GENETIC CONSERVATION OF THE p104 GENE USED FOR PCR-BASED DIAGNOSIS AND SURVEILLANCE OF THE *Theileria parva* PARASITE.

PERIS NYABOKE MAKORI

A THESIS SUBMITTED TO THE SCHOOL OF SCIENCE IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE AWARD OF DEGREE OF MASTER OF SCIENCE IN PARASITOLOGY, UNIVERSITY OF ELDORET, KENYA

FEBRUARY, 2023

DECLARATION

Declaration by the Student

This thesis is my original work and has not been submitted for any academic award in any institution; and shall not be reproduced in part or full, or in any other format without prior written permission from the author and/or University of Eldoret.

Nyaboke Peris Makori

_____ **Date** _____

SC/PGB/012/15

Declaration by the Supervisors

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

_____ **Date** _____

Dr. Jayne J. Binott

University of Eldoret, Eldoret, Kenya

_____ **Date** _____

Dr. Obara Isaiah

Freie Universität – Berlin, Germany.

DEDICATION

This research project is dedicated to my family, friends, and classmates for the selfless support throughout my research. Without them, life could not be what it is. To my mother and father whose patience and support they have shown me all through the study period.

ABSTRACT

Due to cross-border domestic cow movement, the range of *Theileria parva*, the protozoan parasite that causes East Coast disease in cattle, has been spreading to nations where it has not been found before. Cameroon and South Sudan are two recent additions to the list. This calls into doubt the preservation of the p104 antigen gene, which is the foundation of the widely used nested PCR technique for T. parva monitoring in the blood of sick cattle. In this study, parasites circulating naturally in the field in an endemic area of Busia Kenya were sampled and analysed for p104 nucleotide polymorphisms within a 496 bp fragment. p104 sequences derived from archived *T. parva* isolates from distinct regions within the African continent, as well as from medium scale distances within countries were also included in the study. These included isolates originating from Tanzania, Uganda, Rwanda and South Africa. P104 sequence polymorphism was assessed in a total of 56 sequences (of which 23 were obtained from the GenBank) originating from six nations that are widely dispersed throughout the parasite's geographic distribution, which includes eastern, central, and southern Africa. Among them were parasites from the Cape Buffalo wildlife reservoir and livestock. A substantial proportion of the ten allelic variations found in these isolates were also found in the three component stocks of the Muguga cocktail, which is utilized for both the infection and treatment live immunization method to control T. *parva* in the field. p104 variations that were often detected were found in isolates from South Africa, Kenya, Tanzania, and Rwanda. While some isolates showed distinct alleles, the small amount of diversity found did not match the location of origin of the parasite genomes examined in this investigation. The residue-changing p104 mutations did not exhibit any signs of positive selection. The amplicons were produced using nested primer oligonucleotides, which were notably globally conserved. The results show that the p104-based PCR detection technique should be reliable over the whole T. *parva* range, and when paired with amplicon sequencing, it can also yield information on the genotype of the parasite.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS, ACRONYMS, AND SYMBOLS

ECF	East Coast Fever
PIM	Polymorphic Immunodominant Molecule
RLB	Reverse Line Blot
DNA	Deoxyribonucleic Acid
VNTRs	Variable Number Tandem Repeat Sequences
SNPS	Single Nucleotide Polymorphisms
ITM	Infection Treatment Immunization
USD	United States Dollars
PP	Percent Positivity
AIC	Akaike Information Criterion
PCR	Polymerase Chain Reaction
ELISA	Enzyme-Linked Immunosorbent Assay (ELISA)
p104	*T. parva* 104 kDa (p104) rhoptry antigen
kDa	KiloDalton
RNA	Ribonucleic acid
WHO	World Health Organization
p67	67 kDa protein, *T. parva* major sporozoite surface antigen
OD	Optical density
GUTS	ground-up tick supernatants
TAE	Tris base, acetic acid and EDTA
EDTA	Ethylenediaminetetraacetic acid

PAUP Phylogenetic Analysis Using Parsimony

PAML Phylogenetic Analysis by Maximum Likelihood

BEB Bayes empirical Bayes

LnL log likelihod

NNI Nearest-Neighbour Interchange

AICc: Akaike Information Criterion

AICc: Corrected Akaike Information Criterion

c.f.u. Colony Forming Units

CFA Complete Freund's Adjuvant

GFP Green Fluorescent Protein

i.d. Intradermal

IFA Incomplete Freund's Adjuvant

i.m. Intramuscular

LD Lethal Dose

p.f.u. Plaque Forming units

SA Saponin

s.c. Subcutaneous

SR Severe reactor

WOE Water-in-oil emulsion

ACKNOWLEDGEMENT

I wish to thank God, for being the help to my education journey and for committing to be the strength to all my pursuits in education. I also wish to thank the entire faculty of University of Eldoret for their endless commitment to the provision of a conducive environment for academic excellence. Special thanks to my able supervisors Dr. Jayne Chebichii Binot and Dr Isaiah Obara whose patience, guidance, counsel and expertise made this study possible. Finally, to my family and friends for the invaluable guidance, prayers and support throughout the process of this study.

CHAPTER ONE

INTRODUCTION

1.1 Background information

The apicomplexan *Theileria parva* parasite immortalizes bovine lymphocytes resulting in a severe disease known as East Coast fever (ECF) in susceptible cattle, particularly exotic breeds (Chepkwony et al. 2022). The disease results in an estimated one million cattle deaths annually in eastern, central and southern Africa (Reviewed by Nene *et al.,* 2016) and is endemic in 13 countries extending from South Sudan to South Africa (Bishop *et al.,* 2020; Chatanga *et al.,* 2020). Furthermore, un-regulated transboundary movements of cattle with subclinical long-term infections have been shown to be responsible for the spread of *T. parva* from endemic areas North into Southern Sudan (Malak *et al.,* 2012; Marcellino *et al.,* 2017) and West into Cameroon (Silatsa *et al.,* 2020). The presence of *T. parva* infections among cattle in Cameroon and the fact that additional countries such as Central African Republic, Southwest Ethiopia, and Nigeria are also potentially exposed to movements of carrier cattle indicates that *Theileria parva* can be considered an emerging pathogen (Silasta *et al.,* 2020).

The occurrence of *T. parva* infections in previously unknown geographical locations has increased demand for testing procedures to support epidemiological studies on prevalence and distribution (Obara *et al.,* 2023). While microscopy remains an important

diagnostic tool especially in rural laboratories because it is cheap and easy to implement in any standard laboratory infrastructure, it is of limited sensitivity to detect all carrier animals (Mans *et al.,* 2015). A serological test that utilises an indirect ELISA to detect antibodies against the polymorphic immunodominant molecule (PIM), a protein antigen expressed by *T. parva* schizonts and sporozoites is available (Katende *et al.,* 1998). However, the interpretation of the PIM assay is based on expressing the test sera optical density scores as a percentage of the positive control serum, which can sometimes result in false positive, or inconclusive results due to background optical noise. Hence, there is often a need for validation using molecular methods (Obara *et al.,* 2023).

The reverse line blot (RLB) assay is most sensitive for detection of parasites, such as *Theileria mutans* that multiply in erythrocytes, whereas for *T. parva* RLB percentage positivity is likely to represent a minimum figure because most multiplication of *T. parva* occurs in T cells, which are several orders of magnitude less prevalent than erythrocytes in blood (Conrad *et al.,* 1986). Given this quantitative bias, the overall percentage of cattle infected with *T. parva*, as assessed using RLB is typically likely to be an underestimate (Obara et al. 2023).

More sensitive surveillance based on detection of *T. parva* genomic DNA became possible with the development and evaluation of a nested set of primers targeting a 277 bp gene sequence of the *T. parva* 104 kDa (p104) rhoptry antigen (Odongo *et al.,* 2010).

However, currently there is a lack information regarding p104 polymorphisms throughout the entire geo-spatial distribution of the parasite.

Theileria parva diversity has previously been assessed using sequence polymorphism in the 'so called' Tp CD8+ T cell target antigen genes (Graham *et al.,* 2006) and by analysis of a panel of variable number tandem repeat sequences (VNTRs) initially developed by Oura *et al.,* (2003). In addition, analysis of single nucleotide polymorphisms (SNPs) through next generation whole genome sequencing is emerging as a tool for future population genomics studies (Hayashida *et al.,* 2013; Henson *et al.,* 2012). Since whole-genome SNPs analysis is yet to find field applications for *T. parva* population genomic analyses, there is need to understand the utility and limitations of the p104 gene as the standard diagnostic gene and a potential genomic marker for *T. parva* across geo-spatial distribution of the parasite and its expanding range.

1.2 Statement of the problem

Given the limitations associated with microscopy, serology and reverse line blot (RLB) surveys, more sensitive surveillance based on detection of *T. parva* genomic DNA became possible with the development and evaluation of a p104 nested PCR (nPCR) assay targeting the single copy gene encoding the *T. parva* 104 kDa rhoptry antigen. (Odongo *et al.,* 2010). A recent study from Cameroon (Silatsa *et al.,* 2020) surprisingly revealed nine p104 genotypes defined by SNPs resulting into seven variants (Var_1 –

Var_7) based on the predicted amino acid sequences. These findings contrast earlier research suggesting widespread conservation in the section of the gene used for PCR-based diagnostics (Skilton *et al.,* 2002). Some of the less frequent variants detected in the Cameroonian study had not been described previously. Only two isolates of Var 7 in the Cameroonian study are present in the complete genome sequences of the Muguga cocktail component stocks used for the infection and treatment immunization (ITM), while the most frequent genotype (var_1) was identical to that of the vaccine stock *T. parva* Marikebuni that was tested on several farms in the western Kenya highlands (Wanjohi *et al.,* 2001). Given that this data indicates that additional p104 genotypes are identifiable in cattle, there is therefore a need to consider whether or not p104 polymorphism is extensive in the endemic region and can potentially impact the reliability of the PCR assay.

1.3 Justification of the study

ECF is endemic in 13 countries extending from South Sudan to South Africa (Bishop *et al.,* 2020; Chatanga *et al.,* 2020). In addition to being the most significant constraint to cattle productivity within the borders of the current distribution, there have been several reports of geographical spread of *T. parva* to areas initially thought to be free of infection. The most notable ones are the recent ECF outbreak in the Comoros island due to movement of infected cattle from Tanzania (De Deken *et al.,* 2007, the spread at the

borders of the existing distribution in Sudan, and the spread due to trade and trans-humance over great distances to central African countries like Cameroon (Silasta et al. 2020). Taken together, these latest instances of disease spread have the implication that *T. parva* can be considered an emerging parasite that necessitates the upscaling of testing procedures to support epidemiological studies on prevalence and distribution especially in the previously unknown geographical locations.

The currently used testing procedures such as microscopy, the PIM-ELISA and the reverse line blot (RLB) are marred with challenges including sensitivity, subjective operator-dependent interpretation of results, low throughput, and difficulty in standardisation among other problems that limit their application for large scale field surveillance studies (Mans et al. 2015). As such, detection of *T. parva* genomic DNA using a nested PCR targeting the *T. parva* 104 kDa (p104) rhoptry antigen gene offers the most sensitive surveillance in the field (Odongo *et al.,* 2010). However, given the detection of the parasite in hitherto unknown geographical locations, and few polymorphisms recently documented (Silasta *et al.,* 2020), there is a clear need to assess whether this diversity is likely to result in a discrepancy between the prevalence indicated by the p104 PCR detection and the true presence of *T. parva* infections among cattle in the field that may be undetected.

1.4 Objectives

1.4.1 General objective

To evaluate nucleotide polymorphism in the *T. parva* p104 antigen gene used in the standard diagnostic PCR assay and determine whether the polymorphisms can potentially impact the reliability of the PCR assay.

1.4.2 Specific objectives

1. To characterize the nucleotide sequence diversity of *T. parva* p104 gene fragment and the nested primer target sequence in parasites
2. To determine whether the p104 gene nucleotide polymorphism correlates with geographical origin of the isolates using phylogenetic analysis.
3. To assess for positive selection pressure on the p104 antigen gene from sequences across the endemic distribution space of the parasite.

1.5 Research hypotheses

1. There is no variation in nucleotide sequences within the *T. parva* p104 gene in parasite isolates from endemic areas in Kenya, and the archived isolates from distinct regions within the continent

2. The p104 nested primers target sequences are not universally conserved in *T. parva* isolates

3. The phylogenetic clustering of p104 gene does not correlate with the geographical origin of the *T. parva* isolates

CHAPTER TWO

LITERATURE REVIEW

2.1 Molecular taxonomy of the genus *Theileria*

The present taxonomic classification places the genus *Theileria* in the phylum Apicomplexa, which also includes other important human parasites such as *Plasmodium* (causative agent of malaria) and *Toxoplasma* (causative agent of toxoplasmosis) (Bishop *et al.,* 2020b; Sivakumar et al. 2014). This diverse phylum includes all eukaryotes whose members possess a definitive group of structures and secretory organelles called the apical complex. This assembly of secretory organelles at the anterior apex of the cell is responsible for invasion or establishment of the parasite in the cells of their mammalian and invertebrate hosts (Katris *et al.,* 2014).

The phylogenetic analyses have been based on the nuclear small subunit ribosomal RNA gene, termed 16S in prokaryotes and 18S in most eukaryotes (Bishop *et al.,* 1995). Analysis of 18S ribosomal RNA gene sequences of Sporozoans have revealed that the genus *Theileria* is phylogenetically closest to *Babesia*, a genus of tick-transmitted protozoans that infect red blood cells of most eukaryotes and *Plasmodium* (Bishop *et al.,* 2020b). There are an estimated 219 million cases and 660,000 malaria deaths caused by *Plasmodium*, in the tropics and subtropics (WHO World Malaria Report, 2012). A phylogenetic tree depicting the relationships among the apicomplexans is shown in the figure 2. 1.

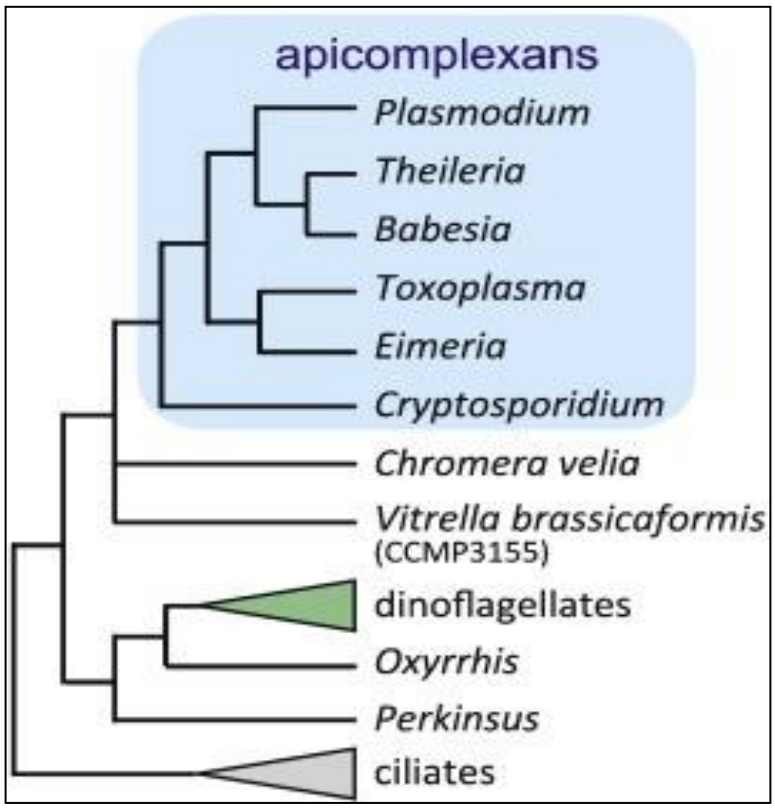


Figure 2.1: Phylogeny and evolution of the apicomplexans (Arisue & Hashimoto, 2015)

2.2 Economically important *Theileria* species globally

2.2.1 Transforming *Theileria* species.

Although a remarkable diversity of tick-borne *Theileria* species infect wild and domestic ruminants, globally the most important ones are *Theileria parva* and *Theileria annulata* (Nene & Morrison, 2016). Both *Theileria parva* and *Theileria annulata* transform host leukocytes resulting in acute lymphoproliferative diseases that are invariably lethal in exotic cattle breeds, but also cause high levels of mortality in indigenous calves (Tretina

et al., 2020). The two are therefore commonly referred to as transforming *Theileria* species (Sivakumar *et al.*, 2014).

For *T. parva*, the major tick vector is predominantly *Rhipicephalus appendiculatus*, but also *Rhipicephalus zambesiensis* in some parts of southern Africa (Morrison *et al.*, 2015). The known geographic distribution of ECF follows the climatic conditions that favour its tick vectors (Olwoch *et al.*, 2008) and the cattle populations at risk are shown in Figure 2.2.

