro-ecological variability, which in turn may affect the transmission system of *T. parva* and the epidemiology of ECF. On the other hand, knowledge of the genetic composition of *T. parva* is crucial prior to the use of live vaccine to prevent the risk of spreading "foreign" parasite variants in new areas.

2.2 General objective

The overall objective of this thesis was to study the genetic variation, population structure and biogeographical dynamics of *R. appendiculatus* and *T. parva* to better understand the epidemiological situation of East Coast Fever (ECF) in different agro-ecological zones of the Great Lakes region (**Figure 2.1**). The results would further contribute to the implementation of appropriate control strategies, especially assuring safe deployment of live vaccine in the Great Lakes region.

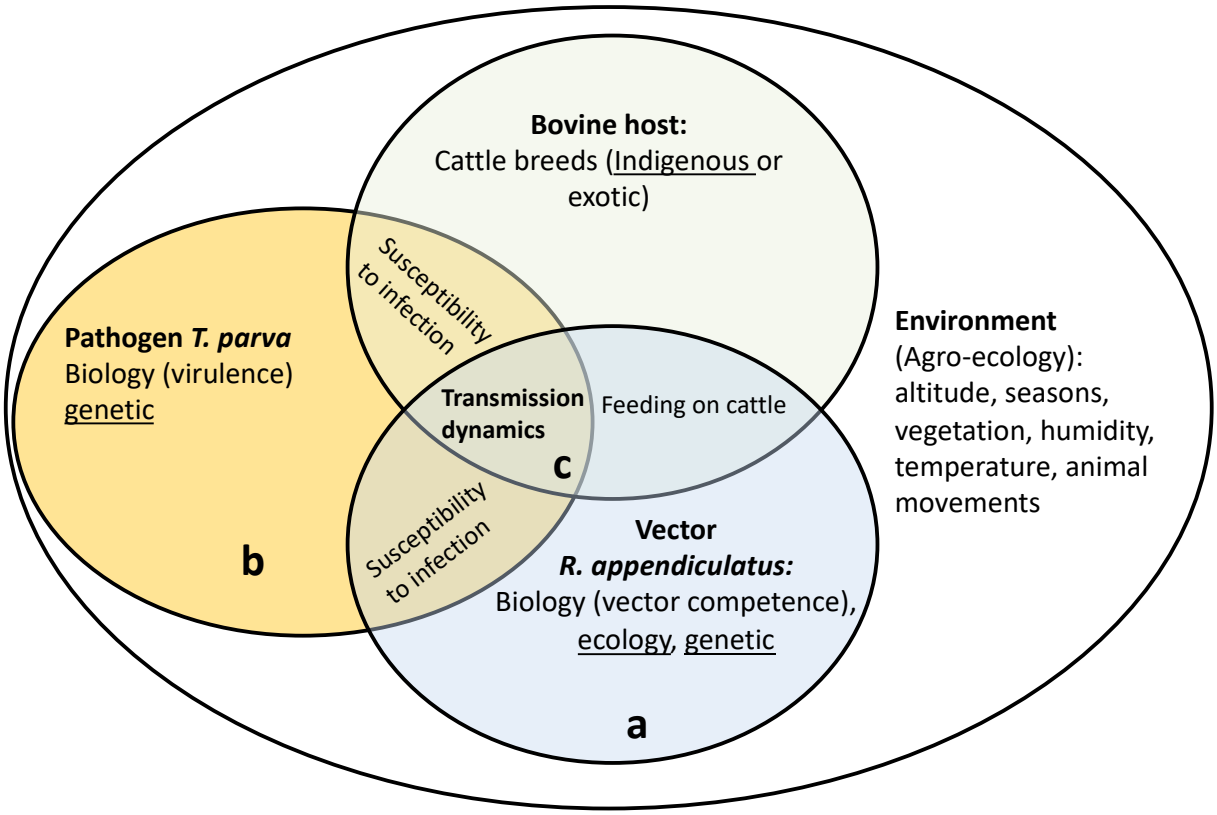

Figure 2.1 Proposed interaction between agro-ecological conditions (lowlands, midlands and highlands) and the genetic of *Theileria parva* and *Rhipicephalus appendiculatus* to explain the epidemiology of ECF: a) Study 1: population genetics of *R. appendiculatus*; b) Study 2: population genetics of *T. parva;* c) Study 3: transmission dynamics of *T. parva.*

2.3 Specific objectives

The specific objectives considered to achieve the main objective were as follows:

1. To analyse the intraspecific genetic diversity and the phylogeographical structure of *Rhipicephalus appendiculatus*
2. To investigate the genetic and antigenic variation and the population structure of *Theileria parva*
3. To model the transmission dynamics of *Theileria parva* in different agro-ecological zones using the vectorial inoculation rate (VIR) as a proxy of ECF incidence

2.4 Thesis outline

The research conducted during this thesis analysed the complexity of the epidemiological situation of ECF in the Great Lakes region in light of the genetic diversity of *R. appendiculatus* and *T. parva.* The results are discussed in relation with agro-ecological variability and the extensive movements of cattle reported in the region. Cross-sectional research and various analytic approaches were applied to study the population genetics of the tick vector and the pathogen it transmits, as well as the epidemiology of ECF. The structure of the study is illustrated in **Figure 2.2**. The contents of the thesis are outlined as follows:

The first study (**Chapter 3)** corresponds to the first specific objective of this thesis. This chapter provides data on the genetic diversity, population structure and phylogeography of *R. appendiculatus.* The tick samples came from cattle in three AEZs of DRC, two AEZs of Burundi and one AEZ of Rwanda. Phylogenetic and phylogeographical analyses were based on sequences of the cytochrome oxidase subunit I (*cox1*) and the 12S ribosomal rRNA (12s RNA). The historical evolutionary dynamics were evaluated based on mismatch models performed on *cox1* DNA sequences. The results allowed to elucidate the colonisation patterns of two lineages of *R. appendiculatus* in the Great Lakes region that have undergone sudden population growth. The analyses were expanded to the overall distribution of the tick in Africa: these data were compared with published sequences from different sub-Saharan African countries to assess their evolutionary relationships and dynamics.

The second study (**Chapter 4)** is related to the second objective of this thesis. In this chapter, we present the results of genetic and antigenic diversity and population structure of *T. parva* samples isolated from cattle blood collected in three AEZs of DRC and one AEZ of Burundi. The AEZs were defined as populations when assessing the population structure of *T. parva.* Phylogenetic, genetic differentiation and antigenic variation analyses were performed based on sequences of two antigen-coding gene markers (*Tp1* and *Tp2)*. In addition to samples collected during this study, published cattle-derived and buffalo-derived sequences were included. These originated from various ecological regions of sub-Saharan Africa to further understand the evolutionary relationships between geographical diverse *T. parva* sequences.

The third study is presented in **Chapter 5** which illustrates the current epidemiological situation related to the transmission dynamics of *T. parva* in three AEZs of DRC during three different seasons (two wet seasons and one dry season). In this study, ticks were collected from cattle

(attached ticks) and from the vegetation (free-living ticks). The attached ticks were used to estimate the tick burden while the free-living ticks were used to assess the proportion of ticks infected with *T. parva*. In addition to tick samples, cattle blood samples were analysed to estimate the prevalence of *T. parva* in the bovine host. The results were discussed in relation with the genetic composition of *R. appendiculatus* and *T. parva* in these AEZs to consolidate the genetic of the vector and the pathogen to the epidemiological data to be able to understand the complexity of ECF in the region.

Finally, **Chapters 6 and 7** summarise and consolidate the main findings obtained from these three main studies under the general discussion and conclusions and perspectives. Their potential implications in the implementation of control measure of ECF are also discussed.

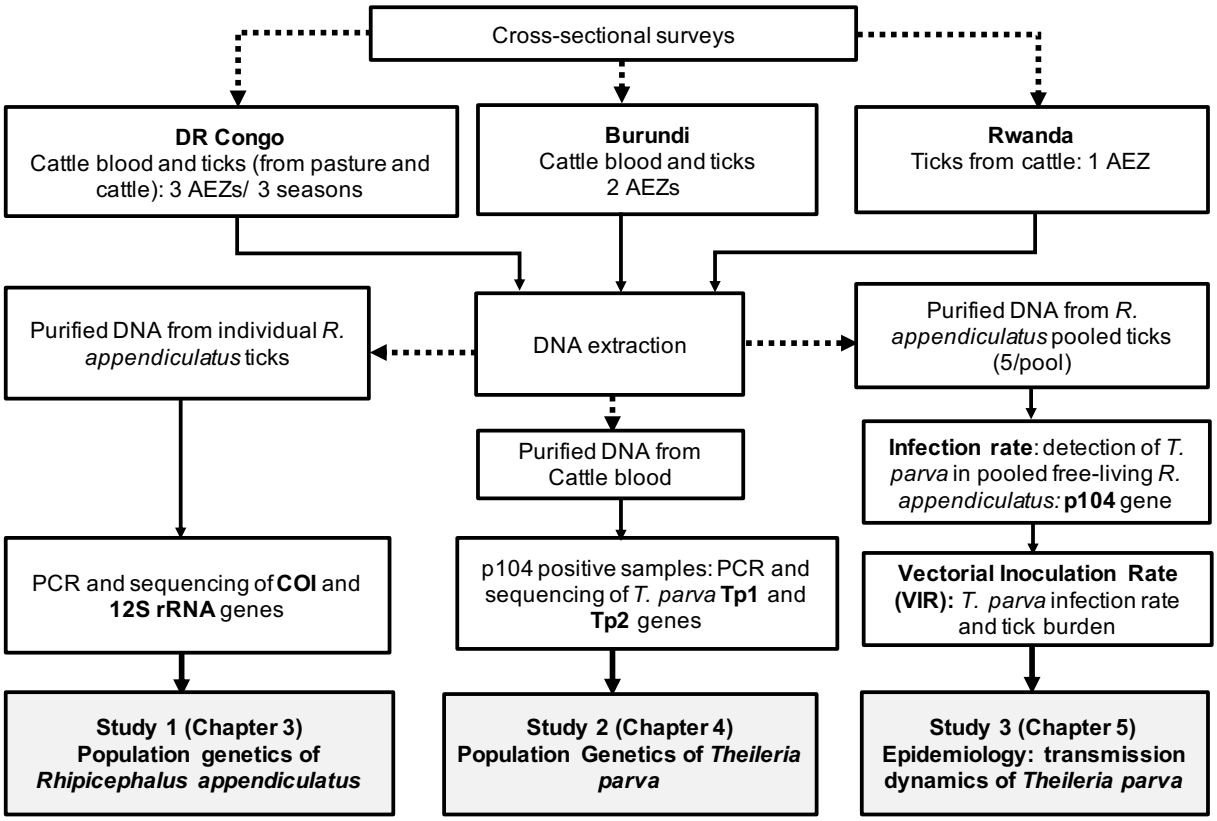

Figure 2.2 Thesis overview and study design

Chapter 3. Phylogeography and population structure of *Rhipicephalus appendiculatus* in the Great Lakes region*

This research was carried out to investigate the genetic composition and agro-ecological distribution of Rhipicephalus appendiculatus *based on mitochondrial genes, using tick samples collected from cattle in the Democratic Republic of Congo, Rwanda and Burundi. It refers to the first objective of the thesis. The results highlighted the occurrence of two mitochondrial lineages of R. appendiculatus in the Great Lakes region, with population admixture of ticks among agro-ecological zones. The "east African" lineage (lineage A) was the most abundant and has been longer well established in the region and the "south African" lineage (lineage B) was more recently introduced through cattle movement, and this is the seminal report of its presence in the region. The observed genetic pattern of the tick vector may have significant implications in the transmission system of the pathogen* Theileria parva*, resulting in the epidemic instability and complexity of ECF observed in the Great Lakes region due to possible reintroduction through cattle movements and temporal subsistence of different tick lineages on cattle.*

* *This chapter was published as:*

Amzati, G.S., Pelle, R., Muhigwa, J.-B.B., Kanduma, E.G., Djikeng, A., Madder, M., Kirschvink, N., Marcotty, T., 2018. Mitochondrial phylogeography and population structure of the cattle tick *Rhipicephalus appendiculatus* in the African Great Lakes region. Parasites & vectors 11, 329-329.

3.1 Abstract

Background: The ixodid tick *Rhipicephalus appendiculatus* is the main vector of *Theileria parva*, which causes the highly fatal cattle disease East Coast fever (ECF) in sub-Saharan Africa. *Rhipicephalus appendiculatus* populations differ in their ecology, diapause behaviour and vector competence. Thus, their expansion in new areas may change the genetic structure and consequently affect the vector-pathogen system and disease outcomes. In this study we investigated the genetic distribution of *R. appendiculatus* across agro-ecological zones (AEZs) in the African Great Lakes region to better understand the epidemiology of ECF and elucidate *R. appendiculatus* evolutionary history and biogeographical colonization in Africa. **Methods:** Sequencing was performed on two mitochondrial genes (*cox*1 and *12S* rRNA) of 218 ticks collected from cattle across six AEZs along an altitudinal gradient in the Democratic Republic of Congo, Rwanda, Burundi and Tanzania. Phylogenetic relationships between tick populations were determined and evolutionary population dynamics models were assessed by mismach distribution. **Results:** Population genetic analysis yielded 22 *cox*1 and 9 *12S* haplotypes in a total of 209 and 126 nucleotide sequences, respectively. Phylogenetic algorithms grouped these haplotypes for both genes into two major clades (lineages A and B). We observed significant genetic variation segregating the two lineages and low structure among populations with high degree of migration. The observed high gene flow indicates population admixture between AEZs. However, reduced number of migrants was observed between lowlands and highlands. Mismatch analysis detected a signature of rapid demographic and range expansion of lineage A. The star-like pattern of isolated and published haplotypes indicates that the two lineages evolve independently and have been subjected to expansion across Africa. **Conclusions:** Two sympatric *R. appendiculatus* lineages occur in the Great Lakes region. Lineage A, the most diverse and ubiquitous, has experienced rapid population growth and range expansion in all AEZs probably through cattle movement, whereas lineage B, the less abundant, has probably established a founder population from recent colonization events and its occurrence decreases with altitude. These two lineages are sympatric in central and eastern Africa and allopatric in southern Africa. The observed colonization pattern may strongly affect the transmission system and may explain ECF endemic instability in the tick distribution fringes.

3.2 Background

The ixodid brown ear tick *Rhipicephalus appendiculatus* is the main vector of the protozoan pathogen *Theileria parva*, the causative agent of a fatal lymphoproliferative cattle disease known as East Coast fever (ECF). East Coast fever is a highly pathogenic and the most economically important tick-borne disease of cattle in 12 sub-Saharan African countries, including Burundi, Democratic Republic of Congo and Rwanda (Mukhebi et al., 1992; Nene et al., 2016). *Rhipicephalus appendiculatus* is the most abundant tick in the Great Lakes region of Central Africa, where its burden and distribution vary significantly among agro-ecological zones (AEZs) (Bazarusanga et al., 2007a; Kaiser et al., 1988; Kalume et al., 2013). The geographical dispersal pattern and population dynamics of this tick are mainly driven by climatic conditions, vegetation, host availability and mobility, grazing system and management practices (Olwoch et al., 2009; Perry et al., 1990).

The Great Lakes region of Central Africa is characterised by cattle movement within and between countries for trade, breeding and pasture (Failly, 1999; Mararo, 2000). During the pre-colonial period, immigrants originating from Rwanda and Burundi settled with their cattle in Congo in search of grazing lands. In addition, political unrest during the past three decades and rapidly increasing demand for animal products increased significantly the importation of live animals (Verweijen and Brabant, 2017; Vlassenroot and Huggins, 2005). This cross-border movement of cattle across geographical areas, together with bioclimatic conditions suitable for *R. appendiculatus*, could play a significant role in spreading ticks and pathogens (Boulinier et al., 2016; De Deken et al., 2007; Maze-Guilmo et al., 2016; Yssouf et al., 2011). Therefore, the spread and establishment of ticks from one geographical region to another might be setting up a complexity in the epidemiological status and control of the disease they transmit (Barre and Uilenberg, 2010; Estrada-Peña and Salman, 2013; Fevre et al., 2006). Thus, predicting vector-borne pathogen dynamics and emergence relies on better understanding of mechanisms underlying the genetic structure of their vectors (Leo et al., 2017; Ogden et al., 2013).

Ecological establishment and population genetic structure of ticks can be affected by founder events and gene flow, largely due to their dispersal across geographical zones through host migration (Criscione et al., 2005; Leo et al., 2017). Arthropod vectors then respond differently to evolutionary forces such as migration, mutation, selection and genetic drift (promoted by bottlenecks) in their new environment (Gandon and Michalakis, 2002; Gooding, 1996). The adaptive mechanism of *R. appendiculatus* to changed environment suggests genetic divergence

between geographical stocks and phenotypic variations, including diapause behaviour and vector competence to transmit *T. parva* (Kubasu, 1992; Odongo et al., 2009). The degree to which different *R. appendiculatus* stocks acquire and transmit *T. parva* is partially genetically dependent because of the heritability of their susceptibility to infection (Young et al., 1995). Furthermore, there is a similar extensive genetic variation among *T. parva* strains in the field associated with different clinical features and disease outcomes, and variable cross-immunity levels. Studies suggest that there are also specific interactions between *R. appendiculatus* and *T. parva* stocks in the transmission dynamic system, significantly affecting the epidemiology of ECF (Ochanda et al., 1998). Thus, phylogenetic and ecological analyses should provide useful information to control ECF by determining: (i) the genetic structure of *R. appendiculatus* populations; (ii) its dispersal pattern; and (iii) its demographic history (Le Roux and Wieczorek, 2009; McCoy, 2008).

The genetic diversity of *R. appendiculatus* has been studied in different African countries using various molecular tools, such as mitochondrial DNA (Mtambo et al., 2007c; Yssouf et al., 2011) and micro- and minisatellite DNA markers (Kanduma et al., 2016b). Two distinct genetic groups have been described, namely the eastern and the southern African lineages (Mtambo et al., 2007c). More recently, Kanduma et al. (Kanduma et al., 2016a) found that the two genetic groups are present in Kenya. These evidences show that *R. appendiculatus* genetic groups may have a wide geographical range, with different ecological preferences and phenology in sub-Saharan Africa (Berkvens et al., 1998; Leta et al., 2013; Speybroeck et al., 2002), due to differences in body size (Chaka et al., 1999) and diapause induction and intensity (Madder et al., 1999; Madder et al., 2002).

Major gaps in knowledge still exist concerning the agro-ecological colonization and establishment of the *R. appendiculatus* lineages in the Great Lakes region of Central Africa, where cattle mobility seems to be the main factor of tick dispersal and epidemic instability of ECF (Bazarusanga et al., 2007b; Kalume et al., 2012; Kalume et al., 2013). Thus, further studies on the population structure and phylogeography of *R. appendiculatus*, including ticks from distinct agro-ecological conditions of DR Congo, Burundi and Rwanda, are important to shed light on the intra and inter population variation, the dispersal pattern and the historical dynamics of the characterised lineages in various ecological situations of Africa. The objective of this study was to assess the evolutionary relationship between *R. appendiculatus* populations and to investigate the impact of geographical locations on its genetic structure, to better understand

the epidemiology of ECF in the Great Lakes region. To achieve this objective, we sequenced and analysed fragments of the cytochrome *c* oxidase subunit 1 (*cox*1) and the *12S* ribosomal RNA (*12S* rRNA) gene loci. The genetic structure of *R. appendiculatus* provides clues to a better understanding of the epidemiology of ECF and insight into the development of targeted selective management and control strategies.

3.3 Methods

3.3.1 Study area

Ticks were collected from cattle in six agro-ecological zones (AEZs) of the Central African Great Lakes region (Democratic Republic of Congo, Rwanda and Burundi) (**Figure 3.1**). The three countries share the Ruzizi Valley, which consists of lowlands along the Ruzizi River flowing between Lake Kivu and Lake Tanganyika. A summary description of the characteristics of the six AEZs is presented in **Table 3.1**. Briefly, although the study area is close to the equator, the wide range of altitudes (700 to 3000 m above sea level) mitigates significantly the tropical climate attributes and therefore the AEZs limits. Generally, the rainfall period is long and bimodal (with variations in the duration between AEZs). An early rainy season starts approximately in September and ends in December, while the late rainy season lasts from February to May, followed by a dry season from June to August. There is an intervening short dry period between the two rainy seasons of about 15 days in January and/or February depending on the AEZs. Rainfall and temperature are strongly influenced by altitude ranges. Rainfall increases while temperature decreases with increasing elevation. The average temperature ranges between 12–25 °C and the annual average rainfall from 800–1100 mm in the Ruzizi Valley to 1300–2000 mm in the highlands.

In eastern DRC, ticks were collected along an altitudinal transect in South-Kivu province. Based on elevation and geographical position within the transect, three major AEZs were defined from lowlands to highlands, namely, DRC AEZ1 (lowlands) in the Ruzizi Valley, DRC AEZ2 (midlands) in the administrative district of Walungu and DRC AEZ3 (highlands) in the district of Kabare and the highlands part of Walungu (Mulumemunene). The lowlands region (DRC AEZ1) is the warmest AEZ, characterised by a tropical dry climate (semi-arid), with a cool dry season of 4–5 months. Cattle are raised generally in an open grazing system in communal pastures of savannah grassland. In contrast, the “upland Kivu” (DRC AEZ2 and

AEZ3) has a montane humid tropical climate and is much cooler. The warm dry season lasts for 3–4 months, with occasional and poorly distributed rainfall.

Ticks from Burundi were collected from two AEZs. The sampled administrative districts included Rugombo and Gihanga in the Imbo lowlands (Burundi AEZ1) and Muramvya, Mwaro and Mugamba districts of the Congo-Nile highlands (Burundi AEZ3). The Imbo region extends along the Ruzizi River and the North of Lake Tanganyika. The vegetation is composed of savannah and small patches of forest.

Ticks from Rwanda were obtained from the eastern low plateau in Nyagatare, Gatsibo and Kayonza districts (Rwanda AEZ2). This region includes the savannah of eastern province of Rwanda. The eastern semi-arid plateau zone is the most important pastoral region in Rwanda, holding 40% of the national herd raised in a communal grazing system. The vegetation type is largely savannah and some river bank woods.

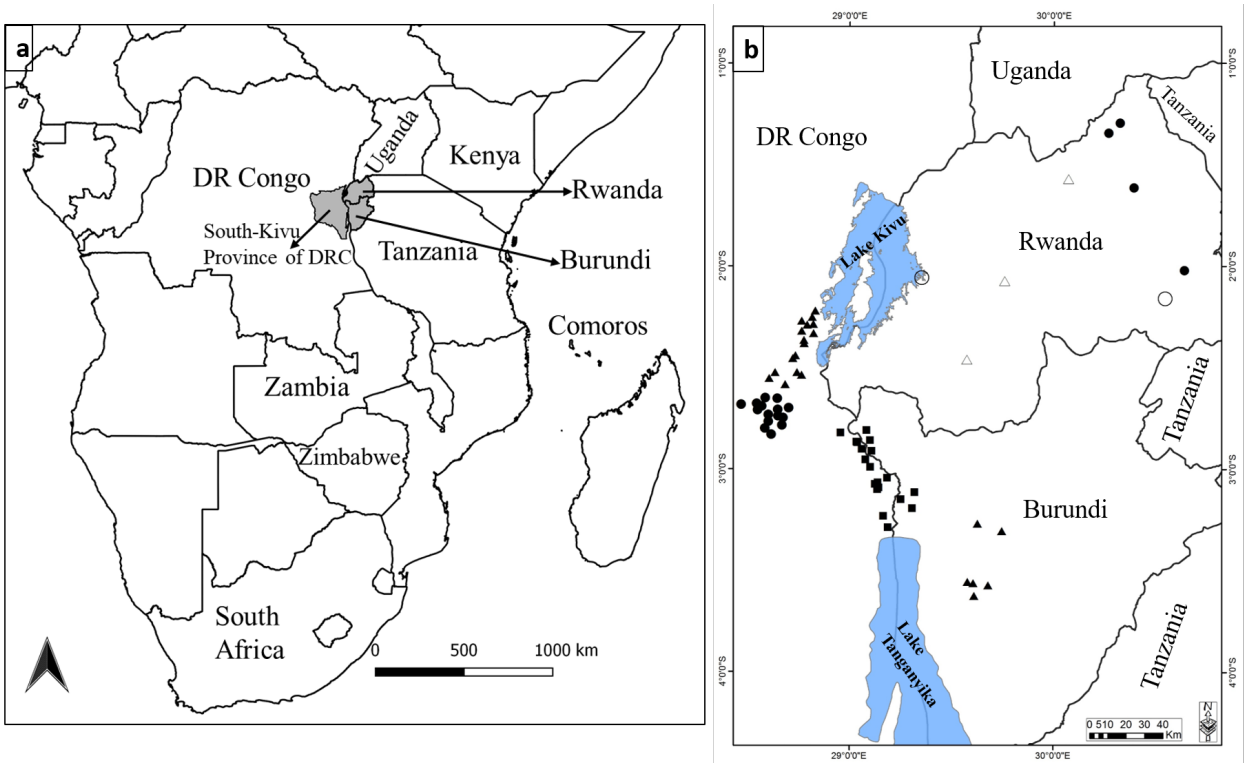

Figure 3.1 Sampling sites of *Rhipicephalus appendiculatus* ticks in DRC, Rwanda and Burundi: **a**) Map of Africa showing the study area (in grey) and other countries where the tick was previously sequenced (indicated by their names); **b**) Sampling localities of *R. appendiculatus* and their altitudes (squares: AEZ1 altitude < 1200 m; circles: AEZ2 altitude 1200–1800 m; and triangles: AEZ3 altitude: 1800-2800 m). The sites represented by empty

circles and triangles show sampling locations described by Mtambo et al. (Mtambo et al., 2007c)

Table 3.1 Geographical and climatic attributes of the six agro-ecological zones (AEZ)

Country	Agro-ecological zone (AEZ)	Altitude (m)	Temperature (°C)	Rainfall (mm/year)	Rainy season	Sample size (no. of ticks)
DRC	AEZ1	780–1100	23–25	800–1000	October-April	46
	AEZ2	1200–1800	17–21	1000–1500	September-May	54
	AEZ3	1800–2800	12–19	1350–2000	September-May	46
Burundi	AEZ1	774–1100	23–24	800–1100	November-May	26
	AEZ3	1700–2800	14–15	1300–2000	September-May	17
Rwanda [a]	AEZ2	1200–1500	21–24	800–950	November-May	20

[a] Sequences previously described in Mtambo at al. (Mtambo et al., 2007c) are not included in the 20 samples from Rwanda and thus were not used in the population genetic analysis presented in Tables 3.2–3.4
Notes: DRC (AEZ1: Lowlands, AEZ2: Midlands, AEZ3: Highlands); Burundi (AEZ1: Lowlands, AEZ3: Highlands); Rwanda (AEZ2: eastern low plateau which is the lowlands of Rwanda as described by Bazarusanga et al. (Bazarusanga et al., 2007a))

3.3.2 Tick sampling and morphological identification

Ticks were collected from cattle during a cross-sectional survey conducted from February to April 2015 (late rainy season) in the six AEZs. In each AEZ, 8 to 12 cattle herds were selected randomly using a random number generator in Microsoft Excel. From ticks collected in each herd representing a location or village, 4 to 5 ticks were selected using the same random process for further analysis. The number of ticks sampled in each population is shown in **Table 3.1**. All ticks were directly immersed in 70% ethanol for preservation and morphologically identified using the identification key of Walker et al. (Walker, 2003). Morphological identification was confirmed at the tick unit at the International Livestock Research Institute (ILRI, Kenya). Ten additional specimens originating from Simanjiro district in northern Tanzania were obtained from the Sokoine University of Agriculture in Tanzania.

3.3.3 DNA extraction, PCR amplification and sequencing

Total genomic DNA was isolated from whole individual ticks using the DNeasy® Blood and Tissue Kit (Qiagen GmbH, Hilden, Germany) according to the standard manufacturer's protocol, except that an additional incubation of 10 min at 56 °C was added after mixing the

sample with 200 µl buffer AL to ensure optimal cell lysis. Prior to extract DNA, ticks were washed twice in double distilled water and left to dry for 10 min at room temperature and homogenized.

Given the suitability of mitochondrial genes to discriminate intraspecific variation in ticks (Kanduma et al., 2016a; Mtambo et al., 2007c), we amplified the *cox*1 and *12S* rRNA gene loci to assess the genetic diversity and phylogenetic relationships of *R. appendiculatus* populations. A 710 bp fragment of *cox*1 gene locus was amplified with the forward primer LCO1490 (5'-GGT CAA CAA ATC ATA AAG ATA TTG G-3') and the reverse primer HC02198 (5'-TAA ACT TCA GGG TGA CCA AAA AAT CA-3') as previously described by Folmer et al. (Folmer et al., 1994); for the *12S* rRNA gene locus we used the primers SR-J-1499 (5'-TAC TAT GTT ACG ACT TAT-3') and SR-N-14594 (5'-AAA CTA GGA TTA GAT ACC C-3') with a fragment size of 380 bp, described in Simon et al. (Simon et al., 1994). PCR amplifications of both *cox*1 and *12S* rRNA gene fragments were performed using 50 ng of genomic DNA, 25 µl of 2× AccuPower® PCR PreMix (Bioneer PCR-PreMix, Seoul, South Korea), 0.2 µM each of forward and reverse primers, and nuclease free water added up to a final reaction volume of 50 µl. The thermal cycling program for *cox*1 consisted of an initial denaturation at 95 °C for 3 min followed by 35 cycles of denaturation at 94 °C for 1 min, annealing at 40 °C for 1 min and extension at 72 °C for 1 min. The final extension was carried out for 10 min at 72 °C. PCR parameters of *12S* rRNA gene fragment were as described for *cox*1, except that the annealing temperature was 52 °C and the extension time was 90 s. PCR products were analysed by electrophoresis on a 1.8% agarose gel. Amplicons were purified using the QIAquick® PCR Purification Kit (Qiagen GmbH, Hilden, Germany) following the manufacturer's instructions. Both strands were sequenced directly with the same primers used for PCR, on an ABI 3730 capillary sequencer (Applied Biosystems, California, USA).

3.3.4 Sequences editing, blasting and multiple alignments

Forward and reverse chromatograms for each individual tick were visually checked. Sequences were manually edited and assembled using CLC Main Workbench software v7.8.1 (CLC Bio, Aarhus, Denmark). Multiple sequences were aligned with CrustalW algorithm using default settings in the same software. The sequences were then trimmed to exclude poor quality bases and obtain uniform sizes. The final sequence sizes were 586 bp and 346 bp for *cox*1 and *12S* rRNA genes, respectively. Aligned and trimmed sequences were subjected to a BLAST search against all NCBI nucleotide databases, with default settings to confirm their species identity. A

total of 219 sequences for *cox*1 and 126 for *12S* were obtained after quality check processing. To check for amplification of nuclear pseudogenes and confirm protein coding, all *cox*1 nucleotide sequences were translated into their amino acid sequences to examine the presence of ambiguous stop codons for correct coding of invertebrate mitochondrial DNA.

3.3.5 Genetic diversity and population genetic structure

Multiple sequences extracted from the alignments were used to construct haplotypes in DnaSP software v5.10.01 (Librado and Rozas, 2009). Genetic variation within populations was estimated on *cox*1 gene sequences. Population genetic indices represented by number of haplotypes (*h*), number of segregating sites (*S*), haplotype diversity (*Hd*), mean number of pairwise nucleotide differences within population (*K*) and nucleotide diversity (π) were calculated for each AEZs (named populations) and for the overall data set using Arlequin v3.5.2.2 (Excoffier and Lischer, 2010) and DnaSP. The population genetic structure among AEZs and among haplogroups was evaluated by an analysis of molecular variance (AMOVA) performed in Arlequin. The significance of AMOVA fixation indices was evaluated based on 1,023 random permutations. In addition, to assess the level of genetic distance/differentiation between populations, we estimated gene flow expressed as of the absolute number of migrants (*Nm*) exchanged among populations, average number of nucleotide differences between populations (K_{XY}) and population pairwise genetic differentiation (F_{ST}) also in Arlequin (Meirmans and Hedrick, 2011). Their significance was tested using 1,000 random permutations. The genetic differentiation and population structure statistics were tested under the hypothesis that different populations are represented by distinct genetic groups or are exchanging migrants. These analyses were performed for combined data (of the main haplogroups identified) to understand the effect of coexistence of the haplogroups in the genetic structure and diversity.

3.3.6 Demographic and spatial expansion history analyses

The historical population dynamics and structure was inferred from mismatch distribution of *cox1* haplotypes implemented in Arlequin (Rogers and Harpending, 1992), which compared the observed distribution of pairwise nucleotide differences between haplotypes and that expected under a population expansion model for each population, haplogroup and overall data. Multimodal mismatch pattern is assumed to define a population of demographic equilibrium or constant size, whereas sudden or expanding population is characterised by unimodal distribution. The sum of square deviation (*SSD*) were estimated to determine if the observed

mismatch deviated significantly from a population expansion model, and the Harpending's raggedness index (*RI*) were used to evaluate the smoothness of mismatch distribution (Harpending, 1994). In addition, to detect deviation from neutrality expectations or mutation-drift equilibrium we performed analysis of Fu's *Fs* (Fu, 1997) and Tajima's *D* (Tajima, 1989) statistics in Arlequin and DnaSP. Tajima's *D* test is based on the difference between the number of polymorphic sites and the mean number of nucleotides pairwise differences, while Fu's *Fs* test is based on haplotype frequencies. The significance of all these statistics was tested by bootstrap resampling of 1000 coalescent simulations. Significant negative values of neutrality statistics should indicate a signature of historical event of population expansion, whereas significantly positive values indicate events such as recent population bottleneck, population subdivision or presence of some recent immigrants in a population. Values near zero and not significant, indicate that population size is constant.

3.3.7 Phylogenetic and phylogeographical reconstruction

Different published haplotype sequences of *R. appendiculatus* from South Africa, Kenya, Grande Comore, Zambia, Zimbabwe, Uganda and Rwanda were retrieved from the National Center for Biotechnology Information (NCBI) database (see Table S3.1) and were compared with sequences obtained in the present study. Phylogenetic reconstruction was performed separately on *cox*1 and *12S* rRNA gene sequences to determine the relationship among populations and possible historical dispersal event. To find the evolutionary substitution model that best describe the evolution of *cox*1 and *12S* rRNA gene sequences, we performed a hierarchical likelihood ratio test, based on the lowest Bayesian information criterion using MEGA v7.0 (Kumar et al., 2016). The nucleotide substitution model selected was then used to perform a Neighbor-joining (NJ) and/or Maximum Likelihood (ML) algorithm in MEGA. The stability and branches support were obtained using 1,000 bootstrap permutations. *Rhipicephalus eversti* and *Rhipicephalus microplus* from this study (GenBank accession numbers MF458972 and MF458973 for *cox*1 and MF479198 and MF479199 for *12S* RNA genes) and *Rhipicephalus turanicus* obtained from the GenBank (accession numbers KU880574 and DQ849231 for *cox*1 and *12S* rRNA genes, respectively) were used as outgroup taxa. A Median-joining (MJ) network was constructed to investigate the phylogenetic and ancestral relationship among haplotypes using PopArt Software (Leigh and Bryant, 2015).

3.4 Results

3.4.1 Morphological and molecular identification of ticks

PCR fragments of the mitochondrial *cox*1 and *12S* rRNA gene loci were successfully generated from 219 and 126 individual ticks, respectively, originating from the three AEZs of DRC, the two AEZs of Burundi, one AEZ of Rwanda and specimens from Tanzania. The 10 sequences from Tanzania and the sequences previously described in Rwanda were not included in the population genetic and structure analyses. Generated nucleotide sequences were aligned with the reference haplotype sequences retrieved from the GenBank (Table S3.1). The final fragment length obtained for *cox*1 was 586 bp and *12S* rRNA yielded a fragment of 346 bp long, with no indels detected in both genes. Their morphological identification has been confirmed by the high molecular identity (99–100%) with known sequences of *R. appendiculatus* found in GenBank (Table S3.2). Haplotype sequences (55 and 23 sequences for *cox*1 and *12S* rRNA, respectively) obtained for each of the six AEZs were deposited and are available in GenBank (GenBank accession numbers: MF458895–MF458949 and MF479166–MF479188 for *cox*1 and *12S* rRNA genes, respectively).

3.4.2 Phylogenetic relationships and haplotypes distribution

The overall sequence analysis revealed that *cox*1 had 27 polymorphic sites, 21 of which were parsimony informative and 6 were singletons, yielding 22 haplotypes (Table S3.3). *cox*1 amino acid sequences did not contain any internal stop codon or indel. Most nucleotide mutations were synonymous, except one non-synonymous change identified at position 32 of the haplotype CH22 (change from an Alanine to a Threonine). The highest number of *cox*1 haplotypes was found in DRC AEZ1, while the lowest was observed in Burundi AEZ3 (**Table 3.2**). The *12S* rRNA gene provided 10 polymorphic sites, five of which were parsimony informative, defining 9 haplotypes. The 22 *cox*1 haplotype sequences obtained in this study were submitted to GenBank under accession numbers MF458950-MF458971 and the nine *12S* rRNA gene haplotypes were deposited under accession numbers MF479189-MF479197. The phylogenetic relationships among *cox*1 haplotypes inferred by a NJ phylogenetic tree (**Figure 3.2**), a ML tree (**Figure 3.3**) and a MJ network (**Figure 3.4**) produced identical topologies and detected two distinct clades or lineages strongly-supported by a NJ bootstrap value of 100%. The two lineages diverged at least by 12 mutational steps (Table S3.3) but shared a wide range of agro-ecological and geographical conditions in the Great Lakes region. The first lineage, named

"haplogroup A", was represented by the most frequent haplotypes (CH1, CH2 and CH5) and comprised in total 19 haplotypes consisting of 189 of the 209 sequences analysed (90%), whereas the second lineage "haplogroup B" included three haplotypes (CH7, CH13, CH20) and had a total of 20 sequences (10%) (**Figure 3.2**, Table S3.3). Haplogroup B comprised of haplotypes present in most AEZs except the highlands zone of Burundi (Burundi AEZ3). These *cox*1 haplogroups presented a star-like pattern in the MJ network with less frequent and single haplotypes connected together to the predominant or ancestral haplotypes generally by single substitutions (**Figure 3.4**), supporting the hypothesis of a recent population expansion. The phylogenetic relationships found in the *cox*1 gene were fully congruent with those revealed by the ML tree performed on *12S* rRNA haplotypes.

The distribution of haplotypes presented in **Table 3.2** showed that there were shared and private *cox*1 haplotypes confined to restricted AEZs. Three haplotypes CH1 (27%), CH2 (28%) and CH5 (19%) were shared between all AEZs and were by far the most ubiquitous in the region, accounting for 74% (154 sequences) of the overall dataset (Table S3.3). Haplotype CH1 was detected in 13 (50%) of the 26 sequences from lowlands of Burundi (Burundi AEZ1). Two haplotypes (CH11 and CH13) defined by 10 and 11 sequences, respectively, were common in all the AEZs of DRC and in Rwanda AEZ2. Haplotype CH12 was exclusive to DRC AEZ2 and AEZ3. Nine out of the 22 haplotypes were found in single individuals; therefore, they belonged to particular AEZs (Table S3.3). Furthermore, two *12S* rRNA haplotypes (12SH1 and 12SH2) were the most abundant, representing 90% of the 126 sequences analysed. Haplotypes 12SH4 and 12SH5 had together 8 (6%) out of the 126 analysed sequences. The presence of single haplotypes indicates high frequency of rare alleles, which suggests a recent population expansion.

Table 3.2 *Rhipicephalus appendiculatus cox*1 haplotypes distribution (%) and genetic diversity indices in six agro-ecological zones of the Democratic Republic of Congo, Burundi and Rwanda

Country	AEZ	*n*	h	Haplotype (frequency in %)[a, b]	Haplogroup[c] (%)		Genetic diversity indices			
					A	B	S (PIS)	*Hd* (SD)	*K*	π (SD)
DRC	AEZ1	46	10	**CH1(30), CH2(28), CH5(6),** CH6(2), CH7(4), CH8(2), CH11(9), CH13(11), CH16(4), CH17(2)	85	15	17 (16)	0.82 (0.03)	4.4 (2.2)	0.007 (0.002)
	AEZ2	54	12	**CH1(18), CH2(30)**, **CH5(28)**, CH7(2), CH11(2), CH12(4), CH13(4), CH18(4), CH19(4), CH20(2), CH21(2), CH22(2)	92	8	20 (17)	0.81 (0.03)	3.1 (1.6)	0.005 (0.001)
	AEZ3	46	7	**CH1(17), CH2(33), CH5(28),** CH11(9), CH12(6), CH13 (4), CH14(2)	96	4	16 (15)	0.79 (0.03)	2.4 (1.3)	0.004 (0.001)
Burundi	AEZ1	26	8	**CH1(50), CH2(19)**, CH3(4), CH4(4), **CH5(11)**, CH6(4), CH7(4), CH10(4)	96	4	18 (3)	0.72 (0.08)	2.01(1.2)	0.003 (0.001)
	AEZ3	17	5	**CH1(29), CH2(35), CH5(23)**, CH8(6), CH 9(6)	100	0	4 (2)	0.77 (0.06)	1.1 (0.9)	0.002 (0.0003)
Rwanda	AEZ2	20	8	**CH1(30), CH2(20)**, CH4(5), **CH5(5)**, CH7(20), CH11(5), CH13(10), CH15(5)	70	30	18 (13)	0.85 (0.05)	6.3 (3.1)	0.011 (0.002)
Total		209	22	–	90	10	27 (21)	0.81(0.01)	3.4(1.7)	0.006 (0.0006)

[a]Haplotypes belonging to the haplogroup B are underlined; [b]Bold indicates shared haplotypes by all agro-ecological zones; [c]A and B are haplogroup names

Abbreviations: AEZ, agro-ecological zones; *n*, number of sequences; h, number of haplotypes; S, segregation sites; PIS, parsimony informative sites; Hd: haplotype diversity; SD, standard deviation; k: mean number of pairwise nucleotide differences; π, nucleotide diversity; CH1-22: names of *cox*1 haplotype

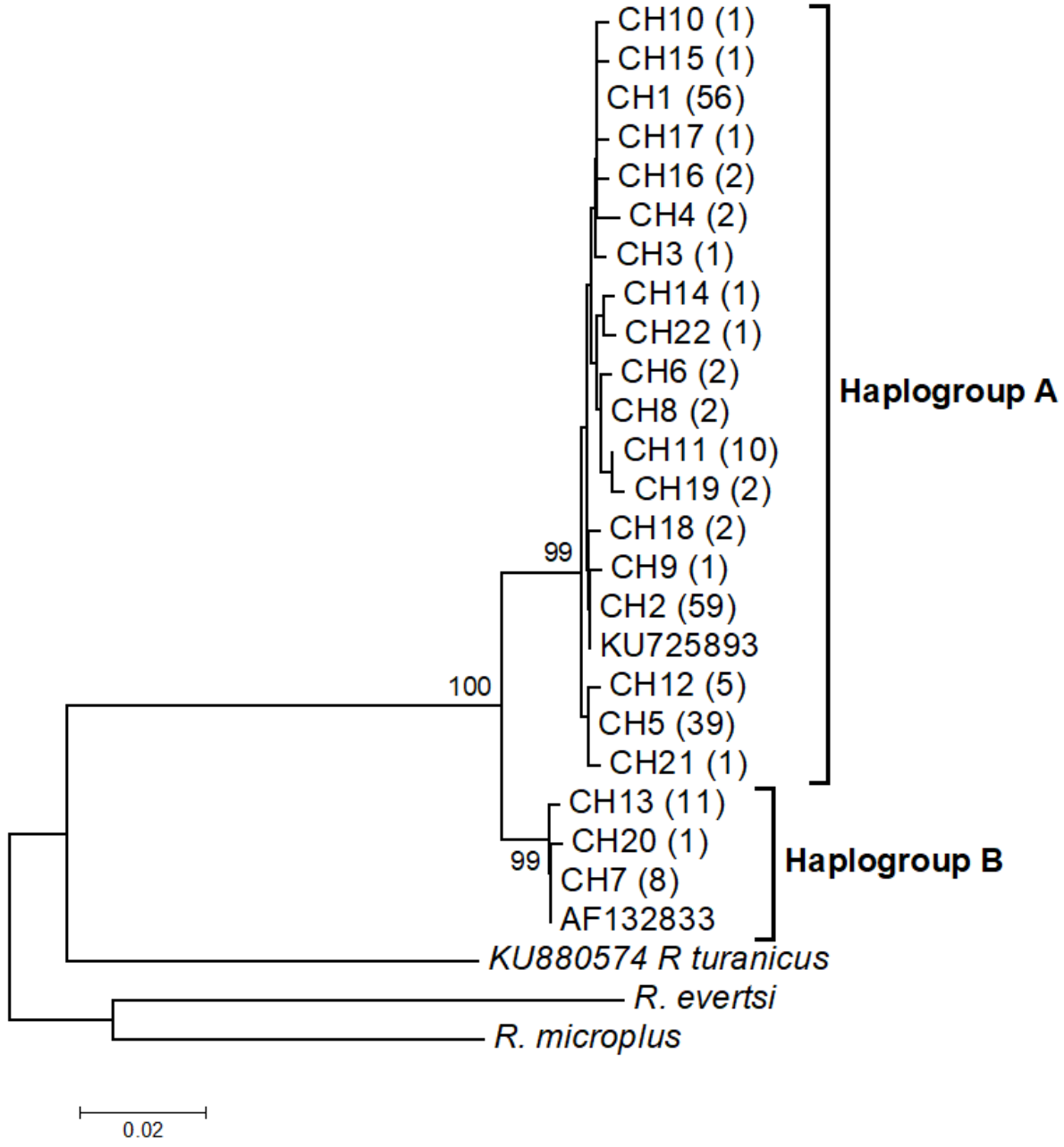

Figure 3.2 Phylogenetic tree of *R. appendiculatus cox*1 haplotypes. The evolutionary history was inferred by using the neighbor-joining method based on the Tamura 3-parameter model. A discrete Gamma distribution was used to model evolutionary rate differences among sites. Bootstrap values (> 80) are displayed above nodes. CH1-22 are haplotype names. The values in parentheses correspond to the frequency of each haplotype. KU725893 and AF132833 are GenBank accession numbers for *R. appendiculatus* sequences used as reference haplotypes from Kenya and Zimbabwe, respectively. Two species (*R. eversti* and *R. microplus*) obtained in this study and *R. turanicus* from GenBank (accession number: KU880574) were included as outgroups

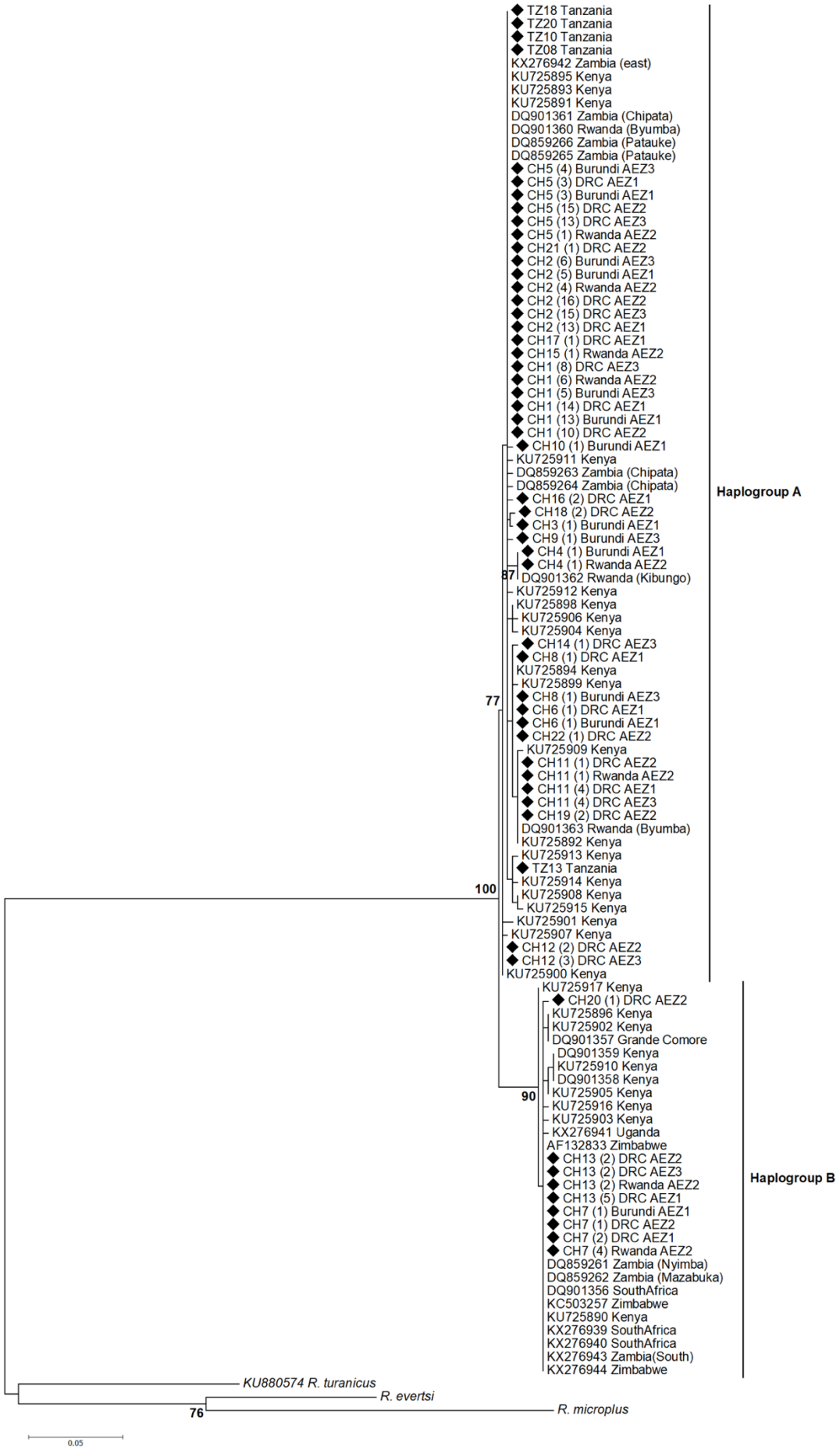
TZ18 Tanzania
TZ20 Tanzania
TZ10 Tanzania
TZ08 Tanzania
KX276942 Zambia (east)
KU725895 Kenya
KU725893 Kenya
KU725891 Kenya
DQ901361 Zambia (Chipata)
DQ901360 Rwanda (Byumba)
DQ859266 Zambia (Patauke)
DQ859265 Zambia (Patauke)
CH5 (4) Burundi AEZ3
CH5 (3) DRC AEZ1
CH5 (3) Burundi AEZ1
CH5 (15) DRC AEZ2
CH5 (13) DRC AEZ3
CH5 (1) Rwanda AEZ2
CH21 (1) DRC AEZ2
CH2 (6) Burundi AEZ3
CH2 (5) Burundi AEZ1
CH2 (4) Rwanda AEZ2
CH2 (16) DRC AEZ2
CH2 (15) DRC AEZ3
CH2 (13) DRC AEZ1
CH17 (1) DRC AEZ1
CH15 (1) Rwanda AEZ2
CH1 (8) DRC AEZ3
CH1 (6) Rwanda AEZ2
CH1 (5) Burundi AEZ3
CH1 (14) DRC AEZ1
CH1 (13) Burundi AEZ1
CH1 (10) DRC AEZ2
CH10 (1) Burundi AEZ1
KU725911 Kenya
DQ859263 Zambia (Chipata)
DQ859264 Zambia (Chipata)
CH16 (2) DRC AEZ1
CH18 (2) DRC AEZ2
CH3 (1) Burundi AEZ1
CH9 (1) Burundi AEZ3
CH4 (1) Burundi AEZ1
CH4 (1) Rwanda AEZ2
87
DQ901362 Rwanda (Kibungo)
KU725912 Kenya
KU725898 Kenya
KU725906 Kenya
KU725904 Kenya
CH14 (1) DRC AEZ3
CH8 (1) DRC AEZ1
KU725894 Kenya
KU725899 Kenya
77
CH8 (1) Burundi AEZ3
CH6 (1) DRC AEZ1
CH6 (1) Burundi AEZ1
CH22 (1) DRC AEZ2
KU725909 Kenya
CH11 (1) DRC AEZ2
CH11 (1) Rwanda AEZ2
CH11 (4) DRC AEZ1
CH11 (4) DRC AEZ3
CH19 (2) DRC AEZ2
DQ901363 Rwanda (Byumba)
KU725892 Kenya
KU725913 Kenya
TZ13 Tanzania
100
KU725914 Kenya
KU725908 Kenya
KU725915 Kenya
KU725901 Kenya
KU725907 Kenya
CH12 (2) DRC AEZ2
CH12 (3) DRC AEZ3
KU725900 Kenya
Haplogroup A
KU725917 Kenya
CH20 (1) DRC AEZ2
KU725896 Kenya
KU725902 Kenya
DQ901357 Grande Comore
DQ901359 Kenya
KU725910 Kenya
DQ901358 Kenya
90
KU725905 Kenya
KU725916 Kenya
KU725903 Kenya
KX276941 Uganda
AF132833 Zimbabwe
CH13 (2) DRC AEZ2
CH13 (2) DRC AEZ3
CH13 (2) Rwanda AEZ2
CH13 (5) DRC AEZ1
CH7 (1) Burundi AEZ1
CH7 (1) DRC AEZ2
CH7 (2) DRC AEZ1
CH7 (4) Rwanda AEZ2
DQ859261 Zambia (Nyimba)
DQ859262 Zambia (Mazabuka)
DQ901356 SouthAfrica
KC503257 Zimbabwe
KU725890 Kenya
KX276939 SouthAfrica
KX276940 SouthAfrica
KX276943 Zambia(South)
KX276944 Zimbabwe
Haplogroup B
KU880574 R. turanicus
R. evertsi
76
R. microplus
0.05

Figure 3.3 Phylogenetic tree of *cox*1 haplotypes displaying the relationship between the *R. appendiculatus* specimens in sub-Saharan African countries. The evolutionary history was inferred by using the maximum likelihood 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 = 0.39)]. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. Bootstrap scores > 70 are displayed to support nodes. The values in bracket behind haplotypes names correspond to the frequency of each haplotype. Haplotype sequences (CH1-22) obtained in the present study are indicated by a black square. *Rhipicephalus eversti* and *R. microplus* obtained in this study and *R. turanicus* (GenBank: KU880574) were used as outgroups.

