Genotyping a novel *Theileria parva* candidate vaccine antigen in cattle- and buffalo-derived parasites

---

Lauren-Leigh Borchers

Supervisor: Prof KP Sibeko-Matjila

Co-supervisor: Dr MA Tjale

2020

Submitted to the Department of Veterinary Tropical Disease, Faculty of Veterinary Science, University of Pretoria, Onderstepoort, Pretoria, South Africa, in partial fulfilment of the requirements for the degree Magister Scientiae Veterinary Science Tropical Disease

Declaration

I, Lauren-Leigh Borchers, declare that the work presented in this dissertation, in partial fulfilment of the requirements for the Magister Scientiae degree at the University of Pretoria is my own. This research has not been previously submitted for examination at any other institution, tertiary or otherwise.

Lauren-Leigh Borchers

5 September 2020

Date signed

Acknowledgments

I would like to give special thanks to my two amazing supervisors, Prof Kgomotso Sibeko-Matjila and Dr Mabotse Tjale. Your knowledge, kindness, patience and understanding have truly been amazing through this degree and I am extremely grateful to them both for all their guidance.

I would like to thank Dr Donald L. Mukolwe and Nosiphiwo Komani for all their help in the laboratory and their help with all the troubleshooting that was done in this study.

Further acknowledgement is extended to the National Research Fund of South Africa, the UP Postgraduate Bursary Programme and the South African Agricultural Sector Education and Training Authority (AgriSETA), awarded to the Department of Veterinary Tropical Diseases, Faculty of Veterinary Science, University of Pretoria for the funding of this study.

TABLE OF CONTENTS

Declaration i

Acknowledgments ii

LIST OF TABLES v

TABLE OF FIGURES vi

SUMMARY 1

CHAPTER 1 2

1. General introduction 2

CHAPTER 2 3

2. Literature Review 3

2.1 *Introduction to* Theileria parva 3

2.2 *Life cycle of* Theileria parva 4

2.3 *Diagnosis of* Theileria parva 6

2.4 *Control and treatment* 7

2.4.1 Physical barriers 7

2.4.2 Tick control 7

2.4.3 Chemotherapy 8

2.4.4 Immunisation 9

2.4.4.1 Immunization with live parasite stocks 9

2.4.4.2 Subunit vaccines 10

2.5 *Immunity to theileriosis* 12

2.6 *Reverse vaccinology and identification of novel vaccine candidates* 13

2.7 *Antigenic diversity in* Theileria parva 15

2.8 *Study justification* 16

CHAPTER 3 18

3. METHODS AND MATERIALS 18

3.1 *Selection of vaccine candidates* 18

3.1.1 Gene expression profiling of the selected candidates 19

3.2 *Genotyping of the selected candidate gene* 19

3.2.1 Primer design and PCR optimisation 20

3.2.2 Sample selection 20

3.2.3 PCR amplification 21

3.2.4 PCR product purification 22

3.2.5 DNA Cloning 22

3.2.5.1 Ligation of TP04_0028 amplicons into pJET 1.2/blunt cloning vector 22
3.2.5.2 Transformation of E. coli competent cells 22
3.2.5.3 Selection and confirmation of recombinants 23
3.2.6 Sequencing and sequence data analysis 23
3.2.7 Sequence alignment and phylogenetic analysis 23
3.2.8 Analysis of sequence variation in the epitope regions 24
CHAPTER 4 26
4. RESULTS 26
4.1 *Selection of a vaccine candidate for genotyping* 26
4.2 *Selected gene amplification* 26
4.2.1 Selection of Primers targeting the partial TP04_0028 gene 26
4.2.2 PCR optimization and amplification of TP04_0028 gene in cattle and buffalo blood samples 29
4.3 *Sequence data analysis* 30
4.3.1 Nucleotide sequence data analysis 30
4.3.2 Amino acid sequence data analysis 37
4.4 *Phylogenetic analysis* 42
CHAPTER 5 46
5. DISCUSSION AND CONCLUSION 46
6. LIMITATIONS AND RECOMMENDATIONS FOR FUTURE STUDIES 51
7. STUDY OUTCOMES 52
8. APPENDIX 53
9. REFERENCES 54

LIST OF TABLES

Table 3.1: The list of *in silico* tools used to identify the vaccine candidates. 18

Table 3.2: The number of DNA samples used in the study, their host from which blood was obtained, geographic origin and the respective p67 allele profiles of each parasite population. 20

Table 3.3: The list of alleles of 20 predicted epitopes from IEDB forming part of TP04_0028 amplicon. 25

Table 4.1: Information on the potential vaccine candidates identified by *in silico* analysis, including expression profile in *T. parva* 7014 and Muguga isolates and predicted protein properties. 27

Table 4.2: Pairwise distance alignment of nucleotide sequences showing the percentage sequence differences between each sequence and the reference gene sequence, TP04_0028 (XM_758570.1), with the average percentage difference of 2%. 33

Table 4.3: Pairwise distance alignment of the nucleotide sequences showing the base pair differences between each sequence and the reference, TP04_0028 (XM_758570.1), with the average number of base pair differences of 13 bp. 34

Table 4.4: Pairwise distance alignment done in MEGA of the amino acid sequences representing the percentage sequence differences between each sequence and the reference (XM_758570.1). The average percentage difference was determined to be 3%. 39

Table 4.5: Pairwise distance alignment done in MEGA of the amino acid sequences representing the number of position differences between each sequence and the reference (XM_758570.1). The average number of position differences is 7. 40

TABLE OF FIGURES

Figure 2.1: The life cycle of *Theileria parva* as depicted in Nene et al., (2016). 5

Figure 3.1: Graphical representation showing the selection criteria of 13 potential vaccine candidates *in silico* analysis 19

Figure 3.2: The distribution of *T. parva* in Africa. 21

Figure 4.1: Primer pair options obtained from Primer Blast (NCBI); Primer set 6 (in a red box) was selected for the amplification of TP08_0028. 28

Figure 4.2: Graphical representation of the entire TP04_0028 translated protein sequence, showing the region of 20 predicted epitopes included in the PCR target region within the two black arrows. The coloured boxes represent the predicted epitope region. 29

Figure 4.3: The agarose gel electrophoresis showing PCR products produced at different annealing temperatures during the optimisation of amplification of the 754 bp target region of TP04_0028 gene. 30

Figure 4.4: Multiple nucleotide sequence alignment of TP04_0028 gene generated using BioEdit, consisting of 34 sequences from cattle- and buffalo-derived T. parva isolates. 37

Figure 4.5: Amino acid multiple sequence alignment of 34 cattle and buffalo sequences generated using ClustalX, in BioEdit. 41

Figure 4.6: The position of seven overlapping predicted epitopes, spanning position 129 to 145 in the protein sequence alignment shown in Figure 4.5. 42

Figure 4.7: Neighbor-Joining phylogenetic tree based on the TP04_0028 **nucleotide (A) and amino acid (B)** sequences, showing two major monophyletic groups. 44

Figure 4.8: Maximum-likelihood phylogenetic tree based on the TP04_0028 **nucleotide (A) and amino acid (B)** sequences, showing two major monophyletic groups 45

SUMMARY

Title: Genotyping a novel *Theileria parva* candidate vaccine antigen in cattle- and buffalo-derived parasites

Degree: Magister Scientiae Veterinary Science Tropical Disease

Student: Lauren-Leigh Borchers

Supervisor: Prof KP Sibeko-Matjila

Co-supervisor: Dr MA Tjale

Department: Veterinary Tropical Diseases

Theileriosis is a lymphoproliferative tick-borne disease of cattle and other wild ruminants, caused by infection with a protozoan, *Theileria parva.* The disease is prevalent in cattle throughout Central, East and southern Africa, where it threatens 50% of the cattle population. There are various control and treatments methods used against theileriosis; however, they all have limitations. The available live immunisation method, the Muguga cocktail, does not confer protection against all field strains, particularly buffalo-derived *T. parva*. Attempts to develop a subunit vaccine have been promising but these have shown limited efficacy due to antigenic and genetic diversity of *T. parva* strains in the field. Thus, there is a need to search for additional vaccine candidates. A related study has identified potential vaccine candidates using a genome-wide *in silico* approach. Consequently, the aim of this study was to genotype one of the identified antigens. TP04_0028 was selected for genotyping among candidate genes with high expression levels in the schizont stage of both cattle- and buffalo-derived *T. parva* isolates. Specific primers were designed and optimised for PCR amplification and sequencing. The comprehensive analysis of sequences from 17 cattle- and 17 buffalo-derived *T. parva*, from East and southern Africa, showed conservation in 12 (60%) of the 20 TP04_0028 predicted epitopes, in both parasite types, irrespective of geographical origin. Eighteen of the 20 predicted epitopes are conserved amongst different BoLA alleles and an area of 7 overlapping epitopes could be the starting point for initial experimental evaluation of the immunogenic properties of TP04_0028. Once the immunogenicity of these epitopes have been tested and the extent to provide protection from cattle- and buffalo-derived infections have been verified, they may be considered for vaccine development.

CHAPTER 1

1. GENERAL INTRODUCTION

Theileria parva causes fatal disease syndromes known as East Coast fever (ECF), January disease and Corridor disease. Cattle theileriosis caused by *T. parva* infections affects many animals in Africa and has a huge economic impact, especially to communities that depend on livestock for their livelihood. Prevention and treatment options for *T. parva* infections include geographical boundaries between livestock and wildlife, use of acaricides for tick control, chemotherapy drugs and live vaccines (Irvin et al., 1989, Peregrine, 1994, Uilenberg, 1999, Di Giulio et al., 2009, Michel and Bengis, 2012, Morrison, 2015). These options have been shown to be effective but each has its own drawbacks ranging from cost of treatment, traces of chemical residue on animal products, resistance to treatment and the risk of development of carrier status. One major drawback, particularly associated with immunisation using live parasite stocks in the infection and treatment method (ITM), is the parasite genetic and antigenic diversity in field strains (Ferraro et al., 2011). The current immunization method only protect against infections with the cattle-derived *T. parva* strains, the causative agents of ECF and not buffalo-derived *T. parva* parasites, which are responsible for Corridor disease (MacHugh et al., 2009). Subunit vaccines are currently being explored as alternatives to alleviate disadvantages suffered by the ITM (reviewed by Nene et al., 2016). Thus, our lab recently identified possible vaccine candidates using *in silico* analysis and the current study investigated the sequence diversity of one of the candidates in cattle- and buffalo-derived *T. parva* isolates, from East and southern Africa.

Study aim

To determine the sequence diversity of a novel vaccine candidate identified by reverse vaccinology in cattle- and buffalo-derived *T. parva* parasites.

Objectives

1. Confirmation of predicted vaccine candidates through *in silico* analysis and selection of a novel protein for genotyping.
2. Screening of cattle and buffalo blood samples from eastern and southern Africa for *T. parva* infection using real-time PCR (qPCR).
3. Design of oligonucleotide primers and amplification of target gene by PCR.
4. Cloning, sequencing and sequence analysis of the candidate gene.
5. Assessment of variations in the predicted epitope regions between sequences obtained from cattle- and buffalo-derived *T. parva* isolates and from isolates from different geographic regions.

CHAPTER 2

2. LITERATURE REVIEW

2.1 *Introduction to* Theileria parva

Theileria parva is an intracellular protozoan parasite that belongs to the family *Theileiriidae* and the order *Piroplasmida* (Rocchi et al., 2006) and causes a lethal lymphoproliferative disease in cattle. Theileriosis caused by *T. parva* infections is of a major economic and social weight in most parts of Africa, with cattle losses of more than one million annually and over US $300 million in economic losses per year, partly due to ineffective control measures (reviewed by Nene et al., 2016). These economic figures are based on studies performed many years ago and have likely increased due to inflation.

Theileria parva infections in cattle cause ECF and Corridor disease in East, Central and southern Africa, and January disease in Zimbabwe (reviewed by Mans et al., 2015). The cattle-derived *T. parva* strains are responsible for ECF and January disease while the buffalo-derived *T. parva* causes Corridor disease (Nene et al., 1996). *Rhipicephalus appendiculatus* and *R. zambeziensis* are the natural tick vectors of *T. parva* and transmit buffalo-derived *T. parva* from infected buffalo to cattle or cattle-derived *T. parva* from infected cattle to cattle (Sibeko et al., 2011). Buffalo are asymptotic carriers of *T. parva,* thus believed to be the original host of this parasite.

In southern Africa, ECF was eradicated in southern Mozambique, Zimbabwe, South Africa and eSwatini, between 1917 and the 1960s (Lawrence, 1992). However, other forms of theileriosis were discovered in Zimbabwe and South Africa, respectively January disease and Corridor disease, which persist to this day in these countries. It also emerged that the African buffalo (*Sycerus caffer)* was the reservoir host to Corridor disease (Chaisi et al., 2011), which is more acute than ECF (reviewed by Nene et al., 2016). Corridor disease is considered self-limiting as the cattle die before the parasite reaches the piroplasm stage of infection which is transmittable and infective to ticks (Yusufmia et al., 2010), however, studies have reported cattle that have recovered and survived from Corridor disease (Mbizeni et al., 2013).

Clinical signs of *T. parva* infections include pyrexia, swollen lymph nodes, laboured breathing, and nasal discharge, corneal opacity and lacrimation which all indicate severe infection (Sitt et al., 2015). In contrast to ECF, Corridor disease and January disease are characterised by low levels of infected leukocytes in the peripheral lymph nodes and fewer schizonts (Lawrence, 1979, Sitt et al., 2015). January disease differs from ECF and Corridor disease in that the disease persists during the rainy

season in Zimbabwe (Matson, 1967). Adult *R. appendiculatus* ticks, the tick life stage that transmits *T. parva* is most active during the wet rainy season of December – March in Zimbabwe (Latif et al., 2001), hence the seasonal occurrence of January disease.

2.2 *Life cycle of* Theileria parva

Theileria is most closely related to the genus *Babesia* hence *Theileria* species follow a similar life cycle (Kuo et al., 2008, Lack et al., 2012) (Figure 2.1). *Rhipicephalus appendiculatus,* the vector responsible for the biological transmission of *T. parva*, is referred to as a 'three-host' tick as three life stages of the tick are vectors for *Theileria*, i.e. the larvae, nymph and adult (Gonder, 1910). *Theileria* enters the cattle bloodstream through sporozoite secretion at tick feeding sites (reviewed by Mans et al., 2015). In the host the sporozoites infect leukocytes and differentiate into the multinucleated schizont stage which initiates a cancer-like phenotype in the infected cells (Nyagwange et al., 2017). While a proportion of schizonts persist in the leukocytes and perpetuate the infection by synchronized replication with the infected cells, some multiply by a process called merogony to form merozoites. These merozoites further invade red blood cells where they develop to piroplasms, which is the tick infective stage (Shaw and Tilney, 1995). The piroplasms are ingested by tick larvae and nymphs during a blood meal (Rocchi et al., 2006). Development occurs in the infected tick gut resulting in the production of gametes which fuse to form a zygote. The zygote migrates to the epithelium of the tick gut where it develops to motile kinetes that are released into the heamocoel and move to the salivary glands. During sporogony, sporozoites are produced in the acini of the tick salivary glands and are released 4-8 days post attachment to the host (Shaw and Young, 1995).

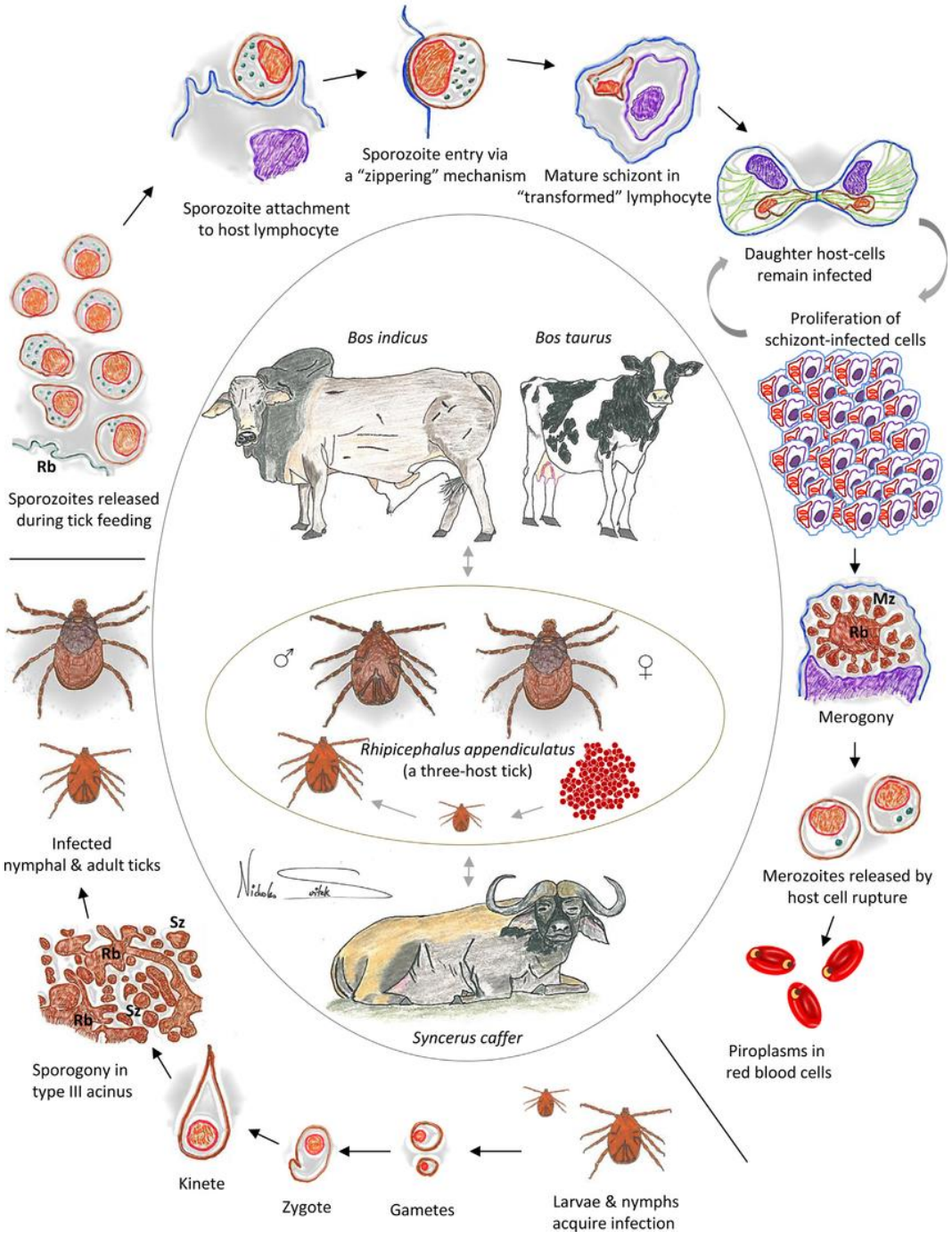

Figure 2.1: *The life cycle of* Theileria parva *as depicted in Nene et al., (2016).*

2.3 *Diagnosis of* Theileria parva

Cattle and buffalo can be infected by multiple species of *Theileria*, some of which may not be of economic importance but may interfere with the diagnosis of pathogenic species (Eygelaar et al., 2015). Hence, specific and sensitive diagnostic methods are crucial for detection of pathogenic species of *Theileria* such as *T. parva. Theileria parva* can be diagnosed using microscopic examination of blood and lymph node smears to respectively determine the presence of piroplasms in erythrocytes and schizonts in leukocytes. Limitations to using microscopic examination include the inability to detect carrier animals and differentiate various *Theileria* species (reviewed by Mans et al., 2015). Xenodiagnostics, serology assays and molecular diagnostic methods are also used to determine *T. parva* infection (Eygelaar et al., 2015). The serology-based indirect fluorescent antibody test (IFAT) utilises whole body antigens and is more sensitive compared to microscopic examination. Schizont antigens are prepared from an *in vitro* cell culture suspension of lymphoid cells that have been previously infected with *T. parva* macroshizonts and are prepared on a slide with the tissue sample of interest (Burridge and Kimber, 1972). The major drawback of this test is lack of specificity; in that many different *Theileria* species may cross-react with the antigen that should detect *T. parva* (Minjauw et al., 1998). Apathogenic species such as *T. tarautragi, T. mutans* and *T. velifera* are commonly found in cattle and buffalo blood samples (Sibeko et al., 2008), and can interfere with the diagnosis of *T. parva* when using microscopy and serological diagnostic tools.

In xenodiagnostics, the tick vector is exposed to either pathogen-infected animal or tissue and is then examined for the ingested pathogen (reviewed by Mans et al., 2015). Since the development of less tedious and more rapid methods, this method is no longer commonly used.

Molecular diagnostic methods are a great improvement for the sensitive, specific and rapid detection of *T. parva.* These polymerase chain reactions (PCR)-based assays include conventional, nested, probe-based reverse line blot and real-time PCR (Sibeko et al., 2008, reviewed by Mans et al., 2015). In South Africa, quantitative real-time PCR (qPCR) has been shown to be the most accurate in the diagnosis of *T. parva* infections in cattle and buffalo blood samples (Sibeko et al., 2008, Papli et al., 2011). These assays amplify a portion of the 18S RNA gene of the parasite and specific detection is based on annealing of the *T. parva*-specific hybridisation or hydrolysis oligonucleotide probes to the compatible amplicon generated from *T. parva* DNA. Hence, the hybridization probe-based qPCR (Sibeko et al., 2008) was selected to screen *T. parva*-positive samples in the current study.

2.4 *Control and treatment*

The occurrence of clinical disease of *T. parva* is rare due to endemic stability; however, the level of stability required to sustain this state is not always attainable (Kimaro et al., 2017), which is why effective education as well as control and treatment methods need to be in place. Cattle theileriosis is controlled through tick control, chemotherapy treatment and immunization. Geographical borders restricting cattle and buffalo interaction are also implemented across South Africa to limit the spread of buffalo-derived *T. parva* infections (Michel and Bengis, 2012).

2.4.1 Physical barriers

The wildlife-livestock interface contributes to the spread of diseases affecting cattle, such as Corridor disease, bovine tuberculosis, bovine brucellosis, and foot-and-mouth disease, in communities around game parks in many African countries. In South Africa one of the biggest control measures against Corridor disease is physical separation of wildlife and livestock by erecting fences to limit interaction between buffalo and cattle. National legislation on buffalo farming dictates that *T. parva*-infected buffalo are restricted to registered game farms within the Corridor disease endemic boundary. All infected buffalo are to be kept within a game farm, within an electrified game-proof fence with a minimum voltage of 5500V at all times (Animal Diseases Act 1984, Act No. 35). In the Eastern Cape, there are buffalo herds which are considered disease-free, such as those from the Addo Elephant Game Farm. However, the increasing prevalence of ticks in this region threatens this status (Smith and Parker, 2010, Yusufmia et al., 2010); hence, strict control of animal movements is also implemented in these areas. Although physical boundaries limit the spread of disease, they are not sufficient in the control of *T. parva* due to tick movement and environmental changes such as the ever changing climatic and seasonal conditions.

