New molecular high throughput methods for *Ehrlichia ruminantium* tick screening and characterization of strain genetic structure in Mozambique and at worldwide scale

Nídia Cangi

Thesis presented on the 30th of January 2017 to obtain the grade of Doctor of Philosophy in Life Science, speciality in Molecular biology and Genetics, from the Université des Antilles

Jury members:

Reviewer: Prof. Christine MARITZ-OLIVIER

Reviewer: Dr Eric DUCHAUD

Examiner: Dr Nicola COLLINS

Examiner: Prof. Jérôme GUERLOTTE

Guest members:

Thesis director: Prof. Olivier GROS

Thesis co-director: Prof. Luís NEVES

Thesis co-director: Dr Nathalie VACHIÉRY

Acknowledgments

I would like to express my gratitude to several people and institutions that contributed directly and indirectly to complete this thesis.

I would like to thank sincerely my supervisors Dr Nathalie Vachiéry and Prof. Luís Neves for all their support and guidance, teaching, kindness and especially patience throughout the project. I would not be able to cross the many barriers on my way without their helping hands.

I also would like to thank all members of CIRAD-Guadeloupe for receiving me, for their friendship, ideas and help in times of need, especially to Laure Bournez, Soledad Castano, Valerie Pinarello, Rosalie Aprelon, Christian Sheikboudou, Isabel Marcelino, Emmanuel Albina, as well as Adela Chavez, Jonathan Gordon and Mathilde Gondard.

To CB-UEM for contributing to my academic development and to my supportive and friendly colleagues.

To Prof. Olivier Gros and the University of Antilles for all the administrative support.

To all my family and friends, especially my mother Balbina Müller and my husband Nilton Vaz that even without understanding the science behind my work always encouraged and loved me.

I am grateful to Hermógenes Mucache, Laure Bournez and Prof. Luís Neves for all the great field trips in our beautiful Mozambique, friendship and support. As well as the Veterinary Services of Mozambique, Coutada de caça 11 and 12 in Sofala and the Veterinary staff of the KNP and SAN Parks (South Africa) for logistic support during sampling.

To God and the Universe for protecting and illuminating my way.

To me, for all my patience, persistence, sacrifice, for all the personal growth, strength to not give up and to recover from mental fatigue and tendinitis. I succeed!

Last, I would like to thank IRD-Doctorants du Sud (Institut de recherche pour le développement), Ministry of Science and Technology in Mozambique, French Embassy in Mozambique, CIRAD and CB-UEM for providing funds, without which this study would not have been possible.

Summary

Ehrlichia ruminantium is the causal agent of heartwater, a ruminant tropical fatal disease transmitted by *Amblyomma* ticks. Up to now, no effective vaccine is available due to a limited cross protection of vaccinal strains on field isolates mainly associated to a high genetic diversity of *E. ruminantium* within geographical locations. Thus, both characterization of *E. ruminantium* genetic population structure at worldwide and regional scale and estimation of *E. ruminantium* tick prevalence are important to delimitate better control strategies and improve heartwater monitoring strategies.

In **Section I**, we developed two new qPCRs, *pCS20* Sol1^{TM} and Sol1^{SG}, to screen *E. ruminantium* in *Amblyomma* ticks, which are powerful tools for: 1) heartwater epidemiological studies, 2) diagnosis in the context of heartwater clinical cases and 3) follow-up of experimental infections, both in ticks and hosts. The *pCS20* Sol1^{TM} qPCR was found as sensitive (up to 30 copies/sample) and specific as the gold standard *pCS20* nested PCR but less prone to sample contamination and less time-consuming. The whole method including the automated DNA extraction and *pCS20* Sol1^{TM} qPCR demonstrated to be sensitive, specific and reproducible. It displayed the same limit of detection of the manual DNA extraction and *pCS20* nested PCR, (60 copies/sample). Moreover, the development of a high-throughput automated DNA/RNA extraction makes the screen of any tick-borne pathogen in several tick species possible.