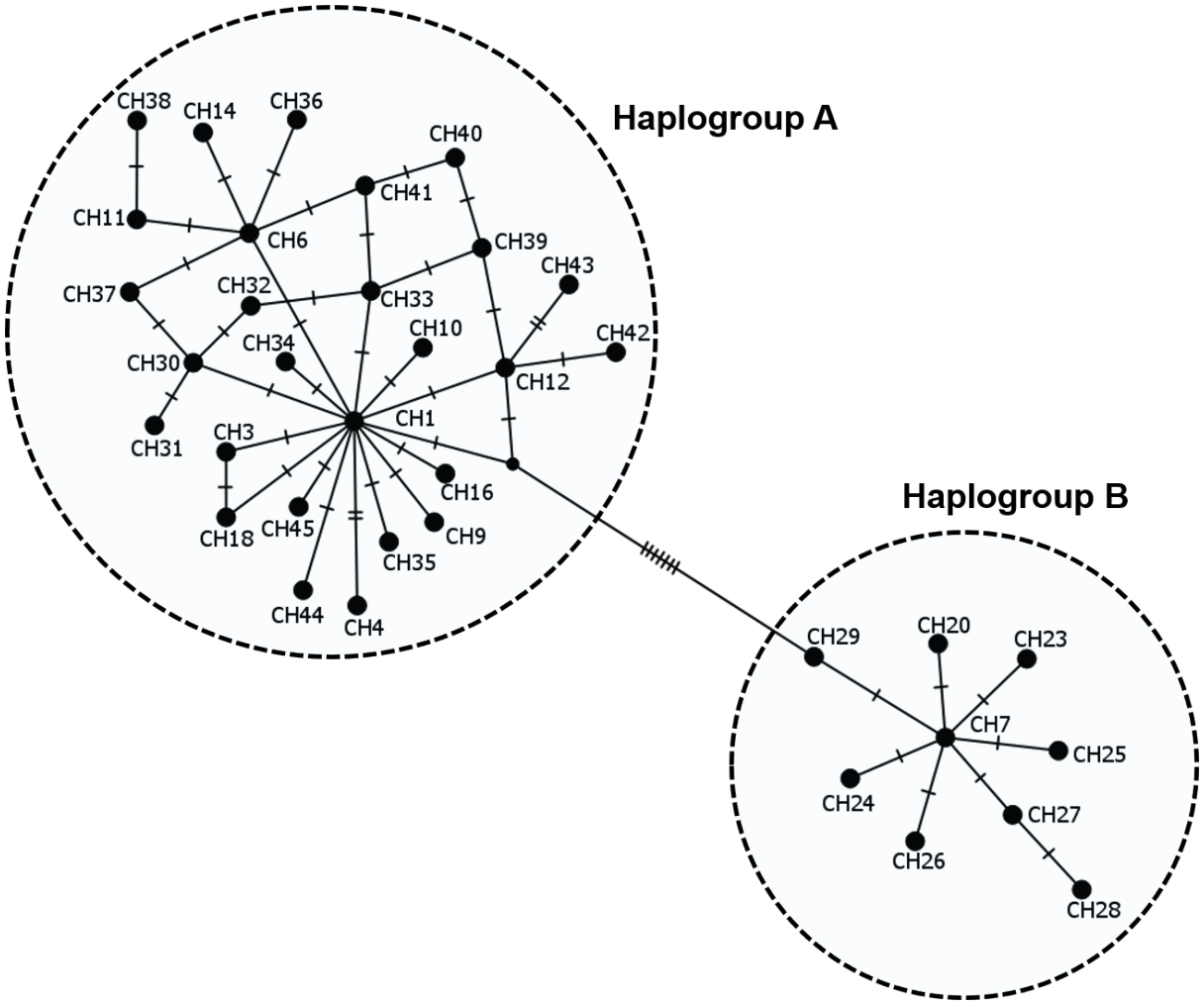

Figure 3.4 Median-joining network of the 36 *cox*1 haplotypes for *R. appendiculatus* ticks across sub-Saharan African countries. Lines represent mutations and the dot corresponds to a possible intermediate haplotype. Each circle denotes a unique haplotype. Haplotype frequencies are not shown here, but their occurrences across Africa are presented in **Table 3.5**

3.4.3 Population genetic diversity

Population genetic indices were estimated using *cox*1 nucleotide sequences and are shown in **Table 3.2**. The overall haplotype diversity (*Hd*) and nucleotide diversity (π) were 0.81 and 0.006, respectively. The haplotype diversity ranged from 0.72 in Burundi AEZ1 to 0.85 in Rwanda AEZ2 and the nucleotide diversity (π) values ranged from 0.002 in Burundi AEZ3 to 0.011 in Rwanda AEZ2. The average number of nucleotides pairwise differences (k) was 3.4 for the overall dataset, with the highest value observed in Rwanda AEZ2 (6.3). Altogether, these population genetic indices showed that the diversity of *R. appendiculatus* is quite different among AEZs. Ticks from Rwanda AEZ2 were more highly diverse than those from three AEZs of DRC and the two AEZs of Burundi. We also observed an excess of singleton mutations in Burundi AEZ1 (15 out of 18 polymorphic sites), suggesting a sudden population expansion. These data analysed separately by haplogroups (**Table 3.3**), showed that the differences of genetic diversity among AEZs was largely affected by the frequency distribution of the two *cox*1 haplogroups (**Table 3.2**).

Table 3.3 Genetic diversity and evolutionary dynamics of the two haplogroups (A and B) identified from *cox*1 sequences of *R. appendiculatus*

Genetic indices and statistics	**Haplogroup A**							**Haplogroup B**	**Overall data set**
	DRC			**Burundi**		**Rwanda**	**Haplogroup A (overall)**		
	AEZ1	**AEZ2**	**AEZ3**	**AEZ1**	**AEZ3**	**AEZ2**			
Diversity indices									
Number of sequences	39	50	44	25	17	14	189	20	209
Number of polymorphic sites	7	9	6	8	4	7	18	2	27
Number of Haplotypes	8	9	6	7	5	6	19	3	22
Haplotype diversity	0.76	0.78	0.77	0.7	0.77	0.77	0.77	0.56	0.81
(SD)	(0.044)	(0.032)	(0.033)	(0.084)	(0.057)	(0.089)	(0.016)	(0.063)	(0.01)
Nucleotide diversity	0.002	0.0025	0.002	0.002	0.002	0.002	0.002	0.001	0.006
(SD)	(0.0003)	(0.0003)	(0.0003)	(0.0004)	(0.0003)	(0.0006)	(0.0001)	(0.0002)	(0.0006)
Neutrality tests									
Tajima's *D*	-0.6	-0.79	0.096	-1.4	-0.21	-1.4	-1.5	0.24	-0.93
(*P*-value)	(0.35)	(0.23)	(0.6)	(0.049)*	(0.45)	(0.07)	(0.032)*	(0.67)	(0.23)
Fu's *Fs*	-2.4	-2.5	-0.1	-2.6	-1.1	-2.1	-10.4	0.2	-3.8
(*P*-value)	(0.08)	(0.097)	(0.49)	(0.027)*	(0.18)	(0.041)*	(0.001)*	(0.49)	(0.15)
Demographic expansion									
Sum of Squared deviation	0.002	0.004	0.002	0.002	0.02	0.001	0.0008	0.029	0.017
(*P*-value)	(0.66)	(0.31)	(0.49)	(0.72)	(0.17)	(0.9)	(0.44)	(0.049)*	(0.1)
Harpending's Raggedness index	0.06	0.079	0.06	0.05	0.16	0.06	0.057	0.21	0.049
(*P*-value)	(0.39)	(0.21)	(0.38)	(0.71)	(0.12)	(0.73)	(0.2)	(0.1)	(0.51)
Spatial expansion									
Sum of Squared deviation	0.002	0.004	0.002	0.002	0.02	0.001	0.0008	0.029	0.036
(*P*-value)	(0.5)	(0.21)	(0.4)	(0.67)	(0.1)	(0.9)	(0.26)	(0.011)*	(0.28)
Harpending's Raggedness index	0.064	0.079	0.062	0.05	0.16	0.06	0.057	0.21	0.049
(*P*-value)	(0.39)	(0.21)	(0.43)	(0.75)	(0.12)	(0.72)	(0.22)	(0.12)	(0.74)

*Values are statistically significant at $P < 0.05$; significance was determined using 1000 coalescent simulations

Abbreviations: *D*, Tajima's neutrality statistic; *Fs*, Fu's neutrality statistic

3.4.4 Population structure and ecological differentiation based on *cox*1 haplotypes

Analysis of molecular variance (AMOVA) based on *cox*1 sequences revealed statistically significant variance among the six AEZs analysed together for both combined haplogroups and haplogroup A alone (6%, $P < 0.001$) (Table S3.4). The genetic variability among individuals within AEZs explained the large majority of molecular variance (94%) for the overall dataset, suggesting an admixture between different populations and the coexistence of the two genetically divergent lineages (**Figure 3.2**, **Table 3.2**). This is consistent with moderate genetic structure of *R. appendiculatus* in the Great Lakes region. The population differentiation indices were then calculated to compare the genetic variation observed among AEZs (**Table 3.4**). Both differentiation statistics based on pairwise estimates of the Inter-population nucleotide differences (K_{XY}), the pairwise genetic distance (F_{ST}), and the number of migrants (Nm) showed evidence of differentiation between some *R. appendiculatus* populations. The population pairwise genetic distance (F_{ST}) varied from a negative and not significant value ($F_{ST} = -0.014$, $P = 0.87$) with infinite value of Nm (between DRC AEZ2 and AEZ3) to the strongest differentiation ($F_{ST} = 0.19$, $P = 0.005$) with low number of migrants ($Nm = 2.1$) between Burundi AEZ3 and Rwanda AEZ2. Pairwise F_{ST} values were found to be low and not significant between DRC AEZ3 and Burundi AEZ3 with infinite Nm and between DRC AEZ2 and Burundi AEZ3 associated with a high number of migrants ($Nm = 11{,}875$), indicating an excess of gene flow between these zones. DRC AEZ1 did not differ significantly with Burundi AEZ1 ($F_{ST} = 0.022$, $P = 0.19$) and Rwanda AEZ2 ($F_{ST} = 0.018$, $P = 0.24$). Tick populations from Rwanda AEZ2 showed a strong genetic differentiation with those from DRC AEZ3 ($F_{ST} = 0.18$, $P < 0.001$) and DRC AEZ2 ($F_{ST} = 0.14$, $P = 0.03$), indicating reduced or lack of gene flow between these populations. These results were confirmed by inter-population nucleotide differences (K_{XY}), which was significant when populations were significantly differentiated by the pairwise F_{ST} statistic. Tick populations were compared by analysing haplogroup A sequences alone. The findings showed that ticks from Rwanda AEZ2 and the two AEZs of Burundi were not genetically different. They shared more migrants belonging to haplogroup A (**Table 3.4**). In general, the population differentiation was observed between lowlands and highlands AEZs.

Table 3.4 Population genetic statistics for pairwise comparison of different populations of *R. appendiculatus* from sequences of *cox*1 gene. Values in parentheses represent the *P*-value statistics

Population 1	Population 2	Haplogroup A and B			Haplogroup A		
		K_{XY}	F_{ST}	*Nm*	K_{XY}	F_{ST}	*Nm*
DRC AEZ1	DRC AEZ2	3.8 (0.044)*	0.036 (0.044)*	13.2	1.5 (< 0.001)*	0.11 (< 0.001)*	4
DRC AEZ1	DRC AEZ3	3.6 (0.032)*	0.057 (0.032)*	8.3	1.5 (0.001)*	0.1 (0.001)*	4.3
DRC AEZ1	Burundi AEZ1	3.3 (0.15)	0.022 (0.19)	22.2	1.2 (0.31)	0.002 (0.33)	243.6
DRC AEZ1	Burundi AEZ3	3 (0.059)	0.056 (0.087)	8.4	1.3 (0.087)	0.043 (0.082)	11.2
DRC AEZ1	Rwanda AEZ2	5.4 (0.3)	0.018 (0.24)	26.8	1.3 (0.78)	-0.026 (0.79)	∞
DRC AEZ2	DRC AEZ3	2.7 (0.85)	-0.014 (0.87)	∞	1.4 (0.93)	-0.016 (0.93)	∞
DRC AEZ2	Burundi AEZ1	2.7 (0.003)*	0.060 (0.011)*	7.9	1.5(< 0.001)*	0.13 (< 0.001)*	3.2
DRC AEZ2	Burundi AEZ3	2.1 (0.23)	0.00004 (0.43)	11876	1.3 (0.46)	-0.012 (0.54)	∞
DRC AEZ2	Rwanda AEZ2	5.2 (0.003)*	0.14 (0.003)*	3.1	1.6 (0.006)*	0.11 (0.004)*	3.8
DRC AEZ3	Burundi AEZ1	2.4 (< 0.001)*	0.076 (0.006)*	6.1	1.5(< 0.001)*	0.14 (< 0.001)*	3.2
DRC AEZ3	Burundi AEZ3	1.8 (0.32)	-0.005 (0.43)	∞	1.3 (0.39)	-0.009 (0.46)	∞
DRC AEZ3	Rwanda AEZ2	5.1 (< 0.001)*	0.18 (< 0.001)*	2.2	1.6 (0.009)*	0.11 (0.011)*	3.9
Burundi AEZ1	Burundi AEZ3	1.6 (0.074)	0.034 (0.08)	12.2	1.2 (0.077)	0.063 (0.073)	7.4
Burundi AEZ1	Rwanda AEZ2	4.8 (0.030)*	0.14 (0.03)*	3.1	1.2 (0.97)	-0.039 (0.98)	∞
Burundi AEZ3	Rwanda AEZ2	4.6 (0.005)*	0.19 (0.005)*	2.1	1.3 (0.16)	0.044 (0.14)	10.9

*Values are statistically significant at $P < 0.05$; significance was determined using 1000 coalescent simulations.

Abbreviations: K_{XY}, average number of nucleotide differences between populations; F_{ST}, pairwise genetic distance *F*-statistic based on nucleotide sequences (Right's fixation index); *Nm*, number of migrants between populations

3.4.5 Demographic and dispersal dynamics of *R. appendiculatus*

Demographic and spatial dynamics inferred from pairwise nucleotide differences revealed bimodal pattern for the total dataset (**Figure 3.5a**). Tajima's *D* and Fu's *Fs* were negative but not significant (**Table 3.3**). However, the sum of square deviation (*SSD*) and raggedness index (*RI*) for both demographic and range expansion did not deviate significantly from that expected under expansion model. The negative values of neutrality tests and the non-significance of *SSD* and *RI* indices, suggest a sudden expansion of *R. appendiculatus* populations in the Great Lakes region. The population dynamics history was also analysed separately for each of the six AEZs (**Figure 3.6**, Table S3.5). Most AEZs showed a bimodal mismatch distribution, except in Burundi AEZ3 where the mismatch pattern was unimodal (**Figure 3.6e**). The observed bimodal pattern suggested population subdivision as shown by the existence of two well resolved haplogroups (**Figure 3.2**).

The population dynamics were then inferred for each haplogroup separately (**Table 3.3**). A unimodal distribution was observed in both haplogroups (**Figure 3.5b, c**). For the haplogroup A, we detected significant evidence of demographic and spatial expansion events from the unimodal mismatch distribution, together with significantly negative values for Tajima's *D* (*D* = -1.5, *P* = 0.032) and Fu's *Fs* (*Fs* = -10.4, *P* = 0.001) statistics. A good fit of sudden population expansion was also observed in this haplogroup, based on sum of squared deviation values that were not significant in all the cases: demographic (*SSD* = 0.0008, *P* = 0.44) and range (*SSD* = 0.0008, *P* = 0.26) expansion, with no significant variation of the raggedness index for both models (**Table 3.3**). In contrast, haplogroup B did not display any significant signature of expansion from the selective neutrality tests. In addition, the observed mismatch pattern for this haplogroup deviated significantly from that expected under population expansion scenario (*SSD* = 0.029, *P* = 0.049 and *P* = 0.011 for demographic and spatial expansion, respectively), implying that haplogroup B did not experience any historical population expansion. This group is characterised either by a constant population size (demographic equilibrium) or had experienced a weak signal of population bottleneck that reduced its diversity. When analysing the demographic dynamics for samples belonging to haplogroup A in each AEZ (**Table 3.3**), population expansion signal was confirmed in all AEZs by mismatch analyses exhibiting unimodal distribution (Figure S3.1). For the six AEZs, none of the statistical comparisons between the observed and the expected distributions rejected the sudden and range expansion assumptions based on the raggedness index and the sum of squared deviation. The neutrality

indices were generally negative but not significant, except in Burundi AEZ1 where both Tajima's *D* (*D* = -1.4, *P* = 0.049) and Fu's *Fs* (*Fs* = -2.6, *P* = 0.027) statistics showed significant negative values and in Rwanda AEZ2 where Fu's *Fs* was negative (*Fs* = -2.1) and significant (*P* = 0.041). According to the mismatch distribution and negative values for neutrality tests, the hypothesis of population growth and spatial expansion models could not be rejected in the six AEZs, which was consistent with a model of sudden expansion for each population subdivision.

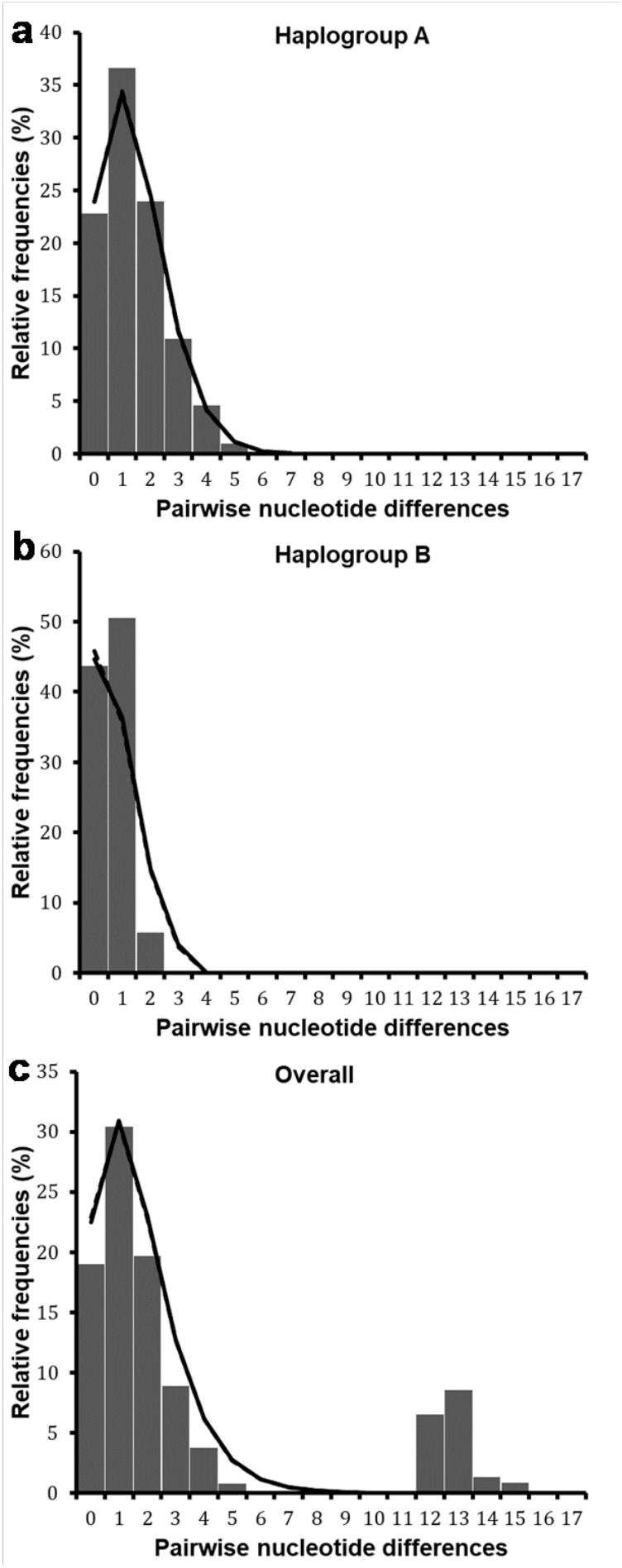
a
Haplogroup A
Relative frequencies (%)
0 5 10 15 20 25 30 35 40
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
Pairwise nucleotide differences
b
Haplogroup B
Relative frequencies (%)
0 10 20 30 40 50 60
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
Pairwise nucleotide differences
c
Overall
Relative frequencies (%)
0 5 10 15 20 25 30 35
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
Pairwise nucleotide differences

Figure 3.5 *cox*1 mismatch distribution pattern for two haplogroups of *R. appendiculatus*: **a)** and **b)** show the mismatch distribution of nucleotide sequences in haplogroups A and B, respectively. **c)** shows the overall mismatch distribution pattern for all AEZs and lineages. The x-axis shows the number of pairwise differences between pairs of haplotype sequences and the y-axis shows their frequencies (in %). The observed frequencies are represented by solid histograms and the simulated mismatch distributions expected under demographic expansion (solid black line) and under spatial expansion (dotted black line). Simulated curves under range and demographic expansion have same pattern in these figures, they overlapped.

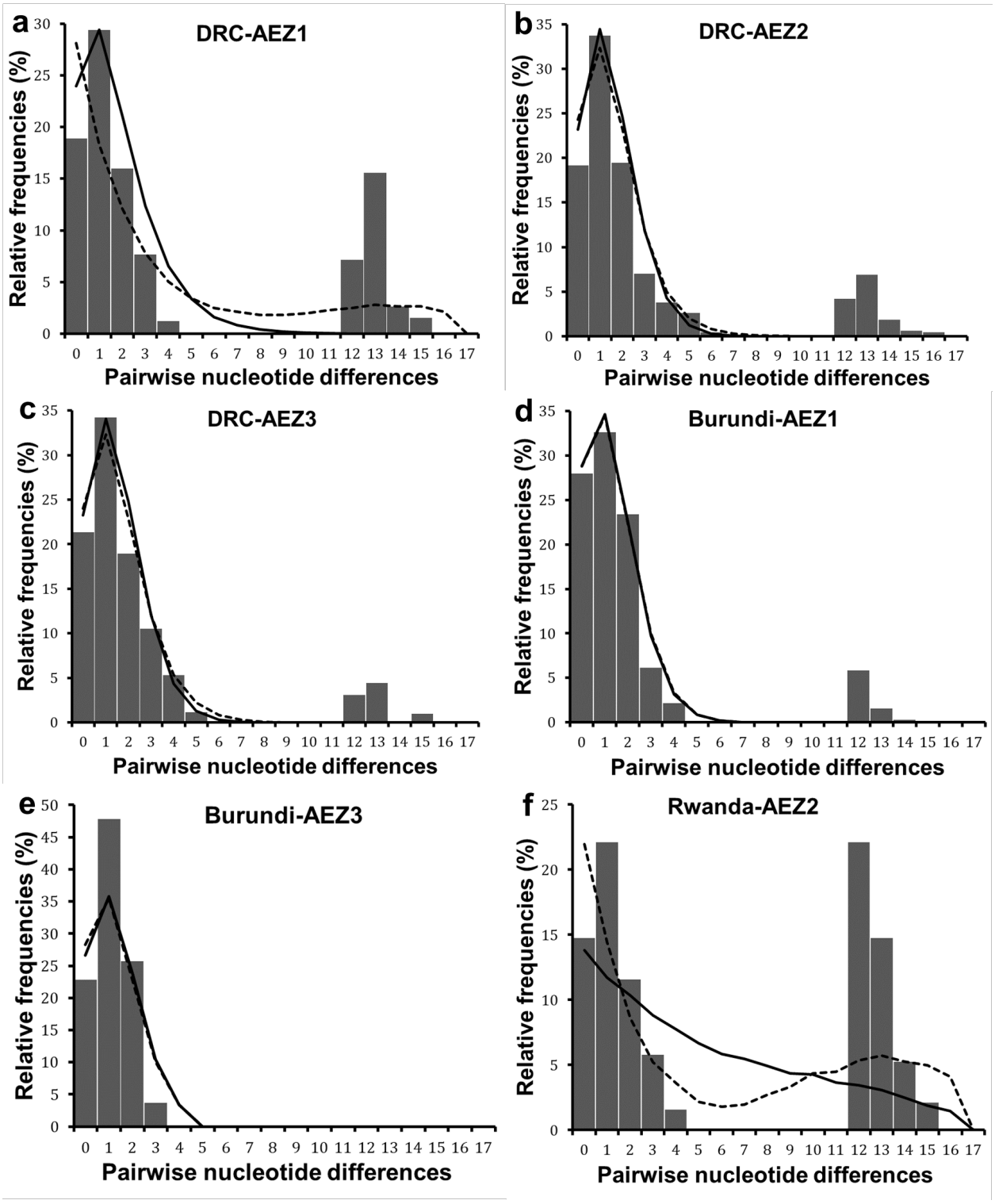

Figure 3.6 *cox*1 mismatch distribution pattern for six populations of *R. appendiculatus*. **a**, **b** and **c** show the mismatch distribution pattern for *R. appendiculatus* from DRC AEZs (AEZ1, 2 and 3, respectively); **d** and **e** represent the mismatch pattern of ticks from Burundi AEZ1 and AEZ3, respectively; **f** depicts the mismatch distribution of ticks from Rwanda AEZ2. The x-axis shows the number of pairwise differences between pairs of haplotype sequences and the y-

axis shows their frequencies in %. The observed frequencies are represented by solid histograms. Black full line represents the expected distribution under sudden expansion model, and dotted line represents the distribution simulated under spatial expansion model. Simulated curves under spatial and demographic expansion have same pattern in (**d**), and they overlapped.

3.4.6 Phylogeographical structure

To study the phylogeographical structure of *R. appendiculatus* in Africa, the haplotype sequences found in this study along with those retrieved from GenBank (Table S3.1) were used to reconstruct the phylogenetic using ML tree and the MJ network methods based on *cox*1 and *12S* rRNA genes. Thirty-six *cox*1 haplotypes were identified in 105 sequences including 50 haplotypes from GenBank and the 55 haplotype sequences obtained in our study for each of the six AEZs and for Tanzania (**Table 3.5**). Twenty-eight haplotypes had been already described in different African countries, and eight haplotypes (CH3, CH9, CH10, CH14, CH16, CH18, CH20 and CH33) were newly described in the present study. Most of these new haplotypes were less abundant (Table S3.3) and diverged from the common ancestral haplotypes generally by only one substitution (**Figure 3.4**). Haplotypes CH1, CH7, CH11 were the most ubiquitous and shared wide geographical distribution in affected African countries. CH1 haplotype was shared by Kenya, Eastern Zambia, DRC (AEZ1, 2 and 3), Burundi (AEZ1 and 3), Tanzania and Rwanda (AEZ2 and GenBank sequences), while haplotype CH7 was reported in Kenya, South Africa, Zimbabwe, Grande Comore, Eastern and southern provinces of Zambia, DRC (AEZ1, 2 and 3), Burundi (AEZ1) and Rwanda (AEZ2). Haplotype CH11 was present in Kenya, Rwanda (AEZ2 and GenBank sequences), Comoros and DRC (AEZ1, 2 and 3). Eighteen haplotypes mostly with unique sequences were found to be restricted to Kenya and have not been reported in any other country. In the same way, haplotype CH23 was particular to Uganda. This country did not share any haplotype with other countries. The ML phylogenetic tree reconstructed using the 105 sequences indicated that globally, the African tick *R. appendiculatus* is consistently clustered into two groups (haplogroups A and B) well-supported by a ML bootstrap value of 100% (**Figure 3.3**). Our 19 haplotypes that had been described for haplogroup A (**Table 3.2**, Table S3.3) formed one clade with 19 haplotypes from Kenya (19/29), all the 3 haplotypes from Rwanda (3/3), six haplotypes from Eastern province of Zambia (6/7), all the five haplotypes from Tanzania, whereas our three haplotypes of haplogroup B were clustered with 10 haplotypes from Kenya, the single haplotype from Grande Comore, all haplotypes from South Africa (3/3), Zimbabwe (3/3), Uganda (2/2), Southern

province of Zambia (2/2), and one haplotype from Eastern province of Zambia. In addition to ML tree, the MJ network also revealed that *R. appendiculatus* is divided into two main groups in Africa, separated by seven mutational steps (**Figure 3.4**).

Similar findings were confirmed by *12S* rRNA gene. Our 23 *12S* rRNA individual haplotypes from each of the six AEZs were analysed together with the 29 sequences obtained from GenBank (Table S3.1). Fourteen haplotypes were observed, two most common (12SH1 and 12SH5) and 12 minors (defined mostly by one sequence or restricted to particular country) (Table S3.6). Haplotype 12SH1 was common in DRC, Burundi, Rwanda, Kenya and Eastern province of Zambia, while haplotype 12SH5 was present in DRC, Rwanda, Zimbabwe, Comoros, South Africa, Eastern Zambia and Kenya. Six new haplotypes were not found outside the Great Lakes region (12SH3, 12SH4 and 12SH6–H9). The NJ phylogenetic resolved these *12S* rRNA haplotypes into two haplogroups (haplogroup A and B) supported by 97% bootstrap value (Figure S3.2). Their pattern was identical to that observed from *cox*1 haplogroups. However, the Ugandan haplotype (12SH10: GenBank AF150028) was clustered in haplogroup A, showing that the haplogroup A is also present in Uganda. Unfortunately, we did not find its corresponding *cox*1 sequence.

Table 3.5 *Rhipicephalus appendiculatus cox*1 haplotypes and their distribution among agro-ecological zones of the Great Lakes region and other sub-Saharan African countries

Haplotype [d]	Haplotypes from GenBank: Country (original haplotype name and GenBank number)	Present study	Haplogroup
CH1[a]	Kenya (H2: KU725891, H4: KU725893, H6: KU725895)[1], Rwanda (H3:DQ901360)[2] Zambia-east (H4: KX276942, H5: DQ859266[3], H3:DQ901361[2], H2: DQ859265[3])	Burundi (AEZ1, AEZ3), DRC (AEZ1, AEZ2, AEZ3), Tanzania (TZ18, TZ10, TZ08, TZ20), Rwanda (AEZ2)	A
CH3	–	Burundi AEZ1	A
CH4	Rwanda (H6: DQ901362)[2]	Burundi AEZ1, Rwanda (AEZ2)	A
CH6[b]	Kenya (H5: KU725894)[1]	Burundi (AEZ1, AEZ3), DRC (AEZ1, AEZ2)	A
CH7[c]	Kenya (H1: KU725890)[1], South Africa (H1: KX276939[1], H1: KX276940[1], H1: DQ901356[2]), Zambia-east (H1: DQ859261)[3], Zambia-south (H1:KX276943[1], H1: DQ859262)[3], Zimbabwe (AF132833[4], KC503257[5], H1: KX276944[1]), Grande Comore (H1) [6]	Burundi AEZ1, DRC (AEZ1, AEZ2, AEZ3), Rwanda (AEZ2)	B
CH9	–	Burundi AEZ3	A
CH10	–	Burundi AEZ1	A
CH11[d]	Kenya (H3: KU725892)[1], Rwanda (H3: DQ901363)[2], Grande Comore (H3)[6]	DRC (AEZ1, AEZ2, AEZ3), Rwanda (AEZ2)	A
CH12	Kenya (H11: KU725900)[1]	DRC (AEZ2, AEZ3)	A
CH14	–	DRC AEZ3	A
CH16	–	DRC AEZ1	A
CH18	–	DRC AEZ2	A
CH20	–	DRC-AEZ2	B
CH23	Uganda (H8: KX276941, KU725897)[1]	–	B
CH24	Kenya (H14: KU725903)[1]	–	B
CH25	Kenya (H27: KU725916) [1]	–	B
CH26	Grande Comore (H2: DQ901357)[2,6], Kenya (H7: KU725896, H13: KU725902) [1]	–	B
CH27	Kenya (H16: KU725905)[1]	–	B
CH28	Kenya (H21: KU725910[1], H9: DQ901359[2], H9: DQ901358[2])	–	B
CH29	Kenya (H28: KU725917) [1]	–	B
CH30	Kenya (H9: KU725898) [1]	–	A
CH31	Kenya (H17: KU725906[1]	–	A
CH32	Kenya (H24: KU725913) [1]	–	A
CH33	–	Tanzania (TZ13)	A
CH34	Kenya (H22: KU725911) [1]	–	A
CH35	Kenya (H23: KU725912) [1]	–	A
CH36	Kenya (H10: KU725899) [1]	–	A
CH37	Kenya (H15:KU725904) [1]	–	A
CH38	Kenya (H20: KU725909) [1]	–	A
CH39	Kenya (H19: KU725908) [1]	–	A
CH40	Kenya (H26: KU725915) [1]	–	A
CH41	Kenya (H25: KU725914[1]	–	A
CH42	Kenya (H18: KU725907) [1]	–	A
CH43	Kenya (H12: KU725901) [1]	–	A
CH44	Zambia-east (H3: DQ859263) [3]	–	A
CH45	Zambia-east (H4: DQ859264)[3]	–	A

[1](Kanduma et al., 2016a); [2](Mtambo et al., 2007c); [3](Mtambo et al., 2007b); [4](Murrell et al., 2000); [5](Burger et al., 2014); [6](Yssouf et al., 2011); [a] CH1: variants CH1, CH2, CH5, CH15, CH17 and CH21 (Table S3.3)
[b] CH6: variants CH6, CH8 and CH22 (Table S3.3); [c] CH7: variants CH7 and CH13; d CH11: CH11 and CH19 (Table S3); [d] Haplotypes underlined are exclusive to the Great Lakes region. Similar data for *12S* rRNA are detailed in Table S3.6

3.5 Discussion

This study analysed the intraspecific variation of mitochondrial DNA to better understand how factors such as agro-ecological zones and anthropogenic movements of cattle may affect population genetic structure and population expansion history of the tick *R. appendiculatus*, the main vector of *T. parva* in sub-Saharan Africa. We expected evidence of *R. appendiculatus* population expansion, gene flow and different colonization patterns of tick lineages, facilitated by the reported cattle mobility in the Great Lakes region.

3.5.1 Two *R. appendiculatus* lineages that are more variable in lowlands than highlands occur in the Great Lakes region

The 22 haplotypes identified by DNA polymorphic analysis of *cox*1 gene locus were clustered into two well-defined major groups, named haplogroup A (the most frequent) and haplogroup B. Similar grouping were obtained with *12S* rRNA analyses. The two haplogroups identified in the present study have been previously described as "east African" and "southern African" genetic groups (Mtambo et al., 2007c), corresponding to our haplogroups A and B, respectively. This genetic grouping fitted well with the phenotypical, physiological and ecological variation studies which have previously distinguished two major subpopulations of *R. appendiculatus* in Africa (Berkvens et al., 1998; Chaka et al., 1999; Leta et al., 2013; Madder et al., 1999; Madder et al., 2002; Speybroeck et al., 2004; Speybroeck et al., 2002). These phenotypic and physiological variations are largely associated with agro-ecological and geographical subdivisions. Tropical areas with prolonged and marked dry seasons are more suitable for larger sized ticks expressing high intensity of diapause and displaying univoltine phenology, corresponding to "southern African group" or haplogroup B. Equatorial areas with bimodal or continuous rainfall rather harbour smaller ticks without diapause with bivoltine or continuous phenology, corresponding to "east African group" or haplogroup A (Madder et al., 2002; Mtambo et al., 2007b).

The highest genetic variability was observed in lowlands, whereas a relatively lower diversity was observed in midlands and highlands. The high genetic diversity in lowlands can be explained by the strong presence of the two lineages A and B observed in these AEZs. The coexistence of these lineages could originate from the dispersal of the tick through livestock movement between AEZs (Excoffier et al., 2009), associated with the suitability of semi-arid climate for lineage B expressing obligatory diapause (Madder et al., 2002).

3.5.2 Moderate genetic structure of *R. appendiculatus* between lowlands and highlands

Population genetic analyses of *cox*1 gene variation in *R. appendiculatus* revealed low to moderate genetic differentiation values and high gene flow rates among AEZs. The two identified *R. appendiculatus* lineages were sympatric in the Great Lakes region, although lineage A was the most abundant and widely distributed in all AEZs and lineage B was particular confined to lowlands, where the climate is tropical dry, more arid with lower annual rainfall and longer dry season than in highlands (short dry season with abundant annual rainfall). These climate conditions in lowlands are quite similar to the described ecological zones of lineage B in southern Africa (Mtambo et al., 2007b; Mtambo et al., 2007c). In addition, altitudinal gradient seems to be a key factor that shapes the distribution pattern and the establishment ability of lineage B. Its frequency decreases with increasing altitude. High degree of genetic similarity was observed between the lowlands of DRC and the low plateau of Rwanda and between the highlands of DRC and Burundi, which are geographically distant from each other. The most likely explanation for this is that the spatial pattern of *R. appendiculatus* lineages is not only driven by geographical separation as described in previous studies (Mtambo et al., 2007b), but also related to their ecological preferences, as observed by the significant genetic differentiation among lowlands and highlands AEZs. On the other hand, adjacent AEZs shared more migrants, especially of lineage A, facilitated by short-distance seasonal movement of cattle (Verweijen and Brabant, 2017), which may have reduced the geographical structuring of the tick (Boulinier et al., 2016; Leo et al., 2017). Analysis of molecular variance (AMOVA) confirmed these findings showing that the variance explained by divergence between the six AEZs was lower (6%), while the largest fraction of genetic variation was observed among individuals within AEZs (94%).

3.5.3 *Rhipicephalus appendiculatus* lineage A has undergone sudden demographic and range expansion in the Great Lakes region

The demographic and spatial dynamics were analysed using multiple algorithms, to elucidate colonization events of the tick *R. appendiculatus* that took place in the Great Lakes region. A strong evidence of recent spatial and demographic expansion was observed for lineage A in all AEZs included in the study. Analyses of *cox*1 sequences revealed relatively high haplotype diversity contrasted with low nucleotide diversity values for each population, suggesting a sudden population expansion (Braverman et al., 1995; Simonsen et al., 1995). This result, together with negative values of Tajima's *D* and Fu's *Fs*, the star-like radiation, the unimodal

mismatch pattern and non-significant *RI* and *SSD* statistics, further support the recent evolutionary history and sudden population growth experienced by lineage A (Emerson et al., 2001; Frankham, 1996; Ray et al., 2003). We did not estimate the expansion time, but we hypothesize that the expansion was recent and not sufficient to increase the nucleotide diversity, probably because of recent coalescence, while rapid population expansion following a selective sweep (bottleneck or genetic drift) could have accumulated new mutations that sufficiently increased the haplotype diversity (Avise, 2000; Robbertse et al., 2016; Rogers and Harpending, 1992). This could explain the excess of singletons polymorphic sites and rare haplotypes that diverge from ancestral haplotypes by only 1–2 mutational steps (Rogers and Harpending, 1992; Slatkin and Hudson, 1991). However, the strong bimodal mismatch pattern observed in Rwanda AEZ2 and DRC AEZ1 suggests the coexistence of *R. appendiculatus* lineages after recent colonization events or exchanging migrants (Ray et al., 2003; Rogers et al., 1996; Rogers and Harpending, 1992). The scenario of lineage B contrasts with that of lineage A. Analyses indicate, for lineage B, an equilibrium situation that is characterised by a weak signal of recent bottleneck and no evidence of population expansion. It was also less diverse than haplogroup A, indicating that haplogroup A is experiencing population expansion independently of haplogroup B and it has been established longer in the Great Lakes region, while haplogroup B was probably introduced more recently and established a founder population. There are three possible explanations of the equilibrium observed in haplogroup B: (i) only few haplotypes were recently introduced; (ii) only the few identified haplotypes persisted after a bottleneck; or (iii) haplogroup B is experiencing an initial selection sweep which has reduced the number of rare haplotypes and singleton mutations (Ray et al., 2003). When we analysed *R. appendiculatus cox*1 sequences available in Africa, the star radiation of the MJ network in each group suggests that the two lineages went through a demographic expansion and evolve independently of each other with limited gene exchange. Unfortunately, we were not able to test the hypothesis of crossbreeding events between the two lineages described here, because of the maternal inheritance of mitochondrial genes used in the present study. More studies, such as extensive biological characterisation, crossbreeding experiments and the use of biparental inheritance markers are necessary to investigate the panmixia of the two lineages and the genetic mechanisms driving their establishment and corresponding phenotypic variations in changed environments.

The fact that *R. appendiculatus* has undergone spatial expansion was in accordance with the expectation that long and short distance movement of cattle are key factors of spreading ticks.

The processes responsible for that evolutionary pattern may have resulted, not only from range expanding of the previously established haplotypes to proximate AEZs, but also from the recolonization events of ticks from other regions and countries (Cristescu, 2015; Excoffier et al., 2009). In addition to cattle movement, the population expansion and establishment of novel haplotypes in previously unoccupied areas could have been driven by recent environmental and climatic changes, affecting vector-borne diseases landscape over recent decades (Biek and Real, 2010; Manel and Holderegger, 2013; Ostfeld, 2009).

3.5.4 Sympatric and allopatric ecological zones of *R. appendiculatus* lineages in Africa

Phylogenetic trees produced two main genetic subpopulations of *R. appendiculatus* that have a wide distribution range in Africa, with large divergence in behavioural diapause (Madder et al., 1999; Madder et al., 2002), spatial variation in body size (Chaka et al., 1999; Speybroeck et al., 2004), ecology and phenology (Berkvens et al., 1998; Leta et al., 2013; Speybroeck et al., 2002). Initially, the two lineages were considered as "east African" and "southern African" genetic groups (Mtambo et al., 2007c). To date, they are sympatric in most eastern and central African ecological zones. For instance, previous studies did not reveal the presence of haplogroup B in Rwanda (Mtambo et al., 2007c). This could be an indication of recent introduction of the tick in the Great Lakes region. The MJ network further elucidate that the initial population of haplogroup B could have come from an ancestral female of haplotype CH7, which is the most prevalent haplotype of this group in areas where it occurs. Consequently, two different eco-genetic zones are shaped in Africa, the sympatric zone where the two lineages are found, which covers central and eastern Africa, and allopatric zone in southern Africa where the two lineages have clear geographical and ecological separation. For instance, in Zambia, lineage B is present in the south (long dry season) and lineage A in the east of the country (shorter dry season) (Mtambo et al., 2007c). In Grande Comore, lineage B has established a stable population, while lineage A was found on imported cattle (Yssouf et al., 2011). In areas where the two lineages are sympatric, their respective abundances differ, mainly driven by their divergence in ecological preferences and plasticity (Speybroeck et al., 2002). The sympatric relationships agree with the observation made by Berkvens et al. (Berkvens et al., 1995), where an east African stock from Kenya expressed diapause, contrasting with the result obtained by Madder et al. (Madder et al., 1999), where another stock from the same region was unable to enter diapause. These evidences show that lineage B has a greater invasive ability into new habitats and better fit wide range of tropical and equatorial conditions, while lineage A is

particularly confined to equatorial conditions. This could be explained by the larger body size and obligatory diapause behaviour for southern African ticks, which allow them to survive hot and dry ecological conditions (Madder et al., 2002; Speybroeck et al., 2002). We hypothesise that these characteristics make the development of lineage B slower than lineage A in sympatric areas, giving an evolutionary advantage to the latter. This could also reduce the abundance of lineage B delaying its oviposition during unfavourable conditions. The processes that took place to divide the two groups need to be further investigated. However, we propose that they could have diverged following genetic drift due to founder events of natural geographical barrier that may result in reproductive isolation. It is demonstrated that biogeographical break again host migration reduces gene exchange and could dictate the spatial and reproductive separation within a species (Avise, 2000).

3.6 Conclusions

This study provided new insights into the genetic structure and colonization events of *R. appendiculatus* in the Great Lakes region and over its distribution range in sub-Saharan Africa. Our results highlighted the occurrence of two major genetic lineages (A and B) in the Great Lakes region. The two lineages are not spatially structured in the study region and they differ in their colonization histories and pattern. Lineage B, not previously reported, was probably introduced recently in the region and its occurrence decreases with increased altitude, whereas lineage A, widely distributed, has been longer established and subjected to sudden demographic and spatial expansion most likely related to short and long-distance cattle movement. *Rhipicephalus appendiculatus* ticks are more diverse in lowlands than highlands with moderate genetic differentiation between the two ecosystems, while more genetic similitude is found in zones with same agro-ecological profiles, in spite of their geographical distance. The genetic distribution of *R. appendiculatus* suggests two different eco-genetic zones in Africa, the sympatric zone (central and eastern Africa) where the two lineages coexist and the allopatric zones (southern Africa) where they have clear geographical divergence. The range expansion pattern of lineages and the genetic admixture of *R. appendiculatus* populations observed in the Great Lakes region can strongly affect the epidemiological dynamics of ECF. This could partially explain the endemic instability and occasional epidemics due to the introduction and temporal subsistence of infected ticks mostly in fringes areas of lowlands.

Chapter 4. Genetic and antigenic variation of *Theileria parva* in the Great Lakes region of Central Africa*

This research refers to the second objective of the thesis and was performed using blood samples collected from cattle in the Democratic Republic of Congo and Burundi. In the previous study of the tick vector Rhipicephalus appendiculatus *(**Chapter 3**), we found that two sympatric lineages of the tick occur in the Great Lakes region and that the genetic diversity was higher in lowlands than highlands. The two lineages diverged in their ecological distribution and abundance in different agro-ecological zones. We hypothesised that the observed genetic composition and distribution pattern of the tick, together with cattle movements and agro-ecological variability, may have significant implications in the transmission dynamics and therefore, the genetic structure of* T. parva*. The present study examined the evolutionary dynamics and the biogeographical distribution of* T. parva *genotypes based on two antigens genes (Tp1 and Tp2), to further understand the epidemiology of ECF in the region and predict vaccine safety. The results revealed high variation of* T. parva *within populations, especially in lowlands and midlands and limited geographical sub-structuring due to the widespread major genotypes. These findings suggest that the biogeographical distribution of* T. parva *genotypes is driven by host dispersal and ecological conditions that affect tick vector distribution and competence. Interestingly, most* T. parva *alleles found in the Great Lakes region were closely similar to the components of the trivalent Muguga vaccine, which justify testing the existing ITM vaccine in the region.*

* *This chapter was published as:*

Amzati, G.S., Djikeng, A., Odongo, D.O., Nimpaye, H., Sibeko, K.P., Muhigwa, J.-B.B., Madder, M., Kirschvink, N., Marcotty, T., 2019. Genetic and antigenic variation of the bovine tick-borne pathogen *Theileria parva* in the Great Lakes region of Central Africa. Parasites & vectors 12, 588.

4.1 Abstract

Background: *Theileria parva* causes East Coast fever (ECF), one of the most economically important tick-borne diseases of cattle in sub-Saharan Africa. A live immunisation approach using the infection and treatment method (ITM) provides a strong long-term strain-restricted immunity. However, it typically induces a tick-transmissible carrier state in cattle and may lead to spread of antigenically distinct parasites. Thus, understanding the genetic composition of *T. parva* is needed prior to the use of the ITM vaccine in new areas. This study examined the sequence diversity and the evolutionary and biogeographical dynamics of *T. parva* within the African Great Lakes region to better understand the epidemiology of ECF and to assure vaccine safety. Genetic analyses were performed using sequences of two antigen-coding genes, *Tp1* and *Tp2*, generated among 119 *T. parva* samples collected from cattle in four agro-ecological zones of DRC and Burundi. **Results:** The results provided evidence of nucleotide and amino acid polymorphisms in both antigens, resulting in 11 and 10 distinct nucleotide alleles, that predicted 6 and 9 protein variants in *Tp1* and *Tp2*, respectively. *Theileria parva* samples showed high variation within populations and a moderate biogeographical sub-structuring due to the widespread major genotypes. The diversity was greater in samples from lowlands and midlands areas compared to those from highlands and other African countries. The evolutionary dynamics modelling revealed a signal of selective evolution which was not preferentially detected within the epitope-coding regions, suggesting that the observed polymorphism could be more related to gene flow rather than recent host immune-based selection. Most alleles isolated in the Great Lakes region were closely related to the components of the trivalent Muguga vaccine. **Conclusions:** Our findings suggest that the extensive sequence diversity of *T. parva* and its biogeographical distribution mainly depend on host migration and agro-ecological conditions driving tick population dynamics. Such patterns are likely to contribute to the epidemic and unstable endemic situations of ECF in the region. However, the fact that ubiquitous alleles are genetically similar to the components of the Muguga vaccine together with the limited geographical clustering may justify testing the existing trivalent vaccine for cross-immunity in the region.

4.2 Background

Theileria parva is an intracellular apicomplexan parasite of cattle transmitted by the ixodid tick *Rhipicephalus appendiculatus*. The parasite infects and transforms bovine lymphocytes and causes an acute lymphoproliferative disease known as East Coast fever (ECF), which remains a major constraint to the improvement of cattle production in sub-Saharan Africa (Nene et al., 2016). The disease is present in 12 countries including the Great Lakes region of Africa, where ticks are invading new areas through the extensive cross-border and seasonal movement of cattle for trade and pasture (Amzati et al., 2018; Bazarusanga et al., 2007b; Bouslikhane, 2015; Kalume et al., 2012; Verweijen and Brabant, 2017). The geographical distribution of *T. parva* is mainly determined by that of its vector, for which predictive models have shown to have a wide range of suitable environments in Africa (Leta et al., 2013; Olwoch et al., 2009; Perry et al., 1991). Thus, host dispersal and ecological traits affecting tick population dynamics and performance may drive the mechanisms for spreading various genotypes of the parasite in different agro-ecological zones (AEZs), which in turn could result in epidemics or disruption in the endemicity of the disease in colonised areas (De Deken et al., 2007; Estrada-Pena et al., 2009; Geysen, 2008; Yssouf et al., 2011).

Current management and control of ECF are achieved by limiting tick infestation through the use of acaricides, as well as by treatment of infected cattle with theilericidal drugs such as buparvaquone. However, the continuous use of acaricides is unsustainable, and treatment is only effective during the early stages of the disease (Nene et al., 2016). In view of these limitations, vaccination remains the most effective control measure. The current approach to immunisation is the Infection and Treatment Method (ITM) of vaccination, which involves inoculation of titrated live sporozoites from three parasite stocks known as “Muguga cocktail” and simultaneous treatment with a long-acting formulation of oxytetracycline (Di Giulio et al., 2009). The Muguga trivalent vaccine provides robust and long-lasting protection against challenge with homologous *T. parva* strains but limited protection against heterologous strains (Bishop et al., 2015; Sitt et al., 2015). It has been demonstrated that ITM vaccination induces strain-specific immunity, mediated by the major histocompatibility complex (MHC) class I-restricted $CD8^+$ T cells killing *T. parva*-infected bovine host cells (Morrison et al., 2015). This suggests that the evolutionary dynamics of genetic diversity, which usually result in antigenic variation of parasites, enable *T. parva* to escape recognition by the host immune system (Connelley et al., 2011; Sivakumar et al., 2014). The genetic diversity of *T. parva* is thought to

be driven by several mechanisms and factors, including gene isolation, genetic drift, mutation, host immunity and genetic exchange through recombination (Katzer et al., 2006; Sivakumar et al., 2014). Furthermore, the deployment of the Muguga cocktail in new areas can introduce new strains and establish locally persistent infections, a source of the permanent spread of the disease through transmission by local ticks, even in the absence of detectable parasitaemia (Olds et al., 2018; Oura et al., 2007; Skilton et al., 2002). The "foreign" vaccine strains may also recombine with local ones and produce new, potentially more virulent genotypes (Katzer et al., 2006). Thus, to reduce the risks of introducing foreign parasite strains, a comprehensive study of parasite genotypes circulating in the region is required prior to the deployment of a live vaccine.

Genetic studies using a panel of DNA mini- and microsatellite markers to characterise the extent of genotypic diversity of *T. parva* have been done across different African countries, including Tanzania (Elisa et al., 2015; Rukambile et al., 2016), Uganda (Muwanika et al., 2016; Oura et al., 2005; Oura et al., 2011b), Kenya (Katzer et al., 2010; Odongo et al., 2006), Zambia (Muleya et al., 2012) and South Sudan (Salih et al., 2018). These studies provided evidence of genetic exchange between some populations and a minimal genetic sub-structuring on a geographical basis. More recently, a number of antigen-coding genes and epitopes that are targets of bovine MHC-I restricted $CD8^+$ T cells were identified in order to develop subunit vaccines against *T. parva* (Graham et al., 2007; Graham et al., 2008). Two of these reported antigens (*Tp1* and *Tp2*), which are immunodominant targets of bovine cytotoxic $CD8^+$ T cells, were shown to be extensively polymorphic in parasite isolates from East Africa, especially in buffalo-derived *T. parva* and those from cattle co-grazing with buffalo (Hemmink et al., 2018; Pelle et al., 2011; Salih et al., 2017; Sitt et al., 2018). The substantial variation previously reported in *Tp1* and *Tp2* antigens proved their great value for studying the antigenic composition and population structure of *T. parva.* In addition, antigenic variability in *T. parva* populations and immunodominance nature of the CD8+ T cell responses are believed to be major determinants of the parasite strain-restricted immunity, although the role of these antigens in immune protection conferred by the ITM vaccination is not clearly demonstrated yet (Bishop et al., 2015; Connelley et al., 2011; MacHugh et al., 2009; Morrison et al., 2015; Pelle et al., 2011; Sitt et al., 2015). It can therefore be hypothesised that the movement of cattle carrying parasites or infected ticks and the agro-ecological variability could define the *T. parva* population structure through continuous introduction of new parasite variants that may affect the epidemiological landscape of ECF in the Great Lakes region of Africa. Thus, population genetic

diversity studies of parasites are useful to better understand the epidemiology of ECF. In our recent genetic study of the tick vector, we identified two sympatric *R. appendiculatus* lineages, that strongly coexist in lowlands grazing areas in the Great Lakes region, where the climate is more arid than in the highlands (Amzati et al., 2018). The colonisation pattern of these two lineages in sympatric zones could result in different transmission dynamics and geographical distributions of *T. parva* genotypes.

Given the reported transboundary cattle movements and the evidence of agro-ecological conditions affecting the population structure of the tick vector in the Great Lakes region, the present study analysed: (i) the level of population genetic diversity and structure of *T. parva* parasites; (ii) their similitude with Muguga cocktail vaccine components; (iii) their phylogenetic relationships and biogeographical patterns; and (iv) their evolutionary dynamics. *Theileria parva* samples originating from three AEZs in the Democratic Republic of Congo and one AEZ in Burundi were analysed using the polymorphic antigens *Tp1* and *Tp2*, in comparison with published sequences, so as to further characterise the genetic relationships between *T. parva* genotypes evolving in different African countries. The knowledge of the population structure and evolution of *T. parva* should provide more insight for better understanding of the epidemiology of ECF and prediction of potential vaccine outcomes and breakthroughs for future sustainable management of ECF in the Great Lakes region.

4.3 Methods

4.3.1 Study area

The study was carried out in three AEZs of the South-Kivu Province of the Democratic Republic of Congo (DRC) and one AEZ in the Imbo valley of Burundi (Rugombo and Gihanga districts). Details of the geographical and climatic attributes of the AEZs and sampling sites characteristics were described in our earlier study (Amzati et al., 2018) and mapped in **Figure 4.1**. Briefly, the South-Kivu Province is covered by three main AEZs based on their altitudes: lowlands (< 1200 m: DRC AEZ1), midlands (1200–1800 m: DRC AEZ2) and highlands (1800-2800 m: DRC AEZ3), while the Imbo valley of Burundi falls into the lowlands area (< 1200 m: Burundi AEZ1). The rainfall period is bimodal and its duration varies significantly between AEZs. The annual rainfall increases with altitude while the temperature decreases. Majority of cattle found in the study area belong to indigenous Ankole breeds (Sanga type), raised mainly under a communal open grazing system and subjected to short and long-distance transhumance

during the dry season. The livestock production system is characterised by irregular use of acaricides in all AEZs, except in more fenced commercial farms in the highlands where acaricides are usually applied on a weekly basis. No vaccination history against ECF had been reported in DRC, whereas a small-scale immunisation programme was introduced in Burundi between 1981 and 1987, using a cocktail vaccine of three local *T. parva* isolates (Gatumba, Gitega and Ngozi) for ITM vaccination (Tama, 1989).