2.4.2 Tick control

One form of control against *T. parva* infections is the constant application of acaricides to control tick infestation; however, this method is not sustainable, especially in low-income settings, as the exercise of continued application becomes costly over a lengthy period of time (Morrison, 2015). Consequently, small farms with smaller herds might choose to use the hand-spray method, as it is pragmatic and cost effective in comparison to a spray race or dipping tanks (Minjauw et al., 1998). Furthermore, excessive use of acaricides poses an environmental and food contamination risk and the prevention of *T. parva* infections is not guaranteed, especially at wildlife-livestock interfaces due to the constant interaction between cattle, buffalo and tick species (Walker et al., 2014, Elisa et al., 2015, Tayebwa et al., 2018). Moreover, acaricides are administered using dipping tanks or spray races, which are expensive to build as well as maintain and run continuously, hence they are usually not ideal in

informal settings (Minjauw et al., 1998, Di Giulio et al., 2009). Development of resistance to acaricides is also another major concern. Three-host ticks such as *R. appendiculatus,* the vector for *T. parva,* are reported to develop acaricide resist slower compared to one-host ticks, however, resistance still exists and is usually the result of acaricide administration at high frequencies and doses (Mekonnen et al., 2002). For example, in Uganda, *R. appendiculatus* ticks were reported to be the second most abundant tick species infected with *T. parva,* in central and east Uganda due to acaricide resistance and resultant acaricide malpractice (Vudriko et al., 2016, Tayebwa et al., 2018). Thus, the use of acaricides as tick control is not efficient for prevention of *T. parva* infections due to the stated limitations and the certain rise in acaricide resistance.

2.4.3 Chemotherapy

The chemotherapeutic drugs, buparvaquone and parvaquone (Irvin et al., 1989, Peregrine, 1994), have been widely used in treating *T. parva* infections. For decades, buparvaquone has been administered with no documented resistance. Similar to cancer, *Theileria* species have evolved to induce proliferation, dissemination of the host cell, resistance to apoptosis and immune invasion (Tretina et al., 2015). Hence the transformation of T-cell and B-cell lines by *T. parva* is reversible using chemotherapy drugs as observed in cancer chemotherapeutic treatment (De Goeyse et al., 2015) and does not cause mutations to the host genome after treatment (Haidar et al., 2018). However, in Sudan, drug resistance in the related *T. annulata* has been reported when using buparvaquone in treatment due to point mutations (Chatanga et al., 2019). Thus, although there are no reports of drug resistance in relation to *T. parva*, the evolutionary possibility may exist. A major drawback to chemotherapeutic treatment is that the drug needs to be administered early in the disease to be completely effective (reviewed by Nene et al., 2016). Cases that are too advanced will not respond to chemotherapeutic treatment (Muraguri et al., 1999). Another concern with chemotherapeutic treatment is the development of carrier status. Animals that recover from treatment with parvaquone have been shown to become carriers of *T. parva* as growth factors such as interleukin-2 (IL-2) continue to induce schizont proliferation following treatment and recovery from the clinical disease (Brown et al., 1989). Other communities have considered other remedies instead of commercial anti-theilerial drugs. It has been reported that in the Maasai pastoralist community in Tanzania, 54.6% of cattle owners use herbal medicines in conjunction with commercial drugs and 93.1% of these, use the specific plant species *'Osukuroi'* (Kimaro et al., 2017). The herbal medicines were noted to reduce the severity of ECF and could be explored as alternatives to the commercially available drugs.

2.4.4 Immunisation

Although the use of acaricides and chemotherapy are effective, the long-term costs and constant application require consideration of other disease control approaches, such as immunisation that would be long lasting and provide broad-spectrum immunity to most *T. parva* strains. According to Uilenberg (1999), immunisation is intended to protect cattle stock in areas where an unstable endemic situation persists or where the disease threatens naïve cattle, as seen in most parts of Africa where *T. parva* is prevalent. Immunisation against *T. parva* is an old phenomenon that began with the concept of acquired immunity. The initial investigations explored several attempts, such as immunization using blood from infected cattle to immunise naïve cattle and inoculation with infected spleen or lymph material, though these efforts did not show any positive results (Uilenberg, 1999). Subsequently, successful protection was realized through the infection and treatment method (ITM) described below, in 2.4.4.1.

2.4.4.1 Immunization with live parasite stocks

Live vaccines have been broadly used in both veterinary and human medicine as they can mobilise the humoral and cellular aspects of the immune response to produce long-term immunity. In the control of *T. parva* infections, the live parasite stocks are used in the ITM, which is widely employed in most African countries as a form of immunisation against cattle theileriosis (Uilenberg, 1999). Immunity against *T. parva* can be acquired from previous infection and recovered animals may become carriers of the disease (Uilenberg, 1999). However, the nature and resolve of the carrier state can vary between different stocks of *T. parva* (Skilton et al., 2002). This phenomenon forms the basis for immunization by ITM. The ITM involves vaccinating cattle with a dilution of sporozoites from particular *T. parva* stocks (e.g. the widely used Mugguga cocktail) followed by the treatment with a long-acting oxytetracycline (Di Giulio et al., 2009). The success of this approach relies on the *T. parva* stock used for infection and if it can induce immunity that can protect against various parasite isolates occuring in the field in a targeted vaccination area. In Zimbabwe, for example, it was shown that employing ITM using the specific *T. parva* stock Boleni resulted in efficient immunisation against January disease while other *T. parva* stocks (such as Muguga) did not have the same effect (Irvin et al., 1989). Hence, the Boleni stock was identified as the ideal parasite stock for use in immunisation programmes against *T. parva* in that country. Although ITM has been employed successfully, *T. parva* infections include multiple strains in the field, which can undermine disease control efforts.

In many cases, immunisation with a single parasite strain results in long-term immunity to the homologous strain but ineffective immunity to heterologous strains is seen (MacHugh et al., 2009). To provide broad protection against field strains, parasite stocks used in ITM have been expanded in

some regions. However, the expansion of the parasite stocks can prove challenging when the goal is to target multiple strains or subtypes of a pathogen (Ferraro et al., 2011). Nonetheless, the Muguga cocktail, a combination of three *T. parva* stocks that include Muguga, Serengeti-transformed and Kiambu-5, was produced in an effort to provide broad-spectrum immunity against ECF (Radley et al., 1975a, Radley et al., 1975b). Even though the cocktail vaccine provides better protection compared to the use of single parasite stocks, it still does not protect against infections from buffalo-derived parasites (Sitt et al., 2015). In addition to the requirement of numerous parasite isolates to broaden immunity, the high production cost and dependence on cryopreservation to maintain the parasite add to the disadvantages of using the cocktail vaccine (McKeever and Morrison, 1994, Perry, 2016). Moreover, vaccinated animals can become carriers that serve as source of future infection. Affordability is also another challenge in some communities. It has been reported that in the Maasai community in Tanzania, 63.8% of cattle owners did not vaccinate their cattle against *T. parva* infections due to the cost involved in purchasing the vaccine (Kimaro et al., 2017).

Considering all the limitations associated with the use of live vaccines, there is a need for an alternative vaccine, which will be safer, more cost-effective and protect against both cattle- and buffalo-derived *T. parva* infections. Parasite genes encoding proteins with cytotoxic lymphocytes (CTL) epitopes, with binding affinity for MHC-class I alleles, have been identified as likely vaccine candidates for the development of a subunit vaccine against *T. parva* infections. *Theileria parva*-specific $CD8^+$ MHC-class I-restricted CTLs are observed in the peripheral blood of immune cattle within 5-7 days post-infection and they are highly effective at killing infected cells *in vitro* (Nene et al., 2012).

2.4.4.2 Subunit vaccines

Instead of the entire pathogen, conventional subunit vaccines utilise only part of the target pathogen which provokes an immune response against that component of the pathogen only (Jorge and Dellagostin, 2017). Recombinant subunit vaccines are based on genes that encode protein antigens, which are recognised by the host's immune system and are utilised as objects for protective immunity (Nascimento and Leite, 2012, Draper et al., 2015). According to Blake et al. (2015), the efficacy and longevity of subunit vaccines in field populations can be predicted through the "knowledge of relevant pre-existing antigenic diversity, population structure, likelihood of co-infection by genetically distinct strains and the efficiency of cross-fertilisation" (Blake et al., 2015). In an effort to produce an efficient malaria vaccine, the subunit vaccine targeting parasite sporozoites, blood stage parasites and blocking the transmission of the resultant gametocytes to mosquitoes have been investigated (Ouattara et al., 2015); all of which are a possibility for the development of a subunit vaccine for *T. parva* due to the similar life cycle shared between *T. parva* and *Plasmodium falciparum.*

For prevention of *T. parva* infections, monoclonal antibodies (mAbs) with the ability to neutralise sporozoite infectivity have been shown to immunise against experimental challenge with *T. parva* (Morrison, 1996). This has been used as the basis to propose that certain sporozoite surface antigens comprise neutralising properties and could make suitable vaccine candidates (reviewed by Musoke et al., 1996). Using mAbs, it was shown that most antibodies bind to p67, a conserved surface antigen that is expressed by *T. parva* during the sporozoite stage of development (Nene et al., 1996), making it an ideal candidate for development of a vaccine for this parasite. A p67-based subunit vaccine has been tested for efficacy against *T. parva* infections using two p67 constructs (p67C and $p67_{635}$). Although the vaccine reduced the incidence of severe ECF in Kenya, it did not fully protect cattle against ECF or slow the progression of the infection (Musoke et al., 2005). It was later discovered that the gene encoding the p67 antigen is not as conserved as previously thought. Four allele types have been reported to date instead of two as previously reported (Sibeko et al., 2010). Cattle-derived *T. parva* parasites seem to strictly possess p67 allele 1 while buffalo-derived parasites possess all four alleles (1, 2, 3 and 4) in South Africa and three (alleles 1, 2 and 3) in Kenya (Sibeko et al., 2010, Sitt et al., 2019). These alleles are discriminated by the presence or absence of two indels, the 129 bp typical of alleles 1 and 2 and 174 bp characteristic of alleles 3 and 4, as defined by Sibeko et al. (2010). Thus, the extent of diversity (especially antigenic) may have implications in the efficacy of a subunit vaccine; hence this phenomenon is important in the identification of vaccine candidates to ascertain that the selected candidate can immunise against all field strains. Accordingly, the demonstration of extensive diversity and the inconsistent reaction to the p67 subunit vaccine has led to various investigations to improve the vaccine and its efficacy.

Other vaccine candidates for a subunit vaccine include the polymorphic immunodominant molecule (PIM), a *T. parva* antigen expressed in both the sporozoite and schizont stages of infection (Toye et al., 1995). MAbs have been shown to recognise PIM antigens and neutralise sporozoite infectivity; however, the recombinant PIM protein did not protect cattle against *T. parva* challenge *in vivo* (Toye et al., 1996, Ververken et al., 2008). Consistently, a recent study used a PIM-based gene gun immunisation against *T. parva* and the results showed ineffective immunisation (Fry et al., 2019).

Theileria parva schizont antigens, Tp1 and Tp2, are also being considered as possible candidates for a subunit vaccine. A study conducted in buffalo-derived and cattle-derived isolates from Kenya buffalo and cattle, have shown allelic diversity in these two antigens, with extensive polymorphism in the epitope regions (Pelle et al., 2011). According to Steinaa et al. (2012), polymorphism may be favourable as all three Tp1 epitope variants investigated had strong cross-reactivity in that study; they suggested that one epitope variant can be sufficient for inclusion in a subunit vaccine. The

investigation on the diversity of these antigens and their potential as vaccine candidates is ongoing (Elisa et al., 2015, Hemmink et al., 2016, Hemmink et al., 2018, Salih et al., 2017, Sitt et al., 2018).

Thus, the search for alternative vaccine candidates that could provide broad-spectrum immunity against both cattle- and buffalo-derived *T. parva* infections remains necessary.

2.5 *Immunity to theileriosis*

Theileria parva causes transformation of host lymphocytes. Infected lymphocytes express high levels of MHC class-I and MHC class-II molecules that could be recognised by both $CD8^+$ and $CD4^+$ T-lymphocytes (Morrison, 2007), which have key roles in immunity against *T. parva.* MHC class-I molecules, critical for the lymphocyte invasion by *T. parva* can be upregulated by IFN-γ which is a mechanism used by the parasite to increase the susceptibility of circulating lymphocytes (Tretina et al., 2015). Although it was previously believed that the IFN-γ pathway might be involved in immune response, Graham et al. (2006) showed a cytolytic immune response, rather than the IFN-γ response, to be associated with immune response and protection to *T. parva* infection.

Nonetheless, MHC class-I restricted $CD8^+$ T-cell responses play a critical role in the immune response of many protozoan infections including *Plasmodium* and *Theileria* infections (Graham et al., 2008). Consistently, immunity against *T. parva* infections is mediated by cytotoxic $CD8^+$ T-cells during the intralymphocytic schizont stage of the parasite development in the host (Connelley et al., 2011). Cytotoxic T-lymphocytes are detected fleetingly in the blood and coincide with removal of parasitized cells from lymphoid tissue. The CTLs specifically target parasitized cells; they recognise the antigens presented on the surface of MHC class-I molecules and reside within the $CD8^+$ subpopulation (Morrison et al., 1995). Although immunity to ECF is long-term compared to the partial immunity to malaria that only develops after years of exposure to *Plasmodium* spp. (Graham et al., 2006), these apicomplexan parasites follow a similar life cycle and host immune response to infection (Morrison et al., 1995). *Plasmodium falcipricum,* the causative agents of malaria, develop through the sporozoite stage as observed in *T. parva.* It has been shown that a number of malarial sporozoites migrate to the proximal lymph nodes where they induce $CD8^+$ T cell-mediated anti-sporozoite responses (Nlinwe et al., 2018). Thus, it is thought that *T. parva* sporozoites may also induce similar immune response because of the close relatedness of these two apicomplexan parasites.

Graham et al. (2006) identified CTL antigens in *T. parva* through two approaches; the first was a targeted gene approach where genes previously predicted from one of the four *T. parva* chromosomes to have a secretion signal were immunoscreened. The second approach used random immunoscreening of schizont cDNA clones. From these methods, six antigens were described; Tp1,

Tp2, Tp4, Tp5, Tp7 and Tp8, all of which were shown to illicit a CTL response (Graham et al., 2006). It was also shown that these *T. parva* antigens are dominant in animals with the corresponding MHC genotypes (Graham et al., 2008).

Strain-specific CTL responses have been observed in some cattle when immunised with a specific parasite stock (MacHugh et al., 2009). This was observed in immunization with the Muguga stock or Marikebuni 3219 stock of *T. parva*, while cross-reactive responses were noted when other stocks were used (Taracha et al., 1995). Strain-specific CTL responses provided limited protection against challenge from a heterologous *T. parva* parasites, while cross-reactive CTL responses provided full protection (Taracha et al., 1995). Using Tp1 and Tp2 antigens from ITM immunised cattle, $CD8^+$ T-cell responses exhibited a dominant response towards both antigens, which was dependent on MHC class-I molecules; thus proposed to play a critical role in the parasite strain specificity of the immune response to *T. parva* infection (MacHugh et al., 2009).

In spite of all the work that has been done to date, immunity against *T. parva* infection is still not fully understood, especially due to infections by mixed population of the parasite. Thus critical questions arise, such as those posed by Lipsitch et al. (2007) including "when a host is confronted with a pathogen with multiple antigens, how does the host decide which antigens to target, which type of immune response to deploy and which of these responses will give enough protection against the pathogen to provide immunity". The answers will be critical in the development of effective vaccines against *T. parva* infections.

2.6 *Reverse vaccinology and identification of novel vaccine candidates*

The search for alternative vaccine candidates that could provide broad-spectrum immunity against both cattle- and buffalo-derived *T. parva* infection is ongoing. One of the approaches that is used for identification of vaccine candidates employs reverse vaccinology. However, before reverse vaccinology conventional approaches were employed for this purpose. The latter involved *in vitro* attenuation of pathogens in order to acquire live-attenuated strains, followed by identification of immunising antigens that could be used in subunit vaccines (Rappuoli, 2000). This approach is time consuming and often unachievable if the pathogen cannot be grown under laboratory conditions.

Reverse vaccinology uses both immunological and genomic information to recognize appropriate protein antigens for use in a vaccine by utilizing $CD8^+$ T-cells or $CD4^+$ T-cells for identification of parasitic epitopes (Sette and Rappuoli, 2010). It also suggests the combination of antigens that could be used for a subunit vaccine (Lipsitch and O'Hagan, 2007). First discovered in 2000, reverse

vaccinology is described as *in vitro* manipulation of the pathogen genomic sequence with a purpose of mining for vaccine candidates (Lew-Tabor and Rodriguez Valle, 2016).

"Immunoinformatics" is the term used to describe the use of computational epitope mapping and reverse vaccinology to identify novel vaccine candidates and is also generally used interchangably with the term "bioinformatics" (Lew-Tabor and Rodriguez Valle, 2016). The predictive tools employed by reverse vaccinology includes prediction of protein domains, transmembrane helices, secretion, GPI anchors, cellular localisation and immune binding epitopes. Furthermore, computer software can be employed to predict peptides with binding affinity to HLA-A, -B and -C molecules (Nene et al., 2012) and has been used to predict and validate the MHC class-I epitopes in *T. parva* to better understand the T-cell responses associated with *T. parva* infection.

There are several studies that have used reverse vaccinology in an effort to identify vaccine candidates for immunization against apicomplexan parasites. The malaria causing *Plasmodium spp.*, is a great example of the effectiveness of reverse vaccinology. The RTS,S subunit vaccine was the product of bioinformatics (Ouattara et al., 2015, Nlinwe et al., 2018) and other studies have reported the use of bioinformatic tools in identifying other possible vaccine candidates for malaria immunisation (Shuaibu et al., 2010). GPI-anchored proteins are regarded as ideal vaccine candidates as they are able to anchor the antigen to the membrance of an infected cell (Kinoshita, 2016, Nathaly Wieser et al., 2019). Using a bioinformatics approach, four GPI-anchored sporozoite antigens (TpMuguga_01g0095, TpMuguga_04g00437, TpMuguga_01g00876 and TpMuguga_01g00939) were identified, whose antisera had moderate to high sporozoite neutralising abilities that could make ideal subunit vaccine candidates against *T. parva* (Nyagwange et al., 2018). Furthermore, another recent study identified 27 CTL epitopes that could induce a CTL response against *Theleria* species (Kar and Srivastava, 2018). Kar and Srivastava (2018) joined all 27 epitope vaccine candidates with an AAY linker to design what they suggested to be a potentially more cost effective multi-epitope vaccine, with enhanced cellular response if used for immunisation. However, the efficacy of this vaccine is yet to be tested.

Two years prior to the current study, our lab also embarked on the search to identify novel vaccine candidates, targeting the pathogenic stage of the parasite, schizont stage, following the reverse vaccinology approach. Predicted secreted antigenic proteins with epitope binding capacity to MHC-1 molecules were identified from the *T. parva* proteome (unpublished data) and one of these is investigated further in the current study.

2.7 *Antigenic diversity in* Theileria parva

The significance of antigenic diversity has been highlighted by the difficulties associated with the selection of the vaccine stock that will provide broad-spectrum immunity against field challenge with ECF and Corridor disease causative agents (Uilenberg, 1999). Thus, antigenic diversity cannot be neglected when identifying a candidate for the development of a subunit vaccine. Antigenic diversity can be defined as "the existence of two or more strains within a pathogen species for which the immune response induced by one strain is more effective against that particular strain (homologous) than against other (heterologous) strains" (Lipsitch and O'Hagan, 2007).

Antigenic diversity has been demonstrated in the comparison of vaccine efficacy of *T. parva* stocks (Morzaria et al., 1987, Taracha et al., 1995). In one of the studies, two groups of cattle were immunised with either Muguga or Marikebuni stocks and then challenged with the opposite stock. It was found that cattle immunised with Muguga were susceptible to challenge from Marikebuni stocks and *vise versa* (Taracha et al., 1995). This varied protection highlights the importance of antigenic diversity in the selection of vaccine stocks, especially the identification of stocks that will provide broad-spectrum immunity.

The host immune response is greatly affected by antigen polymorphism. Candidate antigens that express allelic polymorphism afford the parasite a mechanism of escape from host immune response (Nlinwe et al., 2018). For example, the subunit vaccine currently used for immunisation against malaria, RTS,S vaccine, formulated from the C-terminal of the CSP antigen, has high levels of polymorphism and exhibits limited immune response to some parasite strains (Tanabe et al., 2013, Ouattara et al., 2015, Nlinwe et al., 2018). Accordingly, it has been shown that efficient antibody responses to polymorphic antigens (e.g. apical membrane antigen 1) tend to be allele-specific with only a single allele releasing antibodies for an immune response (Nlinwe et al., 2018).

Parasite genetic and antigenic diversity greatly hinder the efficacy of potential vaccines as seen in many malaria vaccines. With vaccine development, the antigenic variability of vaccine candidates needs to be assessed early in development to improve the potential vaccine efficacy (Ouattara et al., 2015). A total of nine CTL antigens, Tp1 to Tp9, have been detected in *T. parva* (Graham et al., 2006). The genetic and antigenic properties of these antigens have found to be lacking in the Muguga cocktail as it does not incorporate the diversity found in *T. parva* field isolates .(Hemmink et al., 2016). These antigens need to be fully investigated to determine their genetic and antigenic properties to establish which antigens could be suitable candidates for a subunit vaccine. Tp1 and Tp2 have been characterized in *T. parva* field isolates from South Sudan, in comparison to the same antigens in the

Muguga stock, to assess if ITM based on this parasite stock can be used to immunize cattle there (Salih et al., 2017). The sequence difference of the two antigens compared to that of Muguga varied; Tp1 was conserved by 97.7% and Tp2 by 72.6%. Thus, the authors recommended that, based on these findings, a potential vaccine trial using the Muguga vaccine in South Sudan and different geographical areas where *T. parva* infection is present, may need to be investigated before the vaccine is administered.

Since the levels of sequence diversity amongst *T.* parva antigens differ greatly between buffalo-derived and cattle-derived isolates (Pelle et al., 2011), it is important to determine these differences when screening for a vaccine candidate that can protect against infections by both parasite types. Hemmink et al. (2018) recently investigated antigenic diversity between buffalo populations from South Africa and Kenya. The results highlighted that there is widespread allelic diversity amongst the genes Tp1, Tp2, Tp4, Tp5, Tp6 and Tp10 both within and amongst the populations; however, it was also noted that there is a high level of antigen conservation at the amino acid level suggesting a common ancestor preceding geographical isolation (Hemmink et al., 2018). Such information is important when identifying vaccine candidates that can immunise against both buffalo-derived and cattle-derived *T. parva* infections; hence, Ouattara et al. (2015) suggested that future vaccine approaches should be targeted at assessing conserved antigens that would induce broad-spectrum efficacious antibodies. Conserved antigens have been shown to induce specific immune responses in various other vaccine trials (Sitt et al., 2018, Musoke et al., 2005, Nyagwange et al., 2018).

