MOLECULAR DETECTION, GENETIC DIVERSITY AND PHYLOGENETIC ANALYSIS OF *ANAPLASMA MARGINALE* INFECTING CATTLE IN SOUTH AFRICA

By

Awelani Mirinda Mutshembele

Thesis submitted in fulfillment of the requirements for the degree Philosophiae Doctor in the Faculty of Natural and Agricultural Sciences, Department of Zoology and Entomology, University of the Free State

Promoters: Dr. M.S Mtshali and Prof O.M.M Thekisoe

December 2015

PROMOTERS

Dr. Moses S. Mtshali (PhD)

Research and Scientific Services Department

National Zoological Gardens of South Africa

P.O. Box 754

Pretoria

0001

Department of Zoology and Entomology,

University of the Free State, QwaQwa Campus

Private Bag X13

Phuthaditjhaba

9866

Prof. Oriel M.M. Thekisoe (PhD)

Unit for Environmental Sciences and Management

North – West University, Potchefstroom Campus

Private Bag x 6001

Potchefstroom

2520

DECLARATION

I, the undersigned, hereby declare that the work contained in this thesis is my own original work and that I have not previously in it's entirely or in part submitted it at any university for a degree. I furthermore, cede copyright of the thesis in favour of the University of the Free State

06/04/2016

A.M. Mutshembele Date

DEDICATION

This thesis is dedicated to my parents my late father Itani Daniel Mutshembele may his soul rest in peace, and my mother Tshilidzi Mercy Mutshembele for their endless love, support , encouragement and tolerance on my career choice.

ACKNOWLEDGEMENTS

I thank Almighty God for the strength He gave me to do research under supervision of Dr. M. S. Mtshali and Prof O. M. M. Thekisoe, without their assistance and dedicated involvement in every step throughout the process, this study would have never been accomplished. I thank them very much for their support and understanding over these past four years. I also thank our collaborators Prof José de la Fuente and Dr. Alejandro Cabezas-Cruz for their endless support during my research period.

I would also like to show gratitude to my colleagues and friends in the Veterinary Parasitology Programme for their immeasurable contributions and not forgetting the soul mates Dr. Raksha Bhoora, Dr. Tracy Masebe, Zama Khumalo, Lesego Modibedi – Masango, Portia Motheo, Sabath Rathanya, Verdi Giqwa and Seipati Nku for their support and patience during difficult study times and "Sissy Khanyi" Prof Khanyisile Mbatha for her support and advice throughout the year, she was my pillar.

I also thank Moeti Oriel Taioe for his assistance with phylogenetic analysis at the Department of Zoology and Entomology, University of Free State QwaQwa Campus and Dr. Steve Olorunju at the Biostatistics Unit, Medical Research Council of South Africa, Pretoria who kindly assisted me with the statistical analysis and was very patient with me.

I thank my brothers Dr. Theodore Mutshembele, Ndivhoni Mutshembele, my sister Mudzunga Mutshembele and my nieces (Mapitso and Malindi). Every time I was ready to quit, they did not let me to quit and I am forever grateful, this thesis stands as a testament to your unconditional love and encouragement, and Abel Ralinala for his love, patience, jokes and words of inspiration that gave me strength and kept me motivated.

Getting through my dissertation required more than academic support, and I have many, many people to thank for listening to me and, at times, having to tolerate me over the past three years. I cannot begin to express my gratitude and appreciation for their friendship.

I also thank the Directors of Department of Agriculture and Veterinary Institute from Limpopo, Mpumalanga, North West, KwaZulu-Natal, Eastern Cape, Western Cape and Northern Cape (Prof. Nomfundo Mnisi, Drs Mlilo, Songelwayo Chisi, L. Mrwebi, Ronald Sinclair and Francis Babi) with the assistance from the Veterinary Technicians (Toypat Mdluli, Keneilwe Constance Kaotane, Jan Masethe Maime, Anil Suresh and Erwin Lucas) and the farmers who co-operated in sample collection.

I acknowledge the University of the Free State and National Zoological Gardens of South Africa for availing their facilities during this study. This study is based on the research supported in part by the National Research Foundation of South Africa from the DST/NRF Professional Development Programme, grant made available to Dr. M.S. Mtshali; and the National Zoological Gardens of South Africa. Any opinion, finding and conclusion or recommendation expressed in this study is that of the author and the NRF does not accept any liability in this regard.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

APPENDICES

LIST OF ABBREVIATIONS

°C	degree Celsius
BLAST	Basic Local Alignment Search Tool
BLASTn	Basic Local Alignment Search Tool for Nucleotide
bp	base pair
CAT	Card Agglutination Test
cELISA	Competitive enzyme-linked immunosorbent assay
CF	Complement fixation test
DNA	deoxyribonucleic acid
EC	Eastern Cape
EDTA	Ethylenediamine tetra -acetic acid
GDI	Genetic diversity index
GP	Gauteng Province
IFA	Indirect fluorescent antibody
Kb	Kilobase
kDa	Kilodalton
KZN	KwaZulu-Natal
LP	Limpopo Province
MAb	Monoclonal antibody
mM	millimolar
ml	milliliter
MP	Mpumalanga
MSP	Major Surface Protein
NC	Northern Cape
NCBI	National Center for Biotechnology Information
ng	nanogram
ng/µl	nanogram per microliter
NW	North -West
PCR	polymerase chain reaction
TBD	Tick-borne diseases

UV	ultraviolet
WC	Western Cape
µg	microgram(s)
µm	micrometer
µM	micromolar

RESEARCH OUTPUTS

Publications

Awelani M. Mutshembele, Alejandro Cabezas-Cruz, Moses S. Mtshali, Oriel, M. M. Thekisoe, Ruth C. Galindo and José de la Fuente. 2014. Epidemiology and evolution of the genetic variability of *Anaplasma marginale* in South Africa. Ticks and Tick -borne diseases, Volume 5, Issue 6, 624-631.

Conferences

Awelani M. Mutshembele, A. Cabezas-Cruz, M. S. Mtshali, O. M. M. Thekisoe, R. C. Galindo, J de la Fuente. Epidemiology and evolution of the genetic variability of *Anaplasma marginale* in South Africa. 8th International Ticks and Tick-Borne Pathogens and Biennial Social Tropical Veterinary Medicine (TTP-STVM) Conference. Cape Town, South Africa. 24-29 August 2014.

Awelani M. Mutshembele, M. S. Mtshali, O. M. M. Thekisoe, R. C. Galindo, A. Cabezas-Cruz, J de la Fuente. Molecular prevalence, genetic diversity and phylogenetic analysis of *Anaplasma marginale* isolates in cattle in South Africa. 42nd PARSA Conference hosted by the North West University, Stonehenge in Africa, Parys, Free State, 22-24 September 2013.

Seminars and Symposia

Awelani M. Mutshembele, M.S. Mtshali and O. M. Thekisoe. Molecular detection, genetic diversity and phylogenetic analysis of *Anaplasma marginale* isolates in cattle in South Africa. Postgraduate seminar day, University of the Free State- QwaQwa campus. 13-14 November 2013.

Awelani M. Mutshembele, A. Cabezas-Cruz, M. S. Mtshali, O. M. M. Thekisoe, R. C. Galindo, J de la Fuente. Structural and epidemiological analysis of *Anaplasma marginale* using the major surface protein 1a in South Africa. 4th National Zoological Gardens of South Africa (NZG) Research Symposium, November 2013.

Awelani M. Mutshembele, Moses S. Mtshali, Oriel M. M. Thekisoe. The genetic diversity and phylogenetic analysis of *Anaplasma marginale* strains in cattle in South Africa. The Zoology Department Seminar, Faculty of Zoology and Entomology, University of the Free State, Bloemfontein Campus, November 2012.

Awelani M. Mutshembele, Moses S. Mtshali, Oriel M. M. Thekisoe. The genetic diversity and phylogenetic analysis of *Anaplasma marginale* strains in cattle in South Africa. 3rd National Zoological Gardens of South Africa (NZG) Research Symposium, November 2012.

Awelani M. Mutshembele, Moses S. Mtshali, Oriel M. M. Thekisoe. Analysis of phylogeographic relationships of *Anaplasma marginale* from South African isolates using *msp4* gene sequences. The Zoology Department Seminar, Faculty of Zoology and Entomology, QwaQwa campus. November 2011.

ABSTRACT

Bovine anaplasmosis caused by *Anaplasma marginale* is endemic in South Africa. This endemicity is due to presence of tick vectors that transmit *A. marginale* the causal agent of the disease and the high seroprevalence in Limpopo, Free State and North West provinces. To date, the genetic diversity of *A. marginale* isolates infecting cattle in all South African provinces, except Free State, are generally unknown. Recently, vaccines based on the *A. marginale* major surface protein 1a (MSP1a) has been proposed as a strategy for controlling bovine anaplasmosis. However, characterization of genetic diversities of the *A. marginale* isolates in these regions is still needed before this protein can be used for vaccine development. Therefore, the aim of this study was to determine the prevalence, genetic diversity and phylogenetic relationship of *A. marginale* infecting cattle in all South African provinces except the Free State.

A total of 280 whole blood samples were collected from cattle in all provinces with exception of the Free State. Twenty six districts and municipalities were included in this sampling. *Anaplasma marginale* genomic DNA was then extracted from the blood sample using ZR Genomic DNA™ Tissue Miniprep (Zymo Research, CA, USA). A polymerase chain reaction (PCR) was done with primers targeting *msp1a* and *msp4* genes and the PCR products were sequenced using genetic analyser (ABI, Life technologies, CA, USA). The generated sequences were analysed by bioinformatics and their phylogeny as well as genetic diversity index (GDI) was determined based on the sequences of *msp1a* and *msp4* genes.

Overall, the prevalence of *A. marginale* infection in cattle was 76% in all provinces except for Northern Cape Province where the prevalence was zero. The prevalence per province was as follows: Eastern Cape 19.1%, Gauteng 9.6%, KwaZulu-Natal 23.0%, Limpopo 15.3%, Mpumalanga 10.1%, North West 12.4% and Western Cape 10.5%. The msp1a revealed genetic variability with regions of different types of tandem repeats. Some repeats were conserved amongst the *A. marginale* strains and revealed low variable peptides in the MSP1a tandem repeats. A polynomial correlation (R2=0.76) was observed between the GDI and anaplasmosis prevalence per province. Interestingly, provinces with the highest prevalence were not the ones with highest or lowest GDI.

The analysis of *msp4* gene sequences, which provided evolutionary information about geographically distinct *A. marginale* strains, was used in the present study for phylogenetic analysis of samples from Limpopo (LP), Mpumalanga (MP), North West (NW), Gauteng (GP), KwaZulu-Natal (KZN), Eastern Cape (EC) and Western Cape (WC) provinces of South Africa. Two clades were observed which consisted of first clade (LP, NW, GP, KZN and WC) and second clade (MP and EC) isolates.

In addition when DNA sequence variation of *msp4* gene was analysed in combination with isolates from other countries outside South Africa, important phylogeographic information was observed. The South African strains had 100% identity with isolates from Kenya, Zimbabwe and Australia. Good representation of the Southern and Northern Hemispheres was observed and demonstrated that the *msp4* gene was a good phylogeographic marker. These results indicated that *A. marginale* is widespread in South Africa, and suggested that the analysis of *msp1a* and *msp4* gene sequences provided an understanding of the phylogeny and epidemiology of *A. marginale* in South Africa.

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

1.1 Historical background

The genus *Anaplasma* was established in 1910 by Sir Arnold Theiler who first described the "marginal points" in stained erythrocytes of sick cattle as the causative agents of a specific disease (Theiler, 1910; 1911). These erythrocytic inclusions had been seen frequently in red blood cells of anemic cattle often in those suffering from babesiosis (piroplasmosis). The acute phase of the disease is characterized by weight loss, fever, abortion, lowered milk production and often death (Kuttler, 1984).

1.2 Classification of *Anaplasma marginale*

Bovine anaplasmosis is a tick-borne rickettsial disease caused by the hemoparasite *Anaplasma marginale* (Aubry and Geale, 2011). *Anaplasma marginale* is classified within the Order *Rickettsiales* which was recently reorganized into two families, Anaplasmataceae and Rickettsiaceae based on genetic analysis of *16S rRNA, groELS* and surface protein genes (Dumler *et al.,* 2001). Members of the family Anaplasmataceae are obligate intracellular organisms found exclusively within membrane-bound vacuoles in the host cell cytoplasm. The genus *Anaplasma* includes three species that infect ruminants, namely *A. marginale, A. centrale* and *A. ovis* (Dumler *et al.,* 2001), also included within this genus is *A. phagocytophilum*, agent of human granulytic ehrlichiosis [HEG], *A. bovis* and *A. platys* (Kocan *et al.,* 2010).

Anaplasma marginale is a gram- negative rickettsia which is endemic in tropical and subtropical areas throughout the world (Decaro *et al.,* 2008; Howden and Geale, 2010; Kocan *et al.,* 2010; Aubry and Geale, 2011). Bovine anaplasmosis is a major constraint to cattle production in many countries (Kocan *et al.,* 2010).

In order to persist in nature, *A. marginale* infects the mammalian host which usually remains persistently infected (Kieser *et al.,* 1990; Eriks *et al.,* 1993) serving as a reservoir for infection of ticks or mechanical transmission by the transfer of blood from infected to susceptible cattle (Kocan *et al.,* 1992; Ge *et al.,* 1996; Futse *et al.,* 2003).

Taxonomic Classification of *Anaplasma* infecting ruminants (Dumler *et al.,* 2001):

Kingdom: Bacteria

Phylum: Proteobacteria

Class: Alphaproteobacteria

Order: Rickettsiales

Family: Anaplasmataceae

Genus: *Anaplasma*

Species: *A. marginale*
A. centrale
A. bovis
A. ovis
A. phagocytophilum
A. platys

1.3 Epidemiology

Bovine anaplasmosis occurs in tropical and subtropical regions of the world and the disease is a major constraint to cattle production in many countries. *Anaplasma marginale*, in contrast to *A. phagocytophilum*, is quite host specific, infecting only ruminants and causing disease primarily in cattle. In the U.S.A. it is enzootic throughout the southern Atlantic states, Gulf Coast states, and several of the Midwestern and Western states (Kocan *et al.,* 2010).

However, it has been reported in almost every state in the U.S.A. This increasingly wide distribution likely resulted from transport of carrier cattle with subsequent mechanical or biological transmission from asymptomatic persistently infected cattle to susceptible ones (Kocan *et al.*, 2010).

It is also endemic in Mexico, Central and South America, as well as in the Caribbean Islands. It is enzootic in all Latin American countries, with the exception of desert areas and certain mountain ranges, like the Andes. The seroprevalence rates of *A. marginale* vary widely among countries in the America and the variability of these rates contributes to the development of geographically stable enzootic regions (Kocan *et al.*, 2010).

In Europe, *A. marginale* is found mainly in Mediterranean countries with infections having been described in cattle and assorted wildlife species. It is also endemic in regions of Asia and Africa. The distribution of anaplasmosis may be expected to continue to change, in part as a result of global warming, which may influence the movement of the tick hosts. An example of the validity of such a prediction is a confirmed diagnosis of anaplasmosis in a bison herd in Saskatchewan, Canada, during the summer of 2000. The first reported outbreak of anaplasmosis in Canada occurred in 1971, but this outbreak resulted from mechanical transmission from imported carrier cattle to local ones (Kocan *et al.*, 2010).

In South Africa only one study has been conducted to genetically characterize the geographical strains (Mtshali *et al.*, 2007). The results presented by the latter authors indicated the presence of a common genotype between South African, American and European strains of *A. marginale*. However the study focused only on blood samples of cattle collected from Free State province, so there is a need of further documentation of prevalence and genetic diversity of *A. marginale* strains in cattle of South African origin in order to design epidemiological and control strategies.

1.4 Distribution of ticks infesting cattle in South Africa

Several species of ticks are found in different geographical areas in South Africa. The type of vegetation, humidity and temperature influences their distribution. Studies in the eastern Free State revealed that ticks affect mainly cattle of small farmers and the main tick species in the area are *Rhipicephalus decoloratus* (53.1%), *R. evertsi evertsi* (44.7%), *R. follis* (1.0%), *R. gertrudae* (0.7%) and *R. warburtoni* (0.4%). In the south west region of the same province, Fourie and Horak (1991) observed that *Amblyomma marmoreum*, *Hyalomma marginatum rufipes* and *H. truncatum* were the predominant species. A second study by Fourie *et al.* (1996) in the south west region revealed that *Ixodes rubicundus* and *H. m. rufipes* were the most prevalent tick species (Marufu, 2008).

In KwaZulu-Natal, Baker *et al.* (1989) observed that *R. decoloratus*, *R. appendiculatus* and *R. evertsi evertsi* were the most prevalent tick species of cattle raised on commercial farms. *H. m. rufipes*, *R. appendiculatus* and *R. evertsi evertsi* were the most numerous species in the Limpopo Province also present were *R. microplus, R. decoloratus*, *A. hebraeum*, *H. truncatum* and *R. simus*. Bryson *et al.* (2002) noted that the adults of *A. hebraeum*, *R. appendiculatus* and *R. evertsi evertsi* were the most numerous tick species in North West Province. In Mpumalanga, *R. decoloratus* constituted more than 75% of the total tick population. In a five year survey conducted in the Eastern Cape Province, *R. decoloratus*, *A. hebraeum*, *R. appendiculatus* and *R. evertsi evertsi* were found to be the most common tick species infesting cattle (Marufu, 2008).

In contrast to these early findings, Horak (1999) observed *A. hebraeum*, *Haemaphysalis silacea*, *R. appendiculatus* and *R. glabroscutatum* to be most prevalent tick species on yearling commercial cattle on Valley Bushveld. Muchenje *et al.* (2008) in an on station study comparing tick loads on Nguni, Angus and Bonsmara steers grazing on sweet rangeland revealed that *R. decoloratus*, *A. hebraeum*, *R. evertsi evertsi* and *Hyalomma* species were the most common tick infestations. With the exception of *Hyalomma* species, the same tick species composition observed by Muchenje *et al.* (2008), was found to infest cattle and goats in the communal areas of the Eastern Cape (Nyangiwe and Horak, 2007).

Many studies have focused on cattle in the commercial farming system and it is apparent that cattle management and tick control in this farming sector will differ considerably to that in communal farming areas (Bryson *et al.*, 2002). Few studies have focused on cattle kept by resource-poor farmers in communal areas. These studies however, did not compare the prevalence of the parasites in different breeds kept under communal farmer management and across different rangelands types. Other studies have focused on comparing tick loads in indigenous and exotic beef breeds under controlled conditions (Norval *et al.*, 1996; Muchenje *et al.*, 2008). No studies have focused on the comparison of tick loads in the indigenous and nondescript cattle under communal grazing management. Information on the tick loads of cattle can be used in conjunction with sero-diagnostic methods to estimate and compare the level of resistance of different cattle breeds to ticks (Wambura *et al.*, 1998; Mattioli *et al.*, 2000).

Tick distribution and occurrence differs with geographic distribution and vegetation type (Mtshali *et al.*, 2004). No efforts have been made to compare tick loads in cattle on sweet and sour rangelands. Sour rangeland occurs in areas with high water supply and denser vegetation cover (Ellery *et al.*, 1995) and it is more likely to have higher prevalence of ticks than the sweet rangeland which occurs in areas with low water supply and sparse vegetation cover. Comparing the prevalence of ticks in different rangeland types assists policy makers to design appropriate control programmes for each particular rangeland type (Marufu, 2008).

The most common tick species and the disease that they transmit to cattle in South Africa are shown in Table1 (Marufu, 2008).

Table 1: Ticks and the pathogens they transmit to cattle in South African provinces

Tick species	Pathogens transmitted	Geographical distribution (Provinces)
Rhipicephalus (Boophilus) spp	*Anaplasma marginale*, *Babesia bigemina*, *B. bovis*	Limpopo, Mpumalanga, Gauteng, North West, KwaZulu-Natal, Eastern and Northern Free State, Eastern Cape and Coastal strip of Southern and South-western Cape
R. appendiculatus	*Theileria parva* *T. annulata* *T. taurotragi* *T. mutans* *T. velifera*	Limpopo, Mpumalanga, KwaZulu-Natal, Western Cape
R. evertsi evertsi	*A. marginale*	Eastern half of Southern Africa, South-western Cape
Hylomma spp	*A. marginale*	Karoo, Southern Africa except coastal Eastern Cape, KwaZulu-Natal, Eastern Southern and South-western Cape
Amblyomma hebraeum	*Ehrlichia (Cowdria) ruminantium* *T. mutans*	Limpopo, Mpumalanga, KwaZulu-Natal and Eastern and Southern South-western Cape

Sources: Horak *et al.*, 1991; Walker, 1991; Norval *et al.*, 1992; Coetzer *et al.*, 1994.

1.5 Pathogenesis of *A. marginale*

Anaplasma marginale is known to infect only mature, circulating erythrocytes of domestic and wild ruminants *in vivo* (Figure 1A). However, recently *in vitro* and *in vivo* studies indicated that it also infects endothelial cells, which may have implications for both pathogenesis and immune mechanisms. *A. marginale* enters erythrocytes by endocytosis and resides in the membrane-bound vacuole, where it divides by binary fission. The membrane vacuole is derived from erythrocyte membrane and contains 4 to 8 organisms (Figure 1B). In the acute infection as much as 70% of erythrocytes may become infected (Živković, 2010).

During high rickettsiemias (bacteremia) multiple infections of individual erythrocytes are common. The incubation period varies with the infective dose and ranges from 7 to 60 days, with an average of 28 days. *A. marginale* leaves the host cell without disrupting it. Erythrocytes that are physically or chemically altered during the course of the disease are recognized by bovine reticulo-endothelial cells and phagocytized, which will result in the development of mild to severe anemia and icterus, without hemoglobinemia and hemoglobinuria (Živković, 2010).

The acute phase of the disease may also include symptoms such as high fever, dramatic weight loss, abortion, lethargy and often death in animals older than 2 years. Calves less than one year of age develop a relatively mild form of disease. Importantly, cattle that survive acute disease develop a lifelong persistent infection and serve as reservoirs for transmission to new susceptible hosts. Persistence is characterized by sequential rickettsemic cycles, occurring at approximately 5 week intervals, with peaks at 106 bacteria /ml of blood followed by a rapid decline when rickettsemia is controlled by a specific immune response. Persistently infected cattle have a lifelong immunity and are resistant to clinical onset of the disease on challenge exposure (Živković, 2010).

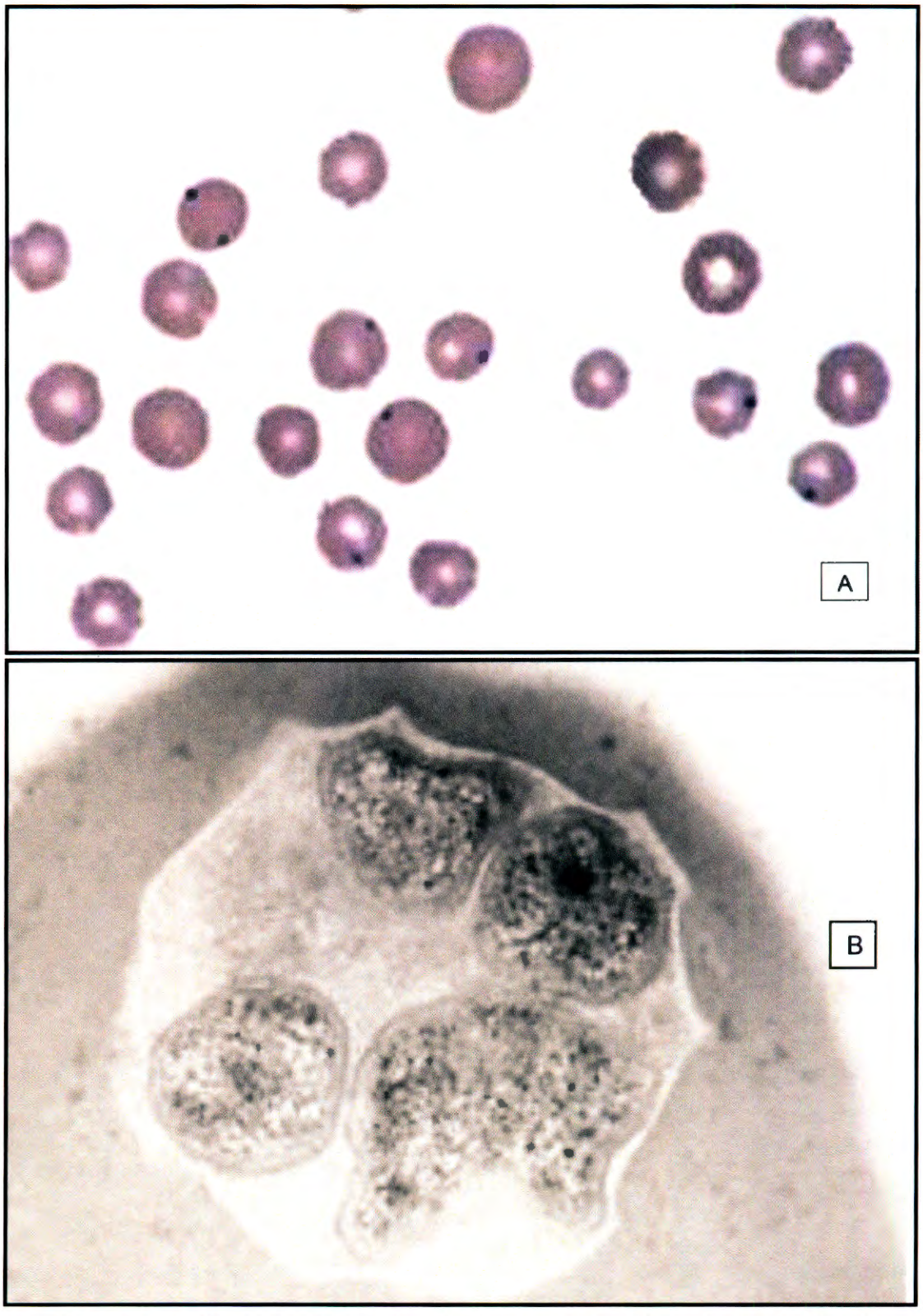


Figure 1: Bovine erythrocytes infected with *A. marginale*. **(A)** Inclusion bodies located at the periphery of the erythrocyte in a stained blood smear. **(B)** Electron micrograph of *A. marginale* inclusion that contains four organisms (Živković, 2010).

1.6 Transmission

Transmission of *A. marginale* can be both mechanical by biting flies or blood-contaminated fomites and biologically by ticks. Mechanical transmission frequently occurs via blood-contaminated fomites, including needles, dehorning saws, nose tongs, tattooing instruments, ear-tagging devices, and castration instruments. Mechanical transmission by arthropods has been reported for bloodsucking diptera of the genera *Tabanus*, *Stomoxys*, and mosquitoes. This form of mechanical transmission is considered to be the major route of dissemination of *A. marginale* in areas of Central and South America and Africa where tick vectors do not occur and where *Rhipicephalus (Boophilus) microplus*, the tropical cattle tick, does not appear to be a biological vector of *A. marginale*. In areas of the United States where geographic isolates of *A. marginale* are not infective for ticks or where ticks have been eradicated by fire ants, mechanical transmission appears to be the major mode of *A. marginale* transmission (reviewed by Kocan *et al.*, 2010).

In addition to mechanical and biological transmission, *A. marginale* can be transmitted from cow to calf transplacentally during gestation. For example, a 15.6% prevalence rate of *in utero* transmission of *Anaplasma* infections was reported in South Africa Transplacental transmission of anaplasmosis may therefore contribute to the epidemiology of this disease in some regions. This obligate intracellular pathogen can be transmitted biological by ticks, and approximately 20 species of ticks have been shown experimentally to transmit *A. marginale* worldwide. Tick transmission can occur from stage to stage (transstadial) or within a stage (intrastadial), while transovarial transmission from one tick generation to the next does not appear to occur. Interstadial transmission of *A. marginale* has been demonstrated by the three-host ticks *Dermacentor andersoni* and *D. variabilis* in the United States and by *R. simus* in South Africa. The one-host tick *R. annulatus* transmits *A. marginale* in Israel, Central America, South America, and Mexico (reviewed by Kocan *et al.*, 2010).

Intrastadial transmission of *A. marginale* is effected by male ticks. Recent studies have demonstrated that male *Dermacentor* ticks may play an important role in the biological transmission of *A. marginale* because they become persistently infected with *A. marginale* and can transmit the parasite repeatedly when they transfer among cattle. Male ticks therefore also serve as reservoirs of *A. marginale* along with persistently infected cattle. Transmission of *A. marginale* by male ticks may be an important mechanism of transmission by one-host ticks, including *Rhipicephalus (Boophilus)* spp. and *D. albipictus*. However, it was shown recently that the co-feeding of adult *Dermacentor* spp. does not appear to influence the dynamics of *A. marginale* transmission (reviewed by Kocan *et al.*, 2010).

1.7 Life cycle of *A. marginale*

The life cycle of *A. marginale* in ticks is complex and well-coordinated with the tick feeding cycle (Figure 2). Infected erythrocytes are ingested by ticks with a blood meal and the first sites of infection are gut and malpighian tubule cells. During the subsequent feeding many other tissues, including salivary glands, become infected from where *A. marginale* can be transmitted to the vertebrate host. At each site of infection two stages of *A. marginale* occur within a membrane bound vacuole in the tick cell cytoplasm (Živković, 2010).

The first form seen within *A. marginale* colonies is the reticulated (vegetative) form, which divides by binary fission and results in formation of large colonies containing hundreds of organisms. The reticulated forms are then transformed into dense forms, which are the infective form and can survive for a short time outside of cells. Cattle become infected when the dense form is transmitted during tick feeding via the salivary glands. Ticks are able to acquire infection after feeding on persistently infected animals with a very low level of rickettsemia. Moreover, once ticks acquire the infection the biological replication of the organism within the ticks makes up for the initial low infective dose (Živković, 2010).

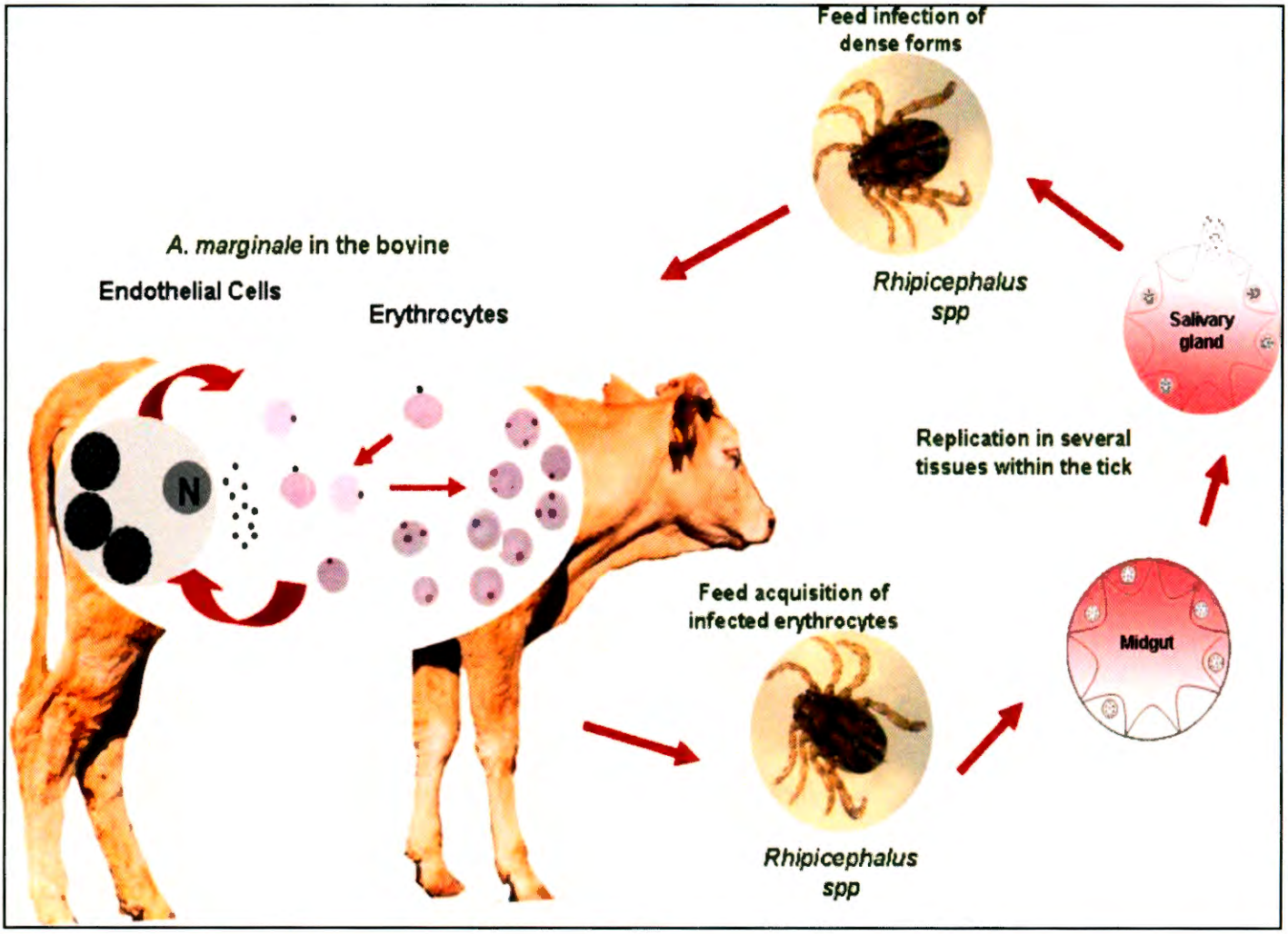


Figure 2: The life cycle of *A. marginale*. The cycle is modified from Kocan (1999), to include the replication of the rickettsia in endothelial cells (Carreño *et al.*, 2007) and *Rhipicephalus microplus* ticks present in Mexico (Rodríguez *et al.*, 2009).