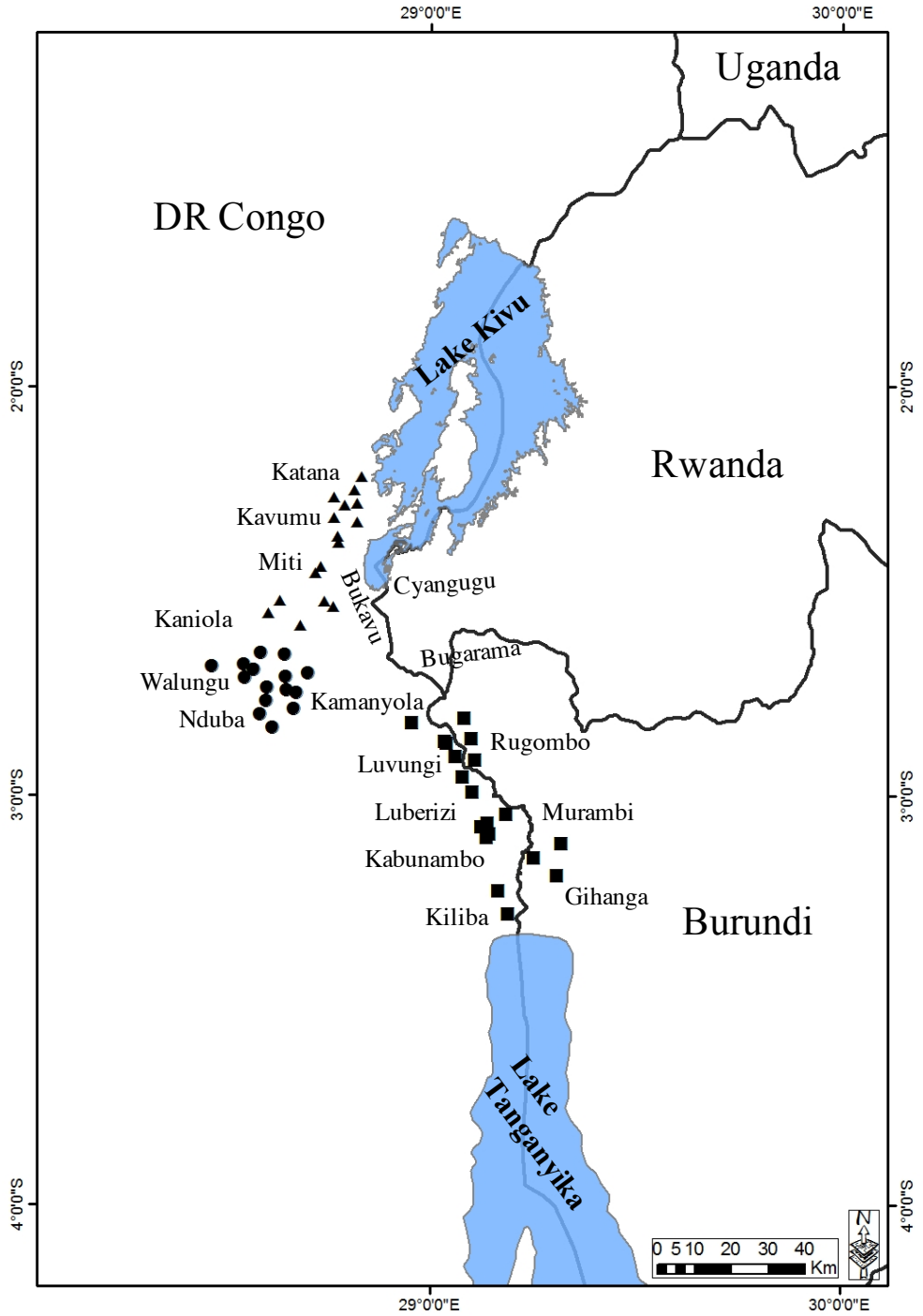

Figure 4.1 Map of the Great Lakes region showing sampling sites and their altitudes in the four agro-ecological zones of DRC and Burundi. Sampling sites located in the lowlands are indicated by squares (AEZ1: altitude < 1200 m), while circles indicate the sampling locations in midlands (AEZ2: altitude 1200–1800 m); and triangles denote sites located in the highlands (AEZ3: altitude 1800-2800 m)

4.3.2 Cattle blood sample collection

A cross-sectional study was conducted during the late rainy season between February and April 2015, as previously described in Amzati et al. (Amzati et al., 2018). Blood samples were collected from indigenous cattle raised under traditional farming systems. Three cattle herds were randomly selected from 8–12 villages in each AEZ. In each herd, 5–10 adult cattle (over 3 years of age) were randomly sampled. The random function of Microsoft Excel program was used as random generator. Cattle blood samples were collected from the jugular vein into EDTA vacutainer tubes, then transferred to the laboratory in a cool box. On the same day, approximately 120 µl of each blood sample was spotted on Whatman FTA cards (Whatman Bioscience, Cambridge, UK), air-dried for 4 h at ambient temperature and stored individually in plastic bags containing dry silica gel packs at room temperature until used. A total of 480 blood samples were collected from the four AEZs (Table S4.1).

4.3.3 DNA isolation and PCR screening for *Theileria parva*

Genomic DNA (gDNA) was extracted from the dried blood spots using a commercial DNA extraction Kit (PureLink® Genomic DNA Mini Kit, Invitrogen, Schwerte, Germany), according to the manufacturer's instructions. DNA concentration was assessed using a Nanodrop spectrophotometer (Wilmington, Delaware, USA). Purified gDNA samples were screened for the presence of *T. parva* DNA using a nested PCR (nPCR) assay targeting the *T. parva*-specific conserved single-copy gene encoding the sporozoite microneme-rhoptry surface antigen, p104-kDa (GenBank: M29954) (Iams et al., 1990; Odongo et al., 2010). Oligonucleotide outer primers IL3231 and IL755 and inner primers were IL4243 and IL3232 were used as previously described (**Table 4.1**) (Odongo et al., 2010; Skilton et al., 2002) to amplify a 277-bp fragment of the p104 antigen gene. Amplification was performed using lyophilized AccuPower® PCR Pre-mixes (Bioneer, Seoul, South Korea) for both primary and secondary PCR. The primary PCR reaction, in addition to AccuPower® PCR Pre-mixes, contained 0.25 µM of each forward and reverse primers, 20 ng of gDNA, and nuclease-free

distilled water added to bring the reaction to a final volume of 20 µl. The reaction mixture for the secondary amplification was as described for the primary PCR, except that the template was 1 µl of 3× diluted primary PCR products. A clean FTA punch was used as negative control for the DNA extraction and the positive control was a *T. parva* Muguga (F100 TpM) DNA obtained from BecA-ILRI Hub. The cycling conditions for the p104 gene have previously been described (Odongo et al., 2010; Skilton et al., 2002), except for some minor modifications (**Table 4.1**). Six microliters of the secondary PCR products were analysed by electrophoresis in 1.8% agarose gels stained with GelRed (Biotium Inc., Hayward, USA) and visualised under UV light.

Table 4.1 PCR oligonucleotide primers used for amplification of p104, *Tp1* and *Tp2* genes with their corresponding annealing temperatures and amplicon sizes

Gene locus	Primer name	Primer sequence (5'-3')	Annealing temperature (°C)	Amplicon size (bp)	Reference
p104	IL3231 (Fw1)	ATTTAAGGAACCTGACGTGACTGC	60	496	(Skilton et al., 2002)
	IL755 (Rev1)	TAAGATGCCGACTATTAATGACACC			
	IL4243 (Fw2)	GGCCAAGGTCTCCTTCAGAATACG	55	277	(Odongo et al., 2010)
	IL3232 (Rev2)	TGGGTGTGTTTCCTCGTCATCTGC			
Tp1	*Tp1*-Fw1	ATGGCCACTTCAATTGCATTTGCC	50	432	(Pelle et al., 2011)
	Tp1-Rev1	TTAAATGAAATATTTATGAGCTTC			
	Tp1-Fw2	TGCATTTGCCGCTGATCCTGGATTCTG	55	405	(Elisa et al., 2015; Salih et al., 2017)
	Tp1-Rev2	TGAGCTTCGTATACACCCTCGTATTCG			
Tp2	*Tp2*-Fw1	ATGAAATTGGCCGCCAGATTA	50	525	(Pelle et al., 2011)
	*Tp*2-Rev1	CTATGAAGTGCCGGAGGCTTC			
	Tp2-Fw2	ATTAGCCTTTACTTTATTATTTWCATTYTAC	54	504	(Elisa et al., 2015; Salih et al., 2017)
	Tp2-Rev2	CTATGAAGTGCCGGAGGCTTCTCCT			

Abbreviations: Fw1, forward outer; Fw2, forward inner; Rev1, reverse outer; Rev2, reverse inner

PCR amplification and sequencing of *Tp1* and *Tp2* gene loci

Samples which tested positive for *T. parva* by an indication of the amplification of the p104 marker were used for the analysis of genetic diversity of *T. parva* samples. Two *T. parva* genes, *Tp1* and *Tp2*, were amplified using a previously described nested PCR (Pelle et al., 2011; Salih et al., 2017). The sizes of amplified regions containing known $CD8^+$ T cell epitopes were 405 bp and 504 bp for *Tp1* and *Tp2* nested amplicons, respectively (Graham et al., 2007; Graham et al., 2008; Pelle et al., 2011). Specific outer and inner primers used to amplify the two genes are presented in **Table 4.1**. The amplification was performed in a 20 µl PCR reaction, with the same components described for p104. Thermal cycling conditions for primary and secondary PCR for *Tp1* and *Tp2* genes were as described in **Table 4.1**. Six microliters of *Tp1* and *Tp2* nPCR products were analysed by electrophoresis in a 1.8% agarose gel. Amplicons obtained from the nPCR were purified using the QIAquick® PCR Purification Kit (Qiagen GmbH, Hilden, Germany) following the manufacturer's instructions and sequenced directly using inner forward and reverse primers on an Applied Biosystems ABI 3730 sequencer (Macrogen Inc. Europe, Amsterdam, The Netherlands).

4.3.4 Prediction of amino acid sequences and epitope identification

Nucleotide sequences of *Tp1* and *Tp2* genes were manually edited, assembled and their consensus translated into amino acid sequences using CLC Main Workbench software v7.9.1 and the online translation tool EMBOSS-Transeq. A BLAST search was then performed to confirm species and gene identity of these sequences against available sequences in the GenBank database. The nucleotide and deduced amino acid sequences were aligned separately for each gene with corresponding Muguga reference sequences, one of the three components of the Muguga cocktail live vaccine (GenBank: JF451936 and JF451856 for *Tp1* and *Tp2*, respectively). Multiple nucleotide sequence alignments were constructed with a codon-based approach under the Muscle algorithm as implemented in the Translator X online platform (http://www.translatorx.co.uk) (Abascal et al., 2010). All alignments were visualised and inspected in the CLC Main Workbench software. Single nucleotide polymorphisms (SNPs), indels and sequence variants were then identified by comparing the generated consensus sequence for each sample to the corresponding Muguga sequence. The $CD8^+$ T cell epitope regions and variants were identified using previously described positions presented in Table S4.2 (Graham et al., 2008; Pelle et al., 2011).

4.3.5 Population genetic analysis and phylogenetic reconstruction

Tp1 and *Tp2* nucleotide sequences were collapsed into alleles using DnaSP software v6.11.01 (Rozas et al., 2017). Genetic diversity indices, including the number of segregating sites (*S*), nucleotide diversity (π), number of distinct alleles (N_A), were estimated for each AEZ and for the entire dataset using the same software. In addition, the nucleotide diversity was estimated throughout the sequenced fragment of the *Tp2* gene based on a sliding-window of 100 nucleotides with a step size of 25 bp, to estimate the stepwise diversity across epitope coding regions. In order to assess the genetic structure within and between populations, pairwise estimates of genetic differentiation among populations based on Wright's fixation index (F_{ST}) and analysis of molecular variance (AMOVA) were implemented in Arlequin v3.5.2.2 (Excoffier and Lischer, 2010). The genetic differentiation (F_{ST}) was interpreted as low (0–0.05), intermediate (0.05–0.15), great (0.15–0.25) and very great ($F_{ST} < 0.25$) (Freeland et al., 2011)[46]. The evolutionary divergence between gene alleles was estimated using proportion genetic distance (p-distance) in MEGA v7.0 (Kumar et al., 2016).

Phylogenetic reconstructions were performed to further investigate the population structure and relationships among *T. parva* alleles in sub-Saharan Africa. Both representative *Tp1* and *Tp2* allele sequences found in the present study and those previously published, obtained from cattle and African buffalo (*Syncerus caffer*) across Africa, were used in the phylogenetic analyses and population differentiation. Published sequences comprised samples from Kenya (cattle-derived, buffalo-derived and buffalo-associated parasites), South Sudan and laboratory isolates from Kenya, Uganda, Tanzania, Zambia and Zimbabwe. The *Tp1* and *Tp2* gene sequences of the three *T. parva* stocks used in the live trivalent Muguga vaccine (Muguga, Serengeti-transformed and Kiambu-5) were also included in the analyses. The obtained overall representative allele datasets were aligned for each gene based on their corresponding amino acid translations using the Translator X server with its Muscle algorithm. Sites that were ambiguously aligned were eliminated from the protein alignment before back-translate to nucleotides using the GBlocks program with default parameters. In addition, the *Tp1* and *Tp2* alignments for individual sequences generated in the present study (from the Great Lakes region) were visually checked and concatenated to generate a data matrix (*Tp1+Tp2*) in order to maximise the phylogenetic signal. Samples with missing data for one locus were excluded from the concatenated dataset. Phylogenetic reconstructions were then performed for the *Tp1* and *Tp2* representative gene alleles separately, as well as for the concatenated nucleotide matrix

in MEGA software using the Neighbor-Joining (NJ) algorithm by performing 1000 bootstrap replications. The best-fitting nucleotide substitution model for each dataset was estimated under the Bayesian information criterion (BIC) using MEGA. The evolutionary distances for each of the three datasets were computed using the Tamura 3-parameter (T92) model of nucleotide sequence evolution in which rate variation among sites was modelled according to gamma distribution. The phylogenetic trees were rooted with orthologous sequences from *Theileria annulata* as the outgroups (GenBank: TA17450 for *Tp1* and TA19865 for *Tp2*). Furthermore, a median-joining (MJ) network was constructed using cattle-derived *Tp2* nucleotide sequences generated during this study as well as published sequences, to investigate the ancestral relationships among *T. parva* alleles on the basis of their geographical origins. The network was computed through the default MJ algorithm described by Bandelt et al. (Bandelt et al., 1999) in the PopArt software (Leigh and Bryant, 2015). Invariant sites were removed from the dataset for network reconstruction.

4.3.6 Molecular evolutionary dynamics

Theileria parva demographic dynamics were analysed using selective neutrality statistics Fu and Li's *D** and *F** (Fu, 1997) and Tajima's *D* (Tajima, 1989) to evaluate the departure from neutral evolution or evidence of natural selection constraint for each studied population as implemented in the DnaSP and Arlequin software. To further assess the selective constraint within *Tp2* epitope coding regions, a sliding-window was estimated for the overall data set within a window of 100 bp using a step size of 25 bp. The significance of these statistics was tested with a coalescence-based approach using 1000 simulations. Statistically significant positive values of neutrality tests indicate an excess of intermediate-frequency alleles in the population than expected, that could be due to balancing selection, population structure or bottlenecks, while negative values denote an excess of rare polymorphisms in a population, which provides evidence of purifying, directional (positive) selection or population expansion. Population dynamics were further assessed with mismatch distribution of pairwise nucleotide differences between sequences in the Arlequin software.

4.4 Results

4.4.1 PCR amplification and gene polymorphisms

Of the 480 samples investigated, 119 produced a p104 amplicon suggesting that they contained *T. parva* DNA. These were subjected to *Tp1* and *Tp2* amplification and sequencing; sequences

were successfully generated from 116 and 96 samples, respectively (Table S4.3). We were unable to obtain amplicons or sequences from 3 samples for the *Tp1* gene and 23 samples for the *Tp2* gene. Novel *Tp1* and *Tp2* sequences were submitted and are available in the GenBank under accession numbers: MF449288-MF449294 for *Tp1*; and MF449295-MF449302 for *Tp2*. The 405-bp sequence region of *Tp1* encodes 134 amino acids (25% of the 543 amino acids of the full-length *Tp1* gene). This region is located between nucleotides 537 and 941 of the reference genome of the *T. parva* Muguga strain (GenBank: XP_762973), while the 504-bp region of the *Tp2* gene encodes 167 amino acids of the 174 amino acid-long protein encoded by the reference *T. parva* Muguga genome (GenBank: XP_765583). Sequence analyses showed moderate synonymous and nonsynonymous nucleotide substitutions randomly distributed along the *Tp1* sequence, including in the single CD8+ T cell target epitope, as well as an indel of 12 nucleotide insertion (TCT GCA CCT CCT) corresponding to the 4 amino acid residues SAPP. In contrast, the analyses revealed extensive polymorphisms in the *Tp2* gene, both at the nucleotide and amino acid levels, which were also identified within the six epitope regions. To further understand the phylogenetic relationships between *T. parva* parasites in sub-Saharan Africa, a comprehensive population genetic analysis was conducted, including *Tp1* and *Tp2* sequences retrieved from the GenBank.

4.4.2 Sequence diversity in the *Tp1* gene locus

Sequence analysis of *Tp1* gene fragment detected 11 distinct alleles in the 116 sequenced DNA samples (**Table 4.2**, Figure S4.1). These alleles were defined by 14 single-nucleotide polymorphisms (SNPs) and one in-frame indel insertion of 12 nucleotides compared with the reference *T. parva* Muguga genome sequence (identical to Serengeti and Kiambu-5 sequences for *Tp1*). The insertion occurred in four samples with the *Tp1* sequences identical to that of allele 45, which is genetically the most distant from the Muguga reference sequence (Table S4.4). On the other hand, *Tp1* allele 1 (present in the three *T. parva* stocks components of the trivalent Muguga vaccine) was the most predominant allele, identified in 76 (65.5%) of the 116 samples. The overall nucleotide polymorphism in the *Tp1* gene was $\pi = 0.5\%$. The lowest genetic diversity was obtained in DRC AEZ3, where all the 25 sequences were represented by the *T. parva* Muguga allele 1. The three other AEZs (DRC AEZ1, 2 and Burundi AEZ1) had very similar levels of nucleotide diversity (**Table 4.2**).

Moreover, the 11 *Tp1* alleles allowed to predict six distinct antigen variants, distinguished by amino acid changes at seven polymorphic residues and one insertion motif of four amino acids

(that contrasted with 92% of conserved amino acid residues) (**Figure 4.2**, **Table 4.2**). The most common antigen variant (var1), which is present in the *T. parva* strains Muguga, Serengeti and Kiambu-5, was found in 69% (80/116 samples) of samples obtained from all AEZs. Furthermore, antigen variants 3 and 31, with the smallest genetic distances to the variant 1, accounted for 10% and 16% of the total *T. parva* samples, respectively; while variants 32–34 were rarely present and were only observed in DRC AEZ1 and 2. In most cases, the predicted protein variants of gene alleles containing unique sequence were identical or nearly similar to the most common antigens due to synonymous substitutions (Figure S4.1, Table S4.5).

The multiple alignment of predicted *Tp1* amino acid sequences revealed the presence of three different $CD8^+$ T cell epitope variants, observed in the defined single *Tp1* $CD8^+$ T cell epitope region (VGYPKVKEEML) (**Figure 4.2**, Table S4.2). The detailed geographical distribution of epitope variants is presented in Table S4.6. Briefly, the epitope variant ending with -ML (which is also present in the three *T. parva* Muguga cocktail vaccine stocks) was found in the majority of samples from all four AEZs (72%: 84 out of 116 samples), followed by the epitope variant -II which was observed in 30 samples (26%) and particularly absent in *T. parva* samples from the highlands of DRC (DRC AEZ3). The third epitope variant (-MI) was only present in two samples from DRC AEZ1 and 2. We noticed the abundance of two of the three epitope variants in Burundi AEZ1, where approximately half of *T. parva* samples carried the epitope -II.

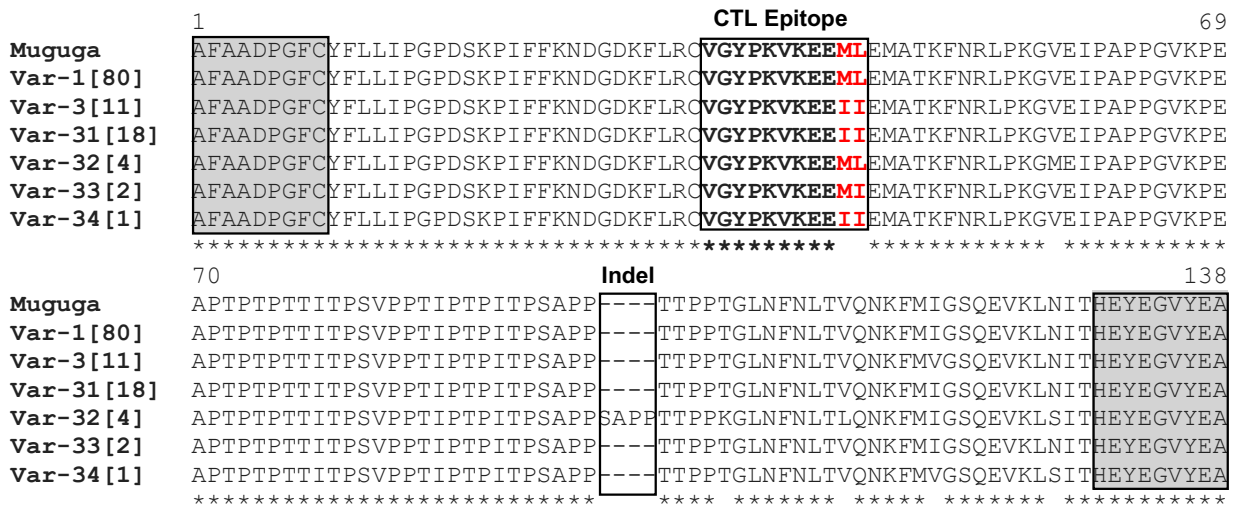

Figure 4.2 Multiple amino acid sequence alignment of six *Tp1* antigen variants in 116 *T. parva* samples obtained from DRC and Burundi. Antigen variants are named var-1 to var-34. The single letter amino acid code is used. The antigen variants nomenclature used in this study was first proposed by Pelle et al. [36]. Variants var-1 and var-3 were first described by Pelle et al. [36] and var-31 by Salih et al. [37]. The numbers in square brackets behind variants names indicate the number of *T. parva* isolates represented by each variant. The single previously identified *T. parva* $CD8^+$ T cell target epitope is bolded and boxed. The polymorphic residues in the T cell epitope are coloured in red. Conserved amino acid residues are denoted by (*) below the alignment, and dashes (-) denote insertion region. Nested PCR primers used for sequencing are shaded and boxed in flanked regions. The Muguga sequence (GenBank: JF451936) was used as the reference sequence; it represents the other component of the Muguga cocktail vaccine (Serengeti-transformed and Kiambu 5) that are identical to Muguga strain sequence for the *Tp1* locus. *Tp1* antigen variant var-1 is found in the three Muguga vaccine strains. Corresponding gene alleles and sample characteristics are presented in Table S4.3

Table 4.2 *Tp1* and *Tp2* sequence diversity in cattle-derived *T. parva* from RDC and Burundi

Gene locus	AEZ	Sample size	Nucleotide sequences			Amino acid sequences	
			Polymorphic sites[a]	No. of alleles[b]	Nucleotide diversity ± SD	Polymorphic sites[a]	No. of antigen variants[b]
Tp1	DRC AEZ1	31	12 + ind	7	0.008 ± 0.002	7 + ind	5
	DRC AEZ2	27	12 + ind	8	0.007 ± 0.002	7 + ind	6
	DRC AEZ3	25	0	1	0	0	1
	Burundi AEZ1	33	4	3	0.005 ± 0.0004	3	3
	Overall *Tp1*	116	14 + ind	11	0.005 ± 0.0007	7 + ind	6
Tp2	DRC AEZ1	25	166	7	0.13 ± 0.025	82	7
	DRC AEZ2	20	165	5	0.17 ± 0.016	81	5
	DRC AEZ3	23	2	2	0.002 ± 0.0004	2	2
	Burundi AEZ1	28	175	6	0.16 ± 0.017	85	5
	Overall *Tp2*	96	181	10	0.14 ± 0.01	88	9

[a]The insertion region (ind) was excluded for the determination of the number of polymorphic sites and the nucleotide diversity

[b]Alleles represent distinct nucleotide sequences diverged at least by one substitution (Figure S4.1, Figure S4.2), while antigen variants represent predicted distinct protein sequences (**Figure 4.2**, **Figure 4.4**)

Abbreviations: AEZ, agro-ecological zone; ind, indels; SD, standard deviation

4.4.3 Sequence diversity in the *Tp2* gene locus

A total of 10 unique *Tp2* alleles were identified among the 96 sequenced *T. parva* samples (**Table 4.2**, Figure S4.2). The 10 alleles were determined by SNPs detected at 181 of the 504 nucleotide positions (variation in 36% of the nucleotide residues). The majority of variable sites (179 of 181) were parsimony informative and there were no deletions or insertions in the *Tp2* sequences analysed. The overall nucleotide polymorphism (π) was 14%, with the highest level of DNA diversity observed among *T. parva* samples obtained from cattle raised in DRC AEZ2 and Burundi AEZ1. In contrast, the nucleotide diversity was lower in DRC AEZ3 where only two *Tp2* alleles (alleles 1 and 2) were described (**Table 4.2**). The sliding-window plot revealed that samples from different AEZs shared similar patterns of diversity through their sequences, with the highest polymorphism observed between nucleotide positions 200–300 in most populations, except in DRC AEZ3 were the diversity was found between positions 50 and 100 (**Figure 4.3a, b**). The number of alleles varied from five to seven among the *T. parva* samples in DRC (AEZ1 and 2) and Burundi AEZ1 (**Table 4.2**). The *Tp2* allele 1 (which is the Muguga and Serengeti allele) was the most ubiquitous, being observed in 39 of the 96 samples (41%) (Table S4.5, Figure S4.2). The next most common *Tp2* allele in the region was allele 2 (the *T. parva* Kiambu-5 allele, a component of the trivalent Muguga cocktail vaccine), and was present in 22 samples (23%). The majority of *T. parva* samples (61 of the 96 samples; 64%) were similar to the three component stocks of the Muguga cocktail, as observed for *Tp1* alleles. Interestingly, although the Muguga/Serengeti type was the most present in the region, it was less abundant than the Kiambu-5 type in Burundi AEZ1. In addition, alleles of the three components of the live vaccine were found in only 6 out of the 20 samples (30%) from DRC AEZ2. Moreover, the most common *Tp2* alleles in this DRC AEZ2 (alleles 56 and 57) were genetically the most distant from Muguga/Serengeti and Kiambu-5 alleles (p-distance > 25%) (Table S4.4).

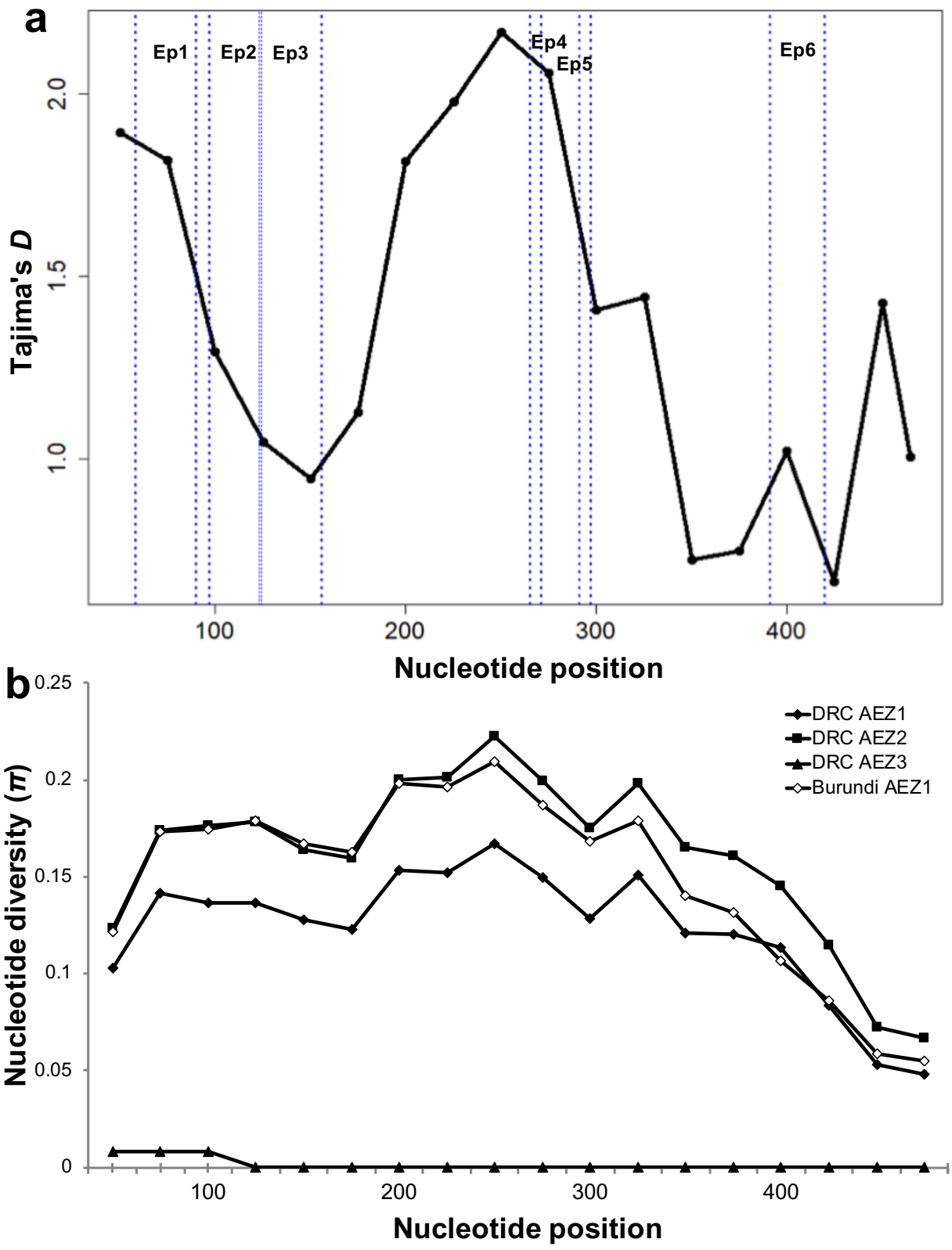

Figure 4.3 *Tp2*-based sliding-window plot of Tajima's *D* statistics (**a**) and nucleotide diversity (**b**) of *T. parva* sequences from the Great Lakes region. A window length of 100 nucleotides and a step size of 25 bp were used. The maximum nucleotide diversity and Tajima's *D* values

are observed between the nucleotide positions 200 and 300, containing the *Tp2* epitopes 4 and 5. *Abbreviation*: Ep1-6, epitope1-6.

The nucleotide variation observed in the 10 *Tp2* alleles resulted in nine distinct protein variants with variations found at 88 amino acid residue positions (53% of variation) (**Figure 4.4**). The results of antigenic variability of the *Tp2* gene are summarised in **Table 4.2**. In general, the number of antigen variants found in *T. parva* samples from different AEZs remained the same as for *Tp2* gene alleles, except in Burundi AEZ1 where the nucleotide substitution observed in allele 59 was synonymous and therefore this allele together with allele 57 were translated into protein variant 54 (Table S4.5). These results reveal that most of SNPs found in the *Tp2* gene were non-synonymous, increasing the antigenic variability for the overall data set.

Furthermore, multiple alignment of *Tp2* amino acid sequences revealed an extensive degree of polymorphism in the six defined $CD8^+$ T cell epitopes within the sequenced gene region (**Table 4.3**, Table S4.2). The numbers of epitope variants ranged from four for epitope 2 to seven for epitope 1 (**Table 4.3**, **Figure 4.4**). Within these epitopes, the number of conserved residues varied from three to five amino acid positions. Epitope 6 had the highest number of conserved amino acid residues, with residues 135–138 found in all protein sequences. Two variants (SHEELKKLGML and SDEELNKLGML) out of the seven variants of epitope 1 were identical to Muguga cocktail vaccine stock variants (Muguga/Serengeti and Kiambu-5) and comprised together 61 out of the 96 *Tp2* sequences studied (64%) (**Table 4.3**, Table S4.6). The majority of sequences described in the present study carried the Muguga/Serengeti epitope variants in most AEZs, except in Burundi AEZ1 where the Kiambu-5 variant was the most prevalent for epitope 1, and in DRC AEZ2 where the most prevalent epitope variants were not present in the Muguga cocktail (Table S4.6). Strikingly, despite the divergence distribution of epitope variants in different AEZs, a large number of overall major variants were common to all AEZs.

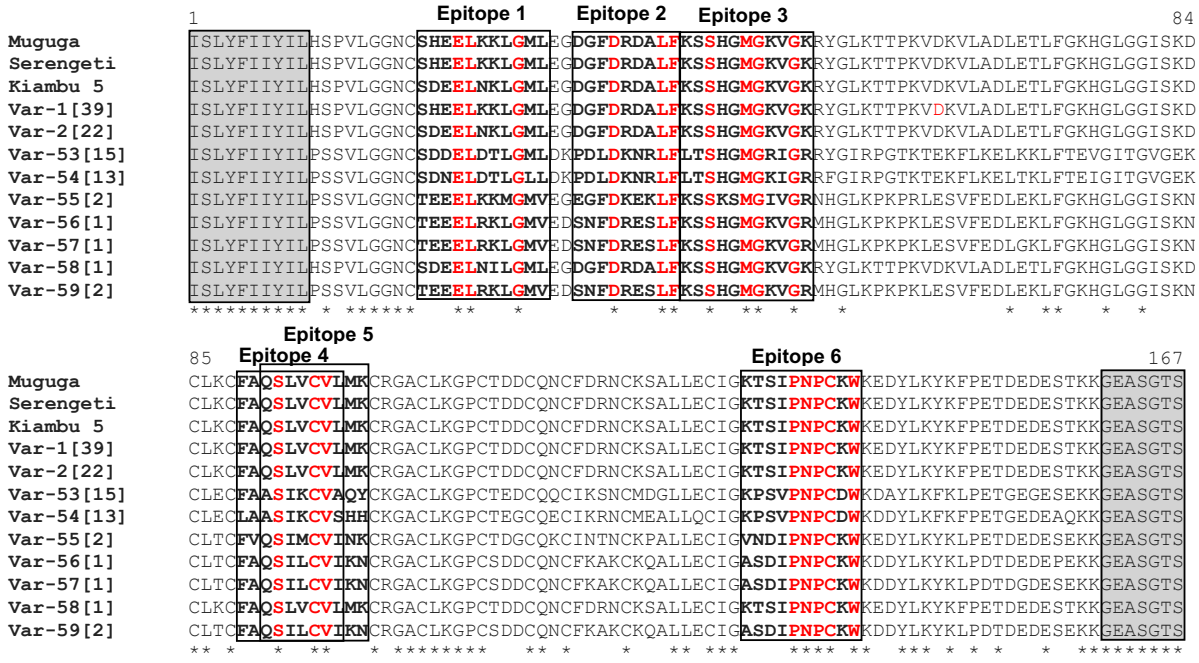

Figure 4.4 Multiple amino acid sequence alignment of nine *Tp2* antigen variants detected in 96 *T. parva* samples from DRC and Burundi. Amino acids are denoted by the single-letter codes. Var-1 to var-59 are variant names. The antigen variants nomenclature used in this study was first initiated by Pelle et al. [36]. Antigen variants var-1 and var-2 were described in Pelle et al. [36] and Salih et al. [37] and are, respectively, Muguga (identical to Serengeti-transformed) and Kiambu-5 strains. Reference sequences component of the Muguga cocktail live vaccine are represented by Muguga (GenBank: JF451856), Serengeti (Serengeti-transformed, GenBank: JF451862) and Kiambu-5 (GenBank: JF451880). The numbers in square brackets behind variants names indicate the number of *T. parva* samples represented by each variant. The six previously described epitopes (epitope1-6), that are the target of the bovine $CD8^+$ T cells immune responses are bolded and boxed. The conserved amino acid residues in the epitopes are coloured in red. The star (*) below the alignment indicates positions of conserved amino acid residues. The shaded and boxed flanked regions denote the inner primers used for sequencing. *Tp2* Antigen variants var-1 and var-2 are found in Muguga/Serengeti and Kiambu-5 strains, respectively. Corresponding gene alleles and sample characteristics are presented in Table S4.3.

Table 4.3 *Tp1* and *Tp2* CD8$^+$ T cell epitope variants identified in cattle-derived *T. parva* from DRC and Burundi

***Tp1* epitope variants ($Tp1_{35–45}$)**	***Tp2* epitope variants**					
	Epitope 1 ($Tp2_{20–30}$)	**Epitope 2 ($Tp2_{33–41}$)**	**Epitope 3 ($Tp2_{42–52}$)**	**Epitope 4 ($Tp2_{89–97}$)**	**Epitope 5 ($Tp2_{91–99}$)**	**Epitope 6 ($Tp2_{131–140}$)**
VGYPKVKEEML	**SHEELKKLGML**	**DGFDRDALF**	**KSSHGMGKVGK**	**FAQSLVCVL**	**QSLVCVLMK**	**KTSIPNPCKW**
(var-1, 32)	(var-1)	(var-1, 2, 58)	(var-1, 2, 58)	(var-1, 2, 58)	(var-1, 2, 58)	(var-1, 2, 58)
VGYPKVKEEII	*TEEELKKMGMV*	*EGFDKEKLF*	*KSSKSMGIVGR*	*FVQSIMCVI*	*QSIMCVINK* (var-	*VNDIPNPCKW*
(var-3, 31, 34)	(var-55)	(var-55)	(var-55)	(var-55)	55)	(var-55)
VGYPKVKEEMI	*TEEELRKLGMV*	*SNFDRESLF*	KSSHGMGKVGR	*FAQSILCVI*	*QSILCVIKN* (var-	*ASDIPNPCKW*
(var-33)	(var-56, 57, 59)	(var-56, 57, 59)	(var-56, 57, 59)	(var-56, 57, 59)	56, 57, 59)	(var-56, 57, 59)
	SDNELDTLGLL	PDLDKNRLF	LTSHGMGKIGR	LAASIKCVS	ASIKCVSHH	KPSVPNPCDW
	(var-54)	(var-53, 54)	(var-54)	(var-54)	(var-54)	(var-53, 54)
	SDDELDTLGML		LTSHGMGRIGR	FAASIKCVA	ASIKCVAQY	
	(var-53)		(var-53)	(var-53)	(var-53)	
	SDEELNKLGML					
	(var-2)					
	SDEELNILGML					
	(var-58)					

Notes: Epitope variants were identified using the reference amino acid positions presented in Table S4.2. Numbers in brackets following the epitope sequences correspond to antigen variants carrying the epitopes (**Figure 4.2**, **Figure 4.4**). Epitope variants described for the first time are in italic and those found in the Muguga cocktail vaccine are in bold. *Tp2* antigen variants var-1 and var-2 are found in Muguga (identical with Serengeti-transformed) and Kiambu-5 strains, respectively. *Tp1* antigen variant var-1 is found in Muguga (identical with Serengeti-transformed and Kiambu-5). *Abbreviations*: Var-1 to var-59, antigen variant names

4.4.4 Phylogenetic and phylogeographical patterns of *T. parva* populations in sub-Saharan Africa

In order to elucidate the phylogeographical structure and the evolutionary relationships among *T. parva* allelic sequences, the nucleotide sequences generated from the present study were analysed together with previously sequenced *T. parva* isolates from cattle and buffalo from different sub-Saharan African countries (Kenya, South Sudan, Tanzania, Uganda, Zambia and Zimbabwe) and the three component stocks of the Muguga cocktail live vaccine obtained from GenBank. In total, 274 *Tp1* and 241 *Tp2* sequences were analysed (Table S4.7, Table S4.8). The allelic analysis yielded 48 distinct alleles for *Tp1* and 61 different *Tp2* alleles. Of these *T. parva* alleles identified in Africa, seven *Tp1* alleles (A43-A49) and eight *Tp2* alleles (A56-A63) were new and exclusive to the Great Lakes region of Central Africa. In addition, 22 *Tp1* alleles (A13-A34) and 36 *Tp2* (A06-A43) were exclusively found in buffalo-derived or buffalo-associated *T. parva* isolates. The phylogenetic tree constructed from *Tp1* gene alleles failed to provide strong phylogenetic signal (Figure S4.3), whereas the one based on *Tp2* alleles showed that *T. parva* parasites are more clustered depending on their mammalian host species than their geographical sub-structuring.

The NJ phylogenetic tree performed on the 61 *Tp2* alleles (representative of 241 individual sequences) distinguished two main phylogenetic groups (clades A and B) (**Figure 4.5**). The two main groups comprised four (A1-A4) and two (B1 and B2) sub-clades for clade A and clade B, respectively. The larger clade (clade A), containing the three component stocks of the Muguga cocktail vaccine, was composed of the majority of sequences. These sequences carried 20 *Tp2* alleles from cattle-derived *T. parva* (162 sequences of the 241 overall individual sequences; 67%) and 31 alleles from buffalo and cattle sharing grazing land with buffalo (35 individual sequences). The cattle-derived sequences found in this clade were clustered within the sub-clade A1 together with the three vaccine strains and were broadly distributed in various geographical areas in Africa (DRC AEZ1, 2 and 3, Burundi AEZ1, Kenya, South Sudan and Katete in the Eastern Province of Zambia) (Table S4.8). In general, the more diverse buffalo-derived isolates found in clade A tend to be clustered in exclusive separate sub-clades (A2, A3 and A4), although sub-clade A1 contained mixed *T. parva* sequences from cattle and buffalo (or buffalo-associated cattle). The minor clade (clade B) contained 10 alleles (44 sequences) and had two independent sub-clades. The first sub-clade (B1) consisted exclusively of *T. parva* sequences from buffalo (6 sequences giving 5 alleles), while the second sub-clade (B2)

comprised only cattle-derived *T. parva* (38 sequences carrying 5 alleles). Cattle-derived parasites found in clade B originated from DRC (AEZ1 and 2), Burundi AEZ1, Kenya, Uganda, Southern Province of Zambia (Chitongo) and Zimbabwe (Boleni). It is worth noting that cattle-derived samples from the lowlands and midlands of DRC and Burundi contained more diverse *T. parva* alleles which were consistently found in the two main *Tp2* clades and had no obvious association between allelic clades or sub-clades and their geographical origins. However, remarkably all the *T. parva* samples from the highlands (DRC AEZ3) were clustered within the major clade (clade A).

The median-joining (MJ) network performed on the 200 cattle-derived *Tp2* individual sequences described in Africa (collapsed into 25 representative alleles) recovered two major genetic groups that diverged at least by 100 nucleotide mutational steps (**Figure 4.6**). The two groups fully corresponded to cattle-derived alleles clustered in clades A1 and B2 detected in the *Tp2* gene tree (**Figure 4.5**) and contained ubiquitous alleles that were shared by two to more populations. The majority of low-frequency alleles occurred in South Sudan and were closely connected to the dominant allele (A01) which was present in all the seven populations. Interestingly, The MJ network showed an extensive admixture of cattle-derived parasite populations from diverse geographical locations with high number of mutational step connections.

The evolution of loci was compared with the evolution of *T. parva* samples using a concatenated phylogenetic analysis performed on 93 *Tp1+ Tp2* individual sequences from cattle in the Great Lakes region. In total, 19 representative alleles were defined in the concatenated sequences. The NJ tree of the concatenated dataset provided similar topologies that fully agreed with that obtained by the *Tp2* phylogenetic analysis (**Figure 4.5**, **Figure 4.6**), resulting into two well-defined main clades of cattle-derived parasites (Figure S4.4). Each of these clades was significantly divided into two sub-clades strongly supported by their bootstrap values. The major clade (clade A), which contained the three Muguga cocktail vaccine alleles, included 70 (75%) concatenated sequences corresponding to alleles found in sub-clade A1 of the *Tp2* NJ tree, while the minor clade (clade B) contained samples clustered in sub-clade B2 of the *Tp2* phylogenetic tree (**Figure 4.5**).

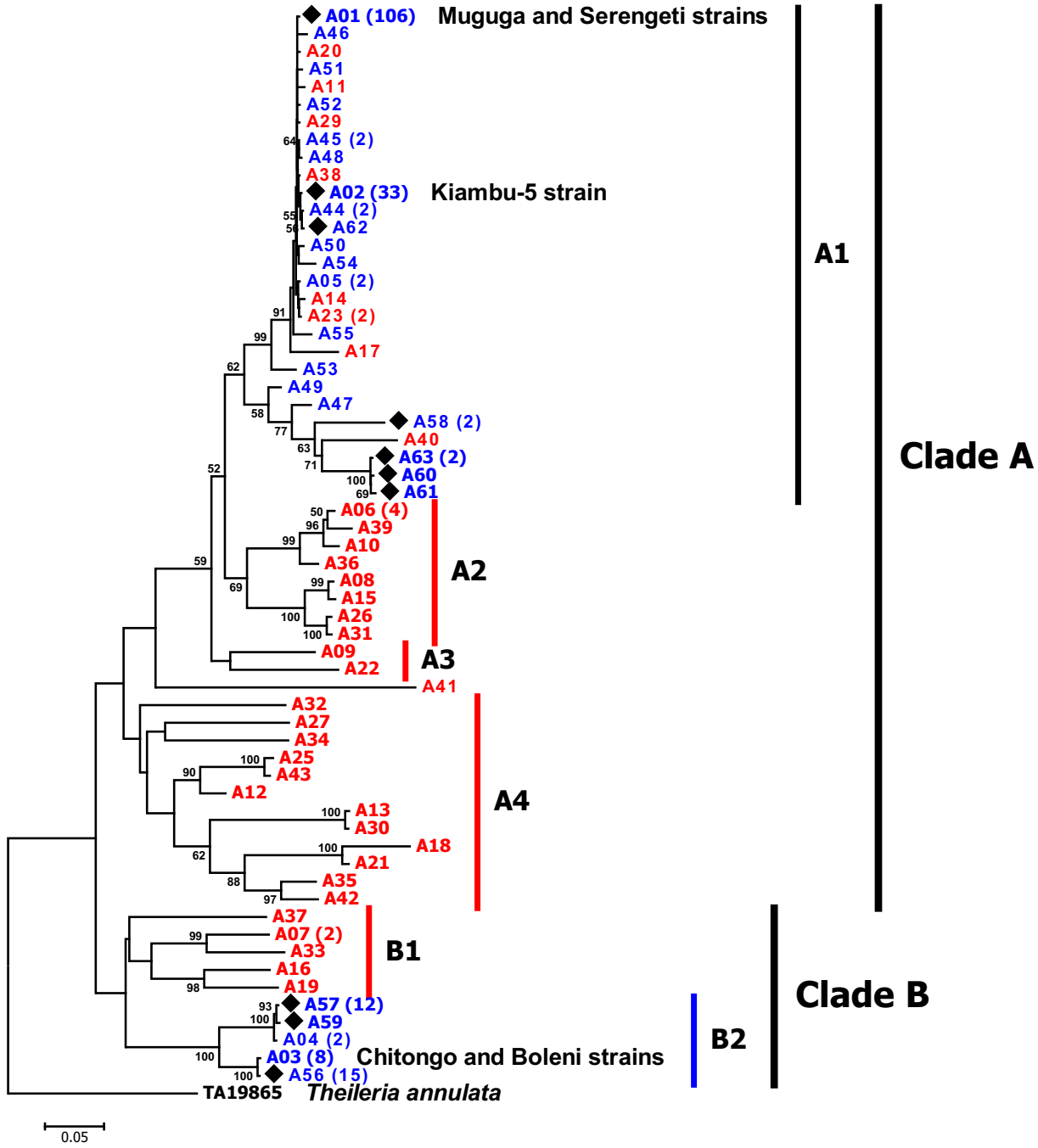

Figure 4.5 Neighbor-Joining tree showing phylogenetic relationships among the 61 *Tp2* gene alleles described in Africa (A01-A63). *Tp2* gene alleles obtained from cattle in the present study are indicated by black diamonds. *Theileria parva* alleles found in cattle with no association with buffalo and in laboratory stocks are coloured in blue, while buffalo-derived and buffalo-associated alleles are depicted in Red. Bootstrap values (> 50%) are shown above branches. The *Tp2* homologous sequence of *T. annulata* (GenBank: TA19865) was used as the outgroup. The number in brackets behind allele names denote the number of *T. parva* isolates carrying the allele. The detailed *Tp2* alleles distribution and their corresponding populations/AEZs are

presented in Table S4.8. *Tp2* allele A01 corresponds to isolates identical to Muguga and Serengeti-transformed strains, while *Tp2* allele A02 represents isolates identical to Kiambu-5 strain

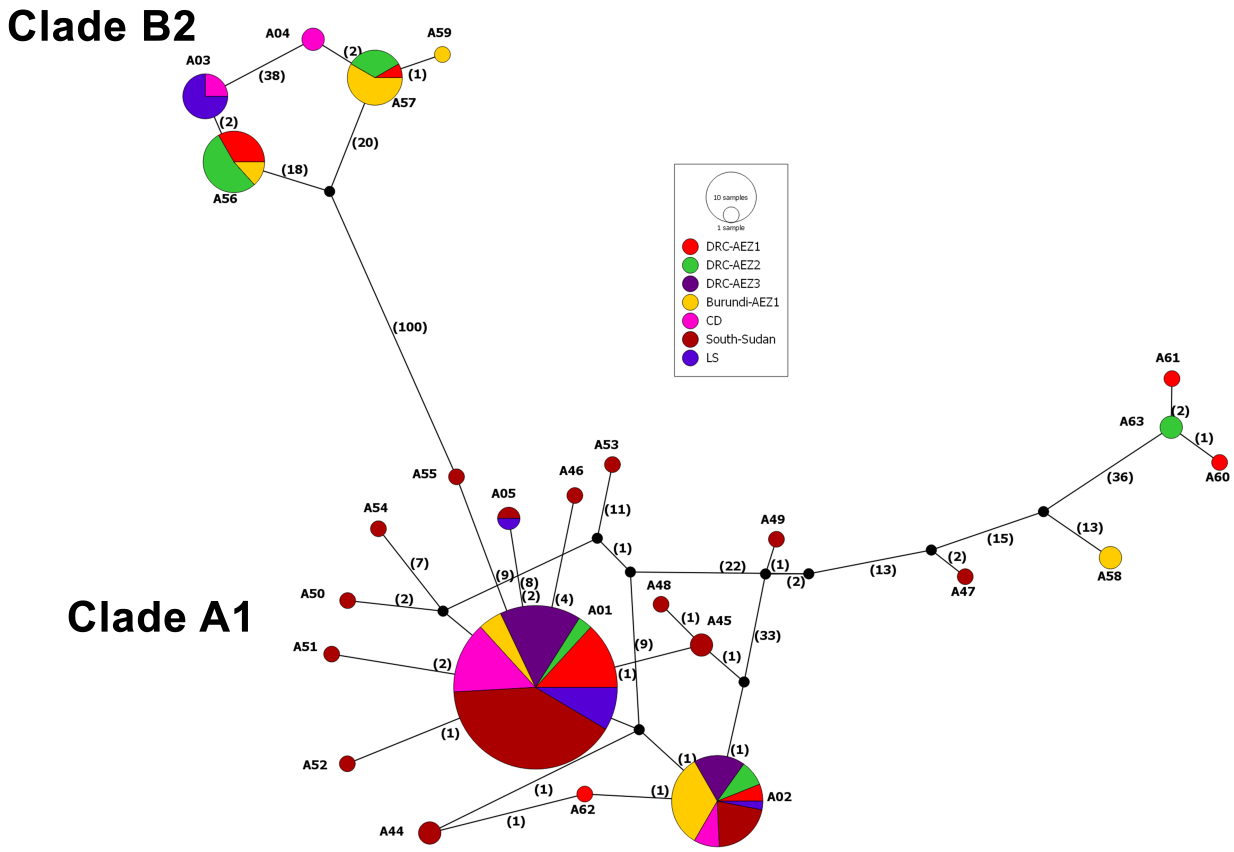

Figure 4.6 Median-joining network representing the phylogeographical distribution of *Tp2* alleles of *T. parva* from cattle in sub-Saharan Africa. Each circle represents a unique allele, with colours depicting the proportion of individuals from different populations sharing the allele. Black nodes represent hypothetical unsampled alleles (or median vectors). Numbers in brackets on connecting lines indicate mutational steps between alleles. The detailed *Tp2* alleles distribution and their corresponding populations/AEZs are presented in Table S4.8. *Tp2* allele A01 corresponds to samples identical to Muguga and Serengeti-transformed strains and *Tp2* allele A02 represents samples identical to Kiambu-5 strain. *Abbreviations*: CD, cattle-derived samples (from Kenya); LS, laboratory samples (ILRI) [36]

4.4.5 Population differentiation

The partition of genetic diversity in *Tp2* sequences was further analysed using analysis of molecular variance (AMOVA) based on allelic variants from the four AEZs of the Great Lakes region and those from South Sudan, Kenya (cattle-derived: CD; buffalo-derived: BD; and buffalo-associated: BA) and the laboratory *T. parva* stocks (LS). The AMOVA results further supported the findings obtained with phylogenetic analyses, showing that most of the variation (75% of the total variation) was found between individuals within populations, whereas a relatively small amount of the total diversity was significantly explained by interpopulation divergence (25%, $P < 0.001$). To examine the degree of gene flow and genetic differentiation levels among *T. parva* populations, Wright's fixation index (F_{ST}) values were computed for each pairwise comparison between *Tp2* sequences from different geographical origins, alongside buffalo-derived *T. parva* sequences obtained from GenBank. Overall, pairwise comparison (F_{ST}) values between different geographical areas and/or populations ranged from -0.02 (between LS and DRC AEZ1) to the greatest genetic divergence of 0.69 (between DRC AEZ2 and South Sudan) (**Table 4.4**). The F_{ST} statistic revealed interesting findings. First, *T. parva* isolates from the highlands of DRC (DRC AEZ3) and those from South Sudan were not genetically different ($F_{ST} = -0.003$), showing a high degree of similarity between alleles. These two populations contained the highest number of samples carrying the Muguga cocktail vaccine component alleles (alleles 1 and 2) (Table S4.8). Secondly, *T. parva* parasite samples from lowlands (DRC AEZ1 and Burundi AEZ1) and midlands (DRC AEZ2) were genetically distant from those of highlands (DRC AEZ3) and the ones from South Sudan. Thirdly, the laboratory isolates (most ancient isolates from different sub-Saharan countries) were neither significantly divergent from those from lowlands and midlands of DRC and Burundi nor from Kenyan field isolates (CD).