2.8 *Study justification*

Cattle still exhibit susceptibility to infection and challenge from some parasite strains, particularly buffalo-derived *T. parva,* after immunization by ITM. Considering the limitations of using live vaccines, it has therefore become imperative to explore other forms of immunisation that will provide broad-spectrum immunity while considering the diverse nature of *T. parva* field strains. Recently, most focus has been on subunit vaccines which are more cost effective (Blake et al., 2015). Thus, in a related study from our lab, an *in silico* approach has been employed to identify novel vaccine candidates for immunisation against *T. parva.* Genes encoding parasite antigens with CTL epitopes with binding affinity to MHC-class I alleles, in line with the mechanism of host immunity against *T.* parva infection have been identified. Antigenic diversity in the epitope regions plays a vital role in the selection of suitable vaccine candidates. Thus it is necessary for putative vaccine candidates to be characterised in both cattle- and buffalo-derived *T. parva* isolates. Subsequently, proteins conserved along *T. parva* strains can be further investigated if they can illicit an immune response and their efficacy assesed for suitability as vaccine candidates.

The *T. parva* pathogenic schizont stage is responsible for the reversible cancer-like phenotype in *T. parva*-infected lymphocytes and transcriptome studies have revealed that the majority of transcriptionally active genes in the genome are found in this developmental stage (Bishop et al., 2005). Thus, antigenic proteins expressed in the schizont stage may be good targets in the identification of potential vaccine candidates for the development of a subunit vaccine.

Accordingly in this study, the genetic and antigenic diversity of one of the vaccine candidates identified from *in silico* analysis was evaluated. Sequence, pairwise and phylogentic analysis was performed to determine if the predicted vaccine candidate is conserved amongst buffalo- and cattle-derived *T. parva* isolates as well as between different geographical locations. The results will form part of determining factors on whether the predicted antigen is a suitable candidate for development of a subunit vaccine against *T. parva.*

CHAPTER 3

3. METHODS AND MATERIALS

3.1 *Selection of vaccine candidates*

Potential vaccine candidates (n = 13) were selected based on *in silico* analysis previously performed in our lab (T Phala, unpublished data). The vaccine candidates were identified based on whether they possess the following: glycosylphosphatidylinositol (GPI) anchor and transmembrane domains (TM) as indicators of secreted proteins, have no homologs in the host proteome, are predicted protective antigens and finally have the MHC class-1 binding affinity. The analyses were performed using tools presented in Table 3.1. Figure 3.1 demonstrates the workflow used in the identification of the potential vaccine candidates.

Table 3.1: *The list of* in silico *tools used to identify the vaccine candidates.*

Analysis	Tool	URL/Reference
GPI Anchor prediction	PredGPI	http://gpcr.biocomp.unibo.it/predgpi/pred.htm
Protein secreted	SignalP v.5.0	Almagro Armenteros et al., 2019
	PrediSi	http://www.predisi.de/home.html
	Signal-BLAST	Frank and Sippl, 2008
	TargetP v.1.1	Emanuelsson et al., 2007
	Phobius	Krogh et al., 2007
Transmembrane domains	TMHMM v. 2.0	Krogh et al., 2001
	Phobius	Krogh et al., 2007
Homologs	Protein BLAST	https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins
	PSI BLAST	https://blast.ncbi.nlm.nih.gov/Blast.cgi?CMD=Web&PAGE=Proteins&PROGRAM=blastp&RUN_PSIBLAST=on
Protective antigens	VaxiJen v. 2.0	Doytchinova and Flower, 2007
	ANTIGENpro	http://scratch.proteomics.ics.uci.edu/index.html
MHC class-1 binding	IEDB	https://www.iedb.org/

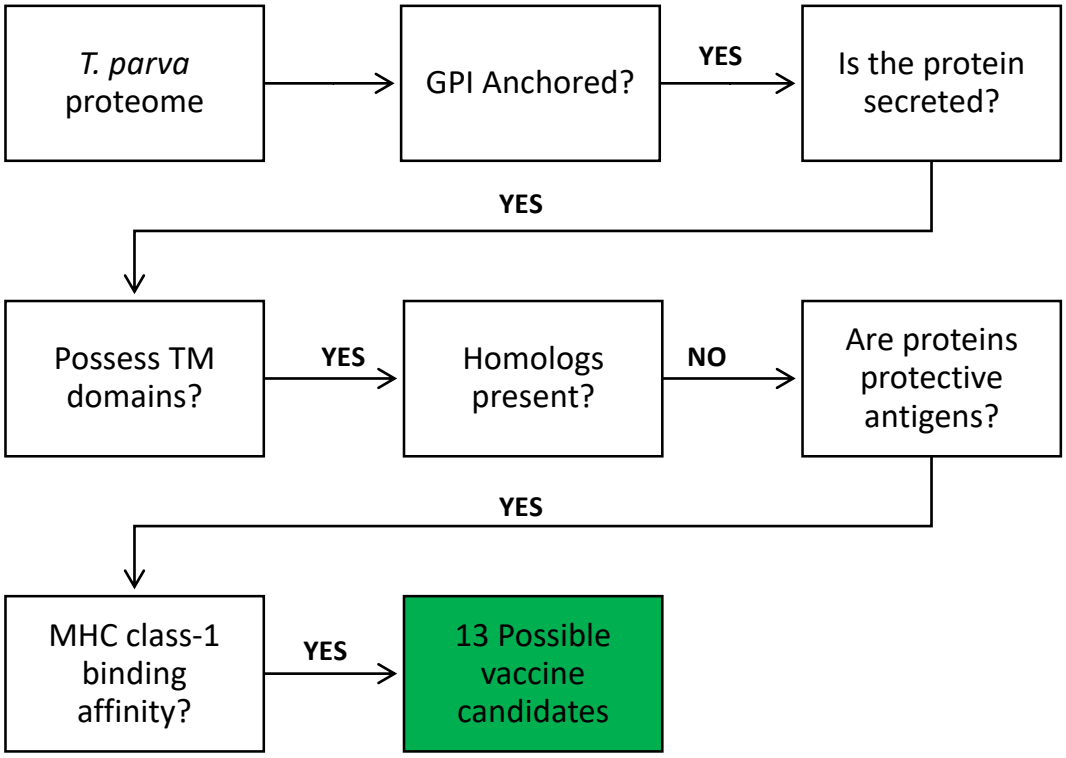

Figure 3.1: *Graphical representation showing the selection criteria of 13 potential vaccine candidates* in silico *analysis.*

3.1.1 Gene expression profiling of the selected candidates

Transcriptome data previously generated from two *T. parva* isolates, Muguga and 7014, representing the cattle-and buffalo-derived parasites respectively (KP Sibeko-Matjila, unpublished data), was analysed to determine whether the candidate genes are expressed in the pathogenic stage, the schizont. The expression levels of genes were normalised by considering both the library size and gene length effects with respect to the RPKM values (reads per kilo base of gene model per million mapped reads) (Mortazavi et al., 2008). The average expression values were normalised in quantiles. Transcripts were considered to be expressed if RPKM values were ≥ 10.

3.2 *Genotyping of the selected candidate gene*

The vaccine candidate, TP04_0028 (accession number XM_758570.1), of 1689 base pairs (bp) in length, was selected for genotyping subsequent to the above analyses. The nucleotide sequence of this gene encoding a hypothetical protein was retrieved from the NCBI Genbank database for primer design for PCR amplification and subsequent sequence analysis.

3.2.1 Primer design and PCR optimisation

Oligonucleotide primers were designed to target TP04_0028 gene using Primer-BLAST (NCBI) (Ye et al., 2012). Sequences of the selected primer pair were submitted to Inqaba Biotec (South Africa) for oligonucleotide primer synthesis. Various temperatures (55°C, 59°C and 65°C) and primer concentrations (0.5 µM, 1 µM and 2 µM) were tested to determine the optimal reaction conditions for specific amplification of the target gene region. The details of the PCR reaction are described in section 3.2.3 below.

3.2.2 Sample selection

The 34 DNA samples used for genotyping were extracted from blood collected from cattle and buffalo from Uganda, Mozambique, Kenya and South Africa as shown in Table 3.2 and Figure 3.3. Extracted DNA samples were made available from a previous PhD study (Mukolwe et al., 2020), with permission from the Department of Agriculture, Forestry and Fisheries (DAFF), under Section 20 number 12/11/1/1. All DNA samples had been previously tested positive for *T. parva* infection using the *T. parva* specific real-time PCR assay, which targets the 18S rRNA gene (Sibeko et al., 2008). The samples used in the current study were selected to represent the different p67 allele types (1, 2, 3 and 4), used to differentiate between cattle and buffalo-derived *T. parva* parasites, as previously determined by Mukolwe et al., (2020). The p67 allele 1 is characteristic of cattle-derived *T. parva* while alleles 2, 3, 4 are specific for buffalo-derived parasite isolates (Sibeko et al., 2010).

Table 3.2: *The number of DNA samples used in the study, their host from which blood was obtained, geographic origin and the respective p67 allele profiles of each parasite population.*

Country	Mammalian host	Number of samples used (n = 34)	Area	p67 allele(s)
Kenya	Cattle	5	Nakuru county	1
	Buffalo	8	Ol Pejeta Conservancy	2, 3
Uganda	Cattle	5	Kaabong (n = 1)	
			Nakapiripirit (n = 2)	1
			Kiruhura district 1 (n = 1)	
			Kiruhura district 2 (n = 1)	
Mozambique	Buffalo	6	Marromeou Game Reserve	1, 2, 3, 4
South Africa	Cattle	7	Hluvukani, Mnisi community	1, 2, 3, 4
	Buffalo	3	Kruger National Park	1, 2, 3, 4

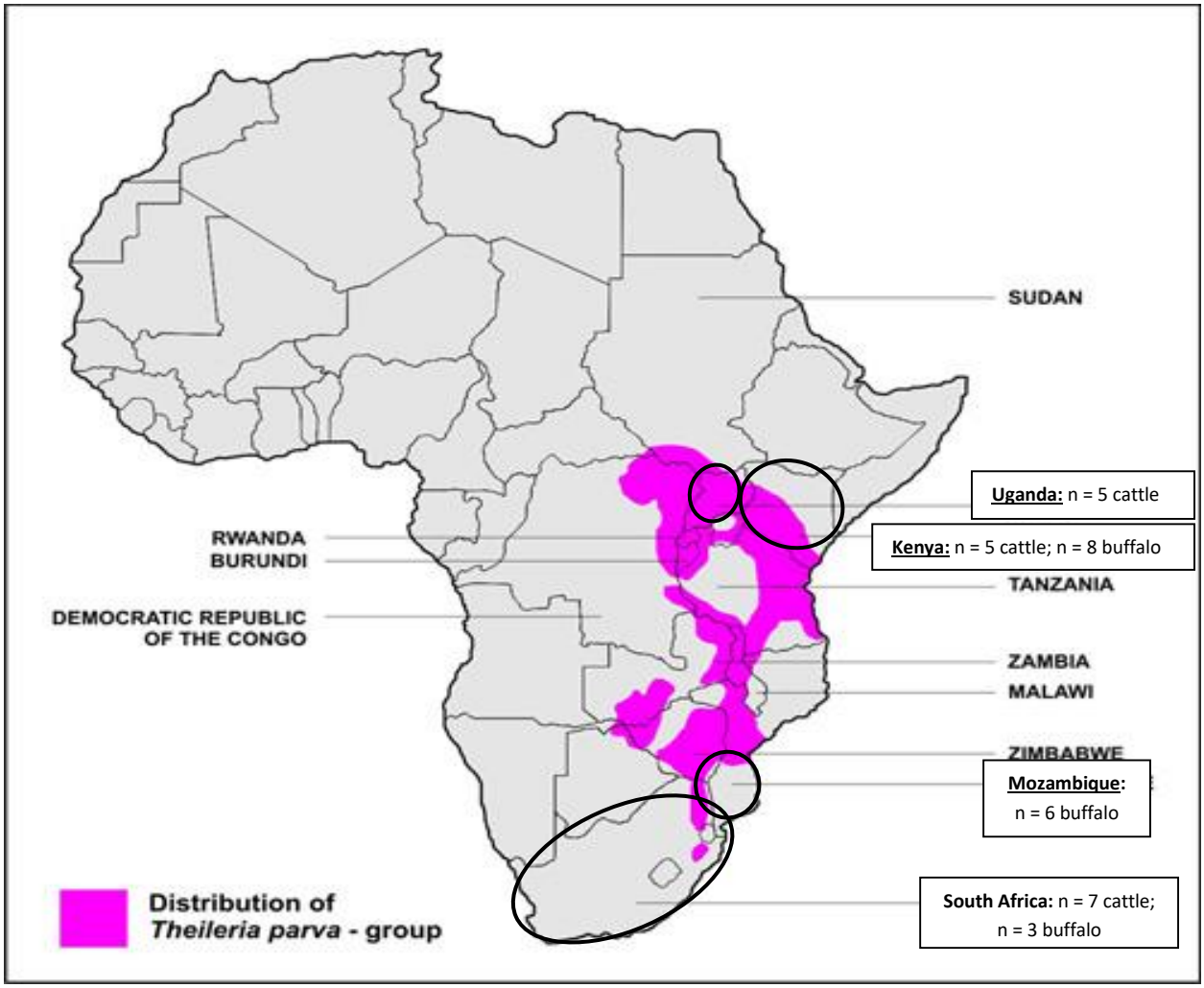

Figure 3.2: *The distribution of T. parva in Africa (adapted from http://www.afrivip.org/sites/default/files/Theilerioses/index.html). The encircled areas indicate the countries where samples used in the current study were obtained and the number of cattle and buffalo from which blood was collected for DNA extraction.*

3.2.3 PCR amplification

The partial TP04_0028 gene, 754 bp in length, was amplified using Phusion Flash High Fidelity master mix (2x) (Thermo Fisher Scientific Inc.) following the manufacturer's instructions. The amplification mixture consisted of 100 ng/µl of *T. parva* genomic DNA, 10 µl Phusion Flash High Fidelity PCR Master mix (Thermo Fisher Scientific Inc.), 0.5 µM of forward and reverse primer, respectively, and nuclease-free H_2O to make up a final reaction volume of 20 µl. The *T. parva* positive (DNA from a known *T. parva*-infected buffalo, 102) and negative controls were included in each experimental run. For the negative control the DNA template was replaced with an equivalent volume of sterile nuclease-free H_2O. The amplification reactions were performed in the PCR Veriti thermocycler (Applied Biosystems) under the following conditions: initial denaturation at 98°C for 10 seconds (s), followed by 30 cycles

of denaturation at 98°C for 1 s, annealing of primers at 65°C for 5 s and extension at 72°C for 15 s; lastly, one cycle of final extension was performed at 72°C for 1 minute. Once amplification was completed, 3 µl of the PCR products was analysed by electrophoresis on a 2% agarose gel stained with 3 µl ethidium bromide (Promega) per 100 ml agarose, at 100V. A 100 bp plus DNA molecular weight marker (GeneRuler, Thermo Fisher Scientific Inc.) was used to determine the size of the PCR products. The presence of the PCR product was visualised under UV light. In the absence of a visible PCR product, re-amplification was performed using the Phusion Flash High Fidelity PCR Master mix as described above, with 1 µl PCR product as a DNA template. The number of amplification cycles was reduced to 20.

3.2.4 PCR product purification

PCR products of the correct base pair size were purified by using QIAquick PCR Purification Kit (Qiagen) following manufacturer's instructions. Briefly, 5 volumes of Buffer PB were added to 1 volume of the PCR reaction and centrifuged at 17 900 x g (13 000 rpm) for 1 minute. The mixture was washed with 750 µl Buffer PE and centrifuged for 1 minute. The purified product was eluted in 30 µl of buffer and analysed by electrophoresis on a 2% agarose gel stained as described in section 3.2.3.

3.2.5 DNA Cloning

In the field *T. parva* infections often occur in populations of mixed strains, furthermore some samples had multiple PCR products even at optimised conditions, especially samples from buffalo, therefore it was necessary to perform cloning for all amplicons.

3.2.5.1 Ligation of TP04_0028 amplicons into pJET 1.2/blunt cloning vector

Purified PCR products were ligated onto a plasmid vector using the CloneJet PCR Cloning Kit (Thermo Scientific Inc.) following the manufacturer's instructions. Briefly, 10 µl ligation reaction was made up of 4 µl 2x Reaction Buffer, 0.5 µl pJET 1.2/blunt cloning vector (50 ng/µl), 0.5 µl T4 DNA ligase, 50 ng/µl purified PCR product and nuclease free water, for each PCR product. The ligation reactions were incubated at room temperature for 7 minutes prior to transformation.

3.2.5.2 Transformation of E. coli competent cells

The Mix & Go Competent cells (*E. coli* strain JM109) (Inqaba Biotec South Africa) were used for transformation. Briefly, 1.5 µl ligation reaction was mixed with 25 µl Mix & Go Competent cells (*E. coli* strain JM109) and immediately plated on Ampicillin agar plates prepared from imMedia Growth Media

(Invitrogen) using a sterile spreader. The plates with transformed bacteria were grown overnight at 37°C.

3.2.5.3 Selection and confirmation of recombinants

A minimum of three colonies from each sample plate were picked and subjected to colony PCR to confirm the presence of the amplicon of interest. Colony PCR was performed using DreamTaq Green PCR Mastermix (Thermo Fisher Scientific Inc.) according to the manufacturer's instructions. Briefly, 20 µl PCR reaction was prepared using 10 µl DreamTaq Green PCR Mastermix (Thermo Fisher Scientific Inc.), 0.4 µl each of the pJET1.2 forward and reverse primer supplied in the cloning kit and 9.2 µl nuclease-free water. A sterile pipette tip was used to pick up each colony and transfer the same purified colony into the respective PCR tube. The amplification reactions were performed in a PCR Veriti thermocycler (Applied Biosystems) under the following conditions: initial denaturation at 95°C for 3 minutes, followed by 25 cycles of denaturation at 94°C for 30s, annealing of primers at 60°C for 30s, extension at 72°C for 1 minute and one cycle of final extension at 72°C for 7 minutes. Once amplification was completed, PCR products were visualised as previously described in section 3.2.3. Only recombinants confirmed by colony PCR to possess the correct PCR product were selected for sequencing.

3.2.6 Sequencing and sequence data analysis

Selected PCR products were sent to Inqaba Biotec (South Africa) for Sanger sequencing. The sequencing was performed in both forward and reverse direction for each sample, using PCR primers. Sequences were analysed for quality and trimmed using Chromas version 2.6.5 (2018) (Technelysium Pty Ltd, South Brisbane, Queensland, Australia). Sequence similarity analysis was performed using BLAST to determine if the correct gene was sequenced. A pairwise alignment of edited forward and reverse sequences was done in BioEdit version 7.2.5 (Hall, 1999) to generate a consensus sequence for each sample.

3.2.7 Sequence alignment and phylogenetic analysis

ClustalX (Thompson et al., 1994) was used to create a multiple alignment in BioEdit version 7.2.5 (Hall, 1999) for both nucleotide and protein sequences. A pairwise distance analysis was done in MEGA7 (Kumar et al., 2016) for both nucleotide and protein sequences to determine the sequence similarities between sequences generated in this study and the reference, based on the number of nucleotide and amino acid differences as well as percentage differences. The sequences were analysed for variance or conservation compared to the reference. Differences and similarities were also compared

between cattle- and buffalo-derived *T. parva* samples as well as between the different geographic origins.

Neighbor-Joining and Maximum likelihood phylogenetic trees were constructed using MEGA7 (Kumar et al., 2016) to determine the relationship of TP04_0028 genotypes from both nucleotide and protein sequences, in cattle- and buffalo-derived *T. parva* samples from different countries. The evolutionary distances of the Neighbor-Joining trees were computed using the Jukes-Cantor method, in units of the number of base substitutions per site. The analysis involved 35 nucleotide sequences and 34 amino acid sequences for each respective phylogenetic tree analysis. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. The number of bootstrap replications was set to 1000.

Initial tree(s) for the heuristic search for the Maximum likelihood phylogenetic trees were obtained automatically by applying Neighbor-Join and BioNJ algorithms to a matrix of pairwise distances estimated using the Maximum Composite Likelihood (MCL) approach and selecting the topology with superior log likelihood value. The trees were drawn to scale, with branch lengths measured in the number of substitutions per site. Once more, the analysis involved 35 nucleotide sequences and 34 amino acid sequences for each respective tree. All positions with less than 95% site coverage were eliminated; thus, fewer than 5% alignment gaps, missing data, and ambiguous bases were allowed at any position. The number of bootstrap replications for this analysis was also set to 1000.

For both Neighbor-Joining and Maximum likelihood phylogenetic tree analyses, branches with bootstrap values of < 50 were collapsed.

3.2.8 Analysis of sequence variation in the epitope regions

To determine amino acid sequence variations in the predicted epitope regions the multiple sequence alignment previously generated, as described in section 3.2.7, was analysed. Epitope regions conserved in different BoLA alleles were identified using MHC Cluster v.2.0 (Thomsen et al., 2013) (Table 3.3). The predicted epitope regions for TP04_0028 were retrieved using IEDB (https://www.iedb.org/), using the BoLA alleles identified above. These predicted epitope regions were analysed for areas of conservation amongst the protein sequences.

Table 3.3: *The list of alleles of 20 predicted epitopes from IEDB forming part of TP04_0028 amplicon.*

Allele	Start	End	Length	Peptide
BoLA-T2c	363	373	11	**ELFIIKL**SNKL
BoLA-N:03401	358	367	10	ESQVT**ELFII**
BoLA-N:03401	438	447	10	KETG**KMPDGI**
BoLA-N:02401	328	338	11	KI**TRLAPSNIY**
BoLA-AW10	*442*	*452*	*11*	**KMPDGILDIIL**
BoLA-N:03001	*442*	*452*	*11*	**KMPDGILDIIL**
BoLA-N:02401	355	365	11	LRNES**QVTELF**
BoLA-N:03401	357	366	10	NES**QVTELF**I
BoLA-AW10	314	324	11	PSG**YGKFSTNL**
BoLA-N:02401	381	391	11	**RHEGVRYKNAY**
BoLA-N:03401	356	365	10	RNE**SQVTELF**
BoLA-N:02401	*386*	*396*	*11*	**RYKNAYYKRTL**
BoLA-N:01301	*386*	*396*	*11*	**RYKNAYYKRTL**
BoLA-N:03401	*232*	*241*	*10*	**SDKIKTKFLL**
BoLA-N:01301	359	369	11	**SQVTELFIIKL**
BoLA-AW10	*231*	*241*	*11*	S**SDKIKTKFLL**
BoLA-N:01301	*354*	*364*	*11*	**TLRNESQVTEL**
BoLA-T2c	*354*	*364*	*11*	**TLRNESQVTEL**
BoLA-AW10	*330*	*340*	*11*	**TRLAPSNIYTV**
BoLA-N:03001	*330*	*340*	*11*	**TRLAPSNIYTV**
BoLA-N:02401	248	258	11	TTKYFRLNADW*
BoLA-N:03401	376	385	10	VEICL**RHEGV**
BoLA-N:03401	261	270	10	VEKDFYDKFL*
BoLA-AW10	317	327	11	**YGKFSTNL**SNI

The peptides in bold font represent epitope regions found in more than one BoLA allele. The two peptides with * are not conserved between BoLA alleles.