The development of this new method allowed processing of a high number of tick samples collected in Mozambique that were then typed by Multi Locus Sequence Typing (MLST) and included into a worldwide *E. ruminantium* strain genetic structure study (**Section II**). Our study reveals the repeated occurrence of recombination between *E. ruminantium* genotypes and its important role in *E. ruminantium* genetic diversity and evolution. Despite the unclear phylogeny and phylogeography due to recombination events, *E. ruminantium* isolates are clustered into two main groups: Group 1 (West Africa) and a Group 2 (worldwide) which is represented by West, East and South Africa, Indian Ocean and Caribbean strains. Common genotypes between West Africa and Caribbean and Southern Africa and Indian Ocean allow to identify two possible ways of *E. ruminantium* introduction in these regions, associated with cattle movement.

In **Section III**, we focused mainly on *E. ruminantium* tick prevalence and genetic diversity and structure of Mozambican isolates from *A. variegatum* and *A. hebraeum* ticks collected in cattle and wildlife. Sampling was performed in 30 localities for Mozambique and in Kruger National Park (KNP, South Africa). *E. ruminantium* tick prevalence in cattle was between 0% [0-23.2 %] and 26.7% [12-45 %], with no infected ticks in 7 localities. In wildlife, tick prevalence was 8.2 [4-14.6 %] % in the KNP and 6.2% [0.2-30.2 %] in hunting concessions of Sofala province. However, no significant difference in prevalence was found between sampling sites and tick species, as well as no linear correlation between *E. ruminantium* prevalence and tick abundance was observed. There was a high genetic diversity of *E. ruminantium*, with 39 different genotypes detected and distribution of identical genotypes in several distant localities. Most genotypes from Mozambique clustered in genetic subgroup G2C (strictly clustering Zimbabwe and Mozambican isolates) and G2E. Interestingly, genotypes from group G1 and G2D associated mainly with West Africa and Caribbean strains were in minority, probably highlighting a recent introduction.

Table of contents

I. Introduction……1

1. Heartwater……...2

1.1. Pathogen: *Ehrlichia ruminantium*……2

1.2. Life cycle……...2

1.3. The disease: Heartwater……3

1.4. Heartwater in Mozambique……...4

1.5. Affected animals……4

1.6. Geographic distribution……6

1.7. Vector species……7

2. Molecular diagnostic…….....7

2.1. DNA extraction of tick and tissue samples for screening *E. ruminantium*……7

2.2. Heartwater diagnostic and *Ehrlichia ruminantium* detection methods……..8

3. *E. ruminantium* genetic characterization……12

3.1. PCR and restriction fragment length polymorphism (RFLP)……...12

3.2. Multi-locus variable numbers of tandem repeats (MLVA)……..13

3.3. Multilocus sequence typing (MLST)……14

3.4. Importance of recombination events……15

II. Aims of the study……....17

III. Section I……...19

Efficient high throughput molecular method to detect *Ehrlichia ruminantium* in ticks (Article 1: submitted to Parasites & Vectors)

IV. Section II……63

Recombination is a major driving force of genetic diversity in theAnaplasmataceae *Ehrlichia ruminantium* (Article 2: published in Frontiers in Cellular and Infection Microbiology)

V. Section III……...64

Ehrlichia ruminantium in Mozambique: a study on prevalence in ticks and isolate genetic diversity (Draft in preparation for publication)

VI. General discussion……91

VII. Conclusions and perspectives…….....99

VIII. References……102

IX. Annexe……...117

Parapatric distribution and sexual competition between two tick species, *Amblyomma variegatum* and *A. hebraeum* (Acari, Ixodidae), in Mozambique (Published in Parasites & Vectors)

I. Introduction

1. Heartwater

1.1. Pathogen: *Ehrlichia ruminantium*

In 1925, Edmund Cowdry named the causal agent of heartwater *Rickettsia ruminantium* (Cowdry, 1925a, Cowdry, 1925b). Later, based on cytological studies, the microorganism was renamed *Cowdria ruminantium* by Moshkovski (1947). Further studies on the biology of the bacteria cultured in bovine umbilical endothelial cells showed that they have a life cycle similar to that of chlamydia species (Jongejan *et al.,* 1991b). With the advance of molecular biology, the 16S rDNA gene from *C. ruminantium* was sequenced and a close phylogenetic relation between the genera *Cowdria* and *Ehrlichia* was demonstrated (van Vliet, Jongejan & van der Zeijst, 1992).