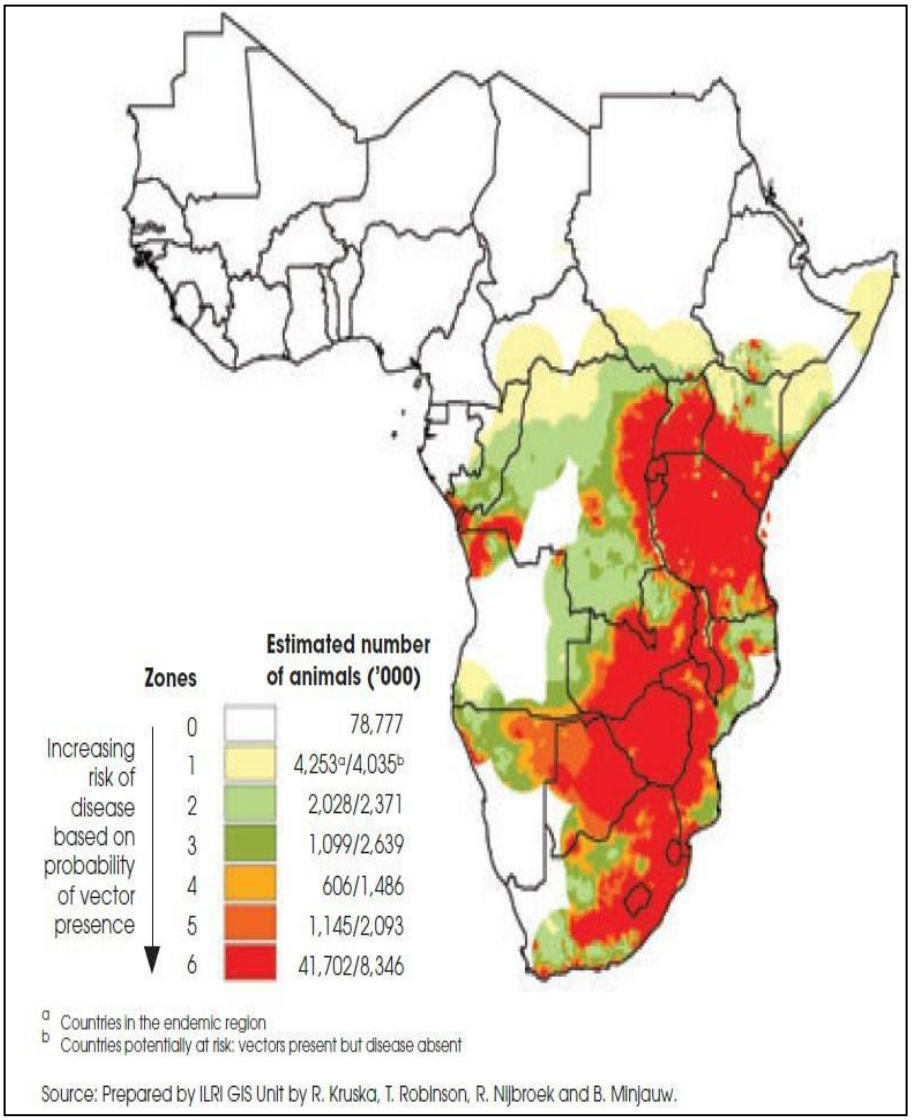


Figure 2.2: The distribution of animals at risk of ECF in eastern, central, and southern Africa.

The other equally important *Theileria* species, *Theileria annulata* causes tropical theileriosis or Mediterranean Coast fever. Transmitted by *Hyalomma anatolicum* and other *Hyalomma species*, it occurs in North Africa, southern Europe, the Near and Middle East, India, China and Central Asia (Liu *et al.*, 2022; Gharbi *et al.*, 2020). It causes both mor-

tality and reduced production and has significant economic impacts as a result (Liu *et al.*, 2022; Gharbi *et al.*, 2015). The geographical distribution of *Theileria annulata* and important *Hyalomma* tick vectors throughout the range of the parasite is shown in Figure 2.3.

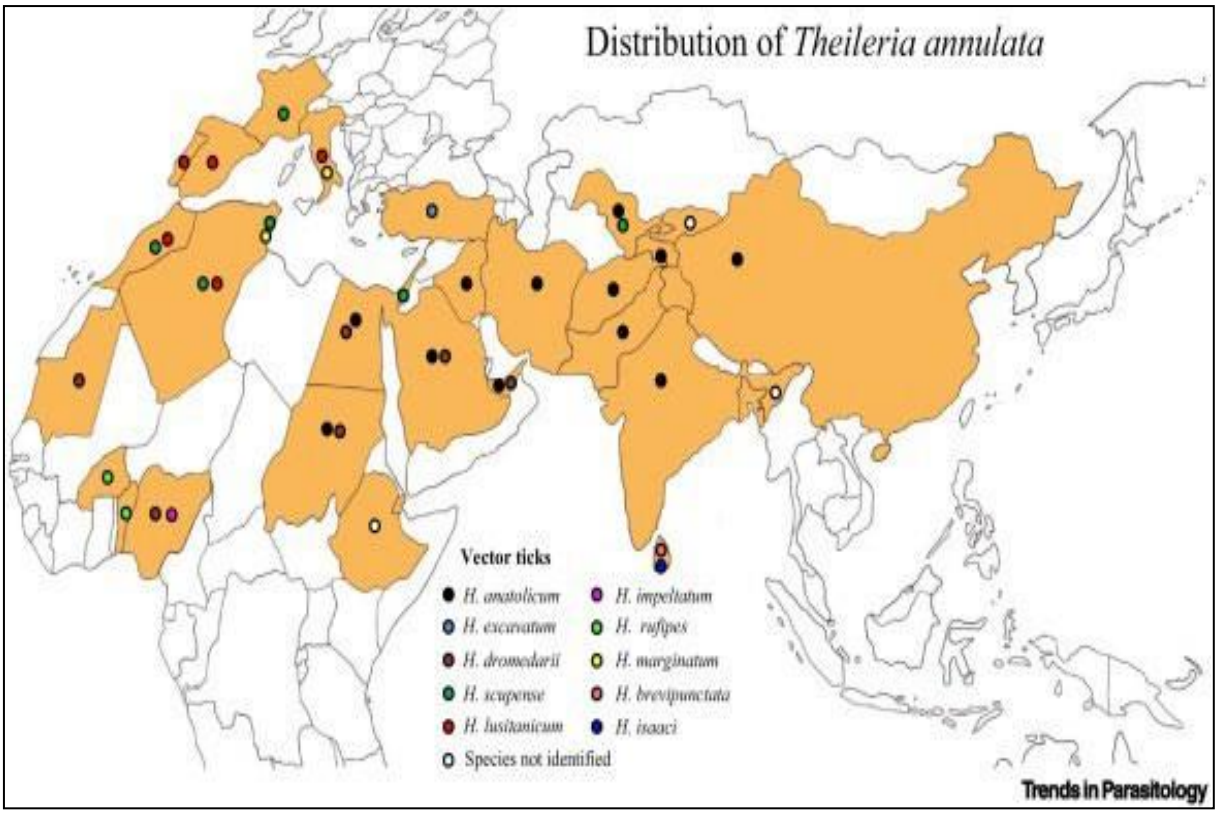


Figure 2.3: The distribution of *Theileria annulata* in Southern Europe, Western, Southern and Eastern Asia and Northern Africa. Source: (Liu *et al.*, 2022).

2.2.2 Non-transforming *Theileria* species

The non-transforming *Theileria* species cause disease in domestic ruminants as a consequence of the anaemia that is induced by the stages of these parasites that are found

in the mammalian erythrocytes. A key non-transforming *Theileria* species is *Theileria mutans* which occurs in Western, Eastern Central and Southern Africa and is transmitted by *Amblyomma* ticks. *Theileria mutans* may be responsible for disease in cattle in sub-Saharan Africa. But since it is frequently found as a mixed infection with other tick-borne pathogens, the impact of the disease it might cause is currently unclear (Mans et al. 2015, Yusufmia et a., 2010). The other less important non-transforming *Theileria* species include, *Theileria velifera*, *Theileria sergenti*, *Theileria buffeli*, *Theileria* ovis and *Theileria separata* (Agina et al. 2020; Sivakumar *et al.*, 2014).

2.3. East Coast fever (ECF), the disease caused by *Theileria parva.*

The apicomplexan *Theileria parva* transforms bovine lymphocytes resulting in a severe disease known as East Coast fever (ECF) (Fry et al. 2016). It has long been known that ECF is invariably lethal in individual high value exotic cattle typical of commercial production systems. This has had the implication that ECF has severely constrained the introduction of the more productive European cattle breeds into eastern, central and southern Africa (Kipronoh *et al.,* 2009. It has also recent been shown in a longitudinal study that ECF is the leading cause of death among indigenous zebu calves in pastoralist systems (Thumbi et al, 2018). This is an important addition to the ECF literature since cattle kept by small holder farmers and pastoralists have not traditionally been the targets

of ECF vaccination programmes (Jumba *et al.,* 2020). The disease is endemic in 13 countries extending from South Sudan to South Africa (Chatanga *et al.,* 2020).

The initial clinical signs of ECF resulting from infection with *Theleria parva* include pyrexia and enlargement of the superficial lymph nodes. Although several lymph nodes are typically enlarged, its parotid and prescapular nodes that show more pronounced enlargement (Mbassa *et al.,* 2006). As the infection progresses, the animals become weak and anorexic and also show signs of severe respiratory distress and a serious reduction in the white blood cells count. Post-mortem examination often reveals infected lymphocytes in the lymph nodes, lungs, liver, kidneys, gastrointestinal tract and sometimes the brain. Within the lungs, a frothy exudate in the trachea results in obstruction of airways due to pulmonary oedema and restriction of lung expansion due to pleural effusion. The ultimate result is respiratory failure and death of the animal, often within three weeks of infection (Fry *et al.,* 2016).

It is important to note that ECF also has a wildlife interface that modulates the epidemiology. This involves the African Cape Buffalo (*Syncerus caffer*) that constitute a wildlife reservoir in which the mammalian host is asymptomatic when infected by *Theileria parva,* but the parasite induces a distinct clinical syndrome, exhibiting low levels of parasitosis and parasitaemia, together with rapid mortality, when transmitted to cattle by *R. appendiculatus* (Morrison *et al.,* 2020). In addition to being the most

significant constraint to cattle productivity within the borders of the current distribution, there have been several reports of geographical spread of *T. parva* to areas initially thought to be free of infection. The most notable ones are the recent ECF breakdown in the Comoros island due to movement of infected cattle from Tanzania (De Deken *et al.*, 2007), the spread at the borders of the existing distribution in Sudan, and the spread due to trade and trans-humance over great distances to central African countries like Cameroon (Silasta *et al.,* 2020). It is also important to note that ECF results in an estimated US $596 million global losses annually from about 50 million cattle (including 10 million calves annually) within the ECF endemic range that are at risk (Allan & Peters, 2021).

2.4 Development of *T. parva* parasites and the disease process

Theileria parva has a complex life cycle involving bovine hosts (domestic cattle and African buffalo and) and the tick vector *Rhipicephalus appendiculatus* (Morrison *et al.*, 2020). The tick feeds on the host three times, as a larva, nymph and adult. *Theileria* sporozoites develop in the salivary glands of infected ticks and are passed to cattle along with tick saliva when the ticks feed. The classification of *Rhipicephalus appendiculatus* as a three-host tick is because the larvae, nymphs and adults feed on different hosts, which are not necessarily cattle.

The transmission of *Theileria parva* from ticks to cattle is described as trans-stadial, which means that larval or nymphal get infected during a blood-meal, and they only transmit to a new host after moulting by nymphs or adults, respectively. This is to be distinguished from transovarian transmission which is seen in the related protozoan *Babesia* transmitted by the one host tick *Boophilus* (Norval *et al.*, 1992).

Feeding ticks infected with *T. parva* inject sporozoites into the bovine host (Nene et al. 2016; Fawcett *et al.,* 1982). These rapidly invade host lymphocytes in a process similar to receptor mediated endocytosis, where they develop into multinucleate intracellular schizonts (Fawcett *et al.,* 1982). The schizont stage is responsible for *T. parva* induced pathology (Fry *et al.,* 2016). This typically involves cancer like immortalization of bovine lymphocytes. *T. parva* schizonts associate with the cell division apparatus to ensures that each daughter cell inherits the infection (Tejeri & Langsley, 2021; Chaussepied *et al.,* 2010). The uncontrolled lymphoproliferation and the invasion of lymphoid and non-lymphoid tissues by parasitized cells results in the death within three weeks of infection in the absence of treatment (Fry *et al.,* 2016). In a proportion of infected cells, the schizonts develop further into merozoites that are released upon rupture of infected cells and these develop into piroplasms, the red blood cell infective stage that is subsequently taken up by feeding ticks to complete the cycle (Morrison *et al.,* 2015; Nene *et al.,* 2016).

Ticks feeding on cattle become infected when they ingest red blood cells containing piroplasms and in their gut, the *Theileria parva* differentiate into male and female gametes, which fuse to form zygotes. The zygotes develop further and eventually migrate to the tick's salivary glands where generation of spores is stimulated by tick feeding leading to the development of between 30,000 to 50,000 sporozoites in an average tick (Jalovecka *et al.,* 2018). It is these thousands of sporozoites that are introduced together with tick saliva into a new bovine host while feeding to initiate a new cycle of parasite development. A summary of the life cycle of *T. parva* is shown in Figure 2.4.

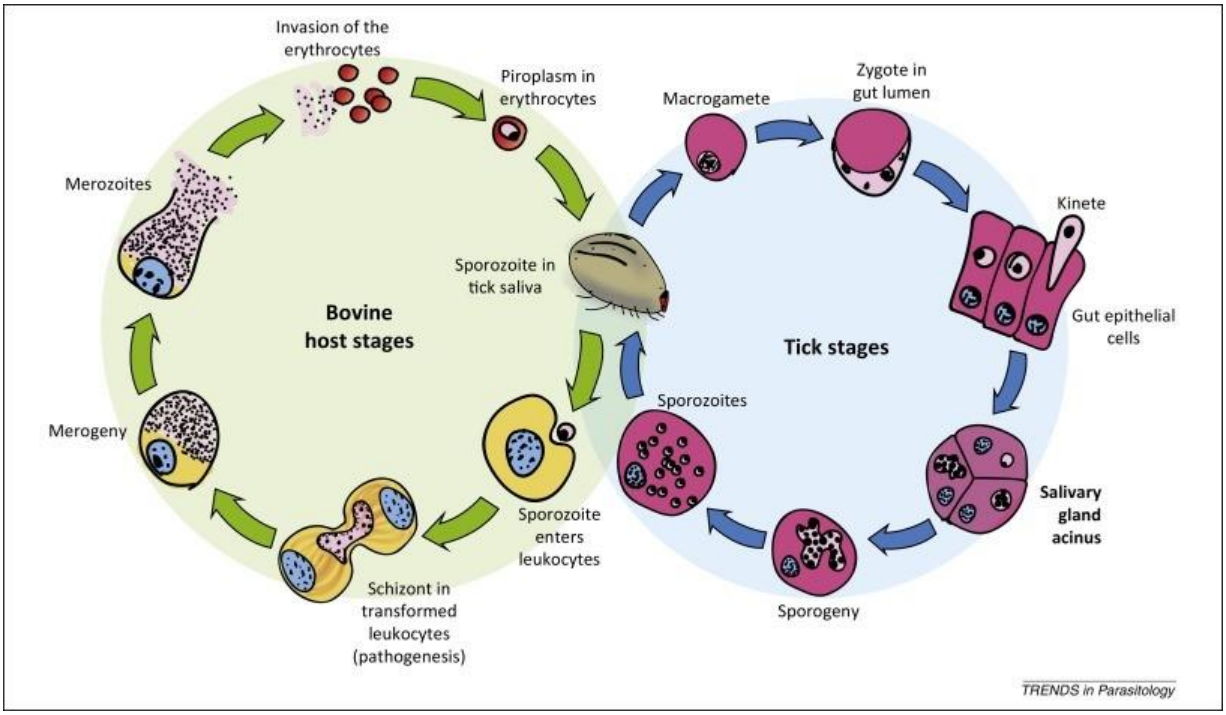


Figure 2.4: Life cycle of *T. parva* in the mammalian host and the tick vector (Tretina *et al.*, 2015)

2.5 Present ECF control methods and their shortcomings

The severe impact of ECF on animal health and the socio-economic dynamics of the particularly resource poor farmers in sub-Saharan Africa has been a driver of persistent efforts to control the disease through a raft of approaches outlined.

2.5.1 Use of acaricides

ECF has traditionally been controlled through the use of acaricides to limit the tick vector *Rhipicephalus appendiculatus*. However, acaricide use is unsustainable both in the medium and long term for reasons including the development of acaricide resistance by ticks, cost implications, and environmental safety concerns (Githaka *et al.,* 2022). Resistance in ticks has been shown to evolve faster than development of new chemicals (Narasimhan et al. 2021). Environmental concerns relate to contamination associated with the chemical residues that often end up in the food chains, and the collateral damage to beneficial arthropod species. There are also cost issues associated with acaricide utilization (Abbas *et al.,* 2014). Furthermore, in the absence of government support for community programmes, facilities for acaricide use like the community dips are now mostly not functional (Mugambi, 2012).