CHAPTER 4

4. RESULTS

4.1 *Selection of a vaccine candidate for genotyping*

The list of vaccine candidates (n = 13) identified using *in silico* analysis is shown in Table 4.1. Six of these (TP01_004, TP01_1056, TP02_0553 TP02_0950, TP03_0512 and TP04_0028) met all criteria (indicated in Chapter 3) for potential vaccine candidates. Following the expression analysis, three candidate genes, TP02_0553, TP02_0950 and TP04_0028, were found to be expressed in the schizont stage, the stage responsible for the pathology of the disease, and at comparable levels in the cattle-(Muguga) and buffalo-derived (7014) *T. parva* isolates (Table 4.1). From these three, TP04_0028, was selected for genotyping based on the length (1689 bp) and predicted epitopes conserved between BoLA alleles, as shown in Table 3.3.

4.2 *Selected gene amplification*

4.2.1 Selection of Primers targeting the partial TP04_0028 gene

The primer design output with various pair options is shown in Figure 4.1. Primer set 6 was selected for amplification of 754 bp of the TP04_0028 gene. The selected primer set consisted of the forward primer CAAGGCCTCCGAAGAGGAAA (5’ to 3’) (Tm 59.68°C) and the reverse primer TTCTGTTTGAGTGGCCGCAT (5’ to 3’) (Tm 60.54°C), with a GC content of 55.00 % and 50.00 %, respectively. Furthermore, it produced the longest amplicon (754 bp), which included 20 out of 44 predicted epitopes of TP04_0028 (Figure 4.2). Among the 20 predicted epitopes within the target region 18 are conserved between BoLA alleles.

***Table 4.1**: Information on the potential vaccine candidates identified by* in silico *analysis, including expression profile in* T. parva *7014 and Muguga isolates and predicted protein properties.*

*Gene ID	Accession number	Annotation status of proteins	Size (bp)	7014 RPKM AVE	Muguga RPKM AVE	GPI anchor	Signal peptide	TM domain	Homologs in host proteome	Protective antigens	MHC binding
TP01_0004	XM_760438.1	hypothetical	1407	151,584	0,425	Y	Y	Y	N	Y	Y
TP01_0049	XM_760483.1	hypothetical	405	3,836	9,902	Y	Y	Y	N	Y	N
TP01_0095	XM_760529.1	hypothetical	762	56,997	56,217	Y	Y	Y	N	Y	N
TP01_1056/32 kDa surface antigen	XM_761484.1	annotated	1045	261,926	742,524	Y	Y	Y	N	Y	Y
TP01_1218	XM_761646.1	hypothetical	498	3,318	2,354	Y	Y	Y	N	N	Y
TP02_0553	XM_760026.1	hypothetical	636	27,257	34,951	Y	Y	Y	N	Y	Y
TP02_0950	XM_760424.1	hypothetical	1293	30,992	22,889	Y	Y	Y	N	Y	Y
TP03_0134	XM_758059.1	hypothetical	576	4,473	10,851	Y	Y	Y	N	N	N
TP03_0135	XM_758060.1	hypothetical	573	6,010	2,088	Y	Y	Y	N	Y	N
TP03_0136	XM_758061.1	hypothetical	927	13,437	305,004	Y	Y	Y	N	Y	N
TP03_0512/acid phosphatase	XM_758447.1	annotated	1215	30,894	107,185	Y	Y	Y	N	Y	Y
TP04_0028	XM_758570.1	hypothetical	1689	8,923	12,416	Y	Y	Y	N	Y	Y
TP04_0251	XM_758793.1	hypothetical	2067	218,017	308,923	Y	Y	Y	Y	N	N

*****Candidates highlighted in green are those that met all criteria stipulated for identification of potential antigenic proteins. The gene highlighted in blue met all criteria and was selected for genotyping.

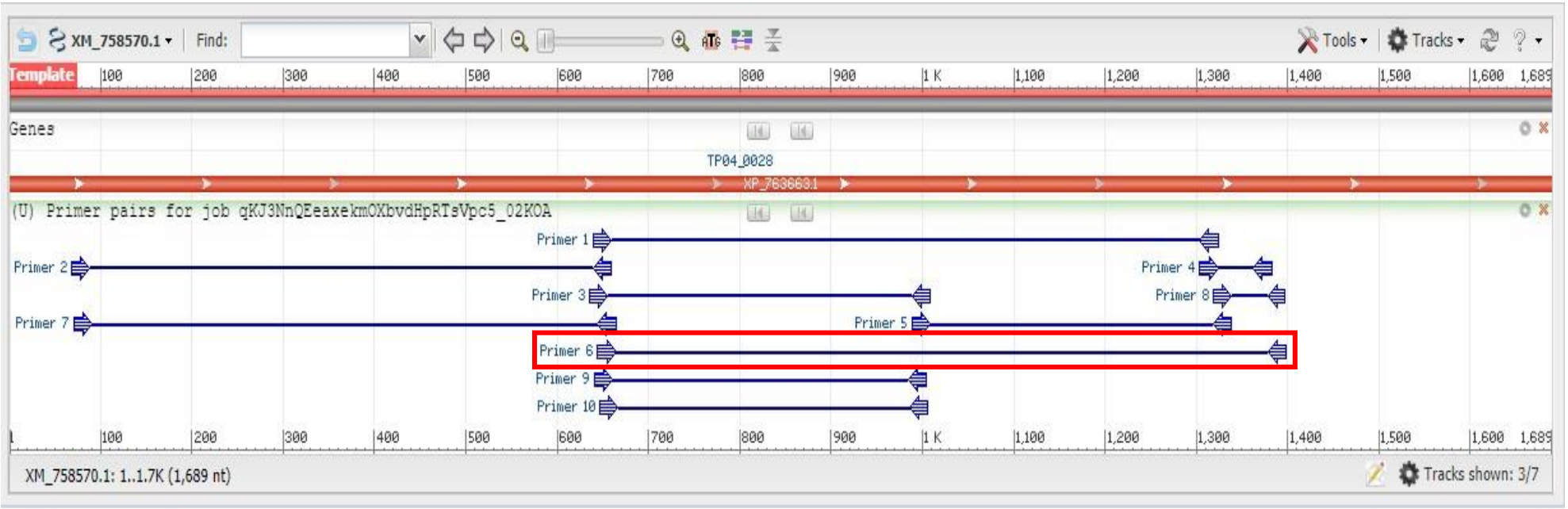

Figure 4.1: *Primer pair options obtained from Primer Blast (NCBI); Primer set 6 (in a red box) was selected for the amplification of TP08_0028.*

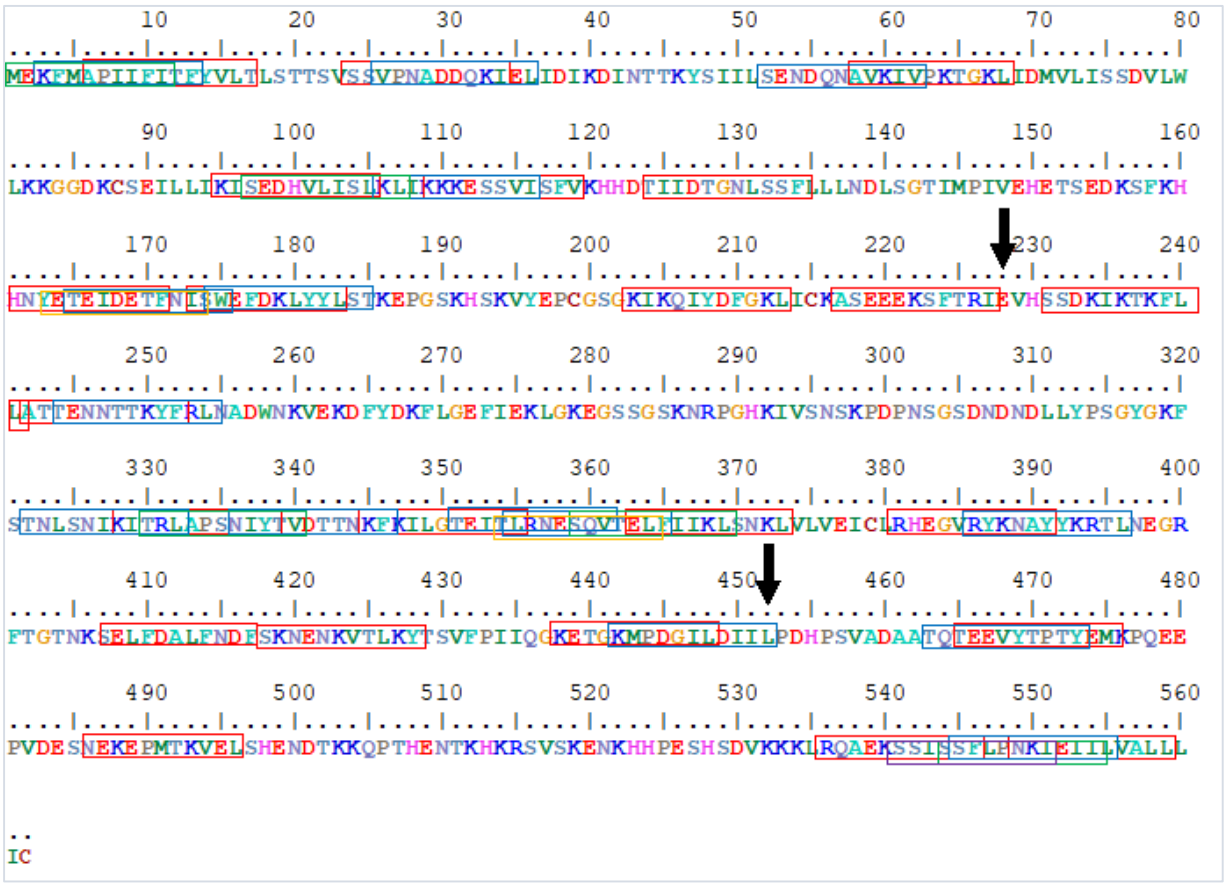

Figure 4.2: *Graphical representation of the entire TP04_0028 translated protein sequence, showing the region of 20 predicted epitopes included in the PCR target region within the two black arrows. The coloured boxes represent the predicted epitope region.*

4.2.2 PCR optimization and amplification of TP04_0028 gene in cattle and buffalo blood samples

Optimal amplification of TP04_0028 gene targeted region was achieved using 0.5 µM for each primer in a 20 µl PCR reaction and the annealing temperature of 65°C for 5 seconds. These conditions generated the desired PCR product of 754 bp (Figure 4.3). The targeted region of TP04_0028 gene was successfully amplified from all DNA samples prepared from both cattle (n = 17) and buffalo (n = 17) blood from South Africa, Mozambique, Kenya, and Uganda.

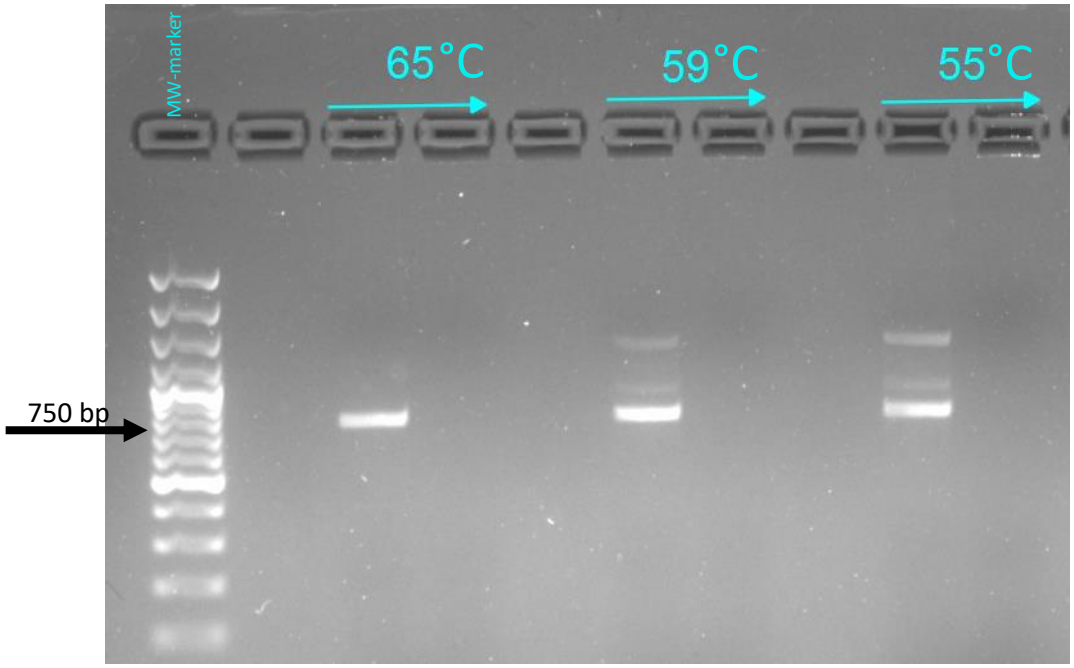

***Figure 4.3**: The agarose gel electrophoresis showing PCR products produced at different annealing temperatures during the optimisation of amplification of the 754 bp target region of TP04_0028 gene; a 100 bp DNA molecular weight (MW) marker (GeneRuler) was used to determine the PCR product size.*

4.3 *Sequence data analysis*

Thirty-four (34) consensus sequences representing 17 samples from cattle and 17 from buffalo, were generated from forward and reverse sequences. All sequences were confirmed by Blast sequence similarity search to be the TP04_0028 (XM_758570.1) gene.

4.3.1 Nucleotide sequence data analysis

The multiple sequence alignment (Figure 4.4) and pairwise analyses (Tables 4.2 and 4.3) of TP04_0028 nucleotide sequences generated in the current study revealed high similarities for some sequences against the reference sequence (XM_758570.1); however, notable differences were also identified for some. Overall, sequences from cattle samples from Kenya and Uganda (representing cattle-derived *T. parva*) were conserved, while several sequence variations were identified from other sample groups, representing buffalo-derived *T. parva,* when compared to the reference sequence (Figure 4.4). Notably, nucleotide base variations occurred in the same positions compared to the reference sequence and these nucleotide mutations were detected in all sequences; with an exception of five sequences from South African (SA) cattle (C71, C84, C941, F369 and 258). These common mutations occurred at nucleotide positions 613 and 615. Furthermore, some sequences from SA cattle and Mozambican (Moz) buffalo had unique mutations (seven base pair insertions) between positions 238 and 245 bp.

In the pairwise distance analysis, sequences with the stringent percentage differences of ≥ 10% were considered diverse; this translates to ≤ 90% similarity compared to the reference sequence. The overall average sequence variation for all sequences analysed was 2%, indicative of 98% similarity between the reference and all other sequences. Sequences obtained from cattle samples from Kenya and Uganda were similar to TP04_0028 reference sequence by 99 - 100% (Table 4.2). On the contrary, most nucleotide differences were detected from buffalo-derived *T. parva* sequences obtained from cattle samples from South Africa and buffalo samples from South Africa, Kenya and Mozambique. The highest percentage difference was identified from SA cattle C91, with a 4% sequence difference to the reference. Although the majority of sequences generated buffalo-derived *T. parva* parasites from buffalo or cattle samples, displayed the highest percentage difference, there were a few (four) sequences from this group that had percentage differences less than 1%, which include, SA cattle C81, SA buffalo 252, Moz buffalo SR4 and SR12. Besides these four and sequences from samples from Kenya and Uganda cattle samples, all other sequences had percentage differences raging between 1.3 and 4%.

Contrary to sequences from cattle samples, sequences from Kenyan buffalo showed the highest percentage nucleotide differences compared to the rest of the samples groups, ranging from 3,0% - 3,5%. Between sequences from buffalo, the least number of nucleotide differences were from sequences from Mozambican samples, at 0,3 - 1,9%. Notably, samples from South Africa (both cattle and buffalo) were the most diverse, with sequence difference from 0,4 - 4,0%, revealing sequences with very high similarity (95,9% - 99,6%) against the reference sequence.

Among buffalo-derived *T. parva* parasites, sequences from cattle samples from South Africa (particularly C71, C84, C89, C91 and F369) showed high differences (1 - 4%) against sequences from buffalo from Mozambique and high similarity (≥ 99,5%) in comparison to sequences from buffalo from Kenya. Conversely, sequences from buffalo samples SR4, SR10 and SR12 from Mozambique, had high (≥ 1,9%) sequence variations against sequences from Kenya. Also notable was the high difference between sequence C81 compared to others from cattle samples from South Africa. Overall, sequences from buffalo-derived *T. parva* from samples from South Africa (7 out of 10 samples analysed) and buffalo samples from Kenya (all analysed samples) were the most diverse in comparison to cattle-derived parasite sequences from Uganda and Kenya cattle samples.

The differences noticed in the percentage pairwise distance analysis was consistent with the differences in the number of nucleotide base (Table 4.3). The average number of base pair differences among all analysed sequences was 13 bp. All the Kenya buffalo samples, South African cattle samples C71, C84, C89, C91, F369 and South African buffalo 258 consistently showed the highest number of

base pair differences compared to the reference sequence, up to 26 bp detected from South African cattle C91 sequence.

Table 4.2: *Pairwise distance alignment of nucleotide sequences showing the percentage sequence differences between each sequence and the reference gene sequence, TP04_0028 (XM_758570.1), with the average percentage difference of 2%.*

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35
1	TP04_0028																																			
2	UG_cat_MA1-52	0,0																																		
3	UG_cat_MK2-12	0,0	0,0																																	
4	UG_cat_KLP5-4	0,0	0,0	0,0																																
5	UG_cat_NA1-45	0,0	0,0	0,0	0,0																															
6	UG_cat_NK1-25	0,1	0,1	0,1	0,1	0,1																														
7	Ken_cat_KC3	0,0	0,0	0,0	0,0	0,0	0,1																													
8	Ken_cat_KC4	0,0	0,0	0,0	0,0	0,0	0,1	0,0																												
9	Ken_cat_KC5	0,0	0,0	0,0	0,0	0,0	0,1	0,0	0,0																											
10	Ken_cat_KC9	0,0	0,0	0,0	0,0	0,0	0,1	0,0	0,0	0,0																										
11	Ken_cat_KC10	0,0	0,0	0,0	0,0	0,0	0,1	0,0	0,0	0,0	0,0																									
12	Ken_buff_BF6	3,3	3,3	3,3	3,3	3,3	3,5	3,3	3,3	3,3	3,3	3,3																								
13	Ken_buff_BF14	3,3	3,3	3,3	3,3	3,3	3,5	3,3	3,3	3,3	3,3	3,3	0,3																							
14	Ken_buff_BF15	3,3	3,3	3,3	3,3	3,3	3,5	3,3	3,3	3,3	3,3	3,3	0,0	0,3																						
15	Ken_buff_BF21	3,3	3,3	3,3	3,3	3,3	3,5	3,3	3,3	3,3	3,3	3,3	0,0	0,3	0,0																					
16	Ken_buff_BF24	3,0	3,0	3,0	3,0	3,0	3,2	3,0	3,0	3,0	3,0	3,0	0,3	0,3	0,3	0,3																				
17	Ken_buff_BF26	3,3	3,3	3,3	3,3	3,3	3,5	3,3	3,3	3,3	3,3	3,3	0,0	0,3	0,0	0,0	0,3																			
18	Ken_buff_BF28	3,5	3,5	3,5	3,5	3,5	3,6	3,5	3,5	3,5	3,5	3,5	0,1	0,4	0,1	0,1	0,4	0,1																		
19	Ken_buff_BF12	3,3	3,3	3,3	3,3	3,3	3,5	3,3	3,3	3,3	3,3	3,3	0,0	0,3	0,0	0,0	0,3	0,0	0,1																	
20	SA_cat_C71	3,2	3,2	3,2	3,2	3,2	3,3	3,2	3,2	3,2	3,2	3,2	0,7	1,0	0,7	0,7	1,0	0,7	0,9	0,7																
21	SA_cat_C81	0,4	0,4	0,4	0,4	0,4	0,6	0,4	0,4	0,4	0,4	0,4	2,9	2,9	2,9	2,9	2,6	2,9	3,0	2,9	3,0															
22	SA_cat_C84	3,8	3,8	3,8	3,8	3,8	3,9	3,8	3,8	3,8	3,8	3,8	1,0	1,3	1,0	1,0	1,3	1,0	1,2	1,0	0,6	3,3														
23	SA_cat_C89	3,3	3,3	3,3	3,3	3,3	3,5	3,3	3,3	3,3	3,3	3,3	0,4	0,7	0,4	0,4	0,4	0,4	0,6	0,4	1,2	2,9	1,2													
24	SA_cat_C91	4,0	4,0	4,0	4,0	4,0	4,1	4,0	4,0	4,0	4,0	4,0	0,6	0,9	0,6	0,6	0,9	0,6	0,7	0,6	0,7	3,5	0,4	0,7												
25	SA_cat_C108	2,6	2,6	2,6	2,6	2,6	2,7	2,6	2,6	2,6	2,6	2,6	0,7	0,7	0,7	0,7	0,4	0,7	0,9	0,7	0,9	2,4	1,2	0,9	1,3											
26	SA_cat_F369	3,3	3,3	3,3	3,3	3,3	3,5	3,3	3,3	3,3	3,3	3,3	0,9	1,2	0,9	0,9	1,2	0,9	1,0	0,9	0,1	3,2	0,4	1,0	0,6	1,0										
27	SA_buff_252	0,9	0,9	0,9	0,9	0,9	1,0	0,9	0,9	0,9	0,9	0,9	2,7	2,7	2,7	2,7	2,4	2,7	2,9	2,7	2,9	0,4	3,2	2,7	3,3	2,2	3,0									
28	SA_buff_258	3,3	3,3	3,3	3,3	3,3	3,5	3,3	3,3	3,3	3,3	3,3	0,9	1,2	0,9	0,9	1,2	0,9	1,0	0,9	0,1	3,2	0,4	1,0	0,6	1,0	0,0	3,0								
29	SA_buff_270	1,5	1,5	1,5	1,5	1,5	1,6	1,5	1,5	1,5	1,5	1,5	1,8	1,8	1,8	1,8	1,5	1,8	1,9	1,8	1,9	1,3	2,3	1,8	2,4	1,0	2,1	1,8	2,1							
30	Moz_buff_SR1	1,3	1,3	1,3	1,3	1,3	1,5	1,3	1,3	1,3	1,3	1,3	2,6	2,6	2,6	2,6	2,2	2,6	2,7	2,6	2,7	1,2	3,0	2,6	3,2	1,8	2,9	0,7	2,9	1,3						
31	Moz_buff_SR3	1,8	1,8	1,8	1,8	1,8	1,9	1,8	1,8	1,8	1,8	1,8	2,1	2,1	2,1	2,1	2,1	2,1	2,3	2,1	2,2	1,3	2,2	2,1	2,4	1,9	2,1	1,8	2,1	0,9	2,2					
32	Moz_buff_SR4	0,7	0,7	0,7	0,7	0,7	0,9	0,7	0,7	0,7	0,7	0,7	3,2	3,2	3,2	3,2	2,9	3,2	3,3	3,2	3,3	0,9	3,6	3,2	3,8	2,4	3,5	0,7	3,5	1,9	1,2	2,2				
33	Moz_buff_SR10	1,3	1,3	1,3	1,3	1,3	1,5	1,3	1,3	1,3	1,3	1,3	3,0	3,0	3,0	3,0	2,9	3,0	3,2	3,0	3,2	0,9	3,5	3,2	3,6	2,7	3,3	0,7	3,3	2,2	1,2	2,1	0,9			
34	Moz_buff_SR11	1,9	1,9	1,9	1,9	1,9	2,1	1,9	1,9	1,9	1,9	1,9	1,9	2,3	1,9	1,9	2,3	1,9	2,1	1,9	2,1	1,5	1,8	1,9	2,2	2,1	1,9	1,9	1,9	1,0	2,4	0,4	2,4	2,2		
35	Moz_buff_SR12	0,3	0,3	0,3	0,3	0,3	0,4	0,3	0,3	0,3	0,3	0,3	3,3	3,3	3,3	3,3	3,0	3,3	3,5	3,3	3,5	0,4	3,8	3,3	4,0	2,6	3,6	0,9	3,6	1,5	1,3	1,8	0,7	1,3	1,9	

The samples enclosed in the red box represent sequences from cattle-derived parasites which were similar to the reference. The green box highlights the sample with sequences that had the highest percentage difference to the reference.