Table 4.4 Pairwise estimates of genetic distance among nine *T. parva* populations using F_{ST} statistic for nucleotide sequences of *Tp2*

Country	Population	DRC AEZ1	DRC AEZ2	DRC AEZ3	Burundi AEZ1	BA	BD	CD	South Sudan	LS
DRC	DRC AEZ1	–	0.008	< 0.001	0.2	< 0.001	0.001	0.4	< 0.001	0.4
	DRC AEZ2	0.17	–	< 0.001	0.06	< 0.001	0.004	< 0.001	< 0.001	0.07
	DRC AEZ3	0.20	0.58	–	< 0.001	< 0.001	< 0.001	0.04	0.3	< 0.001
Burundi	Burundi AEZ1	0.02	0.08	0.3	–	< 0.001	0.008	0.1	< 0.001	0.3
Kenya	BA	0.17	0.19	0.39	0.15	–	0.3	< 0.001	< 0.001	< 0.001
	BD	0.17	0.17	0.45	0.14	0.002	–	< 0.001	< 0.001	0.009
	CD	-0.01	0.28	0.13	0.06	0.22	0.23	–	0.002	0.1
South Sudan	South Sudan	0.29	0.69	-0.003	0.41	0.52	0.59	0.19	–	< 0.001
Laboratory samples	LS	-0.02	0.08	0.35	0.009	0.15	0.14	0.04	0.48	–

Notes: F_{ST} values below the diagonal and *P*-values above the diagonal; The genetic differentiation was considered as low (F_{ST} between 0–0.05), intermediate (F_{ST} between 0.05–0.15), great (F_{ST} between 0.15–0.25) and very great ($F_{ST} < 0.25$). The sample sizes (number of sequences) used in each population are shown in (Table S4.7, Table S4.8)
Abbreviations: AEZ, agro-ecological zones; BD, buffalo-derived; LS, laboratory samples; CD, cattle-derived

4.4.6 Evolutionary population dynamics: evidence of immune selection or demographic processes?

The evolutionary dynamics of *T. parva* isolates from cattle in different geographical areas in sub-Saharan Africa were assessed by neutrality statistics and mismatch analyses to elucidate natural selection and demographic forces responsible for maintaining the observed polymorphism in the *Tp2* gene. We applied Tajima's *D* and Fu and Li's *D** and *F** statistics to assess the mode and significance of any departure from neutral expectations for the entire sequence (**Table 4.5**) and using a sliding-window through the sequence (**Figure 4.3a**). Overall, these analyses showed significant departure from neutral evolution expectations. These statistics, together with the multimodal mismatch pattern, confirmed the significant deviation from population expansion for the majority of studied populations. However, the evidence of population expansion signature was detected only in South Sudan population where the neutrality statistics were negative and significant (**Table 4.5**). For other populations, positive and significant values of Fu and Li's *D** and *F** statistics consistently observed in most areas suggested a significant pressure of balancing selection or diversifying selection that might be the reason of increased allelic frequency and nucleotide diversity, acting to maintain *Tp2* alleles at intermediate frequencies compared with expectations under neutrality (**Table 4.2**). In addition, the sliding-window plot showed that *Tp2* gene region of nucleotide positions 200–300 was the more diverse and subjected to positive values of Tajima's *D* statistic (**Figure 4.3a, b**). This region contains sequences for *Tp2* epitopes 4 and 5 and another part that does not contain defined epitopes, suggesting that the evolutionary pressure signature is randomly distributed within the gene.

Table 4.5 *Tp2*-based demographic structure and natural selection analyses of *T. parva* populations

Country	Population	Sample size	Tajima's *D*	Fu and Li's *D**	Fu and Li's *F**
DRC	DRC AEZ1	25	1.2	0.97	0.94
	DRC AEZ2	20	2.5	1.8**	2.1**
	DRC AEZ3	23	1.1	0.84	1.1
Burundi	Burundi AEZ1	28	2.2	1.9**	2.1**
	Tp2 Clade A	68	-1.2	1.9**	0.6
	Tp2 Clade B	28	3.6**	1.6**	2.6**
Overall (DRC and Burundi)		96	2.5	2.3**	2.3**
Kenya	CD	22	0.17	1.8**	1.4
South Sudan	South Sudan	65	-2.5**	-2.7*	-3.1*
Laboratory samples	LS	17	2.7**	1.7**	2.3**

*$P < 0.05$, **$P < 0.01$

Abbreviations: AEZ, agro-ecological zone; LS, laboratory samples; CD, cattle-derived

4.5 Discussion

The rationalisation and implementation of an effective ECF vaccine-based control require information of the circulating parasite antigenic variants in a region to assure vaccine efficacy (when vaccinated animals are exposed to wild parasites) and safety (cross-immunity is required if vaccine stock is transmitted by ticks to an immune cattle population). Previous genetic studies of *T. parva* schizont-infected cell lines and parasite field isolates from cattle and African buffalo in East Africa using schizont antigen genes revealed an extensive genetic and antigenic diversity in *T. parva* populations, which was much greater in buffalo than in cattle-derived parasites (Elisa et al., 2015; Hemmink et al., 2018; Pelle et al., 2011; Salih et al., 2017; Sitt et al., 2018). In this study, we conducted a comprehensive analysis of *Tp1* and *Tp2* sequences to investigate the extent of diversity, the phylogenetic relationships and the evolutionary dynamics of *T. parva* samples obtained from cattle in four AEZs in the African Great Lakes region and determine how they relate to vaccine stocks and published sequences from various geographical areas of sub-Saharan Africa. We were particularly interested in understanding the role of agro-ecological conditions and anthropogenic movements of cattle in the genetic structuring and evolutionary dynamics of *T. parva*.

4.5.1 *Theileria parva* populations are more variable in lowlands than highlands but ubiquitous alleles are identical to the Muguga vaccine components

The sequence analyses provided evidence of polymorphism at the nucleotide and amino acid levels and within the epitope-containing regions of the two genes in the *T. parva* population from the Great Lakes region. Genetic distance statistics showed particularly a higher level of similarity within *Tp1* sequences and an extensive diversity within *Tp2* sequences, supporting the evidence that the genetic diversity is greater in *Tp2* than in *Tp1* gene as previously reported (Hemmink et al., 2018; Pelle et al., 2011; Sitt et al., 2018). Nevertheless, the major alleles and epitope variants identified in the two genes were identical to those found in the Muguga cocktail vaccine components. Besides, the genetic diversity results further showed that the parasite populations from highlands were less diverse compared to those from lowlands (DRC AEZ1 and Burundi AEZ1) and midlands (DRC AEZ2), which contained the majority of the genetic variation observed in the Great Lakes region. Interestingly, all the AEZs consistently shared the Muguga cocktail vaccine component alleles, that were the most ubiquitous in the region. The fact that sequences identical to the alleles in the Muguga cocktail were the most prevalent and broadly distributed may be associated with the reported unrestricted movement of cattle in the region (Amzati et al., 2018; Bazarusanga et al., 2007b; Bouslikhane, 2015; Kalume et al., 2012; Verweijen and Brabant, 2017). Altogether, these findings indicate that the Muguga cocktail component alleles seem to be endemic and the most transmitted and circulating genotypes in the Great Lakes region, while their coexistence with other genetically distant and more diverse alleles in lowlands and midlands areas might be generating epidemics or unstable endemic situations. Furthermore, nucleotide sequence analysis of *T. parva* at the sub-Saharan African level revealed that cattle-derived *T. parva* populations circulating in the Great Lakes region, especially from lowlands and midlands of DRC and Burundi are more diverse in comparison with those reported in cattle from various ecological zones of sub-Saharan Africa (Elisa et al., 2015; Pelle et al., 2011; Salih et al., 2017). These results further suggest that the level of allelic variation of *T. parva* may be significantly affected by the demographic processes such as broad geographical dispersal of the parasite populations through human population migration with their cattle, which consequently result in high connectivity between cattle populations in Africa (Bouslikhane, 2015; Ndumu et al., 2008; Verweijen and Brabant, 2017).

4.5.2 Limited population structure and geographic separation of *T. parva*

A comprehensive phylogenetic analysis of *T. parva Tp2* sequences from the Great Lakes region and those from other regions across sub-Saharan Africa strongly support the evidence that *T. parva* parasites circulating in cattle from the Great Lakes region are highly diverse, containing individuals similar to those found in cattle from most east African countries and newly described alleles (Elisa et al., 2015; Pelle et al., 2011; Salih et al., 2017). The topologies derived from phylogenetic analysis deduced from the concatenated cattle-derived sequences (*Tp1+ Tp2*) found in the Great Lakes region produced strong congruent results with *Tp2* analysis, suggesting that the two loci co-evolve with similar substitution patterns in cattle-derived *T. parva* samples. The NJ and MJ network algorithms clustered *T. parva* sequences into two main groups (clades). The Muguga reference sequence noticeably clustered together with the majority of sequences in the larger clade while the smaller clade contained alleles that are genetically the most distant from the Muguga reference alleles. However, *T. parva* genetic groups were not clearly separated by geographical sub-structuring, as there were no population-specific clade or sub-clade consistently associated with geographical origins. A particular striking result was that, despite the large overall nucleotide diversity of *Tp2* sequences, the diversity within each clade was lower, strongly suggesting that the overall genetic variation was predominantly affected by the genetic divergence among samples belonging to different clades and poorly among samples from different AEZs. These patterns further reflect a limited geographical segregation of *T. parva* genotypes which seems to be explained by the occurrence of most dominant alleles (Muguga component stocks) in all geographical areas. The reduced population structure could be the evidence that balancing selection acting on the genes studied or gene flow through cattle immigration is maintaining similar ubiquitous alleles in *T. parva* populations from distinct geographical regions (Fijarczyk and Babik, 2015; Ndumu et al., 2008). This was further supported by the AMOVA, which indicate that the sequence variation was substantiality higher between individuals within populations rather than among populations.

The degree of gene flow and genetic differentiation among the populations was assessed by estimating F_{ST} statistics, which investigate the level of population subdivision. Although the phylogenetic analysis did not give a clear population structure or geographical grouping, there is evidence of statistically significant genetic differentiation between *T. parva* populations, mostly due to the presence of some unshared and more diverse alleles that are exclusive to

parasite populations from particular AEZs. Overall, high genetic differentiation was observed between lowlands and highlands, supported by the strong evidence that all *T. parva* samples found in highlands were closely related to the Muguga cocktail vaccine stocks. The lowlands and midlands of DRC and Burundi had similar levels of genotypic distribution and variation. These areas are relatively close and may exchange more genotypes through the dispersal of the parasites during short-distance seasonal movement of cattle. In addition, cattle movements are very extensive in the Ruzizi valley (lowlands of Burundi and DRC) which is an important entry point for imported cattle from neighbour countries (Verweijen and Brabant, 2017; Vlassenroot and Huggins, 2005).

4.5.3 Ecological conditions driving tick population dynamics are suggested to be affecting the biogeographical distribution of *T. parva* genotypes

The observed pattern of genotypic distribution suggests that ecological parameters driving the phenology and establishment ability of tick lineages seem to be further affecting the transmission dynamics of *T. parva* and consequently its genetic diversity and structure. In the African Great Lakes region, AEZs are mainly differentiated by temperature and rainfall (affected by altitude), which are crucial factors underlining the ecology and population dynamics of the tick vector (Berkvens et al., 1998; Leta et al., 2013; Olwoch et al., 2009; Speybroeck et al., 2002; Vajana et al., 2018). Thus, these environmental factors might be involved in determining the population structure of ticks as well as the transmission pattern of specific genotypes of the pathogen in different ecological conditions. Our previous findings of the population structure of the tick vector *R. appendiculatus* allow a direct linking with population structure of the pathogen at the agro-ecological level (Amzati et al., 2018). We found that the diversity of *R. appendiculatus* had a strong altitudinal gradient, being lower in highlands and more extensive in lowlands. With this evidence, we hypothesise that the association between *T. parva* genotypes and biogeographical areas could be explained by the climate factors affecting tick vector capacity (Berkvens et al., 1998; Madder et al., 2002; Speybroeck et al., 2002). The extensive genetic diversity of *T. parva* observed in lowlands and midlands appears to be supported by the intensity of tick activity in cattle which could increase the transmission dynamics of the parasite and multiple reinfection and coinfection events (Bazarusanga et al., 2007a; Bazarusanga et al., 2011; Kalume et al., 2013; Oura et al., 2005). In addition, the lowlands areas are ecologically more suitable for the sympatric coexistence of two lineages of the tick vector with different diapause behaviour, which may allow the temporal

persistence of ticks on cattle and permanent transmission of the parasite (Amzati et al., 2018). Therefore, the repeated and permanent acquisition and continued transmission of parasites may result in genetic recombination among *T. parva* genotypes during their sexual reproduction stage in the tick and generate new genotypes in the parasite population (Katzer et al., 2006). In contrast, the low level of genetic diversity observed in highlands could be a result of reduced tick burden that may consequently reduce the transmission intensity of *T. parva* (Amzati et al., 2018; Bazarusanga et al., 2007a; Kalume et al., 2013). The likely lower transmission intensity in highlands could restrict the effective population size of the parasite and reduce its diversity.

4.5.4 Lack of evidence for recent host immune selective pressure but suggested demographic processes affecting the evolutionary structure of *T. parva*

Theileria parva population dynamics were inferred using neutrality statistics in order to understand the factors underlying the observed genetic variability in Africa. The results of these statistics suggested that balancing selection occurred in most populations except in South Sudan where *T. parva* parasites appear to have experienced a sudden demographic expansion (Salih et al., 2017). However, although neutrality statistics provided positive values suggesting evidence for balancing selection, which might arise as a result of selective pressure of host immunity and might increase the frequency distribution of polymorphisms, it is worth noting that this pattern of departure from neutral evolution can also be caused by demographic processes such as immigration dynamics and population colonisation and admixture (Biswas and Akey, 2006). Previous studies provided evidence of positive selection pressure for amino acid changes acting on *Tp1* and *Tp2* genes, but there was no sufficient evidence of host immune-based selection (Pelle et al., 2011). In addition, it seems that the evolutionary pressure is not predominantly directed to known epitope regions but is randomly distributed across the entire region of the gene. Thus, the observed selection and polymorphism could have arisen either through immune selection acting on epitopes presented by the bovine MHC class I and recognised by $CD8^+$ T cells or most likely from demographic processes of the parasites due to range expansion through cattle movements. Moreover, a recent comparison of the polymorphism in the T cell epitopes of geographically distant *T. parva* parasite populations from buffalo showed that both populations consistently shared a large proportion of epitope variants, suggesting that the majority of variability found in the two genes is more ancient rather than a result of recent immune-based substitutions (Hemmink et al., 2018). This was further supported by the lack of genetic differentiation between the more diverse *T. parva* populations

from DRC and Burundi and the ancient laboratory isolates from cattle in various geographical areas of Africa. It was also suggested that variation observed in cattle-derived parasites may represent the ancient diversity evolved in buffalo and that only a subset founder population have been established within cattle population (Hemmink et al., 2018; Pelle et al., 2011). We can therefore suggest that the genetic distribution and variation of *T. parva* observed in the Great Lakes region are more affected by cattle translocation between populations (gene flow) and ecological traits regulating tick populations than the host immune pressure and other mechanisms such as selection, mutation and genetic drift (or bottlenecks).

4.5.5 The use of the trivalent Muguga vaccine is not expected to introduce new *T. parva* antigenic variants

The findings of this study provided a broad picture of the genetic structure of *T. parva* in the African Great Lakes region as a baseline for future fine scale description of the parasite population and immunisation trials of ITM vaccine. However, the prevalence and the number of *T. parva* genotypes circulating in the region may be underestimated, as some of the strains have a shorter carrier state and low parasitaemia below the detection threshold of antigen markers in asymptomatic cattle sampled during a cross-sectional survey (Geysen, 2008). Longitudinal monitoring of infections could be suggested in order to understand the spatiotemporal dynamics of *T. parva* genotypes in the region and further molecular characterisation could be undertaken using multilocus markers (Katzer et al., 2010; Oura et al., 2003) and high-throughput sequencing approach (Hemmink et al., 2018; Hemmink et al., 2016) or cloning parasites from individual animals to detect all possible diversity profiles. Of interest, the majority of *T. parva* samples analysed in this study have shown to carry alleles identical or nearly similar to Muguga cocktail vaccine strains, although an extensive diversity was observed in lowlands and midlands. The wide distribution of the vaccine alleles in the region may be used as reference point for vaccine trial composed with Muguga cocktail stocks to evaluate cross-immunity in field conditions using local strains as challenge without any risk of introducing new parasite variants. A striking finding was that some samples found in cattle from lowlands and midlands (Tp1: allele-45/var-32) were close to alleles present in buffalo derived and buffalo-associated parasites. These antigenic variants may break through immunity induced by the Muguga vaccine (Bishop et al., 2015; Sitt et al., 2015). Thus, it could be relevant to test an improved alternative vaccine in lowlands and midlands areas that include local parasite stocks to provide broad protection. In order to initiate a vaccination trial, the MHC

class I diversity in cattle from the Great Lakes region could be assessed because of the differential immune responses between cattle of different MHC class I haplotypes (Steinaa et al., 2012).

4.6 Conclusions

The present study sheds light on the strong genetic similarity among major *T. parva* genotypes circulating in the region and Muguga vaccine stocks. The high degree of variation observed within populations and the moderate agro-ecological sub-structuring suggested that *T. parva* genotypes evolving in cattle are circulating within and between African countries through short and long-distance cattle movement. The findings reported in this study also provide insight into factors affecting the population genetic structure and biogeographical distribution of *T. parva* in the African Great Lakes region. It appears that the local persistence and the geographical distribution of *T. parva* genotypes are mainly driven by ecological factors affecting tick vector population dynamics and competence. Furthermore, the widespread of major genotypes and the signature of selection are most probably related to extensive gene flow through cattle immigration and agro-ecological conditions determining the transmission intensity of *T. parva* rather than a recent mutational process of immune selective pressure. The observed patterns of genetic structure and diversity of *T. parva* indicate that the strong genotypic diversity found in the region would be generating ECF endemic instability in lowlands and midlands and an epidemic structure in highlands. However, the fact that ubiquitous alleles are genetically similar to those used in the Muguga vaccine, along with the high level of admixture, partially provides evidence for safe deployment of existing trivalent live vaccine for field trial without any risk of introducing new parasite variants in the Great Lakes region. The Muguga cocktail ITM vaccine trial could be implemented regardless of agro-ecological zone since animal movement plays an important role in the spread of major genotypes. Future efforts should be done to understand the vector-pathogen and host-pathogen genotype relationships in the transmission system and the spatiotemporal dynamics of *T. parva* genotypes.

Chapter 5. Transmission dynamics of *Theileria parva* in the Eastern Democratic Republic of Congo*

This chapter presents the current epidemiological situation of ECF based on the seasonal transmission dynamics of T. parva *in three AEZs of DRC. In the first study (**Chapter 3**), we found that two lineages of* R. appendiculatus *occur in the Great Lakes region with a negative altitudinal gradient of their coexistence. The results of **Chapter 4** revealed that the diversity of* T. parva *and its biogeographical distribution depend on host migration and agro-ecological conditions driving tick population dynamics. It can therefore be postulated that the occurrence and abundance of ticks, together with agro-ecological variability may lead to differences in the transmission dynamics of* T. parva *and consequently different epidemiological profiles of ECF. In this study, ticks were collected from cattle (attached ticks) and from the vegetation (free-living ticks). The attached ticks were used to estimate the tick burden while the free-living ticks were used to assess the infection rate of* T. parva*. The transmission intensity of* T. parva *was then modelled as a function of the tick burden and the transformed infection rate in individual ticks, and expressed directly the vectorial inoculation rate (VIR). In addition to tick samples, cattle blood samples were analysed to estimate the prevalence of* T. parva *infection in the bovine host. The pattern of transmission intensity and dynamics of* T. parva *suggested an endemic situation in lowlands and midlands areas where there was high tick challenge and an epidemic situation in the highlands where tick activity was reduced.*

* Manuscript in preparation

5.1 Abstract

Theileria parva is a protozoan parasite which causes East Coast fever (ECF), a lethal lymphoproliferative disease of cattle in sub-Saharan Africa. It is transmitted by the ixodid tick *Rhipicephalus appendiculatus*. This study was carried out to investigate the agro-ecological and seasonal variations of the transmission dynamics of *T. parva* in the South Kivu Province of the eastern Democratic Republic of Congo. Three cross-sectional studies were conducted during three seasons (two wet seasons and one dry season) in three agroecological zones (AEZ), namely lowlands, midlands and highlands. We assessed *R. appendiculatus* tick burden and molecular prevalence of *T. parva* in cattle as well as in free-living *R. appendiculatus* collected from the vegetation. We then estimated the transmission intensity of *T. parva* by mean of the vectorial inoculation rate (VIR), by expressing the number of infective ticks attached to cattle during a feeding period assumed to be one week. A total of 1424 blood samples and 13,704 *R. appendiculatus* ticks were collected from cattle and 1667 free-living ticks from grazing lands. The overall *T. parva* prevalence in cattle was 32%, with significant variations across AEZs and between seasons. *Rhipicephalus appendiculatus* was the most abundant tick species with a mean burden of 26 ticks per animal, corresponding to 83% of the tick load. The tick burden per animal was significantly higher in lowlands and midlands than highlands. The overall *T. parva* prevalence in individual free-living ticks was 4% with non-significant variation among AEZs and between seasons. The number of infective ticks varied significantly among AEZs and among seasons within AEZs, being higher in lowlands and midlands during the two wet seasons. The agro-ecological and seasonal variation in the transmission intensity of *T. parva* was primarily predicted by the abundance of tick vector rather than the differences in tick infection rate among AEZs. These findings have important implications in the epidemiology of ECF. The pattern of transmission intensity and dynamics of *T. parva* suggested an endemic situation in lowlands and midlands areas where there is high tick challenge and an epidemic situation in the highlands where tick activity and the *T. parva* infection in cattle are reduced. However, epidemic instability is likely to occur in lowlands due to cattle movements and the genetic composition of *T. parva* and *R. appendiculatus.*

5.2 Background

East coast fever (ECF), caused by the intracellular protozoan parasite *Theileria parva* and transmitted by the ixodid tick *Rhipicephalus appendiculatus,* is one of the most pathogenic tick-borne diseases of cattle which hinder cattle production in eastern, central and western Africa (Bett et al., 2019; Boucher et al., 2020; Gachohi et al., 2012; Morrison, 2015). *Theileria parva* transmission occurs transstadially; *R. appendiculatus* larval or nymphal stages acquire the infection while feeding on infected cattle and transmit the pathogen during their subsequent nymphal or adult developmental stages after moulting (Konnai et al., 2006; Olds et al., 2018). *Theileria parva* is widely spread in the Great Lakes region due to the wide distribution and dispersal of its tick vector favoured by cattle movements (Amzati et al., 2018; Bazarusanga et al., 2007a; Kalume et al., 2013). In the South Kivu Province of the Democratic Republic of Congo (DRC), live cattle are usually imported from neighbouring countries for trade and grazing, with very limited veterinary control. These unrestricted cross-border movements of cattle together with local transhumance system are thought to have important implications on the epidemiology of ECF through the spread of non-endemic ticks and pathogens strains, as ticks have potential to become established in new environments (Barre and Uilenberg, 2010; De Deken et al., 2007; Marcellino et al., 2017; Yssouf et al., 2011).

Rhipicephalus appendiculatus ticks undergo a three-host life cycle which occurs both on cattle and on the vegetation. It is believed that the geographical distribution and seasonal population dynamics of such tick vectors are strongly driven by variations in ecological conditions, host availability and density, host movements and farming management practices (Dantas-Torres, 2015; Leta et al., 2013; Olwoch et al., 2003; Perry et al., 1990). The interaction between ticks, pathogens, hosts and environmental conditions have substantial consequences in the establishment ability, vector competence of ticks, intensity of pathogen transmission and disease epidemiology (Leger et al., 2013; Ochanda et al., 1998; Young et al., 1996). In areas where agro-ecological conditions favour optimal interaction, continuous transmission of *T. parva* may lead to endemic state in which only calves are susceptible (Gilioli et al., 2009; Kivaria et al., 2004; Medley et al., 1993). However, agro-ecological variability and movements of cattle with possible introduction and establishment of potentially more virulent exotic *T. parva* strains and more competent ticks may disrupt the endemic situation and lead to epidemics (De Deken et al., 2007; Sitt et al., 2015; Uilenberg, 1999).

Studies suggest that *R. appendiculatus* stocks originating from diverse agro-ecological areas differ in their biological features and their ecological preferences and plasticity. These include

diapause induction and intensity (Madder et al., 1999; Madder et al., 2002), body size (Chaka et al., 1999), infection rate and vector competence (Ochanda et al., 1998; Odongo et al., 2009). Based on these phenotypic and physiological divergences, associated with phylogenetic analyses, two main populations of *R. appendiculatus* have been identified in Africa: the equatorial and tropical lineages (Amzati et al., 2018; Kanduma et al., 2016a; Leta et al., 2013; Mtambo et al., 2007c). The two lineages occur in the Great Lakes region, with differences in their agro-ecological distribution and colonization patterns (a negative altitudinal gradient of their coexistence) (Amzati et al., 2018). It can therefore be postulated that the occurrence and abundance of ticks, together with agro-ecological variability may lead to differences in the transmission dynamics of *T. parva* and consequently different epidemiological profiles of ECF in space and time.

Determining the seasonal dynamics and agro-ecological distribution of *R. appendiculatus* ticks and *T. parva* infection rate is of great significance for the epidemiology of ECF and facilitate the deployment of appropriate control measures either though vector control or cattle vaccination. To assess the epidemiological profile and the dynamics of transmission intensity of *T. parva* in the eastern DRC, we conducted repeated cross-sectional surveys on traditionally managed indigenous cattle population during the dry and the wet seasons in three different AEZs. The study focused on the analyses of the tick burden on cattle and the molecular prevalence of *T. parva* infections in cattle and in pooled free-living ticks. As suggested by (Ostfeld et al., 2006), the risk of exposure to vector-borne pathogens is mainly determined by the abundance of the vector and their infection rate. So, the transmission intensity of *T. parva* was modelled as a function of the tick burden and infection rate of individual ticks, and expressed directly the vectorial inoculation rate (VIR). The VIR estimates, assessed as a proxy measure of incidence, expressed the mean number of infective ticks that successfully attach on an animal during the feeding period. The study aimed to investigate the effect of AEZs and seasonal variation on the abundance of *R. appendiculatus* and the transmission dynamics of *T. parva.* The findings of this study contribute to a better understanding of ECF epidemiology in the South Kivu Province of DRC to guide control measures of the disease.

5.3 Methods

5.3.1 Study area

Sampling was performed from September 2014 to August 2015 across three agro-ecological zones (AEZs) of the South-Kivu province in the eastern Democratic Republic of Congo. The South-Kivu province lies between longitudes 26°47′- 29°20′E and latitudes 1°36′-4°51′S. The landscape morphology is characterised by high variation of altitude ranging from 700 to 3000 m above sea level, with diversified agro-ecological conditions. The ecological diversity in South Kivu is largely defined by the topography, altitude and latitude which subsequently regulate other bioclimatic attributes, including rainfall amount and pattern, temperature and vegetation. In addition, the protected area (Kahuzi Biega National park) and the water bodies including Lakes Kivu and Tanganyika and the Ruzizi river (flowing between the two lakes) also influences the bioclimatic conditions in South Kivu. The climate is essentially tropical and strongly moderated by altitude (the climate is cooler and more humid with increasing altitude). It comprises three major seasons with a bimodal rainfall pattern: (i) the early rainy season (wet1: September to December), (ii) the late rainy season (wet2: February to May) and (iii) the dry season (dry: June-August). The two rainy seasons are alternating with a short dry period of approximatively 15 days (January-February). The duration of seasons varies between AEZs and the rainfall increases while the temperature decreases with altitude. The study area was classified into three major AEZs on the basis of altitudes and climate: (i) the lowlands (< 1200 m) which is located in the Ruzizi valley (along the Ruzizi River), (ii) the midlands (1200-1800 m) located in the district of Walungu and (iii) the highlands (1600-2800 m) in the district of Kabare and part of Walungu (Mulumemunene). Distinctive bioclimatic characteristics of AEZs are described in details in **Table 5.1**. The lowlands AEZ is a semi-arid area characterised by a tropical warm and dry climate, with warmer rainy seasons and a longer cool dry season. The vegetation is strongly dominated by savannah grasslands and small patches forest. The midlands AEZ is considered as an intermediate ecological zone between lowlands and highlands, which consists of a warm and humid tropical climate. The highlands area is much cooler (lower to mild temperature) and experiences abundant and frequent rainfalls, with a warm dry season receiving occasional rainfall. It falls within a montane humid tropical climate. The dominant vegetation is more variable, composed of montane and savannah grasslands, scrubs woodland and degraded forests in the midlands and highlands. In some places of highlands, the vegetation is marked by relict species of seasonal deciduous forests.

Table 5.1 Geographical and climatic characteristics of the three agro-ecological zones

Agro-ecological zone (AEZ)	Altitude (m)	Temperature (°C)	Rainfall (mm/year)	Rainy season
Lowlands	780-1100	23-25	800–1000	October-April
Midlands	1200-1800	17-21	1000-1500	September-May
Highlands	1800-2800	12-19	1350-2000	September-May

5.3.2 Cattle farming system

The majority of cattle herds are composed of Ankole breeds (Sanga type), kept under traditional extensive farming system (open grazing system) where animals are fed on communal natural grazing lands and housed overnight in "kraals". Cattle are subjected to seasonal transhumance for grazing and water resources during the dry season. Adult animals move long distances depending on feed availability while calves are permanently kept in kraals or around homestead during their first year of life. The Ruzizi valley (lowlands), with a denser grass cover, is the most important pastoral area holding the highest cattle densities in South Kivu and the main corridor and destination of cattle originating from Rwanda, Burundi, Tanzania and Uganda and from montane areas in South Kivu. The control of ticks is negligible and irregularly applied, except in commercial farms in the highlands where acaricides are used on a weekly basis.

5.3.3 Cattle blood sample collection

Three cross-sectional surveys were organised from September 2014 to August 2015. Blood samples were collected from a total of 1424 traditionally managed indigenous cattle in three AEZs during three seasons. Sampling was done during planned visits communicated to livestock keepers at least two weeks before. Three cattle herds (or cattle keepers) were randomly selected from 5-12 sites (villages) in each AEZ during each season. Within each herd, 5-10 adult cattle of over 3 years old were randomly sampled using the random number generator in Microsoft Excel. Blood samples were then taken from the jugular vein using EDTA tubes and preserved in a cool box containing ice packs. Samples were transferred the same day to the laboratory and approximatively 120 µl of each sample was spotted on Whatman FTA cards (Whatman Bioscience, Cambridge, UK). FTA cards were allowed to dry and were stored in separate labelled plastic bags at room temperature until further processing. The selected herds (or farmer) from one survey were excluded during the next survey to prevent the likelihood of re-sampling the same animal more than once. Additional information about the number of blood

samples collected in each AEZ and season is detailed in **Table 5.2.** These blood samples were further analysed for molecular detection of the presence of *T. parva*.

Table 5.2 Cattle blood and tick samples collected in three agro-ecological zones during three seasons

Sample type	Season	Agro-ecological zone (AEZ)			Total
		Lowlands	Midlands	Highlands	
Cattle blood: PCR positive samples/ no. of cattle sampled (%)	Wet1	138/261(53)	39/130 (30)	53/274 (19)	230/665 (35)
	Wet2	99/205 (48)	41/79 (51)	16/66 (24)	156/350 (45)
	Dry	65/150 (43)	14/59 (24)	20/200 (10)	99/409 (24)
	Total	302/616 (49)	94/268 (35)	89/540 (16)	485/1,424 (34)
Free-living ticks: PCR positive pools /no. of tested 5 tick pools (%)	Wet1	9/32 (28)	7/31 (23)	7/49 (14)	23/112 (21)
	Wet2	16/46 (35)	14/39 (36)	5/37 (13)	35/122 (29)
	Dry	2/18 (11)	3/20 (15)	1/10 (10)	6/48 (13)
	Total	27/ 96 (28)	24/90 (27)	13/96 (14)	64/282 (23)
Attached ticks: total no. of collected ticks/ no. of cattle sampled (mean no. of ticks per animal)	Wet1	2,247/77 (29)	2,853/73 (39)	1,235/66 (19)	6,335/216 (29)
	Wet2	2,090/54 (39)	2,073/49 (42)	968/42 (23)	5,131/145 (35)
	Dry	1,028/54 (19)	770/44 (18)	440/64 (7)	2,238/162 (14)
	Total	5,365/185 (29)	5,696/166 (34)	2643/172 (15)	13,704/523 (26)

Abbreviations: Dry, the season; Wet1, the early rainy season; Wet2, the late rainy season;

5.3.4 Tick collection from the vegetation

Free-living ticks were collected from the vegetation in open communal grazing areas during three seasons between September 2014 and August 2015 and covering three AEZs. At each AEZ, three sites (villages) were selected from which 2-3 sentinel pastures (grazing areas) were randomly selected for tick collection during 3-4 consecutive visits per season in the same pastures. Free-living ticks were collected by visually searching on the vegetation combined with dragging a white cotton blanket over vegetation in the selected grazing areas. All ticks attached to the cloth and those collected manually were immediately immersed in plastic tubes containing 70% ethanol and maintained at room temperature until identification and DNA extraction. Overall, a total of 1667 free-living ticks was obtained across the three AEZs during three seasons. The collected ticks were placed in separate tubes according to site, pasture and season of collection and were transferred to the laboratory for morphological identification. The adult free-living ticks were subsequently analysed for the presence of *T. parva*.

5.3.5 Sampling of attached ticks on cattle

Attached ticks were collected from 523 cattle among those selected for blood sampling in sites where ticks were collected in the vegetation (**Table 5.2**). Samples were collected based on scheduled visits and farmers were aware to prevent their cattle at least two weeks without acaricide treatment prior to the sampling. Tick collections were made from about 5-10 adult cattle in each herd and from 2-3 herds in each site (3-4 villages per AEZ). Ticks were collected in the afternoon when cattle came back from pastures to "kraals". During each visit, animals were restrained, the whole body was inspected and ticks were manually removed and directly placed into separate labelled vials containing 70% ethanol. After collection, tick samples were brought to the laboratory and stored at room temperature until subsequent identification and counting. Ticks obtained from cattle were used to estimate the tick burden, expressed as the mean number of *R. appendiculatus* ticks per animal.

5.3.6 Morphological identification and molecular confirmation of tick species

Tick samples were morphologically identified to genus and species level under a stereo microscope according to the standard taxonomical identification key proposed by Walker at al. (2003). Subsequently, only adults free-living ticks belonging to *R. appendiculatus* species were grouped in pools of five individuals and were further investigated for the presence of *Theileria parva*. Identified attached *R. appendiculatus* ticks were counted for each animal and recorded by herd, season and collection site. After identification, ticks were stored in 70% ethanol at 4°C

and used for subsequent analysis. Detailed data regarding the tick samples used in this study are shown in **Table 5.2**. Part of the specimens (25-30 ticks) of randomly selected free-living *R. appendiculatus* ticks in each AEZ and additional 5-10 ticks of each of other tick species were confirmed by molecular identification using the *cox*1 gene as described in our previous study (Amzati et al., 2018).

5.3.7 DNA extraction from pooled free-living ticks and cattle blood

Total DNA was extracted from blood and tick samples. Tick DNA extraction was performed on pooled adults free-living *R. appendiculatus*. Pools were made from adult ticks consisting of five ticks per pool according to pasture, collection site and season in each AEZ. A total of 282 pools from the tree AEZs and three seasons were analysed (**Table 5.2**). Ticks from each pool were washed twice in distilled water, air dried on sterile filter paper, crushed and homogenised by grading prior to DNA extraction. Total DNA was then extracted from the homogenates of pooled ticks using the DNeasy Blood and Tissue Kit (Qiagen, Hilden, Germany) according to the manufacturer's guidelines. DNA was eluted from the column in a final volume of 50 µl TE buffer and stored at -20 °C until used for PCR amplification. A negative control DNA was extracted from uninfected laboratory ticks. DNA concentration was measured by spectrophotometry using Nanodrop (Nano Drop® ND-1000, PeqLab Erlangen, Germany) and extraction efficiency was further validated on agarose gel electrophoresis. Besides, cattle blood DNA samples were extracted using the Invitrogen Kit (PureLink® Genomic DNA Mini Kit, Invitrogen, Schwerte, Germany), according to the manufacturer's as described in our previous study (Amzati et al., 2019).

5.3.8 *Theileria parva* detection in DNA from pooled free-living ticks and cattle blood

Purified DNA samples from both pooled free-living ticks (282 pools) and cattle blood (1424 samples) were screened for the presence of *T. parva* DNA using a more sensitive nested PCR (nPCR) approach, that target the *T. parva* specific p104 gene (Iams et al., 1990; Odongo et al., 2010; Skilton et al., 2002).The oligonucleotide primer pair IL3231 and IL755 was used in the first PCR amplification followed by the secondary amplification using the internal forward primer IL4243 and reverse primer IL3232 that amplify a 277 bp fragment of the p104 gene. Primary PCR assays for tick DNA were performed in a final reaction volume of 50 µl containing 25 µl of 2× AccuPower® PCR PreMix (Bioneer PCR-PreMix, Seoul, South Korea), ~20-50 ng of DNA, 0.2 µM of each primer, and nuclease free water added to bring the reaction up to 50 µl. The nPCR was run using the same components, except that the final reaction volume was

20 µl which contained 1µl of DNA template from the primary PCR and 12.5 µl of 2× PCR PreMix. Thermal cycling conditions were as following: 94°C for 1 min (initial denaturation), followed by 30 cycles at 94°C for 1 min (denaturation); 60°C for 1 min (annealing), 72 °C for 1 min (extension) and a final extension at 72 °C for 9 min. For the nested PCR assay, PCR conditions were the same as described above, except that annealing temperature of 55 °C and that 1 µl of the primary PCR product was used as template. DNA samples purified from cattle blood were amplified using the same nPCR approach as described previously (Amzati et al., 2019). A negative control PCR reaction containing nuclease-free water as template instead of DNA was included. Similarly, a positive DNA control obtained from the *T. parva* Muguga isolate was used in each PCR run. The PCR products of the nPCR for both tick and cattle DNA were separated by electrophoresis on a 1.8 % agarose gel and visualised under UV light after dyed with GelRed (Biotium Inc., Hayward, USA). Samples containing PCR products of the expected size (277bp) were treated as positive to *T. parva*. Randomly selected *T. parva* positive samples were sequenced and compared with other published sequence from various areas across Africa.

5.3.9 Statistical analyses

Our goal was to estimate the effects of AEZs and seasons as independent variables on the spatial and temporal variation in cattle exposure to ticks and *T. parva.* Statistical analyses were performed in STATA 11 (Stata Corp, Texas USA) and R software version 3.6.1 (R Core Team, 2019: https://www.r-project.org/) using three separate generalised linear models (GLM) for each data set containing different dependant variables: (i) the prevalence of *T. parva* infection in cattle blood, (ii) the prevalence of *T. parva* in pooled free-living ticks and (iii) the number of ticks per animal. The datasets of *T. parva* prevalence in cattle and in pooled free-living ticks (binary outcome) were analysed using the logistic regression model with logit link function, while the number of ticks per animal (count outcome) was analysed with Poisson regression, followed by negative binomial model to account for overdispersion using the log link function. In both models, fixed explanatory variables were agro-ecological zones with three levels (lowlands, midlands and highlands), seasons (three levels: wet1, wet2 and dry) and their interaction. Basic regression models with only fixed effect terms were extended by robust estimation of standard error for survey, using collection sites as clusters. The design effects parameter (DEFT) was then estimated to evaluate the effect of clustering within sampling sites. The *T. parva* prevalence in cattle and in pooled ticks and their confidence intervals (CI 95%) were computed for different AEZs and seasons. Predictions and CIs were performed based on

the robust model in case of significant intra-cluster correlation (DEFT > 1) and on the basic model in the absence of correlation (DEFT < 1).

5.3.10 Computing the vectorial inoculation rate

The vectorial inoculation rate was estimated as a function of the abundance of ticks on cattle and the infection rate of *T. parva* in free-living ticks. The linear estimators and standards errors were predicted from pooled prevalence and tick burden regression models. These estimations were used to create two separate normal distributions for each season within each AEZ. From these generated distributions, paired but independent random values were simulated 100,000 times. The pool infection rate was expressed as the number of positive pools divided by the total number of pools examined. *Theileria parva* infection detected in pools was expressed as the proportion and minimum infection rate based on the assumption that each PCR positive pool contained at least one positive tick (Sosa-Gutierrez et al., 2016). The distribution of estimated pool prevalence was transformed in individual prevalence based on a simplified algorithm using a probabilistic function: $P_i = 1 - (1 - P_p)^{1/n}$ (with: P_i: Individual prevalence, P_p: pool prevalence and n: number of ticks per pool) (Katholi et al., 1995; Mitchell and Pagano, 2012). The VIR distribution was then generated as a product of the obtained independent and paired tick burden and individual prevalence values. The percentiles values at 2.5, 50 and 97.5 were generated for the prevalence in individual ticks, tick burden and VIR distributions for each combination of seasons and AEZs. The VIR was interpreted as the mean number of infective *R. appendiculatus* ticks that attach to cattle, assuming that attached infected ticks successfully realise the feeding period of one week (Konnai et al., 2007).

5.4 Results

5.4.1 Tick species identification

A total of 1,667 adults free-living *R. appendiculatus* ticks were obtained on the vegetation and 13,704 attached *R. appendiculatus* ticks were collected from cattle in the three AEZs during the dry and the wet seasons (**Table 5.2**). The overall attached ticks (16,421 specimens) belonged to six species grouped into the genera *Rhipicephalus*, *Boophilus sub-genus* of *Rhipicephalus* and *Amblyomma*. The species identified among these 16,421 ticks collected from cattle, in decreasing order of relative abundance, were *Rhipicephalus appendiculatus,* the most abundant, accounting for 83% of the collected ticks, followed by *Rhipicephalus (Boophilus) decoloratus* (8.1%), *Amblyomma variegatum* (3.5%), *Rhipicephalus evertsi evertsi* (2.8%), *Rhipicephalus (Boophilus) microplus* (2.2%) and *Rhipicephalus compositus* (0.4%). *Rhipicephalus appendiculatus* was the most dominant tick species on cattle in all AEZs during the three seasons. The distribution and ecology of other tick species will be handled in a separate paper. Molecular identification confirmed the status of these species and clustered the tested specimens with their corresponding haplotypes previously published. Two lineages of *R. appendiculatus* previously identified in an earlier study were found in the present study (Amzati et al., 2018). The lineage A was present in all AEZs whereas lineage B was found predominantly in lowlands followed by midlands (data not shown).

5.4.2 Agro-ecological and seasonal dynamics in *Rhipicephalus appendiculatus* tick burden on cattle

The abundance of *R. appendiculatus* tick feeding on cattle was analysed with regard to AEZs and seasons. The mean tick count ranged between six and 42 ticks per animal depending on seasons and AEZs. The overall mean *R. appendiculatus* tick burden was 26 ticks/animal corresponding to 83% of the tick load. Summary results of the negative binomial regression model are shown in **Table 5.3**. Statistical analyses revealed significant variation of the abundance of *R. appendiculatus* on cattle across AEZs and seasons ($P < 0.001$) as depicted in **Figure 5.1a**. *Rhipicephalus appendiculatus* was significantly more abundant during the two wet seasons compared to the dry season (**Table 5.3**). There was a markedly higher mean number of ticks per animal in lowlands and midlands compared to highlands where the tick burden significantly decreased (IRR= 0.36 and $P < 0.001$). However, the three AEZs showed similar patterns of seasonal dynamics in *R. appendiculatus* tick burden, being significantly lower during the dry season compared to the wet seasons (wet1 and 2). There was a significant

interaction between AEZs and the seasonal dynamics of *R. appendiculatus*. In lowlands, the tick burden of *R. appendiculatus* averaged 19 ticks per animal during the dry season, with the greatest mean number of ticks per animal obtained during the late wet season. The tick abundance was higher during the late wet season (39 ticks/animal) compared to the early wet season (29 ticks/animal) (**Figure 5.1a**, **Table S5.1**). In the midlands, the tick burden similarly decreased in the dry season (17 ticks/animal) while it increased substantially during the early wet season (39 ticks/animal) and the late wet season (42 ticks/animal). However, there were no significant differences recognised between the tick burden recorded during the two wet seasons in this area. On the other hand, although the abundance of ticks was much lower during the dry season in all AEZs, the mean tick burden found during the dry season in lowlands and midlands was similar to that obtained in highlands during the two wet seasons.

Table 5.3 Summary of the negative binomial model for *R. appendiculatus* tick abundance (mean tick burden per animal) from cattle according to AEZs and seasons

Factor	**Level**	**IRR**[a]	**CI (95%)**		***P-value***	**DEFT**[b]
AEZ	Lowlands	Ref[c]				
	Midlands	0.92	0.61	1.3	0.61	1.01
	Highlands	0.36	0.23	0.55	< 0.001	1.2
Season	Dry	Ref[c]				
	Wet1	1.5	1.1	2.3	0.003	1.1
	Wet2	2.0	1.5	2.7	< 0.001	0.94
Interaction	Midlands & Wet1	1.5	0.86	2.5	0.08	1.1
	Midlands & Wet2	1.2	0.76	1.8	0.45	0.98
	Highlands & Wet1	1.8	0.93	3.4	0.006	1.4
	Highlands & Wet2	1.6	0.96	2.8	0.03	1.1

Notes: AEZ, agro-ecological zones; Seasons: Wet1, early wet season (September-December); Wet2, late wet season (February-May); Dry, dry season (June-August). The model was multivariate, including AEZs (three levels), seasons (three levels) and their interaction. [a]Incidence rate ratio; [b]DEFT, design effects computed using sample sites (villages) as subpopulation clusters assuming intra-village correlation. The model including "sites" as a random effect was a better model (DEFT >1) and was used to predict confidence intervals. [c]Ref., reference level: lowlands and dry season were used as reference for comparison in the model (P=0.05). CI, confident intervals

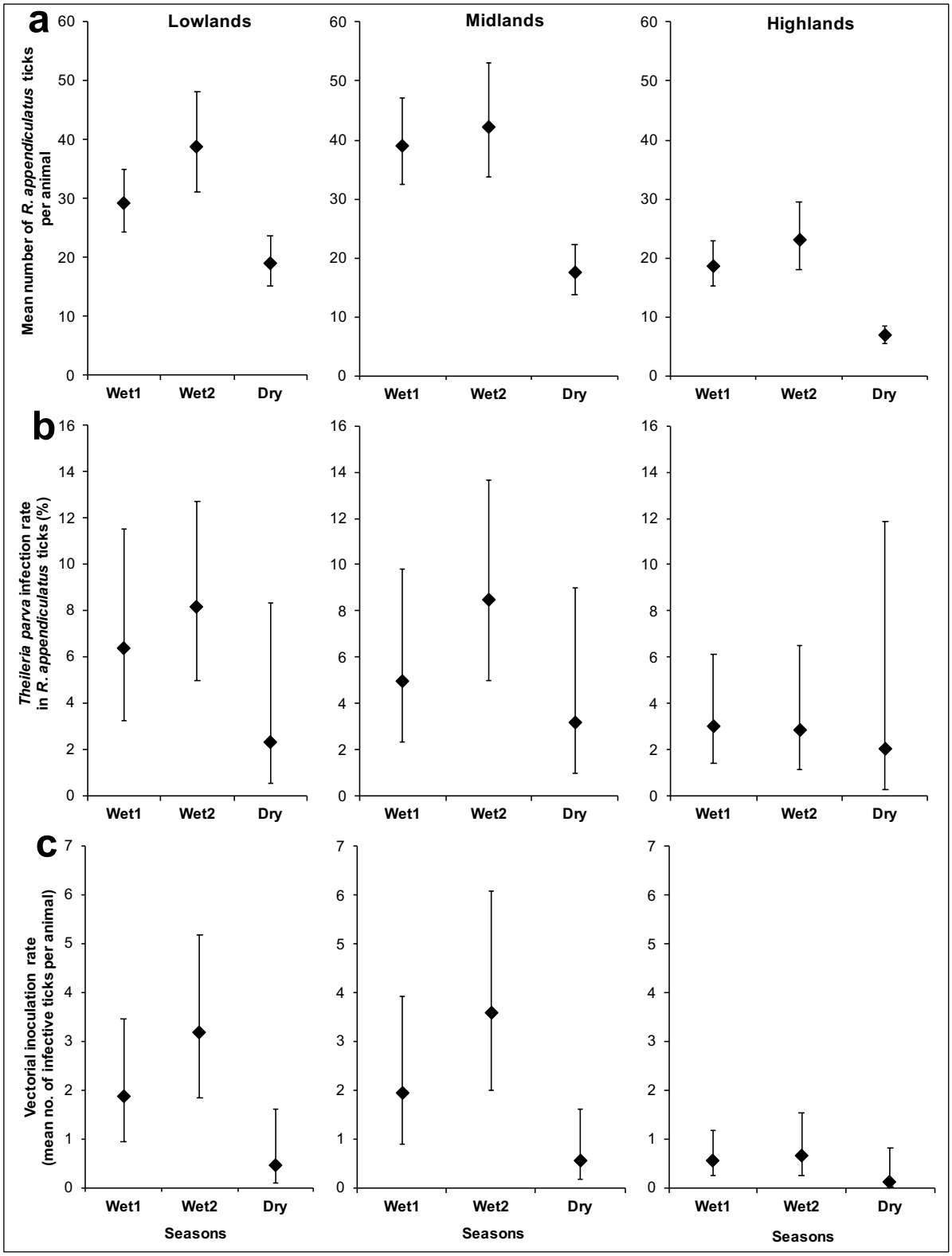

Figure 5.1 Agro-ecological and seasonal dynamics of (**a**) the abundance of *R. appendiculatus* ticks collected from cattle, (**b**) the infection rate of *T. parva* estimated in individual free-living ticks and (**c**) the vectorial inoculation rate (transmission intensity of *T. parva*). The graphs show the means and their 95% confidence intervals. *Notes*: AEZ, agro-ecological zones; Seasons: Wet1, early wet season (September-December); Wet2, late wet season (February-May); Dry, dry season (June-August).

5.4.3 *Theileria parva* infection rate in free-living ticks collected from the vegetation

A total of 282 pools of unfed free-living ticks collected from the vegetation in three AEZs during the three main seasons of the year were analysed for the presence of *T. parva*. Overall, *T. parva* DNA was detected in 64 (pool prevalence of 23%) out of the 282 tested pools (**Table 5.2**). There was no evidence of statistically significant variation in the proportion of *T. parva*-infected free-living *R. appendiculatus* neither among AEZs nor among seasons within each AEZ (**Table S5.3**). The individual infection rate simulated from pool infection showed that free-living *R. appendiculatus* ticks are infected with *T. parva* throughout the year without significant seasonal and agro-ecological patterns (**Figure 5.1b).** The predicted infection rate in individual ticks was moderate and ranged from 2% during the dry season in highlands to 8% during the late wet season in lowlands and midlands with an overall of 4% (Table S5.2).

5.4.4 Vectorial inoculation rate: *Theileria parva* transmission dynamics and intensity

The VIR simulated as the product between the individual tick infection rate and the mean tick burden displayed significant variations among AEZs and among seasons as depicted in **Figure 5.1c.** The overall estimated VIR for the study area was 1.6 infective ticks/animal. The *T. parva* transmission intensity decreased with increasing altitude and the seasonal transmission dynamics also varied significantly within AEZs. A markedly higher mean number of infective *R. appendiculatus* ticks per animal was detected in lowlands and midlands compared to highlands, although it was similar during the dry season in all the three AEZs. Of the three main seasons covered by the present study, the dry season had lowest VIR compared to the wet seasons. In the highlands, the VIR was relatively constant (around one infective tick/animal) throughout the year regardless the seasons, meaning that there was no significant seasonal trend in the transmission intensity of *T. parva* in this area. In contrast, in lowlands and midlands, the transmission intensity was strongly affected by seasons, being lower during the dry season (less than 1 infective tick/animal), and increasing significantly during the wet seasons (2-4 infectives ticks/ animal). Interestingly, the lowest VIR observed during the dry season in lowlands and midlands was closely similar to the ones observed during the dry season and the two wet seasons in highlands. Is it worth noting that, slighter differences in the number of infective ticks were found between the early wet season (2 infective ticks/animal) and the late wet season (3-4 infective ticks/animal) in lowlands and midlands.

5.4.5 Agro-ecological and seasonal dynamics of *Theileria parva* infection in cattle

A total of 1424 cattle blood samples were screened to detect the presence of *T. parva* DNA. Of the 1424 DNA samples analysed, 485 (34%: CI: 32-37%) samples tested positive to *T. parva* (**Table 5.2**). There were statistically significant variations in the prevalence of *T. parva* infection in cattle across AEZs, with different seasonal patterns within AEZs (**Table 5.4**). The prevalence decreased significantly with increasing altitude: midlands (OR = 0.41 and $P = 0.01$) and highlands (OR = 0.15 and $P < 0.001$). The lowest proportion of positive samples was recorded during the dry season in highlands (10%), while there was markedly higher prevalence during the early wet season in lowlands and the late wet season in midlands (Table S5.2). The seasonal pattern of *T. parva* infection in cattle appeared to be very complex in lowlands and midlands (**Table S5.2**). In lowlands, the prevalence was slightly higher during the early wet season (53%) compared to the dry season (43%) which was not statistically different from the late wet season (48%). Besides, in midlands, the risk of infection was significantly higher during the late wet season (53%) compared to the early wet season (30%) and the dry season (24%). Strikingly, the prevalence observed during the dry season was not significantly different from that observed during the early wet season in this AEZ. In highlands, the *T. parva* prevalence was clearly different between the wet seasons and the dry season.

Table 5.4 Summary of the logistic regression model for the prevalence of *T. parva* infection in cattle according to AEZs and seasons

Factor	Level	Odds ratio (OR)	CI (95%)		*P-value*	DEFT[a]
AEZ	Lowlands	Ref[b]				
	Midlands	0.41	0.31	0.54	0.01	0.38
	Highlands	0.15	0.09	0.22	< 0.001	0.69
Season	Dry	Ref[b]				
	Wet1	1.5	0.91	2.4	0.063	1.2
	Wet2	1.2	0.55	2.7	0.35	1.8
Interaction	Midlands & Wet1	0.94	0.48	1.8	0.88	0.76
	Midlands & Wet2	2.8	1.22	6.6	0.017	0.93
	Highlands & Wet1	1.5	0.81	2.6	0.26	0.81
	Highlands & Wet2	2.4	0.87	6.4	0.046	1.1

Notes: AEZ, agro-ecological zones; Seasons: Wet1, early wet season (September-December); Wet2, late wet season (February-May); Dry, dry season (June-August). The model was multivariate, including AEZs (three levels), seasons (three levels) and their interaction. [a]DEFT, design effects computed using sample sites (villages) as subpopulation clusters assuming intra-village correlation. [b]Ref., reference level: lowlands and dry season were used as reference for comparison in the model (P=0.05). CI, confident intervals

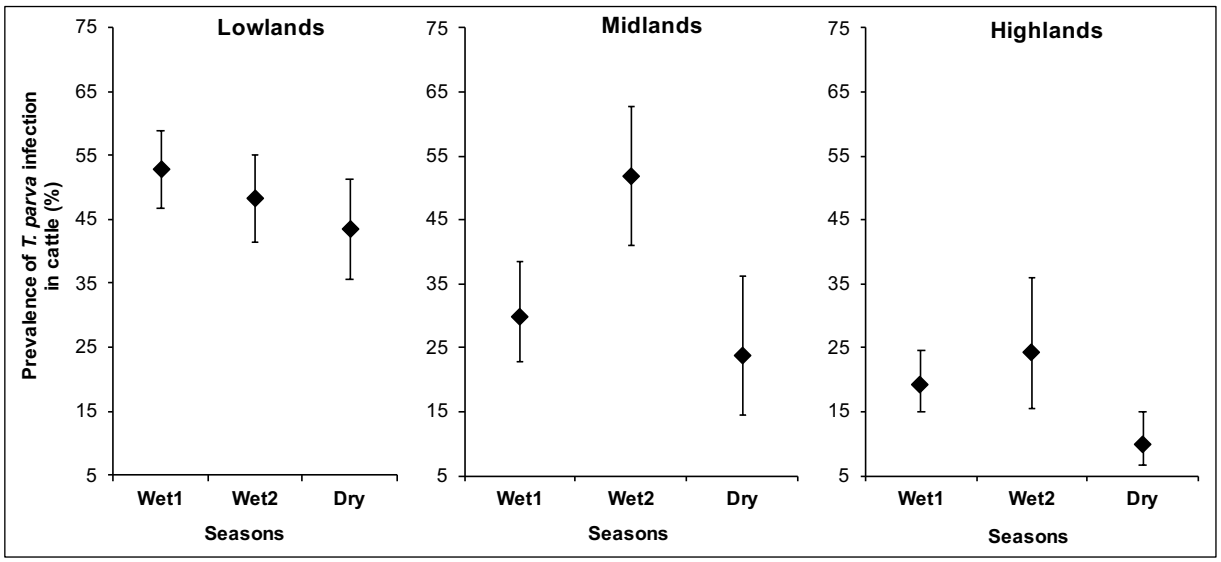

Figure 5.2 Agro-ecological and seasonal dynamics of the prevalence of *T. parva* infections in cattle. The graphs show the means and their 95% confidence intervals. *abbreviations*: AEZ, agro-ecological zones; Seasons: Wet1, early wet season (September-December); Wet2, late wet season (February-May); Dry, dry season (June-August)

5.5 Discussion

This study aimed to investigate the extent of agro-ecological and seasonal dynamics of *T. parva* transmission intensity and the distribution of its main tick vector *R. appendiculatus* in the South Kivu Province (eastern DRC). We conducted repeated cross-sectional surveys in three contrasting AEZs during the wet and dry seasons to determine the *R. appendiculatus* tick burden and the prevalence of *T. parva* infection in free-living ticks and in cattle. We then estimated the vectorial inoculation rate (VIR) based on tick burden and infection rate in ticks. These data provide useful information on tick population dynamics and *T. parva* transmission intensity to better understand the epidemiology of ECF in DRC for future rational and effective control strategies. The abundance of the tick and the transmission dynamics have a direct impact on the degree of endemicity to *T. parva* (Gachohi et al., 2012).