2.5.2 Chemotherapy

Buparvaquone, a trypanocidal drug is available for the treatment of *T. parva* (McHardy *et al.,* 1985). While its use is not widespread, it remains the frontline drug of choice that

is commercially available (Nene *et al.,* 2016). Its use in field situations is limited by the fact that it needs to be administered in the early stages of infection in order to be effective (Martins *et al.,* 2010). But since early diagnosis is not always practical in the field, the effectiveness of chemotherapy is often compromised. Although there are no reports of resistance to this drug yet, drug resistance has been documented in the closely related parasite, *Theileria annulata* correlating with mutations in the cytb gene (Hacılarlıoglu *et al.,* 2023; Mhadhbi *et al.,* 2010), leading to speculations that it is only a matter of time before resistance is detected in *T. parva.* Finally, in common with acaricide use, the costs of these drugs often constrain their accessibility by resource poor farmers in endemic areas.

2.5.3 Live vaccination

For *Theileria parva,* a lethal dose of sporozoites simultaneously administered with a long-acting formulation of oxytetracyline is used to immunize cattle (Nene *et al.,* 2016). The development of ITM was based on the initial observations that animals that recovered from natural *Theileria parva* infections were often immune to subsequent infections (Radley *et al.,* 1975). It subsequent became apparent that this protection could be replicated by artificially infecting animals with *Theileria parva* sporozoites. The ability to harvest and cryopreserve *T. parva* sporozoites from ground up ticks and the commercial availability of long-acting formulations of the oxyteracycline antibiotic

enabled the development of the live infection and treatment immunization procedure (Bishop *et al.*, 2020; Cunningham *et al.*, 1973).

One of the challenges to ITM use was that of parasite strain specificity, usually manifested as the inability of a single parasite strain when used in ITM to protect against all other strains of the parasite. This challenge was eventually addressed, and the currently licensed version of the live-parasite infection and treatment immunization procedure (ITM) is a trivalent cocktail of *T. parva* sporozoites derived from the Muguga, Serengeti-transformed and Kiambu 5 strains that have been well characterised (Radley *et al.*, 1975, Perry, 2016).

Table 1.1: Immunisation against *Theileria parva* by infection and treatment (Radley *et al.* 1975).

Stocks used for immunisation	Challenge stock	Number of animals	Number immune
Muguga	Muguga	5	5 (100%)
-	Kiambu 1	5	0 (0%)
Muguga	Kiambu 1	5	2 (40%)
Serengeti	Kiambu 1	5	1 (20%)
Muguga + Kiambu 5	Kiambu 1	5	2 (40%)
Muguga + Serengeti	Kiambu 1	4	1 (25%)
Muguga + Kiambu 5 + Serengeti	Kiambu 1	5	5 (100%)
Muguga + Kiambu 5 + Serengeti	Solio KB1	5	5 (100%)
Muguga + Kiambu 5 + Serengeti	Entebbe 1	6	6 (100%)
Muguga + Kiambu 5 + Serengeti	Entebbe 2	6	6 (100%)

In Table 1.1, Both Muguga and Kiambu strains are of Kenyan origin that were isolated from cattle, while the Serengeti-transformed was isolated from a buffalo in Tanzania and adapted for cattle transmission through a series of passages in cattle (Irvin *et al.,* 1974; Young & Purnell, 1973; Brocklesby *et al.,* 1961). Despite the ITM vaccine's invention in 1975, it was only deployed to a limited extent over its next 25 years (Di Giulio *et al.,* 2009). However, the success of the vaccine in Tanzanian pastoralist systems, in which over 1 million cattle were vaccinated using the original trivalent Muguga cocktail version of ITM, stimulated demand in other countries and highlighted the potential for vaccine adoption (Patel *et al.,* 2019; Lynen 2006). In Uganda for example, there has not yet been widespread adoption on a large scale as in Tanzania, and there is an estimated 30% calf mortality in the indigenous Ugandan zebu herd which is estimated at approximately 2 million animals (Kasaija *et al.,* 2021).

Of particular interest is the situation in Southern Sudan where vaccination has not yet occurred, but there is a growing demand and a major ECF problem. It has been previously demonstrated that *T. parva* is spreading towards the north of South Sudan into regions north of Juba previously thought to be free of the parasite (Salih *et al.,* 2018; Marcellino *et al.,* 2017). These observations show that ECF is endemic in one region in South Sudan, affects trade cattle in transit in a second region and is spreading further north. The latter is particularly interesting since it makes South Sudan unique as the only

country in the world where *Theileria annulata* and *Theileria parva* both occur and there is preliminary data suggesting co-existence in certain herds (Marcellino *et al.,* 2017).

However, cattle immunized by infection and treatment (ITM) using the trivalent cocktail are often still susceptible to challenge with buffalo-derived parasites (Sitt et al. 2015). This is because buffalos are known to harbour parasites of higher antigenic diversity than those transmitted between cattle by ticks (Conrad et al. 1989; Allan et al. 2021; Obara et al. 2015).

In addition to these efficacy issues, the production of the vaccine is not trivial requiring several nymphal ticks, rabbits and cattle and quality control is very problematic. Furthermore, the current packaging is a 40-dose straw and a single dose costs up to 12 USD, mostly because of the need of a liquid nitrogen cold chain during storage and delivery (Patel *et al.,* 2019). These challenges have constrained the widespread use of the live vaccines. It is worth noting that despite the challenges outlined, the live vaccination procedure has been very useful in dissecting the immune responses relevant to protection in order to inform subunit vaccine design.

Other scientific hurdles that have to be considered to get to the point of wide deployment include: (i) a perception that the introduction of 'foreign' parasite genotypes, because of the persistent, tick transmissible infection ('carrier state') induced by vaccination, could result in enhanced disease problems, and (ii) the fact that the immunity induced by

vaccination is partially strain-specific and that given the known diversity of *Theileria parva* in the natural host Cape buffalo (*Syncerus caffer*), there could be problems with breakthrough in vaccinated animals in areas where there is a cattle-buffalo interface (Di Giulio *et al.,* 2009; Bishop *et al.,* 2004).

2.5.4 Current suite of subunit vaccine candidates

In the development of subunit vaccines, the key questions that needed to be addressed included: (i) what immune responses are required to protect the host from immunopathology associated with *Theileria parva* infection? (ii) what is the choice of antigens and how can antigenic polymorphisms be best dealt with? and (iii) how can the issue of vaccine formulations, including delivery systems be addressed?

(a) Sporozoite-neutralizing vaccine candidates

Theileria parva infections generally induce low levels of antibody against the infective sporozoite stage of the parasite. However, it has been demonstrated that in animals repeatedly challenged by *T. parva* sporozoites, antibodies that completely neutralize the infectivity of sporozoites in *vitro* develop (Dobbelaere *et al.,* 1984; 1985; Musoke *et al.,* 1984). It has further been shown that these neutralizing antibodies are specific for the *T. parva* sporozoite surface called p67 (Musoke *et al.,* 1984).

Based on the findings described above, the gene encoding the *Theileria parva* sporozoite surface antigen, p67, was cloned, and immunization trials performed with baculovirus expressing the p67 recombinant protein (Kaba *et al.,* 2005). Other expression systems were also evaluated. This experimental anti-pathogen subunit vaccine based on a

sporozoite surface antigen (p67) induces consistent 70% protection, regardless of expression system, against experimental *Theileria parva* needle challenge with tick derived stabilates (Kaba *et al.,* 2003, 2005).

There have been several experimental vaccine trials in cattle using different p67 constructs and antigen delivery systems (Musoke et al. 1992, Nene et al. 1995, Nene et al. 1996, Heussler *et al.,* 1998, Gentschev *et al.,* 1998, Honda *et al.,* 1998, Bishop *et al.,* 2003, Kaba et al. 2005, Musoke et al. 2005). What all these studies have in common is that the high level of protection observed in the needle challenge experiments is not replicated following these tick challenge studies in the field. The reasons for the generally low levels of protection in field situations, as opposed to the high levels of protection following needle challenge, it is not known. Two hypotheses have been advanced to explain this discrepancy in protection. The first one suggests a role for immunomodulatory and anti-inflammatory pharmacopeia of tick molecules, while the second relates to antigenic diversity in the field (Obara *et al.,* 2015). More importantly, anti-p67 antibody titres, sporozoite neutralizing antibody titres, and linear epitope mapping have all failed to provide a correlate between anti-p67 responses and immunity to ECF.

The polymorphic immunodominant molecule (PIM) antigen that is expressed by both the infective sporozoite stage of the parasite and the intracellular schizont was identified as another possible target of neutralizing antibodies (Minami et al. 1983, Shapiro et al. 1987, Toye *et al.,* 1996). However, it remains to be seen whether immunization of cattle with different constructs of PIM will be protective.

(b) Schizont epitopes that stimulate CD8+ cytotoxic T-cells in cattle

Studies involving adoptive transfer of different components of the immune system between *T. parva* immune and naive monozygotic twins led to the discovery that CD8 T cells specific for schizont infected lymphocytes are the mediator of live vaccine induced solid protection (Emery, 1981; McKeever *et al.*, 1994). Subsequently, a set of 10 antigens designated Tp 1-10 were shown to recognized by CD8 T cells (Table 2.2) (Graham et al. 2006; Graham *et al.*, 2008).

Table 2.2: Schizont antigens recognized by CD8 T cells (Graham et al. 2006; Graham *et al.*, 2008).

Antigen	Genome reference	Chromosome	Protein	Predicted function	Restricting MHC
Tp1	TP03_0849	3	543aa	Hypothetical	N*01301 (A18)
Tp 2	TP01_0056	1	174aa	Hypothetical	N*01201 (A10 - *Bt*)
Tp 3	TP01_0868	1	265aa	Hypothetical	?
Tp 4	TP03_0210	3	579aa	ε-TCP1	N*00101 (A10 *Bi*)
Tp 5	TP02_0767	2	155aa	Elf-1A	N*00902 (A15)
Tp 6	TP01_0188	1	277aa	Prohibitin	T6
Tp 7	TP02_0244	2	721aa	hsp90	T7
Tp 8	TP02_0140	2	440aa	Cyst. protease	N*00101 (A10- *Bi*)
Tp 9	TP02_0895	2	334aa	Hypothetical	N*02301 (A14)
Tp 10	TP04_0772	4	310aa	Coronin	N*00201 (A10 - *Bt*)
Tp 12	TP01_1091	1	572aa	Hypothetical	N*01801 (A11)

One of these antigens, Tp1 has been shown to be partially protective, although it still requires optimization of delivery. However, it has turned out that these antigens are not

recognized by CD8 T cells from African taurine and indicine cattle. This has the implication that subunit vaccines designed based on these candidate antigens will only be relevant in Holstein and to a limited extent Boran cattle (Graham *et al.,* 2008). The basis for this lack of recognition of the current suite of antigens by African cattle is the phenomenal described as MHC restriction. CD8 T cells only recognize schizont antigens when complexed with cattle class I MHC (BoLA-I) molecules. Recent studies have demonstrated functional divergence between African and exotic taurine BoLA class I molecules (Obara *et al.*, 2016).

The only exception to this lack of recognition is the candidate vaccine antigen designated Tp2. It has been shown that African zebu cattle CD8 T cells can recognize this antigen and produce gamma interferon in an Elispot assay that involved stimulating the CD8 T cells with overlapping peptides derived from Tp2 (Akoolo *et al.,* 2008). This finding is important since at present this is the only antigen that can be pursued further with the goal to produce a vaccine that overcomes the issue of MHC diversity in the field in endemic areas in Africa.

Since the other biggest constrain to the use of Tp2 will be antigenic variation in the field parasites which could result in immune evasion, it will be important to ascertain the extent to which this antigen is variable in field isolates. These studies should particularly

focus on the epitopes encoded within Tp2 that has been shown to be recognized by African zebu cattle that comprise most poor farmers herds (Akoolo *et al.,* 2008).

2.5.5 Integrated pest management (IPM) of ticks and other approaches

To achieve a more sustainable and cost-effective management of ECF, there is an emphasis to combine different control methods to minimize the risks of over-reliance on one method e.g., Acaricides that may be rendered ineffective over time by resistant ticks (Abbass *et al.,* 2014). The IPM approach combines a suit of methods aimed at enhancing human, animal, and environmental health by adapting and integrating both traditional and novel technologies, including biological, chemical, and cultural control approaches (Ehi-Eromosele *et al.,* 2013). Biological control ECF approaches may encompass the use of natural enemies against *Rhipicephalus appendiculatus* ticks — the *T. parva* vector. The prospects for biological control of *R. appendiculatus* have previously been demonstrated by Kaaya *et al*. (1996) using the entomophagous fungi *Beauveria bassiana* and *Metarhizium anisopliae*. Further, biological control may also involve the use of ECF and tick resistant breeds through selective breeding of indigenous cattle breeds that have co-evolved with *T. parva* in ECF-endemically stable areas leading to the development of some level of natural resistance to ECF over time (Nanteza *et al.,* 2023; Vajana *et al.,* 2018; Kabi *et al.,* 2014). On the other hand, cultural control entails the use of practices that reduce the exposure or susceptibility of cattle to ticks or ECF. For example, rotational grazing, pas-

ture burning, bush clearing, and quarantine of infected animals can reduce the tick infestation and transmission of ECF (Norval *et al.,* 1992). Chemical control implements the use of acaricides to kill ticks and prevent ECF transmission. The mechanical or physical control approach involves the use of devices or barriers that prevent or remove ticks from cattle. For example, traps, fences, repellents, and manual removal can reduce the tick burden and ECF risk (Norval *et al.* 1992).

2.6 Testing procedures to support epidemiological studies on *T. parva* prevalence and distribution.

2.6.1 Microscopy

Light microscopy for *T. parva* is a well-established and cost-effective method that is based on examination of Giemsa-stained blood smears (Norval *et al.,* 1992). While this technique is easy to implement and readily available in all standard laboratories, its sensitivity and specificity for parasite detection and identification is dependent on user experience and proficiency (Mans *et al.,* 2015). It has limitations in detecting carrier animals in which parasitaemia is low (Odongo *et al.,* 2010). Moreover, discrimination between piroplasms of other *Theileria* species may be difficult as they are morphologically very similar, thus confusion may arise if mixed infections are present as is normally the case in the field (Mans *et.al* 2015; Odongo *et al.,* 2010).

2.6.2 Serology

Cattle serum samples can be assayed for antibodies to *T. parva* using an indirect ELISA to detect antibodies against the polymorphic immunodominant molecule (PIM), a protein antigen expressed by *T. parva* schizonts and sporozoites (Katende *et al.*, 1998). The interpretation of the test sera optical (OD) scores is based on expression as a percentage of the strong positive control serum, often referred to as percent positivity (PP). The empirically determined threshold is meant to exclude sera whose OD scores are below that of the weak positive serum control pre-diluted to give an OD reading near the cutoff value. However, there are some operational limitations to PIM ELISA. There is often a need to set up all the controls at least in duplicates to address the variability in the ELISA Optical density (OD) readings for internal control replicates (strong positive serum, weak positive serum, known negative serum, as well as conjugate controls). This is required to allow the reliability of the assay to be assessed but background can sometimes result in false positive, or inconclusive results, and there is often a need for validation using molecular methods (Obara et al. 2023).

2.6.3 Molecular diagnosis

A suite of molecular tools has been developed through data mining the *T. parva* genome, that permit investigation of parasite molecular epidemiology in the field. The key tools are a panel of mini and microsatellite markers that are very useful for distinguishing parasite genotypes (Oura et al. 2003; Oura et al. 2005, Odongo et al. 2006). There have

been several molecular epidemiology and population genetics studies using polymorphic variable number tandem repeat (VNTR) sequences and nucleotide sequencing of genes that are targets of the CD8+ T cell responses. Typically, they have combined Tp1 and Tp2 antigens with 12–30 VNTRS. Applying the VNTR data to field populations from endemic locations in Kenya, Tanzania, Uganda, and South Sudan often reveals very high levels of variability (Oura *et al.,* 2005; Odongo *et al.,* 2006; Nanteza *et al.,* 2020; Mwega *et al.,* 2015; Salih *et al.,* 2018).

Regarding parasite diversity ascertained by genes targeted by CD8+ T cell responses, the most comprehensive investigation of its kind employed 82 field isolates of cattle and buffalo for Tp1 and Tp2 variation. According to Pelle et al. ,(2011), these antigens have one or six mapped epitopes, respectively. Both antigens were found to be quite diverse, however the majority of this diversity was found in isolates from buffalo or cattle who had been co-grazing with buffalo and died from ECF (Bishop et al., 2015). Tp antigen sequence study (Pelle et al., 2011) revealed that buffalo has a significantly higher variety of *Theileria parva* genotypes, which has been validated by genome sequencing of partly purified schizonts from cell lines (Hayashida et al 2012).