1.8 Geographic distribution and economic importance of anaplasmosis

The distribution of anaplasmosis may be expected to continue to change in part as a result of global warming, which may influence the movement of the tick hosts (Jonsson and Reid, 2000. Concerns regarding the transmission of infectious agents between wildlife and domestic livestock are increasing especially in areas where free-ranging wildlife and cattle share common grazing grounds (Chomel *et al.,* 1994). Bovine anaplasmosis is endemic almost anywhere the world from the Far East to Australia, Africa, Europe and the Americas (Wen *et al.,* 2002; Ziam and Benaouf, 2004; Naranjo *et al.,* 2006; Vidotto *et al.,* 2006; Stevens *et al.,* 2007). Bovine anaplasmosis causes important economic loss in most countries, mainly due to the high morbidity and mortality in susceptible cattle herds (Marufu, 2014).

Losses due to anaplasmosis are measured through several parameters: low weight gain, reduction in milk production, abortion, the cost of anaplasmosis treatments, and mortality. However, few controlled studies have been carried out to determine the exact annual loss caused by anaplasmosis. The most important economic constraint of anaplasmosis to cattle production in the tropics is on public or private programs for genetic improvement of cattle. Imported *Bos Taurus* cattle brought from temperate nations to the tropics for breed improvement are highly susceptible to tick-borne diseases (TBD), and often do not survive to become part of planned production programs. This constraint is a notable reality for programs for the improvement of cattle in most Latin American countries (Marufu, 2014).

The lack of accurate data on the epidemiology of ticks and TBD makes it difficult to determine their impact. The complexity of determining the direct and indirect economic impact of ticks and TBD, and their control is reflected in the fact that only rough estimates are available for the cost of some of the components. Table 2 shows the estimated costs of ticks and TBD to cattle production in different countries. Although a fairly crude estimate, these values may help to comprehend the importance of ticks and TBD in cattle. These estimates, however, expose the need for more studies on the determination of the economic impact of ticks and TBD in the cattle industry especially in the developing world (Marufu, 2014).

Table 2: Estimated costs of tick and tick-borne diseases to cattle production

Country	Costs (US$)	Reference
Global	13-18 billion	de Castro (1997)
Southern Africa	31.6 million	Minjauw *et al.*, (1998)
Australia	4.09 million	Jonsson *et al.*, (2001)
India	498.7 million	Minjauw and Mcleod (2003)
Australia	170 - 200 million	Playford (2005), Sackett *et al.*, (2006)
Brazil	800 million	Martinez *et al.*, (2006)

Losses are partly due to the direct effects of ticks on cattle, such as damaged hides and skins, anaemia, reduced body weight gains and milk yield, tick toxicoses and mortalities (Gates and Wescott, 2000; Turton, 2001; Mtshali *et al.*, 2004; Kaufman *et al.*, 2006). The damage caused by tick bites also diminishes the value of skins and hides for the manufacture of leather. Ticks with long mouth parts may induce abscesses because of secondary bacterial infections. Depending on the site of infestation, these abscesses can lead to lameness or mastitis resulting in the drop in milk production and subsequent increase in calf mortalities (Marufu, 2014). A large component of the economic cost of ticks in cattle is the application of control measures to reduce infestations (de Castro, 1997; Porto Neto *et al.*, 2011). Conventional tick control is based on the application of acaricides. The practice of intensive tick control spread rapidly throughout Africa following the introduction of imported cattle breeds and, in South Africa, it was enforced through legislation. There are few global reports on the costs involved in tick control and TBD treatments (Marufu, 2014).

Financial analysis revealed that spraying cattle with acaricide twice a week yields a return of 244% and maximised financial benefits to the farmer (Mukhebi *et al.,* 1989). Jonsson *et al.,* (2001), however, estimated the total costs of tick control to contribute up to 49% of the total costs of ticks and TBD on the dairy industry in Australia. Expenditures for tick control were estimated at US$ 8.43, 13.62 and 21.09 per animal per year for plunge dipping, hand spraying and pour-on, respectively (D'haese *et al.,* 1999). The mean annual cost of ticks and TBD control per animal in pastoral and ranch herds was estimated to be US$4.54 (Ocaido *et al.,* 2009). The development of new acaricides is also a lengthy and costly process leading to increasing cost of the newer products. Regular dipping has also led to the loss of resistance to ticks and enzootic stability to TBD. Significant losses also arise indirectly due to the important role of ticks in the transmission of TBD (Marufu, 2014).

1.9 Diagnosis of anaplasmosis

The diagnostic assays (excluding clinical findings) used to identify *A. marginale* can be classified as microscopic, serologic assays such as Indirect fluorescent antibody (IFA), complement fixation (CF) test, capillary, agglutination assay, card agglutination test (CAT), indirect fluorescent antibody (IFA) test, as well as various enzyme linked immunosorbent assays (ELISA) such as a cELISA, indirect ELISA and dot ELISA. The preferred serological tests are cELISA and CAT and molecular diagnostic assay using polymerase chain reaction (PCR) of whole blood samples (revised in Aubrey and Geale, 2011). However, the use of microscopy requires careful examination in cases of low level of parasitemia and *A. marginale* appear as dense, rounded and deeply stained intraerythrocytic bodies, approximately 0.3– 1.0 µm in diameter. Most of these bodies are located on or near the margin of the erythrocyte. This feature distinguishes *A. marginale* from *A. centrale,* as in the latter most of the organisms have a more central location in the erythrocyte. It can be difficult to differentiate *A. marginale* from *A. centrale* in a stained smear, particularly with low levels of rickettsaemia (OIE Terrestrial Manual 2012).

Anaplasma infections usually persist for the life of the animal. However, except for occasional small recrudescences, *Anaplasma* cannot readily be detected in blood smears after acute rickettsaemia. Thus, a number of serological tests have been developed with the aim of detecting persistently infected animals. A feature of the serological diagnosis of anaplasmosis is the highly variable results with regard to both sensitivity and specificity reported for many of the tests from different laboratories. This is due at least in part to inadequate evaluation of the tests using significant numbers of known positive and negative animals. Importantly, the capacity of several assays to detect known infections of long-standing duration has been inadequately addressed. An exception is cELISA, which has been validated using true positive and negative animals defined by nested PCR (Torioni De Echaide *et al.*, 1998), and the card agglutination assay, for which relative sensitivity and specificity in comparison with the cELISA has been evaluated (Molloy *et al.*, 1999).

It should be noted that there is a high degree of cross-reactivity between *A. marginale* and *A. centrale*, as well as cross-reactivity with both *A, phagocytophilum* and *Ehrlichia* spp. in serological tests (Dreher *et al.*, 2005; Al-Adhami *et al.*, 2011). While the infecting species can sometimes be identified using antigens from homologous and heterologous species, equivocal results are obtained on many occasions (OIE Terrestrial Manual 2012).

1.9.1 Competetive enzyme –linked immunosorbent assay

A cELISA using a recombinant antigen termed rMSP5 and MSP5-specific monoclonal antibody (MAb) have proven very sensitive and specific for detection of *Anaplasma*-infected animals (Hofmann-Lehmann *et al.*, 2004; Stik *et al.*, 2007; Reinbold *et al.*, 2010). All *A. marginale* strains tested, along with *A. ovis* and *A. centrale*, express the MSP5 antigen and induce antibodies against the immunodominant epitope recognised by the MSP5-specific MAb. A recent report suggests that antibodies from cattle experimentally infected with *A. phagocytophilum* will test positive in the cELISA (Dreher *et al.*, 2005).

However, in another study no cross-reactivity could be demonstrated, and the MAb used in the assay did not react with *A. phagocytophilum* MSP5 in direct binding assays (Stik *et al.*, 2007). Recently, cross reactivity was demonstrated between *A. marginale* and *Ehrlichia* spp, in naturally and experimentally infected cattle (Al- Adhami *et al.*, 2011). Earlier studies had shown that the cELISA was 100% specific using 261 known negative sera from a non-endemic region, detecting acutely infected cattle as early as 16 days after experimental tick or blood inoculation, and was demonstrated to detect cattle that have been experimentally infected as long as 6 years previously (Knowles *et al.*, 1996). In detecting persistently infected cattle from an anaplasmosis-endemic region that were defined as true positive or negative using a nested PCR procedure, the rMSP5 cELISA had a sensitivity of 96% and a specificity of 95% (Torioni De Echaide *et al.*, 1998).

1.9.2 Card agglutination test

The advantages of the card agglutination test (CAT) are that it is sensitive, may be undertaken either in the laboratory or in the field, and gives a result within a few minutes. Nonspecific reactions may be a problem, and subjectivity in interpreting assay reactions can result in variability in test interpretation. In addition, the CAT antigen, which is a suspension of *A. marginale* particles, can be difficult to prepare and can vary from batch to batch and laboratory to laboratory. Splenectomised calves are infected by intravenous inoculation with blood containing *Anaplasma*-infected erythrocytes. When the rickettsaemia exceeds 50%, the animal is exsanguinated, the infected erythrocytes are washed, lysed, and the erythrocyte ghosts and *Anaplasma* particles are pelleted. The pellets are sonicated, washed, and then resuspended in a stain solution to produce the antigen suspension (OIE Terrestrial Manual 2012).

1.9.3 Complement fixation test

The complement fixation (CF) test has been used extensively for many years; however, it shows variable sensitivity (ranging from 20 to 60%), possibly reflecting differences in techniques for antigen production, and poor reproducibility. In addition, it has been demonstrated that the CF assay fails to detect a significant proportion of carrier cattle (Bradway *et al.*, 2001). It is also uncertain as to whether or not the CF test can identify antibodies in acutely infected animals prior to other assays (Molloy *et al.*, 1999; Coetzee *et al.*, 2007). Therefore, the CF test is no longer recommended as a reliable assay for detecting infected animals (OIE Terrestrial Manual, 2012). The use of the CF test could result in infected cattle being assigned a negative result, which could lead to introduction of persistently infected animals with false-negative results into populations of completely naïve cattle (Coetzee *et al.*, 2007).

1.9.4 Indirect fluorescent antibody test

Other serological tests are generally preferred to the IFA test, because of the number of tests that can be performed daily by one operator. A serious problem encountered with the IFA test is non-specific fluorescence attributed to antibodies adhering to infected erythrocytes. Non-specific fluorescence due to antibodies adhering to infected erythrocytes can be reduced by washing the erythrocytes in an acidic glycine buffer before antigen smears are prepared (OIE Terrestrial Manual 2012).

1.9.5 Polymerase chain reaction

Polymerase chain reaction (PCR) has been the most commonly used method for the diagnosis of *A. marginale* infections (Lew *et al.*, 2002, Shkap *et al.*, 2002). Most of the PCR assays target the *msp4* and *msp1a* genes for differentiating strains of *A. marginale*, which is a useful method for tracking the origin of the outbreak (Bowie *et al.*, 2002; Lew *et al.*, 2002, de la Fuente *et al.*, 2005d; 2007a; Mtshali *et al.*, 2007, Cabezas-Cruz *et al.*, 2013; Mutshembele *et al.*, 2014).

Polymerase chain reaction has been shown to reliably detect *Anaplasma* at the lowest levels of persistent rickettsemia (Lew *et al.*, 2002, Shkap *et al.*, 2002). However, other studies have shown that PCR fails to detect the infection of *A. marginale* at low levels or during the early stages of infection (Molad *et al.*, 2009). The molecular diagnosis of *Anaplasma* infection by PCR using whole blood has become readily available. Various genes are useful for organism identification at genus and species level. Among the most used are the phylogenetically reliable marker genes such as *16S RNA* (Warner and Dawson, 1996; Wen *et al.*, 2002), *gltA* (Inokuma *et al.*, 2001) and *groEL* (Yu *et al.*, 2001).

Polymerase chain reaction assays targeted at the *Anaplasma msp4* and/or *msp1a* genes have been used to differentiate isolates of *A. marginale*, which is useful to track the origin of an outbreak, and to differentiate between different species of *Anaplasma* such as *A. marginale* and *A. centrale* (de la Fuente *et al.*, 2001a; Lew *et al.*, 2002). Because *msp5* expresses epitopes that are conserved among widely divergent strains of *A. marginale*, *A. centrale* and *A. ovis* (Visser *et al.*, 1992) and even A. phagocytophilum (Alleman *et al.*, 2006), a PCR assay targeted only at the *msp5* gene would be expected to have significant specificity issues when used in populations where animals could be infected with other *Anaplasma* spp.

Amplification of related organisms by non-specific primers has been shown to result in false-positive reactions. However, sequencing and sequences comparison may be a solution for this problem. Conversely, false-negatives may occur if extraction procedures fail to remove PCR inhibitors present in a blood sample or if the level of circulating rickettsemia falls below the level of assay. The main problem with the above genes is that they are conserved at species levels so strains identification becomes challenging or impossible using these genes. An alternative is the use of variable genes. The *msp1a* gene for *A. marginale* have become the most used tool for testing strains variability in these two pathogens (Palmer *et al.*, 2001; Almazán *et al.*, 2008; Ruybal *et al.*, 2009).

Nucleic-acid-based tests to detect *A. marginale* infection in carrier cattle have been developed although not yet fully validated. The analytical sensitivity of PCR-based methods has been estimated at 0.0001% infected erythrocytes, but at this level only a proportion of carrier cattle would be detected. A nested PCR has been used to identify *A. marginale* in carrier cattle with a capability of identifying as few as 30 infected erythrocytes per ml of blood, well below the lowest levels in carriers. However, nested PCR poses significant quality control and specificity problems for routine use (Torioni De Echaide *et al.,* 1998).

Real-time PCR based on *msp1b* gene was successfully developed for detection and quantification of *A. marginale* DNA in blood of naturally infected cattle (Carelli *et al.,* 2007; Decaro *et al.,* 2008; Reinbold *et al.,* 2010). The assay was shown to be distinctively sensitive and specific as there were no cross-reactions with other haemoparasites of ruminants (*A. centrale*, *A. bovis*, *A. phagocytophilum*, *B. bigemina* and *T. buffeli*). Then *A. marginale* real-time PCR was modified by adding a primer/probe set specific for *A. centrale,* thus obtaining a duplex assay for simultaneous detection of both *Anaplasma* spp. in the same reaction. Duplex real-time PCR has proven to be a powerful tool for differentiating between closely related infectious agents (Decaro *et al.,* 2006a; 2006b) and between vaccine and field strains of the same genotype (Decaro *et al.,* 2006c; Elia *et al.,* 2008).

1.10 Treatment, prevention and control strategies

Control measures for anaplasmosis vary with different geographic locations, including maintenance of anaplasmosis free herds, control of tick vectors, administration of antibiotics and vaccination (de Waal, 2000). Dairy and beef cattle farmers have relied on dipping measures for control of tick infestation; however, in areas where tick vectors are well-established, the continuous exposure to ticks leads to endemic stability situation (de Waal, 2000). Vector control is labour intensive, expensive and environmental pollution is also a major concern.

Prophylaxis has been through administration of antibiotics. Chemotherapy is intended for prevention of clinical anaplasmosis but it does not prevent cattle from becoming persistently infected with *A. marginale*; however, cattle receiving antibiotics therapy may not be cleared of infection. Tetracycline administration is accompanied by disadvantages of expenses, and the demand of continuous feeding and also the risk of development of resistant *Anaplasma* organisms, although the resistance of *A. marginale* to antibiotics has not been reported (de Waal, 2000). Kocan *et al.,* (2003) reviewed the progression and history of vaccine development which included live and killed vaccines. Both vaccines have relied on *A. marginale* infected bovine erythrocytes as source of antigen.

The live and killed vaccines induce protective immunity that diminishes the clinical signs, but does not prevent cattle from becoming persistently infected with *A. marginale.* Live vaccine (*A. centrale*), which was introduced by Sir Arnold Theiler has been used widely in South Africa but the vaccination renders partial protection against *A. marginale* and its success vary according to genotypes of *A. marginale* (Brown *et al.,* 1998; de Waal, 2000). In contrast, trials conducted in South America and Africa with heterologous strains showed low efficacy of this vaccine (Turton *et al.,* 1998; Bock and de Vos, 2001; Dark *et al.,* 2011).

Major surface proteins (MSP) have been manipulatively used experimentally as vaccine candidates against *A. marginale* infections (Palmer *et al.,* 1987; Allred *et al.,* 1990). MSP1a has an ability to induce T-cell response and contains conserved B-cell epitope in the repeated peptides that is recognized by immunized and protected cattle (Kocan *et al.,* 2003, de la Fuente *et al.,* 2003b). Experiments with recombinant MSP1a have shown partial protection against anaplasmosis in cattle (Torina *et al.,* 2014).

Recent studies have identified members of the *pfam01617* from *msp2* family to be possible vaccine candidates because most of them are found in cross-linked surface antigen complexes with the other members of *pfam01617* encoding conserved outer membrane proteins, which are expressed in *A. marginale* (Noh *et al.*, 2006; Dark *et al.*, 2011).

1.11 *Anaplasma marginale* major surface proteins and their role in host-vector-pathogen interactions

The outer membrane proteins of *A. marginale* have been the focus of direct research to obtain an improved vaccine against bovine anaplasmosis (Palmer *et al.*, 1999). Immunization with purified outer membranes induces protection against acute *A. marginale* infection and disease (Tebele *et al.*, 1991). Various major outer membranes have been described based on the proteomic and genomic approach, and 21 proteins were identified within the outer membrane immunogen (Lopez *et al.*, 2005). Some well-characterized outer membranes proteins designated major surface proteins (MSPs): *msp1a, msp1b, msp2, msp3, msp4* and *msp5* were evaluated as potential candidates for antigens of vaccine production and diagnostic evaluations (Visser *et al.*, 1992; Oberle *et al.*, 1993; McGarey *et al.*, 1994; Alleman and Barbet, 1996; Kano *et al.*, 2002; Garcia-Garcia *et al.*, 2004a).

1.11.1 The MSP1 Complex

The relationship of the rickettsia with the bovine host is complex as it has interaction with the erythrocyte and endothelial cells (Carreño *et al.*, 2007). A number of rickettsial protein and the bovine immune system have not been studied in detail yet, there is no final solution (vaccine or otherwise). Several major surface proteins (MSPs) have been identified in *Anaplasma spp.*, which have been most extensively characterized in *A. marginale* (Palmer *et al.*, 1985; de la Fuente *et al.*, 2001a, 2005b; Kocan *et al.*, 2003; 2004; 2008). *Anaplasma* MSPs are involved in interactions with both vertebrates and invertebrates host cells (de la Fuente *et al.*, 2001a; 2005a; Kocan *et al.*, 2003; 2004; Brayton *et al.*, 2005; Dunning Hotopp *et al.*, 2006; Nelson *et al.*, 2008), and are likely to evolve more rapidly than other genes because they are subjected to selective pressures exerted by host immune systems.

Six MPSs have been identified in *A. marginale* from cattle and ticks of which three, *msp1a, msp4* and *msp5*, are coded by single gene and do not vary within isolates (Bowie *et al.*, 2002). The other three, *msp1b, msp2* and *msp3*, are from multigene families and may vary antigenetically in persistently infected cattle (Barbet *et al.*, 2000; Kocan *et al.*, 2000; Bowie *et al.*, 2002). Major surface protein 1 is a heterodimer composed of two structurally unrelated polypeptides: *msp1a,* (100kDa), which is encoded by a single gene, *msp1a* (Lew *et al.*, 2002) and *msp1b* (105 kDa), which is encoded by at least two genes, *msp1β1* and *msp1β2* (Barbet *et al.*, 1987; Viseshakul *et al.*, 2000; Camacho-Nuez *et al.*, 2000; Bowie *et al.*, 2002; Macmillan *et al.*, 2006). The *msp1a,* and *msp1b* proteins form non-covalent dimers and are exposed on the surface of *A. marginale* (Barbet *et al.*, 1987). However, only a single *msp1b* protein, *msp1b1*, was identified within the MSP1 complex (Macmillan *et al.*, 2006).

The molecular weight of *msp1a* varies in size among isolates due to different numbers of tandemly repeated 23-31 amino acid peptides and has been used for identification of geographic strains (Allred *et al.*, 1990; de la Fuente *et al.*, 2001a; 2003a; 2005a; 2007a). The *msp1a,* tandem repeats are located after a conserved decapeptide in the amino terminal region of the protein and are exposed extracellularly for interaction with host cell receptors (de la Fuente *et al.*, 2003a). The frequency of variable amino acid positions within geographic isolates is higher in this region than in the rest of the protein (de la Fuente *et al.*, 2001a).

Functionally, *msp1a* was shown to be an adhesin for bovine erythrocytes and tick cells and the adhesin domain was identified on the variable N-terminal region containing the repeated peptides (McGarey and Allred, 1994; McGarey *et al.*, 1994; de la Fuente *et al.*, 2001b). It was also shown to be involved in the transmission of *A. marginale* by *Dermacentor* spp. (de la Fuente *et al.*, 2001c) and to be differentially regulated in tick cells and bovine erythrocytes (Garcia-Garcia *et al.*, 2004b). The *msp1a*, although variable in the number of repeated peptides, induces strong T-cell responses and contains a conserved B-cell epitope in the repeated peptides that is recognized by immunized and protected cattle (Kocan *et al.*, 2003; de la Fuente *et al.*, 2005a). The *msp1a* also contains Th1 cell epitopes in the conserved region, which may be involved in immunoprotection (Brown *et al.*, 2001).

The *msp1b*, encoded by at least two genes, *msp1b* and *msp1b2,* is polymorphic among geographic isolates of *A. marginale* (Barbet *et al.*, 1987; Camacho Nuez *et al.*, 2000; Viseshakul *et al.*, 2000; Bowie *et al.*, 2002). Although *msp1b* is encoded by a multigene family, only small variations in protein sequences of *msp1b* and *msp1b2* were observed during the life cycle of the *Rickettsia* in cattle and ticks (Bowie *et al.*, 2002). This protein, which forms a complex with *msp1a,* is an adhesin for bovine erythrocytes (McGarey and Allred, 1994; McGarey *et al.*, 1994). However, MSP1b was recently demonstrated to be adhesin only for bovine erythrocytes and did not prove to be an adhesin for tick cells (de la Fuente *et al.*, 2001a).

The *A. marginale* surface is characterized by the presence of two highly abundant and closely related outer membrane proteins Major Surface Protein 2 (*msp2*) and (*msp3*) (Brayton *et al.*, 2001). The predominant immune responses are generated against these two proteins (McGuire *et al.*, 1991; Brown *et al.*, 2001; 2003). However, both *msp2* and *msp3* are highly antigenically variable, both within an infection and between strains (McGuire *et al.*, 1984; French *et al.*, 1998; 1999; Rodriguez *et al.*, 2005).

Thus, while antibody response to *msp2* and *msp3* antigenic variants plays a key role in how persistent infection is established and the population strain structure, these abundant surface proteins are not targets for development of a widely cross-protective vaccine and anti- *msp2/msp3* immune responses do not associate with protective efficacy of the outer membrane vaccine (Palmer *et al.*, 2009; Noh *et al.*, 2010). Using genomic and proteomic approaches, the minor components of the outer membrane protein immunogen have been identified (Noh *et al.*, 2006; 2008; Brayton *et al.*, 2005; Lopez *et al.*, 2005; 2008; Agnes *et al.*, 2011).

Although markedly less abundant, these minor proteins are invariant during infection and highly conserved among strains-thus representing much more attractive targets for vaccine development. Importantly, the proteomic identification within the outer membrane immunogen and the bioinformatics prediction of surface localization was confirmed for a subset of these proteins by surface-specific cross-linking (Noh *et al.*, 2008). The isolated cross-linked surface protein complex induced protection equal to that of the native outer membrane immunogen (Noh *et al.*, 2008).

The *msp2* and *msp3* are both encoded by large polymorphic, multigene families (Palmer *et al.*, 1994; Alleman *et al.*, 1997). The *msp2* sequence and antigenic composition varies during cyclic rickettsemia in cattle (French *et al.*, 1998; 1999; Barbet *et al.*, 2001) and in persistently infected ticks (de la Fuente *et al.*, 2001a). MSP2 is encoded on a polycistronic mRNA. The *msp2* gene within the expression site is polymorphic. The *msp2* encodes numerous amino acid sequence variants selected in bovine erythrocytic and tick salivary gland populations of *A. marginale* (French *et al.*, 1998; 1999; Barbet *et al.*, 2000; Brayton *et al.*, 2001; de la Fuente *et al.*, 2001a; Meeus and Barbet, 2001).

The *msp3* also varies in antigenic properties and structure among geographic isolates (Alleman and Barbet, 1996). The *msp2* and *msp3* are involved in the induction of a protective bovine immune response to *A. marginale* (Palmer *et al.*, 1999).

The *msp4* and *msp5* are encoded by single copy genes. Although *msp4* is highly conserved (Oberle and Barbet, 1993; de al Fuente *et al.*, 2003a), information about its function is not available.

The *msp5* is highly conserved surface protein that has been proven effective as a diagnostic antigen and used in a competitive enzyme linked immunosorbent assay (cELISA) commercially available in the United States (Torioni De Echaide *et al.*, 1998). The function of *msp5* is also unknown. The *msp2* operon-associated genes OpAG1, OpAG2, and OpAG3, have been identified in *A. marginale* and may encode for surface proteins (Lohr *et al.*, 2002).

1.11.2 Vector-pathogen relationships

1.11.2.1 *Anaplasma* MSPs and vector -pathogen interactions

The evolutionary history of vector-pathogen interactions could be reflected in the sequence variation of the *Anaplasma* major surface proteins. Previous studies demonstrated that *A. marginale msp1a*, but not *msp4* is under positive selection pressure (de la Fuente *et al.*, 2003a). Initial analysis of *A. marginale msp1a* and *Dermacentor variabilis 16S rDNA* sequences from various areas of the United States suggested tick-pathogen co-evolution (de la Fuente *et al.*, 2001d) consistent with the biological function of MSP1a involved in the transmission of *A. marginale* by ticks (de la Fuente *et al.*, 2001c). However, the genetic diversity of sequences complicates the study of tick-pathogen co-evolution (de la Fuente *et al.*, 2010).

Anaplasma marginale strains are apparently not transmissible by ticks and rely on mechanical transmission for completion of the life cycle in nature (de la Fuente *et al.*, 2005d; Hornok *et al.*, 2008). These facts pose the question of what evolutionary adaptations may have occurred to insure an efficient mechanical transmission of non-tick-transmissible *A. marginale* strains. Recently, Scoles *et al.*, (2005) concluded that biological transmission by ticks is more efficient than mechanical transmission of *A. marginale* (de la Fuente *et al.*, 2010).

The adhesin domain of *A. marginale msp1a* has been identified on the extracellular N-terminal region of the protein that contains the repeated peptides (de la Fuente *et al.*, 2003b). The binding of *msp1a* to tick cell extract (TCE) was observed within the *msp1a* tandem repeats, the negatively charged amino acids, aspartic acid (D) and glutamic acid (E), at position 20 demonstrated that peptide containing acidic amino acids D or E at position 20 bound to TCE, while peptides with G as the 20th amino acid were not adhesive to TCE (de la Fuente *et al.*, 2003b).

1.12 Phylogenetic relationships of geographic isolates of *A. marginale*

A phylogenetic tree is an estimate of the relationships among taxa (or sequences) and their hypothetical common ancestors (Nei and Kumar 2000; Felsenstein 2004; Hall 2011). Today most phylogenetic trees are built from molecular data: DNA or protein sequences. Originally, the purpose of most molecular phylogenetic trees was to estimate the relationships among the species represented by those sequences, but today the purposes have expanded to include understanding the relationships among the sequences themselves without regard to the host species, inferring the functions of genes that have not been studied experimentally (Hall *et al.*, 2009), and elucidating mechanisms that lead to microbial outbreaks (Hall and Barlow 2006) among many others. Building a phylogenetic tree requires four distinct steps: (Step 1) identify and acquire a set of homologous DNA or protein sequences, (Step 2) align those sequences, (Step 3) estimate a tree from the aligned sequences, and (Step 4) present that tree in such a way as to clearly convey the relevant information to others (Hall, 2013).

Phylogenetic analysis of *A. marginale* geographic isolates from the United States was performed using the genes *msp1a* and *msp4* (de la Fuente *et al.*, 2001d). The results of these analyses strongly support a southeastern clade of *A. marginale* composed of isolates from Virginia and Florida. Analysis of *16S rDNA* fragment sequences from the tick vector of *A. marginale*, *D. variabilis*, from various areas of the United States was performed and suggested coevolution of the vector and pathogen (de la Fuente *et al.*, 2001d).

Isolates of *A. marginale* from the United States also grouped into two clades, southern clade consisting of isolates from Florida, Mississippi, and Virginia, and a west-central clade consisting of isolates from California, Idaho, Illinois, Oklahoma, and Texas. Although little phylogeographic resolution was detected within any of these higher clades, *msp4* sequences appear to be a good genetic marker for inferring phylogeographic patterns of isolates of *A. marginale* on a broad geographic scale .

In contrast to the phylogeographic resolution provided by *msp4*, DNA and protein sequence variation from *msp1a* representing 20 New World isolates of *A. marginale* failed to provide phylogeographic resolution (de la Fuente *et al.*, 2007b). Most variation in *msp1a* sequences appeared unique to a given isolate. In fact, similar DNA sequence variation in *msp1a* was detected within isolates from Idaho and Florida and from Idaho and Argentina. These results suggest that the *msp1a* sequence may be rapidly evolving and that the *msp1a* gene may provide phylogeographic information only when numerous *msp1a* sequences from a given area are included in the analysis (de la Fuente *et al.*, 2007a).

The analysis of *msp1a* DNA and protein sequences demonstrated extensive genotypic variation among Oklahoma isolates of *A. marginale* and failed to provide phylogeographic resolution within Oklahoma or on a broader scale, including isolates from other United States and Latin America. Furthermore, analysis of codon and amino acid changes over the *msp1a* and *msp4* phylogenies provided evidence that *msp1a* but not *msp4*, is under positive selection pressure. These results suggest that even if *msp1a* sequences are rapidly evolving, MSP1a genotypes reflect the history of cattle movement more than the geographic distribution of *A. marginale* isolates. Research analysis suggest that different *A. marginale* genotypes are maintained within a herd in an area of endemic infection by independent transmission events and that infection with more than one genotype per host is prevented, a phenomenon described as infection exclusion (de la Fuente *et al.*, 2003a).

Therefore, if cattle movements import a new *A. marginale* genotype, it could be established by mechanical and/ or biological transmission to susceptible cattle. In regions with few cattle introductions, like Australia, little genotypic variation are found within *A. marginale* isolates (de la Fuente *et al.*, 2005a).

CHAPTER 2

OBJECTIVES OF THE STUDY

2.1 Statement of the problem

Anaplasma marginale genotypes have been well characterized and are highly variable in endemic areas worldwide (de la Fuente *et al., 2010*). *Anaplasma marginale* is quite host specific for ruminants and anaplasmosis occurs primarily in cattle; other ruminants may serve as reservoirs of infection. While *A. marginale* is transmitted biologically by ticks, mechanical transmission by blood-contaminated mouthparts of biting flies or fomites also frequently occurs. Mechanical transmission may be the only means of spreading anaplasmosis in areas where tick vectors are absent or are unable to transmit the local *A. marginale* strain (Kocan *et al.*, 2010).

Immunization of cattle with affinity- purified native MSP1 complex induced protective immunity in cattle that received homologous or heterologous challenge with *A. marginale* geographic isolates (Palmer *et al.*, 1987; 1989). The *msp1a* has been shown to have a neutralization sensitive epitope (Palmer *et al.*, 1987) and to be an *A. marginale* adhesin for both bovine erythrocytes and tick cells, while *msp1b1* is an adhesin only for bovine erythrocytes (McGarey *et al.*, 1994; McGarey and Allred, 1994; de la Fuente *et al.*, 2010). Killed vaccine marketed previously in the USA used *A. marginale* antigen that was partially purified from bovine erythrocytes (Brock *et al.*, 1965; Hart *et al.*, 1981; McCorkle-Shirley *et al.*, 1985). This vaccine was used effectively until it was removed from the market in 1999 (Kocan *et al.*, 2010).

These blood-derived killed vaccines reduced clinical anaplasmosis and were expensive to produce, difficult to standardize and often not cross-protective in widely separated geographic areas with different endemic *A. marginale* isolates. The blood-derived vaccines also bore risk of being contaminated with bovine cells or pathogens that frequently cause persistent infections in cattle (Palmer, 1989; Kocan *et al.*, 2000).

Despite the importance of anaplasmosis in South Africa and other African countries, *A. marginale* strains have not been genetically characterized in Africa (Mtshali *et al.*, 2007). Vaccination experiments with recombinant *msp1a* have resulted in partial protection against clinical anaplasmosis in cattle and reduced infection levels in ticks, thus supporting the inclusion of *msp1a* in vaccines for the control of bovine anaplasmosis (Kocan *et al.*, 2010).

The phylogenetic analysis of *A. marginale* strains using MSPs has been recently reviewed. These analyses suggest that MSPs may not be good markers for biogeographically studies on a global scale. However, they may be useful for strain comparison in given regions and could provide information about the evolution of host-pathogen and vector-pathogen relationships (Cabezas-Cruz *et al.*, 2013).

While a universal vaccine has yet to be developed for anaplasmosis, preliminary analysis of the phylogeographic relationships of *A. marginale* based on two major surface proteins (*msp1a* and *msp4*) has shown phylogeographic partitioning of parasite isolates in the United States (de la Fuente *et al.*, 2007a). This information supports the inclusion of several geographic isolates of *A. marginale* in development of vaccine formulations for the United States (Kocan *et al.*, 2010). Recent research has focused on MSPs that may be used to elucidate phylogeographic patterns of *A. marginale* (de la Fuente *et al.*, 2001a) and that are involved in interactions with vertebrate and invertebrate host cells (McGarey *et al.*, 1994; McGarey and Allred, 1994; de la Fuente *et al.*, 2001a).

These MSPs involved in host-pathogen interactions may evolve more rapidly than other nuclear genes because of selective pressures exerted by host immune systems (de la Fuente *et al.*, 2010). Of the six *A. marginale* MSPs that have been identified and characterized, only three (*msp1a, msp4* and *msp5*) are each encoded by a single gene. Because these MSPs do not appear to undergo antigenic variation in cattle or ticks, they were reported to be more stable genes for phylogenetic studies (Bowie *et al.*, 2002).