Later, the taxonomy and classification of the order Rickettsiales was clarified with the development of molecular phylogeny, based on the 16S rRNA gene, *groESL* gene and surface protein genes (Dumler *et al.,* 2001). Distinctly, the obligatory intracellular bacteria *Ehrlichia ruminantium* belongs to the class Alphaproteobacteria, order Rickettsiales and family Anaplasmataceae. The genus *Ehrlichia* from the family Anaplasmataceae includes the Gram-negative *E.* (previously *Cowdria*) *ruminantium*, *E. canis*, *E. chaffeensis*, *E. ewingii* and *E. muris* (Dumler *et al.,* 2001).

1.2. Life cycle

The life cycle of *E. ruminantium* occurs in tick gut epithelial cells and subsequently in the cells of the salivary gland and in the reticuloendothelial cells, or endothelial cells of the vertebrate hosts (Figure 1), (Marcelino *et al.,* 2012). Two morphologically distinct forms characterize this life cycle in the vertebrate host: elementary bodies, the infectious extracellular form, and reticulate bodies, the intracellular replicative form. In the mammalian host, the organism (reticulate bodies) begins to replicate by binary fission in reticuloendothelial cells in lymph nodes. The rupture of these cells releases elementary bodies, which then infect endothelial cells (Du Plessis, 1970). After entry into the endothelial cell, possibly by phagocytosis, each organism develops within a cytoplasmatic vacuole to form a colony called “morula”, leading to the rupture of the cell. The rupture disseminates hundreds of elementary bodies into the bloodstream to continue the infection cycle (Prozesky & Du Plessis, 1987). The isolation and culture of *E. ruminantium in vitro* in bovine or caprine endothelial cells allowed for the cycle of development within host cells to be better understood. Microscopic observation of in vitro-

cultivated *E. ruminantium* demonstrated the presence of intracellular reticulate bodies 2 to 4 days after infection and bacterial development for around 4 to 6 days before cell lysis (Jongejan *et al.,* 1991b; Marcelino *et al.,* 2005).