Table 4.3: *Pairwise distance alignment of the nucleotide sequences showing the base pair differences between each sequence and the reference, TP04_0028 (XM_758570.1), with the average number of base pair differences of 13 bp.*

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35
1	TP04_0028																																			
2	UG_cat_MA1-52	0																																		
3	UG_cat_MK2-12	0	0																																	
4	UG_cat_KLP5-4	0	0	0																																
5	UG_cat_NA1-45	0	0	0	0																															
6	UG_cat_NK1-25	1	1	1	1	1																														
7	Ken_cat_KC3	0	0	0	0	0	1																													
8	Ken_cat_KC4	0	0	0	0	0	1	0																												
9	Ken_cat_KC5	0	0	0	0	0	1	0	0																											
10	Ken_cat_KC9	0	0	0	0	0	1	0	0	0																										
11	Ken_cat_KC10	0	0	0	0	0	1	0	0	0	0																									
12	Ken_buff_BF6	22	22	22	22	22	23	22	22	22	22	22																								
13	Ken_buff_BF14	22	22	22	22	22	23	22	22	22	22	22	2																							
14	Ken_buff_BF15	22	22	22	22	22	23	22	22	22	22	22	0	2																						
15	Ken_buff_BF21	22	22	22	22	22	23	22	22	22	22	22	0	2	0																					
16	Ken_buff_BF24	20	20	20	20	20	21	20	20	20	20	20	2	2	2	2																				
17	Ken_buff_BF26	22	22	22	22	22	23	22	22	22	22	22	0	2	0	0	2																			
18	Ken_buff_BF28	23	23	23	23	23	24	23	23	23	23	23	1	3	1	1	3	1																		
19	Ken_buff_BF12	22	22	22	22	22	23	22	22	22	22	22	0	2	0	0	2	0	1																	
20	SA_cat_C71	21	21	21	21	21	22	21	21	21	21	21	5	7	5	5	7	5	6	5																
21	SA_cat_C81	3	3	3	3	3	4	3	3	3	3	3	19	19	19	19	17	19	20	19	20															
22	SA_cat_C84	25	25	25	25	25	26	25	25	25	25	25	7	9	7	7	9	7	8	7	4	22														
23	SA_cat_C89	22	22	22	22	22	23	22	22	22	22	22	3	5	3	3	3	3	4	3	8	19	8													
24	SA_cat_C91	26	26	26	26	26	27	26	26	26	26	26	4	6	4	4	6	4	5	4	5	23	3	5												
25	SA_cat_C108	17	17	17	17	17	18	17	17	17	17	17	5	5	5	5	3	5	6	5	6	16	8	6	9											
26	SA_cat_F369	22	22	22	22	22	23	22	22	22	22	22	6	8	6	6	8	6	7	6	1	21	3	7	4	7										
27	SA_buff_252	6	6	6	6	6	7	6	6	6	6	6	18	18	18	18	16	18	19	18	19	3	21	18	22	15	20									
28	SA_buff_258	22	22	22	22	22	23	22	22	22	22	22	6	8	6	6	8	6	7	6	1	21	3	7	4	7	0	20								
29	SA_buff_270	10	10	10	10	10	11	10	10	10	10	10	12	12	12	12	10	12	13	12	13	9	15	12	16	7	14	12	14							
30	Moz_buff_SR1	9	9	9	9	9	10	9	9	9	9	9	17	17	17	17	15	17	18	17	18	8	20	17	21	12	19	5	19	9						
31	Moz_buff_SR3	12	12	12	12	12	13	12	12	12	12	12	14	14	14	14	14	14	15	14	15	9	15	14	16	13	14	12	14	6	15					
32	Moz_buff_SR4	5	5	5	5	5	6	5	5	5	5	5	21	21	21	21	19	21	22	21	22	6	24	21	25	16	23	5	23	13	8	15				
33	Moz_buff_SR10	9	9	9	9	9	10	9	9	9	9	9	20	20	20	20	19	20	21	20	21	6	23	21	24	18	22	5	22	15	8	14	6			
34	Moz_buff_SR11	13	13	13	13	13	14	13	13	13	13	13	13	15	13	13	15	13	14	13	14	10	12	13	15	14	13	13	13	7	16	3	16	15		
35	Moz_buff_SR12	2	2	2	2	2	3	2	2	2	2	2	22	22	22	22	20	22	23	22	23	3	25	22	26	17	24	6	24	10	9	12	5	9	13	

The samples enclosed in the red box represent sequences from cattle-derived parasites which were similar to the reference. The green box highlights the sample with sequences that had the highest percentage difference to the reference.


TP04_0028
Uganda cattle MA1-52
Uganda cattle MK2-12
Uganda cattle KLP5-4
Uganda cattle NA1-45
Uganda cattle NK1-25
Kenya cattle KC3
Kenya cattle KC4
Kenya cattle KC5
Kenya cattle KC9
Kenya cattle KC10
Kenya buffalo BF6
Kenya buffalo BF12
Kenya buffalo BF14
Kenya buffalo BF15
Kenya buffalo BF21
Kenya buffalo BF24
Kenya buffalo BF26
Kenya buffalo BF28
SA cattle C71
SA cattle C81
SA cattle C84
SA cattle C89
SA cattle C91
SA cattle C108
SA cattle F369
SA buffalo 252
SA buffalo 258
SA buffalo 270
Moz buffalo SR1
Moz buffalo SR3
Moz buffalo SR4
Moz buffalo SR10
Moz buffalo SR11
Moz buffalo SR12


330 340 350 360 370 380 390 400 410 420 430 440 450 460 470 480
TP04_0028 TCACACGACTAGCTCCCAGCAATATATATACAGTTGATACTACTAACAAGTTTAAAATTTTGGGAACTGAGATTACTTTGAGAAACGAATCTCAAGTTACTGAATTATTTATAATAAAACTTTCCAACAAATTAGTTTTAGTTGAGATATGCTTAAGGCA
Uganda cattle MA1-52
Uganda cattle MK2-12
Uganda cattle KLP5-4
Uganda cattle NA1-45
Uganda cattle NK1-25
Kenya cattle KC3
Kenya cattle KC4
Kenya cattle KC5
Kenya cattle KC9
Kenya cattle KC10
Kenya buffalo BF6
Kenya buffalo BF12
Kenya buffalo BF14
Kenya buffalo BF15
Kenya buffalo BF21
Kenya buffalo BF24
Kenya buffalo BF26
Kenya buffalo BF28
SA cattle C71
SA cattle C81
SA cattle C84
SA cattle C89
SA cattle C91
SA cattle C108
SA cattle F369
SA buffalo 252
SA buffalo 258
SA buffalo 270
Moz buffalo SR1
Moz buffalo SR3
Moz buffalo SR4
Moz buffalo SR10
Moz buffalo SR11
Moz buffalo SR12
490 500 510 520 530 540 550 560 570 580 590 600 610 620 630 640
TP04_0028 TGAAGGTGTAAGGTATAAAAATGCTTATTATAAAAGAACTTTAAATGAAGGGAGATTTACAGGTACTAACAAAAGTGAGTTATTTGACGCCTTATTTAATGATTTCAGTAAGAATGAAAACAAGGTAACATTAAAGTATACATCCGTGTTCCCTATAATC
Uganda cattle MA1-52
Uganda cattle MK2-12
Uganda cattle KLP5-4
Uganda cattle NA1-45
Uganda cattle NK1-25
Kenya cattle KC3
Kenya cattle KC4
Kenya cattle KC5
Kenya cattle KC9
Kenya cattle KC10
Kenya buffalo BF6
Kenya buffalo BF12
Kenya buffalo BF14
Kenya buffalo BF15
Kenya buffalo BF21
Kenya buffalo BF24
Kenya buffalo BF26
Kenya buffalo BF28
SA cattle C71
SA cattle C81
SA cattle C84
SA cattle C89
SA cattle C91
SA cattle C108
SA cattle F369
SA buffalo 252
SA buffalo 258
SA buffalo 270
Moz buffalo SR1
Moz buffalo SR3
Moz buffalo SR4
Moz buffalo SR10
Moz buffalo SR11
Moz buffalo SR12

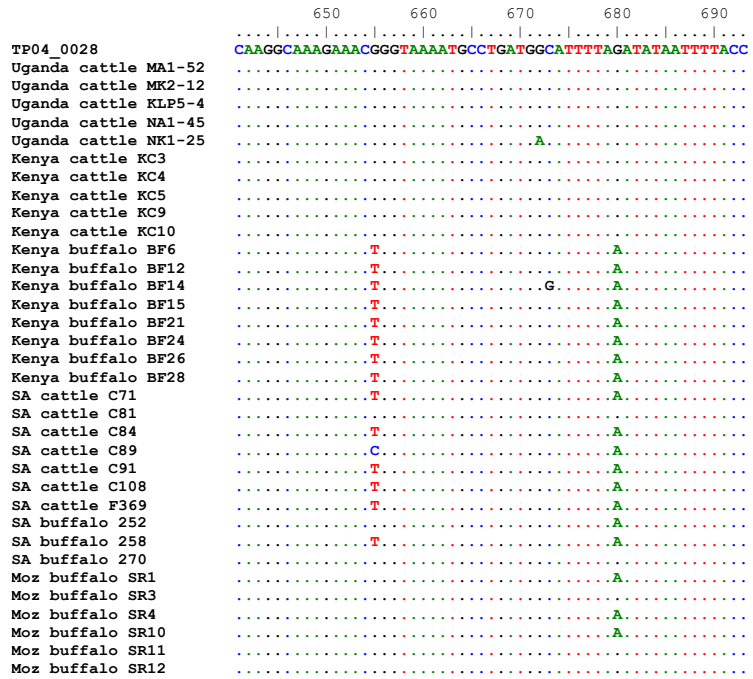

Figure 4.4: *Multiple nucleotide sequence alignment of TP04_0028 gene generated using BioEdit, consisting of 34 sequences from cattle- and buffalo-derived* T. parva *isolates. The DNA samples were prepared from cattle and buffalo blood from Kenya, Uganda, Mozambique and South Africa. The blue box represents the site with the highest number of insertions and the green box represents the site with the highest number of nucleotide substitutions to the reference.*

4.3.2 Amino acid sequence data analysis

Sequence variations appeared to affect nine of the 20 predicted epitopes occurring within the targeted gene region (Figure 4.5). There are three areas where the predicted epitopes are conserved throughout the sequence alignment; amino acid positions 4-15, 129-145, 150-171 and 218-227. The conserved positions 129-145, consists of seven overlapping epitopes (Figure 4.6) making this region one with the most epitopes. Notably, one of the amino acid mutations was detected within this epitope region. Only two of the 20 predicted epitopes within the TP04_0028 amplicon are not conserved amongst different BoLA alleles and show amino acid mutations within epitope groups at positions 21-31 and 34-43 (indicated with black arrows in Figure 4.5).

Consistent with observations from nucleotide sequence analysis, all amino acid sequences from cattle samples from Kenya and Uganda were identical, with 100% similarity to TP04_0028 (XM_758570.1) reference sequence (Table 4.4). The overall average sequence variation for the percentage pairwise distance analysis (3%) was also comparable to that identified from analysis of nucleotide sequences (2%). Notably, some variations identified from nucleotide sequence analysis resulted in differences in the protein sequence as well; hence, the average number of amino acids affected were 7, compared to 13 from the nucleotide sequence analysis. Nucleotide sequence differences (between 238 and 245

bp) that translated into amino acid variations included the unique variations common to sequences from buffalo-derived *T. parva* parasites from cattle from SA and buffalo from SA and Mozambique, at amino acid positions 83 and 84. Also consistent with the nucleotide sequence variations (at position 613 and 615), some sequences from SA cattle and buffalo had unique amino acid mutations at position 207.

The sequence from the Mozambique buffalo sample SR10 was excluded from the amino acid analysis as it could not properly align with the rest of TP04_0028 sequences. Every effort to detect the problem with sequence translation or alignment proved futile.

Table 4.4: *Pairwise distance alignment done in MEGA of the amino acid sequences representing the percentage sequence differences between each sequence and the reference (XM_758570.1). The average percentage difference was determined to be 3%.*

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34
1	TP04_0028																																		
2	UG_cat_MA1-52	0,0																																	
3	UG_cat_MK2-12	0,0	0,0																																
4	UG_cat_KLP5-4	0,0	0,0	0,0																															
5	UG_cat_NA1-45	0,0	0,0	0,0	0,0																														
6	UG_cat_NK1-25	0,4	0,4	0,4	0,4	0,4																													
7	Ken_cat_KC3	0,0	0,0	0,0	0,0	0,0	0,4																												
8	Ken_cat_KC4	0,0	0,0	0,0	0,0	0,0	0,4	0,0																											
9	Ken_cat_KC5	0,0	0,0	0,0	0,0	0,0	0,4	0,0	0,0																										
10	Ken_cat_KC9	0,0	0,0	0,0	0,0	0,0	0,4	0,0	0,0	0,0																									
11	Ken_cat_KC10	0,0	0,0	0,0	0,0	0,0	0,4	0,0	0,0	0,0	0,0																								
12	Ken_buff_BF6	6,4	6,4	6,4	6,4	6,4	6,9	6,4	6,4	6,4	6,4	6,4																							
13	Ken_buff_BF12	6,4	6,4	6,4	6,4	6,4	6,9	6,4	6,4	6,4	6,4	6,4	0,0																						
14	Ken_buff_BF14	6,0	6,0	6,0	6,0	6,0	6,4	6,0	6,0	6,0	6,0	6,0	0,4	0,4																					
15	Ken_buff_BF15	6,4	6,4	6,4	6,4	6,4	6,9	6,4	6,4	6,4	6,4	6,4	0,0	0,0	0,4																				
16	Ken_buff_BF21	6,4	6,4	6,4	6,4	6,4	6,9	6,4	6,4	6,4	6,4	6,4	0,0	0,0	0,4	0,0																			
17	Ken_buff_BF24	5,5	5,5	5,5	5,5	5,5	6,0	5,5	5,5	5,5	5,5	5,5	0,9	0,9	0,4	0,9	0,9																		
18	Ken_buff_BF26	6,4	6,4	6,4	6,4	6,4	6,9	6,4	6,4	6,4	6,4	6,4	0,0	0,0	0,4	0,0	0,0	0,9																	
19	Ken_buff_BF28	6,4	6,4	6,4	6,4	6,4	6,9	6,4	6,4	6,4	6,4	6,4	0,0	0,0	0,4	0,0	0,0	0,9	0,0																
20	SA_cat_C71	6,0	6,0	6,0	6,0	6,0	6,4	6,0	6,0	6,0	6,0	6,0	1,3	1,3	1,8	1,3	1,3	2,2	1,3	1,3															
21	SA_cat_C81	0,4	0,4	0,4	0,4	0,4	0,9	0,4	0,4	0,4	0,4	0,4	6,0	6,0	5,5	6,0	6,0	5,0	6,0	6,0	5,5														
22	SA_cat_C84	6,0	6,0	6,0	6,0	6,0	6,4	6,0	6,0	6,0	6,0	6,0	2,2	2,2	1,8	2,2	2,2	2,2	2,2	2,2	0,9	5,5													
23	SA_cat_C89	5,5	5,5	5,5	5,5	5,5	6,0	5,5	5,5	5,5	5,5	5,5	0,9	0,9	0,4	0,9	0,9	0,0	0,9	0,9	2,2	5,0	2,2												
24	SA_cat_C91	6,9	6,9	6,9	6,9	6,9	7,4	6,9	6,9	6,9	6,9	6,9	1,3	1,3	0,9	1,3	1,3	1,3	1,3	1,3	1,8	6,4	0,9	1,3											
25	SA_cat_C108	4,5	4,5	4,5	4,5	4,5	5,0	4,5	4,5	4,5	4,5	4,5	1,8	1,8	1,3	1,8	1,8	0,9	1,8	1,8	1,3	4,1	1,3	0,9	2,2										
26	SA_cat_F369	4,5	4,5	4,5	4,5	4,5	5,0	4,5	4,5	4,5	4,5	4,5	1,8	1,8	1,3	1,8	1,8	0,9	1,8	1,8	1,3	4,1	1,3	0,9	2,2	0,0									
27	SA_buff_252	1,3	1,3	1,3	1,3	1,3	1,8	1,3	1,3	1,3	1,3	1,3	5,0	5,0	4,5	5,0	5,0	4,1	5,0	5,0	4,5	0,9	4,5	4,1	5,5	3,2	3,2								
28	SA_buff_258	5,5	5,5	5,5	5,5	5,5	6,0	5,5	5,5	5,5	5,5	5,5	1,8	1,8	1,3	1,8	1,8	1,8	1,8	1,8	0,4	5,0	0,4	1,8	1,3	0,9	0,9	4,1							
29	SA_buff_270	2,2	2,2	2,2	2,2	2,2	2,7	2,2	2,2	2,2	2,2	2,2	4,1	4,1	3,6	4,1	4,1	3,2	4,1	4,1	3,6	1,8	3,6	3,2	4,5	2,2	2,2	2,7	3,2						
30	Moz_buff_SR1	2,2	2,2	2,2	2,2	2,2	2,7	2,2	2,2	2,2	2,2	2,2	5,0	5,0	4,5	5,0	5,0	4,1	5,0	5,0	4,5	1,8	4,5	4,1	5,5	3,2	3,2	0,9	4,1	2,7					
31	Moz_buff_SR3	2,2	2,2	2,2	2,2	2,2	2,7	2,2	2,2	2,2	2,2	2,2	4,1	4,1	3,6	4,1	4,1	4,1	4,1	4,1	3,6	1,8	4,0	4,5	5,0	3,2	3,2	2,7	3,6	0,9	3,6				
32	Moz_buff_SR4	1,3	1,3	1,3	1,3	1,3	1,8	1,3	1,3	1,3	1,3	1,3	6,0	6,0	5,5	6,0	6,0	5,0	6,0	6,0	5,5	1,8	5,5	5,0	6,4	4,1	4,1	0,9	5,0	3,6	1,8	3,6			
33	Moz_buff_SR11	2,2	2,2	2,2	2,2	2,2	2,7	2,2	2,2	2,2	2,2	2,2	4,1	4,1	3,6	4,1	4,1	4,1	4,1	4,1	3,6	1,8	3,6	4,0	4,5	3,2	3,2	2,7	3,1	0,9	3,6	0,4	3,6		
34	Moz_buff_SR12	0,4	0,4	0,4	0,4	0,4	0,9	0,4	0,4	0,4	0,4	0,4	6,9	6,9	6,4	6,9	6,9	6,0	6,9	6,9	6,4	0,9	6,4	6,0	7,4	5,0	5,0	1,8	6,0	2,7	2,7	2,7	1,8	2,7	

The samples enclosed in the red box represent sequences from cattle-derived parasites which were similar to the reference. The green box highlights the sample with sequences that had the highest percentage difference to the reference.

Table 4.5: *Pairwise distance alignment done in MEGA of the amino acid sequences representing the number of position differences between each sequence and the reference (XM_758570.1). The average number of position differences is 7.*

		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34
1	TP04_0028																																		
2	UG_cat_MA1-52	0																																	
3	UG_cat_MK2-12	0	0																																
4	UG_cat_KLP5-4	0	0	0																															
5	UG_cat_NA1-45	0	0	0	0																														
6	UG_cat_NK1-25	1	1	1	1	1																													
7	Ken_cat_KC3	0	0	0	0	0	1																												
8	Ken_cat_KC4	0	0	0	0	0	1	0																											
9	Ken_cat_KC5	0	0	0	0	0	1	0	0																										
10	Ken_cat_KC9	0	0	0	0	0	1	0	0	0																									
11	Ken_cat_KC10	0	0	0	0	0	1	0	0	0	0																								
12	Ken_buff_BF6	14	14	14	14	14	15	14	14	14	14	14																							
13	Ken_buff_BF12	14	14	14	14	14	15	14	14	14	14	14	0																						
14	Ken_buff_BF14	13	13	13	13	13	14	13	13	13	13	13	1	1																					
15	Ken_buff_BF15	14	14	14	14	14	15	14	14	14	14	14	0	0	1																				
16	Ken_buff_BF21	14	14	14	14	14	15	14	14	14	14	14	0	0	1	0																			
17	Ken_buff_BF24	12	12	12	12	12	13	12	12	12	12	12	2	2	1	2	2																		
18	Ken_buff_BF26	14	14	14	14	14	15	14	14	14	14	14	0	0	1	0	0	2																	
19	Ken_buff_BF28	14	14	14	14	14	15	14	14	14	14	14	0	0	1	0	0	2	0																
20	SA_cat_C71	13	13	13	13	13	14	13	13	13	13	13	3	3	4	3	3	5	3	3															
21	SA_cat_C81	1	1	1	1	1	2	1	1	1	1	1	13	13	12	13	13	11	13	13	12														
22	SA_cat_C84	13	13	13	13	13	14	13	13	13	13	13	5	5	4	5	5	5	5	5	2	12													
23	SA_cat_C89	12	12	12	12	12	13	12	12	12	12	12	2	2	1	2	2	0	2	2	5	11	5												
24	SA_cat_C91	15	15	15	15	15	16	15	15	15	15	15	3	3	2	3	3	3	3	3	4	14	2	3											
25	SA_cat_C108	10	10	10	10	10	11	10	10	10	10	10	4	4	3	4	4	2	4	4	3	9	3	2	5										
26	SA_cat_F369	10	10	10	10	10	11	10	10	10	10	10	4	4	3	4	4	2	4	4	3	9	3	2	5	0									
27	SA_buff_252	3	3	3	3	3	4	3	3	3	3	3	11	11	10	11	11	9	11	11	10	2	10	9	12	7	7								
28	SA_buff_258	12	12	12	12	12	13	12	12	12	12	12	4	4	3	4	4	4	4	4	1	11	1	4	3	2	2	9							
29	SA_buff_270	5	5	5	5	5	6	5	5	5	5	5	9	9	8	9	9	7	9	9	8	4	8	7	10	5	5	6	7						
30	Moz_buff_SR1	5	5	5	5	5	6	5	5	5	5	5	11	11	10	11	11	9	11	11	10	4	10	9	12	7	7	2	9	6					
31	Moz_buff_SR3	5	5	5	5	5	6	5	5	5	5	5	9	9	8	9	9	9	9	9	8	4	9	10	11	7	7	6	8	2	8				
32	Moz_buff_SR4	3	3	3	3	3	4	3	3	3	3	3	13	13	12	13	13	11	13	13	12	4	12	11	14	9	9	2	11	8	4	8			
33	Moz_buff_SR11	5	5	5	5	5	6	5	5	5	5	5	9	9	8	9	9	9	9	9	8	4	8	9	10	7	7	6	7	2	8	1	8		
34	Moz_buff_SR12	1	1	1	1	1	2	1	1	1	1	1	15	15	14	15	15	13	15	15	14	2	14	13	16	11	11	4	13	6	6	6	4	6	

The samples enclosed in the red box represent sequences from cattle-derived parasites which were similar to the reference. The green box highlights the sample with sequences that had the highest percentage difference to the reference.