Of these three MSPs, *msp1a* has been reported to be an adhesin for bovine erythrocytes and tick cells and to effect infection and transmission of *A. marginale* by *Dermacentor* spp. ticks (McGarey *et al.*, 1994; McGarey and Allred, 1994; de la Fuente *et al.*, 2001b, 2001c). Although the specific function of *msp4* is currently is not known, the previous analysis of the gene from *A. marginale* isolates detected sufficient sequence variation to support its use in phylogeographic studies (de la Fuente *et al.*, 2001a). Therefore, *msp4* warranted further testing for use in detection of phylogeographic patterns among *A. marginale* isolates.

In a previous report of the phylogeographic relationships of *A. marginale* isolates from the United States (de la Fuente *et al.*, 2001a). It was found that DNA sequence variation in *msp4* provided phylogenetic resolution for intraspecific relationships among isolates (de la Fuente *et al.*, 2002b). In addition, when DNA sequence variation of *msp1a* was interpreted in combination with *msp4* phylogeny, important phylogeographic information became apparent. This study utilised MSPs to explore the epidemiology and phylogeny of *A. marginale* causing anaplasmosis in selected South African provinces.

2.2 Aim of the study

The aim of this study is to determine the prevalence, genetic distribution and phylogenetic relationship of *A. marginale* infecting South African-cattle by using *msp1a* and *msp4* gene

2.3 Objectives

2.3.1 To determine the prevalence of *A. marginale* infecting cattle in South Africa by using *msp1a* as a target gene.

2.3.2 To establish genetic diversity of *A. marginale* infecting cattle in South Africa using *msp1a* as marker gene.

2.3.3 To determine phylogenetic relationship of *A. marginale* infecting South African cattle by using *msp1a* and *msp4* genes.

CHAPTER 3

MOLECULAR DETECTION OF *ANAPLASMA MARGINALE* INFECTING CATTLE IN SOUTH AFRICA BY PCR TARGETING *msp1a* GENE

3.1 Introduction

Bovine anaplasmosis is a tick-borne hemoparasitic disease caused by *Anaplasma marginale* and is the most widely distributed disease of cattle in South Africa (de Waal, 2000; Ndou *et al.*, 2010). Economic losses associated with anaplasmosis in cattle include impaired production, mortality and expenses in control measures (Regassa *et al.*, 2003).

Ticks and the diseases they transmit have been identified as the major cause of widespread morbidity and mortality in cattle kept by smallfarm holders in the semiarid areas of South Africa (Dold and Cocks, 2001; Mapiye *et al.*, 2009). Poor cattle health management, resistance of ticks to most acaricides and the use of inappropriate cattle breeds (Dold and Cocks, 2001; Marufu *et al.*, 2011) have increased the prevalence of ticks and tick-borne diseases (TBD) in smallholder cattle herds.

Polymerase chain reaction detection methods have been developed. These tests/assays are extremely sensitive and specific in the detection of *A. marginale* infections in cattle (Lew *et al.*, 2002; de la Fuente *et al.*, 2005b; Molad *et al.*, 2006). Some of these assays, such as the real-time PCR, were developed to enable simultaneous detection and quantification of the *A. marginale* DNA in bovine blood. This is essential in supporting the clinical diagnosis, assessing the carrier status of the cattle and evaluating the efficacy of vaccines and antirickettsial drugs (Carelli *et al.*, 2007).

Among the genes encoding for major surface proteins in *A. marginale*, *msp1a* has been extensively used for strain characterization (Cabezas-Cruz *et al.*, 2013). The protein *msp1a* is involved in the interaction of the bacterium with vertebrate and invertebrate host cells (de la Fuente *et al.*, 2010). Several strains of *A. marginale* have been identified worldwide and these strains differ in their morphology, *msp1a* amino acid sequence, antigenic characteristics, and ability to be transmitted by ticks (de la Fuente *et al.*, 2007a; Estrada-Peña *et al.*, 2009).

The *msp1a* gene is thus a good candidate for the development of a generic *A. marginale* genotyping assay that could be applied globally and also to track strains (Lew *et al.*, 2002). The objective of the study was to determine the prevalence of *A. marginale* infecting cattle from Limpopo, Mpumalanga, North West, Gauteng, KwaZulu-Natal, Eastern Cape, Western Cape and Northern Cape provinces of South Africa by PCR using the *msp1a* gene.

3.2 Materials and methods

3.2.1. Study site and sample collection

Two hundred and eighty samples were collected from cattle in several South African provinces and the exact locations with coordinates are provided in Figure 5 and Table 4. Farms with different productive systems were chosen. Cattle were selected on a random basis following their owner's approval for the collection of blood samples. Blood samples were obtained from cattle and no stratification according to age or sex was applied. No vaccination histories were available for these animals. Tail venupuncture was performed using an 18 gauge needle and blood samples were collected into sterile vacutainer tubes (10 ml) with EDTA and were kept at 4 °C until delivered at the laboratory.

3.2.2 DNA extraction

The genomic DNA was extracted from cattle blood samples using ZR Genomic DNA™ Tissue Miniprep (Zymo Research, CA, USA). Total volume of blood sample was adjusted in a microcentrifuge tube with water to 100 µl prior to adding 95 µl of 2x Digestion buffer and 5 µl of Proteinase K. The mixture was thoroughly mixed and then incubated at 55 °C for 20 minutes. A volume of 700 µl of genomic lysis buffer was added to the tube and mixed thoroughly by vortexing.

The mixture was carefully transferred to a Zymo-Spin™ IIC column in a collection tube and centrifuged at 10,000 x *g* for one minute. A volume of 200 µl of DNA pre-wash buffer was added to the spin column in a new collection tube and centrifuged at 10,000 x *g* for one minute. A volume of 400 µl of g-DNA wash buffer was added to the spin column and centrifuged at 10, 000 x *g* for one minute. The spin columns were carefully transferred to a clean 1.5 ml microcentrifuge tube and 100 µl of DNA Elution buffer added to the spin column. The microcentrifuge tube was incubated for 5 minutes at room temperature, and then centrifuged at 8000 rpm for 30 seconds to elute the DNA. The eluted DNA was stored at -20 °C until used for studies. The concentration of DNA was measured using a NanoDrop® ND-1000 (NanoDrop Technologies Inc., Wilmington, USA).

3.2.3. Polymerase chain reaction (PCR)

A specific set of primers was used to amplify *msp1a* gene. A PCR was performed using the primers that were previously designed to amplify a fragment of *msp1a* gene, containing the tandem repeat region. Primers were synthesized and supplied by Inqaba Biotech (Inqaba Biotechnical Industries (Pty) Ltd, Pretoria, Gauteng, South Africa). The PCR reagents mixture consisted of total volume of 25 µl of Thermo Scientific DreamTaq Green PCR Master Mix (2X), 10 µM of forward and reverse primers, 1 µg Template DNA in 0.2 ml PCR tubes.

The PCR mixture was then amplified by DNA thermocycler (Bio-Rad T100™ DNA Thermal Cycler, Bio-Rad Laboratories, Johannesburg, South Africa). The amplification process involved an initial denaturation at 94 °C for 3 minutes followed by 40 cycles of denaturation at 94 °C for 1 minute at 94 °C, annealing at 65 °C for 1 minute , and an extension at 72 °C for 2 minutes. A cycle of a final extension was done at 72 °C for 7 minutes. The amplified products were then separated by electrophoresis using 1% agarose gel immersed in TBE buffer (89 mM Tris-Borate, 2 mM EDTA, pH 8). One kilo-base pair ladder marker was used to determine the size of the PCR products after staining with 0.5 µg/ml GelRed (Life Sciences, Fermentas GmbH, Germany) and visualization under UV illumination. Gel documentation was done and taking a photograph using a camera (Figure 3).

Table 3: Oligonucleotide sequences of primers used in this study

Primer designation	Oligonucleotide (5'-3')	References
1733F	TGTGCTTATGGCAGACATTTCC	Lew *et al.*, 2002
2957R	AAACCTTGTAGCCCCAACTTATCC	Lew *et al.*, 2002

3.2.4. Statistical analysis

Statistical analysis of the results were performed at the Department of Biostatistics and Statistics at the Medical Research Council, Pretoria Unit using the Fisher's exact and Chi-square test to test whether the prevalence is associated with the type of farm while chi-squared test was used to evaluate any significant difference (P=0.0001). The confidence intervals and standard errors of the prevalence of *Anaplasma marginale* infections in cattle were calculated at 95% using Kruskal-Wallis equality of population rank test. Pearson chi-square test was used to analyse PCR results based on frequency, row, column and cell percentage (Appendix 1).

3.3. Results

3.3.1 Detection of *A. marginale* infections

Two hundred and eighty blood samples were analysed by PCR technique. The primers 1733F and 2957R were used to amplify a fragment of *msp1a* gene, containing the tandem repeat region which varies among isolates. In this study PCR amplicons of variable sizes were obtained, ranging from 630 to 1100 bp corresponding to *msp1a* gene of different molecular weight as shown previously by Lew (2002). Two hundred and nine cattle blood samples out of 280 were PCR positive for *A. marginale* infection as shown in Appendix 1.

3.3.2 Molecular prevalence of *A. marginale* isolates infecting cattle in South African provinces

The PCR results of *A. marginale* infection ranged from 0% to 100% for samples from Mpumalanga and Gauteng. Eastern Cape comprised of 100%, KwaZulu-Natal 87.27%, Western Cape 84.62%, Limpopo 80% and North West 78.79% positive infectious rates. None of the Northern Cape samples were positive by PCR (Appendix 1).

3.3.3 Statistical prevalence of *A. marginale* isolates infecting cattle in South African provinces

Statistical analysis was conducted to obtain the prevalence rates using the PCR results. The highest relative prevalence was observed in KwaZulu-Natal with 23.0%, second Eastern Cape 19.1%, then Limpopo 15.3%, North West 12.4%, Western Cape 10.5%, Mpumalanga 10.1% and Gauteng 9.6%. These were obtained by considering relative rates of prevalence within the province. There was no significant statistical association between prevalence of *A. marginale* with the type of farm (communal or commercial).

PCR products with single bands as shown in Figure 3 were sequenced. The single PCR products from the seven different geographical regions including Limpopo, Mpumalanga, North West, Gauteng, KwaZulu-Natal, Eastern Cape and Western Cape of South African provinces *A. marginale* were sequenced and forty four gene sequences were submitted to GenBank (Appendix 2). BLAST analysis confirmed that the DNA results were indeed *A. marginale msp1a* gene and showed high identity (90-100%) with *A. marginale msp1a* deposited in GenBank.

Although *A. marginale* was detected in some of the samples as indicated by PCR results, most of the positive PCR samples from Mpumalanga, KwaZulu-Natal and Eastern Cape provinces, had non-specific bands and were not further analysed (Figure 3).

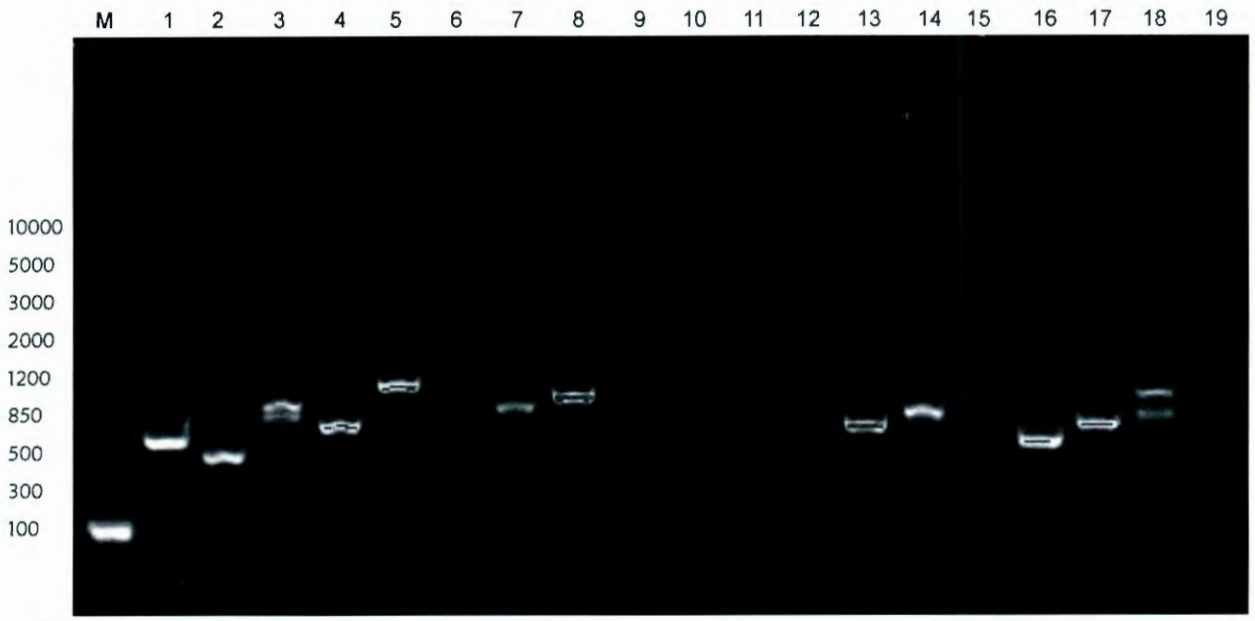


Figure 3: Gel image of 1% agarose gel electrophoresis of PCR amplification products obtained from *A. marginale* isolates from Western Cape Province in South Africa using 1733F and 2973R primers (Lew *et al.,* 2002). Lane Marker (M): GeneRuler™ I Kb DNA ladder, ready-to-use; Lane 1-5, 7-8, 13-14, 16-18 indicate positive results and lane 6, 9-12, 15 indicate negative results; lane 19 negative control.

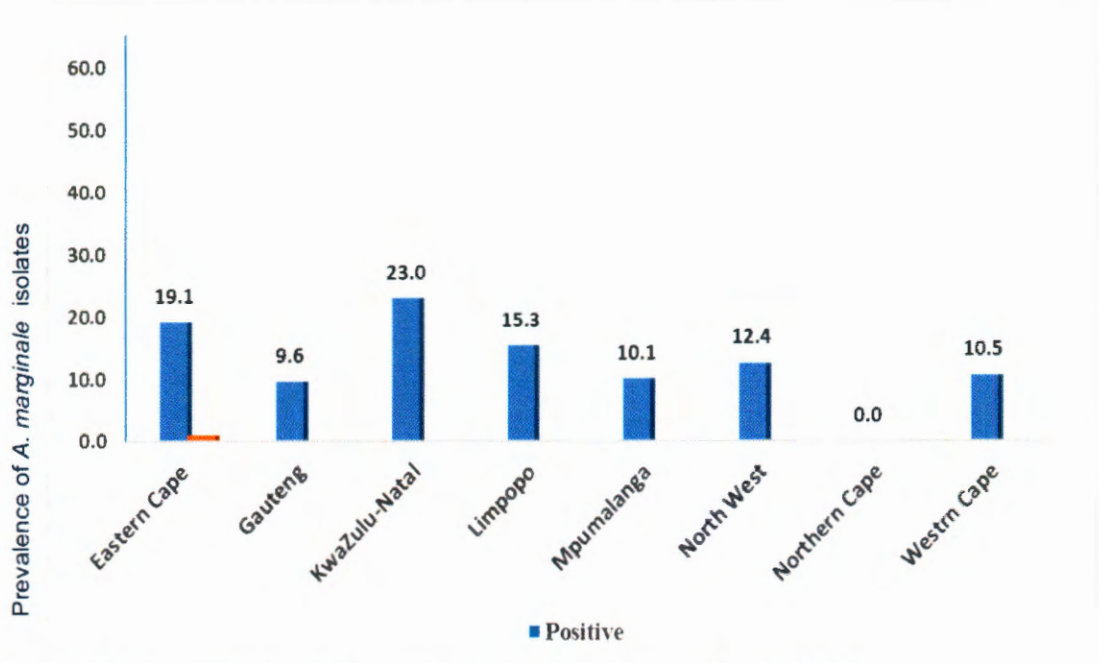


Figure 4: Prevalence of *A. marginale* by province in South Africa.

3.4 Discussion

It has been reported that bovine anaplasmosis is one of the most important cause of cattle mortalities in low-input farming areas in South Africa (Mapiye *et al.*, 2009). Polymerase chain reaction detection methods have been developed, which are extremely sensitive and specific for the detection of *A. marginale* infections in cattle (de la Fuente *et al.*, 2005b; Molad *et al.*, 2006). It is crucial to use sensitive and efficient molecular techniques to generate accurate information on the prevalence of *A. marginale,* which are crucial not only for developing appropriate control measures but for providing an understanding of host resistance in different cattle genotypes (Marufu, 2014).

In this study, a sensitive and specific *msp1a* PCR (Lew *et al.*, 2002) was used to determine the presence of *A. marginale* in eight different provinces of South Africa (Appendix 1). The results indicate the presence of *A. marginale* infection in these provinces with an overall prevalence of 76% except for Northern Cape in which the prevalence was 0%. This prevalence may be considered as high as in other endemic regions of the world.

Recently, 70% prevalence of *A. marginale* in a cattle herd was reported in Brazil (Pohl *et al.*, 2013). Another study by de la Fuente *et al.*, (2005a) showed a prevalence range from 25 to 100% for *A. marginale* infecting cattle herds in Italy. Results of this study confirm that bovine anaplasmosis is endemic and highly prevalent in South Africa.

The absence of positive samples in Northern Cape Province may be explained by special climate conditions in this area that limits the survival of tick vectors (Mtshali and Mtshali, 2013). However, the prevalence of bovine anaplasmosis at certain areas is a complex process also involving bovine population immunity to *A. marginale*, tick control practices, tick transmission, grazing system as well as the genetic variability of *A. marginale* at the population level (Hamou *et al.*, 2012).

Despite high prevalence of *A. marginale* in the studied area, the veterinary authority (State Veterinary officials) has not reported clinical disease in cattle, which is an indication of endemic stability. Endemic stability is an epidemiological state, in which clinical disease is scarce despite high levels of infection in the population (Coleman *et al.*, 2001; Jonsson *et al.*, 2012).

This is the first report of *A. marginale* prevalence at the national level based on molecular data. The results confirm endemicity and high prevalence of *A. marginale* in South Africa. Further investigations should address the influence of such high prevalence in cattle in South Africa. This will help in the understanding of the epidemiology and the use of appropriate control and prevention measures to reduce the high prevalence in cattle in South Africa before experiencing an outbreak.

CHAPTER 4

EPIDEMIOLOGY AND EVOLUTION OF THE GENETIC VARIABILITY OF *ANAPLASMA MARGINALE* IN SOUTH AFRICA

This work has been published as follows:

A.M. Mutshembele, A. Cabezas-Cruz, M.S. Mtshali, O.M.M. Thekisoe, R.C. Galindo and J. de la Fuente. 2014. Epidemiology and evolution of the genetic variability of *Anaplasma marginale* in South Africa. Ticks and Tick -borne diseases, 5 (6): 624-631.

4.1 Introduction

Bovine anaplasmosis is a non-contagious tick-borne disease caused by infection of cattle with *A. marginale*, an obligate intraerythrocytic bacterium classified in the family Anaplasmataceae, order Rickettsiales (Dumler *et al.*, 2001). This pathogen is transmitted biologically by ticks, mechanically by biting insects and blood-contaminated fomites and from cow to calf via transplacental transmission (Aubry and Geale, 2011). Five tick species have been shown experimentally to transmit *A. marginale* in South Africa, including *Rhipicephalus microplus, R. decoloratus, R. evertsi evertsi, R. simus* and *Hyalomma marginatum rufipes* (as reviewed by de Waal., 2000).

Acute disease in cattle is characterized by weight loss, fever, abortion, low milk production and in some cases death; the animals that recover from the disease become persistently infected and serve as reservoir of infection for mechanical transmission and biological transmission by ticks (Kocan *et al.*, 2003). Anaplasmosis is widespread in South Africa and, as estimated by de Waal. (2000), 99% of the total cattle population is at risk of acquiring *A. marginale* infection.

Currently, antimicrobial drugs are not available for the elimination of persistent infections in cattle. Although the World Organization for Animal Health proposed the use of enrofloxacin, imidocarb, and oxytetracycline for the elimination of persistent *A. marginale* infections in cattle, these antimicrobial drugs were not found to eliminate the persistent *A. marginale* infections (Coetzee *et al.*, 2005; 2006).

Vaccines have been used as an alternative method for control of anaplasmosis. A live vaccine using *A. marginale sub sp centrale* (*A. centrale*), a subspecies of relatively low pathogenicity, is available in South Africa, Australia, Israel and Latin America but this vaccine has proven to be only partially effective and reports of vaccination failure are not uncommon (Kocan *et al.*, 2003). While anaplasmosis still constitutes a problem for cattle production in South Africa, sales of vaccines for bovine anaplasmosis in South Africa have reduced from 800, 000 to 200, 000 doses over 22 years (1976-1998) (de Waal., 2000).

MSP1a has been shown to have potential use as a vaccine antigen because this protein contains both neutralization sensitive (Palmer *et al.*, 1987; Allred *et al.*, 1990) and immunodominant epitopes (Garcia-Garcia *et al.*, 2004b). Recently, use of MSP1a for vaccine development for *A. marginale* has regained new attention. The N-terminus tandem repeated region of this protein was used in immunization trials in cattle against *A. marginale* (Torina *et al.*, 2014) and laboratory animal models (Santos *et al.*, 2013; Silvestre *et al.*, 2014), and demonstrated promising results. MSP1a is one of six MSPs that have been described in *A. marginale*. This protein is encoded by a single-copy gene, *msp1a*, which is conserved during the multiplication of the parasite in cattle and ticks (Kocan *et al.*, 2003) and has been useful for epidemiological studies of *A. marginale* in various regions of the world (Ruybal *et al.*, 2009; Almazán *et al.*, 2008).

MSP1a contains tandem repeats at the N-terminus of the protein which present functional residues that serve as adhesins for bovine erythrocytes and tick cells, a prerequisite for infection of host cells (McGarey *et al.*, 1994; de la Fuente *et al.*, 2003a). While the tandem repeats of MSP1a are highly variable, repeats are commonly represented among worldwide strains (Cabezas-Cruz *et al.*, 2013) and have been shown to evolve under positive selection (de la Fuente *et al.*, 2003b).

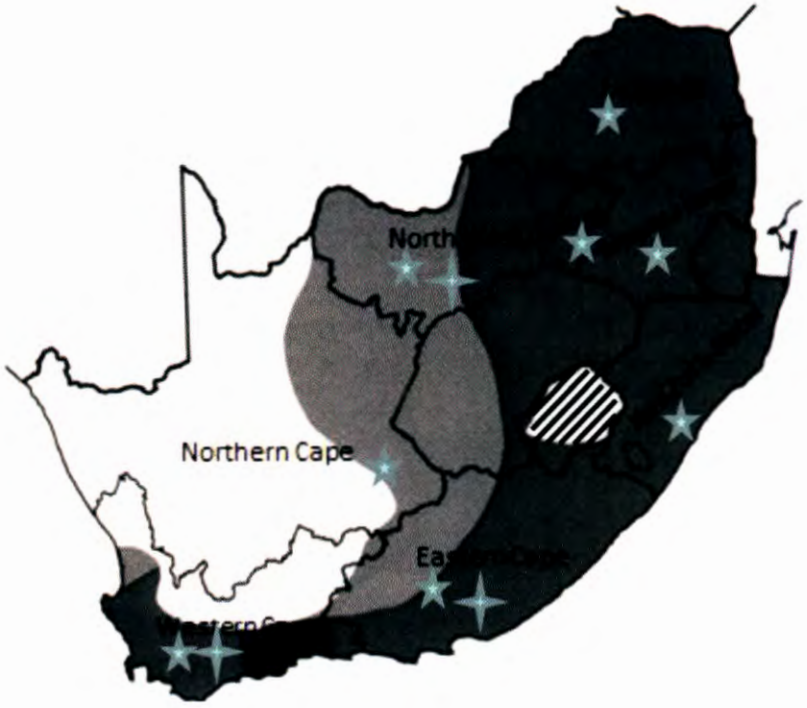
However, the specific codon positions that evolve under positive or negative selection have not been reported. For development of MSP1a-based vaccines, some characteristics of MSP1a should be taken into consideration, including (i) the extant genetic variability of MSP1a, (ii) the evolution of MSP1a genetic diversity, and (iii) the conservative nature of some tandem repeats among different isolates.

In this study molecular evidence is provided regarding the prevalence of *A. marginale* in eight of the nine South African provinces. Additionally, evolution of the genetic diversity of *A. marginale msp1a* in South Africa was studied, demonstrating that different codon positions of this gene evolved under positive or negative selection, likely due to immune selection and transmission fitness. Finally, these studies demonstrated the low variability of some tandem repeats commonly found among *A. marginale* strains. Collectively, results of this research will contribute toward development of new and novel vaccines for control of bovine anaplasmosis in South Africa and other regions of the world.

4.2 Materials and methods

4.2.1 Study site and sample collection

Blood samples were collected for these studies from May 2011 to July 2013 in 26 districts and municipalities from eight South African provinces (Figure 5). Coordinates of the collection sites are provided in Table 4, as well as the farm production systems. Cattle for these studies were randomly selected and blood samples were collected only from adult animals after the owner's consent. While information regarding the age or sex of the animals was not recorded, vaccination histories were not available for these animals. Blood samples were collected by tail venupuncture using an 18 gauge needle and sterile 10 ml vacutainer EDTA tubes and stored at 4 °C.



A. *marginale* epidemic areas
A. *marginale* endemic areas
Areas free of *A. marginale*
Lesotho, area not included in the study

Ticks species involved in *A. marginale* transmission: (*R. evertsi evertsi*)

(*R. microplus* and *R. decolaratus*)

Figure 5: Map of South African areas included in the study. Map of South Africa showing the provinces that were included in the study (Limpopo, Mpumalanga, North West, Gauteng, KwaZulu-Natal, Eastern Cape, Western Cape, and Northern Cape). Endemic, epidemic, and *A. marginale*-free areas are coloured differentially (data collected from de Waal, 2000). The main tick species involved in the transmission of *A. marginale* in the sampled areas are shown: *Rhipicephalus microplus, R. decoloratus*, and *R. evertsi evertsi* (data collected from de Waal, 2000).

Table 4: Sampling sites and their coordinates in South African provinces.

South African provinces	Sampling sites	Map coordinates
Limpopo	Makgodu, Masehlong, Chloen farm, Aganang Municipality	23º 36'S, 29º 18'E, 23º 32'S, 29º 03'E, 23º 45'S, 28º 51'E,
Mpumalanga	Ehlanzeni South District	25º 27'S, 30º 58'59"E
Gauteng	West Rand District, Merafong Municipality, Khutsong South and Carltonville	26º 20' 1" S, 27º 19'39"E 26º 22'S, 27º 24'E
North West	Moretele district, Maubane	25º 16'S, 28º 15'E
KwaZulu-Natal	Pietermaritzburg Chota, Umhlati District, Albert Falls, Shallow drift, UMgungundlovu District, Richmond Municipality, Ndaleni Dip Tank	29º29'17.16"S, 30º26'27.78"E, 29º28' 33.5" S, 30º 26' 11.3" E 29º 52' 55.3" S, 30º 40' 56.2"
Eastern Cape	Amathole District, Nkokobe Municipality (Fort Beaufort, Alice) and Middledrift and Nxuba Municipality (Bedford, Adelaide)	26º38'39"E, 32º46'52"S, 26º51'25"E, 32'41.12S, 27º11'58"E, 32º 52'37"S, 26º26'40"E, 32º44'25"S, 26º 18'18"E, 32º 42'04"S
Western Cape	Stellenbosch district, Boland	33º44'28.6"S, 18º59' 3"E
Northern Cape	John Taolo Gaetsewe District, Ga-Segonyana Municipality, Kuruman, Zero Farm	27º18'53"S, 23º 42'15"E

4.2.2 DNA extraction

Genomic DNA was extracted from cattle blood samples using ZR Genomic DNA™ Tissue Miniprep (Zymo Research, CA, USA). DNA was resuspended in DNA elution buffer and stored at -20 °C. The concentration of DNA was determined using the NanoDrop® ND-1000 (NanoDrop Technologies Inc., Wilmington, USA).

4.2.3 *A. marginale* species-specific PCR

Specific set of primers were used to amplify *msp1a* - 1733F- (5'-TGTGCTTATGGCAGACATTTCC-3'), 2957R- (5'- AAACCTTGTAGCCCCAACTTATCC-3') - gene (Lew *et al.*, 2002). Primers were synthesized and supplied by Inqaba Biotech (Inqaba Biotechnical Industries (Pty) Ltd, Pretoria, Gauteng, South Africa). PCR reactions were prepared using Master Mix (Thermo Scientific DreamTaq Green PCR), 0.1-1.0 µM of Forward and Reverse primers and 1 µg of DNA template in 25 µl final volume. The amplification cycles consisted of 40 cycles of 1 minute at 94 °C, 1 minute at 65 °C, and 2 minutes at 72 °C. Amplified products were separated in 1% TBE (89 mM Tris-Borate, 2 mM EDTA, pH 8) agarose gel using 1 Kb ladder as a DNA size marker (1 Kb DNA ladder, Life Sciences, Fermentas GmbH, Germany). DNA was visualized by gel staining in 0.5 µg/ml GelRed (Fermentas GmbH, Germany) under UV illumination and photographed.

4.2.4 DNA sequencing of *A. marginale msp1a* gene

PCR products containing only single amplification products of *msp1a* gene were sequenced at Inqaba Biotech facilities (Inqaba Biotechnical Industries (Pty) Ltd, Pretoria, Gauteng, South Africa). The termination reactions were performed using BigDye VER3.1 (ABI, Life Technologies, CA, USA) according to manufacturer's instructions. The labelled fragments were purified using Zymo research sequencing clean-up kit (Zymo Research, CA, USA) and subsequently analysed on a 3500xl Genetic Analyzer (ABI, Life Technologies, CA, USA). The sequences obtained in this study were submitted to GenBank and provided with accession numbers for *msp1a* (KC470153-KC470196) gene (Appendix 2).

4.2.5 Sequence analysis of *A. marginale msp1a* gene

To identify the *msp1a* gene sequences obtained in our study the database Nucleotide collection (nr/nt) using Megablast (optimize for highly similar sequences) from the BLAST server was used (Zhang *et al.*, 2000). Protein homology and identity analysis were performed using the multiple-alignment program ClustalW (Thompson *et al.*, 1994). The MSP1a tandem repeats found in this study were reported previously (Cabezas-Cruz *et al.*, 2013).

4.2.6 Codon based phylogenetic analysis of tandem repeats

Codon based alignment was performed using the codon suite server (Schneider *et al.*, 2005; 2007). Detection of selection pressure on individual codons was calculated using two methods, single likelihood ancestor counting (SLAC) and fixed effects likelihood (FEL) (Pond and Frost., 2005), used in the Datamonkey webserver (Delport *et al.*, 2010; Pond and Frost., 2005). Positive and negative selections were assigned to codon where ω=dN (non-synonymous substitutions)/dS (synonymous substitutions) ratio was higher or lower than 1 respectively. The reconstruction of the ancestral amino acid sequence was performed using a neighbour joining tree rooted in tandem repeat 83 under the Dayhoff model of substitutions which was estimated to be the best model fitting the actual data. Three reconstruction methods were used: Joint (Pupko *et al.*, 2000), marginal (Yang *et al.*,1995) and sample (Nielsen, 2002) which are also used in the Datamonkey webserver.

4.2.7 Amino acid variability, composition and genetic diversity index (GDI)

The amino acid variability was calculated in the variability server (Garcia-Boronat *et al.*, 2008) using the Shannon entropy (H) formula (Shannon, 1948) as follow:

$$H = -\sum_{i=1}^{M} P_i \log_2 P_i$$

Where P_i is the fraction of residues with a certain type of amino acid *i*, and *M* is the number of types of amino acid for a position. The proportion of variable over conserved positions was calculated as the number of positions with more than 0 of Shannon variability divided by the number of positions with 0 Shannon variability. Amino acids were also classified regarding biochemistry properties, to know: negative charged, positively charged, uncharged-polar and non-polar. For comparison purposes, Shannon variability was additionally calculated in 28 and 43 MSP1a sequences available in GenBank from Venezuela and USA respectively. To further characterize the genetic diversity of MSP1a, a genetic diversity index (GDI) was calculated for each *A. marginale* strain as follow: number of different MSP1a tandem repeats divided by the total number of tandem repeats per strain. 1 and 0 were considered as maximum and minimum genetic diversity, respectively.

Table 5: Observed prevalences of *Anaplasma marginale* in different provinces of South Africa

South African provinces	No. of blood samples collected per province	Prevalence of *A. marginale*, *msp1a*-positive PCR (%)	No. of *msp1a* sequenced
Limpopo	20	13(65)	8
Mpumalanga	21	21(100)	2
Gauteng	20	20(100)	5
North West	33	24(72)	7
KwaZulu-Natal	55	50(90)	9
Eastern Cape	40	40(100)	3
Western Cape	16	14(87.5)	11
Northern Cape	45	0(0)	0
Free State[a]	215	129(60)	29

[a] Data collected by Mtshali *et al.*, (2007).

4.3 Results and discussion

4.3.1 Molecular evidence of *A. marginale* prevalence in South Africa

Bovine anaplasmosis has been considered to be endemic in South Africa, an assumption based primary on the distribution of the tick vectors (Ndou *et al.*, 2010) and the sero-prevalence of *A. marginale* which was determined only in the Free State (Dreyer *et al.*,1998), Limpopo (Rikhotso *et al.*, 2005) and North West (Ndou *et al.*, 2010) provinces. Molecular evidence of endemic bovine anaplasmosis has not been reported in most of South Africa.

This study was therefore designed to investigate the molecular evidence of *A. marginale* infection in cattle from Mpumalanga, Gauteng, Eastern Cape, Limpopo, North West, KwaZulu-Natal, Eastern Cape and Western Cape provinces (Table 5). Molecular diagnostic of anaplasmosis was determined previously in Free State province (Mtshali *et al.*, 2007). In the present study, species-specific *msp1a* primers were used for PCR assays on 250 DNA samples obtained from cattle blood.