2.6.4 The Reverse line blot (RLB) Hybridization Assay

A PCR-based technology termed as reverse line blotting has been adapted for tick-borne diseases in livestock (Gubbels *et al.,* 1999). This method allows for the simultaneous

detection of multiple parasite species or strains in a single sample which enables surveillance of all currently known tick-borne diseases in the blood of cattle much more comprehensively than was possible previously. RLB is based on PCR amplification of the 18S rRNA gene of *Theileria/Babesia* and hybridization against *T. parva / T. mutans* probes as previously described (Gubbels et al. 1999). However, it should be noted that the RLB assay is more sensitive for detection of parasites, such as *T. mutans* that multiply in erythrocytes, whereas the *T. parva* RLB percentage positivity is likely to represent a minimum figure because most multiplication of *T. parva* occurs in T cells, which are several orders of magnitude less frequent than erythrocytes in blood. Given this bias, the overall percentage of cattle infected with *T. parva* is likely to be an underestimate (Obara et al. 2023). Moreover, among the major draw backs of RLB is that it is time-consuming as the assay involves multiple steps, including DNA extraction, PCR, labelling, and hybridization processes (Mans *et al.,* 2015).

2.6.5 Nested p104 PCR

More sensitive surveillance based on detection of *T. parva* genomic DNA became possible with the development and evaluation of a nested set of primers targeting the gene sequence of the *T. parva* 104 kDa (p104) rhoptry antigen (Odongo *et al.,* 2010). However, the p104 gene has previously received limited use as a polymorphic marker, perhaps in part because it is the basis of the standard diagnostic PCR assay (Skilton *et*

al., 2002, Odongo *et al.,* 2010). *T. parva* molecular epidemiology and population genetics studies have typically used polymorphic variable number tandem repeat (VNTR) sequences and nucleotide sequencing of genes that are targets of the CD8+ T cell responses. These have typically used 12-30 VNTRS often in combination with Tp1 and Tp2 antigens. However, a study from Cameroon (Silatsa *et al.,* 2020), where *T. parva* has only recently become present, revealed nine p104 genotypes, some of the rarer ones not having been described previously. This raises questions regarding the extent of conservation of the p104 gene, including the target regions of the primers, in *T. parva* isolates from multiple geographically separated regions.

CHAPTER THREE

METHODOLOGY

3.1 Study design

An explorative research design was adopted for this study which depended on limited availability of field samples collected from Busia County in the former Western province of Kenya—an ECF endemic region in Kenya (Gachohi *et al*. 2012). The samples were obtained as part of another ongoing research project on animal host diversity. These samples were supplemented with archived parasite material and GenBank™ repository sequences from previous studies (Obara et al., 2015, Mwamuye et al., 2020). The acquisition of the different samples and material is described.

3.2 *Theileria parva* field samples and isolates

Fresh field blood samples (n=50) for this study were collected from asymptomatic cattle from Busia in month of August 2018 with the help of certified animal health technicians. The animals were sampled from five sub-counties, viz., Butula, Teso North, Teso South, Nambale and Matayos. Before sampling, the researcher sought the consent of individual farmers after clarifying the outline and relevance of the study. Up to 10 animals per location were sampled and the blood collected into vacutainers then snap frozen in liquid Nitrogen because the parallel project on animal host diversity was interested in RNA extraction. The Busia samples were shipped to Garissa veterinary laboratory for DNA

extraction and PCR amplifications. In addition, the study utilized archived blood samples from previous studies that were collected from the Marula farm in central Kenya, an areas frequented by buffalo (n = 20) (Bishop *et al.,* 2015) and Tanzania—Monduli (n = 20) (Mwamuye *et al.,* 2020). Further, the study used a range of characterised *T. parva* isolates that were available at the Institute of Parasitology and Tropical Veterinary Medicine, Freie University Berlin that were archived in the form of frozen ground-up tick supernatants (GUTs), salivary glands and cell cultures (n = 11). In total, 101 samples across the geo-spatial distribution of the parasite were processed for this study.

3.3 *Theileria parva* p104 sequences retrieval from GeneBank.

To supplement the p104 sequence data generated from this study and to give a broad coverage analysis across the geographical range of the parasite, p104 sequences from earlier field studies as of May 2022 were downloaded from the GenBank™ repository (https://www.ncbi.nlm.nih.gov/). The downloaded p104 sequences were filtered by length and only sequences of 496 bp and longer were retained as these matched the length of the amplicon generated in the study described herein. The countries from which the samples and GenBank sequences originated are shown in Figure 3.1.

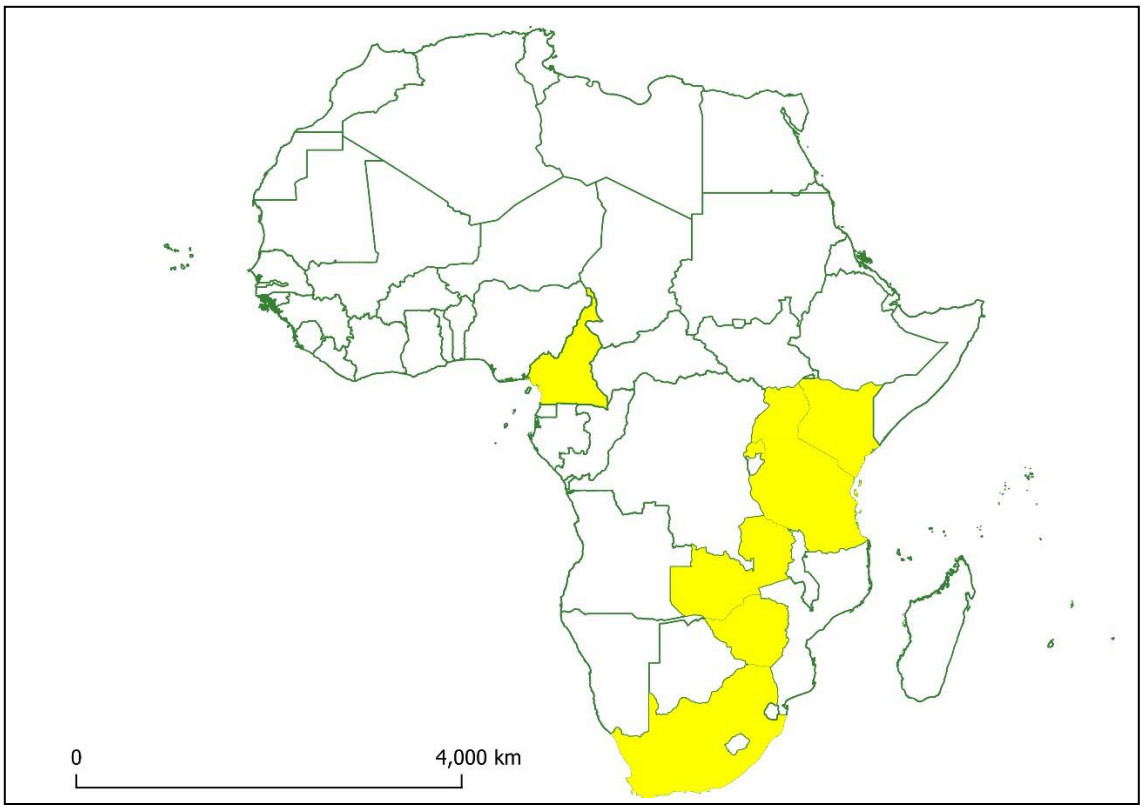


Figure 3.1 A map showing countries of origin and the geographic spread of the *T. parva* samples and GenBank p104 sequences used in this study. The countries are highlighted in yellow (Image credit: Own. 'Made with QGIS).

3.4 Genomic DNA extraction and amplification of the *T. parva* p104 gene

Genomic DNA was extracted from blood, frozen ground-up tick supernatants (GUTS), salivary glands and cell cultures using the Qiagen DNeasy Blood and Tissue Kit according to the manufacturer's instructions and used as template for amplification of the *T. parva* p104 gene as previously described (Skilton *et al.,* 2002).

Briefly, oligonucleotide primers IL3231 5'-ATTTAAGGAACCT GACGTGA CTGC-3' (forward) and IL755 5'-TAAGATGCCGACTATTAATGACACC-3' (reverse) were

used to amplify 496 bp fragment of the p104 gene. The amplifications were performed using the Phusion High-Fidelity PCR reagents (Thermo Scientific) in 50 µL reaction volumes composed of 20 ng of DNA, 1X Phusion HF Buffer, 200 µM of each dNTP, 0.5 µM of forward and reverse primer and 1 unit of Phusion DNA Polymerase. The PCR cycling conditions were 94°C for 1 min, 40 cycles of 1 min at 94°C, 1 min at 60°C, and 1 min at 72°C, followed by a final elongation period of 2 min at 72°C. Known positive and negative controls were included in each run. The amplification products were visualized by UV trans-illumination in 1.5% TAE-agarose gels stained with GRGreen (Excellgene, Monthey, Switzerland).

3.5 Amplicon purification and sequencing.

The amplicons were purified using the Zymoclean Gel DNA Recovery Kit. Briefly, the DNA fragment from the agarose gel was cut using a scalpel blade and transferred into a 1.5 ml microcentrifuge tube. Agarose dissolving buffer (ADB) was added to each volume of agarose excised from the gel and incubated at 55 °C for a minimum of 10 minutes and then briefly mixed by vortexing. This is a crucial step as the gel slice must be completely dissolved before moving on to the subsequent steps. The melted agarose solution was transferred to a Zymo-Spin™ Column and centrifuged for 1 minute before discarding the flow-through. Finally, 200 µl of DNA Wash Buffer was added to the column and centrifuged for 30 seconds. The flow-through was discarded and the wash

step repeated. The purified amplicons were sequenced bidirectionally through the Sanger sequencing technique.

3.6 Analysis of p104 polymorphisms

Editing and assembly of chromatograms for downstream analyses were accomplished using Geneious Prime 2022 software. Next, the bidirectional forward and reverse sanger sequence reads were assembled de novo and translated into their predicted protein sequences. Nucleotide and amino acid sequence alignments were generated using MAFFT V.7 and MUSCLE, respectively, as implemented in Geneious Prime 2022 software, and the percent pairwise identity matrices were generated for both alignments.

3.7 Phylogenetic analysis

The alignments for phylogenetic and positive selection analyses included an additional 23 p104 sequences generated previous field studies in different countries and were archived in the GeneBank. These included Cameroon (Silatsa et al. 2020), South Africa (Sibeko et al. 2011), Tanzania (Ringo *et al.*,2020; Mwega *et al.,* 2014), Uganda (Tayebwa *et al.*, 2018), Zambia (Hayashida *et al.*, 2013), and Central Kenya (Moumouni *et al.,* 2015). For phylogenetic analysis, the best-fitting nucleotide substitution model was selected based on the Akaike Information Criterion corrected for small sample size (AICc). The parameters included: nucleotide substitution rate parameters, equal or unequal base frequencies (+F), a proportion of invariable sites (+I), and rate variation

among sites (+G). Likelihood calculations for the nucleotide substitution models were performed with PhyML_3.0_linux6 (Guindon *et al.,* 2010). A total of 88 models of nucleotide substitution were evaluated using jModelTest 2.1.10 (Darriba *et al.,* 2012). Maximum-likelihood tree-search algorithms were implemented in PAUP 4.0 beta version using the parameter estimates for the best-fit model (Swofford, 2003), and branch support was tested based on 1000 bootstrap replicates.

3.8 Assessing evidence of selection in the *T. parva* p104 gene

The maximum likelihood phylogenetic tree file generated with PAUP was used as an input to run CODEML from the PAML4 package (Yang, 2007) to evaluate if there are positions in the *T. parva* p104 gene coding for residues that show an excess of non-synonymous (dN) over synonymous substitutions (dS) which is the evidence for positive selection for amino acid substitutions. Under the assumptions of neutral evolution, the dN/dS ratio (ω), is expected to have a value of 1. Positive and purifying (negative selection are indicated when $\omega > 1$, and $\omega < 1$, respectively. To assess for selection in this study, the following models were evaluated: M1a – two discrete categories, one for purifying selection where $\omega<1$, and the other for neutral selection where $\omega = 1$ (Nielsen & Yang, 1998; Yang *et al.,* 2005); M2a – an extension of M1a model, with an additional category for positive selection where $\omega >1$ (Nielsen & Yang, 1998; Yang *et al.,* 2005); M7 – a continuous beta distribution of ω restricted to the interval (0;1), no positive selection allowed (Yang *et al.,* 2000) and M8 – extension of M7 model, with additional,

discrete category for positive selection (Yang et al. 2000). Akaike weights were used to evaluate model fit and whenever the best-fit model is M2a or M8, sites under positive selection are determined through the Bayes empirical Bayes (BEB) approach.

CHAPTER FOUR

RESULTS

4.1 Amplification of the p104 gene

A total of 101 samples across the geo-spatial distribution of the parasite in sub-Saharan Africa were processed for this study. However, the targeted 496 bp p104 gene fragment was successfully amplified and purified for sequencing in only 33 samples. Notably, only 2 out of the 50 fresh field samples from Busia were amplified, while Monduli and Marula had 8 and 12 successful amplifications respectively out of 20 samples for each location (Table 4.1). The details of the successfully amplified samples are summarised in Table 4.1, and a representative gel image showing the targeted 496bp fragment from a selected field samples that were amplified is shown in Figure 4.1.

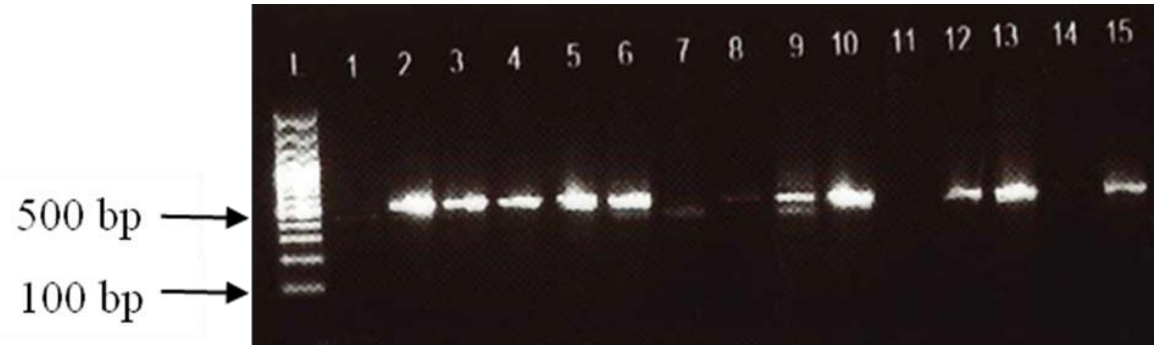


Figure 4.1: A representative gel image of the targeted 496 bp amplification: Lane L = 100 bp DNA Ladder (Thermo Scientific™ O'GeneRuler™). Lane 1 – 13 Samples. Lane 14. Non-Template Negative Control. It's unclear whether the additional faint band on lanes 6 and 9 were indicative of unspecific amplification or the presence of INDELs in the p104 gene. In both cases sequence reads were derived from the prominent 496 bp amplicon. There were no detectable amplicons from lanes 1, and 11and 14.

Table 4.1: Descriptive summary of successful amplified samples in this study and details of their origin (GUTs = ground-up tick supernatants)

No	Sample	Origin	Material used	Reference
1	Ser-(Serengeti)	Tanzania	GUTs	Young & Purnell, 1973
2	Manyara buffalo	Tanzania	GUTs	Schreuder *et al.*, 1977
3	Pug - (Pugu)	Tanzania	Cell culture	Uilenberg *et al.*, 1976
4	M13 – Monduli	Tanzania	Blood	Mwamuye *et al.*, 2020
5	M14 – Monduli	Tanzania	Blood	Mwamuye *et al.*, 2020
6	M15 – Monduli	Tanzania	Blood	Mwamuye *et al.*, 2020
7	M36 – Monduli	Tanzania	Blood	Mwamuye *et al.*, 2020
8	M29– Monduli	Tanzania	Blood	Mwamuye *et al.*, 2020
9	M16 – Monduli	Tanzania	Blood	Mwamuye *et al.*, 2020
10	M136– Monduli	Tanzania	Blood	Mwamuye *et al.*, 2020
11	M141 – Monduli	Tanzania	Blood	Mwamuye *et al.*, 2020
12	Bol-(Boleni)	Zimbabwe	GUTs	Lawrence & Mackenzie, 1980
13	Entebbe	Uganda	Cell culture	Robson *et al.*, 1978
14	Sat- (Satinsyi)	Rwanda	GUTs	-
15	Ond - (Onderstepoort)	South Africa	GUTs	Neitz, 1948
16	B11-(Busia)	Kenya	Blood	-
17	B12-(Busia)	Kenya	Blood	-
18	Mug - (Muguga)	Kenya	Salivary glands	Brocklesby *et al.*, 1961
19	Kia - (Kiambu)	Kenya	Salivary glands	Irvin *et al.*, 1974
20	Mar - (Marikebuni)	Kenya	GUTs	Irvin *et al.*, 1983
21	N4 – (Marula)	Kenya	Blood	Bishop *et al.*, 2015
22	M30 (Marula)	Kenya	Blood	Bishop *et al.*, 2015
23	N62 Marula	Kenya	Blood	Bishop *et al.*, 2015
24	N11 Marula	Kenya	Blood	Bishop *et al.*, 2015
25	N7 Marula	Kenya	Blood	Bishop *et al.*, 2015
26	N75 Marula	Kenya	Blood	Bishop *et al.*, 2015
27	N25 Marula	Kenya	Blood	Bishop *et al.*, 2015
28	N81 Marula	Kenya	Blood	Bishop *et al.*, 2015
29	N49 Marula	Kenya	Blood	Bishop *et al.*, 2015
30	N40 Marula	Kenya	Blood	Bishop *et al.*, 2015
31	N44 Marula	Kenya	Blood	Bishop *et al.*, 2015
32	N12 Marula	Kenya	Blood	Bishop *et al.*, 2015
33	N47 Marula	Kenya	Blood	Bishop *et al.*, 2015

4.2 p104 gene sequence variation from multiple geographically separated regions revealed limited polymorphism.