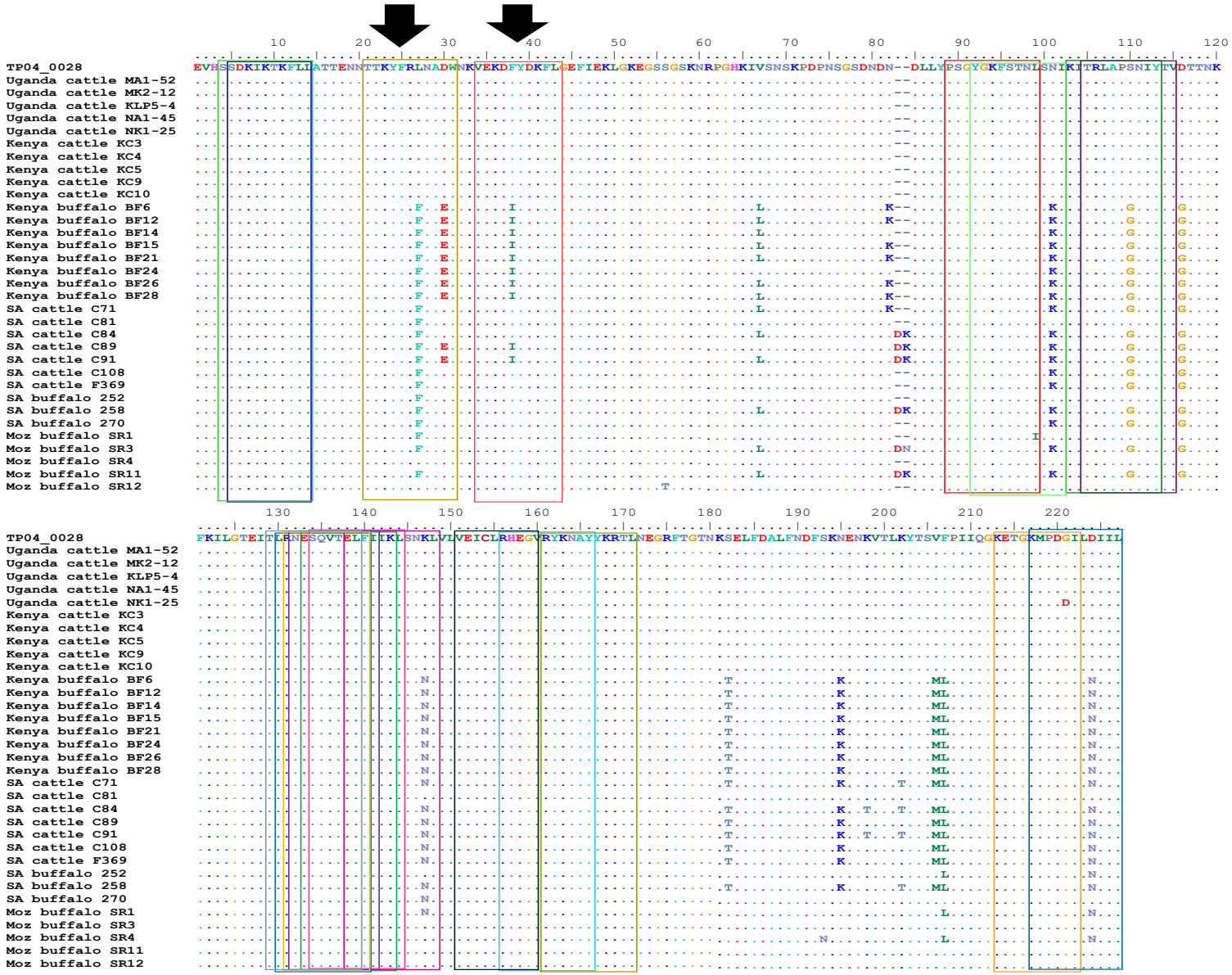

Figure 4.5: *Amino acid multiple sequence alignment of 34 cattle and buffalo sequences generated using ClustalX, in BioEdit. The coloured vertical rectangles represent different predicted epitopes (from IEDB analysis) of TP04_0028. The blak arrows indicate the two predicted epitopes (positions 21-31 and 34-44) that are not conserved amongst BoLA alleles.*

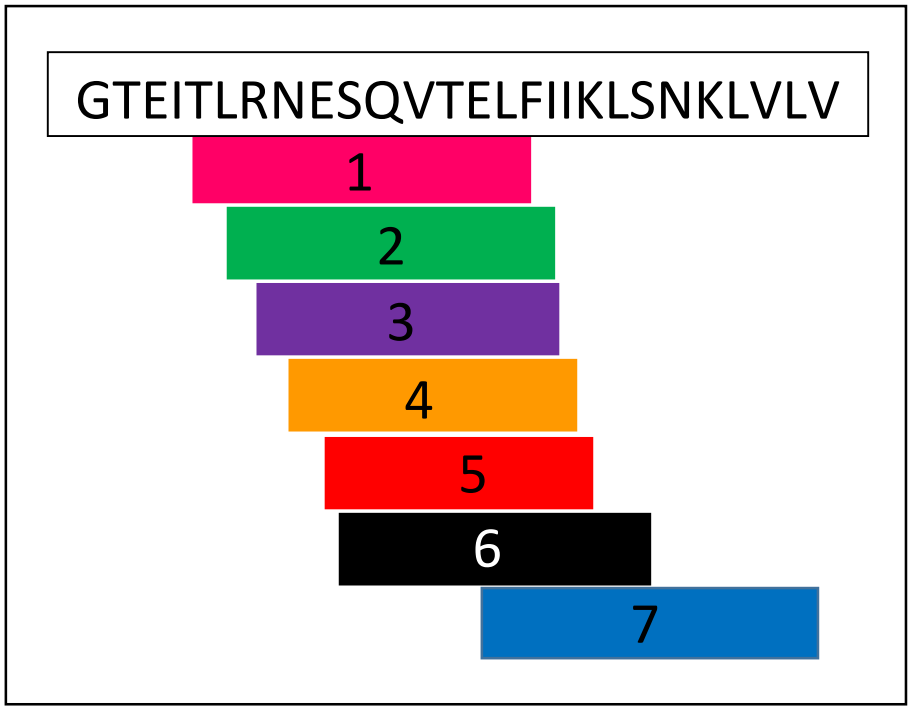

Figure 4.6: *The position of seven overlapping predicted epitopes, spanning position 129 to 145 in the protein sequence alignment shown in Figure 4.5.*

4.4 *Phylogenetic analysis*

The Neighbor-Joining and Maximum likelihood phylogenetic trees were constructed using the TP04_0028 nucleotide and protein sequences, to assess the evolutionary relationship between sequences from cattle- and buffalo-derived *T. parva* isolates from South Africa, Kenya, Uganda and Mozambique (Figure 4.7 - 4.8). Two phylogenetic analyses were performed to ensure that the evolutionary relationship observed was accurate and a true representation of the relationship between cattle- and buffalo-derived samples.

Both the Neighbor-Joining and Maximum likelihood phylogenetic trees produced the same topology, revealing two major monophyletic groups, labelled Groups 1 and 2 (Figures 4.7 - 4.8). Group 1 consists of paraphyletic TP04_0028 sequence groups from both cattle- and buffalo-derived *T. parva* isolates, with bootstrap values ranging from 61 to 97 for the nucleotide sequence phylogenetic analyses (Figures 4.7 A and 4.8 A) respectively and 57 to 71 for amino acid sequence analyses (Figures 4.7 B and 4.8 B). While, sequences in Group 2 were all from buffalo-derived *T. parva*, with a large monophyletic sub-group supported by bootstrap values of 88 and 73 for nucleotide sequence phylogenetic analyses (Figures 4.7 A and 4.8 A) respectively and 79 and 69 for amino acid sequence analyses (Figures 4.7 B and 4.8 B). Although Group 1 consists of sequences from both parasite types, sequences from cattle-derived *T. parva* grouped together and appeared to be derived from a recent speciation event.

Notably, all sequences from buffalo samples from Mozambique were found in Group 1 and those from buffalo from Kenya grouped together in Group 2. Whereas, sequences from cattle and buffalo from South Africa were found in both monophyletic groups. The reference TP04_0028 (XM_758570.1) sequence grouped with sequences from cattle-derived *T. parva* isolates Kenya and Uganda, in both nucleotide and amino acid phylogenetic trees from the two analyses.

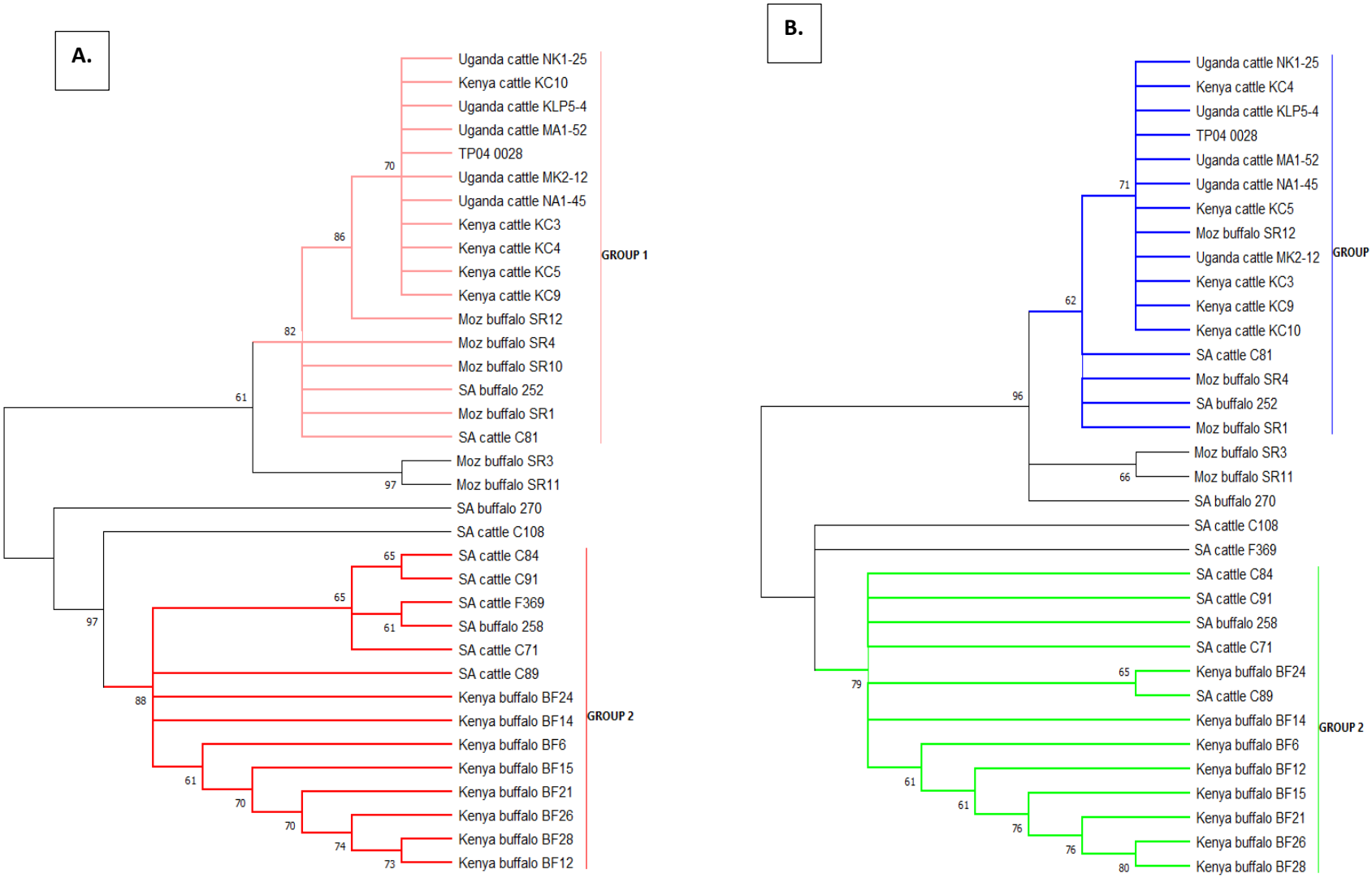

Figure 4.7: *Neighbor-Joining phylogenetic tree based on the TP04_0028* ***nucleotide (A) and amino acid (B)*** *sequences, showing two major monophyletic groups, Group 1 representing sequences from both cattle- and buffalo-derived* T. parva *isolates and Group 2 sequences from buffalo-derived isolates.*

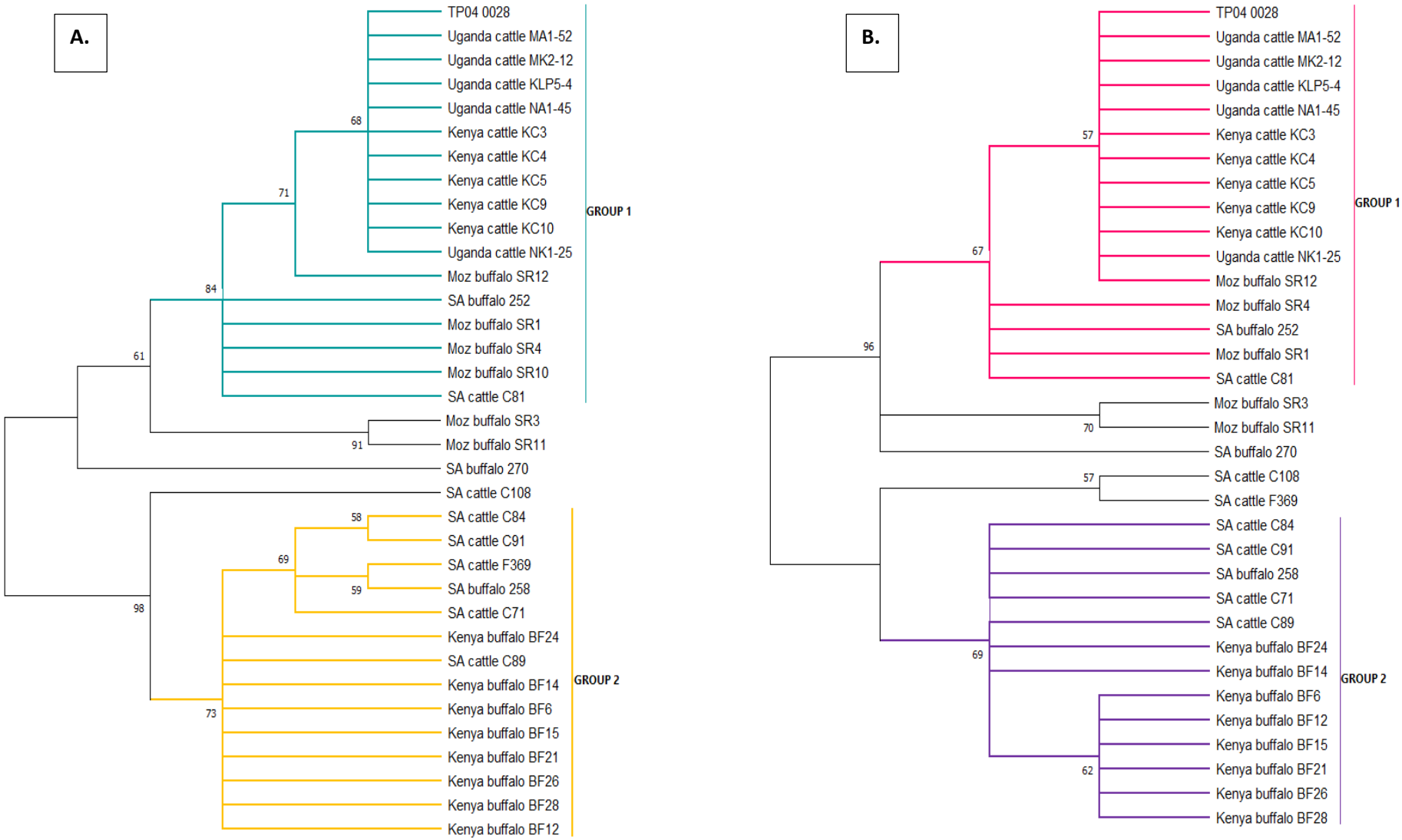

Figure 4.8: *Maximum-likelihood phylogenetic tree based on the TP04_0028* ***nucleotide (A) and amino acid (B)*** *sequences, showing two major monophyletic groups, Group 1 representing sequences from both cattle- and buffalo-derived* T. parva *isolates and Group 2 sequences from buffalo-derived isolates.*

CHAPTER 5

5. DISCUSSION AND CONCLUSION

The genetic and antigenic diversity in *T. parva* field isolates has been one of the major challenges in the development of effective control strategies against bovine theileriosis (Ferraro et al., 2011). Moreover, the ITM currently used for immunization against *T. parva* infections is plagued with numerous limitations (McKeever and Morrison, 1994, Perry, 2016, Sitt et al., 2015). Thus, recombinant/subunit vaccines have become a research trend and the search to find suitable vaccine candidates that can be used to immunise against both buffalo- and cattle- derived *T. parva* infections remains a challenge. The gene encoding the hypothetical protein, TP04_0028, was selected among 13 vaccine candidates identified by *in silico* prediction, to determine its genetic and antigenic diversity by sequence analysis, as part of the process of evaluation of a cohort of candidate genes as suitable antigens for development of a recombinant vaccine for control of *T. parva* infections. The design of recombinant vaccines included identifying candidates with cell adhesion and protein cellular localisation properties, transmembrane helixes, homology and determining their antigenic properties in order to facilitate an immune response (Monterrubio-Lopez et al., 2015, María et al., 2017). The *in silico* analysis revealed that TP04_0028 encodes a secreted hypothetical protein, with a GPI-anchor and possesses transmembrane domains. The *in silico* analysis showed TP04_0028 to possess antigenic properties with MHC Class-I binding affinity, suggesting that the candidate could produce an immune response against *T. parva* infection. The analysis of *T. parva* transcriptome data also showed that TP04_0028 transcript is expressed in the pathogenic stage of the parasite, the schizont.

The schizont stage of *T. parva* infects the mammalian host lymphocytes which causes the reversible cancer-like phenotype (De Goeyse et al., 2015). It has been demonstrated that the schizont-infected cells are destroyed by the MHC class I-restricted $CD8^+$ cytotoxic T lymphocytes (CTL) (Morrison et al., 1987, McKeever and Morrison, 1994). Since immunity against *T. parva* infection is triggered by schizont-infected cells it is logical to target antigens expressed in the schizont life-cycle stage in the search of candidates for development of the recombinant vaccine against *T. parva* infections. Initial transcriptome studies have reported that 60% of *T. parva* genes are expressed in the parasitic schizont stage (Bishop et al., 2005, Gardner et al., 2005), although the recent re-annotation of the *T. parva* genome revealed that 98% of *T. parva* protein-encoding genes are expressed in this parasitic developmental stage (Tretina et al., 2019).

Thus, six out of the thirteen predicted proteins identified from a related *in silico* study (unpublished data) met all the criteria for vaccine candidates, however, five of these were not expressed at

comparable levels between the two *T. parva* isolates, Muguga and 7014, respectively representing cattle- and buffalo-derived parasites, resulting in the selection of TP02_0028 for the purpose of this study. Furthermore, TP02_0028 has a longer encoding gene length with 44 predicted epitope regions, which may offer a high possibility for antigenic capability and consequently, immunity, should it contain immunogenic epitopes. Epitopes are regions in a protein, made up of short amino acid sequences, responsible for inducing immune response (Kao and Hodges, 2009). Notably, TP02_0028 expression levels are low when compared to TP02_0553 and TP02_0950; however, according to Bishop et al., 2005, genes expressed at low transcriptome levels can still be expressed and recognised by the host's immune system. As indicated earlier, TP04_0028 encompasses 44 predicted epitopes and the primer set chosen for amplification of the target region includes 45% of these (20 of the 44).

Several studies on the genetic diversity studies of *T. parva* antigens such as Tp1, Tp2 and Tp9 (MacHugh et al., 2009, Pelle et al., 2011, Sitt et al., 2018) showed varying degrees of polymorphism. The current study investigated the diversity of the putative antigen encoding TP04_0028, in samples collected from various geographical locations, for characterization of the target region in buffalo- and cattle-derived *T. parva* parasites. Ideally, a vaccine candidate should confer protection against all field strains of *T. parva* and should be antigenically conserved in different strains. Conserved antigens also help avoid or reduce immune evasion, parasites are known to use polymorphism to escape the host immune response (Ouattara et al., 2015).