The results of these PCR studies confirmed that *A. marginale* is widespread in South Africa, but with a variable prevalence in all the sampled provinces, except for Northern Cape in which no positive samples were detected (Table 5). The absence of *A. marginale* positive PCR in Northern Cape is not surprising because this area is considered to be free of tick vectors (Mtshali and Mtshali, 2013) and the prevalence of *Babesia* spp., another tick-borne pathogen, was also found to be very low in this province (Mtshali and Mtshali, 2013). The provinces with highest prevalence of *A. marginale* were Mpumalanga, Gauteng, and Eastern Cape with an infection rate of 100%. In the other provinces, *A. marginale* infections in cattle ranged from 65% to 90% (Table 5 and Figure 7).

Differences in the prevalence of *A. marginale* were not observed between commercial and communal farming systems. The prevalence of *A. marginale* in cattle from South Africa can be considered high when compared to the prevalence of *A. marginale* in cattle from other regions of the world, for example, in Brazil a recent study found 70 % prevalence of *A. marginale* in a cattle herd (Pohl *et al.*, 2013). Another study, by de la Fuente *et al.* (2005) showed a range from 25% to 100% of *A. marginale* prevalence in cattle herds from Italy. Analysis of the prevalence and genetic diversity of *A. marginale* MSP1a in different geographic regions constitutes an important step toward development of effective MSP1a based vaccines because the antigenic composition of the vaccine should contain MSP1a variants present in different regions of the world.

```
A   DDSSSASGQQQESSVSSQSE-ASTSSQLG--

83  T...........G.L...GQ......S.--
82  A...........L...DLS..W....--
79  T........P.....LCV.DLS.......--
100 T...........G.L...GQ........-
140 A...........G.L...GQ.......R-
141 T...........G.L...GQ.......R-
142 T.............LP..GH.R....S.-
143 T....G.........L...SQ.RS.....-
144 .............LP..GQD.....S.-
145 G....S.........L...SQ........-
146 A.....GN..........-........-
147 T.....GN..........G-.G......-
148 T.....GD..........G-..A..K.R-
149 T.....GD.........L.G-........-
150 T.....GD.K........IG-.....K..-
151 AN........E....L...DQ........-
152 T.............L...DQ.......R-
153 T...PE-........F...AQ......S.-
154 A.............L...DQ......S.-
155 AN............L...GQ........-
158 T..W..........L...DQ........-
160 A.............P....-.A.....AD
161 ANG...........L...DQ........-
```

Figure 6: Newly reported sequences of MSP1a tandem repeats. The one-letter amino acid code was used to depict MSP1a repeat sequences. Dots indicate identical amino acids, and gaps indicate deletion/insertions. The ID of each repeat form was assigned previously in Cabezas-Cruz *et al.*, (2013). Tandem repeat A was used as a model for amino acid comparison.

4.3.2 *A. marginale* prevalence and *msp1a* genetic diversity

The sequence of *A. marginale msp1a* was analysed in 44 strains (Table 4). We found 52 different types of tandem repeats among South African strains (including those reported by Mtshali *et al.*, 2007 from Free State province) from which 23 were described for first time (Figure 6). Using a genetic diversity index (GDI), the genetic diversity per *A. marginale* strain we described (Table 6) and the average of genetic diversity was calculated per Province. A polynomial correlation (R^2=0.76) occurred between the GDI and the prevalence of anaplasmosis per province (Figure 7). Interestingly, provinces with 100 % of *A. marginale* prevalence were not the provinces with the highest or lowest GDI, rather these provinces were between a range of genetic diversity (0.82 – 0.87).

Immune selection and transmission fitness have been suggested to impact genetic diversification of *A. marginale* (Palmer and Brayton, 2013). Immune pressure in cattle may induce greater genetic diversity in *A. marginale* populations in order to insure persistent infection. At the same time, the development of antigenic variation was suggested to have a transmission cost (Palmer and Brayton, 2013). *A. marginale* strains with lower transmission fitness due to high genetic variability would be at risk of elimination from the population (Palmer and Brayton, 2013). Furthermore, high incidence of ticks correlated to increased MSP1a genetic variability in Argentina (Ruybal *et al.*, 2009).

These apparently contradictory aspects may be reconciled if increased tick transmission contributes to greater circulation of *A. marginale* strains in the cattle population, favouring the interaction between *A. marginale* strains and potentially resistant hosts. In this study, the observed range of GDI in which 100% prevalence (an indicator of transmission efficiency) exist may reflect a positive balance between genetic diversity and biological transmission by ticks taking in account that, as shown in Figure 7, all the studied area, except for Northern Cape, in infected by ticks.

Notably, most of the new tandem repeats shown in Figure 6 were sequenced from KwaZulu-Natal province which has 90 % prevalence of *A. marginale*. When KwaZulu-Natal was excluded from the correlation analysis between *A. marginale* prevalence and GDI (Figure 7), the polynomial regression increase from R^2=0.76 to R^2=0.98. A possible explanation of this result may be that the tandem repeats diversification found among KwaZulu-Natal MSP1a provided a degree of transmission fitness to the strains present in that area. Another explanation could be that cattle immunity is increasing in the area, inducing *A. marginale* diversification for antigenic variation. One interesting question that remains unanswered is how the genetic diversity of *A. marginale msp1a* emerges in a specific region.

4.3.3 Evolution of *msp1a* genetic diversity

In order to analyse the evolution of *msp1a* genetic diversity observed in our samples, the tandem repeats (Table 6) were classified as "frequent" (present more than 22 times) or "rare" (present less than 10 times) based on the frequency of their appearance among South African *A. marginale* strains, including those from Free State province reported by Mtshali *et al.* (2007). The unique tandem repeats found in this study were classified as rare, being repeated, most of them, one time in only one strain (Figure 6). In contrast, tandem repeats 3, **4**, 13, 34, Q and 37 had a high frequency (Table 6) and were also reported in *A. marginale* strains from Israel (3, 4), South America (**4**, 13) and Europe (Q). Repeat sequences 34 and 37 were abundant only in South Africa with rare exceptions (as reviewed by Cabezas-Cruz *et al.*, 2013). Considering this, we wanted to test whether the tandem repeats newly described in this study (Figure 6) originated from extant *A. marginale* MSP1a tandem repeats or had evolved from a tandem repeat that was lost after tandem repeat differentiation.

Table 6: *Anaplasma marginale* strains and putative 2D structures of MSP1a tandem repeats

A. marginale strain	Province of origin	Structure of *msp1α* tandem repeats	GDI[a]	AVE-GDI/STDEV
LP-7	Limpopo	**34** 159	1	0.917/0.144
LP-10	Limpopo	27 **13** **3** 36	1	
LP-30	Limpopo	27 **13** **3**	1	
LP-34	Limpopo	**34** **13** **3** 38	1	
LP-37	Limpopo	27 **13** **13** **37**	0.75	
LP-46	Limpopo	**3** 38	1	
LP-50	Limpopo	**34** **13** **13**	0.667	
MP-C2	Mpumalanga	**34** **1** 15 **37**	1	0.875/0.177
MP-C5	Mpumalanga	15 15 100 83	0.75	
NW-C2	North West	27 **13** **4** **4** **37**	0.8	0.960/0.089
NW-C4	North West	27 **13** **4** **37**	1	
NW-C5	North West	82 **13** 79 **4** **37**	1	
NW-C1-160312	North West	**34** **13** **3** 36 38	1	
NW-C4-160312	North West	**34** 36 38 **3**	1	
GP-C1	Gauteng	82 **13** **4** **4** **37**	0.8	0.826/0.173
GP-C2	Gauteng	**34** 27 **3** 38 **13** **3** 38	0.714	
GP-C5	Gauteng	**3** **4** **4** **4** **37**	0.6	
GP-C1112105	Gauteng	**34** **37**	1	
GP-C4117105	Gauteng	**3** 36 38	1	
GP-C7117105	Gauteng	**34** **13** **13**	0.667	
GP-C1817105	Gauteng	**34** **13** **37**	1	
KZN-D	KwaZulu-Natal	42 43 25 163 31	1	0.919/0.128
KZN-F	KwaZulu-Natal	**4** 43 25 31 31	0.8	
KZN-K	KwaZulu-Natal	**27** **13** **4** **4** **37**	0.8	
KZN-Y	KwaZulu-Natal	143 144 145 146	1	
KZN-MM	KwaZulu-Natal	42 43 25 31	1	
KZN-14	KwaZulu-Natal	142 43 25 31	1	

KZN-19	KwaZulu-Natal	141 140 140	0.667	
KZN-49	KwaZulu-Natal	147 148 149 150	1	
KZN-51	KwaZulu-Natal	147	1	
EC-22	Eastern Cape	**27 13 4 4 37**	0.8	0.867/0.115
EC-23	Eastern Cape	151 152 **4 4** 153	0.8	
EC-24	Eastern Cape	**27 13 4**	1	
WC-4	Western Cape	**4** Q Q m	0.75	0.741/0.286
WC-6	Western Cape	**3 4 4 37**	0.75	
WC-7	Western Cape	M M M M	0.25	
WC-8	Western Cape	**34 4 37**	1	
WC-10	Western Cape	154	1	
WC-11	Western Cape	40 Q Q Q Q Q **37**	0.429	
WC-12	Western Cape	27 **13 37**	1	
WC-13	Western Cape	M Q M Q M	0.4	
WC-14	Western Cape	155 36 38	1	
WC-15	Western Cape	161 **13 37 4** 162	1	
WC-16	Western Cape	**34 13 4 13 13 4 37**	0.571	

More common tandem repeats are highlighted (bold) and mentioned in order of abundance, from Higher to lower: 13,4,37 and 3, 34 have the same frequency, respectively.

[a]Genetic diversity index of MSP1a (GDI) was calculated as follows: number of different MSP1a tandem repeats/total number of tandem repeats per strains. 1 is maximum genetic diversity. The average of GDI (AVE-GDI) and standard deviation (STDEV) for each region is shown.

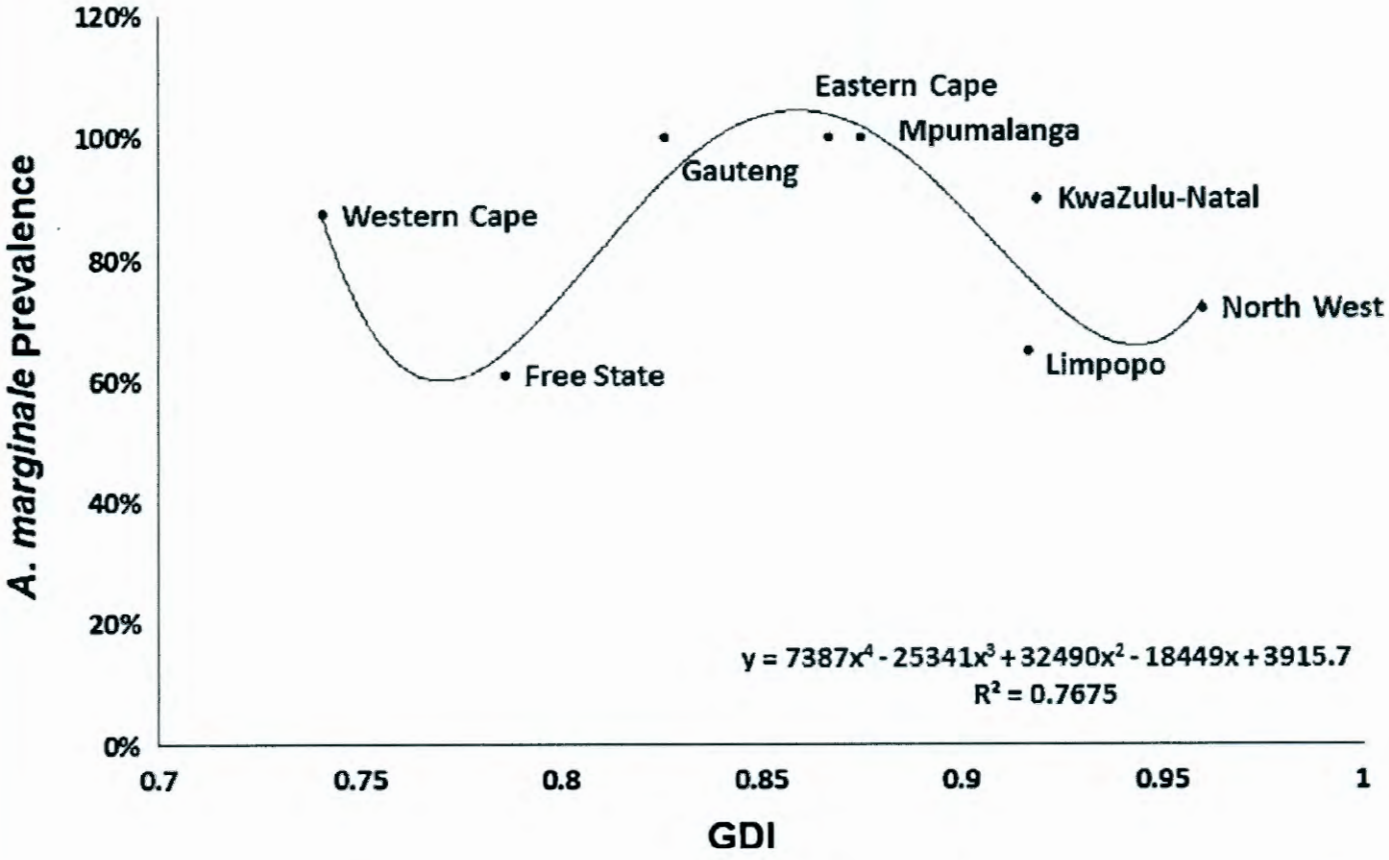


Figure 7: Correlation between *msp1a* genetic diversity and *A. marginale* prevalence in South Africa. The prevalences of anaplasmosis in different South African provinces were plotted against the average of genetic diversity index (GDI). GDI was calculated for each strain as the number of different tandem repeats divided by the total number of tandem repeats. A polynomial correlation was found between these 2 parameters with R2 = 0.76. Provinces with 100% prevalence (Gauteng, Eastern Cape, and Mpumalanga) are in the range of 0.82–0.87 GDI while regions having less than 100% prevalence are out of this range.

To test this hypothesis we reconstructed the ancestral state (Material and Methods) of the new tandem repeats presented in Figure 6. Surprisingly, the ancestral state of all the new tandem repeats was the tandem repeat 4 (Figure 8), suggesting that all of the newly described tandem repeats from South African *A. marginale* strains evolved from this tandem repeat. In addition, this evidence suggests that these repeated sequences may constitute a group of recently evolved tandem repeats specific to South Africa and, in fact, these repeats have not been reported elsewhere (Cabezas-Cruz *et al.*, 2013).

The mechanism for the generation of genetic diversity of *A. marginale* could be gene duplication followed by either mutation or mismatch repair as proposed by Palmer and Brayton. (2013). The variable number of tandem repeats found among MSP1a isolates suggests that tandem repeat duplication due to homologous recombination may occur in order to give a "substrate" for "testing" competitive advantage of new tandem repeats variants. As demonstrated by this research, new tandem repeats shown in Figure 6 evolved from an initial tandem repeat 4 which served as a template for genetic variation.

To test whether tandem repeat 4 evolved to the new forms under selective pressures, the ratio ω (see Materials and Methods) was calculated for each codon position of the new tandem repeats from South Africa. We found that the diversification of tandem repeat 4 in South African strains occurred under both positive and negative selective pressures (Table 7 and Figure 8). Surprisingly, the positions evolving under negative selection (Figure 8 negative signs), 8 and 10, were reported before to be present in an immunodominant B-cell epitope present in MSP1a (Figure 8, first boxed area, Garcia-Garcia *et al.*, 2004) and the position 25, also evolving under negative selection, was found previously in a neutralization-sensitive epitope (Figure 8, second boxed area; Palmer *et al.*, 1987; Allred *et al.*, 1990). In contrast, one of the positions found to be evolving under positive selection was the amino acid position 20 (Figure 8, arrow) which has been implicated in the binding of MSP1a to tick cells extract (de la Fuente *et al.*, 2003a).

These findings support the hypothesis discussed above that immune escape and tick transmission are both driving forces of *A. marginale* MSP1a genetic diversification and suggest that tick transmission and immune escape would be triggers of the observed diversification of tandem repeat 4 in South Africa. This trend of purifying selection/negative selection (or deletion of the unfit) for sites involved in immune recognition by the host was confirmed, in wider frame, by the following data: (i) most of the residues of the MSP1a immunodominant B-cell epitope (Garcia-Garcia *et al.,* 2004) were deleted in the tandem repeats forms α and 108: the tandem repeat α is widespread in strains from Mexico, Brazil, Venezuela, Argentina and Taiwan and is also present in the most common *A. marginale* strain of the world which have the tandem repeat composition (α, β, β, β, Γ) (Cabezas-Cruz *et al.,* 2013), and (ii) the first glutamine (Q) of the neutralization-sensitive epitope (Palmer *et al.,* 1987; Allred *et al.,* 1990) is deleted in several tandem repeats (A, D, E, γ, Σ, φ, 5 - 9, 14, 31, 36, 52, 57, 60 - 66, 69 - 72, 76, 84 – 86, 95 – 99, 105, 107, 116 – 119, 129 – 131, 136 - 139) from *A. marginale* strains reported previously (Cabezas-Cruz *et al.,* 2013). As seen in Figure 8, Q was deleted in some of South African new tandem repeats (146, 147, 148, 149 and 150). Collectively, this evidence suggests that purifying selection is most likely one of the mechanisms that *A. marginale* had evolved to escape immune recognition toward MSP1a.

Tandem repeat 4 from MSP1a was also found in *A. marginale* strains isolated from an anaplasmosis outbreak in Mexico where *R. microplus* was implicated as tick vector (Almazán *et al.,* 2008), and new tandem repeats were reported also in this study. It should be interesting to test whether these newly described Mexican MSP1a variants evolved from tandem repeat 4. We consider these findings to be relevant to the development of MSP1a based vaccines and suggests that MSP1a based vaccination should be combined with tick control strategies in order to minimize genetic diversity of *A. marginale msp1a*. A recent study was reported that combined use of a tick protective antigen with MSP1a in the same vaccine formulation that was directed toward control both tick infestations and anaplasmosis (Torina *et al.,* 2014).

Table 7:Sites that evolved under positive and negative selection in the new tandem repeats from South Africa

Codons	Methods SLAC		Methods FEL		Type of selection
	ω(dN/dS	p value	ω(dN/dS	p value	
1	5.49	0.043	Infinite	0.012	Positive
6	1.58	0.311	Infinite	0.093	Positive
16	1.39	0.371	Infinite	0.196	Positive
20	2.86	0.172	Infinite	0.133	Positive
28	2.22	0.197	Infinite	0.079	Positive
8	-1.66	0.373	-5.14	0.236	Negative
10	-6.81	0.037	-13.89	0.016	Negative
25	-3.57	0.134	-6.56	0.110	Negative

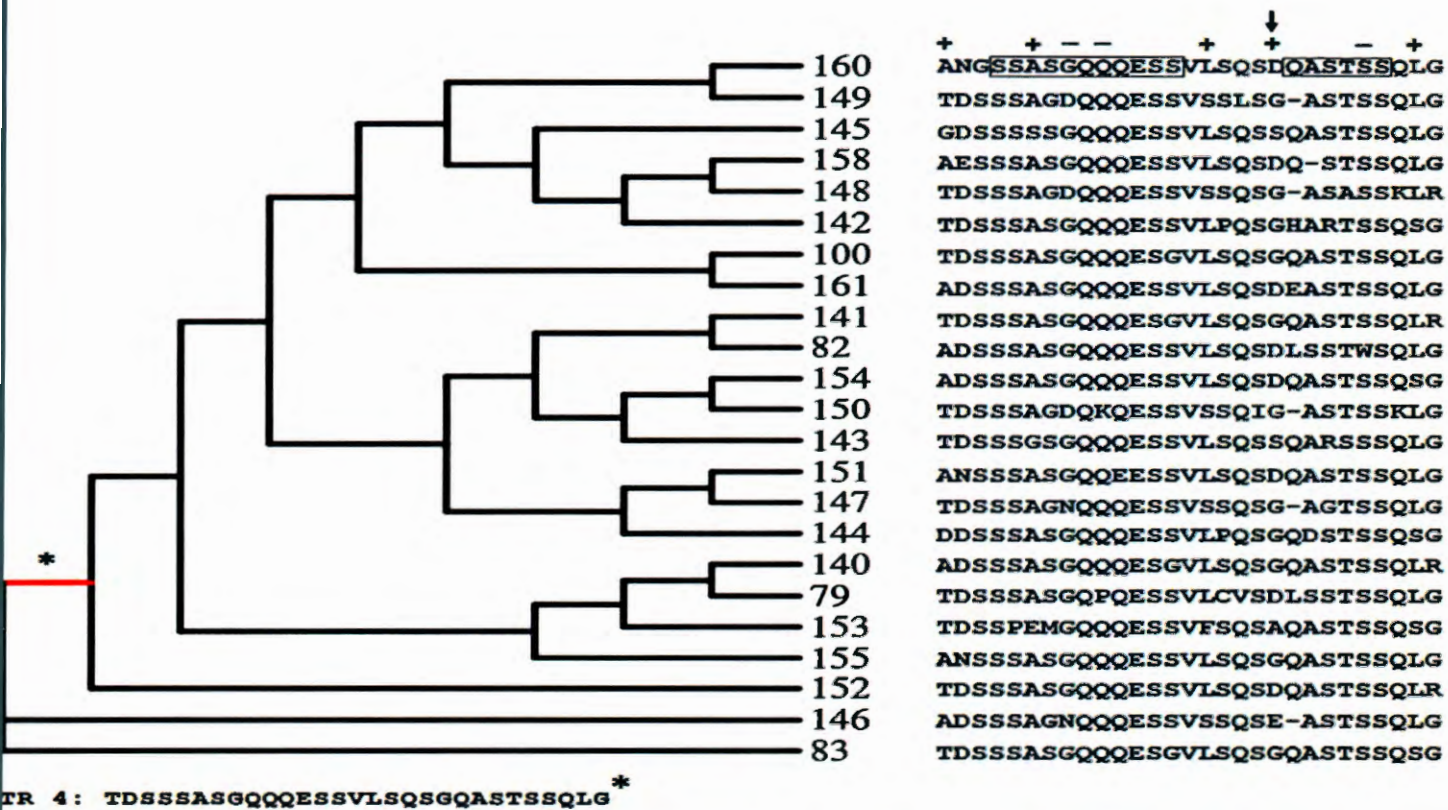


Figure 8: Reconstruction of ancestral amino acid sequence and amino acid positions under positive and negative selection. The reconstruction of the ancestral state (TR 4: Tandem repeat 4, sequence marked with *) of the new tandem repeats found in South Africa (Figure 6) was performed using 3 reconstruction methods, namely: joint, marginal, and sample (see 'Materials and methods' for details). Positions that evolve under negative (−) and positive (+) selection are shown (see Table 7). Amino acid at position 20 is indicated (arrow). The residues of the immunodominant B-cell epitope (Garcia-Garcia *et al.*, 2004) (first box) and the neutralization-sensitive epitope (Palmer *et al.*, 1987; Allred *et al.*, 1990) (second box) are also shown.

4.3.4 Amino acid variability and low variable *msp1a* peptides

Genetic variability could also be analysed by calculating the amino acid variability in each position of the tandem repeat. In order to compare the *msp1a* genetic variability in South Africa with other regions of the world we calculated the amino acid variability for 29 amino acid positions from all the *msp1a* tandem repeats found in South Africa, USA and Venezuela which were available in GenBank and collected by Cabezas-Cruz *et al.*, 2013.

The amino acid variability of *msp1a* tandem repeats from South Africa can be seen in Figure 9 as compared with that from USA (low) and Venezuela (high). In agreement with this, the average of amino acid variability was 0.30, 0.42 and 0.72 for USA, South Africa and Venezuela, respectively (Figure 9A-C). In comparison the proportion of variable over conserved positions in MSP1a tandem repeats was higher Venezuela (29) compared to USA (0.85) and South Africa (2.2). Epitopes from MSP1a were found to induce a protective immune response against *A. marginale* (Santos *et al.*, 2013).

Considering this, using the variability server (Garcia-Boronat *et al.*, 2008), we explored whether some low variable peptides could be found in MSP1a from South Africa, USA and Venezuela. We observed that the amino acid variability of MSP1a tandem repeats from South Africa and USA supported the existence of low variable peptides (Figure 9) which may be useful in development of peptide-based MSP1a vaccines, while the high amino variability of MSP1a tandem repeats from Venezuela do not support for this type of low variable peptide. Interestingly, the low variability peptide in South Africa overlap with the position of the immunodominant B-cell epitope reported previously for MSP1a (Garcia-Garcia *et al.*, 2004).

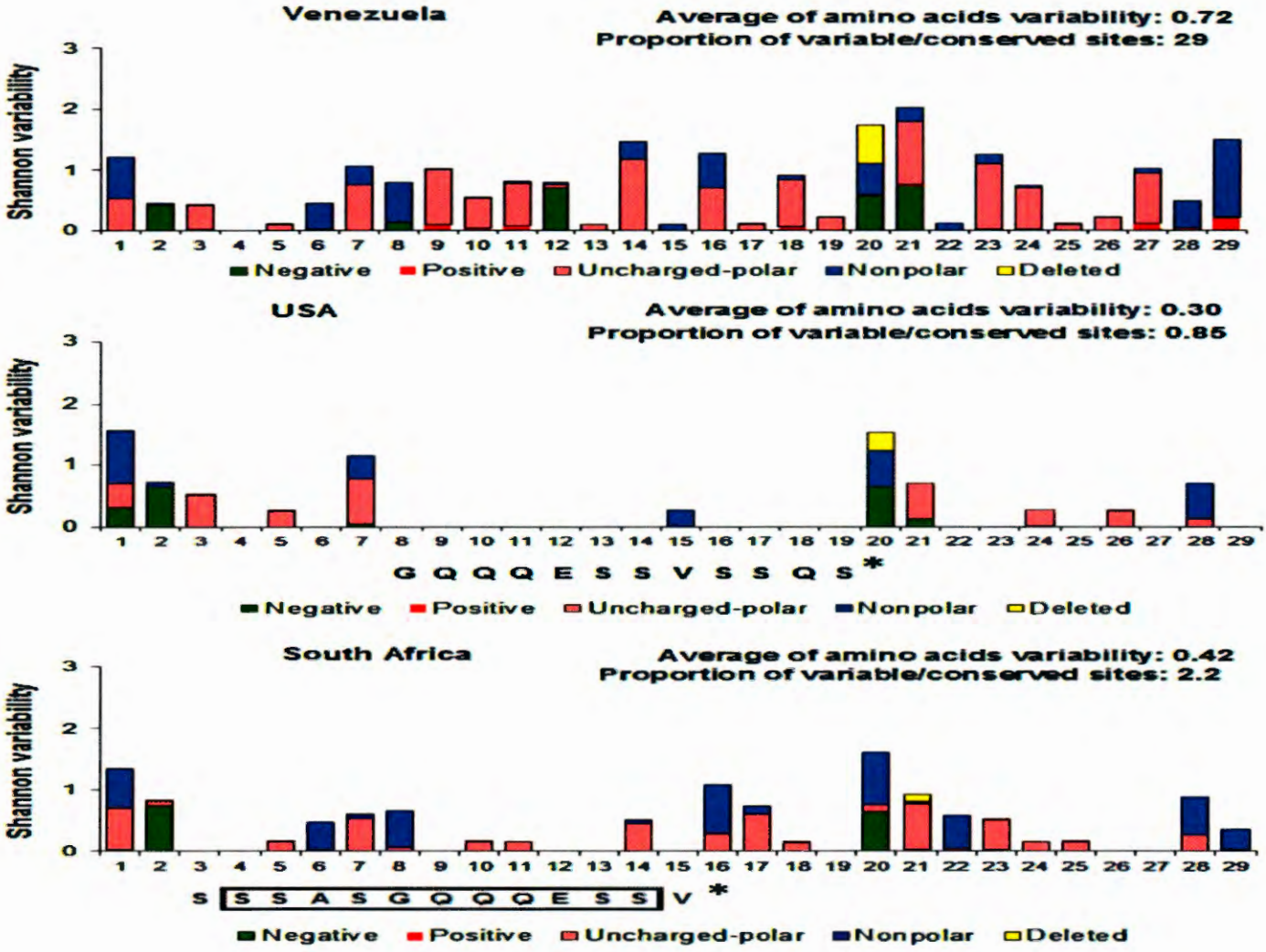


Figure 9: MSP1a amino acid variability, composition, and low-variable peptides. The figure shows the amino acid variability and composition among the tandem repeats in 3 different countries: Venezuela (A), USA (B), and South Africa (C). Venezuela shows a high proportion of variable/conserved sites and a high average of amino acid variability while South Africa and the USA show middle and lower values, respectively. Different colours in columns depict different biochemical properties in the amino acid composition: negative (green), positive (red), uncharged-polar (beige), and non-polar (blue); proportion of deleted positions is shown in yellow. Consensus sequences of low-variable peptides are shown (*) for the USA and South Africa. The region of the immunodominant B-cell epitope from *A. marginale* (Garcia-Garcia *et al.*, 2004) is boxed in the low-variable peptide from South Africa. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

CHAPTER 5

STRUCTURAL AND PHYLOGENETIC ANALYSES OF *ANAPLASMA MARGINALE* INFECTING CATTLE IN SOUTH AFRICA USING *msp1a* AND *msp4* GENES

5.1 Introduction

Phylogenetic analysis establishes the relationships between genes or gene fragments, by inferring the common history of the genes or their fragments. To achieve this, it is essential that homologous sites be compared with each other (*positional homology*) (Nei and Kumar, 2000). Phylogenetic studies have been previously conducted to resolve and understand the *Anaplasma* species and how these species relate to one another (de la Fuente *et al.*, 2007a).

The phylogenetic relationship and evolution of *A. marginale* isolates is important for understanding the biology and the possibilities for control of anaplasmosis (Kocan *et al.*, 2002). Recent research has focused on MSPs that are involved in interactions with vertebrate and invertebrate host cells (McGarey and Allred, 1994; McGarey *et al.*,1994 de la Fuente *et al.*, 2001b) and have been used to elucidate the phylogeographic patterns of *A. marginale* (Kano *et al.*, 2002; Lew *et al.*, 2002; de la Fuente *et al.*, 2007a). MSP1a and MSP4 genes were used previously, for phylogenetic analysis that was found to be involved in interactions with both vertebrate and invertebrate hosts (McGarey and Allred, 1994; McGarey *et al.*, 1994; de la Fuente *et al.*, 2001b). Therefore, these MSPs are also likely to evolve more rapidly than other nuclear genes because they are subjected to selective pressures exerted by host immune systems.

Both *msp1a* and *msp4* genes have been found to be stable genetic markers during multiplication of *A. marginale* (de la Fuente *et al.*, 2001a; Bowie *et al.*, 2002). MSP1a was also shown to effect transmission of *A. marginale* by ticks (de la Fuente *et al.*, 2001a; Kocan *et al.*, 2010). These genes were reported to be more stable for phylogenetic studies because the MSPs do not appear to undergo antigenic variation in cattle or ticks (Bowie *et al.*, 2002). MSP1a, encoded by *msp1a*, has been reported to be an adhesin for bovine erythrocytes and tick cells and to effect adhesin, infection, and transmission of *A. marginale* by ticks of the genus *Dermacentor* (McGarey and Allred, 1994; McGarey *et al.*, 1994; de la Fuente *et al.*, 2001c). Although the specific function of MSP4 is not known, previous analysis of the *msp4* gene from *A. marginale* isolates demonstrated that its sequence varies sufficiently among isolates to support its use in phylogeographic studies (de la Fuente *et al.*, 2003a). In South Africa, phylogenetic profile of *A. marginale* infecting cattle has not been determined meaning that the genetic relationship between South Africa *A. marginale* isolates and those in other African countries is unknown. Therefore, the aim of this study is to establish the phylogenetic profile of *A. marginale* infecting South Africa cattle using *msp1a* and *msp4* as marker genes.

5.2 Materials and methods

5.2.1 DNA extraction

The genomic DNA was extracted from cattle blood samples using ZR Genomic DNA™ Tissue Miniprep (Zymo Research, CA, USA). Total volume of blood sample was adjusted in a microcentrifuge tube with water to 100 µl prior to adding 95 µl of 2x Digestion buffer and 5 µl of Proteinase K. The mixture was thoroughly mixed and then incubated at 55 °C for 20 minutes. A volume of 700 µl of genomic lysis buffer was added to the tube and mixed thoroughly by vortexing.

The mixture was carefully transferred to a Zymo-Spin™ IIC column in a collection tube and centrifuged at 10,000 x *g* for one minute. A volume of 200 µl of DNA pre-wash buffer was added to the spin column in a new collection tube and centrifuged at 10,000 x *g* for one minute. A volume of 400 µl of g-DNA wash buffer was added to the spin column and centrifuged at 10, 000 x *g* for one minute. The spin columns were carefully transferred to a clean 1.5 ml microcentrifuge tube and 100 µl of DNA Elution buffer added to the spin column. The microcentrifuge tube was incubated for 5 minutes at room temperature, and then centrifuged at 8000 rpm for 30 seconds to elute the DNA. The eluted DNA was stored at -20 °C until used for studies. The concentration of DNA was measured using a NanoDrop® ND-1000 (NanoDrop Technologies Inc., Wilmington, USA).