Of the 33 sequences generated in this study, 10 p104 alleles with a mean pairwise identity at the nucleotide sequence level of 98.3% were revealed. The 10 p104 allelic sequences are deposited in the NCBI online database under accession numbers: MZ798149–MZ798158 and are also listed in appendix 1**.** The pairwise identity of the 10 genotypes is summarised in Table 4.2. Notably, the Cameroonian p104 allele (MZ798151) which was 100% identical to a Tanzanian allele had pairwise identities ranging between 98 – 99 % with the other nine genotypes (Table 4.2). The 10 p104 genotypes were shared among the 33 amplified samples as summarised in Table 4.3. Allele MZ798149 was the most frequent being shared with 10/33 (30.3%) and this included the Muguga vaccine p104 amplification and the historical Onderstepoort isolate from South Africa. Moreover, two known buffalo isolates from Tanzania (Pugu and Buffalo-Manyara formed one allele (MZ798151) while the Busia field samples formed one allele with the Kiambu and Marekebuni vaccine isolates.

The nucleotide statistics from the multiple sequence alignment of the 10 unique sequences was as follows: Mean length post-trimming of low-quality bases at the ends of the sequences: 395 bp, identical sites: 391 (95.6%), pairwise identity: 98.5%. The frequencies of the individual nucleotides were. **A**: 36.6%, **C**: 23.3%, **G** 21.5%, **T**: 18.7%

and the **GC** content was 44.7%. Overall, the 10 genotypes had limited polymorphism as depicted in the alignment found in appendix 2, confirming the conservation of the p104 gene.

Table 4.2: Percent pairwise identities (%) among the p104 sequences from this study (CA, KE, RW, TZ, UG, ZA, ZM = Cameroon, Kenya, Rwanda, Tanzania, Uganda, Zambia & Zimbabwe respectively). The accession numbers with more than one country initial indicate the allele was shared as also shown in table 4.3.

	MZ798149	**MZ798150 (ZA,KE,UG,RW)**	**MZ798155 (KE)**	**MZ798156 (KE)**	**MZ798151 (CA, TZ)**	**MZ798157 (KE)**	**MZ798158 (ZM)**	**MZ798152 (ZA, KE)**	**MZ798153 (KE)**	**MZ798154 (KE)**
MZ798149		99	98	98	98	98	98	99	98	98
MZ798150 (ZA,KE,UG,RW)	99		100	99	99	99	99	99	98	99
MZ798155 (KE)	98	100		100	99	99	99	99	98	99
MZ798156 (KE)	98	99	100		99	99	99	99	98	99
MZ798151 (CA, TZ)	98	99	99	99		99	99	98	98	98
MZ798157 (KE)	98	99	99	99	99		99	99	97	98
MZ798158 (ZM)	98	99	99	99	99	99		98	98	98
MZ798152 (ZA, KE)	99	99	99	99	98	99	98		99	99
MZ798153 (KE)	98	98	98	98	98	97	98	99		100
MZ798154	98	99	99	99	98	98	98	99	100	

Table 4.3: p104 alleles, frequency and sharing among the isolates used in this study.

Accession no.	Allele frequency (no. of isolates)	Isolates sharing allele
MZ798149	10	Muguga (Kenya), Serengeti (Tanzania), Onderstepoort (South Africa), Monduli - 13, 15, 16, 29, 36, 136, M141 (Tanzania)
MZ798150	7	Marikebuni (Kenya), Satinsyi (Rwanda), Kiambu (Kenya), Busia (Kenya), Marula -25, 30, 47(Kenya)
MZ798151	2	Pugu (Tanzania), Buffalo-Manyara (Tanzania)
MZ798152	1	Marula 12 (Kenya)
MZ798153	1	Marula 75 (Kenya)
MZ798154	1	Marula 7 (Kenya)
MZ798155	1	Marula 44 (Kenya)
MZ798156	1	Marula 11 (Kenya)
MZ798157	1	Marula 40 (Kenya)
MZ798158	1	Boleni (Zimbabwe)

As shown in figure 4.2, the sequence variation across the 496 bp amplicon as defined by the mean base pair differences across the 10 genotypes from this study was 6 (± 2.2 SD; range = 2 – 10; Median = 5). This limited variation, evident by the number of bases which are not identical, is illustrated by the box plot in figure 4.2 which shows the distribution of base pair differences among the p104 alleles in the isolates studied herein.

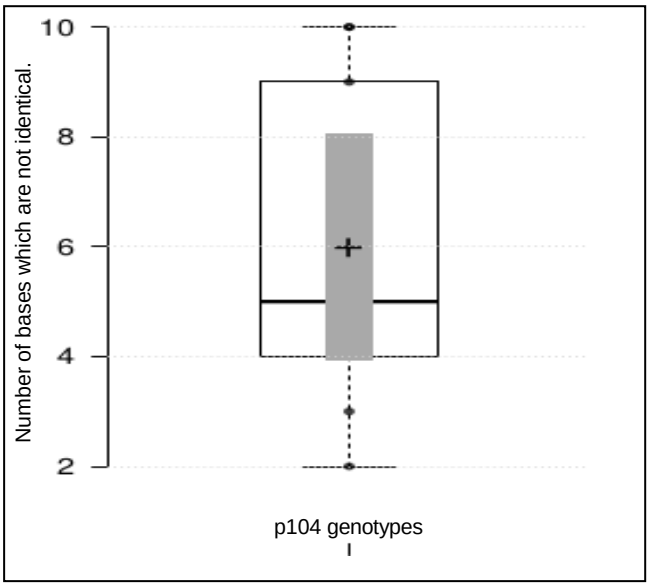


Figure 4.2: Base pair differences among the 10 p104 genotypes. Centre lines show the medians; box limits indicate the 25th and 75th percentiles as determined by R software; whiskers extend 1.5 times the interquartile range from the 25th and 75th percentiles, outliers are represented by dots; crosses represent sample means; bars indicate 95% confidence intervals of the means. n = 10 sample points.

4.3 The phylogenetic relatedness inferred from the p104 sequences does not correlate with the geographical origin of the isolates.

To ascertain the relationships among the isolates genotyped herein and to assess whether the resolved subgroups correlate with geographical origin, our dataset for phylogenetic inference additionally comprised 23 published p104 sequences from across the endemic region.

The AIC ranking of nucleotide substitution models resulted in the most support for the K80+I model (Akaike weight = 0.239072). The likelihood scores, model selection criteria and numerical values for model parameters are summarized in Table 4.4.

Table 4.4: The likelihood scores, model selection criteria and numerical values for model parameters (The full run results of the JmodelTest2 is shown in appendix 3).

Model	-lnL	K	AICc	delta	weight	cumWeight
K80+I	595.59456	68	1371.874834	0.000000	0.239072	0.239072
HKY+I	590.89823	71	1373.187764	1.312930	0.124002	0.363075
JC+I	598.37941	67	1373.943654	2.068820	0.084975	0.448050
TPM1+I	595.10147	69	1374.423036	2.548201	0.066864	0.514914
TrNef+I	95.20330	69	1374.626696	2.751861	0.060391	0.575305
TPM2+I	595.29533	69	1374.810756	2.935921	0.055081	0.630386
F81+I	593.59783	70	1374.984122	3.109287	0.050508	0.680893
TPM3+I	595.52627	69	1375.272636	3.397801	0.043723	0.724616
TPM2uf+I	590.27243	72	1375.573986	3.699152	0.037607	0.762223
TrN+I	590.31083	72	1375.650786	3.775952	0.036190	0.798413

-lnL: negative log likelihood
K: number of estimated parameters
AICc: Corrected Akaike Information Criterion
delta: AICc difference
weight: AICc weight
cumWeight: cumulative AICc weight

The maximum likelihood phylogeny based on the K80+I model resulted in the tree shown in figure 4.3. The phylogeny analysis did not reveal any apparent differentiation based on geographic origin as shown by the multiple clustering together of sequences from diverse origin.

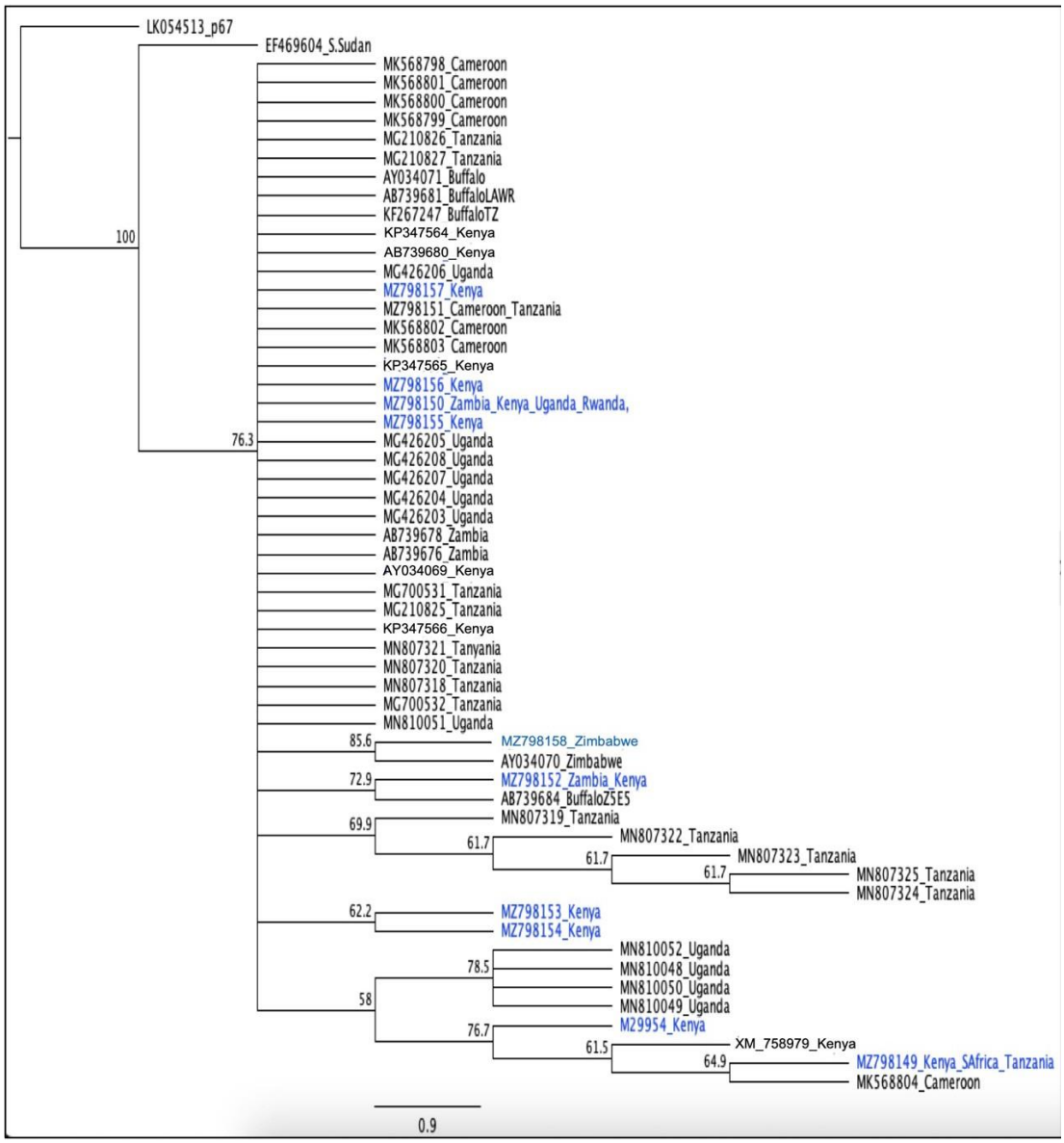


Figure 4.3: Maximum likelihood tree depicting the phylogenetic relationships of *Theileria parva* p104 sequences. The accession numbers in blue font indicates the 10 p104 allelic variants among the 33 isolates investigated in the present study. Also included in the tree are p104 sequences from previous field studies widely distributed across the geographical range of the parasite, including eastern central and southern Africa.

4.4 Evaluation for positive selection in the p104 coding sequences.

The analysis of the positions in the p104 sequences coding for residues that show an excess of non-synonymous (dN) over synonymous substitutions (dS) did not reveal any sites with higher rates of non-synonymous to synonymous nucleotide substitutions than expected under neutral evolution. Both M2a and M8 models used similar numbers of parameters and the same sites under selection were identified (Appendix 4), albeit, not supported by the BEB analysis as depicted by < 95 % posterior probabilities of positive selection. A summary of the selection results is shown in Table 4.5.

Table 4.5: Summary of selection results showing model parameter estimates and positively selected sites under the M2a and M8 models using Bayes empirical Bayes (BEB) analysis.

Model	K	LnL	Parameter estimates				Positive sites
			ω	p0	p1	P2	
M2a	68	-345.2150	2.25648	0.74700	0.00000	0.25300	46 A (0.633), 51 A (0.516), 52 S (0.581)
M8	68	-345.2150	2.25648	0.74700	0.00500	0.25300	46 A (0.752), 51 A (0.648), 52 S (0.707)

K indicates number of estimated parameters; LnL refers to maximised log likelihoods; p0 (purifying), p1(neutral) and p2 (positive) denote the proportion of codons belonging to each site class, while ω represents the dN/dS for the positive selection site class only. The brackets () enclose the posterior probabilities of positive selection. The codon numbers are identified with reference to sequence AB739681 retrieved from GeneBank.

CHAPTER FIVE

DISCUSSION

The p104 gene is the standard diagnostic marker for *T. parva* based on a nested PCR procedure designed to provide enhanced sensitivity for detection of *T. parva* genomic DNA in cattle and buffalo blood, using an outer primer pair for the first amplification, and an additional set of inner primers for the second amplification. In the present study, the outer primers and DNA extracted from a range of *T. parva* samples (n=101) obtained from multiple geographically separated areas in the endemic region were used for the amplification of a 496 bp fragment of the p104 gene.

Only 33 samples were successfully amplified and sequenced with 67.32 % (68/101) failing to generate a PCR product or a clean sequence for downstream analysis. Notably, all the unsuccessful amplifications were field samples. However, the lack of amplification with the outer primers does not necessarily imply that the animals were negative, and a nested PCR would have to be done as this has enhanced sensitivity for detection of the carrier animals in the field (Odongo *et al.,* 2010). However, it's important to note that as described in Odongo *et al.,* (2010), the nested PCR primers, differ from the outer primers and will generate a 277 bp fragment. The 496 bp fragment that is generated by the outer

primer set was the target of this as it presents more opportunities for discovering variation in the gene.

The amplified samples for Marula and Monduli archived blood samples was higher than the Busia samples. Unlike Busia, these two areas from central Kenya and Northern Tanzania are wildlife-livestock interface regions where cattle co-graze with the African cape buffalo, which is the main *T. parva* reservoir host (Bishop *et al.,* 2015, Mwamuye *et al.,* 2020; Morrison *et al.,* 2020). It is, therefore, more probable that cattle from these two wildlife-livestock interface areas are exposed to *T. parva* infected ticks than in Busia.

Nevertheless, the study analysis of the 33 sequences revealed that the most frequent p104 genotype was identical to isolates from multiple geographically widely separated regions. While it is not yet possible, due to the limited depth of coverage of field sampling performed to date, to be certain that all p104 genotypes will be widespread geographically in East, Central and Southern Africa, it seems likely based on this and other findings that the majority will be widespread. Notably, the most frequent allele was identical to the Muguga vaccine p104 isolate from Kenya, the Serengeti isolate from Tanzania and the historical Onderstepoort isolate from South Africa originally isolated in 1937 from the farm Schoonspruit in the Transvaal before ECF was eradicated in the country (Neitz, 1948). Additionally, the Busia (western Kenya) field samples formed one allele with the

Kiambu (central Kenya) and Marekebuni (coastal Kenya) vaccine isolates and Satiyinsi isolate from Rwanda.

It is also interesting to note that the widespread conservation of p104 genotypes extends to buffalo derived parasites in Tanzania and South Africa (Sibeko *et al.*, 2011 and Mwega et al, 2014), yet buffalo are known to harbour a much greater diversity of *T. parva* genotypes (Pelle et al. 2011; Hayashida et al 2012). Furthermore, although the issue of genetic similarity between parasite populations in cattle and buffalo in northern Tanzania is yet to be addressed, previous studies have shown that asymptomatic cattle in northern Tanzania region are infected with buffalo derived *T. parva* genotypes, consistent with recent transmission from buffalo (Mwamuye *et al.*, 2020). In the context of the present study, it is interesting that p104 variants from Monduli, a livestock-wildlife interface area in northern Tanzania were similarly conserved.