Rhipicephalus appendiculatus was the most common tick species recorded in South Kivu with a wide agro-ecological distribution range, accounting for 83% of the overall tick population collected on cattle. These findings are in general consistent with previous results reported in most neighbouring countries and regions, including Rwanda (Bazarusanga et al., 2007a), Nord Kivu in DRC (Kalume et al., 2013), Uganda (Byaruhanga et al., 2015), Tanzania (Hezron E et al., 2012; Laisser et al., 2014) and Burundi (Kaiser et al., 1988; Moran et al., 1996). The wide geographic distribution and the dominant nature of *R. appendiculatus* found in this study represents the ability of this species to colonise most equatorial conditions in the eastern and central Africa were climatic conditions are favourable to tick development and survival (Leta et al., 2013; Olwoch et al., 2003; Perry et al., 1990). *Rhipicephalus appendiculatus* tick burden was significantly affected by AEZs and seasons in South Kivu. It was lower in highlands and higher in lowlands and midlands AEZs. The environmental and livestock management conditions in these AEZs vary in their suitability for *R. appendiculatus*. In addition, migration of people with their cattle, seasonal transhumance, anthropic activities and climatic variability observed in the Great Lakes region may have facilitated the dispersal of the tick species and shaped its current agro-ecological distribution (Bazarusanga et al., 2007a; Byaruhanga et al., 2015; Kalume et al., 2013; Verweijen and Brabant, 2017). In the highlands, there is a montane tropical humid climate characterised by high rainfall but lower temperature. A large proportion of natural grazing lands are degraded and used for crop production due to high human demographic pressure. These factors may have led to reduced exposure of cattle to ticks. According to (Speybroeck et al., 2003), low temperatures compromise the survival capacity and delay the development of early instars of the tick and therefore result in reduced tick activity

on animals. In contrast, the midlands region provides more ecologically optimal conditions for the multiplication and survival of the tick as expected (warm and humid: medium temperature and substantial rainfall). The tick species is known to be most present and abundant at altitudes ranging between 1000-1800 m above the sea level (covering the midlands), where it has more successful reproduction cycles in warm and humid conditions (Gachohi et al., 2012; Rubaire-Akiiki et al., 2006; Yawa et al., 2018). Besides, the obvious high *R. appendiculatus* tick burden on cattle in lowlands could be a result of the establishment of the species in this region despite the relatively semi-arid characteristic of the lowlands AEZ at altitudes below 1200 m (warm and dry: rainfall varying between 800-1000 mm). There are a number of most likely explanations related to this finding: (i) The rainfall amounts which still fall within the range of suitable precipitations (500-2000 mm) for *R. appendiculatu*s and warm conditions prevailing in lowlands likely reflects optimal development and survival of the tick (Berkvens et al., 1998; Rubaire-Akiiki et al., 2006; Speybroeck et al., 2002; Walker et al., 2005); (ii) The lowlands area is subjected to increased cattle movements from neighbouring countries and within the country which could favour permanent connectivity of cattle carrying ticks from different epidemiological situations; (iii) The higher number of cattle concentrated in this area which is the main cattle production zone in the South Kivu Province could contribute to the high activity of ticks in the shared open communal grazing lands; (iv) The availability of vegetation where cattle graze during the dry season and the humid microclimate created along the Ruzizi River and the littoral fringe of Lake Tanganyika likely provide suitable ecological conditions for continuous tick activity. With regards to seasonal variation, *R. appendiculatus* ticks were found throughout the year. However, the tick infestation exhibited a clear seasonal variation in its abundance within all AEZs. The tick burden was lower during the dry season, but large numbers where found during the wet seasons. The lower number of ticks collected from cattle during the dry season naturally reflects the reduced rainfall and vegetation which affect the development, survival rate and behaviour of the tick. Interestingly, although there was lower number of ticks during the dry season, the fact that adult ticks were present on cattle during all seasons throughout the year suggests that *R. appendiculatus* has more than one generation with at least a bivoltine phenology. This is usually observed in the equatorial region of its distribution range (central and eastern Africa) where no synchronisation of the developmental stages is needed as opposed to South African areas which experience marked seasonal pattern of tick instars (Berkvens et al., 1998; Gachohi et al., 2012; Madder et al., 2002; Speybroeck et al., 2002).

The prevalence of *T. parva* infection was estimated in individual free-living ticks collected from the grazing areas. However, the statistical analysis did not provide significant differences of the infection rate neither across AEZs nor among seasons. The low and relatively constant prevalence of *T. parva* in ticks is consistent with other studies that used tick salivary gland staining and PCR methods to assess the infection in ticks from endemic regions (Gitau et al., 2000; Kimaro et al., 2017; Konnai et al., 2006; Ogden et al., 2003; Swai et al., 2006). There is evidence that *T. parva* infection rate in free-living ticks is usually expected to be low in endemic areas where most cattle are carrier (with low-level parasitaemia) and in areas where low prevalence prevails in cattle, which reduce the likelihood that most ticks acquire the infection (Medley et al., 1993; Ogden et al., 2003; Swai et al., 2006; Young et al., 1996). The vectorial inoculation rate (VIR) was much higher in lowlands and midlands than highlands, a situation correlated with the tick burden recorded in these respective AEZs. This is an indication of a stronger predictive power of tick burden to the estimates of *T. parva* transmission intensity. Analysis of seasonal fluctuation in VIR further support these findings showing that the dynamics of transmission intensity was greatly associated with tick burden rather than infection rate in ticks. The mean number of infective ticks was relatively constant throughout the year in highlands, while a significant seasonal pattern was found in lowlands and midlands. The suitability of ecological conditions which determine the extent of tick activity is believed to explain the observed transmission intensity of *T. parva*. Lower transmission may occur during seasons when climatic conditions are unfavourable to *R. appendiculatus* and the intensity of transmission increases with increased activity of adult ticks (Billiouw et al., 2002; Mulumba et al., 2001). Excepted from climatic conditions influencing the tick population dynamics, it could be possible that the observed differences in the transmission intensity and dynamics of *T. parva* are also explained by the genetic composition of *R. appendiculatus* from different ecological conditions, their behavioural diapause and variable levels of vector competence (Amzati et al., 2018; Mtambo et al., 2007b; Ochanda et al., 1998).

The overall prevalence of *T. parva* infection in cattle obtained in this study was generally similar to that recorded previously among indigenous cattle in different endemic areas of eastern and central African countries (Bazarusanga et al., 2007b; Byaruhanga et al., 2016; Kabi et al., 2014; Laisser et al., 2014; Muhanguzi et al., 2014; Tayebwa et al., 2018). The *T. parva* prevalence exhibited a significant variation among AEZs displaying a significant decrease in highlands and increased in lowlands and midlands, an indication of different levels of cattle exposure to *T. parva* infection (Deem et al., 1993; Gitau et al., 2000). These findings appear to

be correlated with the spatial and temporal variations in *R. appendiculatus* tick challenge and the transmission intensity of *T. parva* which were also significantly higher in lowlands and midlands than highlands. So, the differences observed in the extent and seasonal trend of *T. parva* prevalence in the three regions could be attributed to the variation in tick abundance due to changes in agro-ecological conditions and cattle densities (Magona et al., 2011b). However, a more complex seasonal trend of *T. parva* infection was observed in lowlands and midlands. Agro-ecological conditions prevailing in lowlands and highlands seems to favour increased activity of ticks which suggests an optimal interaction between hosts and ticks and subsequently a continuous transmission of *T. parva* (Bazarusanga et al., 2007b; Magona et al., 2011a; Olwoch et al., 2009). Under such conditions where cattle are continuously exposed to high tick challenge throughout the year, together with high transmission intensity and prevalence of *T. parva*, a scenario consistent with endemic status is expected, as reported in most endemic areas in Africa (Gachohi et al., 2012; Jonsson et al., 2012; Kimaro et al., 2017; Kivaria et al., 2004; Magona et al., 2008; Magona et al., 2011b; Rubaire-Akiiki et al., 2006). However, substantial movements that expose cattle to variable tick challenge, temporal variation in tick abundance (limiting early infection of calves) and exposure of cattle to more competence circulating tick stocks and heterologous *T. parva* strains are likely to disrupt the endemic situation and lead to unstable endemic or epidemic situations (Billiouw et al., 2002; De Deken et al., 2007; Latif et al., 2019; Rubaire-Akiiki et al., 2006; Yssouf et al., 2011). The situation observed in the highlands AEZ is totally different: low tick challenge, low and constant transmission intensity throughout the year and low prevalence among cattle. In such circumstances, as an implication to disease epidemiology, a large proportion of cattle are likely maintained susceptible and could be affected during the exposure to *T. parva* infection which usually generate strong epidemic situations (especially during seasons or years when tick abundance increases or moving to higher challenge areas) (Billiouw et al., 2002; Rubaire-Akiiki et al., 2004; Rubaire-Akiiki et al., 2006).

The tick population dynamics and *T. parva* transmission intensity are not the only factors that shape the epidemiological profile of ECF in the region. Other factors related to the genetic composition of parasite and vector populations are thought to bring different levels of interaction and complexity in the transmission system. In our earlier study, we reported an extensive variability of *T. parva* alleles in lowlands and midlands while only alleles identical to the components of the trivalent vaccine occur in highlands (Amzati et al., 2019). Besides, we found that two sympatric *R. appendiculatus* lineages occur in the region with an altitudinal

gradient of their coexistence and diversity. The south African lineage was more prevalent in lowlands and midlands whereas the east African lineage was ubiquitous in all AEZs (Amzati et al., 2018). These genetic discrepancies in *T. parva* and *R. appendiculatus* populations among AEZs, together with ecological conditions which define the differences in *T. parva* transmission dynamics and intensity and the movements of cattle undoubtedly play a crucial role in the complexity of ECF epidemiology reported in the Great Lakes region (Bazarusanga et al., 2007b; Kalume et al., 2012). Although the repeated cross-sectional approach applied in this study provided an overall picture of the epidemiological status of ECF in South Kivu, additional fine-scale longitudinal studies would be important to further assess data on the seroconversion rate, occurrence of clinical cases, mortality and case-fatality among cattle in different AEZs.

5.6 Conclusions

The findings presented in this study provide valuable information on ecological effects on the tick burden and the transmission intensity of *T. parva,* which shed light on the epidemiological profile of ECF in DRC. The exposure of cattle to *T. parva* was found to be strongly affected by the abundance of *R. appendiculatus* ticks and the transmission intensity of *T. parva*. *Rhipicephalus appendiculatus* ticks were present on cattle throughout the year and experience at least two generations. The agro-ecological and seasonal variations in the transmission intensity of *T. parva* was primarily predicted by the abundance of tick vector rather than the differences in tick infection rate among AEZs. The pattern of transmission intensity and dynamics of *T. parva* suggested an endemic situation in lowlands and midlands areas where there is high tick challenge whereas an epidemic situation is suggested in the highlands where tick activity and the *T. parva* infection in cattle are reduced. However, the movements of cattle together with the genetic composition of *R. appendiculatus* and *T. parva* seem to be major factors that are complexifying the epidemiological status of ECF leading to occasional epidemics or disruption in endemicity in lowlands.

Chapter 6. General discussion

6.1 Introduction

East coast fever caused by *Theileria parva* is the most pathogenic and economically important tick-borne disease of cattle in the Great Lakes region. The Great Lakes region of central Africa is characterised by diversified agro-ecological conditions and extensive cattle movements, where the epidemiology of ECF was previously reported to be complex and unstable, restricting the achievement of effective control measures (Amzati, 2011; Bazarusanga et al., 2011; Bazarusanga et al., 2007b; Kalume et al., 2012; Kalume et al., 2013). The control of ECF may rely on acaricide application to reduce tick infestation on cattle, the use of theilericidal drugs to treat seek animals, immunisation of cattle and the potential anti-tick vaccines under evaluation (Nene et al., 2016; Nene and Morrison, 2016; Olds et al., 2016; Parizi et al., 2020). The current immunisation method is based on the use of the Infection and treatment method (ITM) which involve inoculation of live *T. parva* sporozoites and simultaneous treatment with long-acting dose of oxytetracycline. The live vaccination approach provides strong immunity against homologous parasite strains, but limited cross-protection with heterologous strains. The efficacy and safety of live vaccine are limited due to the genetic diversity of *T. parva* in field populations and the risk of spreading "foreign" parasite strains in new areas (Hemmink et al., 2016; Kivaria et al., 2007; McKeever, 2007). Effective control strategies should be preferably designed in light of factors affecting the observed epidemic complexity of ECF in the Great Lakes as the disease is greatly variable among AEZs. The factors that are thought to explain the complexity may include the interplay between the biogeographical genetic composition of the tick vector and the pathogen and the transmission dynamics of *T. parva*. Predicting changes in the epidemiological landscape of vector-borne diseases should rely on an appropriate understanding of the biology, ecology and genetics of tick vector and pathogen populations, the vector-pathogen interface and their implications on the dynamics of the disease (Gooding, 1996; McCoy, 2008; Tabachnick and Black, 1995). These population genetic studies associated with the epidemiology of ECF can reveal and explain more complex epidemiological situations.

The overall aim of the studies presented in this thesis was to assess the population genetics of *R. appendiculatus* and *T. parva* and the transmission and infection dynamics of *T. parva*, as so to further understand the epidemic instability of ECF reported in different agro-ecological zones of the Great Lakes region and to assure the safety and efficacy of existing control measures. To achieve this objective, we examined specifically the intraspecific genetic variation and the phylogeography of *R. appendiculatus* populations (Chapter 3), the genetic and antigenic

variation of *T. parva* (Chapter 4) and the transmission intensity and dynamics of *T. parva* to define the epidemiological states of ECF in contrasting agro-ecological zones by means of vectorial inoculation rate (VIR) (Chapter 5). In the following sections, we will discuss the use of molecular markers and the major findings from our studies with focus on their evolutionary and ecological implications on the epidemiology of ECF.

6.2 Mitochondrial genes for genetic studies of *Rhipicephalus appendiculatus*

The genetic variation within and between tick species by the use of genetic tools and techniques usually provides clues that support the control diseases (Araya-Anchetta et al., 2015; McCoy, 2008; Sivakumar et al., 2014). During the past three decades, genetic markers have been widely used to estimate population genetic structure of ticks at the species level and taxonomic resolution of closely-related tick species (Araya-Anchetta et al., 2015). In the present study, we used mitochondrial COI and 12S rRNA markers to investigate the phylogenetic structure of *R. appendiculatus*. Mitochondrial genes undergo faster evolution due to their higher mutation rate compared to nuclear genes. Their maternal inheritance and clonal evolution support their usefulness for intraspecific evolutionary reconstruction (Shao and Barker, 2007). Our results showed that COI gene was more variable than 12S rRNA because of its high level of intraspecific variation, as previously described for *R. appendiculatus* (Kanduma et al., 2019; Kanduma et al., 2016a; Mtambo et al., 2007b; Mtambo et al., 2007c), *R. microplus* (Burger et al., 2014) and other Rhipicephalus species (Latrofa et al., 2013). Nevertheless, both mitochondrial genes resolved successfully and in congruence the phylogenetic substructuring of *R. appendiculatus*. In contrast, it has been reported previously that the nuclear second internal transcribed spacer (ITS2) of ribosomal DNA did not resolve any clear subgrouping of *R. appendiculatus* probably due to sexual recombination between the nuclear genomes within the species maintained at high level of sequence conservation (Kanduma et al., 2016a; Mtambo et al., 2007c), but was useful to distinguish closely-related species *R. appendiculatus* and *R. zambeziensis* (Mtambo et al., 2007a). Thus, mitochondrial genes were seen to be more powerful than nuclear markers for intraspecific discrimination of *R. appendiculatus*.

6.3 *Theileria parva* antigen-coding genes

The genetic variation within *T. parva* populations is considered to be one of the main survival strategies of this pathogen (Katzer et al., 2010). The availability of whole-genome sequence of

T. parva allowed generating useful molecular tools to discriminate the genetic and antigenic variation between and within *T. parva* populations (Gardner et al., 2005; Hayashida et al., 2013; Henson et al., 2012; Norling et al., 2015). These include variable number of tandem repeat (VNTR) markers (mini and micro-satellites) (Table S6.1) (Odongo et al., 2006; Oura et al., 2007; Oura et al., 2003; Rukambile et al., 2016) and antigen-coding genes that are targets of bovine MHC-I restricted CD8+ T cells (Graham et al., 2007; Graham et al., 2008). In this study, we used two antigen-coding loci that are immunodominant targets of bovine cytotoxic CD8+ T cells in cattle and are substantially polymorphic for typing the genetic composition *T. parva* populations (Table S6.2) (Hemmink et al., 2018; Kerario et al., 2019; Pelle et al., 2011; Salih et al., 2017). It has been shown that the immunodominance nature of the CD8+ T cell responses together with the antigenic variability in *T. parva* populations are major determinants of the phenomenon of parasite strain-specific immunity responses elicited by live vaccination (Connelley et al., 2011; MacHugh et al., 2009; Morrison et al., 2015). However, the direct role of the gene loci used in this study in the immunity provided by the live vaccination is not formally demonstrated to date, but it is believed that the efficacy of the vaccine is affected by the allelic variation and antigenic differences between vaccine and challenge strains (Pelle et al., 2011; Sitt et al., 2018; Sitt et al., 2015). The advantage of using the genes encoding antigens recognised by the CD8+ T cells is that these antigens have been identified as vaccine candidates for the development of subunit vaccine against *T. parva* (Svitek et al., 2018). The extent of their sequence variation is thought to constraint the efficacy of the vaccine (Hemmink et al., 2018; Nene et al., 2016). Furthermore, these gene loci provide an opportunity to study the nature and extent of the signature of selective forces driving the antigenic diversity and the diversity within epitope regions affecting the immune responses (Graham et al., 2007; Graham et al., 2008). Thus, the results of the two genes may give strong indication of potential vaccine outcomes as the success of the live vaccination partially depend on the antigenic variability of *T. parva* which can lead to parasite immune evasion.

6.4 Two divergent *Rhipicephalus appendiculatus* lineages occur in the Great Lakes region: geographical isolation or ecological preferences?

The phylogenetic analyses performed on mitochondrial sequences for *R. appendiculatus* specimens from the Great Lakes region together with published sequences revealed the presence of two genetic groups or “lineages” of the ticks occurring across its distribution range in the sub-Saharan Africa. The two groups are named lineages A and B in the present study,

corresponding to the "east African" and the "south African" stocks, respectively (Mtambo et al., 2007c). The analyses showed extensive polymorphism in individuals that belong to lineage A and low level of diversity within individuals of lineage B. However, they have no clear geographical sub-structuring in the Great Lakes region (sympatric distribution) but strong geographic separation in the southern Africa (allopatric distribution). The lineage A was composed of individuals mainly present in the equatorial region while the lineage B consisted of specimens present in the tropical region and unlike equatorial areas. The ecological and geographical distributions of the two lineages suggested that they have experienced independently different processes and historical dynamics that have shaped their structure. The analysis further revealed low level of genetic differentiation and population subdivision over geographical regions due to extensive gene flow leading to population admixture. A recent analysis of *R. appendiculatus* populations from different geographical areas in Kenya did not show any evidence of isolation by distance, suggesting that the genetic structure of this tick species could be shaped by dynamics other than geographic separation in the eastern African region (Kanduma et al., 2016a; Kanduma et al., 2016b). The evolutionary dynamics analyses revealed a signal of sudden demographic and spatial population expansion, suggesting that the observed genetic distribution is shaped by frequent dispersal of ticks though host mobility in the Great Lakes region, supported by the observed high degree of migration.

Earlier findings from east Africa suggested that one *R. appendiculatus* lineage may have diverged from the other (ancestral population) and experienced sudden population expansion after a bottleneck or genetic drift (Kanduma et al., 2016a). However, factors driving this divergence are unknown and which lineage is the ancestor is still speculative. We hypothesised that the divergence observed between the two *R. appendiculatus* lineages could have arisen following ancient geographical isolation due to founder events. Founder populations that initially belonged to lineage B could have experienced independent selective pressures from lineage A and subjected to genetic drift that shaped their current genetic composition. When, where and how the geographic isolation took place? does this separation represent ongoing speciation? Are there crossbreeding events between the two lineages? Do these lineages maintain their phenotypic features heritable to their offspring in new environments? These questions cannot be answered with the current data generated during this study. The observed current geographical distribution of tick lineages is shaped by cattle movements associated with their ecological preferences and establishment ability in different climatic conditions (Avise, 2000; Avise, 2012; Madder et al., 2002). The fact that the two lineages are sympatric in most

fringe ecological areas in central and eastern Africa show that they share some ecological preferences. The genetic grouping of *R. appendiculatus* stocks in Africa strongly correlated with phenotypic characteristics based on their body size, phenology, diapause behaviour and ecological distribution: the east African stock (lineage A) and the south African stock (lineage B). This support the existence of at least two distinct populations of *R. appendiculatus* in Africa that represent two divergent genetic groups (Leta et al., 2013). The east African stocks are smaller sized, characterised by at least a bivoltine adult phenology and the absence of diapause. The south African stocks are larger sized and have univoltine phenology and obligate diapause (Berkvens et al., 1998; Chaka et al., 1999; Madder et al., 2002; Speybroeck et al., 2004; Speybroeck et al., 2002). The larger body size of these south African stocks contributes to slower developmental and reproductive cycles, but provides strong ability to these ticks to survive harsh environmental conditions (hot and dry) while the smaller sized ticks have lower survival rates under unfavourable conditions. In sympatric areas, the "east African" stock may have faster reproductive ability giving rise to an evolutionary advantage in areas where the oviposition of the "south African" stocks could be delayed due to its obligatory diapause (Madder et al., 2002; Speybroeck et al., 2002).

6.5 Recent introduction of the "south African" *Rhipicephalus appendiculatus* lineage in the Great Lakes region through cattle movements

The south African lineage (tropical lineage) was previously absent from the Great Lakes region during a countrywide genetic study of *R. appendiculatus* conducted in Rwanda in 2007 (Mtambo et al., 2007c). It can be hypothesised that this lineage was recently introduced in the region through livestock movement, but the colonisation trajectory is currently difficult to determine. It could have inherited its diapause behaviour from its region of origin (southern Africa) and only developed the ability to survive in areas where there are pronounced and longer dry seasons like in the lowlands areas of Ruzizi valley (DRC), Imbo valley (Burundi) and the eastern lowlands of Rwanda. The process that reduced its colonisation ability in highlands AEZs needs to be further investigated. However, we can postulate that the reduced abundance of the obligatory diapausing lineage and its quasi-absence in the highlands may be explained by the longer wet season (abundant rainfall) in the highlands which does not give the opportunity to the tick to undergo diapause (Madder et al., 1999; Madder et al., 2002). The lowlands areas where most of specimens of lineage B were found is holding the majority of

cattle imported from neighbouring countries and is ecologically similar to the distribution zone of lineage B in southern Africa (Mtambo et al., 2007b; Mtambo et al., 2007c). In contrast, the east African *R. appendiculatus* lineage may have been established longer in the region since the first report of *R. appendiculatus* and *T. parva* within the Great Lake region (Mortelmans and Kageruka, 1986). The dominant and ubiquitous nature of the distribution of the east African lineage in the Great Lakes region could be an indication of well-established population of this tick in the region. It is known that genetic diversity and distribution generally differ between long-established and recently introduced populations of invasive ticks. Long-established populations are more genetically diverse than recently introduced populations (Nadolny et al., 2015). Although the two groups are sympatric in the Great Lakes region, they displayed contrasting genetic diversity patterns. This provides evidence that the two tick stocks have different life histories as they originated from various ecological conditions and established in the Great Lakes region.

6.6 Extensive genetic diversity among *Theileria parva* in the Great Lakes region

Agro-ecological variations driving tick population dynamics and anthropogenic movements of cattle were expected to be main factors affecting the genetic distribution of *T. parva* in the Great Lakes region. Previous studies on genetic characterisation of *T. parva* in field samples from various ecological areas of Africa revealed extensive diversity in *T. parva* populations, especially in buffalo-derived parasites (Elisa et al., 2015; Hemmink et al., 2018; Kerario et al., 2019; Pelle et al., 2011; Salih et al., 2017; Sitt et al., 2018). In the present study, a number of alleles were not identical to previously reported sequences, which demonstrate the extensive diversity within *T. parva* parasites from the Great Lakes region. *T. parva* populations from the Great Lakes region were extensively more diverse compared to those from various ecological zones in Africa and the majority of cattle-derived alleles circulating in Africa were found in DRC and Burundi (**Figure 6.1**). Interestingly, ubiquitous alleles and epitope variants identified in the two genes were identical or closely related to those found in the trivalent Muguga vaccine components. The evolutionary dynamic analyses did not provide the evidence of immune-based selection driving the observed diversity in the Great Lakes region, rather we proposed the demographic mechanisms such as range expansion of the parasite facilitated by the unidirectional importation of cattle from different countries and the ecological conditions driving tick population dynamics and competence (Estrada-Pena et al., 2009; Hemmink et al.,

2018; Pelle et al., 2011). Other mechanisms such as recombination during the sexual phase in the tick and random mutations could explain the generation of novel alleles (Sivakumar et al., 2014). It can be suggested that only the subset founder population of buffalo-derived *T. parva* established in cattle are circulating in different African countries through the dispersal of infected ticks and cattle. The coexistence of endemic alleles that are ubiquitous with newly introduced genetically distinct alleles might be leading to endemic instability. However, the extent of cross-protection between *T. parva* strains should be determined is such areas characterised by high gene flow. Phylogenetic and population structure analyses revealed a limited geographical separation of *T. parva* populations which was explained by the broad distribution of Muguga alleles associated with cattle movements in the region (Failly, 1999; Verweijen and Brabant, 2017). The cattle movements (gene flow) are maintaining similar ubiquitous cattle-derived *T. parva* alleles in different geographical and ecological areas in Africa, whereas unshared and less abundant alleles are restricted to parasite populations from specific AEZs and explain the significant genetic differentiation between lowlands and highlands areas.

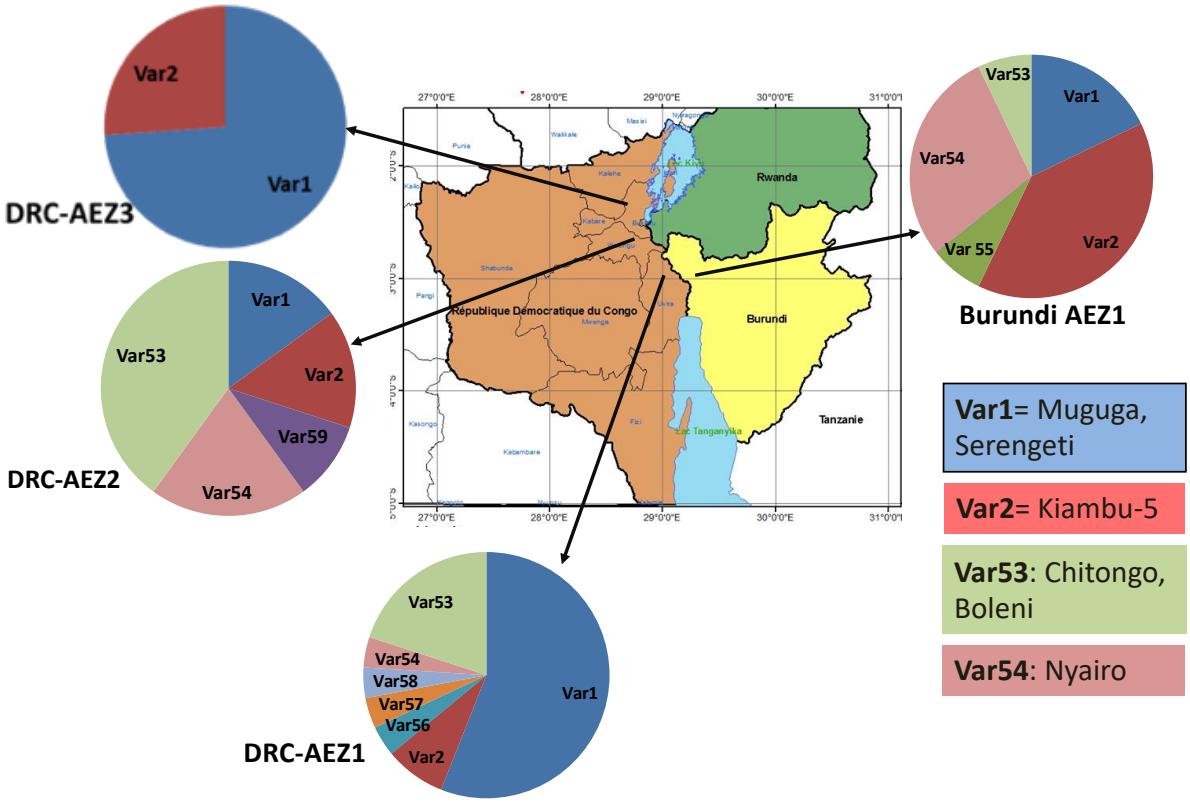

Figure 6.1 *Theileria parva* Tp2 antigenic diversity in the Great Lakes region

A surprising result was that an in-frame indel insertion of 12 nucleotides was found in four Tp1 cattle-derived sequences from DRC (allele-45/var-32) that were genetically the most distant from vaccine strains. This typical insertion in the Tp1 gene and its encoded protein was first reported as only present in buffalo-derived and buffalo-associated isolates (Pelle et al., 2011). However, the transmission of buffalo-derived *T. parva* within cattle by ticks is usually inefficient and cattle infected with these parasites develop a suddenly lethal disease that reduces their transmission sustainability within cattle population (absence of carrier state in cattle) (Bishop et al., 2015; Latif et al., 2019; Mbizeni et al., 2013; Sitt et al., 2015). In addition, cattle are not sharing pasture with African buffalo in the study area. Thus, the presence of these parasites in cattle could result either from the importation of animals carrying infected ticks that have acquired infection from African buffalo in other regions and transmitted the parasite to local cattle in DRC; or that the allele also evolves in cattle-derived parasites but was not described in previous studies in cattle from East Africa. Unfortunately, we were not able to monitor these infected cattle in order to make conclusive statements based on the case-fatality of the infection and other clinical features.

The Tp1 and Tp2 PCR assays used in this study have challenges that still need to be addressed to improve their performance. The sensitivity of these tests is still compromised in cases of mixed infections, especially in endemic areas. In addition, the sensitivity related to the level of parasitaemia and long-term infections is not known. It is known that the detection threshold of molecular essays varies in different field situations according to *T. parva* stocks and the level of parasitaemia (Geysen, 2008; Oura et al., 2004a; Oura et al., 2007). We commanded increasing the DNA template using the current primers to increase the probability of detecting *T. parva* DNA or design improved primers that are specific to different *T. parva* variants.

6.7 Transmission intensity and dynamics of *Theileria parva*

The seasonal dynamics and agro-ecological distribution of *R. appendiculatus* ticks and transmission of *T. parva* are of great epidemiological significance and have strong effect in the transmission of different *T. parva* variants. The infection rate in vector populations is a crucial parameter in assessing the transmission intensity of vector-borne diseases in order to calculate a proxy estimate of the incidence of the disease. In this study, we estimated the transmission intensity by means of the vectorial inoculation rate (VIR), expressing the number of infective ticks that attach on cattle during the feeding period of assumed one week. The p104 PCR-based method used in this study is more sensitive and has the advantage of being specific to detect *T.*

parva either in ticks or in carrier animals (Odongo et al., 2010; Odongo et al., 2009; Olds et al., 2018; Skilton et al., 2002). When a large number of ticks have to be analysed, detecting *T. parva* infection based on individual ticks is time consuming and a costly approach. The pooling approach used in this study seems to be a simple and practical approach to examine the infection rate in free-living ticks. However, the pooling must be random and the number of ticks per pool should allow successful detection of *T. parva* DNA in pools containing at least one infected tick (Katholi et al., 1995).

The differences in agro-ecological conditions were expected to determine the abundance of ticks and the number of generations which in turn are thought to reflect differences in infestation challenge and transmission intensity of *T. parva*. The objective was to assess the epidemiological landscape of ECF based on the level of tick challenge and the transmission intensity of *T. parva* in the three AEZs. Overall, we found that the transmission intensity was strongly predicted by the tick burden rather than the infection rate in ticks which was found to be constant across AEZs and seasons. According to (Norval et al., 1992), the abundance of feeding stages of ticks on cattle determines the level of *T. parva* transmission, increasing the number of infective ticks. The study revealed valuable findings in specific AEZs: in lowlands and midlands, we observed higher number of ticks per animal and higher transmission intensity, while lower tick challenge and lower and constant transmission intensity were found in highlands. The highlands AEZ is less ecological suitable for the survival of different instars of the tick due to the cooler conditions prevailing in this area together with reduced cattle densities whereas the lowlands and midlands area allows the strong occurrence of all instars of the tick throughout the year and provides favourable conditions (warm and humid) for *R. appendiculatus* and *T. parva* survival, facilitating the transmission (Gachohi et al., 2012; Rubaire-Akiiki et al., 2006; Yawa et al., 2018). The development and survival of free-living stages of *R. appendiculatus* are mainly favoured by humidity. The aridity and cooler conditions (lower temperature and dry conditions) compromise the development of earlier stages of the ticks reducing the establishment of tick populations on cattle (Norval et al., 1992; Speybroeck et al., 2003). A seasonal pattern in the transmission intensity was observed in lowlands and midlands, which was increased during the wet seasons. The increased rainfall and vegetation during these seasons favoured the exposure of cattle to ticks and consequently increased transmission intensity. In addition, although data were not collected on a monthly basis, the trends of tick abundance suggest at least two generations per year, a characteristic of the tick in

equatorial conditions (Berkvens et al., 1998; Gachohi et al., 2012; Madder et al., 2002; Speybroeck et al., 2002).

The prevalence of *T. parva* in cattle was higher in lowlands and midlands where the tick challenge and the transmission intensity were also high, especially during the late wet season. This shows that during the earlier wet season, animals are infested by adult ticks that have lost their *T. parva* infection during the dry season. In such conditions, the transmission of *T. parva* during the earlier wet season is obtained from carrier animals and the nymphal tick stage may play a crucial role in the transmission (Mulumba et al., 2001; Mulumba et al., 2000). However, it is difficult to provide clear relationships between the molecular prevalence in cattle and the tick challenge in a cross-sectional survey as the PCR methods usually detect current infections and have fluctuation of the sensitivity in detecting carrier infections (Geysen, 2008; Skilton et al., 2002). The observed pattern of transmission intensity and dynamics of *T. parva* suggested an endemic situation in lowlands and midlands areas where there is high tick challenge and high *T. parva* transmission throughout the year and an epidemic situation in the highlands characterised by low tick activity, low and constant transmission intensity (**Figure 6.2**). Since animals in highlands receive low tick challenge, a large proportion of susceptible animals may exist, which can easily trigger a strong epidemic.

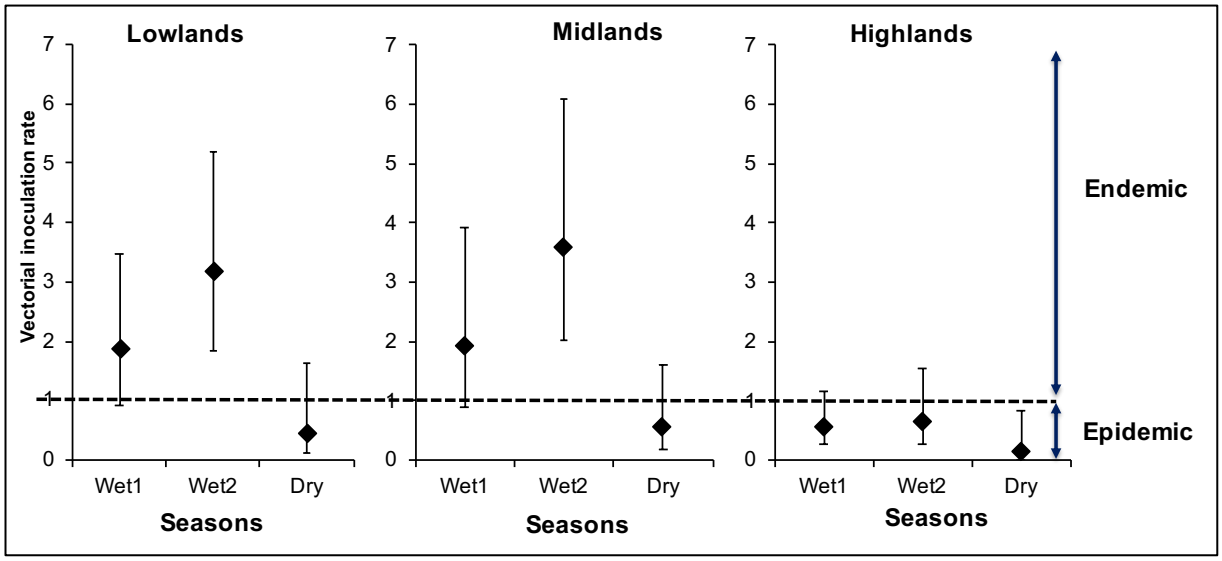

Figure 6.2 East Coast fever epidemiological states based on the vectorial inoculation rate in three agro-ecological zones of DRC

6.8 The genetic composition of *Rhipicephalus appendiculatus* and *Theileria parva* and the transmission dynamics

The evolutionary processes and historical dynamics of tick species are thought to have significant consequences on its ecological adaptability and epidemiological situations (Gooding, 1996; McCoy, 2008). The differences observed in the agro-geographical distribution and ecological preferences of *R. appendiculatus* lineages in the Great Lakes region suggest distinct epidemiological situations of ECF in different AEZs, because of the underlined differences in the vector phenotypes and phenologies (Berkvens et al., 1998; Chaka et al., 1999; Madder et al., 2002; Speybroeck et al., 2002). The geographical variation in the diapause behaviour and phenology of *R. appendiculatus* lineages may result in distinct vector competence (Kubasu, 1992; Ochanda et al., 1998; Odongo et al., 2009), survival and transmission efficiency of various *T. parva* strains in different AEZs, significantly affecting the epidemiology of ECF (Bishop et al., 2008). It has been established previously that the tick stock vector competence is heritable and that specific tick stocks are associated with specific *T. parva* strains in the transmission system (Norval et al., 1992; Ochanda et al., 1998; Tempia, 1997; Young et al., 1995). The biogeographical patterns of the genetic composition of *T. parva* suggested a direct association with ecological parameters driving tick population dynamics which affect the transmission dynamics of *T. parva,* and consequently the genetic structure of *T. parva* and the epidemiological landscape of ECF. The differences in bioclimatic conditions between AEZs (temperature, rainfall, vegetation and cattle population, cattle movements) are believed to underline substantially the ecological dynamics and activity of tick populations (Berkvens et al., 1998; Leta et al., 2013; Pegram and Banda, 1990; Speybroeck et al., 2002). In addition, the spreading ability of *R. appendiculatus* through cattle movement and the genetic distribution if its lineages seem to be altering the disease transmission system. It has been shown that tick population structure and heterogeneity determine parasite diversity (Katzer et al., 2006; McCoy, 2008). In the Great Lakes region, the genetic diversity of *R. appendiculatus* was strongly affected by the altitudes. It was lower in highlands and higher in lowlands and midlands. The east African lineage was present in all AEZs while the newly introduced south African lineage was confined to lowlands and midlands. The extensive genetic diversity of *T. parva* found in lowlands and midlands appears to be correlated with the high transmission intensity of *T. parva* in these areas, the increased number of ticks on cattle and the sympatric coexistence of the two *R. appendiculatus* lineages. The lowlands fringes areas which seem to be ecologically more suitable for the sympatric coexistence of the two lineages allow the

coexistence of tick lineages that have different phenotypes and consequently continuous transmission of *T. parva*. In contrast, the low level of genetic diversity of *T. parva* found in highlands was associated with reduced tick challenge on cattle and consequently, the low transmission intensity of *T. parva* observed in this area. These findings suggest that the genetic structure and diversity of *T. parva* are associated with the mobility of their hosts and the agro-ecological conditions affecting ticks established and competence. Although there was limited geographical substructuring between *R. appendiculatus* populations, the results showed that the occurrence and abundance of tick lineages are driven by agro-ecological and climatic factors facilitated by cattle movements. The factors affecting the tick population dynamics and sudden expansion are found to be major determinant of the epidemiological patterns of ECF in the region. These findings highlighted the strong relationships between the genetic composition of *R. appendiculatus* and *T. parva* and the dynamics of infective ticks.

Chapter 7. Conclusions and perspectives

7.1 Conclusions

This study has contributed substantially to improve our understanding of the relationships between the genetic and biogeographical distribution of *R. appendiculatus* and *T. parva* and the transmission dynamics of *T. parva*. In addition, factors affecting these variations and their nature have been proposed. These findings helped to understand the complex epidemiology of ECF in the Great Lakes region.

Rhipicephalus appendiculatus contains two major genetic groups in the Great Lakes region that differ in they colonisation patterns: lineage A (east African group) and lineage B (south African group). Their distribution in Africa correlated with marked differences in phenotypic features including behaviour diapause, vector competence and body size. The east African lineage (the more diverse and ubiquitous) is widely distributed and has been longer well-established in the Great Lakes region whereas the south African lineage (less abundant) has probably settled a founder population from recent colonisation events and its distribution decreases with altitude. The two genetic groups have been subjected to sudden demographic and spatial population expansion associated with cattle movements in Africa. The genetic divergence of *R. appendiculatus* lineages may have arisen from historical geographic isolation while the current structure is associated with their ecological preferences. *Rhipicephalus appendiculatus* ticks are more diverse in lowlands than highlands, with moderate genetic structure between the two ecosystems, while more genetic similitude is found in areas with same agro-ecological attributes, in spite of their geographical distance. The biogeographical distribution of *R. appendiculatus* suggests a sympatric coexistence in central and eastern Africa and allopatric distribution in south Africa.

Population genetic and phylogenetic analyses provided strong evidence of genetic similarity among ubiquitous *T. parva* genetic variants circulating in the Great Lakes region and the components of the Muguga trivalent vaccine. There was high degree of genetic variation within populations and limited agro-ecological structure due to the widespread major genotypes identical or closely similar to vaccine strains. *Theileria parva* populations from lowlands and midlands were more diverse than those from highlands areas. The majority of cattle-derived *T. parva* alleles circulating in Africa were found in the Great Lakes region, suggesting that *T. parva* populations are extensively more diverse in the Great Lakes region compared to those from various ecological zones in Africa. The genetic structure and biogeographical distribution of *T. parva* genetic variants were found to be strongly driven by gene flow facilitated by cattle

movement and ecological conditions driving tick population dynamics and phenotypes. Importantly, the fact that ubiquitous *T. parva* variants were genetically identical or closely related to the components of Muguga trivalent vaccine, together with the admixture of *T. parva* populations, partially provides evidence for safe use of existing trivalent live vaccine cross-protection field trials without any risk of introducing new parasite variants in the Great Lakes region.

The agro-ecological and seasonal variations in the transmission intensity of *T. parva* were primarily predicted by the abundance of tick vector rather than the differences in tick infection rate. *Rhipicephalus appendiculatus* ticks were present on cattle throughout the year and experience at least two generations in the Great Lakes region. The exposure of cattle to *T. parva* infection was found to be significantly affected by agro-ecological zones. The prevalence of *T. parva* infection in cattle, the tick challenge on cattle and the transmission intensity were significantly higher in lowlands and midlands while the highlands areas exhibited lower tick challenge and lower and constant transmission intensity.

These studies highlighted the strong relationships between the genetic composition of *R. appendiculatus* and *T. parva* and the transmission dynamics of *T. parva* that are complexifying the epidemiology of ECF in the region. Different epidemiological situations were then described in the contrasting agro-ecological zones: (i) In the highlands area, where high cattle mortality was previously reported, only *T. parva* isolates identical to the trivalent Muguga vaccine strains occur, associated with the strong presence of the east African tick lineage and low and constant transmission intensity of *T. parva*. In this area, the situation is suggested to be epidemic; (ii) In midlands and lowlands areas, where previous data reported that the disease was confined to calves, there was high genetic diversity of *T. parva* and *R. appendiculatus*. The two tick lineages are sympatric in these areas and there were high tick challenge and transmission intensity of *T. parva*. This is an indication of an endemic situation in these areas. Particularly, the strong coexistence of two *R. appendiculatus* lineages well adapted to the climatic conditions of the region, together with the continuous introduction of ticks through the extensive cattle movement may disrupt the endemicity and lead to occasional epidemics.

An integrated management of ECF including different control strategies should be implemented to be adapted in each epidemiological situation **(Table 7.1).** In endemic areas (lowlands and midlands), a control strategy using the ITM vaccine on the susceptible animals (calves and exotic) is more indicated. However, the age at first contact and the calving regime need to be

determined to facilitate the choice of optimal period for vaccination. In contrast, strategic tick control during epidemics, treatment of sick animals and a cost-effective vaccination of calves should be recommended in the epidemic area (highlands). In addition, tick control should be applied on cattle moving across agro-ecological zones to prevent the continuous introduction of new tick and parasite strains and reduce the risk of epidemics.

7.2 Perspectives

This study provided the evidence of interaction between the genetic composition of *R. appendiculatus* and *T. parva* and the transmission dynamics which can influence the epidemiology of ECF in the light of ecological conditions (**Table 7.1**). However, it remains to known which *T. parva* genotype is efficiently transmitted by which *R. appendiculatus* lineage. This question should be answered with research directed to assess the vector-pathogen genotype co-evolution and its impact in the transmission system and the spatiotemporal dynamics of *T. parva* genotypes. As demonstrated in this thesis, tick lineages are shifting and expanding their ecological niches. Thus, another important pending research question is to evaluate the heritability of phenotypic characteristics of *R. appendiculatus* in their new habitats and whether the identified two lineages are inbreeding or represent ongoing allopatric or sympatric speciation. These data are valuable to further understand the significance of specific vector-pathogen interaction in the epidemiology of ECF. In addition, although sequencing of mitochondrial genes was used successfully to characterise *R. appendiculatus* lineages, future research should be focused on the identification of tools and techniques (such as RLFP and morphological keys) that are cost-effective and able to easily discriminate the two lineages. This may help to further delineate the ecological and geographical ranges of the two lineages and to study their seasonal dynamics and phenology. On the other hand, an extensive genetic diversity of *T. parva* was found in the region. However, it remains relevant to investigate the relationships between the antigenic variation and the immunological variation of *T. parva* strains to be able to predict at a fine scale their impact in the epidemiology of ECF. Besides, the presence of *T. parva* variants identical to south African stocks (Chitongo and Boleni) that are more genetically distant from the components of Muguga trivalent vaccine remains of concern and their significance in the epidemiology and control of ECF should be further investigated in the ecological context of the Great Lakes region. The fact that majority of *T. parva* parasites circulating in the region were identical to the Muguga trivalent stocks alleviates the concerns of introducing exotic parasite strains through ITM vaccination. Nevertheless, testing the cross-protection with local isolates as challenge should be crucial to define the

composition of the IMT vaccine to be used in the region. A more recent experiment in controlled conditions demonstrated that immunisation with the Muguga strain alone provides similar level of protection compared to trivalent Muguga cocktail (Steinaa et al., 2018a), supporting the evidence of limited antigenic diversity in the three components of the Muguga cocktail vaccine (Hemmink et al., 2016). This raises the research question of testing different compositions of the IMT vaccine in field conditions. Although epidemiological situations were defined in different AEZs, their evolution can be expected to occur because of the reported climate change and changes in landscape use, cattle density and cattle movement trajectories. Thus, it should be of major interest to monitor these changes through a longitudinal study which may help to investigate the temporal distribution of *T. parva* and *R. appendiculatus* genotypes and their contribution to the incidence of ECF. Furthermore, a large-scale cattle blood sampling should be extended to Rwanda and other AEZs of Burundi to further capture the whole picture of *T. parva* strains circulating in these parts of the region. Finally, an ex-ante adoption study involving policymakers in charge of animal health and production, cattle farmers and other stakeholders will be relevant for the success of the deployment of live vaccine and other strategic control measures.

Table 7.1 Conclusions and perspectives

Studies	Conclusions		Perspectives
	Lowlands and midlands	**Highlands**	
Genetic of *R. appendiculatus*	Sympatric coexistence of tropical and equatorial lineages	Equatorial lineage, low diversity	• Characterisation of phenotypic differences and heritability in new environments • Crossbreeding (biological experiments) and panmixia (neutral markers) • Identification tools for the two lineages • Longitudinal survey (phenology and temporal genetic structure)
Genetic of *T. parva*	More diverse and population admixture	Less diverse, identical to vaccine stocks	• Antigenic variation vs immunological variation • *T. parva* strains /variants vs lethality • Temporal dynamics of genetic variants • Neutral markers (SSR) • Additional data for Rwanda
Transmission dynamics	High and seasonal transmission intensity	Low and relatively constant transmission	• Vector-pathogen coevolution and relationships in the transmission system • Longitudinal study: age at first contact • Additional data for Rwanda and other TTBDs
Epidemiological states	Endemic situation	Epidemic situation	• Impact of changing landscape and climate change • Longitudinal study: seroconversion and incidence
Control strategies	ITM Vaccine: susceptible population (calves and exotic cattle)	Strategic tick control, treatment of sick animals and vaccination of calves	• Cross-immunity vaccination trial: monovalent vs trivalent vaccine • Using existing trivalent vaccine or improved vaccine with local isolates? • Possible breakthroughs infection due to more variable *T. parva* variants?

References

Abascal, F., Zardoya, R. and Telford, M.J., 2010. TranslatorX: multiple alignment of nucleotide sequences guided by amino acid translations. Nucleic Acids Research 38: W7-13.

Abouelhassan, E.M., El-Gawady, H.M., Abdel-Aal, A.A., El-Gayar, A.K. and Esteve-Gassent, M.D., 2019. Comparison of Some Molecular Markers for Tick Species Identification. Journal of Arthropod-Borne Diseases 13: 153-164.

Akoolo, L., Pelle, R., Saya, R., Awino, E., Nyanjui, J., Taracha, E.L., Kanyari, P., Mwangi, D.M. and 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: 216-221.

Amzati, G.S., 2011. Enquête rétrospective sur la theilériose bovine au Sud-Kivu dans l'Est de la République Démocratique du Congo. Master thesis. Institute of Tropical Medicine, Antwerp, Belgium.

Amzati, G.S., Djikeng, A., Odongo, D.O., Nimpaye, H., Sibeko, K.P., Muhigwa, J.-B.B., Madder, M., Kirschvink, N. and Marcotty, T., 2019. Genetic and antigenic variation of the bovine tick-borne pathogen *Theileria parva* in the Great Lakes region of Central Africa. Parasites & Vectors 12: 588.

Amzati, G.S., Pelle, R., Muhigwa, J.B., Kanduma, E.G., Djikeng, A., Madder, M., Kirschvink, N. and Marcotty, T., 2018. Mitochondrial phylogeography and population structure of the cattle tick *Rhipicephalus appendiculatus* in the African Great Lakes region. Parasites & Vectors 11: 329.

Araya-Anchetta, A., Busch, J.D., Scoles, G.A. and Wagner, D.M., 2015. Thirty years of tick population genetics: a comprehensive review. Infection, Genetics and Evolution 29: 164-179.

Asiimwe, B.B., Weir, W., Tait, A., Lubega, G.W. and Oura, C.A., 2013. Haemoparasite infection kinetics and the population structure of *Theileria parva* on a single farm in Uganda. Veterinary Parasitology 193: 8-14.

Avise, J.C., 2000. Phylogeography: the history and formation of species. Harvard university press, USA.

Avise, J.C., 2012. Molecular Markers, Natural History and Evolution. Springer Science & Business Media, Chapman and Hall, New York.

Baliraine, F.N., Osir, E.O., Obuya, S.B. and Mulaa, F.J., 2000. Polymorphism in Two Populations of the Brown Ear Tick, *Rhipicephalus appendiculatus* Neumann (Acari: Ixodidae). International Journal of Tropical Insect Science 20: 227-231.

Bandelt, H.J., Forster, P. and Rohl, A., 1999. Median-joining networks for inferring intraspecific phylogenies. Molecular biology and evolution 16: 37-48.

Barker, S.C., 1998. Distinguishing species and populations of rhipicephaline ticks with its 2 ribosomal RNA. The Journal of parasitology 84: 887-892.

Barker, S.C. and Murrell, A., 2004. Systematics and evolution of ticks with a list of valid genus and species names. Parasitology 129: S15-36.

Baron, S., van der Merwe, N.A. and Maritz-Olivier, C., 2018. The genetic relationship between *R. microplus* and *R. decoloratus* ticks in South Africa and their population structure. Molecular Phylogenetics and Evolution 129: 60-69.

Barre, N. and Uilenberg, G., 2010. Spread of parasites transported with their hosts: case study of two species of cattle tick. Revue Scientifique et Technique 29: 135-147.

Bazarusanga, T., Geysen, D., Vercruysse, J. and Madder, M., 2007a. 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 & Applied Acarology 43: 279-291.

Bazarusanga, T., Marcotty, T., Ahou, A., Ntumba, T., Katendi, C. and Geysen, D., 2011. Estimation of the *Theileria parva* entomological inoculation rate (EIR) by means of tick burden and proportion of infected questing ticks in three different farming systems in Rwanda. International Journal of Vocational and Technical Education 3: 99-106.

Bazarusanga, T., Vercruysse, J., Marcotty, T. and Geysen, D., 2007b. Epidemiological studies on Theileriosis and the dynamics of *Theileria parva* infections in Rwanda. Veterinary Parasitology 143: 214-221.

Berkvens, D.L., 1991. Re-assessment of tick control after immunization against East Coast fever in the Eastern Province of Zambia. Annales de la Societe Belge de Medecine Tropicale 71: 87-94.

Berkvens, D.L., Geysen, D.M., Chaka, G., Madder, M. and Brandt, J.R., 1998. A survey of the ixodid ticks parasitising cattle in the Eastern province of Zambia. Medical and Veterinary Entomology 12: 234-240.

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

Bett, B., Lindahl, J. and Delia, G., 2019. Climate change and infectious livestock diseases: The case of Rift Valley fever and tick-borne diseases, In: Rosenstock, T.S., Nowak, A., Girvetz, E. (Eds.) The climate-smart agriculture papers: Investigating the business of a productive, resilient and low emission future. Springer International Publishing, Cham, pp. 29-37.

Biek, R. and Real, L.A., 2010. The landscape genetics of infectious disease emergence and spread. Molecular Ecology 19: 3515-3531.