5.2.2 Set of primers used for amplification of *msp1a* and *msp4* genes

Table 8: Oligonucleotide sequence of primers used in this study and their references

Primer designation	Oligonucleotide (5'-3')	Reference
1733F	TGTGCTTATGGCAGACATTTCC	Lew *et al.*, 2002
2957R	AAACCTTGTAGCCCCAACTTATCC	Lew *et al.*, 2002
MSP45F	GGGAGCTCCTATGAATTACAGAGAATTGTTTAC	de la Fuente *et al.*, 2003a
MSP45R	AAGCTCGAGGCTGAACAGGAATCTTGCTCCAAG	de la Fuente *et al.*, 2003a

5.2.3 *Anaplasma marginale* species-specific PCR of the *msp1a* gene

Specific set of primers were used to amplify *msp1a* gene (Table 8). Primers were synthesized and supplied by Inqaba Biotech (Inqaba Biotechnical Industries (Pty) Ltd, Pretoria, Gauteng, South Africa). The PCR reagents mixture consisted of total volume of 25 µl of Thermo Scientific DreamTaq Green PCR Master Mix (2X), 10 µM of forward and reverse primers, 1 µg Template DNA in 0.2 ml PCR tubes. The PCR mixture was then amplified by DNA thermocycler (Bio-Rad T100™ DNA Thermal Cycler, Bio-Rad Laboratories, Johannesburg, South Africa). The amplification process involved an initial denaturation at 94 °C for 3 minutes followed by 40 cycles of denaturation at 94 °C for 1 minute at 94 °C, annealing at 65 °C for 1 minute, and an extension at 72 °C for 2 minutes. A cycle of a final extension was done at 72 °C for 7 minutes. The amplified products were then separated by electrophoresis using 1% agarose gel immersed in TBE buffer (89 mM Tris-Borate, 2 mM EDTA, pH 8). One kilo-base pair ladder marker was used to determine the size of the PCR products after staining with 0.5 µg/ml GelRed (Life Sciences, Fermentas GmbH, Germany) and visualization under UV illumination. Gel documentation was done and taking a photograph using a camera.

5.2.4 *Anaplasma. marginale* species-specific PCR of the *msp4* gene

The *msp4* gene was amplified from 5 µl per 20 µl reaction volume using 10 pmol of each primer (Table 7) (2x Phusion Flash PCR Master Mix) (Thermo Fisher Scientific, USA). Reactions were performed in a BIO-RAD T100™ DNA thermal cycler (Bio-Rad Laboratories, Johannesburg, South Africa). The amplification cycles following an initial denaturation at 98 °C for 3 minutes for 1 cycle then consisted of 3-step protocol of a denaturation step of 98 °C for 10 seconds, annealing at 63 °C for 30 second and extension at 72 °C for 30 seconds for 5 cycles, 98 °C for 10 seconds, annealing at 65 °C for 30 seconds and extension at 72 °C for 30 seconds for 5 cycles, 98 °C for 10 seconds, annealing at 67 °C for 30 seconds and extension at 72 °C for 30 seconds for 20 cycles, with a final cycle at 72 °C for 7 minutes.

The amplified products were then separated by electrophoresis using 1%agarose gel immersed in TBE (89 mm Tris-Borate, 2 mm EDTA, pH 8). One kilo-base pair ladder marker was used to determine the size of the PCR products after staining with 0.5 µg/ml GelRed (Life Sciences, Fermentas GmbH, Germany) and visualization under UV illumination. Gel documentation was done and taking a photograph using a camera (Figure 10).

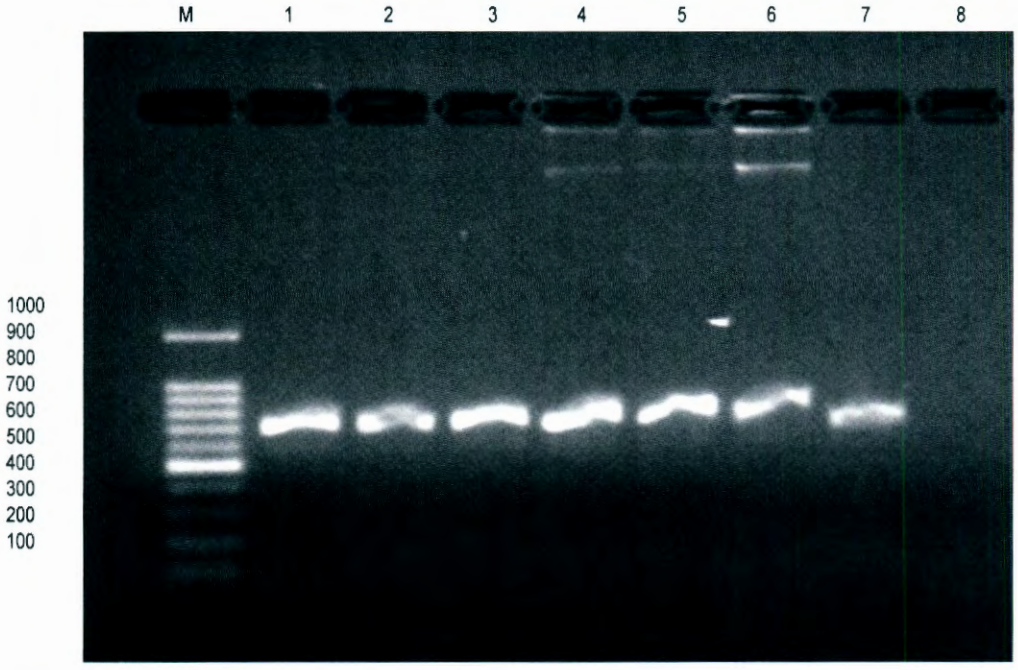


Figure 10: Gel image of 1% agarose gel electrophoresis of PCR amplification products obtained from *A. marginale* from Western Cape Province in South Africa using primer MSP45F and MSP45R. Lane M: GeneRuler™ 1Kb DNA ladder, ready-to-use; Lane 1- 7 indicate positive results and lane 8 indicate negative control.

5.2.5 DNA sequence alignment and phylogenetic analysis

All *A. marginale msp1a* and *msp4* gene sequences and GenBank accessions utilised here are listed in Appendix 2, 3 and 4. The obtained sequences were edited using Chromas V1.49 software and BioEdit (Tom Hall Ibis Biosciences, Carlsbad, USA). Nucleotides were coded as unordered, discrete characters with five possible character states: A, C, G, T or N and gaps were coded as missing data. The percent identity between the tandem repeat in position 1 (R1) and the other tandem repeats was determined using CLUSTALW. The tandem repeats secondary structure was predicted using the PSIPRED server (Buchan *et al.*, 2010).

The sequences obtained in this study were compared with available sequences from the NCBI database using the Basic Local Alignment Search Tool (BLAST) (http://www.ncbi.nlm.nih.gov/BLAST/). The sequences were submitted to GenBank (Appendix 3). Multiple sequence alignment was performed using ClustalW 1.6 program (Thompson *et al.*, 1994). Phylogenetic analysis was conducted in MEGA 5 (Tamura *et al.*, 2007) using Neighbour joining (NJ) and Maximum Likelihood (ML) algorithm (Saitou and Nei, 1987). The evolutionary distances were computed using with Kimura 2-parameters correction (Kimura, 1980) with bootstrap analysis of 1000 replicates. Sequences of *A. marginale msp4* reported previously from West, East and North Africa, North and South America, Europe, Asia and Australia were obtained from GenBank and used for the *msp4* gene phylogenetic analysis of South African isolates sequences (Appendix 3 and 4). The *A. ovis* and *A. phagocytophilum* (GenBank accession number AF393742 and AF530197 respectively) were used as outgroups.

A total of 5 phylogenetic trees were constructed to determine the relatedness of South African *A. marginale* strains with strains from West, East and North Africa, North and South America, Europe, Asia and Australia (Appendix 4). One phylogenetic tree was constructed based on the *msp1a* gene sequences using the maximum likelihood method (Figure 12). Two phylogenetic trees were constructed based on *msp4* gene sequences using neighbor- joining method and the other two using maximum-likelihood method (Figure 13a and 14a). One phylogenetic tree for each neighbor-joining and maximum likelihood method was constructed from South African province isolates and the other with South African province isolates together with the sequences obtained from West, East and North Africa, North and South America, Europe, Asia and Australia the obtained from the GenBank (Figure 13a and 13b) (Figure 14a and 14b) (Appendix 4). *Anaplasma ovis* and *A. phagocytophilum* were used as outgroups for the construction of phylogenetic tree based on *msp4* gene sequences.

5.3 Results

Previous studies revealed that the gene and protein sequences of *msp1a* and *msp4* nucleotide sequences have been used to infer the phylogenetic relationships among Oklahoma and New World isolates from Argentina, Brazil, Mexico and the United States. All the 11 *A. marginale* isolates collected from Oklahoma had different *msp1a* sequences but identical to *msp4* sequences (de la Fuente *et al.*, 2003a). Phylogenetic studies of South African isolates using *msp4* gene sequences demonstrated that *A. marginale* isolates had a 100% and 99% identity with the sequences from West, East and North Africa, North and South America, Europe, Asia and Australia sequences available in the GenBank (Appendix 4).

5.3.1 Phylogenetic analysis of *A. marginale* isolates using *msp1a* gene

5.3.1.1 MSP1a sequence analysis

The *msp1a* sequences of 44 *A. marginale* isolates were analysed as shown in Table 6. Differences were found in the tandem repeat sequences and structure of *msp1a* genes among the isolates included in this study. The *msp1a* tandem repeats gave 23 new sequences with amino acid changes (Figure 6) resulting in identification of forty four different strains of *A. marginale* (Table 6). The strain containing the tandem repeat structure 27/13/4/4/37 was found twice as KZN-K (Table 6). Tandem repeats 3, 13, 27, 34 and 37 were common in the study areas in all provinces and tandem repeat 42 and 43 were highly represented in KwaZulu-Natal (Table 6).

A pattern of proportional decrease in the % of homology of R1 with the subsequent tandem repeats (R1 + *n*) was obtained. The % of homology between R1 and the tandem repeat in position (R1 + *n*) will have a decrease proportional to the % of homology between R1 and the tandem repeat in position (R1 + 1) (Figure 11). Differential pattern of proportional decrease in amino acid homology was found in KwaZulu-Natal (Figure 11). On the other hand, the percent of homology differed among the tandem repeats in the provinces in Figure 11a (approximately 90% and 70% respectively).

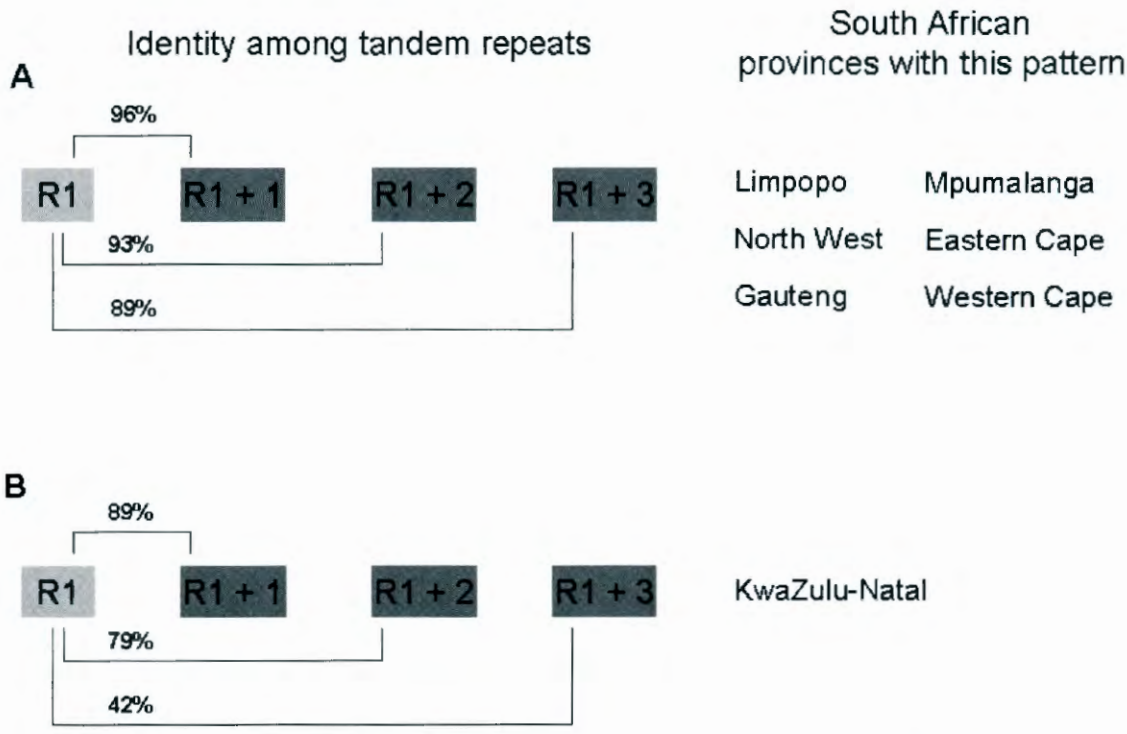


Figure 11: Proportional identity between tandem repeat in position R1 and R1+n.

This pattern was used to suggest a flow of transmission/migration from Eastern Cape to KwaZulu-Natal. The strain containing the tandem repeat structure 27/13/4/4/37 was found twice, once in KwaZulu-Natal and then again in the Eastern Cape (Figure 12). Taking in account that this strain does not follow the pattern of proportional decrease found in KwaZulu-Natal, it was concluded that this strain was introduced from Eastern Cape and this conclusion was in agreement with the maximum likelihood phylogenetic analysis of strains found in KwaZulu-Natal, where KZN-K (27/13/4/4/37) was the only one apart the cluster formed by the other strains from that province (Figure 12).

Table 9: *Anaplasma marginale* strains and putative 2D structures of MSP1a tandem repeats

A. marginale strain	Province of origin	Structure of *msp1a* tandem repeats	MSP1a tandem repeats 2D structure
LP-7	Limpopo	34 15	(α-α) (α-α)
LP-10	Limpopo	27 13 3 36	(α-α)(α-α)(α-c)(c-α)
LP-30	Limpopo	27 13 3	(α-α)(α-α)(α-c)
LP-34	Limpopo	34 13 3 38	(α-α)(α-α)(α-c)(α-c)
LP-37	Limpopo	27 13 13 37	(α-α)(α-α)(α-α)(α-c)
LP-46	Limpopo	3 38	(α-c)(α-c)
LP-50	Limpopo	34 13 13	(α-α)(α-α)(α-α)
MP-C2	Mpumalanga	34 13 158 37	(α-α)(α-α) (α-α)(α-c)
MP-C5	Mpumalanga	15 15 100 83	(α-c)(α-c)(α-c)(α-c)
NW-C2	North West	27 13 4 4 37	(α-α)(α-α)(α-β)(α-β)(α-c)
NW-C4	North West	27 13 4 37	(α-α)(α-α)(α-β)(α-c)
NW-C5	North West	82 13 79 4 37	(α-c)(α-α)(α-α)(α-β)(α-c)
NW-C1-160312	North West	3 13 3 36 38	(α-α)(α-α)(α-c)(c-α)(α-c)
NW-C4-160312	North West	34 36 38 3	(α-α)(c-α)(α-c)(α-c)
GP-C1	Gauteng	82 13 4 4 37	(α-c)(α-α)(α-β)(α-β)(α-c)
GP-C2	Gauteng	34 27 3 38 13 3 38	(α-α)(α-α)(α-c)(α-c)(α-α) (α-c)(α-c)
GP-C5	Gauteng	3 4 4 4 37	(α-c)(α-β)(α-β)(α-β)(α-c)
GP-C1112105	Gauteng	34 37	(α-α)(α-c)
GP-C4117105	Gauteng	3 36 38	(α-c)(c-α)(α-c)
GP-C7117105	Gauteng	34 13 13	(α-α)(α-c)(α-c)
GP-C1817105	Gauteng	34 13 37	(α-α)(α-c)(α-c)
KZN-D	KwaZulu-Natal	42 43 2 1 3 31	(α-c)(α-c)(α-c) (α-α)(α-α)
KZN-F	KwaZulu-Natal	4 43 25 31 31	(α-c)(α-c)(α-c)(α-α)(α-α)
KZN-K	KwaZulu-Natal	27 13 4 37	(α-α) (α-α) (α-β)(α-β)(α-c)
KZN-Y	KwaZulu-Natal	143 144 145 146	(αα)(α-c)(α-c)(α-α)
KZN-MM	KwaZulu-Natal	42 43 25 31	(α-c)(α-c)(α-c)(α-α)
KZN-14	KwaZulu-Natal	142 43 25 31	(c)(α-c)(α-c)(α-α)
KZN-19	KwaZulu-Natal	141 140 140	(α-c)(β-c)(β-c)
KZN-49	KwaZulu-Natal	147 148 149 1 0	(α-c)(α-c)(α-c)(α-c)
KZN-51	KwaZulu-Natal	147	(α-c)
EC-22	Eastern Cape	27 13 4 3	(α-α)(α-α)(α-β)(α-β)(α-c)
EC-23	Eastern Cape	151 1 2 4 4 153	(α-α)(α-α-α)(α-β)(α-β)(αα)
EC-24	Eastern Cape	27 13 4	(α-α)(α-α) (α-β)
WC-4	Western Cape	4 Q Q m	(α-α)(α-α)(α-α)(α-c)
WC-6	Western Cape	3 4 4 37	(α-c)(α-β)(α-β)(α-c)
WC-7	Western Cape	M M M M	(α-c)(α-c)(α-c)(α-c)

WC-8	Western Cape	34 4 37	(α-α)(α-β)(α-c)
WC-10	Western Cape	154	(α-c)
WC-11	Western Cape	40 Q Q Q Q Q 37	(α-α)(α-α)(α-α)(α-α)(α-α) (α-α)(α-c)
WC-12	Western Cape	27 13 37	(α-α)(α-α)(α-c)
WC-13	Western Cape	M Q M Q M	(α-c)(α-α)(α-c)(α-α) (α-c)
WC-14	Western Cape	155 36 38	(αα)(c-α)(α-c)
WC-15	Western Cape	161 13 37 4 162	(α-α)(α-α)(α-c)(α-β)(α-α)
WC-16	Western Cape	34 13 4 13 13 4 37	(α-α)(α-α)(α-β)(α-α)(α-α) (α-β)(α-c)

The *A. marginale* strains detected in this study are presented as *msp1a* tandem repeats organization. Limpopo (LP), Mpumalanga (MP), North West (NW), Gauteng (GP), KwaZulu-Natal (KZN), Eastern Cape (EC) and Western Cape (WC). The code-numbers for tandem repeats are shown in Figure 2. The putative secondary structure of every tandem repeat is shown as two separated alpha helices (α-α), one long alpha helix (αα), beta strand (β) and coiled structure (c).

The pattern found in LP, MP, NW, GP, KZN, EC and WC provinces is listed (Figure 11); reductions of approximately 3% were observed among *A. marginale msp1*a tandem repeats. The *A. marginale* strain NW-C4 served as a model (A). Sequences from KwaZulu-Natal did not follow this pattern and reductions of approximately 10% and 40% were observed. The *A. marginale* strain KZN-14 served as a model (B).

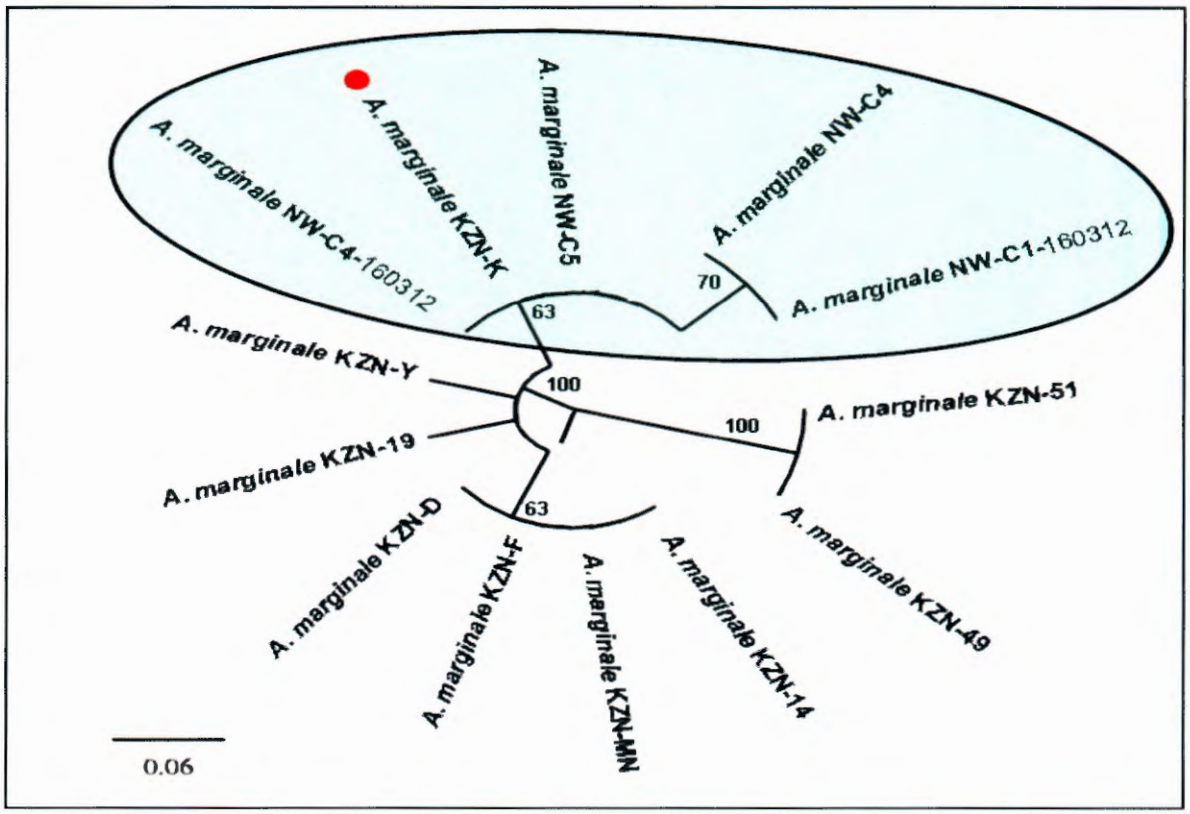


Figure 12: Maximum- likelihood phylogenetic tree based on the MSP1a from the strains present in KwaZulu-Natal and North West provinces. In this tree the strain *A. marginale* KZN-K (red dot) was in the cluster from North West sequences (blue circle). Bootstrap values are shown as a percentage in the internal branch. Only bootstrap values higher than 50% were shown. NJ phylogenetic analysis gave similar results (Data not shown).

5.3.2 Phylogenetic analysis using *msp4* gene sequences

The *msp4* gene of *A. marginale* isolates from Limpopo, Mpumalanga, North West, Gauteng, KwaZulu-Natal, Eastern Cape, Northern Cape and Western Cape provinces of South Africa were amplified by PCR and then sequenced (Figure 10). Positive results were not obtained from Northern Cape samples. Two samples from each province were randomly selected for phylogenetic analysis. Four phylogenetic trees were each constructed for *A. marginale* isolates randomly selected from the LP, MP, GP, NW, GP, KZN, EC and WC provinces of South Africa (Figure 13a and 14a) (Appendix 3), and those from the other countries in Africa and around the world (Figure 13b and 14b) (Appendix 4). *Anaplasma ovis* and *A. phagocytophilum* were included as outgroup taxa for the phylogenetic analysis based on the *msp4* gene.

5.3.3 Phylogenetic analysis of *msp4* gene using neighbor -joining phylogenetic tree

A complete *msp4* gene was obtained after sequencing using the primers designed previously to amplify *msp4* gene (Appendix 3). The *A. marginale msp4* gene was amplified and sequenced isolates infecting cattle from LP, MP, NW, GP, KZN, EC and WC and the results revealed that this nucleotide differed in length and identity.

Anaplasma marginale isolates from LP, GP, NW, KZN and WC had seven different nucleotides from MP and EC isolates at position 163, 197, 223, 239, 251, 281 and 725 with the following changes A to G, G to A, C to A, G to A, A to G, G to A and G to A respectively (Table 10). These changes gave to two clades from South African isolates and were classified as South African first clade and South African second clade. The first South African clade consists of isolates from LP, GP, NW, KZN, and WC and the second South African clade consist of isolates from MP and EC (Figure 13a, 13b, 14a and 14b). The first clade had 99% bootstrap support and was the same as the second clade, thus demonstrating that South African *A. marginale* isolates are more closely genetically related.

Table 10: Differences in the *A. marginale msp4* nucleotide sequences of LP,NW,GP,KZN,WC and MP, EC isolates of South Africa

Nucleotide position							
Isolates	163	197	223	239	251	281	725
LP	A	G	C	G	A	G	G
NW	A	G	C	G	A	G	G
GP	A	G	C	G	A	G	G
KZN	A	G	C	G	A	G	G
WC	A	G	C	G	A	G	G
MP	G	A	A	A	G	A	A
EC	G	A	A	A	G	A	A

In most of the *A. marginale msp4* isolates that have been sequenced, a 849 bp nucleotides with coding sequences were amplified.

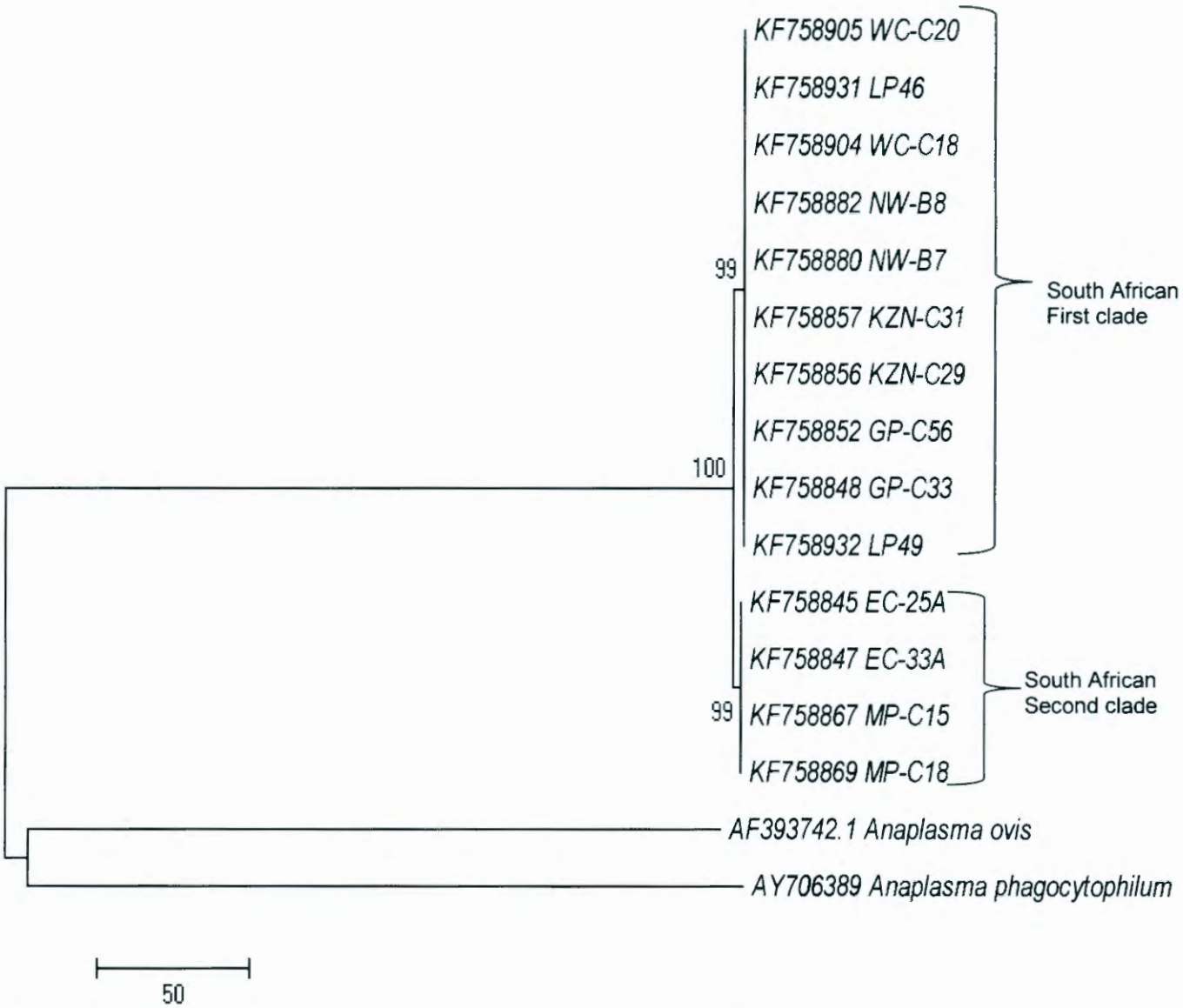


Figure 13a: Neighbor-joining phylogenetic tree of *A. marginale msp4* gene sequences from strains identified in cattle in South Africa. The phylogenetic tree was implemented in the MEGA5 (Tamura *et al.*, 2011). Bootstrap analysis was conducted with 1000 replicates. The GenBank accession numbers of the respective sequences used for the phylogenetic analysis are indicated at the beginning of the name of each sequence.

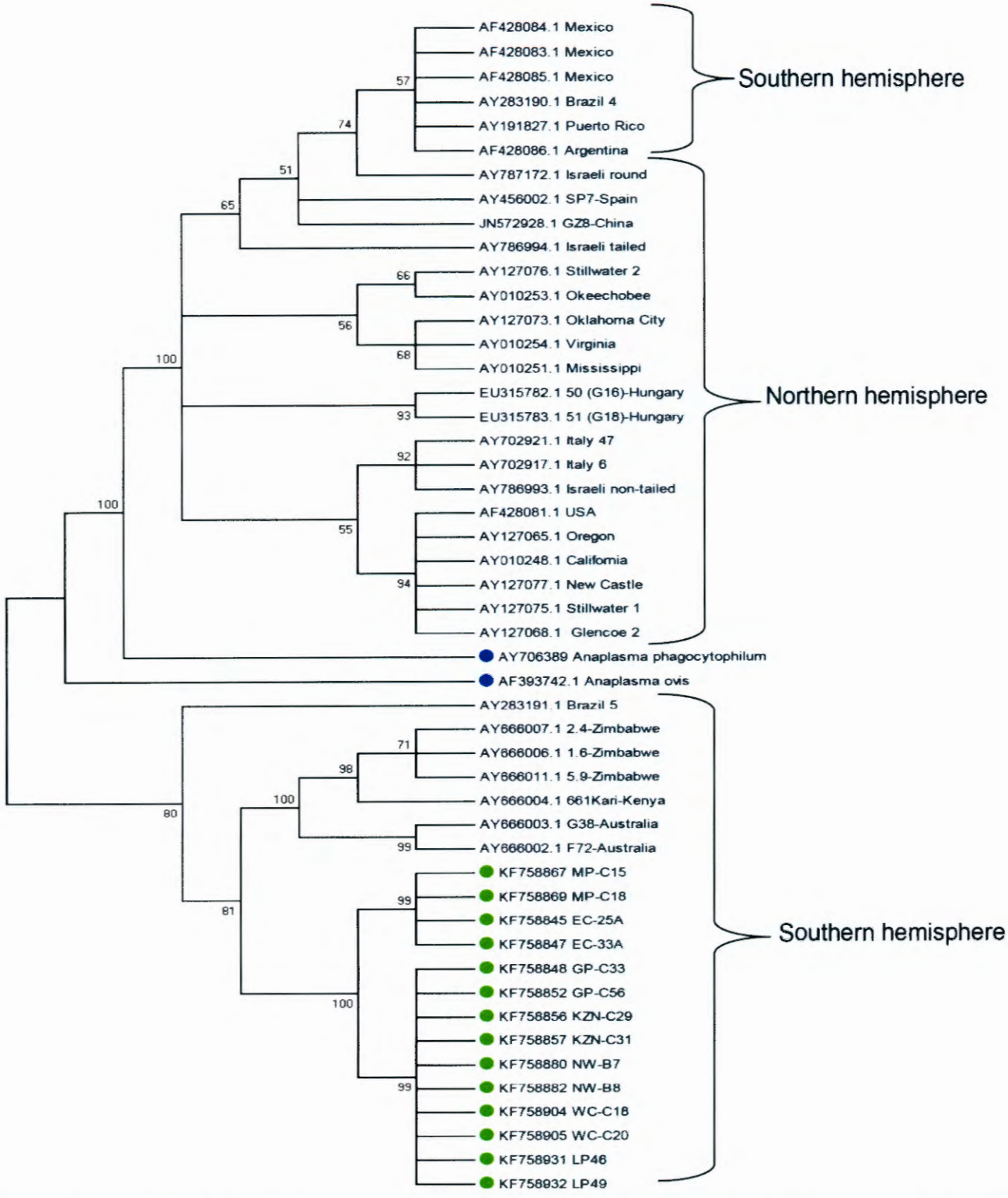


Figure 13b: Neighbor-joining phylogenetic tree of *A. marginale msp4* gene sequences from strains identified in cattle in South Africa represented with a green dot and blue dot representing an outgroup (*A. ovis* and *A. phagocytophilum*) including isolates from East and North Africa, North and South America, Europe and Australia. The phylogenetic tree was implemented in the MEGA5 (Tamura *et al.*, 2011). Bootstrap analysis was conducted with 1000 replicates. The GenBank accession numbers of the respective sequences used for the phylogenetic analysis are indicated at the beginning of the name of each sequence. Only bootstrap values higher than 50% are shown.

5.3.4 The phylogenetic analysis of *msp4* gene using maximum-likelihood phylogenetic tree

Observations made from the maximum-likelihood phylogenetic trees in Figure 13a further verify the results of the neighbor-joining phylogenetic tree. The only difference was that in the second clade the number of bootstrap support significantly decreased. In Figure 13b, the difference was noted in the first and second clade when isolates from other countries were included with a bootstrap support of 97, respectively. However, the South African strains fell into a separate cluster from the rest of the world and other African countries.