The frequently observed p104 variants were present in the three component stocks of the Muguga cocktail used for the ITM live immunisation procedure. This observation is consistent with parasites with similar p104 variants to Muguga cocktail circulating naturally in the field, rather than disseminated following ITM deployment (Bishop *et al.*, 2020). This is because as already described in the results, the most frequent p104 alleles were also present in isolates from multiple geographically widely separated regions in Zambia, Uganda, Kenya, Tanzania, Rwanda and South Africa. Besides, only small-scale ITM tri-

als have been undertaken in Uganda. The frequent p104 variants also match those in the Marikebuni vaccine stabilate which has not been sufficiently widely used for ITM to support the vaccination dissemination hypothesis (Bishop *et al.,* 2020). It is also important to note that previous molecular epidemiology and population genetics studies have similarly shown that parasite diversity does not correlate with geographic origin (Oura *et al.,* 2005; Odongo *et al.,* 2006).

More importantly, no variability was present in the primer sets used for the nested PCR procedure designed to provide enhanced sensitivity for *T. parva* detection based on amplification of the p104 gene. The lack of variability in the primers has the implication that polymorphisms in the p104 gene are unlikely to impact the PCR procedure for detection of *T. parva*.

Phylogenetic analysis resolved the p104 sequences into two evolutionary lineages. The major finding, the observation that the evolutionary lineages do not correlate with geographic origin corroborates findings from earlier studies based on (VNTR) sequences and sequences of genes that are targets of the CD8+ T cell responses (Pelle *et al.,* 2011; Odongo *et al.,* 2006). Furthermore, the present study failed to demonstrate evolutionary patterns consistent with positive selection, a finding that is consistent with previous analyses of selection at other *T. parva* loci (Obara *et al.,* 2015; Mwamuye *et al*. 2020).

CONCLUSION AND RECOMMENDATIONS

This study has evaluated the conservation and variations in the p104 antigen gene from *T. parva* samples and sequences across the wider range of the parasite's geospatial distribution in sub-Saharan Africa. Among the key observations is that there are some mismatches within the primer sequences but with no evidence for positive selection among those p104 mutations that resulted in residue changes. Further, the minor variation observed does not correlate with the geographic origin of the parasite samples considered in this study, although this may need to be confirmed by genotyping additional field samples. Thus, the data indicate that the p104 amplicons can also provide important data on parasite genotype if sequenced.

However, with the observations of variation contrary to the previous belief of the p104 gene conservation, it remains probable that there could exist some divergent p104 alleles that may fail to amplify with the current diagnostic primers. However, any primer mismatches can still be accommodated by the use of degenerate bases as part of ongoing reassessment to accommodate increasing knowledge of genetic heterogeneity of diagnostic targets.

RECOMMENDATION

Therefore, this study recommends further in-depth surveys especially in the regions that parasite considered to be emerging—towards the West and North of Africa.

REFERENCES

Abbas, R. Z., Zaman, M. A., Colwell, D. D., Gilleard, J., & Iqbal, Z. (2014). Acaricide resistance in cattle ticks and approaches to its management: the state of play. *Veterinary parasitology*, *203*(1-2), 6–20. https://doi.org/10.1016/j.vetpar.2014.03.006

Agina OA, Shaari MR, Isa NMM, Ajat M, Zamri-Saad M, Hamzah H. (2020). Clinical Pathology, Immunopathology and Advanced Vaccine Technology in Bovine Theileriosis: A Review. *Pathogens*, 25(9):697. Doi: 10.3390/pathogens9090697.

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

Allan, F. K., & Peters, A. R. (2021). Safety and Efficacy of the East Coast Fever Muguga Cocktail Vaccine: A Systematic Review. *Vaccines*, *9*(11), 1318. http://dx.doi.org/10.3390/vaccines9111318.

Allan, F. K., Jayaraman, S., Paxton, E., Sindoya, E., Kibona, T., Fyumagwa, R., Mramba, F., Torr, S. J., Hemmink, J. D., Toye, P., Lembo, T., Handel, I., Auty, H. K., Morrison, W. I., & Morrison, L. J. (2021). Antigenic Diversity in *Theileria parva* Populations From Sympatric Cattle and African Buffalo Analyzed Using Long

Read Sequencing. *Frontiers in genetics*, 12, 684127. https://doi.org/10.3389/fgene.2021.684127.

Arisue, N., & Hashimoto, T. (2015). Phylogeny and evolution of apicoplasts and apicomplexan parasites. *Parasitology international*, 64(3), 254–259. https://doi.org/10.1016/j.parint.2014.10.005.

Bishop, R.P., Kappmeyer, L.S., Onzere, C.K. *et al.* Equid infective *Theileria* cluster in distinct *18S* rRNA gene clades comprising multiple taxa with unusually broad mammalian host ranges. *Parasites Vectors* **13**, 261 (2020). https://doi.org/10.1186/s13071-020-04131-0

Bishop, R., Allsopp, B. Spooner, P., Sohanpal, B., Morzaria, S., Gobright, E. (1995). *Theileria*: improved species discrimination using oligonucleotides derived from large subunit ribosomal RNA sequences. *Experimental. Parasitolology,* 80 107-115

Bishop, R., Hemmink, J. D., Morrison, W. I., Weir, W., Toye, P. G., Sitt, T., Spooner, P. R., Musoke, A. J., Skilton, R. A., & 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*(3), 333–342. https://doi.org/10.1016/j.ijppaw.2015.08.006

Bishop, R., Musoke, A., Morzaria, S., Gardner, M., & Nene, V. (2004). *Theileria*: intracellular protozoan parasites of wild and domestic ruminants transmitted by ixodid ticks. *Parasitology*, 129 Suppl, S271–S283. https://doi.org/10.1017/s0031182003004748.

Bishop, R., Nene, V., Staeyert, J., Rowlands, J., Nyanjui, J., Osaso, J., Morzaria, S., Musoke, A. (2003). Immunity to East Coast fever in cattle induced by a polypeptide fragment of the major surface coat protein of *Theileria parva* sporozoites. *Vaccine,* 21, 1205-1212

Brocklesby, D. W., Barnett, S. F., & Scott, G. R. (1961). Morbidity and Mortality Rates in East Coast Fever (*Theileria Parva* Infection) and their Application to Drug Screening Procedures. *British Veterinary Journal*, *117*(12), 529–531. https://doi.org/10.1016/s0007-1935(17)43303-3

Chatanga, E., Hayashida, K., Muleya, W., Kusakisako, K., Moustafa, M. A. M., Salim, B., Katakura, K., et al. (2020). Genetic Diversity and Sequence Polymorphism of Two Genes Encoding *Theileria parva* Antigens Recognized by CD8+ T Cells among Vaccinated and Unvaccinated Cattle in Malawi. *Pathogens*, 9(5), 334. MDPI AG. Retrieved from http://dx.doi.org/10.3390/pathogens9050334.

Chaussepied, M., Janski, N., Baumgartner, M., Lizundia, R., Jensen, K., Weir, W., Shiels, B.R.,Weitzman, J.B., Glass, E.J., Werling, D. & Langsley, G. (2010).TGF-β2 induction regulates invasiveness of *Theileria*-transformed leukocytes and disease susceptibility. *PLoS Pathogens*, 6 e1001197.

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., & Young, A. S. (1989). Characterization of buffalo-derived theilerial parasites with monoclonal antibodies and DNA probes. *Parasitology*, 98 Pt 2, 179–188. https://doi.org/10.1017/s0031182000062089

Conrad, P.A., Denham, D., & Brown, C.G. (1986). Intraerythrocytic multiplication of *Theileria parva in vitro*: an ultrastructural study. *International Journal of Parasitology,* 16, 223-229. https://doi.org/10.1016/0020-7519(86)90047-0.

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

Darriba, D., Taboada, G. L., Doallo, R. R., & Posada, D. (2012). JModelTest 2: More models, new heuristics and parallel computing. *Nature Methods*, *9*(8), 772. https://doi.org/10.1038/nmeth.2109

De Deken, R., Martin, V., Saido, A., Madder, M., Brandt, J.& 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

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

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

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

Ehi-Eromosele, C. O., Nwinyi, O. C., & Ajani, O. O. (2013). Integrated Pest Management. *InTech*. https://doi.org/10.5772/54476.

Emery D. L. (1981). Adoptive transfer of immunity to infection with *Theileria parva* (East Coast fever) between cattle twins. *Research in veterinary science*, *30*(3), 364–367.

Fawcett, D. W., Doxsey, S., Stagg, D. A., & Young, A. S. (1982). The entry of sporozoites of *Theileria parva* into bovine lymphocytes in vitro. Electron microscopic observations. *European journal of cell biology*, 27(1), 10–21.

Fry, L. M., Schneider, D. A., Frevert, C. W., Nelson, D. D., Morrison, W. I., & Knowles, D. P. (2016). East Coast Fever Caused by *Theileria parva* Is Characterized by Macrophage Activation Associated with Vasculitis and Respiratory Failure. PloS one, 11(5), e0156004. https://doi.org/10.1371/journal.pone.0156004

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

Gentschev, I., Glaser, I., Goebel, W., McKeever, D.J., Musoke, A. & Heussler, V.T. (1998). Delivery of the p67 sporozoite antigen of *Theileria parva* by using recombinant *Salmonella dublin*: secretion of the product enhances specific antibody responses in cattle. *Infectious Immunology*, 66, 2060-2064

Gharbi, M., Darghouth, M. A., Elati, K., Al-Hosary, A. A. T., Ayadi, O., Salih, D. A., El Hussein, A. M., Mhadhbi, M., Khamassi Khbou, M., Hassan, S. M., Obara, I., Ahmed, L. S., & Ahmed, J. (2020). Current status of tropical theileriosis in Northern Africa: A review of recent epidemiological investigations and implications for control. *Transboundary and emerging diseases, 67 Suppl 1*, 8–25. https://doi.org/10.1111/tbed.13312

Gharbi, M., Rekik, B., Mabrouk, M., Hassni, M., Zroud, W., Mhadhbi, M., Sassi, L., Jedidi, M., & Darghouth, M. (2015). Impact of the carrier state by *Theileria annulata* on milk yield in Tunisian crossbred (*Bos taurus*) cattle. *Asian Pacific Journal of Tropical Medicine*, 5, 884–887.

Githaka, N. W., Kanduma, E. G., Wieland, B., Darghouth, M. A., & Bishop, R. P. (2022). Acaricide resistance in livestock ticks infesting cattle in Africa: Current status and potential mitigation strategies. *Current research in parasitology & vector-borne diseases*, *2*, 100090. https://doi.org/10.1016/j.crpvbd.2022.100090

Graham, S. P., Pellé, 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., Taracha, E. L. N. (2006). *Theileria parva* candidate vaccine antigens recognized by immune bovine cytotoxic T lymphocytes. *Proceedings of the National Academy of Sciences of the United States of America, 103*(9), 3286–3291. https://doi.org/10.1073/pnas.0511273103

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., Taracha, E.L. (2008). Characterization of the fine specificity of bovine CD8 T-cell responses to defined antigens from the protozoan parasite *Theileria parva. Infectious Immunology*, 76, 685-694

Gubbels, M.J., De vos, P.A., Van der weide, M., Viseras, J., Schouls, M.L., De Vries, E. & Jongejan, F. (1999). Simultaneous detection of bovine *Theileria* and *Babesia* species by reverse line bolt hybridization. *Journal of clinical microbiology*, 37(6): 1782-1789.

Guindon, S., Dufayard, J.-F., Lefort, V., Anisimova, M., Hordijk, W., & Gascuel, O. (2010). New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. *Systematic Biology*, *59*(3), 307–321. https://doi.org/10.1093/sysbio/syq010

Hacılarlıoglu, S., Bilgic, H. B., Bakırcı, S., Tait, A., Weir, W., Shiels, B., & Karagenc, T. (2023). Selection of genotypes harbouring mutations in the cytochrome b gene of *Theileria annulata* is associated with resistance to buparvaquone. PloS one, 18(1), e0279925. https://doi.org/10.1371/journal.pone.0279925.

Hayashida, K., Abe, T., Weir, W., Nakao, R., Ito, K., Kajino, K., Suzuki,Y., Jongejan, F., Geysen, D., 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, K., Katayama, S., Wakaguri, T., Inoue, H., Homma,N., 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. & 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-e00212

Henson, S., Bishop, R. P., Morzaria, S., Spooner, P. R., Pelle, R., Poveda, L., Ebeling, M., Küng, E., Certa, U., Daubenberger, C. A., & 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*(1), 503. https://doi.org/10.1186/1471-2164-13-503

Heussler, V.T., Taracha, E.L., Musoke, A., Duchateau, L., McKeever, D.J. (1998). Immunisation with live attenuated *Salmonella dublin* expressing a sporozoite protein confers partial protection against *Theileria parva. Vaccine,* 16, 834-841

Irvin, A. D., Dobbelaere, D. A., Mwamachi, D. M., Minami, T., Spooner, P. R., & Ocama, J. G. (1983). Immunisation against East Coast fever: correlation between monoclonal antibody profiles of *Theileria parva* stocks and cross immunity in vivo. *Research in Veterinary Science*, *35*(3), 341–346. https://doi.org/10.1016/s0034-5288(18)32030-7

Irvin, A. D., Purnell, R. E., Brown, C. G. D., Cunningham, M. P., Ledger, M. A., & Payne, R. C. (1974). The application of an indirect method of infecting ticks with piroplasms for use in the isolation of field infections. *British Veterinary Journal*, *130*(3), 280–287. https://doi.org/10.1016/S0007-1935(17)35894-3

Kaaya, G. P., Mwangi, E. N., & Ouna, E. A. (1996). Prospects for biological control of livestock ticks, *Rhipicephalus appendiculatus* and *Amblyomma variegatum*, using the entomogenous fungi *Beauveria bassiana* and *Metarhizium anisopliae*. *Journal of invertebrate pathology*, 67(1), 15–20. https://doi.org/10.1006/jipa.1996.0003

Kabi, F., Masembe, C., Muwanika, V., Kirunda, H., & 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. https://doi.org/10.1186/1756-3305-7-414.

Kasaija, P. D., Estrada-Peña, A., Contreras, M., Kirunda, H., & de la Fuente, J. (2021). Cattle ticks and tick-borne diseases: a review of Uganda's situation. *Ticks and tick-borne diseases*, 12(5), 101756. https://doi.org/10.1016/j.ttbdis.2021.101756.

Katende, J., Morzaria, S., Toye, P., Skilton, R., Nene, V., Nkonge, C., & Musoke, A. (1998). An enzyme-linked immunosorbent assay for detection of *Theileria parva* antibodies in cattle using a recombinant polymorphic immunodominant molecule. *Parasitology Research*, *84*(5), 408–416. https://doi.org/10.1007/s004360050419

Katris, N. J., van Dooren, G. G., McMillan, P. J., Hanssen, E., Tilley, L., & Waller, R. F. (2014). The apical complex provides a regulated gateway for secretion of invasion factors in Toxoplasma. *PLoS pathogens*, *10*(4), e1004074. https://doi.org/10.1371/journal.ppat.1004074.

Kipronoh, A. K. (2009). The Epidemiology and Control of East Coast fever and other Vector-borne Diseases: Perceptions of the Pastoral Communities in Northern Rift Valley Province Kenya (M.Sc. Thesis). The University of Nairobi.

Lawrence, J., & Mackenzie, P. K. (1980). Isolation of a non-pathogenic Theileria of cattle transmitted by Rhipicephalus appendiculatus. *Zimbabwe Veterinary Journal*, *11*, 27–35.

Liu, J. Guan, G. & Yin, H. (2022). *Theileria annulata*: Parasite of the Month. *Trends in parasitology*. 38 (3): 265-266. https://doi.org/10.1016/j.pt.2021.11.001

Malak, A. K., Mpoke, L., Banak, J., Muriuki, S., Skilton, R. A., Odongo, D., Sunter, J., & Kiara, H. (2012). Prevalence of livestock diseases and their impact on livelihoods in Central Equatoria State, southern Sudan. *Preventive Veterinary Medicine*, *104*(3–4), 216–223. https://doi.org/10.1016/j.prevetmed.2011.12.001

Mans, B. J., Pienaar, R., & Latif, A. A. (2015). A review of Theileria diagnostics and epidemiology. International journal for parasitology. Parasites and wildlife, 4(1), 104–118. https://doi.org/10.1016/j.ijppaw.2014.12.006.

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., & 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*(4), 1229–1235. https://doi.org/10.1111/tbed.12495

Mbassa, G. K., Kipanyula, M. J., Mellau, L. S., Mwamakali, E. D., Bulegeya, F. R., & 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(3-4), 260–270. https://doi.org/10.1016/j.vetpar.2006.07.010.

McHardy, N., Wekesa, L. S., Hudson, A. T., & Randall, A. W. (1985). Antitheilerial activity of BW720C (buparvaquone): a comparison with parvaquone. *Research in veterinary science*, *39*(1), 29–33.