Billiouw, M., Mataa, L., Marcotty, T., Chaka, G., Brandt, J. and Berkvens, D., 1999. The current epidemiological status of bovine theileriosis in eastern Zambia. Tropical Medicine & International Health 4: A28-A33.

Billiouw, M., Vercruysse, J., Marcotty, T., Speybroeck, N., Chaka, G. and Berkvens, D., 2002. *Theileria parva* epidemics: a case study in eastern Zambia. Veterinary Parasitology 107: 51-63.

Bishop, R., Geysen, D., Skilton, R., Odongo, D., Nene, V., Allsopp, B., Mbogo, S., Spooner, P. and Morzaria, S., 2002. Genomic polymorphism, sexual recombination and molecular epidemiology of *Theileria parva*. In: Dobbelaere D.A.E., McKeever D.J. (eds) *Theileria*. World Class Parasites, vol 3 pp. 23-39. Springer, Boston, MA

Bishop, R., Geysen, D., Spooner, P., Skilton, R., Nene, V., Dolan, T. and Morzaria, S., 2001. Molecular and immunological characterisation of *Theileria parva* stocks which are components of the 'Muguga cocktail' used for vaccination against East Coast fever in cattle. Veterinary Parasitology 94: 227-237.

Bishop, R., Musoke, A., Morzaria, S., Gardner, M. and Nene, V., 2004. *Theileria*: intracellular protozoan parasites of wild and domestic ruminants transmitted by ixodid ticks. Parasitology 129: S271-283.

Bishop, R., Musoke, A., Morzaria, S., Sohanpal, B. and Gobright, E., 1997. Concerted evolution at a multicopy locus in the protozoan parasite *Theileria parva*: extreme divergence of potential protein-coding sequences. Molecular and Cellular Biology 17: 1666-1673.

Bishop, R., Musoke, A., Skilton, R., Morzaria, S., Gardner, M. and Nene, V., 2008. *Theileria*: life cycle stages associated with the ixodid tick vector, In: Bowman, A.S., Nuttall, P. (Eds.) Biology, Disease and Control. Cambridge University Press, New York, pp. 308-324.

Bishop, R., Sohanpal, B., Kariuki, D.P., Young, A.S., Nene, V., Baylis, H., Allsopp, B.A., Spooner, P.R., Dolan, T.T. and Morzaria, S.P., 1992. Detection of a carrier state in *Theileria parva*-infected cattle by the polymerase chain reaction. Parasitology 104: 215-232.

Bishop, R.P., Hemmink, J.D., Morrison, W.I., Weir, W., Toye, P.G., Sitt, T., Spooner, P.R., Musoke, A.J., Skilton, R.A. and Odongo, D.O., 2015. The African buffalo parasite *Theileria*. sp. (buffalo) can infect and immortalize cattle leukocytes and encodes divergent orthologues of *Theileria parva* antigen genes. International Journal for Parasitology: Parasites and Wildlife 4: 333-342.

Bishop, R.P., Sohanpal, B.K., Allsopp, B.A., Spooner, P.R., Dolan, T.T. and Morzaria, S.P., 1993. Detection of polymorphisms among *Theileria parva* stocks using repetitive, telomeric and ribosomal DNA probes and anti-schizont monoclonal antibodies. Parasitology 107: 19-31.

Bishop, R.P., Spooner, P., Kanhai, G., Kiarie, J., Latif, A., Hove, T., Masaka, S. and Dolan, T., 1994. Molecular characterization of *Theileria* parasites: application to the epidemiology of theileriosis in Zimbabwe. Parasitology 109: 573-581.

Biswas, S. and Akey, J.M., 2006. Genomic insights into positive selection. Trends in Genetics 22: 437-446.

Boucher, F., Moutroifi, Y., Ali, M., Moindjie, Y., Soulé, M., Charafouddine, O., Cêtre-Sossah, C. and Cardinale, E., 2019. Impact of East Coast fever on Grande Comore: assessment taking a participatory epidemiology approach. Tropical Animal Health and Production 51: 99-107.

Boucher, F., Moutroifi, Y., Peba, B., Ali, M., Moindjie, Y., Ruget, A.S., Abdouroihamane, S., Madi Kassim, A., Soulé, M., Charafouddine, O., Cêtre-Sossah, C. and Cardinale, E., 2020. Tick-borne diseases in the Union of the Comoros are a hindrance to livestock development: Circulation and associated risk factors. Ticks and Tick-borne Diseases 11: 101283.

Boulinier, T., Kada, S., Ponchon, A., Dupraz, M., Dietrich, M., Gamble, A., Bourret, V., Duriez, O., Bazire, R., Tornos, J., Tveraa, T., Chambert, T., Garnier, R. and McCoy, K.D., 2016. Migration, Prospecting, Dispersal? What Host Movement Matters for Infectious Agent Circulation? Integrative and Comparative Biology 56: 330-342.

Bouslikhane M. Cross border movements of animals and animal products and their relevance to the epidemiology of animal diseases in africa. In: Technical item of the 21st

conference of the OIE Regional Commission for Africa: Rabat, Maroc, 16–20 February 2015.

Brabant, J., and Nzweve, J.K., 2013. La houe, la vache et le fusil—Conflits liés à la transhumance en territoires de Fizi et Uvira (Sud-Kivu, RDC): État des lieux et leçons tirées de l'expérience de LPI. Uppsala: Life and Peace Institute, Uppsala, Sweden.

Braverman, J.M., Hudson, R.R., Kaplan, N.L., Langley, C.H. and Stephan, W., 1995. The hitchhiking effect on the site frequency spectrum of DNA polymorphisms. Genetics 140: 783-796.

Burger, T.D., Shao and R., and Barker, S.C., 2014. Phylogenetic analysis of mitochondrial genome sequences indicates that the cattle tick, *Rhipicephalus (Boophilus) microplus*, contains a cryptic species. Molecular Phylogenetics and Evolution 76: 241-253.

Byaruhanga, C., Collins, N.E., Knobel, D., Chaisi, M.E., Vorster, I., Steyn, H.C. and Oosthuizen, M.C., 2016. Molecular investigation of tick-borne haemoparasite infections among transhumant zebu cattle in Karamoja Region, Uganda. Veterinary Parasitology: Regional Studies and Reports 3-4: 27-35.

Byaruhanga, C., Collins, N.E., Knobel, D., Kabasa, W. and Oosthuizen, M.C., 2015. Endemic status of tick-borne infections and tick species diversity among transhumant zebu cattle in Karamoja Region, Uganda: Support for control approaches. Veterinary Parasitology: Regional Studies and Reports 1-2: 21-30.

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 & International Health 4: A43-A48.

Coimbra-Dores, M.J., Maia-Silva, M., Marques, W., Oliveira, A.C., Rosa, F. and Dias, D., 2018. Phylogenetic insights on Mediterranean and Afrotropical *Rhipicephalus* species (Acari: Ixodida) based on mitochondrial DNA. Experimental & Applied Acarology 75: 107-128.

Connelley, T.K., MacHugh, N.D., Pelle, R., Weir, W. and Morrison, W.I., 2011. Escape from CD8+ T cell response by natural variants of an immunodominant epitope from *Theileria parva* is predominantly due to loss of TCR recognition. Journal of Immunology 187: 5910-5920.

Conrad, P.A., Iams, K., Brown, W.C., Sohanpal, B. and ole-MoiYoi, O.K., 1987. DNA probes detect genomic diversity in *Theileria parva* stocks. Molecular and Biochemical Parasitology 25: 213-226.

Conrad, P.A., ole-MoiYoi, O.K., Baldwin, C.L., Dolan, T.T., O'Callaghan, C.J., Njamunggeh, R.E., Grootenhuis, J.G., Stagg, D.A., Leitch, B.L. and Young, A.S., 1989. Characterization of buffalo-derived theilerial parasites with monoclonal antibodies and DNA probes. Parasitology 98: 179-188.

Cox, T.P., 2012. Farming the battlefield: the meanings of war, cattle and soil in South Kivu, Democratic Republic of the Congo. Disasters 36: 233-248.

Criscione, C.D., Poulin, R. and Blouin, M.S., 2005. Molecular ecology of parasites: elucidating ecological and microevolutionary processes. Molecular Ecology 14: 2247-2257.

Cristescu, M.E., 2015. Genetic reconstructions of invasion history. Molecular Ecology 24: 2212-2225.

Cruickshank, R.H., 2002. Molecular markers for the phylogenetics of mites and ticks. Systematic and Applied Acarology 7: 3-15.

Cumming, G.S., 2000. Using habitat models to map diversity: pan-African species richness of ticks (Acari: Ixodida). Journal of Biogeography 27: 425-440.

Dabert, M., 2006. DNA markers in the phylogenetics of the Acari. Biological Letters 43: 97-107.

Dantas-Torres, F., 2015. Climate change, biodiversity, ticks and tick-borne diseases: The butterfly effect. International Journal for Parasitology: Parasites and Wildlife 4: 452-461.

De Deken, R., Martin, V., Saido, A., Madder, M., Brandt, J. and Geysen, D., 2007. An outbreak of East Coast Fever on the Comoros: a consequence of the import of immunised cattle from Tanzania? Veterinary Parasitology 143: 245-253.

Deem, S.L., Perry, B.D., Katende, J.M., McDermott, J.J., Mahan, S.M., Maloo, S.H., Morzaria, S.P., Musoke, A.J. and Rowlands, G.J., 1993. Variations in prevalence rates of tick-borne diseases in Zebu cattle by agroecological zone: implications for East Coast fever immunization. Preventive Veterinary Medicine 16: 171-187.

Deitsch, K.W., Lukehart, S.A. and Stringer, J.R., 2009. Common strategies for antigenic variation by bacterial, fungal and protozoan pathogens. Nature Reviews Microbiology 7: 493-503.

Desiere, S., Niragira, S. and D'Haese, M., 2015. Cow or Goat? Population pressure and livestock keeping in Burundi. Agrekon 54: 23-42.

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: 85-92.

Dobbelaere, D. and Heussler, V., 1999. Transformation of leukocytes by *Theileria parva* and *T. annulata*. Annual Review of Microbiology 53: 1-42.

Dobbelaere, D.A., Shapiro, S.Z. and Webster, P., 1985. Identification of a surface antigen on *Theileria parva* sporozoites by monoclonal antibody. Proceedings of the National Academy of Sciences of the United States of America 82: 1771-1775.

Elisa, M., Gwakisa, P., Sibeko, K., Oosthuizen, M. and Geysen, D., 2014. Molecular characterization of *Theileria parva* field strains derived from cattle and buffalo sympatric populations of northern Tanzania. American Journal of Research Communication 2: 10-22.

Elisa, M., Hasan, S.D., Moses, N., Elpidius, R., Skilton, R. and Gwakisa, P., 2015. Genetic and antigenic diversity of *Theileria parva* in cattle in Eastern and Southern zones of Tanzania. A study to support control of East Coast fever. Parasitology 142: 698-705.

Ellegren, H. and Galtier, N., 2016. Determinants of genetic diversity. Nature reviews. Genetics 17: 422-433.

Emerson, B.C., Paradis, E. and Thébaud, C., 2001. Revealing the demographic histories of species using DNA sequences. Trends in Ecology and Evolution 16: 707-716.

Estrada-Pena, A., 2003. Climate change decreases habitat suitability for some tick species (Acari: Ixodidae) in South Africa. The Onderstepoort Journal of Veterinary Research 70: 79-93.

Estrada-Pena, A., 2008. Climate, niche, ticks, and models: what they are and how we should interpret them. Parasitology Research 103: S87-95.

Estrada-Pena, A., Naranjo, V., Acevedo-Whitehouse, K., Mangold, A.J., Kocan, K.M. and de la Fuente, J., 2009. Phylogeographic analysis reveals association of tick-borne pathogen, *Anaplasma marginale*, MSP1a sequences with ecological traits affecting tick vector performance. BMC Biology 7: 57.

Estrada-Peña, A. and Salman, M., 2013. Current Limitations in the Control and Spread of Ticks that Affect Livestock: A Review. Agriculture 3: 221-235

Excoffier, L., Foll, M. and Petit, R.J., 2009. Genetic Consequences of Range Expansions. Annual Review of Ecology, Evolution and Systematics 40: 481-501.

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 : 564-567.

Failly, D.d., 1999. L'économie du Sud-Kivu 1990–2000: mutations profondes cachées par une panne. In: Marysse S, Reyntjens F, editors. L'Afrique des Grands Lacs: annuaire 1999–2000. Paris: L'Harmattan; 2000. p. 163–92.

Fevre, E.M., Bronsvoort, B.M., Hamilton, K.A. and Cleaveland, S., 2006. Animal movements and the spread of infectious diseases. Trends in Microbiology 14: 125-131.

Fijarczyk, A. and Babik, W., 2015. Detecting balancing selection in genomes: limits and prospects. Molecular Ecology 24: 3529-3545.

Folmer, O., Black, M., Hoeh, W., Lutz, R. and 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.

Frankham, R., 1996. Relationship of genetic variation to population size in wildlife. Conservation Biology 10: 1500-1508.

Freeland, J.R., Petersen, S.D. and Kirk H., 2011. Molecular Ecology. 2nd ed. Chichester: Wiley.

Freeman, A.R., Hoggart, C.J., Hanotte, O. and Bradley, D.G., 2006. Assessing the relative ages of admixture in the bovine hybrid zones of Africa and the Near East using X chromosome haplotype mosaicism. Genetics 173: 1503-1510.

Fu, Y.X., 1997. Statistical tests of neutrality of mutations against population growth, hitchhiking and background selection. Genetics 147: 915-925.

Gachohi, J., Skilton, R., Hansen, F., Ngumi, P. and Kitala, P., 2012. Epidemiology of East Coast fever (*Theileria parva* infection) in Kenya: past, present and the future. Parasites & Vectors 5: 194.

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: 451-462.

Gardner, M.J., Bishop, R., Shah, T., de Villiers, E.P., Carlton, J.M., Hall, N., Ren, Q., Paulsen, I.T., Pain, A., Berriman, M., Wilson, R.J., Sato, S., Ralph, S.A., Mann, D.J., Xiong, Z., Shallom, S.J., Weidman, J., Jiang, L., Lynn, J., Weaver, B., Shoaibi, A., Domingo, A.R., Wasawo, D., Crabtree, J., Wortman, J.R., Haas, B., Angiuoli, S.V., Creasy, T.H., Lu, C., Suh, B., Silva, J.C., Utterback, T.R., Feldblyum, T.V., Pertea, M., Allen, J., Nierman, W.C., Taracha, E.L., Salzberg, S.L., White, O.R., Fitzhugh,

H.A., Morzaria, S., Venter, J.C., Fraser, C.M. and Nene, V., 2005. Genome sequence of *Theileria parva*, a bovine pathogen that transforms lymphocytes. Science 309: 134-137.

Gauer, M., Mackenstedt, U., Mehlhorn, H., Schein, E., Zapf, F., 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.

George, J.E., Pound, J.M. and Davey, R.B., 2004. Chemical control of ticks on cattle and the resistance of these parasites to acaricides. Parasitology 129: S353-366.

Geysen, D., 1999. Molecular epidemiology of *Theileria parva*: contribution of PCR to disease control. In: Bedreiging van insleep van ziekten uit de derdewereldlanden, detectiesystemen en economische implicaties; proceedings of the 7th Annual Meeting of the Flemish Society for Veterinary Epidemiology and Economics, 28 October 1999, Antwerp, Belgium.

Geysen, D., 2008. Live immunisation against *Theileria parva*: spreading the disease? Trends in Parasitology 24: 245-246.

Geysen, D., Bazarusanga, T., Brandt, J. and Dolan, T.T., 2004. An unusual mosaic structure of the PIM gene of *Theileria parva* and its relationship to allelic diversity. Molecular and Biochemical Parasitology 133: 163-173.

Geysen, D., Bishop, R., Skilton, R., Dolan, T.T. and Morzaria, S., 1999. Molecular epidemiology of *Theileria parva* in the field. Tropical Medicine & International Health 4: A21-27.

Gifford-Gonzalez, D. and Hanotte, O., 2011. Domesticating animals in Africa: Implications of genetic and archaeological findings. Journal of World Prehistory 24: 1-23.

Gilioli, G., Groppi, M., Vesperoni, M.P., Baumgärtner, J. and Gutierrez, A.P., 2009. An epidemiological model of East Coast Fever in African livestock. Ecological Modelling 220: 1652-1662.

Gitau, G.K., McDermott, J.J., Katende, J.M., O'Callaghan, C.J., Brown, R.N. and Perry, B.D., 2000. Differences in the epidemiology of theileriosis on smallholder dairy farms in contrasting agro-ecological and grazing strata of highland Kenya. Epidemiology and Infection 124: 325-335.

Githaka, N., Konnai, S., Bishop, R., Odongo, D., Lekolool, I., Kariuki, E., Gakuya, F., Kamau, L., Isezaki, M., Murata, S. and Ohashi, K., 2014. Identification and sequence characterization of novel *Theileria* genotypes from the waterbuck (*Kobus defassa*) in a *Theileria parva*-endemic area in Kenya. Veterinary Parasitology 202: 180-193.

Goddeeris, B.M., Morrison, W.I., Toye, P.G. and Bishop, R., 1990. Strain specificity of bovine *Theileria parva*-specific cytotoxic T cells is determined by the phenotype of the restricting class I MHC. Immunology 69: 38-44.

Gooding, R.H., 1996. Genetic variation in arthropod vectors of disease-causing organisms: obstacles and opportunities. Clinical Microbiology Reviews 9: 301-320.

Graham, S.P., Honda, Y., Pelle, R., Mwangi, D.M., Glew, E.J., de Villiers, E.P., Shah, T., Bishop, R., van der Bruggen, P., Nene, V. and Taracha, E.L., 2007. A novel strategy for the identification of antigens that are recognised by bovine MHC class I restricted cytotoxic T cells in a protozoan infection using reverse vaccinology. Immunome Research 3: 2.

Graham, S.P., Pelle, R., Honda, Y., Mwangi, D.M., Tonukari, N.J., Yamage, M., Glew, E.J., de Villiers, E.P., Shah, T., Bishop, R., Abuya, E., Awino, E., Gachanja, J., Luyai, A.E., Mbwika, F., Muthiani, A.M., Ndegwa, D.M., Njahira, M., Nyanjui, J.K., Onono, F.O., Osaso, J., Saya, R.M., Wildmann, C., Fraser, C.M., Maudlin, I., Gardner, M.J., Morzaria, S.P., Loosmore, S., Gilbert, S.C., Audonnet, J.C., van der Bruggen, P., Nene, V. and 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: 3286-3291.

Graham, S.P., Pelle, R., Yamage, M., Mwangi, D.M., Honda, Y., Mwakubambanya, R.S., de Villiers, E.P., Abuya, E., Awino, E., Gachanja, J., Mbwika, F., Muthiani, A.M., Muriuki, C., Nyanjui, J.K., Onono, F.O., Osaso, J., Riitho, V., Saya, R.M., Ellis, S.A., McKeever, D.J., MacHugh, N.D., Gilbert, S.C., Audonnet, J.C., Morrison, W.I., van der Bruggen, P. and Taracha, E.L., 2008. Characterization of the fine specificity of bovine CD8 T-cell responses to defined antigens from the protozoan parasite *Theileria parva*. Infection and Immunity 76: 685-694.

Guglielmone, A.A., Robbins, R.G., Apanaskevich, D.A., Petney, T.N., Estrada-Penã, A. and Horak, I.G., 2014. The hard ticks of the world (Acari: Ixodida: Ixodidae). Springer Science & Business Media Press, Dordrecht, Netherland.

Harpending, H.C., 1994. Signature of ancient population growth in a low-resolution mitochondrial DNA mismatch distribution. Human Biology 66 : 591-600.

Hatungumukama, G., Hornick, J.-L. and Detilleux, J., 2007. Aspects zootechniques de l'élevage bovin laitier au Burundi: Présent et futur. Annales de Médecine Vétérinaire 151 : 150-165.

Hayashida, K., Abe, T., Weir, W., Nakao, R., Ito, K., Kajino, K., Suzuki, Y., Jongejan, F., Geysen, D. and Sugimoto, C., 2013. Whole-genome sequencing of *Theileria parva* strains provides insight into parasite migration and diversification in the African continent. DNA Research 20: 209-220.

Hayashida, K., Hara, Y., Abe, T., Yamasaki, C., Toyoda, A., Kosuge, T., Suzuki, Y., Sato, Y., Kawashima, S., Katayama, T., Wakaguri, H., Inoue, N., Homma, K., Tada-Umezaki, M., Yagi, Y., Fujii, Y., Habara, T., Kanehisa, M., Watanabe, H., Ito, K., Gojobori, T., Sugawara, H., Imanishi, T., Weir, W., Gardner, M., Pain, A., Shiels, B., Hattori, M., Nene, V. and Sugimoto, C., 2012. Comparative genome analysis of three eukaryotic parasites with differing abilities to transform leukocytes reveals key mediators of *Theileria*-induced leukocyte transformation. MBio 3: e00204-00212.

Hemmink, J.D., Sitt, T., Pelle, R., de Klerk-Lorist, L.M., Shiels, B., Toye, P.G., Morrison, W.I. and Weir, W., 2018. Ancient diversity and geographical sub-structuring in African buffalo *Theileria parva* populations revealed through metagenetic analysis of antigen-encoding loci. International Journal for Parasitology 48: 287-296.

Hemmink, J.D., Weir, W., MacHugh, N.D., Graham, S.P., Patel, E., Paxton, E., Shiels, B., Toye, P.G., Morrison, W.I. and Pelle, R., 2016. Limited genetic and antigenic diversity within parasite isolates used in a live vaccine against *Theileria parva*. International Journal for Parasitology 46: 495-506.

Henson, S., Bishop, R.P., Morzaria, S., Spooner, P.R., Pelle, R., Poveda, L., Ebeling, M., Kung, E., Certa, U., Daubenberger, C.A. and Qi, W., 2012. High-resolution genotyping and mapping of recombination and gene conversion in the protozoan *Theileria parva* using whole genome sequencing. BMC Genomics 13: 503.

Herrero, M., Grace, D., Njuki, J., Johnson, N., Enahoro, D., Silvestri, S. and Rufino, M.C., 2013. The roles of livestock in developing countries. Animal 7: 3-18.

Hezron E, N., Adrian, M. and Robinson H, M., 2012. Tick infestations in extensively grazed cattle and efficacy trial of high-cis cypermethrin pour-on preparation for control of ticks in Mvomero district in Tanzania. BMC Veterinary Research 8: 224-224.

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

Iams, K.P., Young, J.R., Nene, V., Desai, J., Webster, P., ole-MoiYoi, O.K. and Musoke, A.J., 1990. Characterisation of the gene encoding a 104-kilodalton microneme-rhoptry protein of *Theileria parva*. Molecular and Biochemical Parasitology 39: 47-60.

Jalovecka, M., Hajdusek, O., Sojka, D., Kopacek, P. and Malandrin, L., 2018. The Complexity of Piroplasms Life Cycles. Frontiers in Cellular and Infection Microbiology 8: 248.

Jenicek, V. and Grofova, S., 2015. Least developed countries-the case of Burundi. Agricultural Economics 61: 234-247.

Jongejan, F. and Uilenberg, G., 2004. The global importance of ticks. Parasitology 129: S3-14.

Jonsson, N.N., Bock, R.E., Jorgensen, W.K., Morton, J.M. and Stear, M.J., 2012. Is endemic stability of tick-borne disease in cattle a useful concept? Trends in Parasitology 28: 85-89.

Kabi, F., Masembe, C., Muwanika, V., Kirunda, H. and Negrini, R., 2014. Geographic distribution of non-clinical *Theileria parva* infection among indigenous cattle populations in contrasting agro-ecological zones of Uganda: implications for control strategies. Parasites & Vectors 7: 414.

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. and Bourne, A.S., 1991. Tick (Acarina: Ixodidae) infestations on zebu cattle in northern Uganda. Bulletin of Entomological Research 81: 257-262.

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.

Kalume, M., Saegerman, C., Marcotty, T., Duchatel, J. and Losson, B., 2012. Statut épidémiologique de l'East Coast fever dans deux troupeaux de bovins issus de systèmes d'élevage distincts au Nord-Kivu, République démocratique du Congo. Annales de Médecine Vétérinaire 156 : 99-108.

Kalume, M.K., Saegerman, C., Mbahikyavolo, D.K., Makumyaviri, A.M.P., Marcotty, T., Madder, M., Caron, Y., Lempereur, L. and Losson, B., 2013. Identification of hard ticks (Acari: Ixodidae) and seroprevalence to *Theileria parva* in cattle raised in North Kivu Province, Democratic Republic of Congo. Parasitology Research 112: 789-797.

Kanduma, E.G., Bishop, R.P., Githaka, N.W., Skilton, R.A., Heyne, H. and Mwacharo, J.M., 2019. Mitochondrial and nuclear multilocus phylogeny of *Rhipicephalus* ticks from Kenya. Molecular Phylogenetics and Evolution 140: 106579.

Kanduma, E.G., Mwacharo, J.M., Githaka, N.W., Kinyanjui, P.W., Njuguna, J.N., Kamau, L.M., Kariuki, E., Mwaura, S., Skilton, R.A. and Bishop, R.P., 2016a. Analyses of mitochondrial genes reveal two sympatric but genetically divergent lineages of *Rhipicephalus appendiculatus* in Kenya. Parasites & Vectors 9: 353.

Kanduma, E.G., Mwacharo, J.M., Mwaura, S., Njuguna, J.N., Nzuki, I., Kinyanjui, P.W., Githaka, N., Heyne, H., Hanotte, O., Skilton, R.A. and Bishop, R.P., 2016b. Multi-locus genotyping reveals absence of genetic structure in field populations of the brown ear tick (*Rhipicephalus appendiculatus*) in Kenya. Ticks and Tick-Borne Diseases 7: 26-35.

Kariuki, D.P., Young, A.S., Morzaria, S.P., Lesan, A.C., Mining, S.K., Omwoyo, P., Wafula, J.L. and Molyneux, D.H., 1995. *Theileria parva* carrier state in naturally infected and artificially immunised cattle. Tropical Animal Health and Production 27: 15-25.

Katholi, C.R., Toé, L., Merriweather, A. and Unnasch, T.R., 1995. Determining the prevalence of *Onchocerca volvulus* infection in vector populations by Polymerase Chain Reaction screening of pools of black flies. The Journal of Infectious Diseases 172: 1414-1417.

Katzer, F., Lizundia, R., Ngugi, D., Blake, D. and McKeever, D., 2011. Construction of a genetic map for *Theileria parva*: identification of hotspots of recombination. International Journal for Parasitology 41: 669-675.

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

Katzer, F., Ngugi, D., Walker, A.R. and McKeever, D.J., 2010. Genotypic diversity, a survival strategy for the apicomplexan parasite *Theileria parva*. Veterinary Parasitology 167: 236-243.

Keirans, J.E., 1992. Systematics of the Ixodida (Argasidae, Ixodidae, Nuttalliellidae): An Overview and Some Problems, In: Fivaz, B., Petney, T., Horak, I. (Eds.) Tick Vector Biology: Medical and Veterinary Aspects. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 1-21.

Kerario, II, Chenyambuga, S.W., Mwega, E.D., Rukambile, E., Simulundu, E. and Simuunza, M.C., 2019. Diversity of two *Theileria parva* CD8+ antigens in cattle and buffalo-derived parasites in Tanzania. Ticks and Tick-Borne Diseases 10: 1003-1017.

Kerfua, S.D., Shirima, G., Kusiluka, L., Ayebazibwe, C., Mwebe, R., Cleaveland, S. and Haydon, D., 2018. Spatial and temporal distribution of foot-and-mouth disease in four districts situated along the Uganda-Tanzania border: Implications for cross-border efforts in disease control. The Onderstepoort Journal of Veterinary Research 85: 1-8.

Kibwana, D.K., Makumyaviri, A. and Hornick, J.-L., 2012. Pratiques d'élevage extensif et performances de bovins de race locale, et croisée avec des races laitières exotiques en République démocratique du Congo. Revue d'Elevage et de Médecine Vétérinaire des Pays Tropicaux 65 : 67-74.

Kimaro, E.G., Mor, S.M., Gwakisa, P. and Toribio, J.-A., 2017. Seasonal occurrence of *Theileria parva* infection and management practices amongst Maasai pastoralist communities in Monduli District, Northern Tanzania. Veterinary Parasitology 246: 43-52.

Kivaria, F.M., Heuer, C., Jongejan, F., Okello-Onen, J., Rutagwenda, T., Unger, F. and Boehle, W., 2004. Endemic stability for *Theileria parva* infections in Ankole calves of the Ankole ranching scheme, Uganda. The Onderstepoort Journal of Veterinary Research 71: 189-195.

Kivaria, F.M., Ruheta, M.R., Mkonyi, P.A. and Malamsha, P.C., 2007. Epidemiological aspects and economic impact of bovine theileriosis (East Coast fever) and its control: a preliminary assessment with special reference to Kibaha district, Tanzania. Veterinary Journal 173: 384-390.

Klapwijk, C.J., Bucagu, C., van Wijk, M.T., Udo, H.M.J., Vanlauwe, B., Munyanziza, E. and Giller, K.E., 2014. The 'One cow per poor family' programme: Current and potential fodder availability within smallholder farming systems in southwest Rwanda. Agricultural Systems 131: 11-22.

Klapwijk, C.J., Schut, M., van Asten, P.J.A., Vanlauwe, B., Giller, K.E. and Descheemaeker, K., 2019. Micro-livestock in smallholder farming systems: the role, challenges and opportunities for cavies in South Kivu, eastern DR Congo. Tropical Animal Health and Production 52: 1167-1177

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: 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: 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.

Kubasu, S.S., Makokkah, G.L. and Kaaya, G., 2007. Biological differences within *Rhipicephalus appendiculatus* Neumann (Acari: Ixodidae populations in Kenya. Journal of the Egyptian Society of Parasitology 37: 411-418.

Kumar, S., Stecher, G. and Tamura, K., 2016. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. Molecular Biology and Evolution 33: 1870-1874.

Lacasta, A., Mwalimu, S., Kibwana, E., Saya, R., Awino, E., Njoroge, T., Poole, J., Ndiwa, N., Pelle, R., Nene, V. and Steinaa, L., 2018. Immune parameters to p67C antigen adjuvanted with ISA206VG correlate with protection against East Coast fever. Vaccine 36: 1389-1397.

Lack, J.B., Reichard, M.V. and Van Den Bussche, R.A., 2012. Phylogeny and evolution of the Piroplasmida as inferred from 18S rRNA sequences. International Journal for Parasitology 42: 353-363.

Laisser, E.L.K., Chenyambuga, S.W., Karimuribo, E.D., Msalya, G., Kipanyula, M.J., Mwilawa, A.J., Mdegela, R.H. and Kusiluka, L.J.M., 2017. A review on prevalence, control measure, and tolerance of Tanzania Shorthorn Zebu cattle to East Coast fever in Tanzania. Tropical Animal Health and Production 49: 813-822.

Laisser, E.L.K., Kipanyula, M.J., Msalya, G., Mdegela, R.H., Karimuribo, E.D., Mwilawa, A.J., Mwega, E.D., Kusiluka, L. and Chenyambuga, S.W., 2014. Tick burden and

prevalence of *Theileria parva* infection in Tarime zebu cattle in the lake zone of Tanzania. Tropical Animal Health and Production 46: 1391-1396.

Latif, A.A. and Hove, T., 2011. History and critical review of *Theileria parva* (Boleni), the vaccine stock against Zimbabwean cattle theileriosis. Ticks and Tick-Borne Diseases 2: 163-167.

Latif, A.A., Hove, T., Kanhai, G.K. and Masaka, S., 2002. Buffalo-associated *Theileria parva*: the risk to cattle of buffalo translocation into the Highveld of Zimbabwe. Annals of the New York Academy of Sciences 969: 275-279.

Latif, A.A., Punyua, D.K., Capstick, P.B. and Newson, R.M., 1991. Tick infestations on Zebu cattle in western Kenya: host resistance to *Rhipicephalus appendiculatus* (Acari: Ixodidae). Journal of Medical Entomology 28: 127-132.

Latif, A.A., Troskie, P.C., Peba, S.B., Maboko, B.B., Pienaar, R. and Mans, B.J., 2019. Corridor disease (buffalo-associated *Theileria parva*) outbreak in cattle introduced onto a game ranch and investigations into their carrier-state. Veterinary Parasitology: Regional Studies and Reports 18: 100331.

Latrofa, M.S., Dantas-Torres, F., Annoscia, G., Cantacessi, C. and Otranto, D., 2013. Comparative analyses of mitochondrial and nuclear genetic markers for the molecular identification of *Rhipicephalus* spp. Infection, Genetics and Evolution 20: 422-427.

Le Roux, J. and Wieczorek, A.M., 2009. Molecular systematics and population genetics of biological invasions: towards a better understanding of invasive species management. Annals of Applied Biology 154: 1-17.

Leger, E., Vourc'h, G., Vial, L., Chevillon, C. and McCoy, K.D., 2013. Changing distributions of ticks: causes and consequences. Experimental & Applied Acarology 59: 219-244.

Leigh, J.W. and Bryant, D., 2015. popart: full-feature software for haplotype network construction. Methods in Ecology and Evolution 6: 1110-1116.

Lempereur, L., Geysen, D. and Madder, M., 2010. Development and validation of a PCR-RFLP test to identify African *Rhipicephalus (Boophilus)* ticks. Acta Tropica 114: 55-58.

Leo, S.S., Gonzalez, A. and Millien, V., 2017. The genetic signature of range expansion in a disease vector-the black-legged tick. The Journal of Heredity 108: 176-183.

Lessard, P., L'Eplattenier, R., Norval, R.A., Kundert, K., Dolan, T.T., Croze, H., Walker, J.B., Irvin, A.D. and Perry, B.D., 1990. Geographical information systems for studying the epidemiology of cattle diseases caused by *Theileria parva*. The Veterinary Record 126: 255-262.

Leta, S., De Clercq, E.M. and Madder, M., 2013. High-resolution predictive mapping for *Rhipicephalus appendiculatus* (Acari: Ixodidae) in the Horn of Africa. Experimental & Applied Acarology 60: 531-542.

Librado, P. and Rozas, J., 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. Bioinformatics 25: 1451-1452.

Lv, J., Wu, S., Zhang, Y., Chen, Y., Feng, C., Yuan, X., Jia, G., Deng, J., Wang, C., Wang, Q., Mei, L. and Lin, X., 2014a. Assessment of four DNA fragments (COI, 16S rDNA, ITS2, 12S rDNA) for species identification of the Ixodida (Acari: Ixodida). Parasites & Vectors 7: 93.

Lv, J., Wu, S., Zhang, Y., Zhang, T., Feng, C., Jia, G. and Lin, X., 2014b. Development of a DNA barcoding system for the Ixodida (Acari: Ixodida). Mitochondrial DNA 25: 142-149.

Maass, B.L., Katunga Musale, D., Chiuri, W.L., Gassner, A. and Peters, M., 2012. Challenges and opportunities for smallholder livestock production in post-conflict South Kivu, eastern DR Congo. Tropical Animal Health and Production 44: 1221-1232.

MacHugh, N.D., Connelley, T., Graham, S.P., Pelle, R., Formisano, P., Taracha, E.L., Ellis, S.A., McKeever, D.J., Burrells, A. and Morrison, W.I., 2009. CD8+ T-cell responses to *Theileria parva* are preferentially directed to a single dominant antigen: Implications for parasite strain-specific immunity. European Journal of Immunology 39: 2459-2469.

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

Madder, M., Speybroeck, N., Brandt, J., Tirry, L., Hodek, I. and Berkvens, D., 2002. Geographic variation in diapause response of adult *Rhipicephalus appendiculatus* ticks. Experimental & Applied Acarology 27: 209-221.

Magona, J.W., Walubengo, J. and Kabi, F., 2011a. Response of Nkedi Zebu and Ankole cattle to tick infestation and natural tick-borne, helminth and trypanosome infections in Uganda. Tropical Animal Health Production 43: 1019-1033.

Magona, J.W., Walubengo, J., Olaho-Mukani, W., Jonsson, N.N., Welburn, S.C. and Eisler, M.C., 2008. Clinical features associated with seroconversion to *Anaplasma marginale, Babesia bigemina* and *Theileria parva* infections in African cattle under natural tick challenge. Veterinary Parasitology 155: 273-280.

Magona, J.W., Walubengo, J., Olaho-Mukani, W., Jonsson, N.N., Welburn, S.W. and Eisler, M.C., 2011b. Spatial variation of tick abundance and seroconversion rates of indigenous cattle to *Anaplasma marginale, Babesia bigemina* and *Theileria parva* infections in Uganda. Experimental & Applied Acarology 55: 203-213.

Male, D., Brostoff, J., Roth, D. and Roitt, I., 2007. Immunologie: Campus références: Paris, Elsevier Masson.

Manel, S. and Holderegger, R., 2013. Ten years of landscape genetics. Trends in Ecology & Evolution 28: 614-621.

Manirakiza, J., Hatungumukama, G., Thévenon, S., Gautier, M., Besbes, B., Flori, L. and Detilleux, J., 2017. Effect of genetic European taurine ancestry on milk yield of Ankole-Holstein crossbred dairy cattle in mixed smallholders system of Burundi highlands. Animal Genetics 48: 544-550.

Mans, B.J. and Neitz, A.W., 2004. Adaptation of ticks to a blood-feeding environment: evolution from a functional perspective. Insect Biochemistry and Molecular Biology 34 : 1-17.

Mararo, S.B., 2000. Pouvoirs, élevage bovin et la question foncière au Nord-Kivu. In: Marysse S, Reyntjens F, editors. L'Afrique des Grands Lacs: annuaire 2000– 2001. Paris: L'Harmattan; 2001. p. 219–250.

Marcellino, W.L., Salih, D.A., Njahira, M.N., Ndiwa, N., Araba, A., El Hussein, A.M., Seitzer, U., Ahmed, J.S., Bishop, R.P. and Skilton, R.A., 2017. The Emergence of

Theileria parva in Jonglei State, South Sudan: Confirmation Using Molecular and Serological Diagnostic Tools. Transboundary and Emerging Diseases 64: 1229-1235.

Marcotty, T., Brandt, J., Billiouw, M., Chaka, G., Losson, B. and 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., Young, A.S., Lesan, A.C., Ndungu, S.G., Stagg, D.A. and Ngumi, P.N., 1992. Transformation of *Theileria parva* derived from African buffalo (*Syncerus caffer*) by tick passage in cattle and its use in infection and treatment immunization. Veterinary Parasitology 43: 1-14.

Maze-Guilmo, E., Blanchet, S., McCoy, K.D. and Loot, G., 2016. Host dispersal as the driver of parasite genetic structure: a paradigm lost? Ecology Letters 19: 336-347.

Mazimpaka, E., Mbuza, F., Michael, T., Gatari, E.N., Bukenya, E.M. and James, O.-A., 2017. Current status of cattle production system in Nyagatare District-Rwanda. Tropical Animal Health and Production 49: 1645-1656.

Mbassa, G.K., Kipanyula, M.J., Mellau, L.S., Mwamakali, E.D., Bulegeya, F.R. and Kauto-Mboni, K., 2006. *Theileria parva* infection in calves causes massive lymphocyte death in the thymus, spleen and lymph nodes without initial proliferation. Veterinary Parasitology 142: 260-270.

Mbizeni, S., Potgieter, F.T., Troskie, C., Mans, B.J., Penzhorn, B.L. and Latif, A.A., 2013. Field and laboratory studies on Corridor disease (*Theileria parva* infection) in cattle population at the livestock/game interface of uPhongolo-Mkuze area, South Africa. Ticks and Tick-Borne Diseases 4: 227-234.

Mbogo, S.K., Kariuki, D.P., Ngumi, P.N. and McHardy, N., 1996. A mild *Theileria parva* parasite with potential for immunisation against East Coast fever. Veterinary Parasitology 61: 41-47.

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

McHardy, N., Wekesa, L.S., Hudson, A.T. and Randall, A.W., 1985. Antitheilerial activity of BW720C (buparvaquone): a comparison with parvaquone. Research in Veterinary Science 39: 29-33.

McKeever, D.J., 2001. Cellular immunity against *Theileria parva* and its influence on parasite diversity. Research in Veterinary Science 70: 77-81.

McKeever, D.J., 2006. *Theileria parva* and the bovine CTL response: down but not out? Parasite Immunology 28: 339-345.

McKeever, D.J., 2007. Live immunisation against *Theileria parva*: containing or spreading the disease? Trends in Parasitology 23: 565-568.

McKeever, D.J., 2009. Bovine immunity - a driver for diversity in *Theileria* parasites? Trends in Parasitologgy 25: 269-276.

McKeever, D.J., Taracha, E.L., Innes, E.L., MacHugh, N.D., Awino, E., Goddeeris, B.M. and 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 of the United States of America 91: 1959-1963.

McKeever, D.J., Taracha, E.L., Morrison, W.I., Musoke, A.J. and Morzaria, S.P., 1999. Protective immune mechanisms against *Theileria parva*: evolution of vaccine development strategies. Parasitology Today 15: 263-267.

Medley, G.F., Perry, B.D. and Young, A.S., 1993. Preliminary analysis of the transmission dynamics of *Theileria parva* in eastern Africa. Parasitology 106: 251-264.

Meirmans, P.G. and Hedrick, P.W., 2011. Assessing population structure: FST and related measures. Molecular Ecology Resources 11: 5-18.

Minami, T., Spooner, P.R., Irvin, A.D., Ocama, J.G., Dobbelaere, D.A. and Fujinaga, T., 1983. Characterisation of stocks of *Theileria parva* by monoclonal antibody profiles. Research in Veterinary Science 35: 334-340.

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 and eastern and southern Africa. Research report, DFID Animal Health Programme. Edinburg: Centre for Tropical Veterinary Medicine, University of Edinburgh, UK.

Minshull, J.I. and Norval, N., 1982. Factors influencing the spatial distribution of *Rhipicephalus appendiculatus* in Kyle Recreational Park, Zimbabwe. South African Journal of Wildlife Research 12: 118-123.

Mitchell, S. and Pagano, M., 2012. Pooled testing for effective estimation of the prevalence of *Schistosoma mansoni*. The American Journal of Tropical Medicine and Hygiene 87: 850-861.

Moran, M.C., Nigarura, G. and Pegram, R.G., 1996. An assessment of host resistance to ticks on cross-bred cattle in Burundi. Medical and Veterinary Entomology 10: 12-18.

Morrison, D.A., 2009. Evolution of the Apicomplexa: where are we now? Trends in Parasitology 25: 375-382.

Morrison, W.I., 2007. The biological and practical significance of antigenic variability in protective T cell responses against *Theileria parva*. Veterinary Parasitology 148: 21-30.

Morrison, W.I., 2015. The aetiology, pathogenesis and control of theileriosis in domestic animals. Revue Scientifique et Technique 34: 599-611.

Morrison, W.I., Connelley, T., Hemmink, J.D. and MacHugh, N.D., 2015. Understanding the basis of parasite strain-restricted immunity to *Theileria parva*. Annual Review of Animal Biosciences 3: 397-418.

Morrison, W.I. and Goddeeris, B.M., 1990. Cytotoxic T cells in immunity to *Theileria parva* in cattle. Current Topics in Microbiology and Immunology 155: 79-93.

Morrison, W.I., Goddeeris, B.M. and Teale, A.J., 1987. Bovine cytotoxic T cell clones which recognize lymphoblasts infected with two antigenically different stocks of the protozoan parasite *Theileria parva*. European Journal of Immunology 17: 1703-1709.

Morrison, W.I., MacHugh, N.D. and Lalor, P.A., 1996. Pathogenicity of *Theileria parva* is influenced by the host cell type infected by the parasite. Infection and Immunology 64: 557-562.

Morrison, W.I. and McKeever, D.J., 2006. Current status of vaccine development against *Theileria* parasites. Parasitology 133: S169-187.

Morrison, W.I., Taracha, E.L.N. and McKeever, D.J., 1995. Contribution of T-cell responses to immunity and pathogenesis in infections with *Theileria parva*. Parasitology Today 11: 14-18.

Mortelmans, J. and Kageruka, P., 1986. L'histoire de 80 années d'observations et de recherches sur les theilerioses au Zaïre, Rwanda et Burundi. Annales de la Société Belge de Médecine Tropicale 66: 199-212.

Mtambo, J., 2008. *Rhipicephalus appendiculatus/zambeziensis* complex from southern and eastern Zambia: genetic and phenotypic diversity related to the associated variation of the epidemiology of bovine theileriosis. PhD Thesis. Ghent University, Belgium.

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 & Applied Acarology 41: 115-128.

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

Mtambo, J., Madder, M., Van Bortel, W., Geysen, D., Berkvens, D. and Backeljau, T., 2007c. 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: 168-175.

Muhanguzi, D., Picozzi, K., Hatendorf, J., Thrusfield, M., Welburn, S.C., Kabasa, J.D. and Waiswa, C., 2014. Prevalence and spatial distribution of *Theileria parva* in cattle under crop-livestock farming systems in Tororo District, Eastern Uganda. Parasites & Vectors 7: 91.

Mukhebi, A.W., Perry, B.D. and Kruska, R., 1992. Estimated economics of theileriosis control in Africa. Preventive Veterinary Medicine 12: 73-85.

Muleya, W., Namangala, B., Simuunza, M., Nakao, R., Inoue, N., Kimura, T., Ito, K., Sugimoto, C. and Sawa, H., 2012. Population genetic analysis and sub-structuring of *Theileria parva* in the northern and eastern parts of Zambia. Parasites & Vectors 5: 255.

Mulumba, M., Speybroeck, N., Berkvens, D.L., Geysen, D.M. and Brandt, J.R., 2001. Transmission of *Theileria parva* in the traditional farming sector in the Southern Province of Zambia during 1997-1998. Tropical Animal Health and Production 33: 117-125.

Mulumba, M., Speybroeck, N., Billiouw, M., Berkvens, D.L., Geysen, D.M. and Brandt, J.R., 2000. Transmission of theileriosis in the traditional farming sector in the southern province of Zambia during 1995-1996. Tropical Animal Health and Production 32: 303-314.

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

Murrell, A., Campbell, N.J. and Barker, S.C., 2000. Phylogenetic analyses of the rhipicephaline ticks indicate that the genus *Rhipicephalus* is paraphyletic. Molecular Phylogenetics and Evolution 16: 1-7.

Murrell, A., Campbell, N.J. 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: 244-258.

Musoke, A., Morzaria, S., Nkonge, C., Jones, E. and Nene, V., 1992. A recombinant sporozoite surface antigen of *Theileria parva* induces protection in cattle. Proceedings of the National Academy of Sciences of the United States of America 89: 514-518.

Musoke, A., Rowlands, J., Nene, V., Nyanjui, J., Katende, J., Spooner, P., Mwaura, S., Odongo, D., Nkonge, C., Mbogo, S., Bishop, R. and Morzaria, S., 2005. Subunit vaccine based on the p67 major surface protein of *Theileria parva* sporozoites reduces severity of infection derived from field tick challenge. Vaccine 23: 3084-3095.

Muwanika, V., Kabi, F. and Masembe, C., 2016. Population genetic structure of *Theileria parva* field isolates from indigenous cattle populations of Uganda. Ticks and Tick-Borne Diseases 7: 291-297.

Mwangi, E.N., Newson, R.M. and Kaaya, G.P., 1991. Drop-Off Patterns for Engorged Adult Females, Nymphs and Larvae of *Rhipicephalus appendiculatus*. International Journal of Tropical Insect Science 12: 629-633.

Nadolny, R., Gaff, H., Carlsson, J. and Gauthier, D., 2015. Comparative population genetics of two invading ticks: Evidence of the ecological mechanisms underlying tick range expansions. Infection, Genetics and Evolution 35: 153-162.

Nava, S., Guglielmone, A.A. and Mangold, A.J., 2009. An overview of systematics and evolution of ticks. Frontiers in Bioscience 14: 2857-2877.

Ndumu, D.B., Baumung, R., Hanotte, O., Wurzinger, M., Okeyo, M.A., Jianlin, H., Kibogo, H. and Sölkner, J., 2008. Genetic and morphological characterisation of the Ankole Longhorn cattle in the African Great Lakes region. Genetics Selection Evolution 40: 467.

Ndungu, S.G., Brown, C.G. and Dolan, T.T., 2005. In vivo comparison of susceptibility between *Bos indicus* and *Bos taurus* cattle types to *Theileria parva* infection. The Onderstepoort Journal of Veterinary Research 72: 13-22.

Nene, V., Iams, K.P., Gobright, E. and Musoke, A.J., 1992. Characterisation of the gene encoding a candidate vaccine antigen of *Theileria parva* sporozoites. Molecular and Biochemical Parasitology 51: 17-27.

Nene, V., Kiara, H., Lacasta, A., Pelle, R., Svitek, N. and Steinaa, L., 2016. The biology of *Theileria parva* and control of East Coast fever - Current status and future trends. Ticks and Tick-Borne Diseases 7: 549-564.

Nene, V. and Morrison, W.I., 2016. Approaches to vaccination against *Theileria parva* and *Theileria annulata*. Parasite Immunology 38: 724-734.

Nene, V., Morzaria, S. and Bishop, R., 1998. Organisation and informational content of the *Theileria parva* genome. Molecular and Biochemical Parasitology 95: 1-8.

Nene, V., Musoke, A., Gobright, E. and Morzaria, S., 1996. Conservation of the sporozoite p67 vaccine antigen in cattle-derived *Theileria parva* stocks with different cross-immunity profiles. Infection and Immunology 64: 2056-2061.

Nilsson, P., Backman, M., Bjerke, L. and Maniriho, A., 2019. One cow per poor family: Effects on the growth of consumption and crop production. World Development 114: 1-12.

Norling, M., Bishop, R.P., Pelle, R., Qi, W., Henson, S., Drabek, E.F., Tretina, K., Odongo, D., Mwaura, S., Njoroge, T., Bongcam-Rudloff, E., Daubenberger, C.A. and Silva, J.C., 2015. The genomes of three stocks comprising the most widely utilized live sporozoite *Theileria parva* vaccine exhibit very different degrees and patterns of sequence divergence. BMC Genomics 16: 729.

Norval, R.A., 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., 1992. The epidemiology of theileriosis in Africa. Harcourt Brace Jovanovich, London.

Nyagwange, J., Awino, E., Tijhaar, E., Svitek, N., Pelle, R. and Nene, V., 2019. Leveraging the Medicines for Malaria Venture malaria and pathogen boxes to discover chemical inhibitors of East Coast fever. International Journal for Parasitology: Drugs and Drug Resistance 9: 80-86.

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.

Ochanda, H., Young, A.S., Mutugi, J.J., Mumo, J. and Omwoyo, P.L., 1988. The effect of temperature on the rate of transmission of *Theileria parva parva* infection to cattle by its tick vector, *Rhipicephalus appendiculatus*. Parasitology 97: 239-245.

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

Odongo, D.O., Oura, C.A., Spooner, P.R., Kiara, H., Mburu, D., Hanotte, O.H. and Bishop, R.P., 2006. Linkage disequilibrium between alleles at highly polymorphic mini- and micro-satellite loci of *Theileria parva* isolated from cattle in three regions of Kenya. International Journal for Parasitology 36: 937-946.

Odongo, D.O., Sunter, J.D., Kiara, H.K., Skilton, R.A. and Bishop, R.P., 2010. A nested PCR assay exhibits enhanced sensitivity for detection of *Theileria parva* infections in bovine blood samples from carrier animals. Parasitology Research 106: 357-365.

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: 888-894.

Ogden, N.H., Gwakisa, P., Swai, E., French, N.P., Fitzpatrick, J., Kambarage, D. and Bryant, M., 2003. Evaluation of PCR to detect *Theileria parva* in field-collected tick and bovine samples in Tanzania. Veterinary Parasitology 112: 177-183.

Ogden, N.H., Mechai, S. and Margos, G., 2013. Changing geographic ranges of ticks and tick-borne pathogens: drivers, mechanisms and consequences for pathogen diversity. Frontiers in Cellular and Infection Microbiology 3: 46.

Okello-Onen, 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.

Olds, C.L., Mason, K.L. and Scoles, G.A., 2018. *Rhipicephalus appendiculatus* ticks transmit *Theileria parva* from persistently infected cattle in the absence of detectable parasitemia: implications for East Coast fever epidemiology. Parasites & Vectors 11: 126.

Olds, C.L., Mwaura, S., Odongo, D.O., Scoles, G.A., Bishop, R. and Daubenberger, C., 2016. Induction of humoral immune response to multiple recombinant *Rhipicephalus appendiculatus* antigens and their effect on tick feeding success and pathogen transmission. Parasites & Vectors 9: 484.

Olwoch, J.M., Rautenbach, C.J.d.W., Erasmus, B.F.N., Engelbrecht, F.A. and Van Jaarsveld, A.S., 2003. Simulating tick distributions over sub-Saharan Africa: the use of observed and simulated climate surfaces. Journal of Biogeography 30: 1221-1232.

Olwoch, J.M., Reyers, B., Engelbrecht, F.A. and 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: 108-120.

Olwoch, J.M., Reyers, B. and van Jaarsveld, A.S., 2009. Host–parasite distribution patterns under simulated climate: implications for tick-borne diseases. International Journal of Climatology 29: 993-1000.

Ostfeld, R.S., 2009. Climate change and the distribution and intensity of infectious diseases. Ecology 90: 903-905.

Ostfeld, R.S., Canham, C.D., Oggenfuss, K., Winchcombe, R.J. and Keesing, F., 2006. Climate, Deer, Rodents, and Acorns as Determinants of Variation in Lyme-Disease Risk. PLOS Biology 4: e145.

Oura, C., Bishop, R., Wampande, E., Lubega, G. and Tait, A., 2004a. The persistence of component *Theileria parva* stocks in cattle immunized with the 'Muguga cocktail'live vaccine against East Coast fever in Uganda. Parasitology 129: 27-42.

Oura, C.A., Asiimwe, B.B., Weir, W., Lubega, G.W. and Tait, A., 2005. Population genetic analysis and sub-structuring of *Theileria parva* in Uganda. Molecular and Biochemical Parasitology 140: 229-239.

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

Oura, C.A., Bishop, R.P., Wampande, E.M., Lubega, G.W. and Tait, A., 2004b. Application of a reverse line blot assay to the study of haemoparasites in cattle in Uganda. International Journal for Parasitology 34: 603-613.

Oura, C.A., 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 characterisation of field isolates of *Theileria parva*. International Journal for Parasitology 33: 1641-1653.

Oura, C.A., Tait, A., Asiimwe, B., Lubega, G.W. and Weir, W., 2011a. Haemoparasite prevalence and *Theileria parva* strain diversity in Cape buffalo (*Syncerus caffer*) in Uganda. Veterinary Parasitology 175: 212-219.

Oura, C.A., Tait, A., Asiimwe, B., Lubega, G.W. and Weir, W., 2011b. *Theileria parva* genetic diversity and haemoparasite prevalence in cattle and wildlife in and around Lake Mburo National Park in Uganda. Parasitology Research 108: 1365-1374.

Pain, A., Renauld, H., Berriman, M., Murphy, L., Yeats, C.A., Weir, W., Kerhornou, A., Aslett, M., Bishop, R., Bouchier, C., Cochet, M., Coulson, R.M., Cronin, A., de Villiers, E.P., Fraser, A., Fosker, N., Gardner, M., Goble, A., Griffiths-Jones, S., Harris, D.E., Katzer, F., Larke, N., Lord, A., Maser, P., McKellar, S., Mooney, P., Morton, F., Nene, V., O'Neil, S., Price, C., Quail, M.A., Rabbinowitsch, E., Rawlings, N.D., Rutter, S., Saunders, D., Seeger, K., Shah, T., Squares, R., Squares, S., Tivey, A., Walker, A.R., Woodward, J., Dobbelaere, D.A., Langsley, G., Rajandream, M.A., McKeever, D., Shiels, B., Tait, A., Barrell, B. and Hall, N., 2005. Genome of the host-cell transforming parasite *Theileria annulata* compared with *T. parva*. Science 309: 131-133.