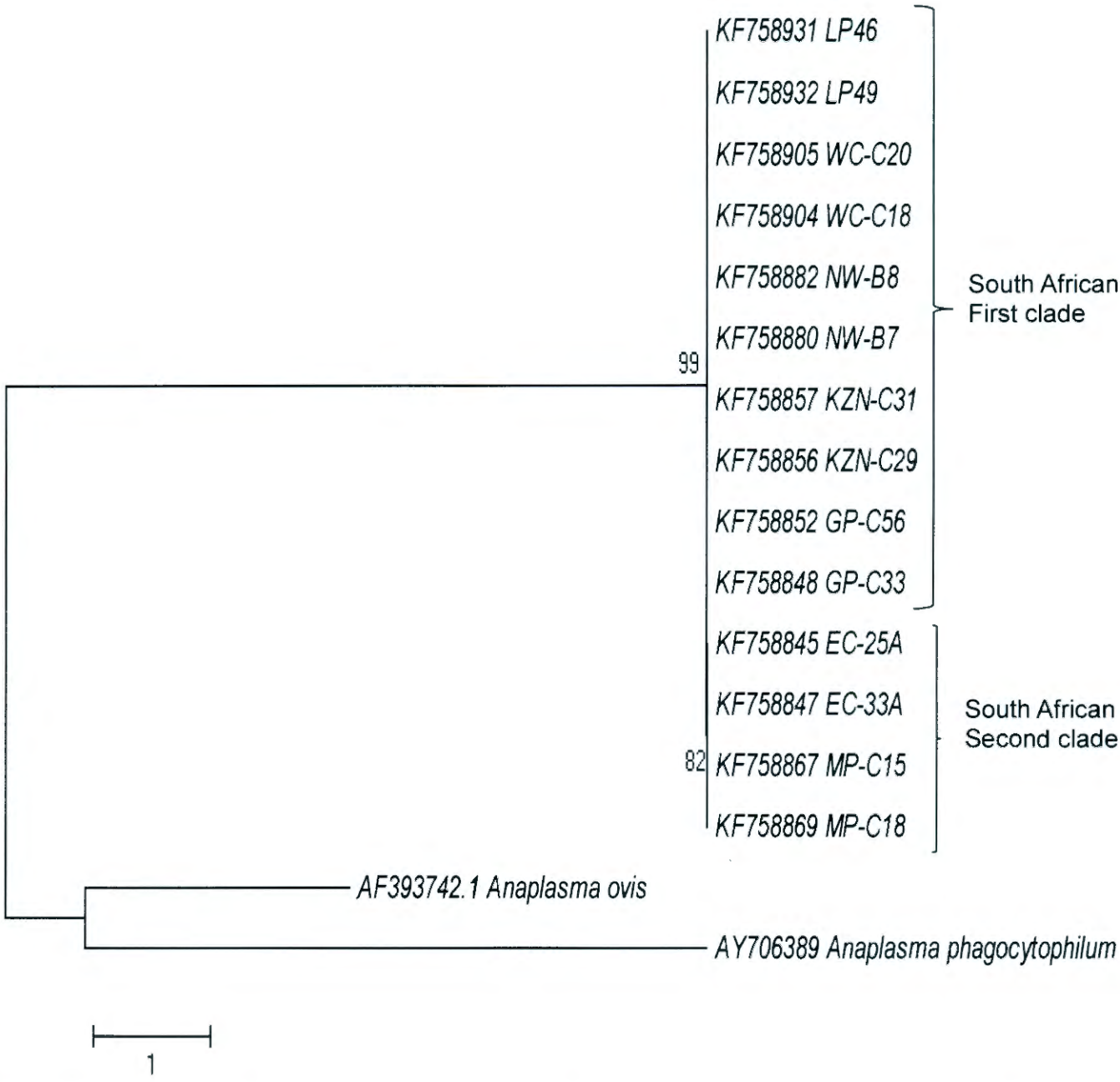


Figure 14a: Maximum-likelihood phylogenetic tree of *A. marginale msp4* gene sequences from strains identified in cattle in South Africa. The phylogenetic tree was implemented in the MEGA5 (Tamura *et al.*, 2011). Bootstrap analysis was conducted with 1000 replicates. The GenBank accession numbers of the respective sequences used for the phylogenetic analysis are indicated at the beginning of the name of each sequence.

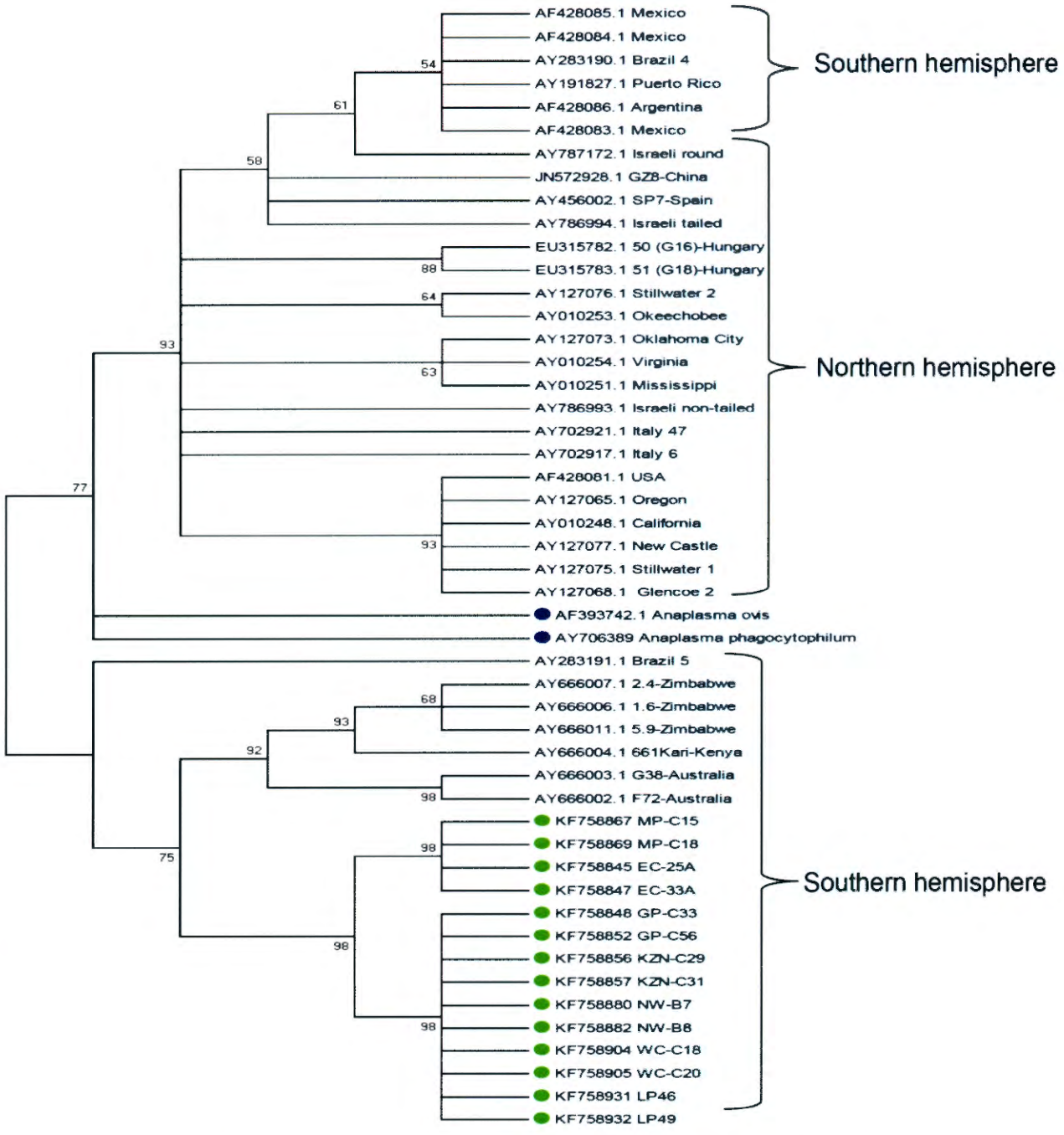


Figure 14b: Maximum-likelihood phylogenetic tree of *A.marginale msp4* gene sequences from strains identified in cattle in South Africa represented with a green dot and blue dot representing an outgroup including isolates from East and North Africa, North and South America, Europe, Asia and Australia. The phylogenetic tree was implemented in the MEGA5 (Tamura *et al.*, 2011). Bootstrap analysis was conducted with 1000 replicates. The GenBank accession numbers of the respective sequences used for the phylogenetic analysis are indicated at the beginning of the name of each sequence. Only bootstrap values higher than 50% are shown.

5.4 Discussion

The use of molecular data for inferring phylogenetic trees has now gained considerable interest among biologists of different disciplines, and is often used in addition to morphological data to study relationships in further detail. The number of repeats found in *msp1a* has been used to characterize geographic isolates of *A. marginale* (Allred *et al.,* 1990; de la Fuente *et al.,* 2002a). In a previous study of *A. marginale* isolates from the United States, de la Fuente and colleagues (2003a) concluded that *msp1a* is not a good marker for the characterization of *A. marginale* geographic isolates and suggested that the genetic heterogeneity observed among *A. marginale* isolates of Oklahoma could be better characterized by use of *msp4* gene and protein sequences.

Furthermore, concurrent analysis of *msp1a* and *msp4* gene sequences provided phylogeographic information (de la Fuente *et al.,* 2003a). On a broader geographic scale, analysis of New World isolates demonstrated that *msp4* sequences, but not *msp1a* DNA or protein sequences, provide phylogeographic information (de la Fuente *et al.,* 2007a). These results, together with the finding that multiple *msp1a* genotypes circulate in some geographic regions (e.g., an area of endemicity in Oregon; (Palmer *et al.,* 2001), questioned the use of *msp1a* sequences for identification of geographic isolates of *A. marginale* and suggested that *msp1a* sequences may evolve more rapidly (de la Fuente *et al.,* 2007a).

Overall, these results suggested that *msp1a* genotypes most likely reflect the history of cattle movement rather than the geographic distribution of *A. marginale* isolates. Recent results have shown that multiple *A. marginale* genotypes were maintained within a herd in an area of endemicity by independent transmission events and that infection with more than one genotype per host may have been prevented by a mechanism of infection exclusion (Palmer *et al.,* 2001; de la Fuente *et al.,* 2007a; Pohl *et al.,* 2013; Cabezas-Cruz *et al.,* 2013; Mutshembele *et al.,* 2014).

Therefore, if cattle persistently infected with a new *A. marginale* genotype were introduced to a herd, these genotypes could become established by mechanical and/or biological transmission to susceptible cattle. In regions with few introductions of *A. marginale* isolates such as Australia, genotypic variation was found to be minimal (Bock and de Vos, 2001; Lew *et al.*, 2002). In regions like Oklahoma, where the movement of cattle has been extensive, a highly heterogeneous population of *A. marginale* isolates would be expected (de la Fuente *et al.*, 2007a).

In summary, as demonstrated in previous studies (de la Fuente *et al.*, 2005c), the *msp4* sequences of *A. marginale* provided sufficient variation to provide phylogeographic relationships of *A. marginale* isolates on a geographic scales. In contrast to the results obtained with *msp4*, DNA and deduced amino acid sequences of *msp1a* failed to resolve phylogeographic patterns of *A. marginale* isolates and suggested that cattle movement and the maintenance of different genotypes by independent transmission events explain the genetic heterogeneity observed among isolates of *A. marginale* in Oklahoma and, possibly, in other regions of endemicity (de la Fuente *et al.*, 2007a).

5.4.1 Phylogenetic analysis of *A. marginale* using *msp1a* gene

The tandem repeat region of *msp1a* is useful as a genetic marker for analysis of *A. marginale* evolutionary and diversity in endemic areas (de la Fuente *et al.*, 2007a). As shown previously, *A. marginale* is endemic in South Africa (de Waal, 2000). This study used a sensitive and specific *msp1a* PCR primer pair (Lew *et al.*, 2002) to study the epidemiology of *A. marginale* in 8 different provinces of South Africa as well as the structural analysis of *msp1a*. As was found in previous studies (Pohl *et al.*, 2013; Cabezas-Cruz *et al.*, 2013; Mutshembele *et al.*, 2014) the *msp1a* sequences of *A. marginale* isolates from South Africa contained a variable number of tandem repeats in the amino-terminal region of the protein, while the remainder of the protein was highly conserved.

Newly described tandem repeats from South African isolates were reported for first time in this study (Figure 6) suggesting strains that are unique to this geographical region. The presence of different *msp1a* genotypes identified in different regions (Cabezas-Cruz *et al.*, 2013) suggest that *msp1a* sequences, although conserved during multiplication of the parasite in the bovine host and after transmission by ticks (Palmer *et al.*, 2001; Bowie *et al.*, 2002) are very variable and may contribute to the definition of genotype variations within cattle. This heterogeneity among *msp1a* gene sequences could be explained by the tandem repeats within a region of high mutability, which would be supported by the frequency of variable amino acid positions within geographical strain that are higher in this region than in the rest of the protein (Bowie *et al.*, 2002).

Nonetheless analysis of this diversity should take into account cattle movement and the consecutive multiple introductions of *A. marginale* strains in a given region (de la Fuente *et al.*, 2010) which could be a non-evolutive mechanism that contributes to the local genetic variability. Despite the fact that no geographic phylogenetic relationship was found using *A. marginale msp1a*, the percent of identity among *msp1a* tandem repeats in two different regions differed in this study, which suggest that different genetic mechanisms may have been involved in the *msp1a* amino acid variability.

Using this phylogenetic pattern South African *A. marginale* strains EC22 and NW-C2 (27/13/4/4/37) were likely introduced in KwaZulu-Natal from Eastern Cape or North West provinces. This supports the evidence that the strains do not fall in the phylogenetic cluster formed by the other sequences from KwaZulu-Natal and notably, these strains contained tandem repeat number 4 with primarily ß-strand as secondary structure. Recently, the presence of ß- strand in *msp1a* tandem repeats was found to be phylogenetically correlated with non-tick transmission phenotype of *A. marginale* (Cabezas-Cruz *et al.*, 2013).

This is in contrast with the predominance of α-helices in South African strains that correlate with tick transmission phenotype of *A. marginale* (Cabezas-Cruz *et al.*, 2013). Several tick species have been incriminated in *A. marginale* transmission in South Africa (Marufu, 2014) with mechanical and transplacental transmission being reported (Aubry and Geale, 2011), which may likely also contribute to transmission of South African EC22-(27/13/4/4/37) strain.

Anaplasmosis is widespread in KwaZulu-Natal province (du Plessis *et al.*, 1994) and some of the tandem repeats found in our study has also been reported in other regions of the world (Vidotto *et al.*, 2006; de la Fuente *et al.*, 2007a; Mtshali *et al.*, 2007). Although tandem repeats including 25, 31, 42 and 43 are common in *A. marginale* isolates in KwaZulu-Natal, a higher proportion of new South African tandem repeats were found in this study. This data suggested that conditions in this region promoted high genetic variability of *A. marginale*, a finding in agreement with the discovery that tandem repeats in this province had the lowest amino acid identity. High infection rates by *A. marginale*-isolates were observed in Mpumalanga, Gauteng and Eastern Cape. The high prevalence is attributed to heavy-tick infestation previously reported in the areas by Horak *et al* (1991), Walker (1991) and Coetzer *et al* (1994). In particular *R. decoloratus, R. microplus* and *R. evertsi evertsi* have been reported in these regions and some of these species transmit *A. marginale*.

The 2D structure of the *msp1a* tandem repeats in these provinces was predominantly α-helix, which would be expected in areas heavily infested by these. In contrast, the lack of *A. marginale* in Northern Cape would be expected in areas where conditions and intensive tick control programs contributed to low tick populations. Amino acid variability in individual positions along the 29 amino acids of all strains suggested that this position is the most variable among isolates worldwide (Cabezas-Cruz *et al.*, 2013). In addition, this position has a high level of negatively-charged amino acids, a region which is associated with transmissibility by ticks (Mutshembele *et al.*, 2014).

This study has used structural epidemiology to establish for the first time *A. marginale msp1a* structural patterns for different geographic regions. Putative secondary structure, identity, and phylogenetic analysis of the *msp1*a tandem repeats were used to infer movement of *A. marginale* strains and regional specific properties of *A. marginale msp1a.* Therefore, the findings of the present study will form a basis for future research where the epidemiology of *A. marginale* can be correlated with the structure of *msp1a* and tick transmissibility.

5.4.2 Phylogenetic analysis of *A. marginale* using *msp4* gene

The analysis of *msp4* gene, which has provided evolutionary information about geographically distinct *A. marginale* strains (de la Fuente *et al.,* 2007a), was used in this study for phylogenetic analysis for isolates infecting cattle from Limpopo, Mpumalanga, North West, Gauteng, KwaZulu-Natal, Eastern Cape and Western Cape. Two clades were observed, which consisted of the first (LP, NW, GP, KZN and WC) and the second (MP and EC) clade isolates.

Figure 13a represent the neighbor-joining tree of the isolates from MP and EC formed a cluster which was isolated from LP, NW, GP, KZN and WC with well supported bootstrap values of 99%. However, it shared high genetic similarities with both clades with a high bootstrap support value of 99%.

A maximum likelihood phylogenetic tree revealed strong bootstrap of 99% for a clade containing isolates of *A. marginale msp4* gene from LP, NW, GP, KZN and WC. A strong bootstrap support of 82% from MP and EC isolates of South Africa (Figure 13b). These results were also observed in phylogenetic tree of *A. marginale msp4* gene sequences from first clade (LP, NW, GP, KZN and WC) and the second clade (MP and EC) isolates of South Africa, North Africa, North and South America, Europe, Australia as well as Asia.

These phylogenetic trees were constructed using neighbor-joining and maximum-likelihood method (Figure 13a and 13b). Sequence comparison of the *msp4* gene was recognized as one of the most powerful and precise methods for determining the phylogenetic relationships of *A. marginale* (de la Fuente *et al.*, 2007a).

South African strain had 100% nucleotide identity with isolates from Kenya, Zimbabwe and Australian. Even though they had highly shared the identity, South African strains remained in separate clusters. These strains shared 99% identity with isolates from North and South America, Europe and Asia (Appendix 4). Good representation of southern and northern hemisphere strains was observed, which further demonstrated that *msp4* gene may serve as good marker for phylogeographic analysis (Figure 13a and 13b). In addition, these result suggest that South African cattle could be be infected with the same strains that infect cattle from West, East and North Africa, North and South America, Europe, Asia as well as Australia.

The DNA sequences of *msp4* gene revealed phylogeographical segregation with South African isolates and other isolates submitted in the NCBI database. MSP4 gene sequences provided phylogenetic resolution for intraspecific relationships among isolates. South African isolates demonstrated apparent phylogeographic information. The present study revealed that *msp4* nucleotide sequences are sufficiently variable and like in other studies could be used to detect phylogeographic patterns broadly (de la Fuente *et al.*, 2010). Therefore, the msp4 gene may be used as a marker for understanding the phylogeographic patterns and phylogenetic relationships of *A. marginale* in South Africa.

CHAPTER 6

GENERAL DISCUSSION, CONCLUSIONS AND RECOMMENDATIONS

6.1 General discussion

The prevalence of bovine anaplasmosis in the various South African regions is often complex and may involve the bovine population immunity to *A. marginale*, the means of transmission (mechanical, biological by ticks, transplacental or a combination thereof), as well as the genetic variability of *A. marginale* strains within a given area. In this study, data was not collected on the immunity of cattle populations for *A. marginale*. Certainly, tick transmission plays a very important role in the prevalence of *A. marginale*, and the distribution of tick species in South Africa has been determined by de Waal, (2000). The provinces with 100% prevalence of *A. marginale* infected cattle did not correlate with provinces with the highest or lowest *A. marginale* genetic diversity, but instead represented a range of diversity (Figure 7). It has been hypothesized that increased genetic variability may have an impact on *A. marginale* strains transmission (Palmer and Bryton, 2013).

Alternately, *A. marginale* strains with low genetically variability may be lost after immune clearance. Interestingly, in the Western Cape the prevalence of *A. marginale* was high while the genetic diversity was low, which could be due to low immunity of cattle population in the particular sample points. Overall, the prevalence of *A. marginale* in cattle from South Africa is high when compared with other regions of the world. In Brazil a recent study found prevalence of *A. marginale* in a cattle herd to be 70% (Pohl *et al.*, 2013). Another study by de la Fuente *et al.* (2005a) reported the prevalence of *A. marginale* in cattle herds from Italy to range from 25 - 100%.

In these studies the prevalence and genetic diversity of *A. marginale* were found to have a polynomial, and nonlinear correlation rather than a linear one. Subsequently, this finding most likely reflects the complexity of factors that may be influencing the prevalence of *A. marginale* in the various regions of South Africa as evident in the province of Limpopo, North West and KwaZulu-Natal. This polynomial correlation suggests that increased genetic diversity impacted on the transmissibility of *A. marginale* strains. The other major factor that could influence *A. marginale* prevalence in a given region is the immunity of a population.

This factor is also related to genetic diversity because, as shown previously, a mechanism of infection exclusion may prevent cattle infected with one genotype from becoming infected with a second closely related one. However, widely separated genotypes apparently can be maintained in cattle as co-infections which could contribute to wider genetic diversity at the population level and impact on the prevalence of bovine anaplasmosis. *Anaplasma marginale* with low genetic variability will likely become highly prevalent in a cattle population with predominately naïve animals, without immunity as evidenced in the cattle population in the Western Cape.

6.2 Conclusions

The regional prevalence of bovine anaplasmosis appears to be a result of complex factors, including the immunity of cattle populations to *A. marginale*, the mode of transmission (biological by ticks, mechanical and/or transplacental) and the genetic variability of *A. marginale* strains within cattle populations. These factors have contributed to the challenges of developing effective diagnostic and control measures for *A. marginale*. Molecular tools are required for evaluating the genetic diversity of *A. marginale*. Developing an understanding of the molecular mechanisms underlying the generation of genetic diversity is crucial in implementation of effective control measures that are developed in concern with evolutionary changes of the pathogen which do not stay one step behind in pathogen evolution.

The first step in this molecular study of bovine anaplasmosis was to confirm that *A. marginale* was widespread in various regions in South Africa. The genetic diversity of *A. marginale msp1a* was then characterized, and found to arise from the evolution of extent tandem repeats, which under positive and negative selection, diversified as noted by the presence of newly reported tandem repeats variants. These new variants were most likely constricted by two forces: (1) the immune system and (2) biological transmission of *A. marginale* by ticks. Common *msp1a* tandem repeats were found among *A. marginale* strains in South Africa and other regions of the world which, in combination with the use of low variable *msp1a* peptides, will likely be useful in the development of *msp1a* based vaccine for more effective regional control of anaplasmosis.

Gene sequences of *msp4* provided sufficient variation to allow for the characterization of phylogeographic patterns on a broad scale. The South African clade of *A. marginale msp4* gene sequences included isolates from West, East and North Africa, North and South America, Europe, Australia and Asia (Appendix 4). Finally the detection of diverse clades of *A. marginale* representing different geographic regions including the results that were obtained based on *msp1a* gene sequences may provide vital information for developing new novel vaccine approaches for control of anaplasmosis.

This work presents information on the prevalence, genetic diversity and phylogenetic analysis of *A. marginale* in cattle throughout South Africa. Data compiled in this study will aid in the future management of *A. marginale* infection and may provide a molecular basis for targeted vaccine development. The information gained from this research provides an understanding of the epidemiology of *A. marginale* in South Africa which will be fundamental toward the development of diagnostic assays and control measures for bovine anaplasmosis.

6.3 Recommendations

6.3.1 To further determine the species, distribution and population densities of the known tick vectors of *A. marginale* in the provinces of South Africa, and to correlate these tick populations with the prevalence of bovine anaplasmosis.

6.3.2 To characterize *A. marginale* strains in cattle populations in order to determine the strain diversity, and to conduct phylogenetic studies of these strains to gain insight into the origin and evolution of *A. marginale* in South Africa.

6.3.3 To identify the tick species infesting cattle in anaplasmosis endemic regions, and to determine the immunity of cattle in these areas for *A. marginale* infections.

6.3.4 To evaluate genes and gene products of the *A. marginale* strains, and to determine whether the diversity of *msp4* gene is related to selective pressures.

REFERENCES

AGNES, J. A., BRAYTON, K. A., LAFOLLETT, M., NORIMINE, J., BROWN, W. C. and PALMER, G.H. (2011). Identification of *Anaplasma marginale* outer membrane protein antigens conserved between sensu stricto strains and the live *Anaplasma marginale* ss. *centrale* vaccine. *Infection and Immunity*. **79:** 1311–1318.

AL-ADHAMI, B., SCANDRETT, W. W. B., LOVANOV, V. and GAJADHAR, A. A. (2011). Serological cross reactivity between *Anaplasma marginale* and *Ehrlichia* species in naturally and experimentally infected cattle. *Journal of Veterinary Diagnostic Investigation*. **23:** 1181-1188.

ALLEMAN, A. R. and BARBET, A. F. (1996). Evaluation of *Anaplasma marginale* major surface protein 3 (MSP3) as a diagnostic test antigen. *Journal of Clinical Microbiology*. **34:** 270–276.

ALLEMAN, A. R., PALMER, G. H., MCGUIRE, T. C., MCELWAIN, T. F., PERRYMAN, L. E. and BARBET, A. F. (1997). *Anaplasma marginale* major surface protein 3 is encoded by a polymorphic, multigene family. *Infection and Immunity*. **65:**156-163.

ALLRED, D. R., MCGUIRE, T. C., PALMER, G. H., LEIB, S. R., HARKINS, T. M., MCELWAIN, T. F. and BARBET, A. F. (1990). Molecular basis for surface antigen size polymorphisms and conservation of a neutralization-sensitive epitope in *Anaplasma marginale*. *Proceedings of the National Academy of Science USA*. **87**:3220-3224.

ALMAZÁN, C., MEDRANO, C., ORTIZ, M. and DE LA FUENTE, J. (2008). Genetic diversity of *Anaplasma marginale* strains from an outbreak of bovine anaplasmosis in an endemic area. *Veterinary Parasitology*. **158:** 103-109.

AUBRY, P. and GEALE, D. W. (2011). A review of bovine anaplasmosis. *Transboundary and Emerging Diseases*. **58:** 1–30.

BAKER, M. K., DUCASSE, F. B. W., SUTHERST, R. W. and MAYWALD, G. F. (1989). The seasonal tick populations on traditional and commercial cattle grazed at four altitudes in Natal. *Journal of South African Veterinary Association*. **60:** 95–101.

BARBET, A. F., PALMER, G. H., MYLER, P. J. and MCGUIRE, T. C. (1987). Characterization of an immunoprotective protein complex of *Anaplasma marginale* by cloning and expression of the gene coding for polypeptide AM 105L. *Infection and Immunity*. **55:** 2428-2435.

BARBET, A. F., LUNDGREN, A., JOOYOUNG, Y. I., RURANGIRWA, F. R. and PALMER, G. H. (2000). Antigenic variation of *Anaplasma marginale* by expression of MSP2 mosaics. *Infection and Immunity*. **68:** 6133-6138.

BARBET, A. F., JOOYOUNG, Y. I., LUNDGREN, A., MCEWEN, B. R., BLOUIN, E. F. and KOCAN, K. M. (2001). Antigenic variation of *Anaplasma marginale*: major surface protein 2 diversity during cyclic transmission between ticks and cattle. *Infection and Immunity*. **69:** 3057-3066.

BOCK, R. E. and DE VOS, A. J. (2001). Immunity following use of Australian tick fever vaccine: a review of the evidence. *Australian Journal of Veterinary.* **79:** 832-839.

BOWIE, M. V., DE LA FUENTE, J., KOCAN, K. M., BLOUIN, E. F. and BARBET, A. F. (2002). Conservation of major surface protein 1 genes of *Anaplasma marginale* during cyclic transmission between ticks and cattle. *Gene.* **282:** 95– 102.

BRADWAY D. S., TORIONI DE ECHAIDE, S., KNOWLES, D. P., HENNAGER, S. G. and MCELWAIN T. F. (2001). Sensitivity and specificity of the complement fixation test for detection of cattle persistently infected with *Anaplasma marginale*. *Journal of Veterinary Diagnostic Investigation.* **13**, 79–81.

BRAYTON, K. A., KNOWLES, D. P., MCGUIRE, T. C. and PALMER, G. H. (2001). Efficient use of a small genome to generate antigenic diversity in tick-borne *ehrlichial* pathogens. *Proceedings of the National Academy of Science.* **98:** 4130–4135.

BRAYTON, K. A., KAPPMEYER, L. S., HERNDON, D. R., DARK, M. J., TIBBALS, D. L., PALMER, G. H., MCGUIRE, T. C. and KNOWLES JR., D. P. (2005). Complete genome sequencing of *Anaplasma marginale* reveals that the surface is skewed to two superfamilies of outer membrane proteins. *Proceedings of the National Academy of Science.* **102:** 844–849.

BROCK, W. E., KLIEWER, I. O. and PEARSON, C. C. (1965). A vaccine for anaplasmosis. *Journal of the American Veterinary Medical Association.* **147:** 948–951.

BROWN, W. C., ZHU, D., SHKAP, V., MCGUIRE, T. C., BLOUIN, E. F., KOCAN, K. M. and PALMER, G. H. (1998). The repertoire of *Anaplasma marginale* antigens recognized by CD4þ T-lymphocyte clones from protectively immunized cattle is diverse and includes major surface protein 2 (MSP-2) and MSP-3. *Infection and Immunity.* **66:** 5414–5422.

BROWN, W. C., MCGUIRE, T. C., ZHU, D., LEWIN, H. A., SOSNOW, J. and PALMER, G. H. (2001). Highly conserved regions of the immunodominant major surface protein 2 of the gene group II *ehrlichial* pathogen *Anaplasma marginale* are rich in naturally derived CD4+T lymphocyte epitopes that elicit strong recall responses. *Journal of Immunology.* **166:** 1114-1124.

BROWN, W. C., BRAYTON, K. A., STYER, C. M. and PALMER, G. H. (2003). The hypervariable region of *Anaplasma marginale* major surface protein 2 (MSP2) contains multiple immunodominant CD4+ T lymphocyte epitopes that elicit variant-specific proliferative and IFN-γ responses in MSP2 vaccinates. *Journal of Immunology.* **170:** 3790–3798.

BRYSON, N. R., TICE, G. A., HORAK, I. G., STEWART, C. G. and DU PLESSIS, B. J. A. (2002). Ixodid ticks on cattle belonging to small-scale farmers at 4 communal grazing areas in South Africa. *Journal of the South African Veterinary Association.* **73:** 98-103.

BUCHAN, D. W., WARD, S. M., LOBLEY, A. E., NUGENT, T. C., BRYSON, K. and JONES, D. T. (2010). Protein annotation and modelling servers at University College London. *Nucleic Acids Research.* **38:** W563-568.

CABEZAS-CRUZ, A., PASSOS, L. M. F., LIS, K., KENNEIL, R., VALDÉS, J. J., FERROLHO, J., TONK, M., POHL, A. E., GRUBHOFFER, L., ZWEYGARTH, E., SHKAP, V., RIBEIRO, M. F. B., ESTRADA-PEÑA, A., KOCAN, K. M. and DE LA FUENTE, J. (2013). Functional and Immunological Relevance of *Anaplasma marginale* Major Surface Protein 1a Sequence and Structural Analysis. *PloS One.* **8:** 1-13.

CAMACHO NUEZ, M., MŪNOZ, M. L., SUÁREZ, C. E., MCGUIRE, T. C., BROWN, W. C. and PALMER, G. H. (2000). Expression of polymorphic msp1α genes during acute *Anaplasma marginale* rickettsemia. *Infection and Immunity.* **68:**1946-1952.

CARELLI, G., DECARO, N., LORUSSO, A., ELIA, G., LORUSSO, E., MARI, V., CECI, L. and BUONAVOGLIA, C. (2007). Detection and quantification of *Anaplasma marginale* DNA in blood samples of cattle by real-time PCR. *Veterinary Microbiology.* **124:** 107–114.

CARREŃO, A. D., ALLEMAN, A. R., BARBET, A. F., PALMER, G. H., NOH, S. M. and JOHNSON, C. M. (2007). In vivo endothelial cell infection by *Anaplasma marginale*. *Veterinary Pathology.* **44:**116-118.

CHOMEL, B. B., CARNICIU, M. L., KASTEN, R. W., CASTELLI, P. M., WORK, T. M. and JESSUP, D. A. (1994). Antibody prevalence of eight ruminant infectious diseases in California mule and black-tailed deer (Odocoileus hemionus). *Journal of Wildlife Diseases.* **30:** 51–59.

COETZEE, J. F., APLEY, M. D., KOCAN, K. M., RURANGIRWA, F. R., and VAN DONKERSGOED, J. (2005). Comparison of three oxytetracycline regimes for the treatment of persistent *Anaplasma marginale* infections in beef cattle. *Veterinary Parasitology*. **127:** 61–73.

COETZEE, J. F., APLEY, M. D., and KOCAN, K. M. (2006). Comparison of the efficacy of enrofloxacin, imidocarb, and oxytetracycline for clearance of persistent. *Anaplasma marginale* infections in cattle. *Journal of Veterinary Therapeutics*. **7:** 347–360.

COETZEE, J. F., SCHMIDT, P. L., APLEY, M. D., REINBOLD, J. B. and KOCAN, K. M. (2007). Comparison of the complement fixation test and competitive ELISA for serodiagnosis of *Anaplasma marginale* infection in experimentally infected steers. *American Journal of Veterinary Research*. **68:** 872–878.

COETZER, J. A. W., THOMSON, G. R. and TUSTIN, R. C. (1994). Infectious Diseases of Livestock with Special Reference to Southern Africa. Oxford University Press, Cape Town.

COLEMAN, P. G., PERRY, B. D. and WOOLHOUSE, M. E. J. (2001). Endemic stability—a veterinary idea applied to human public health. *Lancet*. **357:** 1284–1286.

D'HAESE, L., PENNE, K. and ELYN, R. (1999). Economics of theileriosis control in Zambia. *Tropical Medicine and International Health*. **4:** 49-57.

DARK, M. J., AL-KHEDERY, B. and BARBET, A. F. (2011) Multistrain genome analysis identifies candidate vaccine antigens of *Anaplasma marginale*. *Vaccine*. **29:** 4923–4932.