McKeever, D. J., Evans L. N. Taracha, Innes, E. L., MacHugh, N. D., Awino, E., Goddeeris, B. M., & Morrison, W. I. (1994). Adoptive Transfer of Immunity to *Theileria parva* in the CD8+ Fraction of Responding Efferent Lymph. *Proceedings of the National Academy of Sciences of the United States of America*, 91(5), 1959–1963.

Mhadhbi, M., Chaouch, M., Ajroud, K., Darghouth, M. A., & BenAbderrazak, S. (2015). Sequence Polymorphism of Cytochrome b Gene in *Theileria annulata* Tunisian Isolates and Its Association with Buparvaquone Treatment Failure. *PloS one*, 10(6), e0129678. https://doi.org/10.1371/journal.pone.0129678.

Minami, T., Spooner, P. R., Irvin, A. D., Ocama, J. G., Dobbelaere, D. A., & Fujinaga, T. (1983). Characterisation of stocks of *Theileria parva* by monoclonal antibody profiles. *Research in veterinary science,* 35(3), 334–340.

Morrison, W. I., Hemmink, J. D., & Toye, P. G. (2020). *Theileria parva*: a parasite of African buffalo, which has adapted to infect and undergo transmission in cattle. *International journal for parasitology*, 50(5), 403–412. https://doi.org/10.1016/j.ijpara.2019.12.006.

Mwamuye, M. M., Odongo, D., Kazungu, Y., Kindoro, F., Gwakisa, P., Bishop, R. P., Nijhof, A. M., & Obara, I. (2020). Variant analysis of the sporozoite surface antigen gene reveals that asymptomatic cattle from wildlife-livestock interface areas in northern Tanzania harbour buffalo-derived *T. parva*. *Parasitology research,* 119(11), 3817–3828. https://doi.org/10.1007/s00436-020-06902-1.

Mugambi, I. R. (2012). The factors influencing the collapse of communal cattle dips after privatization of veterinary services in Kenya: a case of Maara District, Tharaka-Nithi County. Master's thesis, University of Nairobi.

Musoke, A. J., Nantulya, V. M., Rurangirwa, F. R., & Buscher, G. (1984). Evidence for a common protective antigenic determinant on sporozoites of several *Theileria parva* strains. *Immunology*, *52*(2), 231–238.

Nanteza, A., Nsadha, Z., Nsubuga, J., Oligo, S., Kazibwe, A., Terundaja, C., Matovu, E., & Lubega, G. W. (2023). Assessment of the Impact of Early Diagnosis and Early Treatment in the Integrated Control of East Coast Fever (ECF) Involving Acquired Immunity Induced by Natural Infection in Ankole Cattle. *Pathogens*, 12(1), 115. https://doi.org/10.3390/pathogens12010115

Nanteza, A., Obara, I., Kasaija, P., Mwega, E., Kabi, F., Salih, D.A., Njahira, M., Njuguna, J., Odongo, D., Bishop, R.P., Skilton, R.A., Ahmed, J., Clausen, P-H. & Lubega, G.W. (2020). "Antigen gene and variable number tandem repeat (VNTR) diversity in Theileria parva parasites from Ankole cattle in south-western Uganda: Evidence for conservation in antigen gene sequences combined with extensive polymorphism at VNTR loci." *Transbound Emerging Diseases*. ;67 Suppl 1:99-107.

Narasimhan, S., Kurokawa, C., DeBlasio, M., Matias, J., Sajid, A., Pal, U., Lynn, G., & Fikrig, E. (2021). Acquired tick resistance: The trail is hot. *Parasite immunology*, *43*(5), e12808. https://doi.org/10.1111/pim.12808

Neitz, W. O. (1948). Studies on East Coast Fever. *S.Afr. Sci*, *1*, 133–135.

Nene, V., & Morrison, W. I. (2016). Approaches to vaccination against *Theileria parva* and *Theileria annulata*. *Parasite immunology*. 38 (12): 724-734. https://doi.org/10.1111/pim.12388

Nene. V, Kiara, H., Lacasta, A., Pelle, R., Svitek, N., Steinaa, L. (2016) The biology of *Theileria parva* and control of East Coast fever - Current status and future trends. *Ticks Tick Borne Diseases*. 7(4):549-64. doi: 10.1016/j.ttbdis.2016.02.001.

Nielsen, R., & Yang, Z. (1998). Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. *Genetics*, *148*(3), 929–936. https://doi.org/10.1093/genetics/148.3.929

Norval, R. A. I., Perry, B. D., & Young, A. (1992). *The epidemiology of theileriosis in Africa*. Academic Press.

Obara, I., Ulrike, S., Musoke, T., Spooner, P. R., Jabbar, A., Odongo, D., Kemp, S., Silva, J. C., & Bishop, R. P. (2015). Molecular evolution of a central region containing B cell epitopes in the gene encoding the p67 sporozoite antigen within a field population of *Theileria parva*. *Parasitology research*, 114(5), 1729–1737. https://doi.org/10.1007/s00436-015-4358-6.

Obara, I., Nielsen, M., Jeschek, M., Nijhof, A., Mazzoni, C. J., Svitek, N., Steinaa, L., Awino, E., Olds, C., Jabbar, A., Clausen, P. H., & Bishop, R. P. (2016). Sequence diversity between class I MHC loci of African native and introduced Bos taurus cattle in *Theileria parva* endemic regions: in silico peptide binding prediction identifies distinct functional clusters. *Immunogenetics*, 68(5), 339–352. https://doi.org/10.1007/s00251-016-0902-5

Odongo, D. O., Sunter, J. J. D., Kiara, H. H. K., Skilton, R. R. A., & Bishop, R. P. R. (2010). A nested PCR assay exhibits enhanced sensitivity for detection of Theileria parva infections in bovine blood samples from carrier animals. *Parasitology Research*, *106*(2), 357–365. https://doi.org/10.1007/s00436-009-1670-z

Odongo, D. O., Oura, C. A., Spooner, P. R., Kiara, H., Mburu, D., Hanotte, O. H., & 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(8), 937–946. https://doi.org/10.1016/j.ijpara.2006.01.008.

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. https://doi.org/10.1016/j.jaridenv.2007.04.003.

Oura, C. A. L., Odongo, D. O., Lubega, G. W., Spooner, P. R., Tait, A., & 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*(14), 1641–1653. https://doi.org/10.1016/S0020-7519(03)00280-7

Oura, C. A., Asiimwe, B. B., Weir, W., Lubega, G. W., & Tait, A. (2005). Population genetic analysis and sub-structuring of *Theileria parva* in Uganda. *Molecular and biochemical parasitology*, *140*(2), 229–239. https://doi.org/10.1016/j.molbiopara.2004.12.015

Patel, E., Mwaura, S., Di Giulio, G., Cook, E. A. J., Lynen, G., & Toye, P. (2019). Infection and treatment method (ITM) vaccine against East Coast fever: reducing the number of doses per straw for use in smallholder dairy herds by thawing, diluting and refreezing already packaged vaccine. *BMC Veterinary Research*, 15, 46. https://doi.org/10.1186/s12917-019-1787-y.

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., & 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(4), e19015. https://doi.org/10.1371/journal.pone.0019015.

Radley, D. E., Brown, C. G., Cunnigham, M. P., Kimber, C. D., Musisi, F. L., Purnell, R. E., & Stagg, S. M. (1975). East Coast fever: challenge if immunised cattle by prolonged exposure to infected ticks. *The Veterinary record*, 96(24), 525–527. https://doi.org/10.1136/vr.96.24.525.

Robson, J., Pedersen, V., Odeke, G. M., Kamya, E. P., & Brown, C. G. D. (1978). East coast fever immunisation trials in Uganda: Field exposure of zebu cattle immunised with three isolates of *Theileria parva*. *Tropical Animal Health and Production*, *10*(1), 62. https://doi.org/10.1007/BF02235307.

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., & Skilton, R. A. (2018). Genetic diversity and population structure of *Theileria parva* in South Sudan. *Ticks and tick-borne diseases*, *9*(4), 806–813. https://doi.org/10.1016/j.ttbdis.2018.03.002.

Schreuder, B. E., Uilenberg, G., & Tondeur, W. (1977). Studies on Theileriidae (Sporozoa) in Tanzania. VIII. Experiments with African buffalo (*Syncerus caffer*). *Tropenmedizin Und Parasitologie*, *28*(3), 367–371. https://europepmc.org/article/med/410128.

Shapiro, S. Z., Fujisaki, K., Morzaria, S. P., Webster, P., Fujinaga, T., Spooner, P. R., & Irvin, A. D. (1987). A life-cycle stage-specific antigen of Theileria parva recognized by anti-macroschizont monoclonal antibodies. Parasitology, 94 (Pt 1), 29–37. https://doi.org/10.1017/s0031182000053427.

Silatsa, B. A., Simo, G., Githaka, N., Kamga, R., Oumarou, F., Keambou Tiambo, C., Machuka, E., Domelevo, J. B., Odongo, D., Bishop, R., Kuiate, J. R., Njiokou, F., Djikeng, A., & Pelle, R. (2020). First detection of *Theileria parva* in cattle from Cameroon in the absence of the main tick vector *Rhipicephalus appendiculatus*. *Transboundary and Emerging Diseases*, *67*(S1), 68–78. https://doi.org/10.1111/tbed.13425

Sivakumar, T., Hayashida, K., Sugimoto, C., & Yokoyama, N. (2014). Evolution and genetic diversity of *Theileria*. Infection, genetics and evolution: journal of molecular epidemiology and evolutionary genetics in infectious diseases, 27, 250–263. https://doi.org/10.1016/j.meegid.2014.07.013.

Skilton, R. A., Bishop, R. P., Katende, J. M., Mwaura, S., & 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*(3), 265–276. https://doi.org/10.1017/S0031182001001196

Swofford, D. (2003). PAUP*. Phylogenetic analysis using parsimony (*and other methods). Version 4. In *Sinauer Associates, Sunderland, Massachusetts.* Sinauer Associates, Sunderland, Massachusetts.

Tajeri, S. and Langsley, G. (2021). *Theileria* secretes proteins to subvert its host leukocyte. Biol. Cell, 113: 220-233. https://doi.org/10.1111/boc.202000096

Tretina, K., Haidar, M., Madsen-Bouterse, S. A., Sakura, T., Mfarrej, S., Fry, L., Chaussepied, M., Pain, A., Knowles, D. P., Nene, V. M., Ginsberg, D., Daubenberger, C. A., Bishop, R. P., Langsley, G., & Silva, J. C. (2020). *Theileria* parasites subvert E2F signaling to stimulate leukocyte proliferation. Scientific reports, 10(1), 3982. https://doi.org/10.1038/s41598-020-60939-x.

Tretina, K., Gotia, H. T., Mann, D. J., & Silva, J. C. (2015). *Theileria*-transformed bovine leukocytes have cancer hallmarks. Trends in parasitology, 31(7), 306–314. https://doi.org/10.1016/j.pt.2015.04.001.

Toye, P., Nyanjui, J., Goddeeris, B., & Musoke, A. J. (1996). Identification of neutralization and diagnostic epitopes on PIM, the polymorphic immunodominant molecule of *Theileria parva*. Infection and immunity, 64(5), 1832–1838. https://doi.org/10.1128/iai.64.5.1832-1838.1996.

Uilenberg, G., Schreuder, B. E., Silayo, R. S., & Mpangala, C. (1976). Studies on Theileriidae (Sporozoa) in Tanzania. IV. A field trial on immunization against East Coast fever (*Theileria parva* infection of cattle). *Tropenmedizin Und Parasitologie*, *27*(3), 329–336.

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., NextGen Consortium, Joost, S., & 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. https://doi.org/10.3389/fgene.2018.00385

Wanjohi, J. M., Ngeranwa, J. N., Rumberia, R. M., Muraguri, G. R., & Mbogo, S. K. (2001). Immunization of cattle against East Coast fever using *Theileria parva* (Marikebuni) and relaxation of tick control in North Rift, Kenya. In *Kenya. Onderstepoort Journal of Veterinary Research*, 68 (3):217-223

Yang, Z. (2007). PAML 4: phylogenetic analysis by maximum likelihood. *Molecular Biology and Evolution*, *24*(8), 1586–1591. https://doi.org/10.1093/molbev/msm088.

Yang, Z., Nielsen, R., Goldman, N., Mette, A.-, & Pedersen, K. (2000). Codon-Substitution Models for Heterogeneous Selection Pressure at Amino Acid Sites. *Genetics*, 155 (1), 431–449. https://doi.org/10.1093/genetics/155.1.431.

Yang, Z., Wong, W. S. W., & Nielsen, R. (2005). Bayes empirical Bayes inference of amino acid sites under positive selection. *Molecular Biology and Evolution*, *22*(4), 1107–1118. https://doi.org/10.1093/molbev/msi097

Yusufmia, S. B.A.S., Collins, N. E., Nkuna, R., Troskie, M., Van Den Bossche, P., Penzhorn, B. L. (2010). Occurrence of *Theileria parva* and other haemoprotozoa in cattle at the edge of Hluhluwe-iMfolozi Park, KwaZulu-Natal, *South Africa Journal of the South African Veterinary Association,* 81, (1)a95 | DOI: https://doi.org/10.4102/jsava.v81i1.95.

Young, A. S., & Purnell, R. E. (1973). Transmission of *Theileria lawrencei* (Serengeti) by the ixodid tick, *Rhipicephalus appendiculatus*. *Tropical Animal Health and Production*, *5*(3), 146–152. https://doi.org/10.1007/BF02251383

APPENDICES

Appendix I. Sequences data set generated in this study

>MZ798158.1 *Theileria parva* isolate Boleni 104 kDa microneme-rhoptry antigen (p104) gene, partial cds

```
agg cca agg tct cct tca gaa tac gaa gat aca tca cca gga gac tac cca tct ctt cca
 R   P   R   S   P   S   E   Y   E   D   T   S   P   G   D   Y   P   S   L   P
atg aaa cgt cat cgc ttg gaa aga ctc aga cta act act act gaa atg gaa act gat cct
 M   K   R   H   R   L   E   R   L   R   L   T   T   T   E   M   E   T   D   P
ggt aga atg gcc aaa gat gca agt gga aag cct gtt aaa ctt aaa cgt tcc aag agt ttt
 G   R   M   A   K   D   A   S   G   K   P   V   K   L   K   R   S   K   S   F
gat gat ctt aca act gtt gaa ctt gcg cct gag cca aaa gct agt agg att gtt gtg gac
 D   D   L   T   T   V   E   L   A   P   E   P   K   A   S   R   I   V   V   D
gat gaa ggc act gag gca gat gac gag gaa aca cac cca cca gaa gaa aga caa aaa aca
 D   E   G   T   E   A   D   D   E   E   T   H   P   P   E   E   R   Q   K   T
gaa gtc aga cgc aga cgt cca cca aag aaa cca tcc aaa tca ccg agg cca tcg aag cca
 E   V   R   R   R   R   P   P   K   K   P   S   K   S   P   R   P   S   K   P
aag aaa cca aag aag cca gat tct gca tat att cct tca att ctc gcc
 K   K   P   K   K   P   D   S   A   Y   I   P   S   I   L   A
```

> MZ798157.1 *Theileria parva* isolate Marula 40 104 kDa microneme-rhoptry antigen (p104) gene, partial cds

```
gga aag cct gtt aaa ctt aaa cgt tcc aag agt ttt gat gat ctt aca act gtt gaa ctt
 G   K   P   V   K   L   K   R   S   K   S   F   D   D   L   T   T   V   E   L
gcg cct gag cca aaa ggt agg aag att gtt gtg gac aat gaa ggc act gag gca gat gac
 A   P   E   P   K   G   R   K   I   V   V   D   N   E   G   T   E   A   D   D
gag gaa aca cac cca cca gaa gaa aga caa aaa aca gaa gtc aga cgc aga cgt cca cca
 E   E   T   H   P   P   E   E   R   Q   K   T   E   V   R   R   R   R   P   P
aag aaa cca tcc aaa tca ccg agg cca tcg aag cca aag aaa cca aag aag cca gat tct
 K   K   P   S   K   S   P   R   P   S   K   P   K   K   P   K   K   P   D   S
gca tat att cct tca att gtc gcc
 A   Y   I   P   S   I   V   A
```

>MZ798156.1*Theileria parva* isolate Marula 11 104 kDa microneme-rhoptry antigen (p104) gene, partial cds

```
agg cca agg tct cct tca gaa tac gaa gat aca tca tca gga gac tac cca gct ctg cca
 R   P   R   S   P   S   E   Y   E   D   T   S   S   G   D   Y   P   A   L   P
atg aaa cgt cat cgc ttg gaa aga ctc aga cta act act act gaa atg gaa act gat cct
```

M K R H R L E R L R L T T T E M E T D P
ggt aga atg gcc aaa gat gca agt gga aag cct gtt aaa ctt aaa cgt tcc aag agt ttt
G R M A K D A S G K P V K L K R S K S F
gat gat ctt aca act gtt gaa ctt gag cct gag cca aaa ggt agg aag att gtt gtg gac
D D L T T V E L E P E P K G R K I V V D
gat gaa ggc act gag gca gat gac gag gaa aca cac cca cca gaa gaa aga caa aaa aca
D E G T E A D D E E T H P P E E R Q K T
gaa gtc aga cgc aga cgt cca cca aag aaa cca tcc aaa tca ccg agg cca tcg aag cca
E V R R R R P P K K P S K S P R P S K P
aag aaa cca aag aag cca gat tct gca tat gtt cct tca att gtc gcc
K K P K K P D S A Y V P S I V A