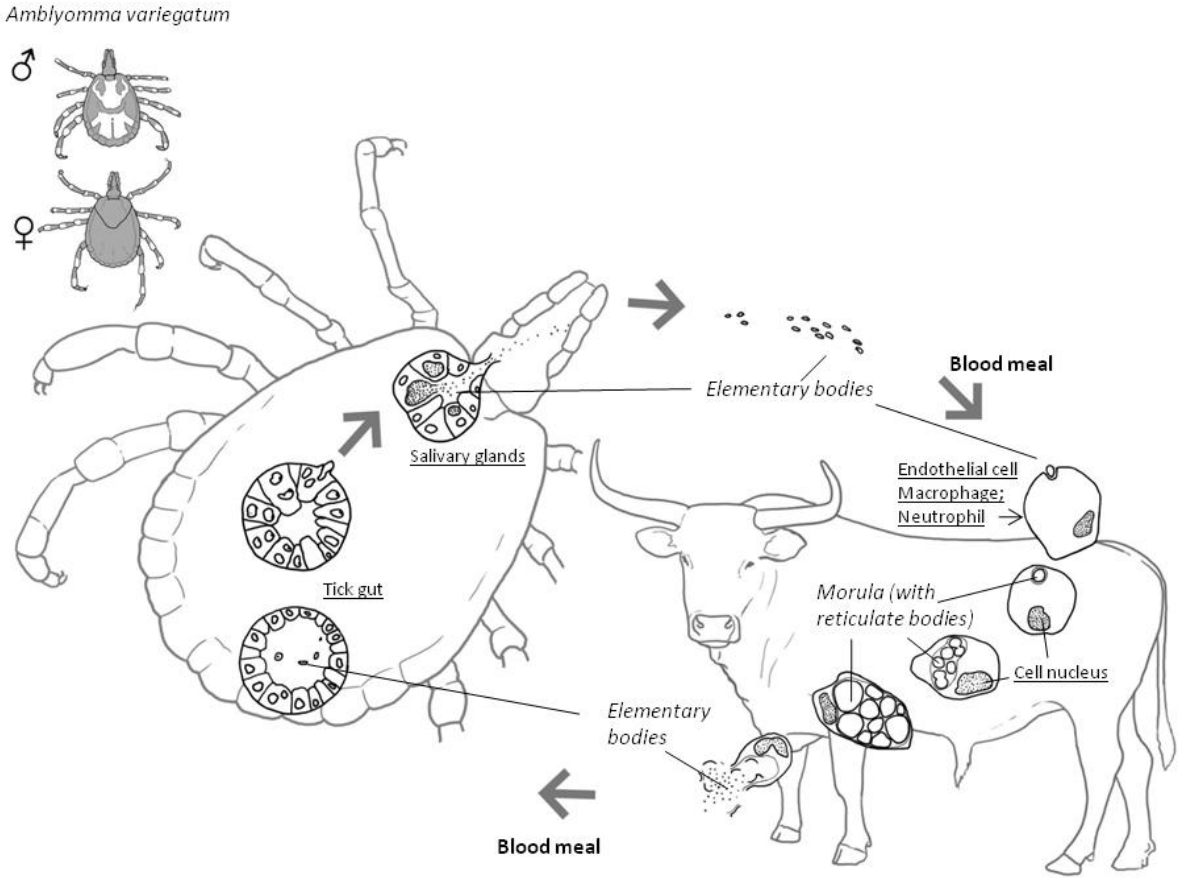


Figure 1 Life cycle of *E. ruminantium* in the tick vector (gut epithelial cells and salivary gland cells) and in the host (vascular endothelial cells, neutrophils and macrophages), (Marcelino *et al.,* 2012).

1.3. The disease: Heartwater

Heartwater or Cowdriosis is an infectious, virulent, transmissible and non-contagious disease caused by *E. ruminantium*. Depending on the susceptibility of animals, different forms of the disease varying from peracute to chronic can be found. Generally, the infection causes a high fever, nervous signs, hydropericardium and hydrothorax, and leads to death in susceptible animals. The susceptibility depends on the animal species and breed, with goats being more susceptible than sheep and cattle. Furthermore a high mortality rate is observed in exotic breeds

(up to 90%). Natural incubation can take from 10 days to 1 month, with an average of 2-3 weeks (Allsopp 2009; Martinez & Uilenberg, 2010). This tropical disease was described for the first time in South Africa on 17th February 1838, and recognized to be a tick-borne disease in 1900 (Lounsbury, 1900). It is one of the major obstacles, in some instances the most important one, to the introduction of high-producing animals (exotic breeds) into Africa with the aim of upgrading or replacing local stock (Uilenberg, 1982).

Currently, heartwater is included in the World Organization for Animal Health list of multiple species diseases, infections and infestations (OIE, 2016) and is also considered to be the 12th most important animal transboundary disease listed by the US Homeland Security department for American mainland (Roth, Richt & Morozov, 2013).

The economic impact of heartwater in the SADC region (Southern Africa Development Community) has been estimated to amount to an expenditure of US$ 44.7 million due to the loss of production and the costs of control, including antibiotic treatment and acaricide use (Minjauw 2000; Minjauw & Mcleod, 2003).

1.4. Heartwater in Mozambique

In Mozambique, heartwater was first reported in 1969, but its importance received only casual attention (Valadão, 1969). However, some serological studies showed a wide difference in prevalence between the South (63.5%) and North and Center regions (10%) of the country (Asselbergs *et al.,* 1993; Bekker *et al.,* 2001). These values indicate that this disease is present throughout the country and is transmitted, at least, by the ticks' *A. hebraeum* to the south of the Save River and by *A. variegatum* in central and northern Mozambique (Dias, 1991). Heartwater is considered to be a major cause of morbidity and mortality in ruminant production systems in the country, being associated to outbreaks with mortality rates above 80%, paritcularly when susceptible animals are introduced in endemic areas, as is the case of livestock development programs, where animals from Tete, are introduced in the provinces of Maputo and Gaza (Bekker *et al.,* 2001; Bila *et al*., 2003).