DE CASTRO J. J. (1997). Sustainable tick and tick-borne disease control in livestock improvement in developing countries. *Veterinary Parasitology*. **77:** 77–97.

DE LA FUENTE, J., GARCIA-GARCIA, J. C., BLOUIN, E. F., RODRIGUEZ, S. D., GARCIA, M. A., and KOCAN, K. M. (2001a). Evolution and function of tandem repeats in the major surface protein 1a of the *ehrlichial* pathogen *Anaplasma marginale*. *Animal Health Research and Review*. **2:** 163–173.

DE LA FUENTE, J., GARCIA-GARCIA, J. C., BLOUIN, E. F., and KOCAN, K. M. (2001b). Differential adhesin of major surface proteins 1a and 1b of the *ehrlichial* cattle pathogen *Anaplasma marginale* to bovine erythrocytes and tick cells. *International Journal of Parasitology*. **31:** 145–153.

DE LA FUENTE, J., GARCIA-GARCIA, J. C., BLOUIN, E. F., and KOCAN, K. M. (2001c). Major surface protein 1a effects tick infection and transmission of the *ehrlichial* pathogen *Anaplasma marginale*. *International Journal of Parasitology*. **31:** 1705–1714.

DE LA FUENTE, J., VAN DEN BUSSCHE, R. A., and KOCAN, K. M. (2001d). Molecular phylogeny and biogeography of North American strains of *Anaplasma marginale* (Rickettsiaceae: Ehrlichieae). *Veterinary Parasitology*. **97:** 65–76.

DE LA FUENTE, J., VAN DEN BUSSCHE, R. A., GARCIA-GARCIA, J. C., RODRIGUEZ, S. D., GARCIA, M. A., GUGLIELMONE, A. A., MANGOLD, A. J., FRICHE PASSOS, L. M., BLOUIN, E. F. and KOCAN, K. M. (2002a). Phylogeography of New World strains of *Anaplasma marginale* (Rickettsiaceae: *Ehrlichieae*) based on major surface protein sequences. *Veterinary Microbiology*. **88:** 275–285.

DE LA FUENTE, J., GARCIA-GARCIA, J. C., BLOUIN, E. F., SALIKI, J. T., and KOCAN, K. M. (2002b). Infection of tick cells and bovine erythrocytes with one genotype of the intracellular ehrlichia *Anaplasma marginale* excludes infection with other genotypes. *Clinical and Diagnostic Laboratory Immunology*. **9:** 658–668.

DE LA FUENTE, J., GARCIA-GARCIA, J. C., BLOUIN, E. F. and KOCAN, K. M. (2003a). Characterization of the functional domain of major surface protein 1a involved in adhesin of the rickettsia *Anaplasma marginale* to host cells. *Veterinary Microbiology*. **91:** 265–283.

DE LA FUENTE, J., VAN DEN BUSSCHE, R. A., PRADO, T. M. and KOCAN, K. M. (2003b). *Anaplasma marginale msp1a* genotypes evolved under positive selection pressure but are not markers for geographic isolates. *Journal of Clinical Microbiology*. **41:** 1609–1616.

DE LA FUENTE, J., LEW, A., LUTZ, H., MELI, M. L., HOFMANN-LEHMANN, R., SHKAP, V., MOLAD, T., MANGOLD, A. J., ALMAZAN, C., NARANJO, V., GORTAZAR, C., TORINA, A., CARACAPPA, S., GARCIA-PEREZ, A. L., BARRAL, M., OPORTO, B., CECI, L., CARELLI, G., BLOUIN, E. F. and KOCAN, K. M. (2005a). Genetic diversity of *Anaplasma* species major surface proteins and implications for anaplasmosis serodiagnosis and vaccine development. *Animal Health Research and Review*. **6:** 75–89.

DE LA FUENTE, J., TORINA, A., CARACAPPA, S., TUMINO, G., FURLA, R., ALMAZAN, C. and KOCAN, K. M. (2005b). Serologic and molecular characterization of *Anaplasma* species infection in farm animals and ticks from Sicily. *Veterinary Parasitology.* **133:** 357–362.

DE LA FUENTE, J., MASSUNG, R. B., WONG, S. J., CHU, F. K., LUTZ, H., MELI, M., VON LOEWENICH, F. D., GRZESZCZUK, A., TORINA, A., CARACAPPA, S., MANGOLD, A. J., NARANJO, V., STUEN, S. and KOCAN, K. M. (2005c). Sequence analysis of the *msp4* gene of *Anaplasma phagocytophilum* strains. *Journal of Clinical Microbiology.* **43:** 1309–1317.

DE LA FUENTE, J., NARANJO, V., RUIZ-FONS, F., HOFLE, U., FERNANDEZ DE MERA, I. G., VILLANUA, D., ALMAZAN, C., TORINA, A., CARACAPPA, S., KOCAN, K. M. and GORTAZAR, C. (2005d). Potential vertebrate reservoir hosts and invertebrate vectors of *Anaplasma marginale* and *A. phagocytophilum* in central Spain. *Vector-Borne Zoonosis and Diseases.* **5:** 390–401.

DE LA FUENTE, J., RUYBAL, P., MTSHALI, M. S., NARANJO, LI SHUQING, L., MANGOLD, A. J., RODRIGUEZ, S. D., JIMENEZ, R., VICENTE, J., MORETTA, R., TORINA, A., ALMAZAN, C., MBATI, P. M., FARBER, M., GORTAZAR, C. and KOCAN, K. M. (2007a). Analysis of world strains of *Anaplasma marginale* using major surface protein 1a repeat sequences. *Veterinary Microbiology.* **119:** 382–390.

DE LA FUENTE, J., ATKINSON, M. W., NARANJO, V., FERNANDEZ DE MERA, I. G., MANGOLD, A. J., KEATING, K. A. and KOCAN, K. M. (2007b). Sequence analysis of the *msp4* gene of *Anaplasma ovis* strains. *Veterinary Microbiology.***119:** 375–81.

DE LA FUENTE, J., KOCAN, K. M., BLOUIN, E. F., ZIVKOVIC, Z., NARANJO, V., ALMAZAN, C., ESTEVES, E., JONGEJAN, F., DAFFRE, S. and MANGOLD, A. J. (2010). Functional genomics and evolution of tick- *Anaplasma* interactions and vaccine development. *Veterinary Parasitology.* **167:** 175– 186.

DE WAAL, D. T. (2000). Anaplasmosis control and diagnosis in South Africa. *Annals of the New York Academy of Science.* **916:** 474-483.

DECARO, N., ELIA, G., DESARIO, C., ROPERTO, S., MARTELLA, V., CAMPOLO, M., LORUSSO, A., CAVALLI, A. and BUONAVOGLIA, C. (2006a). A minor groove binder probe real-time PCR assay for discrimination between type 2-based vaccines and field strains of canine parvovirus. *Journal of Virology Methods.* **136:** 65–70.

DECARO, N., ELIA, G., MARTELLA, V., CAMPOLO, M., DESARIO, C., CAMERO, M., CIRONE, F., LORUSSO, E., LUCENTE, M.S., NARCISI, D., SCALIA, P. and BUONAVOGLIA, C. (2006b). Characterisation of the canine parvovirus type 2 variants using minor groove binder probe technology. *Journal of Virology Methods.* **133:** 92–99.

DECARO, N., MARTELLA, V., ELIA, G., DESARIO, C., CAMPOLO, M., BUONAVOGLIA, D., BELLACICCO, A.L., TEMPESTA, M. and BUONAVOGLIA, C. (2006c). Diagnostic tools based on minor groove binder technology for rapid identification of vaccine and field strains of canine parvovirus type 2b. *Journal of Virology Methods.* **138:**10–16.

DECARO, N., CARELLI, G., LORUSSO, E., LUCENTE, M.S., GRECO G., LORUSSO, A., RADOGNA, A., CECI, L. and BUONAVOGLIA, C. (2008). Duplex real-time polymerase chain reaction for simultaneous detection and quantification of *Anaplasma marginale* and *Anaplasma centrale*. *Journal of Veterinary Diagnostic Investigation*. **20:** 606–611.

DELPORT, W., POON, A. F., FROST, S. D. and KOSAKOVSKY POND, S. L. (2010). Datamonkey 2010: suite of phylogenetic analysis tools for evolutionary biology. *Bioinformatics*. **26:** 2455–2457.

DOLD, A. P. and COCKS, M. L. (2001). Traditional veterinary medicine in the Alice district of the Eastern Cape Province, South Africa. *South African Journal of Science*. **97:** 375-379.

DREHER, U. M., DE LA FUENTE, J., HOFMANN-LEHMANN, R., MELI, M. K., PUSTERIA, N., KOCAN, K. M., WOLDEHIWET, A., REGULA, G. and STAERK, K. D. C. (2005). Serologic cross reactivity between *Anaplasma marginale* and *Anaplasma phagocytophilum*. *Clinical Vaccine Immunology*. **12:** 1177–1183.

DREYER, K., FOURIE, L. J., and KOK, D. J. (1998). Epidemiology of tick-borne diseases of cattle in Botshabelo and Thaba Nchu in the Free State Province. *Onderstepoort Journal of Veterinary Research*. **65:** 285–289.

DU PLESSIS, J. L., DE WAAL, D. T. and STOLTSZ, W. H. (1994). A survey of the incidence and importance of the tick-borne diseases heartwater, redwater and anaplasmosis in the heartwater –endemic regions of South Africa. *Onderstepoort Journal of Veterinary Research.* **61:** 295-301.

DUMLER, J., BARBET, A., BEKKER, C., DASCH, G., PALMER, G., RAY, S., RIKIHISA, Y. and RURANGIRWA, F. (2001). Reorganization of genera in the families Rickettsiaceae and Anaplasmataceae in the order Rickettsiales: unification of some species of *Ehrlichia* with *Anaplasma*, *Cowdria* with *Ehrlichia* and *Ehrlichia* with *Neorickettsia*, descriptions of six new species combinations and designation of *Ehrlichia equi* and 'HE agent' as subjective synonyms of *Ehrlichia phagocytophila*. *International Journal of Systematics and Evolution Microbiology.* **51:** 2145-2165.

DUNNING HOTOPP, J. C., LIN, M., MADUPU, R., CRABTREE, J., ANGIUOLI, S. V., EISEN, J., SESHADRI, R., REN, Q., WU, M., UTTERBACK, T. R., SMITH, S., LEWIS, M., KHOURI, H., ZHANG, C., NIU, H., LIN, Q., OHASHI, N., ZHI, N., NELSON, W., BRINKAC, L. M., DODSON, R. J., ROSOVITZ, M. J., SUNDARAM, J., DAUGHERTY, S. C., DAVIDSEN, T., DURKIN, A. S., GWINN, M., HAFT, D. H., SELENGUT, J. D., SULLIVAN, S. A., ZAFAR, N., ZHOU, L., BENAHMED, F., FORBERGER, H., HALPIN, R., MULLIGAN, S., ROBINSON, J., WHITE, O., RIKIHISA, Y. and TETTELIN, H. (2006). Comparative genomics of emerging human ehrlichiosis agents. *PLoS Genetics.* **2 (2):** e21.

ELIA, G., SAVINI, G., DECARO, N., MARTELLA, V., TEODORI, L., CASACCIA, C., DI GIALLEONARDO, L., LORUSSO, E., CAPORALE, V. AND BUONAVOGLIA, C. (2008). Use of real-time RT-PCR as a rapid molecular approach for differentiation of field and vaccine strains of bluetongue virus serotypes 2 and 9. *Molecular Cellular and Probes.* **22:** 38–46.

ELLERY, W. N., SCHOLES, R. J. and SCHOLES, M. C. (1995). The distribution of sweetveld and sourveld in South Africa's grassland biome in relation to environmental factors. *African Journal of Range and Forage Science.* **12:** 38-45.

ERIKS, I. S., STILLER, D. and PALMER, G.H. (1993). Impact of persistent *Anaplasma marginale* rickettsemia on tick infection and transmission. *Journal of Clinical Microbiology.* **31:** 2091.

ESTRADA-PEÑA, A., NARANJO, V., ACEVEDO-WHITEHOUSE, K., MANGOLD, A. J., KOCAN, K. M. and DE LA FUENTE, J. (2009). Phylogeographic analysis reveals association of tick-borne pathogen, *Anaplasma marginale*, *Msp1a* sequences with ecological traits affecting tick vector performance. *BioMed Central Biology.* **57:** 1-13.

ExPaSy Translation Tool 2011. Available from: http://expasy.hcuge.ch/ tools/dna.html.

FELSENSTEIN, J. (2004). Inferring Phylogenies. Sunderland Massachusetts: Sinauer Associates.

FOURIE, L. J. and HORAK, I. G. (1991). The seasonal activity of adult ixodid ticks on Angora goats in the south western Orange Free State. *Journal of the South African Veterinary Association.* **62:** 104–106.

FOURIE, L. J., KOK, D. J. and HEYNE, H. (1996). Adult ixodid ticks on two cattle breeds in the southwestern Free State, and their seasonal dynamics. *Onderstepoort Journal of Veterinary Research.* **63:**19–23.

FRENCH, D. F., MCELWAIN, T. F., MCGUIRE, T. C. and PALMER, G. H. (1998). Expression of *Anaplasma marginale* Major Surface Protein 2 variants during persistent cyclic rickettsemia. *Infection and Immunity*. **66:** 1200–1207.

FRENCH, D. M., BROWN, W. C. and PALMER, G. H. (1999). Emergence of *Anaplasma marginale* antigenic variants during persistent rickettsemia. *Infection and Immunity*. **67:** 5834–5840.

FUTSE, J. E., UETI, M. W., KNOWLES, D. P., JR. and PALMER, G. H. (2003). Transmission of *Anaplasma marginale* by *Boophilus microplus*: retention of vector competence in the absence of vector-pathogen interaction. *Journal of Clinical Microbiology*. **41:** 3829-3834.

GARCIA-GARCIA, J. C., DE LA FUENTE, J., BLOUIN, E. F., HALBUR, T., ONET, V. C., SALIKI, J. T. and KOCAN, K. M. (2004a). Differential expression of the *msp1a* gene of *Anaplasma marginale* occurs in bovine erythrocytes and tick cells. *Veterinary Microbiology*. **98:** 261–272.

GARCIA-GARCIA, J. C., DE LA FUENTE, J., KOCAN, K. M., BLOUIN, E. F., HALBUR, T., ONET, V. C. and SALIKI, J. T. (2004b). Mapping of B-cell epitopes in the N-terminal repeated peptides of *Anaplasma marginale* major surface protein 1a and characterization of the humoral immune response of cattle immunized with recombinant and whole organism antigens. *Veterinary Immunology and Immunopathology*. **98:** 137–151.

GARCIA-BORONAT, M., DIEZ-RIVERO, C. M., REINHERZ, E. L., and RECHE, P. A. (2008). PVS: a web server for protein sequence variability analysis tuned to facilitate conserved epitope discovery. *Nucleic Acids Research.* **36:** 35–41.

GATES, N. I. and WESCOTT, R. B. (2000). Parasites of Cattle. WSU, cru84.cahe.wsu.edu/cgibin/pubs/EB1742.html, Accessed 14 August 2013.

GE, N., KOCAN, K. M., BLOUIN, E. F. and MURPHY, G. L. (1996). Developmental studies of *Anaplasma marginale* (Rickettsiales: Anaplasmataceae) in male *Dermacentor andersoni* (Acari: Ixodidae) infected as adults by using nonradioactive in situ hybridization and microscopy. *Journal of Medical Entomology.* **33:** 911–920.

HALL, B. G. (2011). Phylogenetic Trees Made Easy: A How-To Manual. 4th Edition. Sinauer, Assoc. Sunderland, MA.

HALL, B. G. (2013). Building Phylogenetic Trees from Molecular Data with MEGA. *Molecular Biology Evolution.* **30:** 1229-1235.

HALL, B. G. and BARLOW, M. (2006). Phylogenetic Analysis as a Tool for Molecular Epidemiology of Infectious Diseases. *Annals of Epidemiology.* **16:** 157-169.

HALL, B. G., PIKIS, A. and THOMPSON, J. (2009). Evolution and biochemistry of Family 4 glycosidases: implications for assigning enzyme function in sequence annotations. *Molecular Biology Evolution.* **26:** 2487-2497.

HAMOU, S. A., RAHALI, T., SAHIBI, H., BELGHYTI, D., LOSSON, B., GOFF, W. and RHALEM, A. (2012). Molecular and serological prevalence of *Anaplasma marginale* in cattle of North Central Morocco. *Research in Veterinary Science.* **93:** 1318-1323.

HART, L. T., LARSON, A. D., DECKER, J. L., WEEKS, J. P., and CLANCY, P. L. (1981). Preparation of intact *Anaplasma marginale* devoid of host cell antigen. *Current Microbiology.* **5:** 95–100.

HOFMANN-LEHMANN, R., MELI, M. L., DREHER, U. M., GÖNCZI, E., DEPLAZES, P., BRAUN, U., ENGELS, M., SCHÜPBACH, J., JÖRGER, K., THOMA, R., GRIOT, C., STÄRK, K. D. C., WILLI, B., SCHMIDT, J., KOCAN, K. M. and LUTZ H. (2004). Concurrent infections with vector-borne pathogens associated with fatal haemolytic anemia in a cattle herd in Switzerland. *Journal of Clinical Microbiology.* **42:** 3775–3780.

HORAK, I. G., KNIGHT, M. M. and WILLIAMS, E. J. (1991). Parasites of domestic and wild animals in South Africa. XXVIII. Helminth and arthropod parasites of Angora goats and kids in Valley Bushveld. *Onderstepoort Journal of Veterinary Research.* **58:** 253–260.

HORAK, I.G. (1999). Parasites of domestic and wild animals in South Africa. XXXVII. Ixodid ticks on cattle on Kikuyu grass pastures and in Valley Bushveld in the Eastern Cape Province. *Onderstepoort Journal of Veterinary Research.* **66:** 175– 184.

HORNOK, S., FOLDVARI, G., ELEK, V., NARANJO, V., FARKAS, R., and DE LA FUENTE, J. (2008). Molecular identification of *Anaplasma marginale* and rickettsial endosymbionts in blood-sucking flies (Diptera: Tabanidae, Muscidae) and hard ticks (Acari: Ixodidae). *Veterinary Parasitology.* **154:** 354–359.

HOWDEN, K. J., and GEALE, D. W. (2010). An update on bovine anaplasmosis (*Anaplasma marginale*) in Canada. *Journal of Canada Veterinary Research.* **51:** 837–840.

INOKUMA, H., YUTAKA, T., KAMIO, T., RAOULT, D. and BROUQUI, P. (2001). Analysis of the 16S rRNA gene sequence of *Anaplasma centrale* and its phylogenetic relatedness to other Ehrlichiae. *Clinical and Diagnostic Laboratory Immunology.* **8:** 241-244.

JONSSON, N. N. and REID, S. W. J. (2000). Global climate change and vector borne diseases. Guest editorial. *Veterinary Journal.* **160:** 87–89.

JONSSON, N. N., DAVIS, R. and DE WITT, M. (2001). An estimate of the economic effects of cattle tick (*Boophilus microplus*) infestation on Queensland dairy farms. *Australian Veterinary Journal.* **79:** 826–831.

JONSSON, N. N., BOCK, R. E., JORGENSEN, W. K., MORTON, J. M. and STEAR, M. J. (2012). Is endemic stability of tick-borne disease in cattle a useful concept? *Trends in Parasitology.* **28:** 85–89.

KANO, F. S., VIDOTTO, O., PACHECO, R. C. and VIDOTTO, M. C. (2002). Antigenic characterization of *Anaplasma marginale* isolates from different regions of Brazil. *Veterinary Microbiology.* **87:** 131–138.

KAUFMAN, P. E., KOEHLER, P. G. and BUTLER, J. F. (2006). External Parasites on Beef Cattle. Entomology and Nematology Department, Florida Cooperative Extension Service, Institute of Food and Agricultural Sciences, University of Florida, http://edis.ifas.ufl.edu/IG130, Accessed 12 May 2013.

KIESER, S. T., ERIKS, I. S. and PALMER, G. H. (1990). Cyclic rickettsemia during persistent *Anaplasma marginale* infection of cattle. *Infection and Immunity.* **58:** 1117–1119.

KIMURA, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution.* **16:** 111-120.

KNOWLES, D., TORIONI DE ECHAIDE, S., PALMER, G., MCGUIRE, T., STILLER, D. and MCELWAIN, T. (1996). Antibody against an *Anaplasma marginale* MSP5 epitope common to tick and erythrocyte stages identifies persistently infected cattle. *Journal of Clinical Microbiology.* **34:** 2225–2230.

KOCAN, K. M., GOFF, W. L., STILLER, D., CLAYPOOL, P. L., EDWARDS, W., EWING, S. A., HAIR, J. A. and BARRON, S. J. (1992). Persistence of *Anaplasma marginale* (Rickettsiales: Anaplasmataceae) in male *Dermacentor andersoni* (Acari: Ixodidae) transferred successively from infected to susceptible calves. *Journal of Medical Entomology.* **29:** 657–668.

Kocan, K.M., (1999). Parasitol. Today 15 (7) poster.

KOCAN, K. M., BLOUIN, E. F. and BARBET, A. F. (2000). Anaplasmosis control: past, present and future. *Annals of the New York Academy of Science.* **916:** 501-509.

KOCAN, K. M., DE LA FUENTE, J., BLOUIN, E.F., and GARCIA-GARCIA, J. C. (2002). Adaptations of the tick-borne pathogen, *Anaplasma marginale*, for survival in cattle and ticks. *Experimental and Applied Acarology.* **28:** 9–25.

KOCAN, K. M., DE LA FUENTE, J., GUGLIELMONE, A. A. and MELENDEZ, R. D. (2003). Antigens and alternatives for control of *Anaplasma marginale* infection in cattle. *Clinical Microbiology Review.* **16:** 698–712.

KOCAN, K. M., DE LA FUENTE, J., BLOUIN, E. F. and GARCIA-GARCIA, J. C. (2004). *Anaplasma marginale* (Rickettsiales: Anaplasmataceae): recent advances in defining host pathogen adaptations of a tick-borne rickettsia. *Parasitology.* **129:** S285– 300.

KOCAN, K. M., DE LA FUENTE, J. and BLOUIN, E. F. (2008). Advances toward understanding the molecular biology of the *Anaplasma*-tick interface. *Front Biosci Journal.* **13:** 7032-7045.

KOCAN, K. M., DE LA FUENTE, J., BLOUIN, E. F., COETZEE, J. F. and EWING, S. A. (2010). The natural history of *Anaplasma marginale*. *Veterinary Parasitology*. **167:** 95-107.

KUTTLER, K.L. (1984). *Anaplasma* infections in wild and domestic ruminants: a review. *Journal of Wildlife Diseases*. **20:** 12–20.

LEW, A. E., BOCK, R. E., MINCHIN, C. M. and MASAKA, S. (2002). A *msp1α* polymerase chain reaction assay for specific detection and differentiation of *Anaplasma marginale* isolates. *Veterinary Microbiology*. **86:** 325-335.

LOHR, C. V., BRAYTON, K. A., SHKAP, V., MOLDA, T., BARBET, A. F., BROWN, W. C. and PALMER, G. H. (2002). Expression of *Anaplasma marginale* major surface protein 2 operon-associated proteins during mammalian and arthropod infection. *Infection and Immunity*. **70:** 6005-6012.

LOPEZ, J. E., SIEMS, W. F., PALMER, G. H., BRAYTON, K. A., MCGUIRE, T. C., NORIMINE, J. and BROWN, W. C. (2005). Identification of novel antigenic proteins in a complex *Anaplasma marginale* outer membrane immunogen by mass spectrometry and genomic mapping. *Infectious Immunology*. **73:** 8109–8118.

LOPEZ, J. E., BEARE, P. A., HEINZEN, R. A., NORIMINE, J., LAHMERS, K. K., PALMER, G. H. and BROWN, W. C. (2008). High-throughput identification of T-lymphocyte antigens from *Anaplasma marginale* expressed using in vitro transcription and translation. *Journal of Immunological Methods*. **332:** 129–141.

MACMILLAN, H., BRAYTON, K.A., PALMER, G.H., MCGUIRE, T.C., MUNSKE, G., SIEMS, W.F. and BROWN, W.C. (2006). Analysis of the *Anaplasma marginale* major surface protein 1 complex protein composition by tandem mass spectrometry. *Journal of Bacteriology.* **188:** 4983-4991.

MAPIYE C., CHIMONYO M., DZAMA K., RAATS J. G. and MAPEKULA M. (2009). Opportunities for improving Nguni cattle production in the smallholder farming systems of South Africa. *Livestock Science.* **124:** 196-204.

MARTINEZ, M. L., MACHADO, M. A., NASCIMENTO, C. S., SILVA, M. V. G. B., TEODORO, R. L., FURLONG, J., PRATA, M. C. A., CAMPOS, A. L., GUIMARÃES, M. F. M., AZEVEDO, A. L. S., PIRES, M. F. A. and VERNEQUE, R. S. (2006). Association of BoLA-DRB3.2 alleles with tick (*Boophilus microplus*) resistance in cattle. *Genetics and Molecular Research.* **5:** 513–524.

MARUFU, M. C. (2008). Prevalence of ticks, tick-borne diseases in cattle on communal rangelands in the highland areas of the Eastern Cape Province. *M.Sc., Thesis,* University of Fort Hare, Alice, South Africa.

MARUFU, M. C., CHIMONYO, M., DZAMA, K. and MAPIYE C. (2011). Tick prevalence in cattle raised on sweet and sour rangelands in semi-arid areas. *Tropical Animal Health and Production.* **43:** 307-313.

MARUFU, M. C. (2014). Mechanisms of resistance to *Rhipicephalus* ticks in Nguni cattle reared in the semiarid areas of South Africa. *Doctoral Thesis,* University of KwaZulu-Natal, South Africa.

MATTIOLI, R. C., PANDEY, V. S., MURRAY, M. and FITZPATRICK, J. L. (2000). Review: Immunogenetic influences on tick resistance in African cattle with particular reference to trypanotolerant N'Dama (*Bos taurus*) and trypanosusceptible Gobra zebu (*Bos indicus*) cattle. *Acta Tropica.* **75:** 263-277.

MCCORKLE-SHIRLEY, S., HART, L. T., LARSON, A. D., TODD, W. J. and MYHAND, J.D. (1985). High-yield preparation of purified *Anaplasma marginale* from infected bovine red blood cells. *American Journal of Veterinary Research.* **46:** 1745–1747.

MCGAREY, D. J. and ALLRED, D. R. (1994). Characterization of hemagglutinating components on the *Anaplasma marginale* initial body surface and identification of possible adhesins. *Infection and Immunity.* **62:** 4587–4593.

MCGAREY, D. J., BARBET, A. F., PALMER, G. H., MCGUIRE, T. C. and ALLRED, D. R. (1994). Putative adhesins of *Anaplasma marginale*: major surface polypeptides 1a and 1b. *Infection and Immunity.* **62:** 4594–4601.

MCGUIRE, T. C., PALMER, G. H., GOFF, W. L., JOHNSON, M. I. and DAVIS, W. C. (1984). Common and isolate restricted antigens of *Anaplasma marginale* detected with monoclonal antibodies. *Infection and Immunity.* **45:** 697–700.

MCGUIRE, T. C., DAVIS, W. C., BRASSFIELD, A. L., MCELWAIN, T.F. and PALMER, G. H. (1991). Identification of *Anaplasma marginale* long-term carrier cattle by detection of serum antibody to isolated MSP-3. *Journal of Clinical Microbiology.* **29:** 788–793.

MEEUS, P. F. M. and BARBET, A. F. (2001). Ingenious gene generation. *Trends Microbiology*. **9:** 353-355.

MINJAUW, B., PERRY, B. D., KRUSKA, R., PETER, T. F., CHAMBOKO, T., MAHAN, S. M., MEDLEY, G. F. and O'CALLAGHAN, C. J. (1998). Economic impact assessment of heartwater in southern Africa. In: Proceedings of the 9th International Conference of the Association of Institutions of Tropical Veterinary Medicine (AITVM), Harare, Zimbabwe.

MINJAUW, B. and MCLEOD, A. (2003). The impact of ticks and tick-borne diseases on the livelihood of small-scale and marginal livestock owners in India and eastern and southern Africa. Research report, DFID Animal Health Programme, Centre for Tropical Veterinary Medicine, University of Edinburgh, UK. *Tick-borne diseases and poverty*.

MOLAD, T., MAZUZ, M. L., FLEIDEROVITZ, L., FISH, L., SAVITSKY, I., KRIGEL, Y., LEIBOVITZ, B., MOLLOY, J., JONGEJAN, F. and SHKAP, V. (2006). Molecular and serological detection of *A. centrale*- and *A. marginale*-infected cattle grazing within an endemic area. *Veterinary Microbiology*. **113:** 55–62.

MOLAD, T., FLEIDROVICH, L., MAZUZ, M., FISH, L., LEIBOVITZ, B., KRIGEL, Y., SHKAP, V. (2009). Genetic diversity of major surface protein 1a of *Anaplasma marginale* in beef cattle. *Veterinary Microbiology*. **136:** 54-60.

MOLLOY, J. B., BOWLES, P. M., KNOWLES, D. P., MCELWAIN, T. F., BOCK, R. E., KINGSTON, T. G., BLIGHT, G. W. and DALGLIESH, R. J. (1999). Comparison of a competitive inhibition ELISA and the card agglutination test for detection of antibodies to *Anaplasma marginale* and *Anaplasma centrale* in cattle. *Australian Veterinary Journal.* **77:** 245–249.

MTSHALI, M. S., DE WAAL, D. T. and MBATI, P. A. (2004). A sero-epidemiological survey of blood parasites in cattle in the north-eastern Free State, South Africa. *Onderstepoort Journal of Veterinary Research.* **71:** 67–75.

MTSHALI, M. S., DE LA FUENTE, J., RUYBAL, P., KOCAN, K. M., VICENTE, J., MBATI, P. A., SHKAP, V., BLOUIN, E. F., MOHALE, N. E., MOLOI, T. P., SPICKETT, A. M. and LATIF, A. A. (2007). Prevalence and genetic diversity of *Anaplasma marginale* strains in cattle in South Africa. *Zoonotic Public Health.* **54:** 23-30.

MTSHALI, M. S. and MTSHALI, P. S. (2013). Molecular diagnosis and phylogenetic analysis of *Babesia bigemina* and *Babesia bovis* hemoparasites from cattle in South Africa. *BioMed Central Veterinary Research.* **9:** 154.

MUCHENJE, V., DZAMA, K., CHIMONYO, M., RAATS J. G. and STRYDOM, P. E. (2008). Tick susceptibility and its effects on growth performance and carcass characteristics of Nguni, Bonsmara and Angus steers raised on natural pasture. *Animal.* **2:** 298-304.

MUKHEBI, A. W., WATHANGA, J., PERRY, B. D., IRVIN, A. D. and MORZARIA, S. P. (1989). Financial analysis of East Coast fever control strategies on beef production under farm conditions. *Veterinary Record.* **125:** 456–459.

MUTSHEMBELE, A.M., CABEZAS-CRUZ, A., MTSHALI, M.S., THEKISOE, O.M.M., GALINDO, R.C. AND DE LA FUENTE, J. (2014). Epidemiology and evolution of the genetic variability of *Anaplasma marginale* in South Africa. *Ticks and Tick -borne diseases.* **5:** 24-631.

NARANJO, V., RUIZ-FONS, F., HOFLE, U., FERNANDEZ DE, M. I. G., VILLANUA, D., ALMAZAN, C., TORINA, A., CARACAPPA, S., KOCAN, K. M., GORTAZAR, C. and DE LA FUENTE, J. (2006). Molecular epidemiology of human and bovine anaplasmosis in southern Europe. *Annals of the New York Academy Science.* **1078:** 95–99.

NDOU, R. V., DIPHAHE, T. P., DZOMA, B. M. and MOTSEI, L. E. (2010). The seroprevalence and endemic stability of anaplasmosis in cattle around Mafikeng in the North West Province, South Africa. *Veterinary Research.* **3:** 1–3.

NEI, M. and KUMAR, S. (2000). Molecular evolution and phylogenetics. Oxford University Press, New York. Pg. 10-50.

NELSON, C. M., HERRON, M. J., FELSHEIM, R. F., SCHLOEDER, B. R., GRINDLE, S. M., CHAVEZ, A. O., KURTTI, T. J. and MUNDERLOH, U. G. (2008). Whole genome transcription profiling of *Anaplasma phagocytophilum* in human and tick host cells by tiling array analysis. *BioMed Central Genomics.* **9:** 364.

NIELSEN, R. (2002). Mapping mutations on phylogenies. Syst. Biol. 51, 729–739.

NOH, S. M., BRAYTON, K. A., KNOWLES, D. P., AGNES, J. T., DARK, M. J, BROWN, W, C., BASZLER, T. V. and PALMER, G. H. (2006). Differential expression and sequence conservation of the *Anaplasma marginale* msp2 gene superfamily outer membrane proteins. *Infection and Immunity*. **74:** 3471–3479.

NOH, S. M., BRAYTON, K. A., BROWN, W. C., NORIMINE, J., MUNSKE, G. R., DAVITT, C. M. and PALMER, G. H. (2008). Composition of the surface proteome of *Anaplasma marginale* and its role in protective immunity induced by outer membrane immunization. *Infection and Immunity*. **76:** 2219–2226.

NOH, S. M., ZHUANG, Y., FUTSE, J. E., BROWN, W. C., BRAYTON, K. A. and PALMER, G. H. (2010). The immunization-induced antibody response to the *Anaplasma marginale* Major Surface Protein 2 and its association with protective immunity. *Vaccine*. **28:** 3741–3747.

NORVAL, R. A. I., PERRY, B. D. and YOUNG, A. S. (1992). The Epidemiology of theileriosis in Africa. Academic Press, London.

NORVAL, R. A. I., SUTHERST, R. W. and KERR, J. D. (1996). Infestations of the bont tick *Amblyomma hebraeum* (Acari: Ixodidae) on different breeds of cattle in Zimbabwe. *Experimental and Applied Acarology*. **20:** 599-605.