The first step in the genotyping of TP04_0028 was specific amplification of the target region, encompassing 20 predicted epitopes. Optimal PCR conditions for specific amplification of TP04_0028 was achieved through several steps of troubleshooting. The initial amplification using the standard PCR reaction recommended for 20 µl PCR mix resulted in a combination of results raging from no amplicon to multiple bands of non-specific PCR products obtained together with the expected amplicon. For troubleshooting of results that showed no amplicon, the manufacturer of the Phusion Flash High Fidelity PCR master-mix (Thermo Fisher Scientific Inc.) recommends optimisation of the annealing temperature, increasing the number of cycles in the amplification conditions or increasing the extension time based on the expected amplicon size. Where non-specific products are obtained they recommend shortening the extension time, decreasing primer concentration, performing a temperature gradient PCR and/or increasing the annealing temperature. The troubleshooting criteria given by the manufacturer are in line with recommendations made in literature (Lorenz, 2012, Zhang et al., 2019). In the current study, different annealing temperatures (including using a temperature gradient) and primer concentrations were explored. The extension time was altered to improve both 'no amplicon' and 'non-specific products' reactions; however, the change in extension time did not

change the results originally observed. The specific amplification of the target region was subsequently achieved with the primer annealing temperature at 65 °C and concentration of 0.5 µM each primer. However, even at optimised conditions, most of the samples investigated, especially from buffalo, showed faint bands from gel electrophoresis analysis and with some showing no visible amplicon or primer dimers. Where no band or primer dimers were visualised, it was assumed that amplification had taken place but the amplicon produced at low quantities invisible on the gel, possibly due to low infection levels in affected samples. Primer dimers are normally absent when the PCR primers are of good design and have been used up during a reaction (Hengen, 1995, Brownie et al., 1997, Canene-Adams, 2013). Thus it was necessary to re-amplify the primary PCR product to obtain reasonable quantities for downstream analysis. Furthermore, field samples, especially from buffalo, are usually infected with multiple *T. parva* strains (Sibeko et al., 2010, Tayebwa et al., 2018) therefore cloning was essential to ensure sequencing of all products, from various strains.

Sequence analysis was performed using programmes and methods that have been applied successfully previously (Habibi, 2016). It has been reported that extensive polymorphism at nucleotide level does not always translate to extensive polymorphism at amino acid level (Sitt et al., 2018). Consistently in this study, limited changes were detected in amino acid sequences compared to nucleotide sequences. Nonetheless, both nucleotide and amino acid pairwise distance analyses showed that TP04_0028 sequences from cattle samples from Kenya and Uganda had the highest percent similarity (99 - 100%) to the reference (XM_758570.1). According to the p67 characterization (Mukolwe et al., 2020), these samples were infected with the cattle-derived *T. parva,* the causative agent of ECF. The high percent sequence similarity to the reference can be attributed to the fact that the TP04_0028 reference sequence is derived from the Muguga stock, a cattle-derived *T. parva* stock from Kenya, known to cause ECF. These sequences were also conserved when compared to each other. Consistently, the phylogenetic analysis showed a close relationship of sequences from these samples, as sequences from Kenya and Uganda cattle samples grouped with TP04_0028 reference in the same sub-group. This close relationship has been previously reported in genome studies where cattle-derived sequences from Kenya (Muguga) and Uganda (Entebbe) grouped together (Hayashida et al., 2013).

On the contrary, variations (93,1 - 99,7%) were identified from TP04_0028 sequences from cattle samples from South Africa and buffalo samples from Kenya, South Africa and Mozambique, when compared to the reference. These samples were infected with buffalo-derived *T. parva,* based on the p67 sequence analysis (Mukolwe et al., 2020). It is known that antigens in buffalo-derived *T. parva* are generally more polymorphic compared to cattle-derived isolates (Pelle et al., 2011), due to the

heterogeneous population of *T. parva* strains in the buffalo. In the characterization of Tp1 and Tp2 antigens, Hemmink et al., (2018) also showed extreme allelic diversity in buffalo-derived *T. parva* parasites, at population level and between individual buffalo samples. Consistently in the current study, the sequence variations detected within buffalo-derived parasites was also supported by sub-groups observed from phylogenetic analysis. Since it has been accepted that all *T. parva* parasites originated from buffalo, it has been speculated that *T. parva* strains that are classically buffalo-derived are parasites which have not yet evolved into the more conserved cattle-derived populations such as those of Eastern Africa (Sitt et al., 2015), which explains the two very distinct buffalo-derived and cattle-derived *T. parva* phylogenetic groups observed in this study.

Sequences from buffalo-derived *T. parva* samples from South Africa (cattle and buffalo samples) and Kenya (buffalo samples) were the most diverse compared to the reference, and accordingly similar to each other. In agreement, the phylogenetic analysis of these sequences showed a close relationship as they grouped together in Group 2, although in different subgroups which largely appear to be geographic-based. This result suggests that the buffalo-derived parasites from East and South Africa shared a common ancestor. According to Hemmink et al., (2018), the observed similarity between buffalo-derived *T. parva* isolates from these two regions could be due to conservation in these isolates before geographical movement of the Cape buffalo occurred in Africa. Unfortunately, samples from Ugandan buffalo-derived *T. parva* were not available for our study. It would be interesting to see if their TP04_0028 sequences would group with sequences from Kenyan buffalo consistent with the phylogenetic relationship of the cattle-derived parasites from these two countries or if they would group with other buffalo-derived sequences in Group 2. It has been reported that Cape buffalo in Uganda represented a high multiplicity of infection, expressing multiple *T. parva* genotypes (Oura et al., 2011).

Notably, sequences from buffalo samples from Mozambique had the highest percent similarity to the reference, compared to sequences from other buffalo-derived *T. parva* samples. In fact, the phylogenetic analysis showed that they grouped together with sequences from cattle-derived parasites from Kenya and Uganda, although in different sub-groups. This heterogeneous phylogenetic group further supports reports that *T. parva* was originally a buffalo parasite and isolates circulating in cattle originated from buffalo; hence this close relationship is also observed through TP04_0028 genotypes. Interestingly, according to the sequence analysis of the p67 antigen gene buffalo-derived isolates in Group 1 have different alleles, including all four *T. parva* p67 alleles previously characterized by Mukolwe et al., (2020). The four p67 alleles represent various strains of *T. parva* in both cattle- and buffalo-derived parasite groups and the close phylogenetic relationship of these isolates in Group 1

may suggest that this specific genotype of TP04_0028 could be a suitable subunit vaccine candidate which can confer protection against a broad group of *T. parva* parasites. In addition to sequences from buffalo-derived isolates from Mozambique, two TP04_0028 sequences from samples from South Africa also showed a close phylogenetic relationship with cattle-derived *T. parva* sequences in Group 1, further suggesting that TP04_0028 is conserved between *T. parva* isolates from different geographical locations. Thus, this putative antigen may have the potential to protect even against from infections by strains across various geographic origin. However, the individual predicted genotypes of TP04_0028 identified in the current study would need to be characterised further using a variety of immune assays to determine if they can induce T-cell response and confer broad protection in the field. Although CD8 T-cell epitope prediction methods are usually accurate, it is reported that 90% of predicted epitopes may not be immunogenic (Zhong et al., 2003).

The stringent criterion of the pairwise distance analysis of < 10% was applied for characterization of a highly conserved vaccine target. Consistently, only sequences with > 10% differences to the reference were thus considered to be variant. Although some sequence variations were detected, especially from buffalo-derived *T. parva* samples when compared to TP04_0028, the overall analysis of TP04_0028 sequences produced in this study showed < 5% difference against the reference, suggesting reasonable conservation between cattle- and buffalo-derived *T. parva* parasite sequences. The aim of this study was to confirm if TP04_0028 was conserved between cattle- and buffalo-derived *T.* parva isolates for consideration as a possible subunit vaccine candidate. Thus, the conserved TP04_0028 may have the ability to provide immune protection against both cattle- and buffalo-derived infections, should it be found to be immunogenic. This would be a major improvement in immunisation against *T. parva* infections considering that the current widely used ITM immunization is not able to confer broad protection in the field, especially against buffalo-derived parasites (Sitt et al., 2015) and the extensively researched p67 subunit vaccine has shown incomplete protection in field studies (Musoke et al., 2005, Sitt et al., 2015).

Notably, 11 out of the 20 predicted epitopes (55%) within the amplified target region of TP04_0028 gene were conserved with no sequence variations, when compared to the reference; while the rest (n = 8) had single amino acid mutations, except for one epitope region which had two mutations. Furthermore, the majority (n = 18) of the 20 predicted epitopes are conserved between different BoLA alleles. Also notable, is that among the predicted epitope regions there is an area with seven overlapping epitopes. This area could be targeted in the initial search for regions with immunogenic efficiency should TP04_0028 be further characterised as a vaccine candidate. Overlapping epitopes increase the antigenic structure for MHC recognition and binding during immune response (Mateu et

al., 1990, Lohia and Baranwal, 2014). Thus, their conservation could be a good quality for a vaccine candidate that can provide broad protection against cattle- and buffalo-derived *T. parva* infections. It has been reported that some specific $CD8^+$ T cell clones can tolerate up to three amino acid exchanges within the epitope sequence, while others fail to recognize a variant epitope with a single substitution (MacHugh et al., 2011). This would need to be validated amongst the predicted epitopes observed in this study as some predicted epitopes have between 1-2 amino acid substitutions when compared to TP04_0028 reference. Antigenic polymorphism has been shown to result in differential recognition by $CD8^+$ T-cell clones where some clones varied in their capacity to recognize specific alleles (MacHugh et al., 2011) leading to incomplete immune protection. On the contrary, the analysis of the malaria antigen, merozoite surface protein-2 demonstrated that extensive diversity in DNA sequences does not necessarily translate to antigenic polymorphism, at least for this antigen (Franks et al., 2003). Thus, the limited variations in TP04_0028 do not exclude its gene product from being a contender for vaccine candidate.

CONCLUSION

The aim of this study was to determine the sequence diversity of TP04_0028 in both cattle- and buffalo-derived *T. parva* isolates. The results of this study showed that the putative antigen TP04_0028 sequences are highly conserved among cattle-derived *T. parva* parasites investigated in this study and has limited variations among buffalo-derived parasites, irrespective of geographic origin. Overall, amongst both nucleotide and amino acid sequences, the pairwise distances were < 5% compared to the reference sequence. Conserved antigens are the best choice when searching for possible candidates for subunit vaccines (Sitt et al., 2018). A close phylogenetic relationship of sequences from buffalo-derived samples from Mozambique and South Africa with those from cattle-derived samples from Kenya and Uganda was also revealed, further support for the suggestion that this putative antigen may have potential to confer protection against both cattle- and buffalo-derived *T. parva* infections. The ability of a vaccine to cross-protect remains crucial when searching for vaccine candidates. Cattle immunised with a cattle-derived *T. parva* vaccine show incomplete protection but those immunised with buffalo-derived parasites are protected in part against cattle-derived *T. parva* challenge (Morrison et al., 2020). This could be due to a pre-dated evolutionary relationship between buffalo-derived and cattle-derived *T. parva* infections.

6. LIMITATIONS AND RECOMMENDATIONS FOR FUTURE STUDIES

A major limitation to this study was not designing overlapping primers that cover the entire gene and thus all the predicted epitopes. Perhaps the overlapping primers would have influenced the outcome

of this study showing more diversity in different areas of the gene or perhaps even extreme polymorphism in some areas. It would have also increased the probability for identification of immunogenic epitopes since only 10% of predicted epitopes are found to be immunogenic. However due to time constraints, it was not possible to characterize the entire gene. It would also be interesting to include known January disease samples in future studies to determine how TP04_0028 sequences from these isolates compare to those detected from ECF and Corridor disease considered in this study. Perhaps, future studies should also include a wide variety of samples from other African countries. Most important for future studies is the experimental evaluation of the immunogenic properties of this putative antigen.

7. STUDY OUTCOMES

This research was presented at University of Pretoria, Faculty of Veterinary Science Faculty Day 2019 as a poster presentation.

A poster presentation was presented at the 48th Annual Conference of the Parasitological Society of Southern Africa held in Windhoek, Namibia, 2019 and was awarded Third Place First Time Poster Presentation.

8. APPENDIX

UNIVERSITEIT VAN PRETORIA
UNIVERSITY OF PRETORIA
YUNIBESITHI YA PRETORIA

Faculty of Veterinary Science

Research Ethics Committee

Project Title	**Investigation of sequence diversity in novel vaccine candidates for Theileria parva infections**
Project Number	**REC117-19**
Researcher / Principal Investigator	**Miss L Borchers**

Dlissertation / Thesis submitted for	**Masters**

Supervisor	**Dr KP Sibeko-Matjila**

APPROVED	Date: **2019-07-04**
CHAIRMAN: UP Research Ethics Committee	Signature:

Office of the Chairman: Research Ethics Committee
Room 2-24, Pathology Building, Onderstepoort
University of Pretoria, Private Bag X04
Onderstepoort 0110, South Africa
Tel +27 (0)12 529 8052
Email
www.up.ac.za

9. REFERENCES

ALMAGRO ARMENTEROS, J. J., TSIRIGOS, K. D., SØNDERBY, C. K., PETERSEN, T. N., WINTHER, O., BRUNAK, S., VON HEIJNE, G. & NIELSEN, H. 2019. SignalP 5.0 improves signal peptide predictions using deep neural networks. *Nature Biotechnology,* 37**,** 420-423.

BISHOP, R., SHAH, T., PELLE, R., HOYLE, D., PEARSON, T., HAINES, L., BRASS, A., HULME, H., GRAHAM, S. P., TARACHA, E. L. N., KANGA, S., LU, C., HASS, B., WORTMAN, J., WHITE, O., GARDNER, M. J., NENE, V. & DE VILLIERS, E. P. 2005. Analysis of the transcriptome of the protozoan *Theileria parva* using MPSS reveals that the majority of genes are transcriptionally active in the schizont stage. *Nucleic Acids Research,* 33**,** 5503-5511.

BLAKE, D. P., CLARK, E. L., MACDONALD, S. E., THENMOZHI, V., KUNDU, K., GARG, R., JATAU, I. D., AYOADE, S., KAWAHARA, F., MOFTAH, A., REID, A. J., ADEBAMBO, A. O., ALVAREZ ZAPATA, R., SRINIVASA RAO, A. S., THANGARAJ, K., BANERJEE, P. S., DHINAKAR-RAJ, G., RAMAN, M. & TOMLEY, F. M. 2015. Population, genetic, and antigenic diversity of the apicomplexan *Eimeria tenella* and their relevance to vaccine development. *Proceedings of the National Academy of Sciences of the United States of America,* 112**,** E5343-50.

BROWN, W. C., SHAW, M. K., CONRAD, P. A. & DOLAN, T. T. 1989. *Theileria parva*: Reappearance of schizonts in infected lymphoblastoid cells treated with parvaquone is dependent on interleukin 2-like growth factors. *Experimental Parasitology,* 68**,** 308-325.

BROWNIE, J., SHAWCROSS, S., THEAKER, J., WHITCOMBE, D., FERRIE, R., NEWTON, C. & LITTLE, S. 1997. The elimination of primer-dimer accumulation in PCR. *Nucleic acids research,* 25**,** 3235-3241.

BURRIDGE, M. J. & KIMBER, C. D. 1972. The indirect fluorescent antibody test for experimental East Coast fever (*Theileria parva* infection of cattle). Evaluation of a cell culture schizont antigen. *Research in Veterinary Science,* 13**,** 451-455.

CANENE-ADAMS, K. 2013. Explanatory chapter: troubleshooting PCR. *Methods Enzymology,* 529**,** 271-8.

CHAISI, M. E., SIBEKO, K. P., COLLINS, N. E., POTGIETER, F. T. & OOSTHUIZEN, M. C. 2011. Identification of *Theileria parva* and *Theileria sp.* (buffalo) 18S rRNA gene sequence variants in the African Buffalo (*Syncerus caffer*) in southern Africa. *Veterinary Parasitology,* 182**,** 150-62.

CHATANGA, E., MOSSSAD, E., ABDO ABUBAKER, H., AMIN ALNOUR, S., KATAKURA, K., NAKAO, R. & SALIM, B. 2019. Evidence of multiple point mutations in *Theileria annulata* cytochrome b gene incriminated in buparvaquone treatment failure. *Acta Tropica,* 191**,** 128-132.

CONNELLEY, T. K., MACHUGH, N. D., PELLE, R., WEIR, W. & MORRISON, W. I. 2011. Escape from CD8+ T cell response by natural variants of an immunodominant epitope from *Theileria parva* is predominantly due to loss of TCR recognition. *The Journal of Immunology,* 187**,** 5910-20.

DE GOEYSE, I., JANSEN, F., MADDER, M., HAYASHIDA, K., BERKVENS, D., DOBBELAERE, D. & GEYSEN, D. 2015. Transfection of live, tick derived sporozoites of the protozoan Apicomplexan parasite *Theileria parva*. *Veterinary Parasitology,* 208**,** 238-41.

DI GIULIO, G., LYNEN, G., MORZARIA, S., OURA, C. & BISHOP, R. 2009. Live immunization against East Coast fever--current status. *Trends in Parasitology,* 25**,** 85-92.

DOYTCHINOVA, I. A. & FLOWER, D. R. 2007. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. *BMC Bioinformatics,* 8**,** 4.

DRAPER, S. J., ANGOV, E., HORII, T., MILLER, L. H., SRINIVASAN, P., THEISEN, M. & BISWAS, S. 2015. Recent advances in recombinant protein-based malaria vaccines. *Vaccine,* 33**,** 7433-43.

ELISA, M., HASAN, S. D., MOSES, N., ELPIDIUS, R., SKILTON, R. & GWAKISA, P. 2015. Genetic and antigenic diversity of *Theileria parva* in cattle in Eastern and Southern zones of Tanzania. A study to support control of East Coast fever. *Parasitology,* 142**,** 698-705.

EMANUELSSON, O., BRUNAK, S., VON HEIJNE, G. & NIELSEN, H. 2007. Locating proteins in the cell using TargetP, SignalP and related tools. *Nature Protocols,* 2**,** 953.

EYGELAAR, D., JORI, F., MOKOPASETSO, M., SIBEKO, K. P., COLLINS, N. E., VORSTER, I., TROSKIE, M. & OOSTHUIZEN, M. C. 2015. Tick-borne haemoparasites in African buffalo (*Syncerus caffer*) from two wildlife areas in Northern Botswana. *Parasites & Vectors,* 8**,** 26.

FERRARO, B., MORROW, M. P., HUTNICK, N. A., SHIN, T. H., LUCKE, C. E. & WEINER, D. B. 2011. Clinical applications of DNA vaccines: current progress. *Clinical Infectious Diseases,* 53**,** 296-302.

FRANK, K. & SIPPL, M. J. 2008. High-performance signal peptide prediction based on sequence alignment techniques. *Bioinformatics,* 24**,** 2172-2176.

FRANKS, S., BATON, L., TETTEH, K., TONGREN, E., DEWIN, D., AKANMORI, B. D., KORAM, K. A., RANFORD-CARTWRIGHT, L. & RILEY, E. M. 2003. Genetic diversity and antigenic polymorphism in *Plasmodium falciparum*: extensive serological cross-reactivity between allelic variants of merozoite surface protein 2. *Infection and Immunity,* 71**,** 3485-95.

FRY, L. M., BASTOS, R. G., STONE, B. C., WILLIAMS, L. B., KNOWLES, D. P. & MURPHY, S. C. 2019. Gene gun DNA immunization of cattle induces humoral and CD4 T-cell-mediated immune responses against the *Theileria parva* polymorphic immunodominant molecule. *Vaccine,* 37**,** 1546-1553.

GARDNER, M. J., BISHOP, R., SHAH, T., DE VILLIERS, E. P., CARLTON, J. M., HALL, N., REN, Q., PAULSEN, I. T., PAIN, A., BERRIMAN, M., WILSON, R. J. M., SATO, S., RALPH, S. A., MANN, D. J., XIONG, Z., SHALLOM, S. J., WEIDMAN, J., JIANG, L., LYNN, J., WEAVER, B., SHOAIBI, A., DOMINGO, A.

R., WASAWO, D., CRABTREE, J., WORTMAN, J. R., HAAS, B., ANGIUOLI, S. V., CREASY, T. H., LU, C., SUH, B., SILVA, J. C., UTTERBACK, T. R., FELDBLYUM, T. V., PERTEA, M., ALLEN, J., NIERMAN, W. C., TARACHA, E. L. N., SALZBERG, S. L., WHITE, O. R., FITZHUGH, H. A., MORZARIA, S., VENTER, J. C., FRASER, C. M. & NENE, V. 2005. Genome Sequence of *Theileria parva*, a Bovine Pathogen That Transforms Lymphocytes. *Science,* 309**,** 134-137.

GONDER, R. 1910. The life cycle of *Theileria parva*: The cause of East Coast fever in cattle in South Africa. *Journal of Comparative Pathology and Therapeutics,* 23**,** 328-335.

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., ONONO, F. O., OSASO, J., SAYA, R. M., WILDMANN, C., FRASER, C. M., MAUDLIN, I., GARDNER, M. J., MORZARIA, S. P., LOOSMORE, S., GILBERT, S. C., AUDONNET, J.-C., VAN DER BRUGGEN, P., NENE, V. & 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**,** 3286.

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*. *Infection and Immunity,* 76**,** 685-94.

HABIBI, G. 2016. In silico analysis of Ta9 gene polymorphism in an Iranian *Theileria annulata* schizont-infected cell line S15 vaccine strain and native isolates. *Archives of Razi Institute,* 71**,** 71-80.

HAIDAR, M., LATRE DE LATE, P., KENNEDY, E. J. & LANGSLEY, G. 2018. Cell penetrating peptides to dissect host-pathogen protein-protein interactions in *Theileria*-transformed leukocytes. *Bioorganic & Medicinal Chemistry,* 26**,** 1127-1134.

HALL, T. A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series,* 41**,** 95-98.

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**,** 209-20.

HEMMINK, J. D., SITT, T., PELLE, R., DE KLERK-LORIST, L. M., SHIELS, B., TOYE, P. G., MORRISON, W. I. & WEIR, W. 2018. Ancient diversity and geographical sub-structuring in African buffalo *Theileria parva* populations revealed through metagenetic analysis of antigen-encoding loci. *International Journal for Parasitology,* 48**,** 287-296.

HEMMINK, J. D., WEIR, W., MACHUGH, N. D., GRAHAM, S. P., PATEL, E., PAXTON, E., SHIELS, B., TOYE, P. G., MORRISON, W. I. & PELLE, R. 2016. Limited genetic and antigenic diversity within parasite isolates used in a live vaccine against *Theileria parva*. *International Journal for Parasitology,* 46**,** 495-506.

HENGEN, P. N. 1995. Methods and reagents: Reamplification of PCR fragments. *Trends in Biochemical Sciences,* 20**,** 124-125.

IRVIN, A. D., MORZARIA, S. P., MUNATSWA, F. C. & NORVAL, R. A. I. 1989. Immunization of cattle with a *Theileria parva bovis* stock from zimbabwe protects against challenge with virulent *T.p. parva* and *T.p. lawrencei* stocks from Kenya. *Veterinary Parasitology,* 32**,** 271-278.

JORGE, S. & DELLAGOSTIN, O. A. 2017. The development of veterinary vaccines: a review of traditional methods and modern biotechnology approaches. *Biotechnology Research and Innovation,* 1**,** 6-13.

KAO, D. J. & HODGES, R. S. 2009. Advantages of a Synthetic Peptide Immunogen Over a Protein Immunogen in the Development of an Anti-Pilus Vaccine for *Pseudomonas aeruginosa*. *Chemical Biology & Drug Design,* 74**,** 33-42.