1.5. Affected animals

Heartwater is mainly a disease of ruminants, affecting all domestic Bovidae such as cattle, sheep and goats and approximately fifteen species of wild Bovidae (Allsopp, 2010).

African wild ruminants are most likely the original reservoirs of the disease (Neitz, 1967). In southern Africa, the most important wild ruminant reservoirs of the disease are probably blesbuck, black wildebeest (Neitz, 1935), African buffalo (Allsopp *et al.,* 1999) and Eland (Wesonga, Mukolwe & Grootenhuis, 2001). Peter, Burridge & Mahan (2002) described the wildlife host range and their susceptibility to natural or experimental infection (Table 1). Knowledge on the susceptibility of wild ruminants is important for the introduction of game species in areas that are endemic to heartwater. In game and farming interface areas, wild animals are also an important source of tick infection, especially if strict acaricide control is applied to domestic animals (Peter *et al.,* 1999).

Furthermore, a few cases of heartwater in South Africa have been reported in humans (Louw, Allsopp & Meyer, 2005). In all these cases, DNA extracted from tissue samples and serum were positive to *E. ruminantium* and subsequent nucleotide sequencing confirmed the diagnostic (Louw, Allsopp & Meyer, 2005). Attention should be given to heartwater as a potential emerging human disease and sensitive and specific diagnostic assays for *E. ruminantium* detection should be available in endemic areas, where people are exposed to ticks.

Table 1. Wildlife species susceptible to *E. ruminantium* (Peter, Burridge & Mahan 2002).

Species	Diagnostic basis for susceptibility
African ruminants	
African buffalo (Syncerus caffer)	
Black wildebeest (Connochaetes gnou)	
Blesbok (Damaliscus pygargus)	
Blue wildebeest (Connochaetes taurinus)	Experimental infection[1]
Eland (Taurotragus oryx)	
Giraffe (Giraffa camelopardalis)	
Greater kudu (Tragelaphus strepsiceros)	
Sable antelope (Hippotragus niger)	
Lechwe (Kobus leche kafuensis)	
Sitatunga (Tragelaphus spekii)	Natural infection[2]
Springbok (Antidorcas marsupialis)	
Steenbok (Raphicerus campestris)	
Non-African ruminants	
White-tailed deer (Odocoileus virginianus)	Experimental infection[1]
Chital (Axis axis)	Natural infection[2]
Timor deer (Cervus timorensis)	Natural infection[2]
Rodents	
Four-striped grass mouse (Rhabdomys pumilio)	Natural infection[2]
Southern multimammate mouse (Mastomys coucha)	

1.6. Geographic distribution

Heartwater has a widespread distribution, occurring in almost all of sub-Saharan Africa, except for the very dry southwest. It has also been detected in the islands around the African continent such as Madagascar, Zanzibar, Reunion, Mauritius, Grande Comoros and São Tomé and Principe (Provost & Bezuidenhout, 1987). Heartwater also occurs in the Caribbean islands of, Guadeloupe and Antigua (Barré, Garris & Camus 1995; Vachiéry *et al.,* 2008). In the Caribbean, the existence of migratory birds infested with potentially infected *Amblyomma* ticks as well as the presence of endemic *Amblyomma* species that are able to transmit the disease, represent a risk of introduction of heartwater to the American mainland (Estrada-Pena *et al.,* 2007; Barré *et al.,* 1987).

1.7. Vector species

E. ruminantium is transmitted by three-host ticks from the genus *Amblyomma*. The transmission occurs from stage to stage, with nymphs and adults being the infective stage (Bezuidenhout, 1987). Ticks become infected after 2-4 days of attachment (Camus & Barré, 1992). Currently, 13 species of ticks belonging to the genus *Amblyomma* are able to transmit the disease naturally or experimentally, including *A. americanum* (Martinez & Uilenberg, 2010). The most important vectors of *E. ruminantium* are *Amblyomma hebraeum* in southern Africa and *A. variegatum*, the widest spread vector, transmitting the disease to the rest of the African mainland, Indian Ocean islands and the Caribbean (Walker & Olwage, 1987; Stachurski *et al.,* 2013).