Paling, R.W. and Geysen, D., 1981. Observations on Rwandan Strains of *Theileria parva* and the Value of *T. parva* Nyakizu as a Possible Vaccine Strain, In: Irvin, A.D., Cunningham, M.P., Young, A.S. (Eds.) Advances in the Control of Theileriosis: Proceedings of an International Conference held at the International Laboratory for Research on Animal Diseases in Nairobi, 9–13th February, 1981. Springer Netherlands, Dordrecht, pp. 238-241.

Paling, R.W., Mpangala, C., Luttikhuizen, B. and Sibomana, G., 1991. Exposure of Ankole and crossbred cattle to theileriosis in Rwanda. Tropical Animal Health and Production 23: 203-214.

Parizi, L.F., Rangel, C.K., Sabadin, G.A., Saggin, B.F., Kiio, I., Xavier, M.A., da Silva Matos, R., Camargo-Mathias, M.I., Seixas, A., Konnai, S., Ohashi, K., Githaka, N.W. and da Silva Vaz, I., 2020. *Rhipicephalus microplus* cystatin as a potential cross-protective tick vaccine against *Rhipicephalus appendiculatus*. Ticks and Tick-Borne Diseases 11: 101378.

Patel, E.H., Lubembe, D.M., Gachanja, J., Mwaura, S., Spooner, P. and Toye, P., 2011. Molecular characterization of live *Theileria parva* sporozoite vaccine stabilates reveals extensive genotypic diversity. Veterinary Parasitology 179: 62-68.

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., Toye, P.G., Spooner, P.R., Musoke, A.J., Mwangi, D.M., Taracha, E.L., Morrison, W.I., Weir, W., Silva, J.C. and 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. PLoS One 6: e19015.

Perry, B.D., 2016. The control of East Coast fever of cattle by live parasite vaccination: A science-to-impact narrative. One health 2: 103-114.

Perry, B.D., Kruska, R., Lessard, P., Norval, R.A.I. and Kundert, K., 1991. Estimating the distribution and abundance of *Rhipicephalus appendiculatus* in Africa. Preventive Veterinary Medicine 11: 261-268.

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

Pienaar, R., Potgieter, F.T., Latif, A.A., Thekisoe, O.M. and Mans, B.J., 2011. Mixed *Theileria* infections in free-ranging buffalo herds: implications for diagnosing

Theileria parva infections in Cape buffalo (*Syncerus caffer*). Parasitology 138: 884-895.

Radley, D.E., Brown, C.G.D., Cunningham, M.P., Kimber, C.D., Musisi, F.L., Payne, R.C., Purnell, R.E., Stagg, S.M. and Young, A.S., 1975. East coast fever: 3. Chemoprophylactic immunization of cattle using oxytetracycline and a combination of theilerial strains. Veterinary Parasitology 1: 51-60.

Randolph, S.E., 2000. Ticks and tick-borne disease systems in space and from space. Advances in Parasitology 47: 217-243.

Randolph, S.E., 2004. Tick ecology: processes and patterns behind the epidemiological risk posed by ixodid ticks as vectors. Parasitology 129: S37-S65.

Randolph, S.E. and Rogers, D.J., 1997. A generic population model for the African tick *Rhipicephalus appendiculatus*. Parasitology 115: 265-279.

Ray, N., Currat, M. and Excoffier, L., 2003. Intra-deme molecular diversity in spatially expanding populations. Molecular Biology and Evolution 20: 76-86.

Robbertse, L., Baron, S., van der Merwe, N.A., Madder, M., Stoltsz, W.H. and Maritz-Olivier, C., 2016. Genetic diversity, acaricide resistance status and evolutionary potential of a *Rhipicephalus microplus* population from a disease-controlled cattle farming area in South Africa. Ticks and Tick-Borne Diseases 7: 595-603.

Rogers, A.R., Fraley, A.E., Bamshad, M.J., Watkins, W.S. and Jorde, L.B., 1996. Mitochondrial mismatch analysis is insensitive to the mutational process. Molecular Biology and Evolution 13: 895-902.

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.

Rozas, J., Ferrer-Mata, A., Sanchez-DelBarrio, J.C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S.E. and Sanchez-Gracia, A., 2017. DnaSP 6: DNA Sequence Polymorphism Analysis of Large Data Sets. Molecular Biology and Evolution 34: 3299-3302.

Rubaire-Akiiki, C., Okello-Onen, J., Nasinyama, G.W., Vaarst, M., Kabagambe, E.K., Mwayi, W., Musunga, D. and Wandukwa, W., 2004. The prevalence of serum antibodies to tick-borne infections in Mbale District, Uganda: the effect of agro-ecological zone, grazing management and age of cattle. Journal of Insect Science 4: 8.

Rubaire-Akiiki, C.M., Okello-Onen, J., Musunga, D., Kabagambe, E.K., Vaarst, M., Okello, D., Opolot, C., Bisagaya, A., Okori, C., Bisagati, C., Ongyera, S. and Mwayi, M.T., 2006. Effect of agro-ecological zone and grazing system on incidence of East Coast Fever in calves in Mbale and Sironko Districts of Eastern Uganda. Preventive Veterinary Medicine 75: 251-266.

Rukambile, E., Machuka, E., Njahira, M., Kyalo, M., Skilton, R., Mwega, E., Chota, A., Mathias, M., Sallu, R. and Salih, D., 2016. Population genetic analysis of *Theileria parva* isolated in cattle and buffaloes in Tanzania using minisatellite and microsatellite markers. Veterinary Parasitology 224: 20-26.

Salih, D.A., Mwacharo, J.M., Pelle, R., Njahira, M.N., Odongo, D.O., Mbole-Kariuki, M.N., Marcellino, W.L., Malak, A.K., Kiara, H., El Hussein, A.R.M., Bishop, R.P. and Skilton, R.A., 2018. Genetic diversity and population structure of *Theileria parva* in South Sudan. Ticks and Tick-Borne Diseases 9: 806-813.

Salih, D.A., Pelle, R., Mwacharo, J.M., Njahira, M.N., Marcellino, W.L., Kiara, H., Malak, A.K., El Hussein, A.R., Bishop, R. and Skilton, R.A., 2017. Genes encoding two *Theileria parva* antigens recognized by CD8+ T-cells exhibit sequence diversity in South Sudanese cattle populations but the majority of alleles are similar to the Muguga component of the live vaccine cocktail. PLoS One 12: e0171426.

Shao, R. and Barker, S.C., 2007. Mitochondrial genomes of parasitic arthropods: implications for studies of population genetics and evolution. Parasitology 134: 153-167.

Shapiro, S.Z., Fujisaki, K., Morzaria, S.P., Webster, P., Fujinaga, T., Spooner, P.R. and Irvin, A.D., 1987. A life-cycle stage-specific antigen of *Theileria parva* recognized by anti-macroschizont monoclonal antibodies. Parasitology 94: 29-37.

Shaw, M.K., 2003. Cell invasion by *Theileria* sporozoites. Trends in Parasitology 19, 2-6.

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

Sibeko, K.P., Collins, N.E., Oosthuizen, M.C., Troskie, M., Potgieter, F.T., Coetzer, J.A. and Geysen, D., 2011. Analyses of genes encoding *Theileria parva* p104 and polymorphic immunodominant molecule (PIM) reveal evidence of the presence of cattle-type alleles in the South African *T. parva* population. Veterinary Parasitology 181: 120-130.

Sibeko, K.P., Geysen, D., Oosthuizen, M.C., Matthee, C.A., Troskie, M., Potgieter, F.T., Coetzer, J.A. and Collins, N.E., 2010. Four p67 alleles identified in South African *Theileria parva* field samples. Veterinary Parasitology 167: 244-254.

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 the Entomological Society of America 87: 651-701.

Simonsen, K.L., Churchill, G.A. and Aquadro, C.F., 1995. Properties of statistical tests of neutrality for DNA polymorphism data. Genetics 141: 413-429.

Sitt, T., Henson, S., Morrison, W.I. and Toye, P., 2019. Similar levels of diversity in the gene encoding the p67 sporozoite surface protein of *Theileria parva* are observed in blood samples from buffalo and cattle naturally infected from buffalo. Veterinary Parasitology 269: 21-27.

Sitt, T., Pelle, R., Chepkwony, M., Morrison, W.I. and Toye, P., 2018. *Theileria parva* antigens recognized by CD8+ T cells show varying degrees of diversity in buffalo-derived infected cell lines. Parasitology 145: 1430-1439.

Sitt, T., Poole, E.J., Ndambuki, G., Mwaura, S., Njoroge, T., Omondi, G.P., Mutinda, M., Mathenge, J., Prettejohn, G., Morrison, W.I. and Toye, P., 2015. Exposure of vaccinated and naive cattle to natural challenge from buffalo-derived *Theileria parva*. International Journal for Parasitology: Parasites and Wildlife 4: 244-251.

Sivakumar, T., Hayashida, K., Sugimoto, C. and Yokoyama, N., 2014. Evolution and genetic diversity of *Theileria*. Infection, Genetics and Evolution 27: 250-263.

Skilton, R.A., Bishop, R.P., Katende, J.M., Mwaura, S. and Morzaria, S.P., 2002. The persistence of *Theileria parva* infection in cattle immunized using two stocks which differ in their ability to induce a carrier state: analysis using a novel blood spot PCR assay. Parasitology 124: 265-276.

Skilton, R.A., Bishop, R.P., Wells, C.W., Spooner, P.R., Gobright, E., Nkonge, C., Musoke, A.J., Macklin, M. and Iams, K.P., 1998. Cloning and characterization of a 150 kDa microsphere antigen of *Theileria parva* that is immunologically cross-reactive with the polymorphic immunodominant molecule (PIM). Parasitology 117: 321-330.

Skilton, R.A., Musoke, A.J., Wells, C.W., Yagi, Y., Nene, V., Spooner, P.R., Gachanja, J., Osaso, J., Bishop, R.P. and Morzaria, S.P., 2000. A 32 kDa surface antigen of *Theileria parva*: characterization and immunization studies. Parasitology 120: 553-564.

Slatkin, M. and Hudson, R.R., 1991. Pairwise comparisons of mitochondrial DNA sequences in stable and exponentially growing populations. Genetics 129: 555-562.

Sonenshine, D.E. and Roe, R.M., 2013. Biology of ticks, Vol 2. Oxford University Press, New York.

Sosa-Gutierrez, C.G., Vargas-Sandoval, M., Torres, J. and Gordillo-Perez, G., 2016. Tick-borne rickettsial pathogens in questing ticks, removed from humans and animals in Mexico. Journal of Veterinary Science 17: 353-360.

Speybroeck, N., Madder, M., Brandt, J., Chungu, H., Van Den Bossche, P., Mbao, V. and Berkvens, D., 2003. Questing activity of *Rhipicephalus appendiculatus* (Acari: Ixodidae) nymphs: a random process? Physiological Entomology 28: 356-361.

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: 347-354.

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: 430-441.

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

Steinaa, L., Saya, R., Awino, E. and Toye, P., 2012. Cytotoxic T lymphocytes from cattle immunized against *Theileria parva* exhibit pronounced cross-reactivity among different strain-specific epitopes of the Tp1 antigen. Veterinary Immunology and Immunopathology 145: 571-581.

Steinaa, L., Svitek, N., Awino, E., Njoroge, T., Saya, R., Morrison, I. and Toye, P., 2018a. Immunization with one *Theileria parva* strain results in similar level of CTL strain-specificity and protection compared to immunization with the three-component Muguga cocktail in MHC-matched animals. BMC Veterinary Research 14: 145.

Steinaa, L., Svitek, N., Awino, E., Saya, R. and Toye, P., 2018b. Cytotoxic T lymphocytes from cattle sharing the same MHC class I haplotype and immunized with live *Theileria parva* sporozoites differ in antigenic specificity. BMC Research Notes 11: 44.

Svitek, N., Saya, R., Awino, E., Munyao, S., Muriuki, R., Njoroge, T., Pelle, R., Ndiwa, N., Poole, J., Gilbert, S., Nene, V. and Steinaa, L., 2018. An Ad/MVA vectored *Theileria parva* antigen induces schizont-specific CD8(+) central memory T cells and confers partial protection against a lethal challenge. NPJ Vaccines 3: 35.

Swai, E.S., Karimuribo, E.D., Rugaimukamu, E.A. and Kambarage, D.M., 2006. Factors influencing the distribution of questing ticks and the prevalence estimation of *T. parva* infection in brown ear ticks in the Tanga region, Tanzania. Journal of Vector Ecology 31: 224-228.

Tabachnick, W.J. and Black, W.C., 1995. Making a case for molecular population genetic studies of arthropod vectors. Parasitology Today 11: 27-30.

Tajima, F., 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics 123: 585-595.

Tama, E., 1989. East Coast fever immunization in Burundi. In: Dolan TT, editor. Theileriosis in eastern, central and southern Africa. Proceedings of a Workshop on East Coast Fever Immunization, 1988 September 20–22; Lilongwe, Malawi. Nairobi: The International Laboratory for Research on Animal Diseases. p. 37–8.

Taracha, E.L., Goddeeris, B.M., Morzaria, S.P. and Morrison, W.I., 1995a. Parasite strain specificity of precursor cytotoxic T cells in individual animals correlates with cross-protection in cattle challenged with *Theileria parva*. Infection and Immunology 63: 1258-1262.

Taracha, E.L., Goddeeris, B.M., Teale, A.J., Kemp, S.J. and Morrison, W.I., 1995b. Parasite strain specificity of bovine cytotoxic T cell responses to *Theileria parva* is determined primarily by immunodominance. Journal of Immunology 155: 4854-4860.

Tayebwa, D.S., Vudriko, P., Tuvshintulga, B., Guswanto, A., Nugraha, A.B., Gantuya, S., Batiha, G.E., Musinguzi, S.P., Komugisha, M., Bbira, J.S., Okwee-Acai, J., Tweyongyere, R., Wampande, E.M., Byaruhanga, J., Adjou Moumouni, P.F., Sivakumar, T., Yokoyama, N. and Igarashi, I., 2018. Molecular epidemiology of *Babesia* species, *Theileria parva*, and *Anaplasma marginale* infecting cattle and the tick control malpractices in Central and Eastern Uganda. Ticks and Tick-Borne Diseases 9: 1475-1483.

Tempia, S., 1997. Comparison of the vectorial capacity of *Rhipicephalus appendiculatus* (Acari: Ixodidae) nymphs and adults originating from eastern and southern provinces of Zambia infected by *Theileria parva* Katete stock. Institute of Tropical Medicine, Antwerp, Belgium.

Theiler, A., 1912. The immunisation of cattle against East Coast fever. Second Report of the Director of Veterinary Research, 216–314.

Tibayrenc, M., Kjellberg, F. and Ayala, F.J., 1990. A clonal theory of parasitic protozoa: the population structures of *Entamoeba*, *Giardia*, *Leishmania*, *Naegleria*, *Plasmodium*, *Trichomonas*, and *Trypanosoma* and their medical and taxonomical consequences. Proceedings of the National Academy of Sciences of the United States of America 87: 2414-2418.

Toye, P., Gobright, E., Nyanjui, J., Nene, V. and Bishop, R., 1995a. Structure and sequence variation of the genes encoding the polymorphic, immunodominant molecule (PIM), an antigen of *Theileria parva* recognized by inhibitory monoclonal antibodies. Molecular and Biochemical Parasitology 73: 165-177.

Toye, P.G., Goddeeris, B.M., Iams, K., Musoke, A.J. and Morrison, W.I., 1991. Characterization of a polymorphic immunodominant molecule in sporozoites and schizonts of *Theileria parva*. Parasite Immunology 13: 49-62.

Toye, P.G., Metzelaar, M.J., Wijngaard, P., Nene, V., Iams, K., Roose, J., Nyanjui, J.K., Gobright, E., Musoke, A.J. and Clevers, H.C., 1995b. Characterization of the gene encoding the polymorphic immunodominant molecule, a neutralizing antigen of *Theileria parva*. The Journal of Immunology 155: 1370-1381.

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

Vajana, E., Barbato, M., Colli, L., Milanesi, M., Rochat, E., Fabrizi, E., Mukasa, C., Del Corvo, M., Masembe, C., Muwanika, V.B., Kabi, F., Sonstegard, T.S., Huson, H.J., Negrini, R., Joost, S. and Ajmone-Marsan, P., 2018. Combining landscape genomics and ecological modelling to investigate local adaptation of indigenous Ugandan cattle to East Coast Fever. Frontiers in Genetics 9: 385.

Van Saceghem, R., 1925. Propagation de *Theileria parva* par les tiques. Bulletin Agricole du Congo Belge 16: 582-591.

Verweijen, J. and Brabant, J., 2017. Cows and guns. Cattle-related conflict and armed violence in Fizi and Itombwe, eastern DR Congo. The Journal of Modern African Studies 55: 1-27.

Verweijen, J. and Vlassenroot, K., 2015. Armed mobilisation and the nexus of territory, identity, and authority: the contested territorial aspirations of the Banyamulenge in eastern DR Congo. Journal of Contemporary African Studies 33: 191-212.

Vlassenroot, K. and Huggins, C., 2005. Land, migration and conflict in eastern DRC. In: Huggins C, Clover J, editors. From the ground up: land rights, conflict and peace in sub-Saharan Africa. Institute for Security Studies, Pretoria. p. 115–194.

Vudriko, P., Okwee-Acai, J., Byaruhanga, J., Tayebwa, D.S., Okech, S.G., Tweyongyere, R., Wampande, E.M., Okurut, A.R.A., Mugabi, K., Muhindo, J.B., Nakavuma, J.L., Umemiya-Shirafuji, R., Xuan, X. and Suzuki, H., 2018. Chemical tick control practices in southwestern and northwestern Uganda. Ticks and Tick-Borne Diseases 9: 945-955.

Vudriko, P., Okwee-Acai, J., Tayebwa, D.S., Byaruhanga, J., Kakooza, S., Wampande, E., Omara, R., Muhindo, J.B., Tweyongyere, R., Owiny, D.O., Hatta, T., Tsuji, N., Umemiya-Shirafuji, R., Xuan, X., Kanameda, M., Fujisaki, K. and Suzuki, H., 2016. Emergence of multi-acaricide resistant *Rhipicephalus* ticks and its implication on chemical tick control in Uganda. Parasites & Vectors 9: 4.

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, Edinburg Scotland, UK.

Walker, J.B., 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, UK.

Wang, H., Kaufman, W.R., Cui, W.W. and Nuttall, P.A., 2001. Molecular individuality and adaptation of the tick *Rhipicephalus appendiculatus* in changed feeding environments. Medical and Veterinary Entomology 15: 403-412.

Wang, T., Zhang, S., Pei, T., Yu, Z. and Liu, J., 2019. Tick mitochondrial genomes: structural characteristics and phylogenetic implications. Parasites & Vectors 12: 451.

Wanjohi, J.M., Ngeranwa, J.N., Rumberia, R.M., Muraguri, G.R. and Mbogo, S.K., 2001. Immunization of cattle against East Coast fever using *Theileria parva* (Marikebuni)

and relaxation of tick control in North Rift, Kenya. The Onderstepoort Journal of Veterinary Research 68: 217-223.

Wurzinger, M., Ndumu, D., Baumung, R., Drucker, A., Okeyo, A.M., Semambo, D.K., Byamungu, N. and Sölkner, J., 2006. Comparison of production systems and selection criteria of Ankole cattle by breeders in Burundi, Rwanda, Tanzania and Uganda. Tropical Animal Health and Production 38: 571-581.

Yawa, M., Nyangiwe, N., Muchenje, V., Kadzere, C.T., Mpendulo, T.C. and Marufu, M.C., 2018. Ecological preferences and seasonal dynamics of ticks (Acari: Ixodidae) on and off bovine hosts in the Eastern Cape Province, South Africa. Experimental and Applied Acarology 74: 317-328.

Young, A.S., Dolan, T.T., Morzaria, S.P., Mwakima, F.N., Norval, R.A., Scott, J., Sherriff, A. and Gettinby, G., 1996. Factors influencing infections in *Rhipicephalus appendiculatus* ticks fed on cattle infected with *Theileria parva*. Parasitology 113: 255-266.

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.

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

Yssouf, A., Lagadec, E., Bakari, A., Foray, C., Stachurski, F., Cardinale, E., Plantard, O. and Tortosa, P., 2011. Colonization of Grande Comore Island by a lineage of *Rhipicephalus appendiculatus* ticks. Parasites & Vectors 4: 38.

Appendices

Appendices (Chapter 3)

Table S3.1 *Rhipicephalus appendiculatus cox*1 and *12S* rRNA haplotype sequences retrieved from GenBank

Gene locus	Geographical origin	Number of sequences	GenBank accession number	Reference
Cox1	Kenya	29	KU725890-917	[31]
			DQ901358-59	[29]
	Zimbabwe	3	AF132833	[68]
			KX276944	[31]
			KC503257	[69]
	Comoros	1	DQ901357	[29]
	Rwanda	3	DQ901360, DQ901362-63	[29]
	Zambia (Southern province)	2	DQ859262	[52]
			KX276943	[31]
	Zambia (Eastern province)	7	DQ859261, DQ859263-66	[52]
			DQ901361	[29]
			KX276942	[31]
	South Africa	3	DQ901356	[29]
			KX276939-40	[31]
	Uganda	2	KX276941, KU725897	[31]
12S rRNA	Kenya	6	KX276945-49	[31]
			DQ901320	[29]
	Zimbabwe	2	AF031859, AF150027	[73]
	Comoros	1	DQ901317	[29]
	Rwanda	5	DQ901279, DQ901281-82, DQ901284, DQ901286	[29]
	Zambia (Southern province)	6	DQ849203-05, DQ849208, DQ901309, DQ901311	[29]
	Zambia (Eastern province)	6	DQ849207, DQ849210, DQ849212, DQ849214, DQ901277, DQ901288	[29]
	South Africa	2	DQ849233, DQ849235	[29]
	Uganda	1	AF150028	[70]

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

[31]. Kanduma EG, Mwacharo JM, Githaka NW, Kinyanjui PW, Njuguna JN, Kamau LM, et al. Analyses of mitochondrial genes reveal two sympatric but genetically divergent lineages of *Rhipicephalus appendiculatus* in Kenya. Parasit Vectors. 2016;9:353.

[52]. Mtambo J, Madder M, Van Bortel W, Chaka G, Berkvens D, Backeljau T. Further evidence for geographic differentiation in *R. appendiculatus* (Acari: Ixodidae) from Eastern and Southern provinces of Zambia. Exp Appl Acarol. 2007;41:129-38.

[68]. Murrell A, Campbell NJH, Barker SC. Phylogenetic Analyses of the Rhipicephaline Ticks Indicate That the Genus Rhipicephalus Is Paraphyletic. Mol Phylogenet Evol. 2000;16:1-7.

[69]. Burger TD, Shao R, Barker SC. Phylogenetic analysis of mitochondrial genome sequences indicates that the cattle tick, *Rhipicephalus (Boophilus) microplus*, contains a cryptic species. Mol Phylogenet Evol. 2014;76:241-53.

[70] Beati L, Keirans JE. 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. J Parasitol. 2001;87:32-48.

Table S3.2 *cox*1 and *12S* rRNA BLAST results for species identification and confirmation

Gene locus	Haplotype	GenBank accession no.	BLAST results			Reference
			GenBank match	Identity (%)	Source	
cox1	CH1	MF458950	KU725895	100	Kenya	[31]
	CH2	MF458951	KU725893	100	Kenya	[31]
	CH3	MF458952	KU725895	99	Kenya	[31]
	CH4	MF458953	KU725895	99	Kenya	[31]
	CH5	MF458954	KU725891	100	Kenya	[31]
	CH6	MF458955	KU725894	99	Kenya	[31]
	CH7	MF458956	KC503257	100	Zimbabwe	[69]
	CH8	MF458957	KU725894	100	Kenya	[31]
	CH9	MF458958	KU725893	99	Kenya	[31]
	CH10	MF458959	KU725895	99	Kenya	[31]
	CH11	MF458960	KU725892	100	Kenya	[31]
	CH12	MF458961	KU725900	99	Kenya	[31]
	CH13	MF458962	AF132833	100	Zimbabwe	[68]
	CH14	MF458963	KU725894	99	Kenya	[31]
	CH15	MF458964	KU725895	99	Kenya	[31]
	CH16	MF458965	KU725895	99	Kenya	[31]
	CH17	MF458966	KU725895	99	Kenya	[31]
	CH18	MF458967	KU725893	99	Kenya	[31]
	CH19	MF458968	KU725892	99	Kenya	[31]
	CH20	MF458969	KC503257	99	Zimbabwe	[70]
	CH21	MF458970	KU725891	99	Kenya	[31]
	CH22	MF458971	KU725894	99	Kenya	[31]
12S rRNA	12SH1	MF479189	DQ849209	100	Zambia(east)	[29]
	12SH2	MF479190	AF150028	100	Uganda	[70]
	12SH3	MF479191	DQ849209	99	Zambia(east)	[29]
	12SH4	MF479192	DQ849203	99	Zambia(south)	[29]
	12SH5	MF479193	DQ849203	99	Zambia(south)	[29]
	12SH6	MF479194	DQ849209	99	Zambia(east)	[29]
	12SH7	MF479195	DQ849209	99	Zambia(east)	[29]
	12SH8	MF479196	DQ849209	99	Zambia(east)	[29]
	12SH9	MF479197	DQ849209	99	Zambia(east)	[29]

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

[31]. Kanduma EG, Mwacharo JM, Githaka NW, Kinyanjui PW, Njuguna JN, Kamau LM, et al. Analyses of mitochondrial genes reveal two sympatric but genetically divergent lineages of *Rhipicephalus appendiculatus* in Kenya. Parasit Vectors. 2016;9:353.

[69]. Burger TD, Shao R, Barker SC. Phylogenetic analysis of mitochondrial genome sequences indicates that the cattle tick, *Rhipicephalus (Boophilus) microplus*, contains a cryptic species. Mol Phylogenet Evol. 2014;76:241-53.

[68]. Murrell A, Campbell NJH, Barker SC. Phylogenetic Analyses of the Rhipicephaline Ticks Indicate That the Genus Rhipicephalus Is Paraphyletic. Mol Phylogenet Evol. 2000;16:1-7.

[70] Beati L, Keirans JE. 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. J Parasitol. 2001;87:32-48.

Table S3.3 Polymorphism in the 22 haplotypes of the *cox*1 gene fragment of *R. appendiculatus*

Haplotype	No. of sequences (%)[a]	Polymorphic sites (27 positions) *																												Haplogroup
		21	41	56	80	93	95	104	107	143	164	167	179	233	249	302	314	320	329	395	401	458	470	498	500	527	539	548		
CH1	56 (27)	C	A	A	T	G	T	C	T	A	T	C	G	T	T	C	T	T	T	G	T	G	C	C	A	G	A	A		A
CH2	59 (28)	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.		A
CH3	1 (0.5)	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.		A
CH4	2 (1)	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G		A
CH5	39 (19)	.	G	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.		A
CH6	2 (1)	.	.	.	.	.	.	T	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.		A
CH8	2 (1)	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.		A
CH9	1 (0.5)	.	G	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.		A
CH10	1 (0.5)	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	G	.		A
CH11	10 (5)	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	A	.	.	.	.	.	.	.	.		A
CH12	5 (2)	.	G	.	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.		A
CH14	1 (0.5)	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	T	.	.	.	.		A
CH15	1 (0.5)	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.		A
CH16	2 (1)	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.		A
CH17	1 (0.5)	.	.	.	C	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.		A
CH18	2 (1)	.	G	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.		A
CH19	2 (1)	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	C	.	A	.	.	.	.	.	.	.	.		A
CH21	1 (0.5)	.	G	G	.	.	.	.	.	G	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.		A
CH22	1 (0.5)	.	G	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	.	A	.	.	.	.	.	.	.	.		A
CH7	8 (4)	T	T	.	.	.	.	T	.	.	.	T	.	.	C	T	C	.	.	.	C	A	T	.	G	T	.	.		B
CH13	11 (5)	T	T	.	.	.	.	T	.	G	.	T	.	.	C	T	C	.	.	.	C	A	T	.	G	T	.	.		B
CH20	1 (0.5)	T	T	.	.	.	.	T	.	.	.	T	.	C	C	T	C	.	.	.	C	A	T	.	G	T	.	.		B

[a] Number of specimens sharing the same haplotype

*The number represents the position of variable sites (bp)

dots (.) represent nucleotides identical to haplotype 1

Table S3.4 Population genetic structure inferred by analysis of molecular variance (AMOVA) based on *cox*1 sequences of *R. appendiculatus* from different agro-ecological zones

Haplogroup	**Level of partitioning**	**Source of variation**	**d.f.**	**Percentage of variation**	***P-value***
Overall data	AEZs	Among AEZ	5	6	< 0.001
		Within AEZ	203	94	-
Haplogroup A	AEZs	Among AEZ	5	6	< 0.001
		Within AEZ	183	94	-
	DRC	Among AEZ	2	6	< 0.001
		Within AEZ	130	94	
	Burundi	Among AEZ	1	6	0.06
		Within AEZ	40	94	-

d.f., degree of freedom

Table S3.5 Evolutionary neutrality, demographic and spatial history of mitochondrial *cox*1 gene

Statistics	DRC			Burundi		Rwanda	Overall
	AEZ1	**AEZ2**	**AEZ3**	**AEZ1**	**AEZ3**	**AEZ2**	
Selective neutrality tests							
Tajima's D	0.41	-0.95	-1.3	-2.1	-0.21	0.92	-0.93
(P-value)	(0.71)	(0.17)	(0.14)	(0.005)*	(0.44)	(0.83)	(0.23)
Fu's Fs	1	-1.4	0.84	-1.6	-1.1	1.6	-3.8
(P-value)	(0.71)	(0.31)	(0.7)	(0.048)*	(0.15)	(0.8)	(0.15)
Demographic expansion							
Harpending's Raggedness index (RI)	0.061	0.064	0.055	0.047	0.16	0.086	0.049
P (Simulated RI ≥ Observed RI)	0.89	0.29	0.47	0.79	0.11	0.28	0.51
Sum of Squared deviation (SSD)	0.13	0.014	0.007	0.006	0.022	0.084	0.017
P (Simulated SSD ≥ Observed SSD)	0.004*	0.047*	0.19	0.5	0.16	0.12	0.1
Spatial expansion							
Harpending's Raggedness index (RI)	0.061	0.064	0.055	0.047	0.16	0.086	0.049
P (Simulated RI ≥ Observed RI)	0.73	0.39	0.54	0.78	0.12	0.59	0.74
Sum of Squared deviation (SSD)	0.043	0.014	0.006	0.006	0.022	0.062	0.036
P (Simulated SSD ≥ Observed SSD)	0.23	0.044*	0.22	0.45	0.082	0.13	0.28

*Values are statistically significant at $p < 0.05$; Significance was determined using 1000 coalescent simulations

Abbreviations: D, Tajima's neutrality statistic; *Fs*, Fu's neutrality statistic

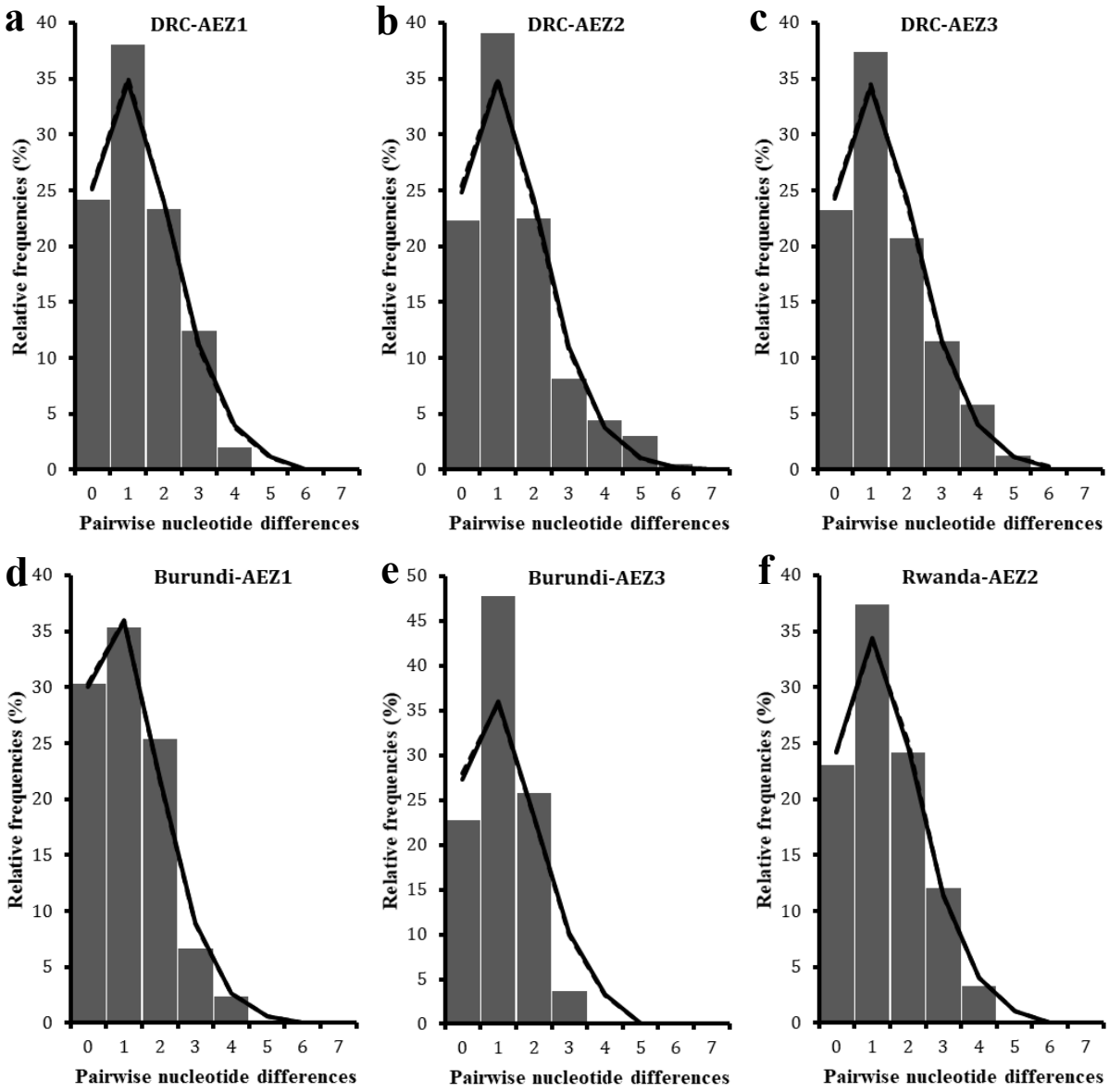

Figure S3.1 *cox*1 mismatch distribution pattern for *R. appendiculatus* haplogroup A in different agro-ecological zones. The x-axis shows the number of pairwise differences between pairs of haplotype sequences and the y-axis shows their frequencies. The observed frequencies are represented by solid histograms and the simulated mismatch distributions expected under demographic expansion (solid black line) and under spatial expansion (dotted black line)

Table S3.6 *Rhipicephalus appendiculatus 12S* rRNA haplotypes and their distribution across agro-ecological zones of the Great Lakes region and other sub-Saharan African countries

Haplotype	Haplotypes from GenBank: Country (original haplotype name and GenBank number)	This study	Haplogroup
12SH1	Kenya (H2: KX276946, H5: KX276949) [1], Rwanda (H5: DQ901279, H5: DQ901281, H5: DQ901282, H5: DQ901284)[2], Zambia-east (H5: DQ849210, DQ901288)[2]	Burundi (AEZ1, AEZ3), DRC AEZ1, AEZ2, AEZ3), Rwanda AEZ2	A
12SH2	Kenya (H3: KX276947) [1]	Burundi (AEZ1, AEZ3), DRC AEZ1, AEZ2, AEZ3), Rwanda AEZ2	A
12SH3	-	Burundi AEZ1	A
12SH4	-	Burundi AEZ1, DRC AEZ1, Rwanda AEZ2	B
12SH5	Zimbabwe (AF031859, AF150027)[3], Grande Comore (H1: DQ901317) [2], Zambia-south (H1: DQ849203, H1: DQ849205, H1: DQ901311, H1: DQ849208, H1: DQ849204, H1: DQ901309) [2], Zambia-East (H1: DQ849207) [2], South Africa (H1: DQ849233, H1: DQ849235) [2], Kenya (DQ901320[2], H1: KX276945[1]),	DRC (AEZ1, AEZ3), Rwanda AEZ2	B
12SH6	-	DRC AEZ3	A
12SH7	-	DRC AEZ3	A
12SH8	-	DRC AEZ3	A
12SH9	-	DRC AEZ3	A
12SH10	Uganda (AF150028)[4]	-	A
12SH11	Zambia-east (H2: DQ901277, H2: DQ849214) [2]	-	A
12SH12	Zambia-east (H3: DQ849212) [2]	-	A
12SH13	Rwanda (H4: DQ901286) [2]	-	A
12SH14	Kenya (H4: KX276948) [1]	-	B

[1][31]; [2][29]; [3][73], [4][74]

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

[31]. Kanduma EG, Mwacharo JM, Githaka NW, Kinyanjui PW, Njuguna JN, Kamau LM, et al. Analyses of mitochondrial genes reveal two sympatric but genetically divergent lineages of *Rhipicephalus appendiculatus* in Kenya. Parasit Vectors. 2016;9:353.

[73]. Murrell A, Campbell NJH, Barker SC. Phylogenetic Analyses of the Rhipicephaline Ticks Indicate That the Genus Rhipicephalus Is Paraphyletic. Mol Phylogenet Evol. 2000;16:1-7.

[74]. Beati L, Keirans JE. 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. J Parasitol. 2001;87:32-48.

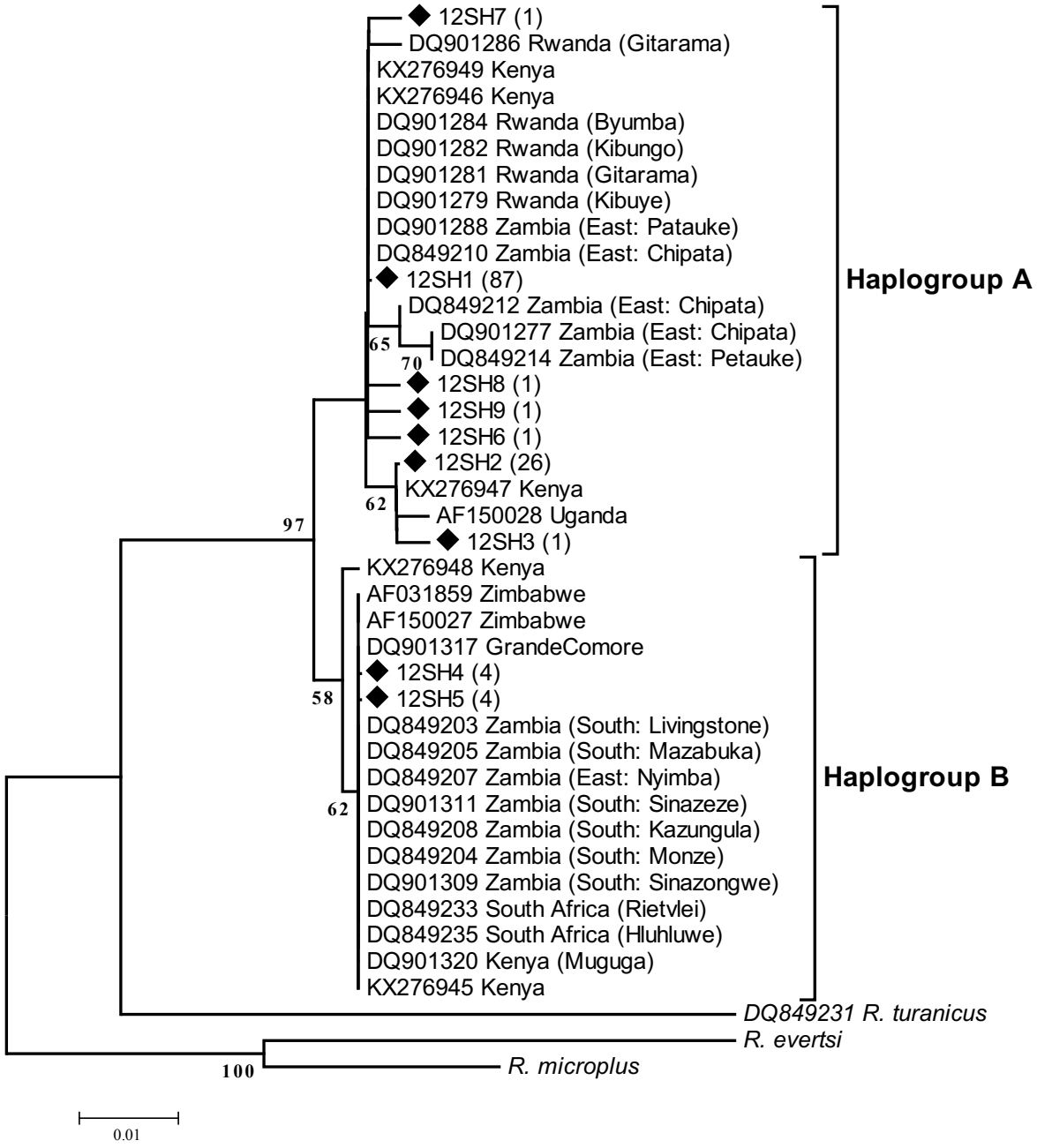

Figure S3.2 Neighbor-joining tree of *12S* haplotype sequences for *R. appendiculatus* across African countries. The evolutionary distances were computed using the Tamura 3-parameter method. Bootstrap values (>60) are displayed above nodes. The values in bracket behind haplotypes names correspond to the frequency of each haplotype. Haplotype sequences (12SH1-9) obtained in the present study are indicated by a black square. *Rhipicephalus eversti* and *R. microplus* obtained in the present study and *R. turanicus* from GenBank (accession number: DQ849231) were used as outgroups.

Appendices (Chapter 4)

Table S4.1 Cattle blood sample distribution across agro-ecological zones

Country	Agro-ecological zone (AEZ)	Blood sample (no. of cattle)
DRC	AEZ1	110
	AEZ2	114
	AEZ3	130
Burundi	AEZ1	126
Total		480

DRC (AEZ1: Lowlands, AEZ2: Midlands, AEZ3: Highlands); Burundi (AEZ1: Lowlands)

Table S4.2 Nucleotide and amino acid sequences of *Tp1* and *Tp2* antigen epitopes from *T. parva* Muguga reference sequence

Gene locus	Epitope positions with inner primers used in the present study			Published epitopes using outer primers (GenBank JF451936 and JF451856 for Tp1 and Tp2, respectively) [34, 36]
	Epitope	Amino acid sequence*	Nucleotide sequence*	
Tp1	$Tp1_{35-45}$	^{35}VGYPKVKEEML^{45}	^{104}GTAGGGTATCCAAAGGTTAAAGAAGAAATGCTA^{136}	$Tp1_{214-224}$
Tp2	$Tp2_{20-30}$	^{20}SHEELKKLGML^{30}	^{58}AGTCATGAAGAACTAAAAAAATTGGGAATGCTA^{90}	$Tp2_{27–37}$
	$Tp2_{33-41}$	^{33}DGFDRDALF^{41}	^{97}GATGGTTTCGACAGGGATGCATTGTTC^{123}	$Tp2_{40–48}$
	$Tp2_{42-52}$	^{42}KSSHGMGKVGK^{52}	^{124}AAATCATCACATGGTATGGGAAAGGTAGGAAAA^{156}	$Tp2_{49–59}$
	$Tp2_{89-97}$	^{89}FAQSLVCVL^{97}	^{265}TTTGCACAAAGCCTAGTGTGCGTATTA^{291}	$Tp2_{96–104}$
	$Tp2_{91-99}$	^{91}QSLVCVLMK^{99}	^{271}CAAAGCCTAGTGTGCGTATTAATGAAA^{297}	$Tp2_{98–106}$
	$Tp2_{131-140}$	^{131}KTSIPNPCKW^{140}	^{391}AAAACAAGTATTCCAAATCCATGTAAATGG^{420}	$Tp2_{138–147}$

*The numbers flanking epitope sequences represent their positions in the *Tp1* and *Tp2* antigen gene fragment

34. Graham SP, Pelle R, Yamage M, Mwangi DM, Honda Y, Mwakubambanya RS, et al. Characterization of the fine specificity of bovine CD8 T-cell responses to defined antigens from the protozoan parasite *Theileria parva*. Infect Immun. 2008;76(2):685-94.

36. Pelle R, Graham SP, Njahira MN, Osaso J, Saya RM, Odongo DO, et al. Two *Theileria parva* CD8 T cell antigen genes are more variable in buffalo than cattle parasites, but differ in pattern of sequence diversity. PLoS One. 2011;6(4):e19015.

Table S4.3 Characteristics of 119 *T. parva* samples obtained from cattle in different agro-ecological zones (AEZs) of The Democratic Republic of Congo and Burundi

N	Sample ID	Sample origin			*Tp1* locus				*Tp2* locus			
		District	Country	AEZ[a]	Gene allele	Antigen variant	GenBank No. [b]	Cluster	Gene allele	Antigen variant	GenBank No. [b]	Cluster
1	B01	Rugombo	Burundi	1	4	3	JF451973	1	1	1	JF451856	1A
2	B02	Rugombo	Burundi	1	1	1	JF451936	1	58	55	MF449297	1B
3	B03	Rugombo	Burundi	1	1	1	JF451936	1	57	54	MF449296	2A
4	B04	Rugombo	Burundi	1	nd	nd	nd	nd	58	55	MF449297	1B
5	B05	Rugombo	Burundi	1	4	3	JF451973	1	2	2	JF451880	1A
6	B06	Rugombo	Burundi	1	37	31	KJ566597	1	2	2	JF451880	1A
7	B07	Rugombo	Burundi	1	37	31	KJ566597	1	59	54	MF449298	2A
8	B08	Rugombo	Burundi	1	1	1	JF451936	1	2	2	JF451880	1A
9	B09	Rugombo	Burundi	1	1	1	JF451936	1	2	2	JF451880	1A
10	B10	Rugombo	Burundi	1	1	1	JF451936	1	1	1	JF451856	1A
11	B11	Rugombo	Burundi	1	37	31	KJ566597	1	1	1	JF451856	1A
12	B12	Rugombo	Burundi	1	37	31	KJ566597	1	57	54	MF449296	2A
13	B13	Rugombo	Burundi	1	1	1	JF451936	1	2	2	JF451880	1A
14	B14	Rugombo	Burundi	1	1	1	JF451936	1	2	2	JF451880	1A
15	B15	Rugombo	Burundi	1	1	1	JF451936	1	2	2	JF451880	1A
16	B16	Rugombo	Burundi	1	4	3	JF451973	1	56	53	MF449295	2B
17	B17	Rugombo	Burundi	1	37	31	KJ566597	1	57	54	MF449296	2A
18	B18	Rugombo	Burundi	1	4	3	JF451973	1	2	2	JF451880	1A
19	B19	Rugombo	Burundi	1	1	1	JF451936	1	2	2	JF451880	1A
20	B20	Gihanga	Burundi	1	1	1	JF451936	1	56	53	MF449295	2B
21	B21	Gihanga	Burundi	1	1	1	JF451936	1	1	1	JF451856	1A
22	B22	Gihanga	Burundi	1	1	1	JF451936	1	2	2	JF451880	1A
23	B23	Gihanga	Burundi	1	1	1	JF451936	1	2	2	JF451880	1A
24	B24	Gihanga	Burundi	1	37	31	KJ566597	1	57	54	MF449296	2A
25	B25	Gihanga	Burundi	1	37	31	KJ566597	1	57	54	MF449296	2A
26	B26	Gihanga	Burundi	1	37	31	KJ566597	1	57	54	MF449296	2A
27	B27	Gihanga	Burundi	1	37	31	KJ566597	1	57	54	MF449296	2A
28	B28	Gihanga	Burundi	1	1	1	JF451936	1	1	1	JF451856	1A
29	B29	Gihanga	Burundi	1	1	1	JF451936	1	nd	nd	nd	nd
30	B30	Gihanga	Burundi	1	1	1	JF451936	1	nd	nd	nd	nd
31	B31	Gihanga	Burundi	1	1	1	JF451936	1	nd	nd	nd	nd

32	B32	Gihanga	Burundi	1	1	1	JF451936	1		nd	nd	nd	nd
33	B33	Gihanga	Burundi	1	37	31	KJ566597	1		nd	nd	nd	nd
34	B34	Gihanga	Burundi	1	37	31	KJ566597	1		nd	nd	nd	nd
35	K01	Kabare	DRC	3	1	1	JF451936	1		1	1	JF451856	1A
36	K02	Kabare	DRC	3	1	1	JF451936	1		2	2	JF451880	1A
37	K03	Kabare	DRC	3	1	1	JF451936	1		2	2	JF451880	1A
38	K04	Kabare	DRC	3	1	1	JF451936	1		1	1	JF451856	1A
39	K05	Kabare	DRC	3	1	1	JF451936	1		1	1	JF451856	1A
40	K06	Kabare	DRC	3	1	1	JF451936	1		1	1	JF451856	1A
41	K07	Kabare	DRC	3	1	1	JF451936	1		1	1	JF451856	1A
42	K08	Kabare	DRC	3	1	1	JF451936	1		1	1	JF451856	1A
43	K09	Kabare	DRC	3	1	1	JF451936	1		2	2	JF451880	1A
44	K10	Kabare	DRC	3	1	1	JF451936	1		1	1	JF451856	1A
45	K11	Kabare	DRC	3	1	1	JF451936	1		2	2	JF451880	1A
46	K12	Kabare	DRC	3	1	1	JF451936	1		2	2	JF451880	1A
47	K13	Kabare	DRC	3	nd	nd	nd	nd		1	1	JF451856	1A
48	K14	Kabare	DRC	3	1	1	JF451936	1		1	1	JF451856	1A
49	K15	Kabare	DRC	3	1	1	JF451936	1		1	1	JF451856	1A
50	K16	Kabare	DRC	3	1	1	JF451936	1		1	1	JF451856	1A
51	K17	Kabare	DRC	3	1	1	JF451936	1		1	1	JF451856	1A
52	K18	Kabare	DRC	3	1	1	JF451936	1		1	1	JF451856	1A
53	K19	Kabare	DRC	3	1	1	JF451936	1		2	2	JF451880	1A
54	K20	Kabare	DRC	3	1	1	JF451936	1		1	1	JF451856	1A
55	K21	Kabare	DRC	3	1	1	JF451936	1		1	1	JF451856	1A
56	K22	Kabare	DRC	3	1	1	JF451936	1		1	1	JF451856	1A
57	K23	Kabare	DRC	3	1	1	JF451936	1		1	1	JF451856	1A
58	K24	Kabare	DRC	3	1	1	JF451936	1		nd	nd	nd	nd
59	K25	Kabare	DRC	3	1	1	JF451936	1		nd	nd	nd	nd
60	K26	Kabare	DRC	3	1	1	JF451936	1		nd	nd	nd	nd
61	U01	Uvira	DRC	1	1	1	JF451936	1		1	1	JF451856	1A
62	U02	Uvira	DRC	1	1	1	JF451936	1		1	1	JF451856	1A
63	U03	Uvira	DRC	1	1	1	JF451936	1		56	53	MF449295	2B
64	U04	Uvira	DRC	1	1	1	JF451936	1		1	1	JF451856	1A
65	U05	Uvira	DRC	1	1	1	JF451936	1		1	1	JF451856	1A
66	U06	Uvira	DRC	1	1	1	JF451936	1		1	1	JF451856	1A
67	U07	Uvira	DRC	1	37	31	KJ566597	1		1	1	JF451856	1A
68	U08	Uvira	DRC	1	37	31	KJ566597	1		1	1	JF451856	1A

69	U09	Uvira	DRC	1	1	1	JF451936	1		1	1	JF451856	1A
70	U10	Uvira	DRC	1	1	1	JF451936	1		1	1	JF451856	1A
71	U11	Uvira	DRC	1	1	1	JF451936	1		1	1	JF451856	1A
72	U12	Uvira	DRC	1	4	3	JF451973	1		56	53	MF449295	2B
73	U13	Uvira	DRC	1	4	3	JF451973	1		56	53	MF449295	2B
74	U14	Uvira	DRC	1	4	3	JF451973	1		2	2	JF451880	1A
75	U15	Uvira	DRC	1	1	1	JF451936	1		1	1	JF451856	1A
76	U16	Uvira	DRC	1	1	1	JF451936	1		1	1	JF451856	1A
77	U17	Uvira	DRC	1	1	1	JF451936	1		1	1	JF451856	1A
78	U18	Uvira	DRC	1	37	31	KJ566597	1		57	54	MF449296	2A
79	U19	Uvira	DRC	1	1	1	JF451936	1		2	2	JF451880	1A
80	U20	Uvira	DRC	1	4	3	JF451973	1		62	58	MF449301	1A
81	U21	Uvira	DRC	1	1	1	JF451936	1		1	1	JF451856	1A
82	U22	Uvira	DRC	1	4	3	JF451973	1		56	53	MF449295	2B
83	U23	Uvira	DRC	1	37	31	KJ566597	1		56	53	MF449295	2B
84	U24	Uvira	DRC	1	45	32	MF449290	2		60	56	MF449299	1B
85	U25	Uvira	DRC	1	45	32	MF449290	2		61	57	MF449300	1B
86	U26	Uvira	DRC	1	43	1	MF449288	1		nd	nd	nd	nd
87	U27	Uvira	DRC	1	1	1	JF451936	1		nd	nd	nd	nd
88	U28	Uvira	DRC	1	1	1	JF451936	1		nd	nd	nd	nd
89	U29	Uvira	DRC	1	1	1	JF451936	1		nd	nd	nd	nd
90	U30	Uvira	DRC	1	44	3	MF449289	1		nd	nd	nd	nd
91	U31	Uvira	DRC	1	46	33	MF449291	1		nd	nd	nd	nd
92	W01	Walungu	DRC	2	1	1	JF451936	1		1	1	JF451856	1A
93	W02	Walungu	DRC	2	1	1	JF451936	1		56	53	MF449295	2B
94	W03	Walungu	DRC	2	39	1	KJ566599	1		1	1	JF451856	1A
95	W04	Walungu	DRC	2	1	1	JF451936	1		2	2	JF451880	1A
96	W05	Walungu	DRC	2	4	3	JF451973	1		56	53	MF449295	2B
97	W06	Walungu	DRC	2	37	31	KJ566597	1		56	53	MF449295	2B
98	W07	Walungu	DRC	2	37	31	KJ566597	1		56	53	MF449295	2B
99	W08	Walungu	DRC	2	1	1	JF451936	1		56	53	MF449295	2B
100	W09	Walungu	DRC	2	1	1	JF451936	1		56	53	MF449295	2B
101	W10	Walungu	DRC	2	1	1	JF451936	1		56	53	MF449295	2B
102	W11	Walungu	DRC	2	37	31	KJ566597	1		56	53	MF449295	2B
103	W12	Walungu	DRC	2	1	1	JF451936	1		2	2	JF451880	1A
104	W13	Walungu	DRC	2	nd	nd	nd	nd		1	1	JF451856	1A
105	W14	Walungu	DRC	2	45	32	MF449290	2		63	59	MF449302	1B

106	W15	Walungu	DRC	2	45	32	MF449290	2		63	59	MF449302	1B
107	W16	Walungu	DRC	2	1	1	JF451936	1		57	54	MF449296	2A
108	W17	Walungu	DRC	2	1	1	JF451936	1		57	54	MF449296	2A
109	W18	Walungu	DRC	2	1	1	JF451936	1		57	54	MF449296	2A
110	W19	Walungu	DRC	2	1	1	JF451936	1		2	2	JF451880	1A
111	W20	Walungu	DRC	2	39	1	KJ566599	1		57	54	MF449296	2A
112	W21	Walungu	DRC	2	48	34	MF449293	1		nd	nd	nd	nd
113	W22	Walungu	DRC	2	47	1	MF449292	1		nd	nd	nd	nd
114	W23	Walungu	DRC	2	1	1	JF451936	1		nd	nd	nd	nd
115	W24	Walungu	DRC	2	1	1	JF451936	1		nd	nd	nd	nd
116	W25	Walungu	DRC	2	1	1	JF451936	1		nd	nd	nd	nd
117	W26	Walungu	DRC	2	1	1	JF451936	1		nd	nd	nd	nd
118	W27	Walungu	DRC	2	1	1	JF451936	1		nd	nd	nd	nd
119	W28	Walungu	DRC	2	49	33	MF449294	1		nd	nd	nd	nd

Notes: Ninety-six and 116 reliable sequences were obtained for *Tp2* and *Tp1* gene loci respectively; Twenty-three samples failed to be amplified by *Tp2* gene and three samples failed for *Tp1* gene

[a]AEZ-1, lowlands valleys of DRC (Ruzizi valley) and Burundi (Imbo valley); AEZ-2, midlands of DRC in the Walungu district; AEZ-3, highlands of DRC in the Kabare district

[b] GenBank accession numbers with JF and KJ prefixes were previously described in Kenya (Pelle et al., 2011) and South Sudan (Salih et al., 2017), while GenBank accession numbers for newly described alleles are represented by MF prefix.