NYANGIWE, N. and HORAK, I. G. (2007). Goats as alternative hosts of cattle ticks, *Onderstepoort Journal of Veterinary Research*. **74:**1–7.

OBERLE, S. M. and BARBET, A. F. (1993). Derivation of the complete msp4 gene sequence of *Anaplasma marginale* without cloning. *Gene*. **136:** 291-294.

OBERLE, S. M., PALMER, G. H. and BARBET, A. F. (1993). Expression and immune recognition of the conserved MSP4 outer membrane protein of *Anaplasma marginale*. *Infection and Immunity*. **61:** 5245– 5251.

OCAIDO, M., MUWAZI, R. and OPUDA-ASIBO, J. (2009). Economic impact of ticks and tick-borne diseases on cattle production systems around Lake Mburo National Park in South Western Uganda. *Tropical Animal Health and Production*. **41:** 731-739.

OIE. (2012). Bovine Anaplasmosis. OIE Manual of Diagnostic Tests and Vaccines for Terrestrial Animals, seventh ed. Office International des Epizooties, Paris (Chapter 2.4.1).

PALMER, G. H., KOCAN, K. M., BARRON, S. J., HAIR, J. A., BARBET, A. F., DAVIS, W. C. and MCGUIRE, T. C. (1985). Presence of common antigens, including major surface protein epitopes, between the cattle (intraerythrocytic) and tick stages of *Anaplasma marginale*. *Infection and Immunity*. **50:** 881–886.

PALMER, G. H., WAGHELA, S. D., BARBET, A. F., DAVIS, W. C. and MCGUIRE, T. C. (1987). Characterization of a neutralization-sensitive epitope on the Am 105 surface protein of *Anaplasma marginale*. *International Journal of Parasitology*. **17:** 1279–1285.

PALMER, G. H. (1989). *Anaplasma* vaccines, In I. G. Wright (ed.), Veterinary protozoan and hemoparasite vaccines. CRC Press, Boca Raton, Fla. p. 2-29.

PALMER, G. H., EID, G., BARBET, A. F., MCGUIRE, T. C. and MCELWAIN, T. F. (1994). The immunoprotective *Anaplasma marginale* major surface protein 2 is encoded by a polymorphic multigene family. *Infect. Immun.* **62:**3808-3816.

PALMER, G. H., RURANGIRWA, F. R., KOCAN, K. M. and BROWN, W. C. (1999). Molecular basis for vaccine development against the ehrlichial pathogen *Anaplasma marginale*. *Parasitology Today*. **15:** 253-300.

PALMER, G. H., RURANGIRWA, F. R. and MCELWAIN, T. F. (2001). Strain composition of the ehrlichia *Anaplasma marginale* within persistently infected cattle, a mammalian reservoir for tick transmission. *Journal of Clinical Microbiology*. **39:** 631–635.

PALMER, G. H., BANKHEAD, T. and LUKEHART, S. A. (2009). Nothing is permanent but change: Antigenic variation in persistent bacterial pathogens. *Cell Microbiology*. **11:** 1697–1705.

PALMER, G. H. and BRAYTON, K.A. (2013). Antigenic variation and transmission fitness as drivers of bacterial strain structure. *Cell. Microbiology*. **15:** 1969-75.

PLAYFORD, M. (2005). Review of Research Needs for Cattle Tick Control Phases I and II Meat and Livestock Australia Limited, North Sydney, 1–162.

POHL, A. E., CABEZAS-CRUZ, A., RIBEIRO, M. F. B., SILVEIRA, J. A. G., SILAGHI, C., PFISTER, K. and PASSOS, L. M. F. (2013). Detection of genetic diversity of *Anaplasma marginale* isolates in Minas Gerais, Brazil. *Brazilian Journal of Veterinary Parasitology*. **22:** 129-135.

POND, S. L. and FROST, S. D. (2005). Datamonkey: rapid detection of selective pressure on individual sites of codon alignments. *Bioinformatics*. **21:** 2531–2533.

PORTO NETO, L. R., JONSSON, N. N., D'OCCHIO, M. J. and BARENDSE, W. (2011). Molecular genetic approaches for identifying the basis of variation in resistance to tick infestation in cattle. *Veterinary Parasitology*. **180:** 165-172.

PUPKO, T., SHAMIR, I. P. R. and GRAUR, D. (2000). A fast algorithm for joint reconstruction of ancestral amino acid sequences. *Molecular Biology and Evolution*. **17:** 890–896.

REGASSA, A., PENZHORN, B. L. and BRYSON, N. R. (2003). Attainment of endemic stability to *Babesia bigemina* in cattle on a South African ranch where non-intensive tick control was applied. *Veterinary Parasitology*. **116:** 267-274.

REINBOLD, J. B., COETZEE, J. F., SIRIGIREDDY, K. R. and GANTA, R. R. (2010). Detection of *Anaplasma marginale* and *A. phagocytophilum* in bovine peripheral blood samples by duplex real-time reverse transcriptase PCR assay. *Journal of Clinical Microbiology*. **48:** 2424–2432.

RIKHOTSO, B. O., STOLTSZ, W. H., BRYSON, N. R. and SOMMERVILLE, J. E. (2005). The impact of 2 dipping systems on endemic stability to bovine babesiosis and anaplasmosis in cattle in 4 communally grazed areas in Limpopo Province, South Africa. *Journal of South African Veterinary Association*. **76:** 217–223.

RODRIGUEZ, J. L., PALMER, G. H., KNOWLES, D. P. and BRAYTON, K .A. (2005). Distinctly different *msp2* pseudogene repertoires in *Anaplasma marginale* strains that are capable of superinfection. *Gene*. **361:** 127–132.

RODRÍGUEZ, S, D., ORTIZ, M.A.G., OCAMPO, R.J., VEGA Y. and MURGUÍA, C.A. (2009). Molecular epidemiology of bovine anaplasmosis with a particular focus in Mexico. *Infection Genetics and evolution*. **9:** 1092-1101.

RUYBAL, P., MORETTA, R., PEREZ, A., PETRIGH, R., ZIMMER, P., ALCARAZ, E., ECHAIDE, I., TORIONI DE ECHAIDE, S., KOCAN, K. M., DE LA FUENTE, J. and FARBER, M. (2009). Genetic diversity of *Anaplasma marginale* in Argentina. *Veterinary Parasitology*. **162:** 176- 180.

SACKETT, D., HOLMES, P., ABBOTT, K., JEPHCOTT, S. and BARBER, M. (2006). Assessing the Economic Cost of Endemic Disease on the Profitability of Australian Beef Cattle and Sheep Producers. Meat and Livestock Australia Limited, North Sydney, 1–133.

SAITOU, N. and NEI, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*. **4**: 406–425.

SANTOS, P. S., SENA, A. A., NASCIMENTO, R., ARAÚJO, T. G., MENDES, M. M., MARTINS, J. R., MINEO, T. W., MINEO, J. R. and GOULART, L. R. (2013). Epitope-based vaccines with the *Anaplasma marginale Msp1a* functional motif induce a balanced humoral and cellular immune response in mice. *PLoS One.* 8**:** e60311.

SCHNEIDER, A., CANNAROZZI, G. M. and GONNET, G. H. (2005). Empirical codon substitution matrix. *BioMed Central Bioinformatics.* **6:** 134.

SCHNEIDER, A., GONNET, G. and CANNAROZZI, G. (2007). SynPAM – a distance measure based on synonymous codon substitutions. IEEE/ACM Trans. *Computational Biology Bioinformatics.* **4:** 553– 560.

SCOLES, G. A., BROCE, A. B., LYSYK, T. J. and PALMER, G. H. (2005). Relative efficiency of biological transmission of *Anaplasma marginale* (Rickettsiales: Anaplasmataceae) by *Dermacentor andersoni* (Acari: Ixodidae) compared with mechanical transmission by Stomoxys calcitrans (Diptera: Muscidae). *Journal of Medical Entomology.* **42:** 668– 675.

SHANNON, C.E. (1948). The mathematical theory of communication. *Bell System Technical Journal.* **27:** 379–423.

SHKAP, V., BRAYTON, K., BROWN, W. C. and PALMER, G. H. (2002). Expression of major surface protein 2 variants with conserved T-cell epitopes in *Anaplasma centrale* vaccinates. *Infection and Immunity.* **70:** 642–648.

SILVESTRE, B. T., RABELO, E. M., VERSIANI, A. F., DA FONSECA, F. G., SILVEIRA, J. A., BUENO, L. L., FUJIWARA, R. T. and RIBEIRO, M. F. (2014). Evaluation of humoral and cellular immune response of BALB/c mice immunized with a recombinant fragment of *Msp1a* from *Anaplasma marginale* using carbon nanotubes as a carrier molecule. Vaccine, http://dx.doi.org/10.1016/j.vaccine.2014.02.062 (Epub ahead of print).

STEVENS, K. B., SPICKETT, A. M., VOSLOO, W., PFEIFFER, D. U., DYASON, E. and DU PLESSIS, B. (2007). Influence of dipping practices on the seroprevalence of babesiosis and anaplasmosis in the foot-and-mouth disease buffer zone adjoining the Kruger National Park in South Africa. *Onderstepoort Journal of Veterinary Research.* **74:** 87–95.

STIK, N. I., ALLEMAN, A. R., BARBET, A. F., SORENSON, H. L., WANSLEY, H. L., GASCHEN, F. P., LUCKSCHANDER, N., WONG, S., CHU, F., FOLEY, J. E., BJOERSDORFF, A., STUEN, S. and KNOWLES, D.P. (2007). Characterization of *Anaplasma phagocytophilum* major surface protein 5 and the extent of its cross-reactivity with *A. marginale. Clinical Vaccine Immunology.* **14:** 262–268.

TAMURA, K., DUDLEY, J., NEI, M. and KUMAR, S. (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution.* **24**: 1596-1599.

TAMURA, K., PETERSON, D., PETERSON, N., STECHER, G., NEI, M. and KUMAR, S. (2011). MEGA5: Molecular Evolutionary Genetics Analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution.* **28(10):** 2731-2739.

TEBELE, N., MCGUIRE, T. C. and PALMER, G. H. (1991). Induction of protective immunity by using *Anaplasma marginale* initial body membranes. *Infection and Immunity.* **59:** 3199–3204.

THEILER, A. (1910). *Anaplasma marginale* (gen and spec., nov.). The marginal points in the blood of cattle suffering from a specific disease. Report of the Government on Veterinary Bacteriology in Transvaal, Department of Agriculture 1908-1909.

THEILER, A. (1911). Further investigations into anaplasmosis of South African cattle. In: 1st Report of the Director of Veterinary Research. Department of Agriculture of the Union of South Africa, 7–46.

THOMPSON, J. D., HIGGINS, D. G. and GIBSON, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. *Nucleic. Acids Research.* **22:** 4673–4680.

TORINA, A., MORENO-CID, J. A., BLANDA, V., FERNÁNDEZ DE MERA, I. G., DE LA LASTRA, J. M., SCIMECA, S., BLANDA, M., SCARIANO, M. E., BRIGANÒ, S., DISCLAFANI, R., PIAZZA, A., VICENTE, J., GORTÁZAR, C., CARACAPPA, S., LELLI, R.C. and DE LA FUENTE, J. (2014). Control of tick infestations and pathogen prevalence in cattle and sheep farms vaccinated with the recombinant Subolesin-Major Surface Protein 1a chimeric antigen. *Parasitology Vectors Biology.* 7: 10.

TORIONI DE ECHAIDE, S., KNOWLES, D. P., MCGUIRE, T. C., PALMER, G. H., SUAREZ, C. E. and MCELWAIN, F. F. (1998). Detection of cattle naturally infected with *Anaplasma marginale* in a region of endemicity by nested PCR and a competitive enzyme-linked immunosorbent assay using recombinant major surface protein 5. *Journal of Clinical Microbiology.* **36:** 777–782.

TURTON J. A., KATSANDE T. C., MATINGO M. B., JORGENSEN W. K., USHEWOKUNZE-OBATOLU U. and DALGLIESH R. J. (1998). Observations on the use of *Anaplasma centrale* for immunization of cattle against anaplasmosis in Zimbabwe. *Onderstepoort Journal of Veterinary* .65:81–86.

TURTON, J. (2001). External parasites of cattle. Department of Agriculture, S.A., http://www.nda.agric.za/docs/parasites/parasites.htm, 18 May 2013.

VIDOTTO, M. C., KANO, S. F., GREGORI, F., HEADLEY, S. A. and VIDOTTO, O. (2006). Phylogenetic analysis of *Anaplasma marginale* strains from Paraná State, Brazil, using the msp1α and msp4 genes. *Journal Veterinary Medicine.* **53:** 404-411.

VISESHAKUL, N., KAMPER, S., BOWIE, M. V. and BARBET, A. F. (2000). Sequence and expression analysis of a surface antigen gene family of the rickettsia *Anaplasma marginale*. *Gene*. **253:** 45-53.

VISSER, E. S., MCGUIRE, T. C., PALMER, G. H., DAVIS, W. C., SHKAP, V., PIPANO, E. and KNOWLES, D.P. (1992). The *Anaplasma marginale msp5* gene encodes a 19-kilodalton protein conserved in all recognized *Anaplasma* species. *Infection and Immunity*. **60:** 5139–5144.

WALKER, J. B. 1991. A review of the ixodid ticks (Acari, Ixodidae) occurring in southern Africa. *Onderstepoort Journal of Veterinary Research*. **58:** 81–105.

WAMBURA, P. N., GWAKISA P. S., SILAYO R. S. and RUGAIMUKAMU E. A. (1998). Breed-associated resistance to tick infestation in *Bos indicus* and their crosses with *Bos taurus*. *Veterinary Parasitology*. **77:** 63-70.

WARNER, C. K. and DAWSON, J. E. (1996). Genus- and species-level identification of *Ehrlichia* species by PCR and sequencing. In PCR protocols for emerging infectious diseases. Edited by Persing DH. Washington DC: ASM Press; 100–105.

WEN, B., JIAN, R., ZHANG, Y. and CHEN, R. (2002): Simultaneous detection of *Anaplasma marginale* and a new *Ehrlichia* species closely related to *Ehrlichia chaffeensis* by sequence analysis of 16S ribosomal DNA in Boophilus microplus ticks from Tibet. *Journal of Clinical Microbiology*. **40:** 3286–3290.

YANG, Z., KUMAR, S. and NEI, M. (1995). A new method of inference of ancestral nucleotide and amino acid sequences. *Genetics*. **14:** 1641–1650.

YU, X. J., ZHANG, X. F., MCBRIDE, J. W., ZHANG, Y. and WALKER, D. H. (2001).Phylogenetic relationships of *Anaplasma marginale* and '*Ehrlichia platys*' to other *Ehrlichia* species determined by *GroEL* aa sequences. *International Journal of System Evolution Microbiology*. **51(3):** 1143–1146.

ZHANG, Z., SCHWARTZ, S., WAGNER, L. and MILLER, W. (2000). A greedy algorithm for aligning DNA sequences. *J. Comput. Biol.* **7:** 203–214.

ZIAM, H. and BENAOUF, H. (2004). Prevalence of blood parasites in cattle from wilayates of Annaba and El Tarf east Algeria. *Arch. Inst. Pasteur Tunis*. **81:** 27–30.

ŽIVKOVIĆ, Z. (2010). Tick-pathogen interactions in bovine anaplasmosis. *Doctoral Thesis,* University of Utrecht, Netherland.

APPENDICES

Appendix1: Raw data of statistical analysis of *A. marginale* in cattle in studied provinces of South Africa

The table below gives the general outcome

```
. tab Anaplasma_results_nw

 ANAPLASMA |
 MARGINALE |
PCR RESULTS |      Freq.     Percent        Cum.
------------+-----------------------------------
   Negative |         71       25.36       25.36
   Positive |        209       74.64      100.00
------------+-----------------------------------
      Total |        280      100.00
```

The table above is presented in another format to give your 95% confidence interval for outcome.

```
. prop Anaplasma_results_nw

Proportion estimation                 Number of obs     =       280

--------------------------------------------------------------------------------
                   |       Proportion   Std. Err.     [95% Conf. Interval]
-------------------+------------------------------------------------------------
Anaplasma_results_nw |
          Negative |          .2535714   .0260461      .2057817    .3081534
          Positive |          .7464286   .0260461      .6918466    .7942183
--------------------------------------------------------------------------------
```

It shows that over the entire sample, there was 74.6% prevalence over all

The table below gives the cross-tabulation of outcome by farm types.

```
. tab Anaplasma_results_nw farm_nw, exact

 ANAPLASMA |
 MARGINALE |
    PCR |      FARM
  RESULTS | Commercia   Communal |     Total
----------+----------------------+----------
 Negative |         4         67 |        71
 Positive |        61        148 |       209
----------+----------------------+----------
    Total |        65        215 |       280

      Fisher's exact =                 0.000
  1-sided Fisher's exact =             0.000
```

The hypothesis that prevalence is independent of farm is rejected. The result above showed that prevalence is associated with the type of farm.

```
. prop Anaplasma_results_nw, over( farm_nw)

Proportion estimation        Number of obs   =   280

   Negative: Anaplasma_results_nw = Negative
   Positive: Anaplasma_results_nw = Positive

  Commercial: farm_nw = Commercial
   Communal: farm_nw = Communal

------------------------------------------------------------------------
    Over | Proportion  Std. Err.    [95% Conf. Interval]
-------------+----------------------------------------------------------
Negative     |
 Commercial |  .0615385  .0300394    .0230112  .1543787
  Communal |  .3116279  .0316609    .2529283  .3770745
-------------+----------------------------------------------------------
Positive     |
 Commercial |  .9384615  .0300394    .8456213  .9769888
  Communal |  .6883721  .0316609    .6229255  .7470717
------------------------------------------------------------------------


. kwallis Anaplasma_results_nw  ,by( location_nw)

Kruskal-Wallis equality-of-populations rank test

 +---------------------------------+
 |  location_nw | Obs | Rank Sum |
 |--------------+-----+----------|
 | Eastern Cape | 40 | 7040.00 |
 |    Gauteng | 20 | 3520.00 |
 | KwaZulu-Natal | 55 | 8700.00 |
 |    Limpopo | 40 | 5920.00 |
 |  Mpumalanga | 21 | 3696.00 |
 |--------------+-----+----------|
 |  North West | 33 | 4828.00 |
 | Northern Cape | 45 | 1620.00 |
 | Western Cape | 26 | 4016.00 |
 +---------------------------------+

chi-squared =  94.425 with 7 d.f.
probability =   0.0001

chi-squared with ties =  166.293 with 7 d.f.
probability =   0.0001
```

The hypothesis that the farms are similar with respect to prevalence is rejected by the results above

```
. tab location_nw Anaplasma_results_nw  , row col cell chi2
+-------------------+
| Key               |
|-------------------|
|   frequency       |
| row percentage    |
| column percentage |
| cell percentage   |
+-------------------+

               | ANAPLASMA MARGINALE
               |   PCR RESULTS
   LOCATION    | Negative  Positive |   Total
---------------+--------------------+----------
 Eastern Cape  |        0        40 |        40
               |     0.00    100.00 |    100.00
               |     0.00     19.14 |     14.29
               |     0.00     14.29 |     14.29 Second
---------------+--------------------+----------
    Gauteng    |        0        20 |        20
               |     0.00    100.00 |    100.00
               |     0.00      9.57 |      7.14
               |     0.00      7.14 |      7.14
---------------+--------------------+----------
KwaZulu-Natal  |        7        48 |        55
               |    12.73     87.27 |    100.00
               |     9.86     22.97 |     19.64
               |     2.50     17.14 |     19.64 Highest prevalence
---------------+--------------------+----------
    Limpopo    |        8        32 |        40
               |    20.00     80.00 |    100.00
               |    11.27     15.31 |     14.29
               |     2.86     11.43 |     14.29
---------------+--------------------+----------
  Mpumalanga   |        0        21 |        21
               |     0.00    100.00 |    100.00
               |     0.00     10.05 |      7.50
               |     0.00      7.50 |      7.50
---------------+--------------------+----------
  North West   |        7        26 |        33
               |    21.21     78.79 |    100.00
               |     9.86     12.44 |     11.79
               |     2.50      9.29 |     11.79
---------------+--------------------+----------
Northern Cape  |       45         0 |        45
               |   100.00      0.00 |    100.00
               |    63.38      0.00 |     16.07
               |    16.07      0.00 |     16.07 Lowest(Northern Cape)
---------------+--------------------+----------
```

```
 Western Cape |      4      22 |     26
              |   15.38   84.62 |  100.00
              |    5.63   10.53 |    9.29
              |    1.43    7.86 |    9.29
--------------+----------------------+----------
        Total |     71     209 |     280
              |   25.36   74.64 |  100.00
              |  100.00  100.00 |  100.00
              |   25.36   74.64 |  100.00

     Pearson chi2(7) = 166.8890  Pr = 0.000
```

Prevalence level by Farm types.

```
. prop Anaplasma_results_nw if farm_nw==1(Commercial)

Proportion estimation        Number of obs   =    65

------------------------------------------------------------------------------
                    | Proportion  Std. Err.    [95% Conf. Interval]
--------------------+---------------------------------------------------------
Anaplasma_results_nw |
            Negative |  .0615385  .0300394    .0226719   .1563737
            Positive |  .9384615  .0300394    .8436263   .9773281
------------------------------------------------------------------------------

. prop Anaplasma_results_nw if farm_nw==2 (Communal)

Proportion estimation        Number of obs   =    215

------------------------------------------------------------------------------
                    | Proportion  Std. Err.    [95% Conf. Interval]
--------------------+---------------------------------------------------------
Anaplasma_results_nw |
            Negative |  .3116279  .0316609    .2528555    .377165
            Positive |  .6883721  .0316609     .622835   .7471445
------------------------------------------------------------------------------

. tab farm_nw

    FARM |    Freq.    Percent     Cum.
-----------+-----------------------------------
 Commercial |      65      23.21     23.21
  Communal |     215      76.79    100.00
-----------+-----------------------------------
   Total |     280    100.00


. kwallis location_nw if Anaplasma_results_nw==0   , by( farm_nw)

Kruskal-Wallis equality-of-populations rank test

 +----------------------------+
 |  farm_nw | Obs | Rank Sum |
```

```
 |------------+-----+----------|
 | Commercial |   4 |   278.00 |
 |  Communal | 67 | 2278.00 |
 +----------------------------+

chi-squared =  11.167 with 1 d.f.
probability =   0.0008

chi-squared with ties =  15.049 with 1 d.f.
probability =   0.0001


. kwallis location_nw if Anaplasma_results_nw==1(Pos) , by( farm_nw)

Kruskal-Wallis equality-of-populations rank test

 +----------------------------+
 |  farm_nw | Obs | Rank Sum |
 |------------+-----+----------|
 | Commercial | 61 | 5766.50 |
 |  Communal | 148 | 16178.50 |
 +----------------------------+

chi-squared =   2.580 with 1 d.f.
probability =   0.1082

chi-squared with ties =   2.654 with 1 d.f.
probability =   0.1033

. tab Anaplasma_results_nw

 ANAPLASMA |
 MARGINALE |
PCR RESULTS |   Freq.   Percent    Cum.
------------+-----------------------------------
  Negative |     71    25.36    25.36
  Positive |    209    74.64   100.00
------------+-----------------------------------
   Total |    280   100.00


. prop location_nw if Anaplasma_results_nw==1(Pos), over( farm_nw)

Proportion estimation        Number of obs   =   209

  Commercial: farm_nw = Commercial
   Communal: farm_nw = Communal

--------------------------------------------------------------
    Over | Proportion  Std. Err.    [95% Conf. Interval]
------------+-------------------------------------------------
Eastern Cape |
 Commercial |  .3114754  .0597855   .2070474  .4393882
```

```
  Communal |   .1418919    .02878    .0940079   .2085517
-----------+------------------------------------------------
Gauteng    |
 Commercial |   .3278689   .060604    .2209735   .4561933
  Communal |      . (no observations)
-----------+------------------------------------------------
KwaZulu-Natal|
 Commercial |      . (no observations)
  Communal |   .3243243    .03861    .2532568    .404531
-----------+------------------------------------------------
Limpopo    |
 Commercial |      . (no observations)
  Communal |   .2162162   .0339534    .1567198   .2905182
-----------+------------------------------------------------
Mpumalanga |
 Commercial |      . (no observations)
  Communal |   .1418919    .02878    .0940079   .2085517
-----------+------------------------------------------------
 North West |
 Commercial |      . (no observations)
  Communal |   .1756757   .0313867    .1220444   .2462639
-----------+------------------------------------------------
Western Cape |
 Commercial |   .3606557   .0619924    .2492643    .489377
  Communal |      . (no observations)
-------------------------------------------------------------

. prop location_nw if Anaplasma_results_nw==1 & farm_nw==1 , over(
farm_nw)

Proportion estimation          Number of obs   =       61
  Commercial: farm_nw = Commercial

-------------------------------------------------------------
    Over | Proportion   Std. Err.    [95% Conf. Interval]
-----------+-------------------------------------------------
Eastern Cape |
 Commercial |   .3114754   .0597855    .2057294   .4413712
-----------+-------------------------------------------------
Gauteng    |
 Commercial |   .3278689   .060604    .2196101   .4581632
-----------+-------------------------------------------------
Western Cape |
 Commercial |   .3606557   .0619924     .247815   .4913163
-------------------------------------------------------------


. prop location_nw if Anaplasma_results_nw==1 & farm_nw==2 , over(
farm_nw)

  Communal: farm_nw = Communal
-------------------------------------------------------------
    Over | Proportion   Std. Err.    [95% Conf. Interval]
```

```
-------------+----------------------------------------------------
Eastern Cape |
  Communal |  .1418919   .02878   .0939113   .208739
-------------+----------------------------------------------------
KwaZulu-Natal|
  Communal |  .3243243   .03861    .253097  .4047346
-------------+----------------------------------------------------
Limpopo      |
  Communal |  .2162162  .0339534   .1565927  .2907164
-------------+----------------------------------------------------
Mpumalanga   |
  Communal |  .1418919   .02878   .0939113   .208739
-------------+----------------------------------------------------
 North West  |
  Communal |  .1756757  .0313867    .121933  .2464571
-------------------------------------------------------------------

. prop location_nw if Anaplasma_results_nw==0, over( farm_nw)

Proportion estimation        Number of obs   =    71

  Commercial: farm_nw = Commercial
   Communal: farm_nw = Communal

-------------------------------------------------------------------
    Over | Proportion  Std. Err.   [95% Conf. Interval]
-------------+-----------------------------------------------------
KwaZulu-Natal|
 Commercial |      . (no observations)
  Communal |  .1044776  .0376511   .0496877   .206552
-------------+-----------------------------------------------------
Limpopo      |
 Commercial |      . (no observations)
  Communal |  .119403  .0399139    .059794  .2242624
-------------+-----------------------------------------------------
 North West  |
 Commercial |      . (no observations)
  Communal |  .1044776  .0376511   .0496877   .206552
-------------+-----------------------------------------------------
Northern Cape|
 Commercial |      . (no observations)
  Communal |  .6716418  .0578057   .5480652  .7752826
-------------+-----------------------------------------------------
Western Cape |
 Commercial |      1     0        .        .
  Communal |      . (no observations)
-------------------------------------------------------------------
```

Appendix 2: GenBank accession number of *Anaplasma marginale msp1a* gene isolates and their origin in South Africa

Sample names	Origin	GenBank accession numbers
LP-37	Limpopo	(KC470153)
LP-46	Limpopo	(KC470154)
LP-7	Limpopo	(KC470155)
LP-10	Limpopo	(KC470156)
LP-50	Limpopo	(KC470157)
LP-30	Limpopo	(KC470158)
LP-34	Limpopo	(KC470159)
MP-C2	Mpumalanga	(KC470160
MP-C5	Mpumalanga	(KC470161)
NW-C2	North West	(KC470162)
NW-C4	North West	(KC470163)
NW-C5	North West	(KC470164)
NW-C1-160312	North West	(KC470165)
NW-C4-160312	North West	(KC470166)
GP-C1	Gauteng	(KC470167)
GP-C2	Gauteng	(KC470168)
GP-C5	Gauteng	(KC470169)

GP-C1112105	Gauteng	(KC470170)
GP-C4117105	Gauteng	(KC470171)
GP-C1817105	Gauteng	(KC470172)
GP-C7117105	Gauteng	(KC470173)
KZN-F	KwaZulu-Natal	(KC470174)
KZN-MN	KwaZulu-Natal	(KC470175)
KZN-K	KwaZulu-Natal	(KC470176)
KZN-D	KwaZulu-Natal	(KC470177)
KZN-Y	KwaZulu-Natal	(KC470178)
KZN-14	KwaZulu-Natal	(KC470179)
KZN-19	KwaZulu-Natal	(KC470180)
KZN-49	KwaZulu-Natal	(KC470181)
KZN-51	KwaZulu-Natal	(KC470182)
EC-22	Eastern Cape	(KC470183)
EC-23	Eastern Cape	(KC470184)
EC-24	Eastern Cape	(KC470185)
WC-4	Western Cape	(KC470186)
WC-6	Western Cape	(KC470187)
WC-7	Western Cape	(KC470188)
WC-8	Western Cape	(KC470189)
WC-10	Western Cape	(KC470190)

WC-11	Western Cape	(KC470191)
WC-12	Western Cape	(KC470192)
WC-13	Western Cape	(KC470193)
WC-14	Western Cape	(KC470194)
WC-15	Western Cape	(KC470195)
WC-16	Western Cape	(KC470196)

Appendix 3: GenBank accession number of *Anaplasma marginale msp4* gene isolates and their origin in South Africa

Samples name	Origin	GenBank accession number
LP-C1	Limpopo Province	KF758907
LP-C2	Limpopo Province	KF758909
LP-C3	Limpopo Province	KF758910
LP-C7	Limpopo Province	KF758911
LP-C9	Limpopo Province	KF758912
LP-C10	Limpopo Province	KF758913
LP-C11	Limpopo Province	KF758914
LP-C14	Limpopo Province	KF758915
LP-C19	Limpopo Province	KF758916
LP-C21	Limpopo Province	KF758917
LP-C23	Limpopo Province	KF758918
LP-C27	Limpopo Province	KF758919
LP-C29	Limpopo Province	KF758920
LP-C30	Limpopo Province	KF758921
LP-C32	Limpopo Province	KF758922
LP-C33	Limpopo Province	KF758923
LP-C34	Limpopo Province	KF758924

LP-C35	Limpopo Province	KF758925
LP-C38	Limpopo Province	KF758926
LP-C40	Limpopo Province	KF758927
LP-C41	Limpopo Province	KF758928
LP-C43	Limpopo Province	KF758929
LP-C46	Limpopo Province	KF758930
LP-C49	Limpopo Province	KF758931
LP-C50	Limpopo Province	KF758932
LP-C77	Limpopo Province	KF758933
LP-C85	Limpopo Province	KF758934
MP-C1	Mpumalanga Province	KF758860
MP-C5	Mpumalanga Province	KF758862
MP-C7	Mpumalanga Province	KF758864
MP-C8	Mpumalanga Province	KF758866
MP-C15	Mpumalanga Province	KF758867
MP-C18	Mpumalanga Province	KF758869
MP-C19	Mpumalanga Province	KF758871
MP-C20	Mpumalanga Province	KF758872
MP-C21	Mpumalanga Province	KF758874
NW-CB3	North West Province	KF758842
NW-CB4	North West Province	KF758876

NW-CB5	North West Province	KF758878
NW-CB7	North West Province	KF758880
NW-CB8	North West Province	KF758882
NW-C28	North West Province	KF758883
GP-C33	Gauteng Province	KF758847
GP-C48	Gauteng Province	KF758849
GP-C56	Gauteng Province	KF758852
KZN-C26	KwaZulu-Natal Province	KF758854
KZN-C29	KwaZulu-Natal Province	KF758856
KZN-C31	KwaZulu-Natal Province	KF758857
KZN-C37	KwaZulu-Natal Province	KF758858
KZN-C42	KwaZulu-Natal Province	KF758859
EC-C6A	Eastern Cape Province	KF758843
EC-C25A	Eastern Cape Province	KF758845
EC-C33A	Eastern Cape Province	KF758847
WC-C4	Western Cape Province	KF758885
WC-C5	Western Cape Province	KF758886
WC-C6	Western Cape Province	KF758888
WC-C7	Western Cape Province	KF758890
WC-C8	Western Cape Province	KF758892
WC-C9	Western Cape Province	KF758894

WC-C10	Western Cape Province	KF758896
WC-C12	Western Cape Province	KF758898
WC-C15	Western Cape Province	KF758900
WC-C16	Western Cape Province	KF758902
WC-C18	Western Cape Province	KF758904
WC-C20	Western Cape Province	KF758905

Appendix 4: Amplified sequences of *Anaplasma spp.* isolates , their origin and GenBank accession number

Isolates and their origin	GenBank accession number
California	AY010248.1
GZ8-China	JN572928.1
51 (G18)-Hungary	EU315783.1
New Castle	AY127077.1
Glencoe 2	AY127068.1
Stillwater 1	AY127075.1
AY851150.1	Switzerland
Oregon	AY127065.1
661Kari-Kenya	AY666004.1
USA	AF428081.1
1.6-Zimbabwe	AY666006.1
2.4-Zimbabwe	AY666007.1
5.9-Zimbabwe	AY666011.1
Italy 6	AY702917.1
Italy 47	AY702921.1
Israeli non-tailed	AY786993.1
50 (G16)-Hungary	EU315782.1
Mississippi	AY010251.1

Virginia	AY010254.1
Oklahoma City	AY127073.1
Israeli tailed	AY786994.1
Okeechobee	AY010253.1
Stillwater 2	AY127076.1
Israeli round	AY787172.1
SP7-Spain	AY456002.1
TWN1-Taiwan	EU677383.1
Mexico	AF428083.1
Mexico	AF428084.1
Mexico	AF428085.1
Brazil 4	AY283190.1
Brazil 5	AY283191.1
Puerto Rico	AY191827.1
F72-Australia	AY666002.1
G38-Australia	AY666003.1
Argentina	AF428086.1
Anaplasma phagocytophilum	AY706389
Anaplasma ovis	AF393742.1