KAR, P. P. & SRIVASTAVA, A. 2018. Immuno-informatics Analysis to Identify Novel Vaccine Candidates and Design of a Multi-Epitope Based Vaccine Candidate Against *Theileria* parasites. *Frontiers in Immunology,* 9**,** 2213.

KIMARO, E. G., MOR, S. M., GWAKISA, P. & TORIBIO, J. A. 2017. Seasonal occurrence of *Theileria parva* infection and management practices amongst Maasai pastoralist communities in Monduli District, Northern Tanzania. *Vet Parasitology,* 246**,** 43-52.

KINOSHITA, T. 2016. Glycosylphosphatidylinositol (GPI) Anchors: Biochemistry and Cell Biology: Introduction to a Thematic Review Series. *Journal of Lipid Research,* 57**,** 4-5.

KROGH, A., LARSSON, B., VON HEIJNE, G. & SONNHAMMER, E. L. L. 2001. Predicting transmembrane protein topology with a hidden markov model: application to complete genomes11Edited by F. Cohen. *Journal of Molecular Biology,* 305**,** 567-580.

KROGH, A., SONNHAMMER, E. L. L. & KÄLL, L. 2007. Advantages of combined transmembrane topology and signal peptide prediction—the Phobius web server. *Nucleic Acids Research,* 35**,** W429-W432.

KUMAR, S., STECHER, G. & TAMURA, K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Molecular Biology and Evolution,* 33**,** 1870-1874.

KUO, C.-H., WARES, J. P. & KISSINGER, J. C. 2008. The Apicomplexan Whole-Genome Phylogeny: An Analysis of Incongruence among Gene Trees. *Molecular Biology and Evolution,* 25**,** 2689-2698.

LACK, J. B., REICHARD, M. V. & VAN DEN BUSSCHE, R. A. 2012. Phylogeny and evolution of the *Piroplasmida* as inferred from 18S rRNA sequences. *International Journal for Parasitology,* 42**,** 353-363.

LATIF, A. A., HOVE. T., KANHAI, G. K., MASAKA, S. & PEGRAM, R. G. 2001. Epidemiological observations of Zimbabwean theileriosis: disease incidence and pathogenicity in susceptible cattle during *Rhipicephalus appendiculatus* nymphal and adult seasonal activity. *Onderstepoort Journal of Veterinary Research,* 68**,** 187-195.

LAWRENCE, J. 1992. History of bovine theileriosis in southern Africa. *The epidemiology of theileriosis in Africa***,** 1-40.

LAWRENCE, J. A. 1979. The differential diagnosis of the bovine *theilerias* of southern Africa. *Journal of the South African Veterinary Association,* 50**,** 311-313.

LEW-TABOR, A. E. & RODRIGUEZ VALLE, M. 2016. A review of reverse vaccinology approaches for the development of vaccines against ticks and tick borne diseases. *Ticks and Tick-borne Diseases,* 7**,** 573-85.

LIPSITCH, M. & O'HAGAN, J. J. 2007. Patterns of antigenic diversity and the mechanisms that maintain them. *Journal of the Royal Society Interface,* 4**,** 787-802.

LOHIA, N. & BARANWAL, M. 2014. Conserved peptides containing overlapping CD4+ and CD8+ T-cell epitopes in the H1N1 influenza virus: an immunoinformatics approach. *Viral Immunology,* 27**,** 225-34.

LORENZ, T. C. 2012. Polymerase Chain Reaction: Basic Protocol Plus Troubleshooting and Optimization Strategies. *Journal of Visualized Experiments***,** e3998.

MACHUGH, N. D., CONNELLEY, T., GRAHAM, S. P., PELLE, R., FORMISANO, P., TARACHA, E. L., ELLIS, S. A., MCKEEVER, D. J., BURRELLS, A. & MORRISON, W. I. 2009. CD8+ T-cell responses to *Theileria parva* are preferentially directed to a single dominant antigen: Implications for parasite strain-specific immunity. *European Journal of Immunology,* 39**,** 2459-69.

MACHUGH, N. D., WEIR, W., BURRELLS, A., LIZUNDIA, R., GRAHAM, S. P., TARACHA, E. L., SHIELS, B. R., LANGSLEY, G. & MORRISON, W. I. 2011. Extensive polymorphism and evidence of immune selection in a highly dominant antigen recognized by bovine CD8 T cells specific for *Theileria annulata*. *Infection and Immunity,* 79**,** 2059-69.

MANS, B. J., PIENAAR, R. & LATIF, A. A. 2015. A review of *Theileria* diagnostics and epidemiology. *International Journal for Parasitology: Parasites and Wildlife.* 2015/04/02 ed.

MARÍA, R. A. R., ARTURO, C. V. J., ALICIA, J. A., PAULINA, M. L. G. & GERARDO, A. O. 2017. The Impact of Bioinformatics on Vaccine Design and Development. *Vaccines.*

MATEU, M. G., MARTÍNEZ, M. A., CAPUCCI, L., ANDREU, D., GIRALT, E., SOBRINO, F., BROCCHI, E. & DOMINGO, E. 1990. A single amino acid substitution affects multiple overlapping epitopes in the major antigenic site of foot-and-mouth disease virus of serotype C. *Journal of General Virology,* 71**,** 629-637.

MATSON, B. A. 1967. Theileriosis in Rhodesia: 1. A study of diagnostic specimens over two seasons. *Journal of the South African Veterinary Association,* 38**,** 93-102.

MBIZENI, S., POTGIETER, F. T., TROSKIE, C., MANS, B. J., PENZHORN, B. L. & LATIF, A. A. 2013. Field and laboratory studies on Corridor disease (*Theileria parva* infection) in cattle population at the livestock/game interface of uPhongolo-Mkuze area, South Africa. *Ticks and Tick-borne Diseases,* 4**,** 227-34.

MCKEEVER, D. I. & MORRISON, W. I. 1994. Immunity to a parasite that transforms T lymphocytes. *Current Opinion in Immunology,* 6**,** 564-567.

MEKONNEN, S., BRYSON, N. R., FOURIE, L. J., PETER, R. J., SPICKET, A. M., TAYLOR, R. J., STRYDOM, T. & HORAK, I. G. 2002. Acaricide resistance profiles of single- and multi-host ticks from communal and commercial farming areas in the Eastern Cape and North-West Provinces of South Africa. *Onderstepoort Journal of Veterinary Research,* 69**,** 99-105.

MICHEL, A. L. & BENGIS, R. G. 2012. The African buffalo: a villain for inter-species spread of infectious diseases in southern Africa. *Onderstepoort Journal of Veterinary Research,* 79**,** 453.

MINJAUW, B., OTTE, M. J., JAMES, A. D., DE CASTRO, J. J. & SINYANGWE, P. 1998. Effect of different East Coast fever control strategies on disease incidence in traditionally managed Sanga cattle in Central Province of Zambia. *Preventive Veterinary Medicine,* 35**,** 101-113.

MONTERRUBIO-LOPEZ, G. P., GONZALEZ, Y. M. J. A. & RIBAS-APARICIO, R. M. 2015. Identification of Novel Potential Vaccine Candidates against Tuberculosis Based on Reverse Vaccinology. *Biomed Research International,* 2015**,** 483150.

MORRISON, I. W., 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*.

MORRISON, W. I. 1996. Subunit vaccines for complex tick bornepathogens: Dream or reality? *British Veterinary Journal,* 152**,** 617-619.

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

MORRISON, W. I. 2015. The aetiology, pathogenesis and control of *theileriosis* in domestic animals. *Scientific and Technical Review of the Office International des Epizooties,* 32.

MORRISON, W. I., GODDEERIS, B. M., TEALE, A. J., GROOCOCK, C. M., KEMP, S. J. & STAGG, D. A. 1987. Cytotoxic T-cells elicited in cattle challenged with *Theileria parva* (Muguga): evidence for restriction by class I MHC determinants and parasite strain specificity. *Parasite Immunology,* 9**,** 563-578.

MORRISON, W. I., TARACHA, E. L. N. & MCKEEVER, D. J. 1995. Theileriosis: progress towards vaccine development through understanding immune responses to the parasite. *Veterinary Parasitology,* 57**,** 177-187.

MORTAZAVI, A., WILLIAMS, B. A., MCCUE, K., SCHAEFFER, L. & WOLD, B. 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods,* 5**,** 621-628.

MORZARIA, S. P., IRVIN, A. D., TARACHA, E., SPOONER, P. R., VOIGT, W. P., FUJINAGA, T. & KATENDE, J. 1987. Immunization against east coast fever: The use of selected stocks of *Theileria parva* for immunization of cattle exposed to field challenge. *Veterinary Parasitology,* 23**,** 23-41.

MUKOLWE, L. D., ODONGO, D. O., BYARUHANGA, C., SNYMAN, L. P. & SIBEKO-MATJILA, K. P. 2020. Analysis of p67 allelic sequences reveals a subtype of allele type 1 unique to buffalo-derived *Theileria parva* parasites from southern Africa. *PLOS ONE,* 15**,** e0231434.

MURAGURI, G. R., KIARA, H. K. & MCHARDY, N. 1999. Treatment of East Coast fever: a comparison of parvaquone and buparvaquone. *Veterinary Parasitology,* 87**,** 25-37.

MUSOKE, A., ROWLANDS, J., NENE, V., NYANJUI, J., KATENDE, J., SPOONER, P., MWAURA, S., ODONGO, D., NKONGE, C., MBOGO, S., BISHOP, R. & MORZARIA, S. 2005. Subunit vaccine based on the p67 major surface protein of *Theileria parva* sporozoites reduces severity of infection derived from field tick challenge. *Vaccine,* 23**,** 3084-3095.

MUSOKE, A. J., PALMER, G. H., MCELWAIN, T. F., NENE, V. & MCKEEVER, D. 1996. Prospects for subunit vaccines against tick borne diseases. *British Veterinary Journal,* 152**,** 621-639.

NASCIMENTO, I. P. & LEITE, L. C. C. 2012. Recombinant vaccines and the development of new vaccine strategies. *Brazilian Journal of Medical and Biological Research,* 45**,** 1102-1111.

NATHALY WIESER, S., SCHNITTGER, L., FLORIN-CHRISTENSEN, M., DELBECQ, S. & SCHETTERS, T. 2019. Vaccination against babesiosis using recombinant GPI-anchored proteins. *International Journal for Parasitology,* 49**,** 175-181.

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 and Tick-borne Diseases,* 7**,** 549-64.

NENE, V., MUSOKE, A., GOBRIGHT, E. & MORZARIA, S. 1996. Conservation of the Sporozoite p67 Vaccine Antigen in Cattle-Derived *Theileria parva* Stocks with Different Cross-Immunity Profiles. *Infection and Immunity,* 64**,** 2056-2061.

NENE, V., SVITEK, N., TOYE, P., GOLDE, W. T., BARLOW, J., HARNDAHL, M., BUUS, S. & NIELSEN, M. 2012. Designing bovine T cell vaccines via reverse immunology. *Ticks and Tick-borne Diseases,* 3**,** 188-192.

NLINWE, O. N., KUSI, K. A., ADU, B. & SEDEGAH, M. 2018. T-cell responses against Malaria: Effect of parasite antigen diversity and relevance for vaccine development. *Vaccine,* 36**,** 2237-2242.

NYAGWANGE, J., NENE, V., MWALIMU, S., HENSON, S., STEINAA, L., NZAU, B., TIJHAAR, E. & PELLE, R. 2018. Antibodies to in silico selected GPI-anchored *Theileria parva* proteins neutralize sporozoite infection in vitro. *Veterinary Immunology and Immunopathology,* 199**,** 8-14.

NYAGWANGE, J., TIJHAAR, E., TERNETTE, N., MOBEGI, F., TRETINA, K., SILVA, J. C., PELLE, R. & NENE, V. 2017. Characterization of the *Theileria parva* sporozoite proteome. *International Journal for Parasitology*.

OUATTARA, A., BARRY, A. E., DUTTA, S., REMARQUE, E. J., BEESON, J. G. & PLOWE, C. V. 2015. Designing malaria vaccines to circumvent antigen variability. *Vaccine,* 33**,** 7506-12.

OURA, C. A. L., TAIT, A., ASIIMWE, B., LUBEGA, G. W. & WEIR, W. 2011. Haemoparasite prevalence and *Theileria parva* strain diversity in Cape buffalo (*Syncerus caffer*) in Uganda. *Veterinary Parasitology,* 175**,** 212-219.

PAPLI, N., LANDT, O., FLEISCHER, C., KOEKEMOER, J. O., MANS, B. J., PIENAAR, R., JOSEMANS, A., ZWEYGARTH, E., POTGIETER, F. & LATIF, A. A. 2011. Evaluation of a TaqMan real-time PCR for the detection of *Theileria parva* in buffalo and cattle. *Veterinary Parasitology,* 175**,** 356-359.

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**,** e19015.

PEREGRINE, A. S. 1994. Chemotherapy and delivery systems: haemoparasites. *Veterinary Parasitology,* 54**,** 223-248.

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

RADLEY, D. E., BROWN, C. G. D., BURRIDGE, M. J., CUNNINGHAM, M. P., KIRIMI, I. M., PURNELL, R. E. & YOUNG, A. S. 1975a. East coast fever: 1. Chemoprophylactic immunization of cattle against *Theileria parva* (Muguga) and five theilerial strains. *Veterinary Parasitology,* 1**,** 35-41.

RADLEY, D. E., BROWN, C. G. D., CUNNINGHAM, M. P., KIMBER, C. D., MUSISI, F. L., PAYNE, R. C., PURNELL, R. E., STAGG, S. M. & YOUNG, A. S. 1975b. East coast fever: 3. Chemoprophylactic immunization of cattle using oxytetracycline and a combination of theilerial strains. *Veterinary Parasitology,* 1**,** 51-60.

RAPPUOLI, R. 2000. Reverse vaccinology. *Current Opinion in Microbiology,* 3**,** 445-450.

ROCCHI, M. S., BALLINGALL, K. T., MACHUGH, N. D. & MCKEEVER, D. J. 2006. The kinetics of *Theileria parva* infection and lymphocyte transformation in vitro. *International Journal of Parasitology,* 36**,** 771-8.

SALIH, D. A., PELLE, R., MWACHARO, J. M., NJAHIRA, M. N., MARCELLINO, W. L., KIARA, H., MALAK, A. K., EL HUSSEIN, A. R., BISHOP, R. & SKILTON, R. A. 2017. Genes encoding two *Theileria parva* antigens recognized by CD8+ T-cells exhibit sequence diversity in South Sudanese cattle populations but the majority of alleles are similar to the Muguga component of the live vaccine cocktail. *PLoS One,* 12**,** e0171426.

SETTE, A. & RAPPUOLI, R. 2010. Reverse vaccinology: developing vaccines in the era of genomics. *Immunity,* 33**,** 530-41.

SHAW, M. K. & TILNEY, L. G. 1995. The entry of *Theileria parva* merozoites into bovine erythrocytes occurs by a process similar to sporozoite invasion of lymphocytes. *Parasitology,* 111**,** 455-461.

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

SHUAIBU, M. N., KIKUCHI, M., CHERIF, M. S., HELEGBE, G. K., YANAGI, T. & HIRAYAMA, K. 2010. Selection and identification of malaria vaccine target molecule using bioinformatics and DNA vaccination. *Vaccine,* 28**,** 6868-75.

SIBEKO, K. P., COLLINS, N. E., OOSTHUIZEN, M. C., TROSKIE, M., POTGIETER, F. T., COETZER, J. A. & GEYSEN, D. 2011. Analyses of genes encoding *Theileria parva* p104 and polymorphic immunodominant molecule (PIM) reveal evidence of the presence of cattle-type alleles in the South African *T. parva* population. *Vet Parasitology,* 181**,** 120-30.

SIBEKO, K. P., GEYSEN, D., OOSTHUIZEN, M. C., MATTHEE, C. A., TROSKIE, M., POTGIETER, F. T., COETZER, J. A. & COLLINS, N. E. 2010. Four p67 alleles identified in South African *Theileria parva* field samples. *Vet Parasitology,* 167**,** 244-54.

SIBEKO, K. P., OOSTHUIZEN, M. C., COLLINS, N. E., GEYSEN, D., RAMBRITCH, N. E., LATIF, A. A., GROENEVELD, H. T., POTGIETER, F. T. & COETZER, J. A. 2008. Development and evaluation of a real-time polymerase chain reaction test for the detection of *Theileria parva* infections in Cape buffalo (*Syncerus caffer*) and cattle. *Vet Parasitology,* 155**,** 37-48.

SITT, T., HENSON, S., MORRISON, W. I. & TOYE, P. 2019. Similar levels of diversity in the gene encoding the p67 sporozoite surface protein of *Theileria parva* are observed in blood samples from buffalo and cattle naturally infected from buffalo. *Vet Parasitology,* 269**,** 21-27.

SITT, T., PELLE, R., CHEPKWONY, M., MORRISON, W. I. & TOYE, P. 2018. *Theileria parva* antigens recognized by CD8+ T cells show varying degrees of diversity in buffalo-derived infected cell lines. *Parasitology,* 145**,** 1430-1439.

SITT, T., POOLE, E. J., NDAMBUKI, G., MWAURA, S., NJOROGE, T., OMONDI, G. P., MUTINDA, M., MATHENGE, J., PRETTEJOHN, G., MORRISON, W. I. & TOYE, P. 2015. Exposure of vaccinated and naive cattle to natural challenge from buffalo-derived *Theileria parva*. *International Journal for Parasitology: Parasites and Wildlife,* 4**,** 244-51.

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**,** 265-76.

SMITH, E. R. & PARKER, D. M. 2010. Tick communities at the expanding wildlife/cattle interface in the Eastern Cape Province, South Africa: implications for Corridor disease. *Journal of the South African Veterinary Association,* 81**,** 237-240.

STEINAA, L., SAYA, R., AWINO, E. & TOYE, P. 2012. Cytotoxic T lymphocytes from cattle immunized against *Theileria parva* exhibit pronounced cross-reactivity among different strain-specific epitopes of the Tp1 antigen. *Veterinary Immunology and Immunopathology,* 145**,** 571-81.

TANABE, K., MITA, T., PALACPAC, N. M., ARISUE, N., TOUGAN, T., KAWAI, S., JOMBART, T., KOBAYASHI, F. & HORII, T. 2013. Within-population genetic diversity of *Plasmodium falciparum* vaccine candidate antigens reveals geographic distance from a Central sub-Saharan African origin. *Vaccine,* 31**,** 1334-9.

TARACHA, E. L. N., GODDEERIS, B. M., MORZARIA, S. P. & MORRISON, W. I. 1995. Parasite Strain Specificity of Precursor Cytotoxic T Cells in Individual Animals Correlates with Cross-Protection in Cattle Challenged with *Theileria parva*. *Infection and Immunity,* 63.

TAYEBWA, D. S., VUDRIKO, P., TUVSHINTULGA, B., GUSWANTO, A., NUGRAHA, A. B., GANTUYA, S., BATIHA, G. E., MUSINGUZI, S. P., KOMUGISHA, M., BBIRA, J. S., OKWEE-ACAI, J., TWEYONGYERE, R., WAMPANDE, E. M., BYARUHANGA, J., ADJOU MOUMOUNI, P. F.,

SIVAKUMAR, T., YOKOYAMA, N. & IGARASHI, I. 2018. Molecular epidemiology of *Babesia species*, *Theileria parva*, and *Anaplasma marginale* infecting cattle and the tick control malpractices in Central and Eastern Uganda. *Ticks and Tick-borne Diseases,* 9**,** 1475-1483.

THOMPSON, J. D., HIGGINS, D. G. & GIBSON, T. J. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic acids research,* 22**,** 4673-4680.

THOMSEN, M., LUNDEGAARD, C., BUUS, S., LUND, O. & NIELSEN, M. 2013. MHCcluster, a method for functional clustering of MHC molecules. *Immunogenetics,* 65**,** 655-665.

TOYE, P., GOBRIGHT, E., NYANJUI, J., NENE, V. & BISHOP, R. 1995. Structure and sequence variation of the genes encoding the polymorphic, immunodominant molecule (PIM), an antigen of *Theileria parva* recognized by inhibitory monoclonal antibodies. *Molecular and Biochemical Parasitology,* 73**,** 165-177.

TOYE, P., 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**,** 1832-1838.

TRETINA, K., GOTIA, H. T., MANN, D. J. & SILVA, J. C. 2015. *Theileria*-transformed bovine leukocytes have cancer hallmarks. *Trends in Parasitology,* 31**,** 306-14.

TRETINA, K., PELLE, R., ORVIS, J., GOTIA, H. T., IFEONU, O. O., KUMARI, P., PALMATEER, N. C., IQBAL, S. B. A., FRY, L., NENE, V. M., DAUBENBERGER, C., BISHOP, R. P. & SILVA, J. C. 2019. Re-annotation of the *Theileria parva* genome refines 53% of the proteome and uncovers essential components of N-glycosylation, a conserved pathway in many organisms. *bioRxiv***,** 749366.

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

VERVERKEN, C., GEYSEN, D., LOOTS, K., JANSSENS, M. E., GUISEZ, Y. & GODDEERIS, B. M. 2008. Orientation of bovine CTL responses towards PIM, an antibody-inducing surface molecule of *Theileria parva*, by DNA subunit immunization. *Veterinary Immunology and Immunopathology,* 124**,** 253-263.

VUDRIKO, P., OKWEE-ACAI, J., TAYEBWA, D. S., BYARUHANGA, J., KAKOOZA, S., WAMPANDE, E., OMARA, R., MUHINDO, J. B., TWEYONGYERE, R., OWINY, D. O., HATTA, T., TSUJI, N., UMEMIYA-SHIRAFUJI, R., XUAN, X., KANAMEDA, M., FUJISAKI, K. & SUZUKI, H. 2016. Emergence of multi-acaricide resistant *Rhipicephalus* ticks and its implication on chemical tick control in Uganda. *Parasites & Vectors,* 9**,** 4.

WALKER, J. G., KLEIN, E. Y. & LEVIN, S. A. 2014. Disease at the wildlife-livestock interface: acaricide use on domestic cattle does not prevent transmission of a tick-borne pathogen with multiple hosts. *Vet Parasitology,* 199**,** 206-14.

YE, J., COULOURIS, G., ZARETSKAYA, I., CUTCUTACHE, I., ROZEN, S. & MADDEN, T. L. 2012. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC bioinformatics,* 13**,** 134-134.

YUSUFMIA, S. B. A. S., COLLINS, N. E., NKUNA, R., TROSKIE, M., VAN DEN BOSSCHE, P. & PENZHORNA, 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**,** 45-49.

ZHANG, H., LI, H., ZHU, H., PEKÁREK, J., PODEŠVA, P., CHANG, H. & NEUŽIL, P. 2019. Revealing the secrets of PCR. *Sensors and Actuators B: Chemical,* 298**,** 126924.

ZHONG, W., RECHE, P. A., LAI, C. C., REINHOLD, B. & REINHERZ, E. L. 2003. Genome-wide characterization of a viral cytotoxic T lymphocyte epitope repertoire. *Journal of Biological Chemistry,* 278**,** 45135-44.