Abbreviations: AEZ, agro-ecological zones; nd, not determined.

36. Pelle R, Graham SP, Njahira MN, Osaso J, Saya RM, Odongo DO, et al. Two *Theileria parva* CD8 T cell antigen genes are more variable in buffalo than cattle parasites, but differ in pattern of sequence diversity. PLoS One. 2011;6(4):e19015.

37. Salih DA, Pelle R, Mwacharo JM, Njahira MN, Marcellino WL, Kiara H, et al. Genes encoding two *Theileria parva* antigens recognized by CD8+ T-cells exhibit sequence diversity in South Sudanese cattle populations but the majority of alleles are similar to the Muguga component of the live vaccine cocktail. PLoS One. 2017;12(2):e0171426

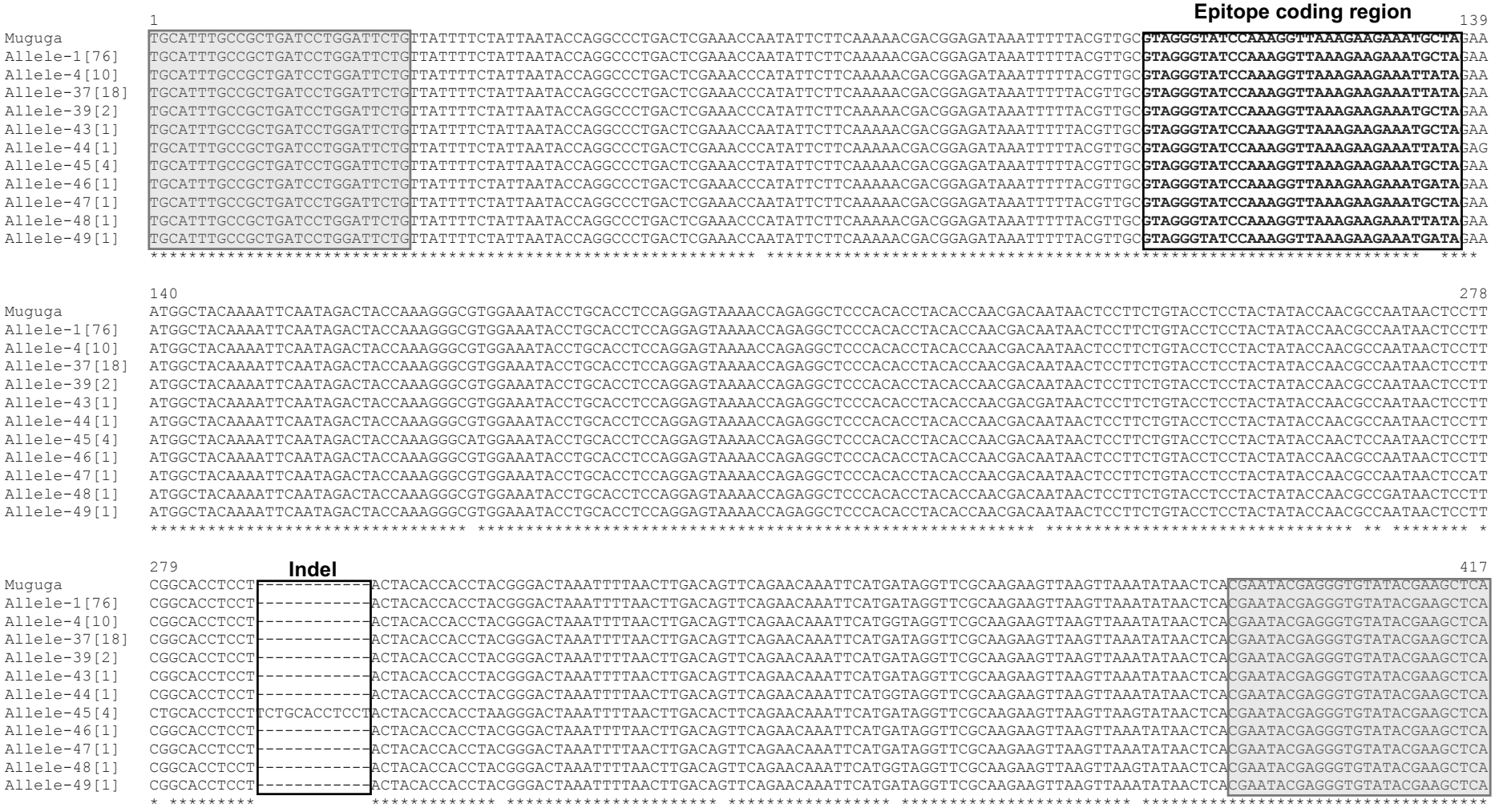

Figure S4.1 Multiple sequence alignment of the 11 *Tp1* gene alleles obtained in this study. The flanked primer regions are shaded and boxed. The CD8+ T cell target epitope coding region and indel insertion are bolded and boxed. *Tp1* allele1 corresponds to samples identical with the three vaccine strains (Muguga, Serengeti-transformed and Kiambu-5).

Table S4.4 Estimates of evolutionary divergence between gene alleles for *Tp1* and *Tp2*, using proportion nucleotide distance

Tp1	Allele-1	Allele-4	Allele-37	Allele-39	Allele-43	Allele-44	Allele-45	Allele-46	Allele-47	Allele-48	Allele-49
Allele-1											
Allele-4	0.01										
Allele-37	0.007	0.002									
Allele-39	0.002	0.007	0.005								
Allele-43	0.002	0.012	0.010	0.005							
Allele-44	0.012	0.002	0.005	0.010	0.015						
Allele-45	0.017	0.025	0.022	0.017	0.020	0.027					
Allele-46	0.005	0.005	0.002	0.002	0.007	0.007	0.020				
Allele-47	0.002	0.012	0.010	0.005	0.005	0.015	0.020	0.007			
Allele-48	0.015	0.005	0.007	0.012	0.017	0.007	0.025	0.010	0.017		
Allele-49	0.002	0.007	0.005	0.005	0.005	0.010	0.020	0.002	0.005	0.012	

Tp2	Allele-1	Allele-2	Allele-56	Allele-57	Allele-58	Allele-59	Allele-60	Allele-61	Allele-62	Allele-63
Allele-1										
Allele-2	0.004									
Allele-56	0.25	0.25								
Allele-57	0.26	0.25	0.075							
Allele-58	0.15	0.15	0.26	0.28						
Allele-59	0.26	0.25	0.077	0.002	0.28					
Allele-60	0.14	0.14	0.24	0.26	0.10	0.27				
Allele-61	0.14	0.14	0.25	0.27	0.10	0.27	0.006			
Allele-62	0.006	0.002	0.25	0.25	0.15	0.25	0.14	0.14		
Allele-63	0.14	0.14	0.24	0.26	0.10	0.27	0.002	0.004	0.14	

Notes: Evolutionary divergence between genes alleles was estimated using proportion nucleotide distance in MEGA.

Tp1 allele-1 corresponds to isolates identical to the three vaccine strains (Muguga, Serengeti-transformed and Kiambu-5).

Tp2 allele-1 corresponds to isolates identical to Muguga and Serengeti-transformed strains and *Tp2* allele-2 represents samples identical to Kiambu-5 strain.

Table S4.5 *Tp1* and *Tp2* genes alleles with their corresponding antigen variants

Marker	Gene alleles	Antigen variants	Number of *T. parva* samples	Frequency (%)
Tp1	Allele-1	Var-1	76	65.5
	Allele-4	Var-3	10	8.6
	Allele-37	Var-31	18	15.5
	Allele-39	Var-1	2	1.7
	Allele-43	Var-1	1	0.9
	Allele-44	Var-3	1	0.9
	Allele-45	Var-32	4	3.4
	Allele-46	Var-33	1	0.9
	Allele-47	Var-1	1	0.9
	Allele-48	Var-34	1	0.9
	Allele-49	Var-33	1	0.9
	Total no. of *Tp1* sequences		116	100.0
Tp2	Allele-1	Var-1	39	40.6
	Allele-2	Var-2	22	22.9
	Allele-56	Var-53	15	15.6
	Allele-57	Var-54	12	12.5
	Allele-58	Var-55	2	2.1
	Allele-59	Var-54	1	1.0
	Allele-60	Var-56	1	1.0
	Allele-61	Var-57	1	1.0
	Allele-62	Var-58	1	1.0
	Allele-63	Var-59	2	2.1
	Total no. of *Tp2* sequences		96	100.0

Table S4.6 Amino acid variants of *Tp1* and *Tp2* $CD8^+$ T cell target epitopes of *T. parva* from DRC and Burundi

Epitope variants	DRC AEZ1 (25)	DRC AEZ2 (20)	DRC AEZ3 (23)	Burundi AEZ1 (28)	Overall (96)
$Tp2_{20\text{-}30}$					
SHEELKKLGML	14	3	17	5	39
SDEELNKLGML	2	3	6	11	22
SDDELDTLGML	5	8	0	2	15
SDNELDTLGLL	1	4	0	8	13
TEEELRKLGMV	2	2	0	0	4
TEEELKKMGMV	0	0	0	2	2
SDEELNILGML	1	0	0	0	1
$Tp2_{33\text{-}41}$					
DGFDRDALF	17	5	23	16	61
PDLDKNRLF	6	13	0	10	29
SNFDRESLF	2	2	0	0	4
EGFDKEKLF	0	0	0	2	2
$Tp2_{42\text{-}52}$					
KSSHGMGKVGK	17	6	23	16	62
LTSHGMGRIGR	5	8	0	2	15
LTSHGMGKIGR	1	4	0	8	13
KSSHGMGKVGR	2	2	0	0	4
KSSKSMGIVGR	0	0	0	2	2
$Tp2_{89\text{-}97}$					
FAQSLVCVL	17	6	23	16	62
FAASIKCVA	5	8	0	2	15
LAASIKCVS	1	4	0	8	13
FAQSILCVI	2	2	0	0	4
FVQSIMCVI	0	0	0	2	2
$Tp2_{91\text{-}99}$					
QSLVCVLMK	17	6	23	16	62
ASIKCVAQY	5	8	0	2	15
ASIKCVSHH	1	4	0	8	13
QSILCVIKN	2	2	0	0	4
QSIMCVINK	0	0	0	2	2
$Tp2_{131\text{-}140}$					
KTSIPNPCKW	17	6	23	16	62
KPSVPNPCDW	6	12	0	10	28
ASDIPNPCKW	2	2	0	0	4
VNDIPNPCKW	0	0	0	2	2
$Tp1_{35\text{-}45}$	**31**	**27**	**25**	**33**	**116**
VGYPKVKEEML	20	21	25	18	84
VGYPKVKEEII	10	5	0	15	30
VGYPKVKEEMI	1	1	0	0	2

Notes: Numbers in bracket are sample size in each AEZ; Epitope sequences in bold are present in the reference sequences of *T. parva* stocks component of the live vaccine.

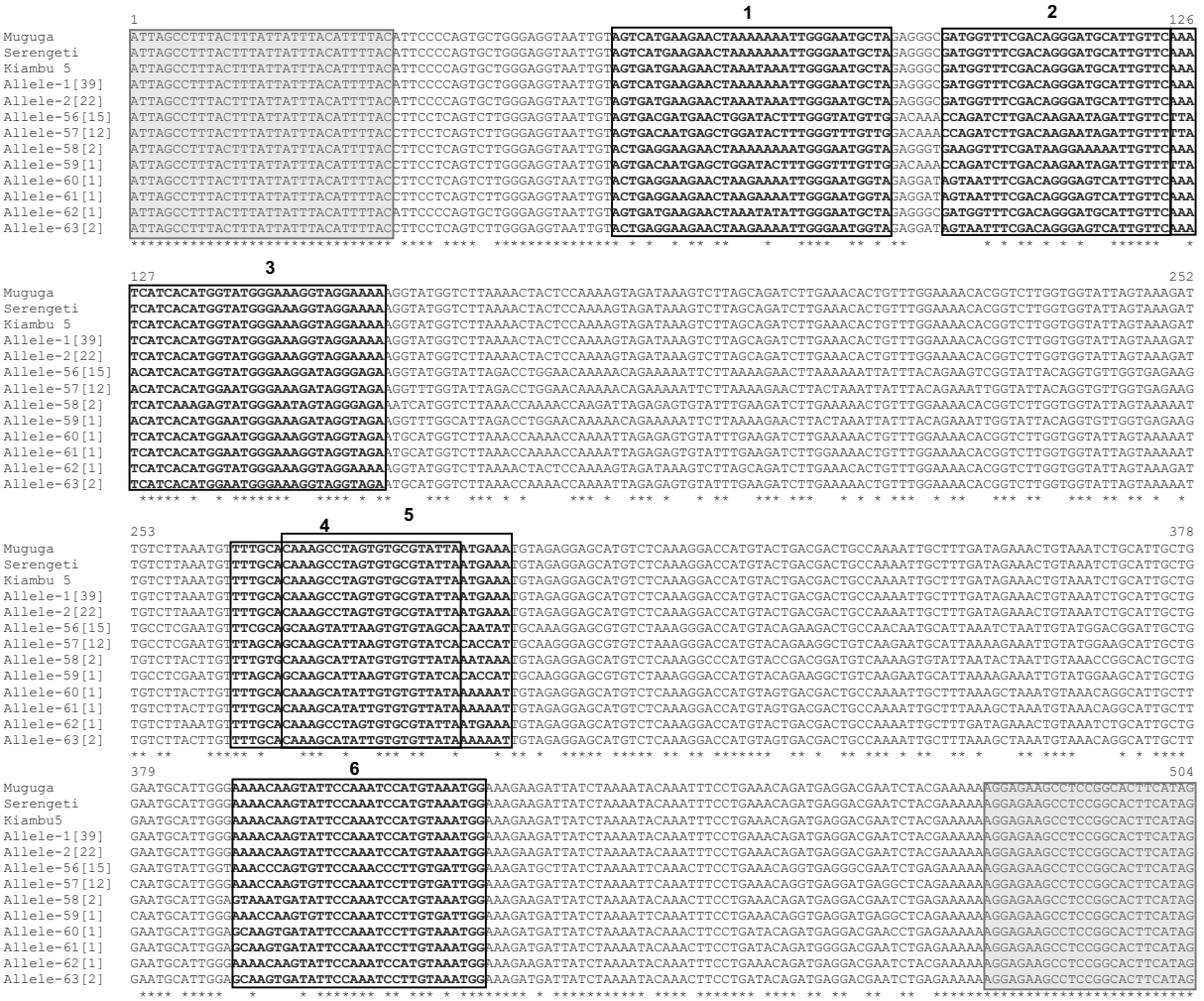

Figure S4.2 Multiple sequence alignment of the 10 *Tp2* gene alleles obtained in this study. The flanked primer regions are shaded and boxed. The $CD8^+$ T cell target epitope coding regions (1-6) are bolded and boxed. *Tp2* allele-1 corresponds to *T. parva* samples identical to Muguga and Serengeti-transformed and *Tp2* allele-2 represents samples similar to Kiambu-5.

Table S4.7 Distribution of *Tp1* gene alleles of *T. parva* from cattle and buffalo in the sub-Saharan region of Africa

***Tp1* alleles**	***Tp1* variants**	**The Great Lakes region**				**Kenya**			**South Sudan**	**Laboratory samples**	**Total**
		DRC AEZ1	DRC AEZ2	DRC AEZ3	Burundi AEZ1	BA	BD	CD			
A01	Var-1	17	16	25	18	1	5	12	53	6	**153**
A02	Var-2	0	0	0	0	0	0	3	0	8	**11**
A03	Var-2	0	0	0	0	0	0	1	0	1	**2**
A04	Var-3	5	1	0	4	0	0	2	14	0	**26**
A05	Var-4	0	0	0	0	0	0	0	0	1	**1**
A06	Var-5	0	0	0	0	0	0	1	0	0	**1**
A07	Var-6	0	0	0	0	0	0	1	0	0	**1**
A08	Var-7	0	0	0	0	0	0	1	0	0	**1**
A09	Var-1	0	0	0	0	0	0	1	0	0	**1**
A10	Var-8	0	0	0	0	0	0	1	0	0	**1**
A11	Var-9	0	0	0	0	0	0	1	0	0	**1**
A12	Var-10	0	0	0	0	0	0	1	0	0	**1**
A13 (2nd indel)	Var-11	0	0	0	0	7	0	0	0	0	**7**
A14	Var-12	0	0	0	0	1	2	0	0	0	**3**
A15 (2nd indel)	Var-13	0	0	0	0	2	0	0	0	0	**2**
A16 (2nd indel)	Var-14	0	0	0	0	0	1	0	0	0	**1**
A17 (1st & 2nd indel)	Var-15	0	0	0	0	0	1	0	0	0	**1**
A18 (2nd indel)	Var-16	0	0	0	0	0	1	0	0	0	**1**
A19 (1st & 2nd indel)	Var-17	0	0	0	0	0	1	0	0	0	**1**
A20	Var-18	0	0	0	0	0	1	0	0	0	**1**
A21 (2nd indel)	Var-19	0	0	0	0	0	1	0	0	0	**1**
A22	Var-18	0	0	0	0	0	1	0	0	0	**1**
A23 (2nd indel)	Var-20	0	0	0	0	0	1	0	0	0	**1**
A24 (1st & 2nd indel)	Var-21	0	0	0	0	0	1	0	0	0	**1**
A25	Var-22	0	0	0	0	1	0	0	0	0	**1**
A26 (2nd indel)	Var-23	0	0	0	0	1	0	0	0	0	**1**
A27 (2nd indel)	Var-14	0	0	0	0	1	0	0	0	0	**1**
A28 (2nd indel)	Var-24	0	0	0	0	1	0	0	0	0	**1**
A29	Var-25	0	0	0	0	1	0	0	0	0	**1**
A30 (2nd indel)	Var-26	0	0	0	0	1	0	0	0	0	**1**
A31 (2nd indel)	Var-13	0	0	0	0	1	0	0	0	0	**1**
A32	Var-27	0	0	0	0	1	0	0	0	0	**1**
A33	Var-28	0	0	0	0	1	0	0	0	0	**1**
A34	Var-29	0	0	0	0	1	0	0	0	0	**1**
A35 (2nd indel)	Var-30	0	0	0	0	1	0	0	0	0	**1**
A36	Var-31	0	0	0	0	0	0	0	4	0	**4**
A37	Var-31	4	3	0	11	0	0	0	3	0	**21**
A38	Var-9	0	0	0	0	0	0	0	1	0	**1**
A39	Var-1	0	2	0	0	0	0	0	1	0	**3**
A40	Var-31	0	0	0	0	0	0	0	1	0	**1**
A41	Var-31	0	0	0	0	0	0	0	1	0	**1**
A42	Var-1	0	0	0	0	0	0	0	1	0	**1**
A43	Var-1	1	0	0	0	0	0	0	0	0	**1**
A44	Var-3	1	0	0	0	0	0	0	0	0	**1**
A45 (2nd indel)	Var-32	2	2	0	0	0	0	0	0	0	**4**
A46	Var-33	1	0	0	0	0	0	0	0	0	**1**
A47	Var-1	0	1	0	0	0	0	0	0	0	**1**
A48	Var-34	0	1	0	0	0	0	0	0	0	**1**
A49	Var-33	0	1	0	0	0	0	0	0	0	**1**
Total		**31**	**27**	**25**	**33**	**22**	**16**	**25**	**79**	**16**	**274**

Notes: *Tp1* allele A01 corresponds to *T. parva* alleles identical to the one in the three vaccine strains (Muguga, Serengeti-transformed and Kiambu-5). *Abbreviations*: BD, buffalo-derived; BA, buffalo-associated; CD, cattle derived

Table S4.8 Distribution of *Tp2* gene alleles of *T. parva* from cattle and buffalo in the sub-Saharan region of Africa

***Tp2* alleles**	**The Great Lakes region**				**Kenya**			**South Sudan**	**Laboratory samples**	**Total**
	DRC AEZ1	DRC AEZ2	DRC AEZ3	Burundi AEZ1	BA	BD	CD			
A01	14	3	17	5	0	0	15	43	9	**106**
A02	2	3	6	11	0	0	3	7	1	**33**
A03	0	0	0	0	0	0	2	0	6	**8**
A04	0	0	0	0	0	0	2	0	0	**2**
A05	0	0	0	0	0	0	0	1	1	**2**
A06 & 28	0	0	0	0	4	0	0	0	0	**4**
A07	0	0	0	0	0	2	0	0	0	**2**
A08	0	0	0	0	0	1	0	0	0	**1**
A09	0	0	0	0	0	1	0	0	0	**1**
A10	0	0	0	0	0	1	0	0	0	**1**
A11	0	0	0	0	0	1	0	0	0	**1**
A12	0	0	0	0	0	1	0	0	0	**1**
A13	0	0	0	0	0	1	0	0	0	**1**
A14	0	0	0	0	0	1	0	0	0	**1**
A15	0	0	0	0	0	1	0	0	0	**1**
A16	0	0	0	0	0	1	0	0	0	**1**
A17	0	0	0	0	0	1	0	0	0	**1**
A18	0	0	0	0	0	1	0	0	0	**1**
A19	0	0	0	0	0	1	0	0	0	**1**
A20	0	0	0	0	0	1	0	0	0	**1**
A21	0	0	0	0	0	1	0	0	0	**1**
A22	0	0	0	0	1	0	0	0	0	**1**
A23 & 24	0	0	0	0	2	0	0	0	0	**2**
A25	0	0	0	0	1	0	0	0	0	**1**
A26	0	0	0	0	1	0	0	0	0	**1**
A27	0	0	0	0	1	0	0	0	0	**1**
A29	0	0	0	0	1	0	0	0	0	**1**
A30	0	0	0	0	1	0	0	0	0	**1**
A31	0	0	0	0	1	0	0	0	0	**1**
A32	0	0	0	0	1	0	0	0	0	**1**
A33	0	0	0	0	1	0	0	0	0	**1**
A34	0	0	0	0	1	0	0	0	0	**1**
A35	0	0	0	0	1	0	0	0	0	**1**
A36	0	0	0	0	1	0	0	0	0	**1**
A37	0	0	0	0	1	0	0	0	0	**1**
A38	0	0	0	0	1	0	0	0	0	**1**
A39	0	0	0	0	1	0	0	0	0	**1**
A40	0	0	0	0	1	0	0	0	0	**1**
A41	0	0	0	0	1	0	0	0	0	**1**
A42	0	0	0	0	1	0	0	0	0	**1**
A43	0	0	0	0	1	0	0	0	0	**1**
A44	0	0	0	0	0	0	0	2	0	**2**
A45	0	0	0	0	0	0	0	2	0	**2**
A46	0	0	0	0	0	0	0	1	0	**1**
A47	0	0	0	0	0	0	0	1	0	**1**
A48	0	0	0	0	0	0	0	1	0	**1**
A49	0	0	0	0	0	0	0	1	0	**1**

A50	0	0	0	0	0	0	0	1	0	**1**
A51	0	0	0	0	0	0	0	1	0	**1**
A52	0	0	0	0	0	0	0	1	0	**1**
A53	0	0	0	0	0	0	0	1	0	**1**
A54	0	0	0	0	0	0	0	1	0	**1**
A55	0	0	0	0	0	0	0	1	0	**1**
A56	5	8	0	2	0	0	0	0	0	**15**
A57	1	4	0	7	0	0	0	0	0	**12**
A58	0	0	0	2	0	0	0	0	0	**2**
A59	0	0	0	1	0	0	0	0	0	**1**
A60	1	0	0	0	0	0	0	0	0	**1**
A61	1	0	0	0	0	0	0	0	0	**1**
A62	1	0	0	0	0	0	0	0	0	**1**
A63	0	2	0	0	0	0	0	0	0	**2**
Total	**25**	**20**	**23**	**28**	**25**	**16**	**22**	**65**	**17**	**241**

Notes: *Tp2* allele A01 is identical to the one from *T. parva* Muguga and Serengeti-transformed strains and *Tp2* allele A02 represents *T. parva* samples similar to Kiambu-5 strain. *Abbreviations*: BD, buffalo-derived; BA, buffalo-associated; CD, cattle derived

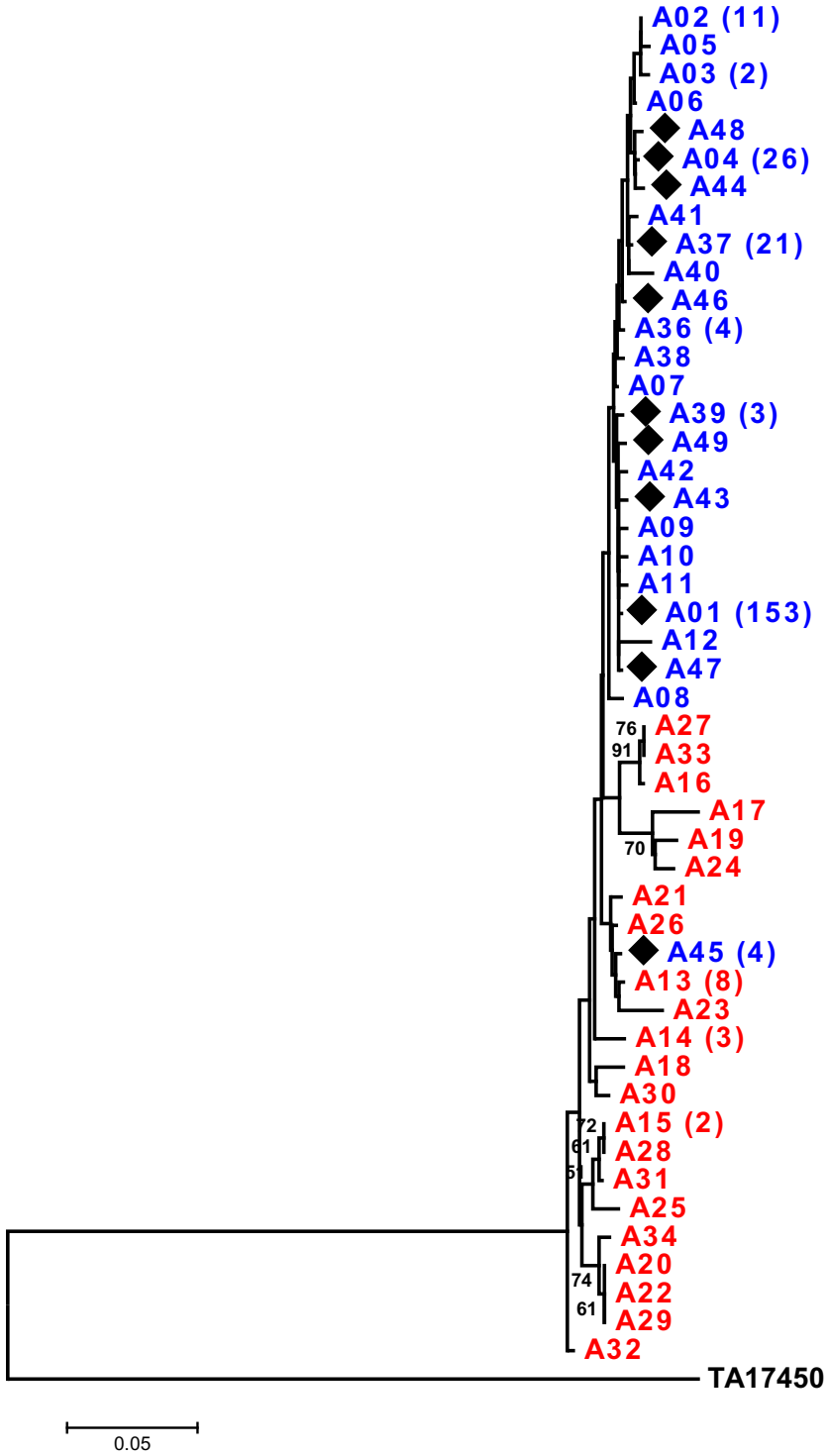

Figure S4.3 Neighbor-Joining tree showing phylogenetic relationships among the 48 *Tp1* gene alleles described in Africa (A01-A49). *Tp1* gene alleles obtained in the present study are indicated by black diamonds. *Theileria parva* alleles found in cattle with no association with buffalo and in laboratory stocks are coloured in blue and those from buffalo and buffalo-associated cattle are shown in Red. Bootstrap values (>50%) are shown above branches. The *Tp1* homologous sequence of *Theileria annulata* (GenBank accession no. TA17450) was used as outgroup. The number in brackets behind alleles names denote the number of *T. parva* isolates carrying the allele. The frequencies of *Tp1* alleles and their corresponding populations/AEZs are detailed in Table S7. *Tp1* allele A01 corresponds to isolates identical to the three Muguga cocktail vaccine strains (Muguga, Serengeti-transformed and Kiambu-5).

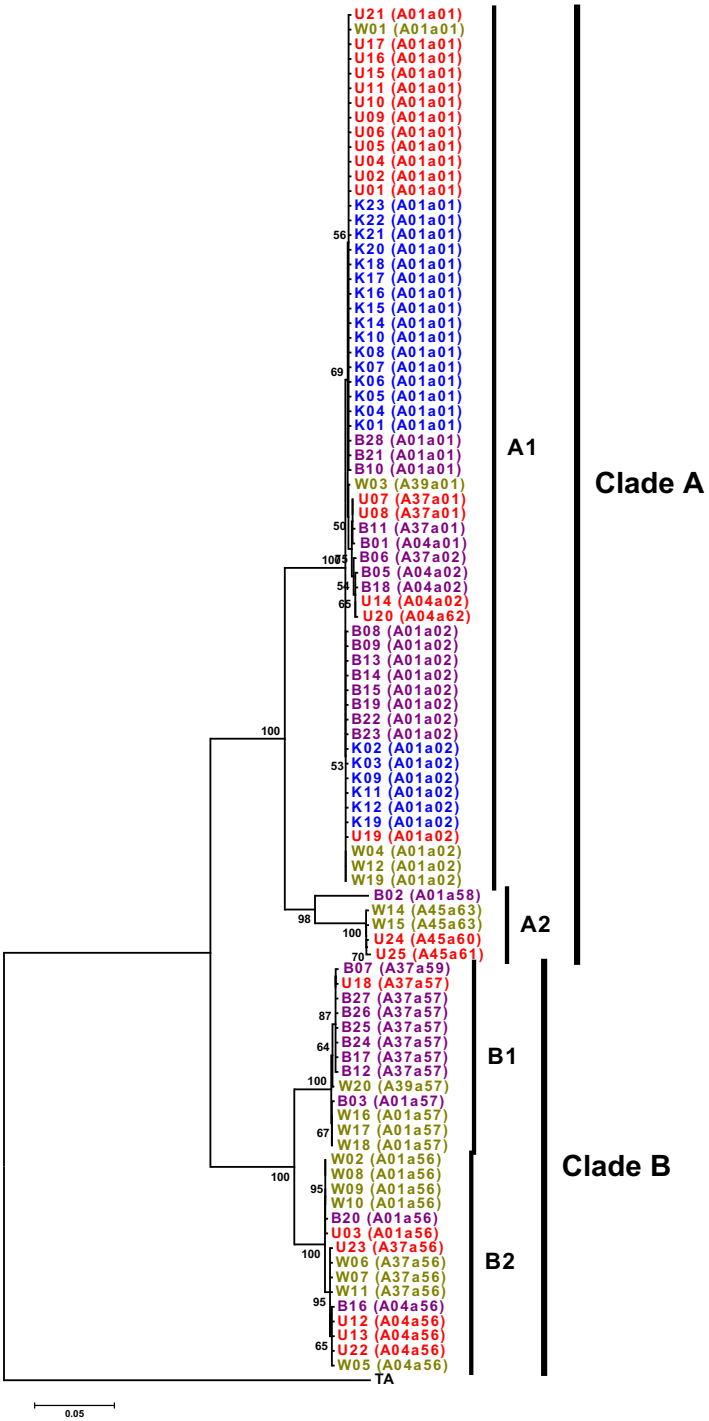

Figure S4.4 Phylogenetic tree showing the relationships among concatenated *Tp1* and *Tp2* nucleotide sequences of 93 *T. parva* samples from cattle in DRC and Burundi. The evolutionary history was constructed using the Neighbor-Joining method with 1000 bootstrap replicates. Bootstrap values (>50%) are shown above branches and indicate the degree of support of each node. The concatenated homologous *Theileria annulata* sequence (GenBank accession no. TA17450 and TA19865 for *Tp1* and *Tp2*, respectively) was used as the outgroup. The codes in brackets behind sample names correspond to their respective concatenated *Tp1* (A) and *Tp2* (a) alleles (Aa). Allele A01a01 corresponds to Muguga and Serengeti-transformed vaccine strains, while allele A02a02 is carried by Kiambu-5 strain. Samples are colour-coded based on their agro-ecological origin (U and Red=DRC AEZ1; W and Green = DRC AEZ2; K and Blue=DRC AEZ3; B and Purple =Burundi AEZ1). Detailed sample characteristics and corresponding gene alleles and protein variants are shown in Table S3.

Appendices (Chapter 5)

Table S5.1 Predictions and confidence intervals (95%) for tick abundance on cattle and the vectorial inoculation rate (VIR) according to AEZs and seasons

AEZ	Season	Mean tick burden on cattle[a]			Vectorial inoculation rate (VIR)[b]		
		Prediction	CI (95%)		Prediction	CI (95%)	
			lower	upper		lower	upper
Lowlands	Wet1	29.2	24.3	35.0	1.9	0.93	3.5
	Wet2	38.7	31.1	48.0	3.2	1.8	5.2
	Dry	19.1	15.3	23.7	0.44	0.11	1.6
Midlands	Wet1	39.1	32.5	47.1	1.9	0.89	3.9
	Wet2	42.3	33.7	53.1	3.6	2.01	6.1
	Dry	17.5	13.7	22.3	0.56	0.17	1.6
Highlands	Wet1	18.7	15.3	22.8	0.57	0.26	1.2
	Wet2	23.1	17.9	29.5	0.66	0.26	1.5
	Dry	6.9	5.5	8.5	0.14	0.02	0.83

[a]Expressed as the mean number of ticks per animal
[b]VIR expressed as the number of infective ticks feeding on one animal during the feeding period of one week, which is used as the transmission intensity of *T. parva*
Notes: AEZ, agro-ecological zones (AEZ1: lowlands, AEZ2: midlands and AEZ3: highlands). Seasons: Wet1, early wet season (September-December); Wet2, late wet season (February-May); Dry, dry season (June-August)
Abbreviations: IC confident intervals

Table S5.2 Predictions and confidence intervals (95%) for *T. parva* prevalence in cattle blood and in free-living ticks according to AEZs and seasons

AEZ	Season	***T. parva* prevalence in cattle blood**			***T. parva* prevalence in pooled ticks**[a]			***T. parva* prevalence in individual ticks**[b]		
		Prediction	CI (95%)		Prediction	CI (95%)		Prediction	CI (95%)	
			lower	upper		lower	upper		lower	upper
Lowlands	Wet1	0.53	0.47	0.59	0.28	0.15	0.46	0.064	0.033	0.12
	Wet2	0.48	0.41	0.55	0.35	0.22	0.49	0.082	0.05	0.13
	Dry	0.43	0.36	0.51	0.11	0.028	0.35	0.023	0.006	0.084
Midlands	Wet1	0.3	0.23	0.38	0.23	0.11	0.4	0.05	0.023	0.098
	Wet2	0.51	0.41	0.63	0.36	0.22	0.51	0.085	0.05	0.14
	Dry	0.24	0.15	0.36	0.15	0.049	0.38	0.032	0.01	0.09
Highlands	Wet1	0.19	0.15	0.24	0.14	0.07	0.27	0.03	0.014	0.061
	Wet2	0.24	0.15	0.36	0.13	0.057	0.29	0.029	0.012	0.065
	Dry	0.1	0.065	0.15	0.1	0.014	0.47	0.021	0.003	0.12

[a]Ticks were grouped in pools of 5 ticks for *T. parva* DNA detection

[b]The prevalence in individual ticks was simulated using the probability formula applied on pooled prevalence

Notes: AEZ, agro-ecological zones (AEZ1: lowlands, AEZ2: midlands and AEZ3: highlands). Seasons: Wet1, early wet season (September-December); Wet2, late wet season (February-May); Dry, dry season (June-August).

Abbreviations: IC confident intervals

Table S5.3 Summary of the logistic regression model showing the association between *T. parva* infection rate in free-living ticks and ecological variables (AEZs and seasons)

Factor	Level	Odds ratio (OR)	CI (95%)		*P-value*	DEFT[a]
AEZ	Lowlands	Ref[b]				
	Midlands	1.4	0.21	9.6	0.72	0.78
	Highlands	0.88	0.07	11.2	0.92	0.77
Season	Dry	Ref[b]				
	Wet1	3.1	0.59	16.5	0.17	0.53
	Wet2	4.3	0.86	20.8	0.07	0.71
Interaction	Midlands & Wet1	0.52	0.05	4.9	0.57	0.62
	Midlands & Wet2	0.74	0.08	6.1	0.78	0.69
	Highlands & Wet1	0.47	0.03	7.6	0.6	0.39
	Highlands & Wet2	0.32	0.02	5.3	0.43	0.55

Notes: AEZ, agro-ecological zones; Seasons: Wet1, early wet season (September-December); Wet2, late wet season (February-May); Dry, dry season (June-August). The model was multivariate, including AEZs (three levels), seasons (three levels) and their interaction. [a]DEFT, design effects computed using sample sites (villages) as subpopulation clusters assuming intra-village correlation. [b]Ref., reference level: lowlands and dry season were used as reference for comparison in the model (P=0.05). CI, confident intervals

Appendices (Chapter 6)

Table S6.1 Linkage disequilibrium of *T. parva* populations from cattle in different geographic locations of Africa based on microsatellite markers

Country	Region/AEZ	N	N_L	Linkage equilibrium/disequilibrium			
				I_A^S	V_D	L_{MC}	*Structure*
South Sudan (Salih et al., 2018)	Kajo Keji	41	14	0.0544	2.7536	2.2393	LD
	Juba	45	14	0.0436	2.1546	1.8188	LD
	Yei	30	14	0.0150	2.1309	2.3414	LE
	Bor	62	14	0.0295	1.9931	1.7692	LD
	Overall	178	14	0.0499	2.1846	1.6119	LD
Tanzania (Elisa et al., 2015)	Eastern	21	14	-0.0016	2.876	3·4409	LE
	Southern	14	14	0.0211	3.269	4·2228	LE
	Overall	35	14	0.0066	3.267	3·7396	LE
Uganda (Oura et al., 2005)	Western (Mbarara)	34	12	0.0679	4.09	2.54	LD
	Northern (Lira)	40	12	0.0164	2.80	2.63	LD
	Central (Kayunga)	30	12	0.1439	5.87	2.59	LD
	Overall	104	12	0.0405	3.10	2.29	LD
Uganda (Muwanika et al., 2016)	Western	30	8	0.1424	2.9373	1.405	LD
	Northern	19	8	0.1647	3.2166	1.700	LD
	Central	11	8	0.1846	3.0424	1.964	LD
	Eastern	18	8	0.1302	2.8625	1.811	LD
	Overall	78	8	0.0749	2.0591	1.886	LD
Zambia (Muleya et al., 2012) [a]	Eastern (Petauke)	28	9	0.027	1.979	1.910	LD
	Northern (Isoka)	33	9	0.104	3.418	2.254	LD
	Overall	61	9	0.078	2.457	1.677	LD

Sample size (N), number of loci (N_L), Standardized index of association (I_A^S); V_D: Mismatch variance which is compared to the variance expected (V_E) for linkage equilibrium (null hypothesis linkage equilibrium: $V_D = V_E$); L_{MC}: 95% confidence limit of Monte-Carlo simulation: The null hypothesis of linkage disequilibrium is rejected when V_D is significantly greater than L.

[a] Northern Zambia: Kanyelele and Kalembe areas (Isoka district), and Eastern Zambia: Saukani area (Petauke district) (Muleya et al., 2012)

LD: linkage disequilibrium; LE: linkage equilibrium

Table S6.2 Parameters of the analysis of genetic diversity and population growth based on Tp2 gene sequences

Genetic indices and statistics	Populations					
	Burundi	**DRC**	**Kenya**	**LS**	**South Sudan**	**Tanzania**
Number of sequences	28	68	22	17	65	71
Number of variable sites	175	166	145	129	79	177
Number of mutations	199	194	153	130	83	208
Number of gene alleles	6	8	4	4	15	27
Haplotype diversity (SD)	0.77 (0.051)	0.69 (0.046)	0.52 (0.12)	0.62 (0.083)	0.55 (0.073)	0.85 (0.029)
Nucleotide diversity (SD)	0.16 (0.017)	0.13 (0.015)	0.08 (0.028)	0.12 (0.019)	0.01 (0.003)	0.13 (0.007)
Number of pairwise nucleotide differences	69.7	57.2	41.5	62.7	4.6	66.5
Tajima's *D*	1.4	1.4	-0.05	2.7**	-2.5***	1.9
Fu and Li's *D**	1.9**	2.1**	1.8**	1.7**	-2.7*	1.5
Fu and Li's *F**	2.05**	2.2**	1.5**	2.3**	-3.1*	1.9*

Table S6.3 Frequency distribution (%) of Tp1 antigen variants from cattle-derived *T. parva* among sub-Saharan African countries

Tp1 antigen variant	Tp1 gene alleles	Burundi[1]	DRC[2]	Kenya (CD)[3]	South Sudan[4]	Tanzania[5]	Overall
Var-1 [AEG42617] [a]	Allele-1, 9, 39, 42, 43, 47	55	75	52	70	64	66.1
Var-2 [AEG42641] [b]	Allele-2, 3,	0	0	16	0	0	1.2
Var-3 [AEG42654]	Allele-4, 44	12	8	8	18	23	16.1
Var-4 [AEG42657]	Allele-5	0	0	4	0	0	0.3
Var-5 [AEG42659]	Allele-6	0	0	4	0	0	0.3
Var-6 [AEG42662]	Allele-7	0	0	4	0	0	0.3
Var-7 [AEG42661]	Allele-8	0	0	4	0	0	0.3
Var-8 [AEG42663]	Allele-10	0	0	4	0	0	0.3
Var-9 [AEG42658]	Allele-11, 38	0	0	4	1	0	0.6
Var-31 [AIA64003]	Allele-36, 37, 40, 41	33	8	0	11	8	10.8
Var-32 [AVM41544]	Allele-45	0	5	0	0	0	1.2
Var-33 [AVM41545]	Allele-46, 49	0	2	0	0	4	2.0
Var-34 [AVM41547]	Allele-48	0	1	0	0	0	0.3
Var-35 [QBL98042]		0	0	0	0	1	0.3
No. of sequences		33	83	25	79	122	342

1: Burundi (Amzati et al., 2019); **2**: RDC(Amzati et al., 2019); **3**: CD Kenya (Pelle et al., 2011), **4**: South Sudan (Salih et al., 2017); **5**: Tanzania (Elisa et al., 2015; Kerario et al., 2019). [a]**Var-1**: Muguga (Kenya: AEG42617), Mariakani (Kenya: AEG42618), Kiambu-5 (Kenya: AEG42620), Serengeti (Tanzania: AEG42621), Nanyuki (Kenya: AEG42622); [b]**Var-2**: Marikebuni (Kenya: AEG42641), Boleni (Zimbabwe: AEG42644), Uganda (AEG42652), Mariakani (Kenya: AEG42646).
Notes: Var-10 (not shown in the table) is only found in the laboratory stock Chitongo (Zambia: AEG42656); Var-11 to var-30 (not shown in the table) are only described in buffalo-derived and buffalo-associated *T. parva* (Pelle et al., 2011; Sitt et al., 2018)

Table S6.4 Frequency distribution (%) of Tp2 antigen variants from cattle-derived *T. parva* among sub-Saharan African countries

Tp2 Antigen variant	Burundi[1]	DRC[2]	Kenya (CD)[3]	South Sudan[4]	Tanzania[5]	Overall
Var-1 [AEG42537] [a]	18	50	68	66	50	51.9
Var-2 [AEG42561] [b]	39	16	14	11	16	17.0
Var-3 [AEG42565] [c]	0	0	9	0	0	0.7
Var-4 [AEG42574]	0	0	9	0	0	0.7
Var-5 [AEG42575] [d]	0	0	0	2	0	0.4
Var-42 [AIA64009]	0	0	0	3	0	0.7
Var-43 [AIA64013]	0	0	0	5	0	1.1
Var-44 [AIA64011]	0	0	0	2	0	0.4
Var-45 [AIA64012]	0	0	0	2	0	0.4
Var-46 [AIA64018]	0	0	0	2	0	0.4
Var-47 [AIA64014]	0	0	0	2	0	0.4
Var-48 [AIA64017]	0	0	0	2	0	0.4
Var-49 [AIA64016]	0	0	0	2	0	0.4
Var-50 [AIA64015]	0	0	0	2	0	0.4
Var-51 [AIA64019]	0	0	0	2	0	0.4
Var-52 [AIA64020]	0	0	0	2	0	0.4
Var-53 [AVM41549]	7	19	0	0	0	5.3
Var-54 [AVM41550]	29	7	0	0	0	4.6
Var-55 [AVM41551]	7	0	0	0	0	0.7
Var-56 [AVM41553]	0	1	0	0	0	0.4
Var-57 [AVM41554]	0	1	0	0	0	0.4
Var-58 [AVM41555]	0	1	0	0	0	0.4
Var-59 [AVM41556]	0	3	0	0	0	0.7
Var-60 [QBL98216]	0	0	0	0	1	0.4
Var-61 [QBL98165]	0	0	0	0	2	0.7
Var-62 [QBL98164]	0	0	0	0	23	8.1
Var-63 [QBL98170]	0	0	0	0	5	1.8
Var-64 [QBL98175]	0	0	0	0	3	1.1
No. of sequences	28	68	22	65	100	283

1: Burundi (Amzati et al., 2019); **2**: RDC (Amzati et al., 2019); **3**: CD Kenya (Pelle et al., 2011), **4**: South Sudan (Salih et al., 2017); **5**: Tanzania (Elisa et al., 2015); [a]**var-1**: Muguga (Kenya: AEG42537), Marikebuni (Kenya: AEG42538), Mariakani (Kenya: AEG42541), Serengeti (Tanzania: AEG42543), Katete (Zambia: AEG42544); [b]**var-2**: Kiambu-5 (Kenya: AEG42561); [c]**var-3**: Chitongo (Zambia: AEG42565), Boleni (Zimbabwe: AEG42566), Uganda (AEG42567), Mariakani (AEG42569); [d]**var-5**: Nanyuki (Kenya: AEG42575).
Notes: Var-6 to var-41 (not shown in the table) are only described in buffalo-derived *T. parva* (Pelle et al., 2011; Sitt et al., 2018)

Table S6.5 CD8+ T-cell epitope variants identified so far in two immunodominant antigens (Tp1 and Tp2) from cattle in Africa

Tp1$_{214–224}$	**Tp2$_{27–37}$**	**Tp2$_{40–48}$**	**Tp2$_{49–59}$**	**Tp2$_{96–104}$**	**Tp2$_{98–106}$**	**Tp2$_{138–147}$**
VGYPKVKEEML (1, 2, 3, 4, 5) [a]	SHEELKKLGML (1, 2, 3, 4, 5) **var-1,5**	DGFDRDALF (1, 2, 3, 4, 5) **var-1,2,5**	KSSHGMGKVGK (1, 2, 3, 4, 5) **var-1,2,5**	FAQSLVCVL (1, 2, 3, 4, 5) **var-1,2,5**	QSLVCVLMK (1, 2, 3, 4, 5) **var-1,2,5**	KTSIPNPCKW (1, 2, 3, 4, 5) **var-1,2,5**
VGYPKVKEEII (1, 2, 3, 4, 5) [b]	SDEELNKLGML (1, 2, 3, 4, 5) **var-2**	PDLDKNRLF (1, 2, 3, 5) **var-3**	LTSHGMGRIGR (1, 2, 3) **var-3**	FAASIKCVA (1, 2, 3, 5) **var-3**	ASIKCVAQY (1, 2, 3, 5) **var-3**	KPSVPNPCDW (1, 2, 3, 5) **var-3**
VGYPKVKEEIL (3, 4, 5)	SDDELDTLGML (1, 2, 3, 5) **var-3**	SNFDRESLF (2)	LTSHGMGKIGR (1, 2, 3, 5)	LAASIKCVS (1, 2, 3, 5)	ASIKCVSHH (1, 2, 3, 5)	ASDIPNPCKW (2)
VGYPKVKEEMI (2, 5)	SDNELDTLGLL (1, 2, 3, 5)	EGFDKEKLF (4)	KSSHGMGKVGR (2)	FAQSILCVI (2)	QSILCVIKN (2)	VNDIPNPCKW (1)
VGYPKVKEEMV (5)	SDEELNILGML (2)	EGFDRETLF (4)	KSSKSMGIVGR (1)	FVQSIMCVI (1)	QSIMCVINK (1)	KNDIPNPCKW (4)
	TEEELRKLGMV (2)	GNFDRELLF (5)	KSSKSMGKVGK (4)	FAQSIMCVL (4)	QSIMCVLKK (4)	KPSIPNPCKW (4)
	TEEELKKMGMV (1)		KSSQSMGKVGK (4)	FAQSLMCVL (4)	QSLMCVLMK (4)	KTCFPNPCKW (4)
	SEEELKKLGML (4)		KSSHGMGKIGR (5)	LAQSIVCVV (5)	QSIVCVVSK (5)	KTDIPNPCKW (4)
	SHEELNILGML (4)					VSDIPNPCKW (5)
	SQEELKKMGML (4)					
	SYEELKKLGML (4)					
	THEELKKMGML (4)					
	SHDGLKKLGML (5)					
	SDDELNKLGML (5)					
	SEAELRKMGMI (5)					

1: Burundi (Amzati et al., 2019); **2**: RDC(Amzati et al., 2019); **3**: CD Kenya (Pelle et al., 2011), **4**: South Sudan (Salih et al., 2017); **5**: Tanzania (Elisa et al., 2015); **var-1**: Muguga (Kenya: AEG42537), Marikebuni (Kenya: AEG42538), Mariakani (Kenya: AEG42541), Serengeti (Tanzania: AEG42543), Katete (Zambia: AEG42544); **var-2**: Kiambu-5 (Kenya: AEG42561); **var-3**: Chitongo (Zambia: AEG42565), Boleni (Zimbabwe: AEG42566), Uganda (AEG42567), Mariakani (AEG42569); **var-5**: Nanyuki (Kenya: AEG42575).

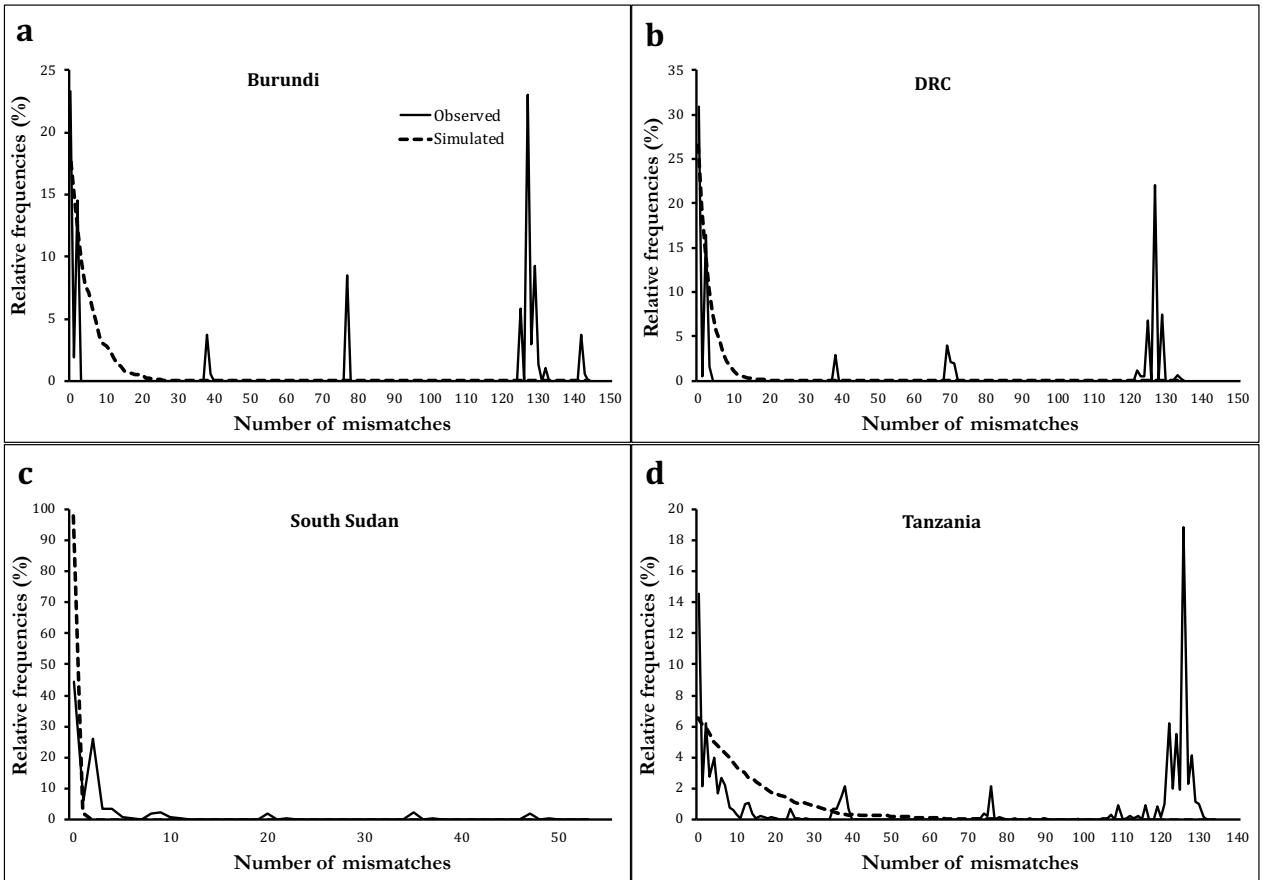

Figure S6.6 Mismatch distribution of frequencies of pairwise differences among *T. parva* Tp2 alleles assuming population expansion