>MZ798155.1 *Theileria parva* isolate Marula 44 104 kDa microneme-rhoptry antigen (p104) gene, partial cds

agg cca agg tct cct tca gaa tac gaa gat aca tca tca gga gac tac cca tct ctg cca
R P R S P S E Y E D T S S G D Y P S L P
atg aaa cgt cat cgc ttg gaa aga ctc aga cta act act act gaa atg gaa act gat cct
M K R H R L E R L R L T T T E M E T D P
ggt aga atg gcc aaa gat gca agt gga aag cct gtt aaa ctt aaa cgt tcc aag agt ttt
G R M A K D A S G K P V K L K R S K S F
gat gat ctt aca act gtt gaa ctt gag cct gag cca aaa ggt agg aag att gtt gtg gac
D D L T T V E L E P E P K G R K I V V D
gat gaa ggc act gag gca gat gac gag gaa aca cac cca cca gaa gaa aga caa aaa aca
D E G T E A D D E E T H P P E E R Q K T
gaa gtc aga cgc aga cgt cca cca aag aaa cca tcc aaa tca ccg agg cca tcg aag cca
E V R R R R P P K K P S K S P R P S K P
aag aaa cca aag aag cca gat tct gca tat att cct tca att gtc gcc
K K P K K P D S A Y I P S I V A

>MZ798154.1 *Theileria parva* isolate Marula 7 104 kDa microneme-rhoptry antigen (p104) gene, partial cds

agg cca agg tct cct tca gaa tac gaa gat aca tca tca gga gac tac cca gct ctg cca
R P R S P S E Y E D T S S G D Y P A L P
atg aaa cgt cat cgc ttg gaa aga ctc aga cta act act act gaa atg gaa act gat cct
M K R H R L E R L R L T T T E M E T D P
ggt aga atg gcc aaa gat gca agt gga aag cct gtt aaa ctt aaa cgt tcc aag agt ttt
G R M A K D A S G K P V K L K R S K S F
gat gat ctt aca act gtt gaa ctt gag cct gag cca aaa gct agt aag att gtt gtg gac
D D L T T V E L E P E P K A S K I V V D
gat gaa ggc act gag gca gat gac gag gaa aca cac cca cca gaa gaa aga caa aaa aca

D E G T E A D D E E T H P P E E R Q K T
gaa gtc aga cgc aga cgt cca cca aag aaa ccg tcc aaa tca ccg agg cca tcg aag cca
E V R R R R P P K K P S K S P R P S K P
aag aaa ccg aag aag cca gat tct gca tat att cct tca att gtc gcc
K K P K K P D S A Y I P S I V A

>MZ798153.1 *Theileria parva* isolate Marula 75 104 kDa microneme-rhoptry antigen (p104) gene, partial cds

agg cca agg tct cct tca gaa tac gaa gat aca tca tca gga gac tac cca gct ctg cca
R P R S P S E Y E D T S S G D Y P A L P
atg aaa cgt cat cgc ttg gaa aga ctc aga cta act act act gaa atg gaa act gat cct
M K R H R L E R L R L T T T E M E T D P
ggt aga atg gcc aaa gat gca agt gga aag cct gtt aaa ctt aaa cgt tcc aag agt ttt
G R M A K D A S G K P V K L K R S K S F
gat gat ctt aca act gtt gaa ctt gag cct gag cca aaa gct agt aag att gtt gtg gac
D D L T T V E L E P E P K A S K I V V D
gat gaa ggc act gag gca gat gac gag gaa aca cac cca cca gaa gaa aga caa aaa aca
D E G T E A D D E E T H P P E E R Q K T
gaa gtc aga cgc aga cgt cca cca aag aaa ccg tcg aaa tca ccg agg cca tcg aag cca
E V R R R R P P K K P S K S P R P S K P
aag aaa ccg aag aag cca gat tct gca tat att cct tca att ctc gcc
K K P K K P D S A Y I P S I L A

>MZ798152.1 *Theileria parva* isolate Marula 12 104 kDa microneme-rhoptry antigen (p104) gene, partial cds

agg cca agg tct cct tca gaa tac gaa gat aca tca tca cga gac tac cca gct ctg cca
R P R S P S E Y E D T S S R D Y P A L P
atg aaa cgt cat cgc ttg gaa aga ctc aga cta act act act gaa atg gaa act gat cct
M K R H R L E R L R L T T T E M E T D P
ggt aga atg gcc aaa gat gca agt gga aag cct gtt aaa ctt aaa cgt tcc aag agt ttt
G R M A K D A S G K P V K L K R S K S F
gat gat ctt aca act gtt gaa ctt gcg cct gag cca aaa gct agt aag att gtt gtg gac
D D L T T V E L A P E P K A S K I V V D
gat gaa ggc act gag gca gat gac gag gaa aca cac cca cca gaa gaa aga caa aaa aca
D E G T E A D D E E T H P P E E R Q K T
gaa gtc aga cgc aga cgt cca cca aag aaa cca tcc aaa tca ccg agg cca tcg aag cca
E V R R R R P P K K P S K S P R P S K P
aag aaa cca aag aag cca gat tct gca tat att cct tca att gtc gcc
K K P K K P D S A Y I P S I V A

>MZ798151.1 *Theileria parva* isolate Manyara 104 kDa microneme-rhoptry antigen (p104) gene, partial cds

```
agg tca agg tct cct tca gaa tac gaa gat aca tca tca gga gac tac cca act ctg cca
 R   S   R   S   P   S   E   Y   E   D   T   S   S   G   D   Y   P   T   L   P
atg aaa cgt cat cgc ttg gaa aga ctc aga cta act act act gaa atg gaa acc gat cct
 M   K   R   H   R   L   E   R   L   R   L   T   T   T   E   M   E   T   D   P
ggt aga atg gcc aaa gat gca agt gga aag cct gtt aaa ctt aaa cgt tcc aag agt ttt
 G   R   M   A   K   D   A   S   G   K   P   V   K   L   K   R   S   K   S   F
gat gat ctt aca act gtt gaa ctt gag cct gag cca aaa ggt agg aag att gtt gtg gac
 D   D   L   T   T   V   E   L   E   P   E   P   K   G   R   K   I   V   V   D
gat gaa ggc act gag gca gat gac gag gaa aca cac cca cca gaa gaa aga caa aaa aca
 D   E   G   T   E   A   D   D   E   E   T   H   P   P   E   E   R   Q   K   T
gaa gtc aga cgc aga cgt cca cca aag aaa cca tcc aaa tca ccg agg cca tcg aag cca
 E   V   R   R   R   R   P   P   K   K   P   S   K   S   P   R   P   S   K   P
aag aaa cca aag aag cca gat tct gca tat att cct tca att gtc gcc
 K   K   P   K   K   P   D   S   A   Y   I   P   S   I   V   A
```

>MZ798150.1 *Theileria parva* isolate Satinsyi 104 kDa microneme-rhoptry antigen (p104) gene, partial cds

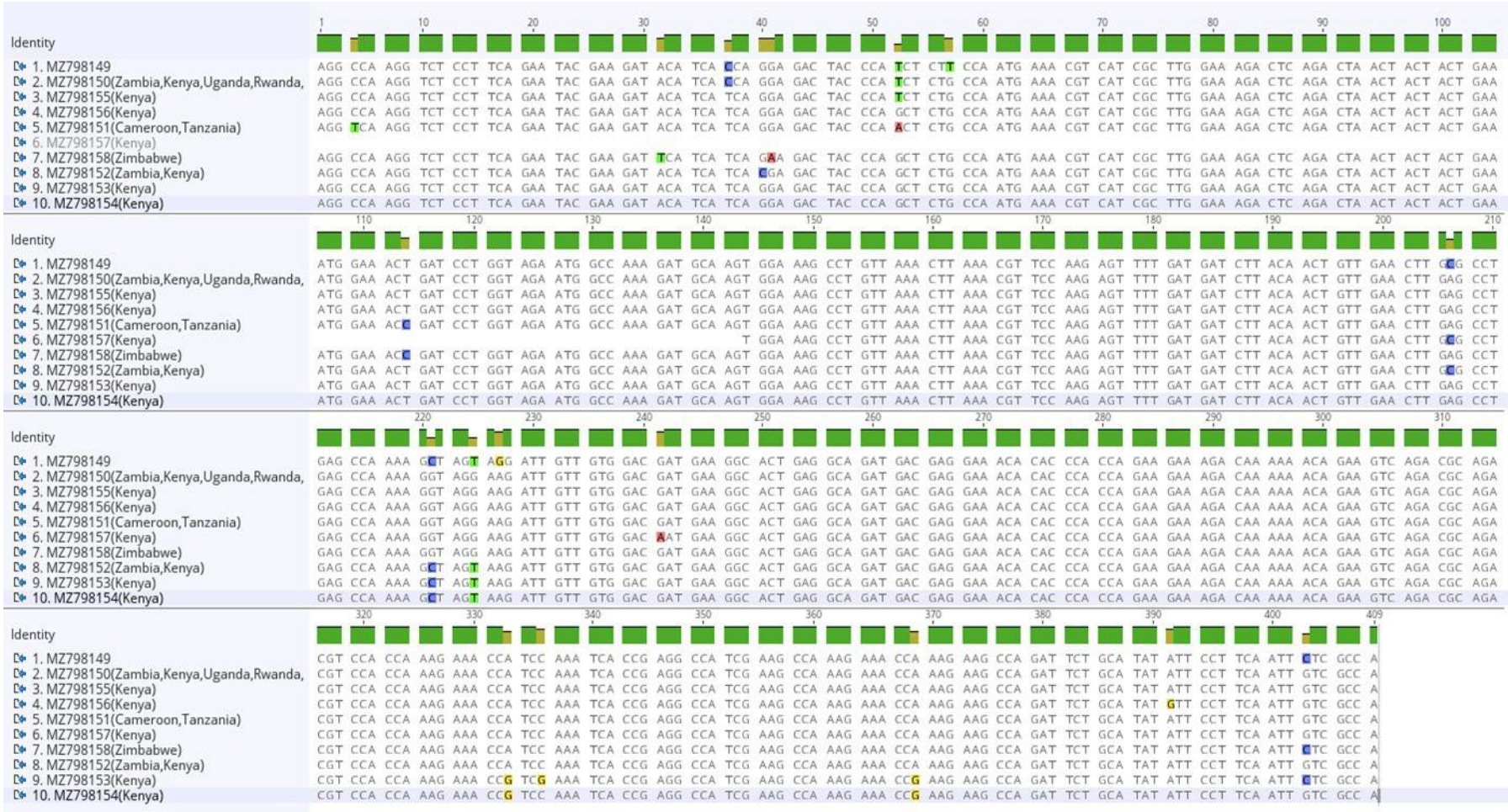
```
agg cca agg tct cct tca gaa tac gaa gat aca tca cca gga gac tac cca tct ctg cca
 R   P   R   S   P   S   E   Y   E   D   T   S   P   G   D   Y   P   S   L   P
atg aaa cgt cat cgc ttg gaa aga ctc aga cta act act act gaa atg gaa act gat cct
 M   K   R   H   R   L   E   R   L   R   L   T   T   T   E   M   E   T   D   P
ggt aga atg gcc aaa gat gca agt gga aag cct gtt aaa ctt aaa cgt tcc aag agt ttt
 G   R   M   A   K   D   A   S   G   K   P   V   K   L   K   R   S   K   S   F
gat gat ctt aca act gtt gaa ctt gag cct gag cca aaa ggt agg aag att gtt gtg gac
 D   D   L   T   T   V   E   L   E   P   E   P   K   G   R   K   I   V   V   D
gat gaa ggc act gag gca gat gac gag gaa aca cac cca cca gaa gaa aga caa aaa aca
 D   E   G   T   E   A   D   D   E   E   T   H   P   P   E   E   R   Q   K   T
gaa gtc aga cgc aga cgt cca cca aag aaa cca tcc aaa tca ccg agg cca tcg aag cca
 E   V   R   R   R   R   P   P   K   K   P   S   K   S   P   R   P   S   K   P
aag aaa cca aag aag cca gat tct gca tat att cct tca att gtc gcc
 K   K   P   K   K   P   D   S   A   Y   I   P   S   I   V   A
```

>MZ798149.1 *Theileria parva* isolate Onderstepoort 104 kDa microneme-rhoptry antigen (p104) gene, partial cds

```
agg cca agg tct cct tca gaa tac gaa gat aca tca cca gga gac tac cca tct ctt cca
 R   P   R   S   P   S   E   Y   E   D   T   S   P   G   D   Y   P   S   L   P
atg aaa cgt cat cgc ttg gaa aga ctc aga cta act act act gaa atg gaa act gat cct
 M   K   R   H   R   L   E   R   L   R   L   T   T   T   E   M   E   T   D   P
```

```
ggt aga atg gcc aaa gat gca agt gga aag cct gtt aaa ctt aaa cgt tcc aag agt ttt
 G   R   M   A   K   D   A   S   G   K   P   V   K   L   K   R   S   K   S   F
gat gat ctt aca act gtt gaa ctt gcg cct gag cca aaa gct agt agg att gtt gtg gac
 D   D   L   T   T   V   E   L   A   P   E   P   K   A   S   R   I   V   V   D
gat gaa ggc act gag gca gat gac gag gaa aca cac cca cca gaa gaa aga caa aaa aca
 D   E   G   T   E   A   D   D   E   E   T   H   P   P   E   E   R   Q   K   T
gaa gtc aga cgc aga cgt cca cca aag aaa cca tcc aaa tca ccg agg cca tcg aag cca
 E   V   R   R   R   R   P   P   K   K   P   S   K   S   P   R   P   S   K   P
aag aaa cca aag aag cca gat tct gca tat att cct tca att ctc gcc
 K   K   P   K   K   P   D   S   A   Y   I   P   S   I   L   A
```

Appendix II: Multiple sequence alignment for the p104 sequences generated using MAFFT in this study.



Multiple sequence alignment for the p104 sequences generated using MAFFT in Geneious Prime 2022.

Appendix III: Best fit nucleotide substitution model selection results

Candidate models = 88
number of substitution schemes = 11
including models with equal/unequal base frequencies (+F)
including models with/without a proportion of invariable sites (+I)
including models with/without rate variation among sites (+G) (nCat = 4)
Optimized free parameters (K) = substitution parameters + 65 branch lengths + topology
Base tree for likelihood calculations = ML tree
Tree topology search operation = NNI

Best fit model
Model selected:
Model = K80+I
partition = 010010
-lnL = 595.5946
K = 68
kappa = 2.5027 (ti/tv = 1.2513)
p-inv = 0.8730

PAUP* Commands Block:
Lset base=equal nst=2 tratio=1.2513 rates=equal pinvar=0.8730;

Appendix IV: Results of the analysis of positive selection based on likelihood ratio tests

Site model (SM)									
Model	**np**	**Ln L**	**Estimates of parameters**				**Model compared**	**LRT P-value**	**Positive sites**
M3	63	-4907.943623	p:	0.90755	0.09245	0.00000			
			ω:	0.00075	0.15511	40.69742			
M0	59	-5214.294242	ω0:	0.01109			M0 vs. M3	0.000000000	Not Allowed
M2a	62	-5058.261484	p:	0.91597	0.08403	0.00000			
			ω:	0.00206	1.00000	29.20183			
M1a	60	-5058.261463	p:	0.91597	0.08403				
			ω:	0.00206	1.00000		M1a vs. M2a	0.999979000	Not Allowed
M8	62	-4924.382958	p0=0.99999	p=0.11213	q=1.19849				
			(p1= 0.00001)	ω= 1.00000					
M7	60	-4924.381768	p=	0.11213	q=	1.19849	M7 vs.M8	0.998810708	Not Allowed
M8a	61	-4898.182645	p0=0.99999	p=0.04240	q=1.57375				
			(p1= 0.00001)	ω= 1.00000			M8a vs.M8	0.000000000	Not Allowed

Appendix V: Similarity Report

University of Eldoret

Certificate of Plagiarism Check for Thesis

Author Name	PERIS NYABOKE MAKORI (SC/PGB/012/15)
Course of Study	Type here...
Name of Guide	Type here...
Department	Type here...
Acceptable Maximum Limit	Type here...
Submitted By	titustoo@uoeld.ac.ke
Paper Title	GENETIC CONSERVATION OF THE p104 GENE USED FOR PCR-BASED DIAGNOSIS AND SURVEILLANCE OF THE Theileria parva PARASITE.
Similarity	12%
Paper ID	1086283
Submission Date	2023-11-07 08:32:03

Signature of Student

Signature of Guide

University of Eldoret
UNIVERSITY LIBRARIAN
P.O. Box 1125, ELDORET-30100

Head of the Department

University Librarian

Director of Post Graduate Studies

* This report has been generated by DrillBit Anti-Plagiarism Software