The effectiveness of *Amblyomma* ticks as vectors of heartwater in a region depends on their vector efficiency, distribution, activity, abundance, and adaptation to local wild or domestic carriers of *E. ruminantium* (Uilenberg, 1983). There are records of differences in vector competence between *A. hebraeum* and *A. variegatum* and the severity of heartwater in the southern African region (Norval, 1983; Karrar, 1986; Mahan *et al.,* 1995), thus affecting the epidemiology of the disease. However, there is no data concerning the association of genotypes to one of the mentioned tick species, particularly in areas where both species are present such as certain areas of Mozambique (discussed later in furthere detail).

2. Molecular diagnostic

2.1. DNA extraction of tick and tissue samples for screening *E. ruminantium*

DNA extraction is one of the first steps for sample preparation in molecular diagnostics. It is crucial to remove inhibitors from the samples and to obtain enough DNA to be tested (Radstrom *et al.,* 2004; Hajibabaei *et al.,* 2005). *E. ruminantium* DNA can be extracted from infected cell cultures, blood, ticks and organs, preserved either frozen at –20°C or in 70% ethanol, using classical methods or commercially available DNA extraction kits (Peter *et al.,* 1995; Martinez *et al.,* 2004a). A classical and very common DNA extraction method from infected blood, cell culture and ticks is based on phenol-chloroform-isoamyl-alchool (Sambrook, Fritschi & Maniatis, 1989; Mahan *et al.,* 1992; Waghela *et al.,* 1991). This method usually provides a good DNA concentration with a low cost per sample. However, DNA purity is not always satisfactory, decreasing the efficiency of PCR amplification. Moreover, it is a labour-intense method that can be hazardous if no adequate protection is used (Javadi *et al.,*

2014). Commercial kits have been developed to reduce the drawbacks of classical DNA extraction. In fact, work done by Halos *et al.* (2004) has demonstrated a more efficient technique for DNA extraction of ticks using crushing with a beads beater, proteinase K digestion followed by DNA extraction using a commercial kit (Halos *et al.,* 2004). Presently, various methods for manual extraction of tick DNA are available (Ammazzalorso *et al.,* 2015) and they result in high DNA yields. However, all of these methods have a low sample processing capacity. A few automatic DNA extraction methods have been tested and successfully optimized for arthropods such as spiders and fleas (Allender *et al.,* 2015; Vidergar, Toplak & Kuntner, 2014; Rodriguez-Perez *et al.,* 2013). Specifically for ticks, Moriarity, Loftis & Dasch (2005) developed a high throughput DNA extraction method for *Ixodes scapularis* using the Promega Wizard SV96 genomic DNA purification system. Further, Crowder *et al.* (2010) was able to automatize a Qiagen MiniElute Virus extraction kit and detect the presence of *B. burgdorferi* and Powassan virus in *I. scapularis* ticks. However, an automatic DNA extraction method has never been developed for *Amblyomma* ticks and further screening of *E. ruminantium*.

2.2. Heartwater diagnostic and *Ehrlichia ruminantium* detection methods

For Heartwater diagnostic and *E. ruminantium* detection methods, a broad range of tests are available with different uses depending on the methods: serological tests, probes, conventional and nested PCR, reverse line blotting, restriction fragment length polymorphism and qPCR. In particular, molecular diagnostics based on DNA amplification and polymerase chain reaction have revolutionized the detection of parasites because of their specificity and sensitivity (Collins, Allsopp & Allsopp, 2002).

2.2.1 Serology

The diagnostic of *E. ruminantium* from blood was first driven by serology three decades ago. The first serological test developed for detection of *E. ruminantium* antibodies was the indirect fluorescence antibody test (IFAT), where the antigen based on peritoneal macrophages of mice, infected with the Kümm strain, bound with an antibody, could be seen using fluorescence microscopy (Du Plessis & Malan, 1987). Afterwards, an indirect ELISA, a competitive ELISA (C-ELISA), (Jongejan *et al.,* 1991a) and a Western blot (Mahan *et al.,* 1993) were developed for serological diagnosis of *E.ruminantium*. These IFA and ELISA tests have limited reliability,

giving false positive and false negative reactions, with low sensitivity and specificity due to cross-reaction with other *Ehrlichia sp*. (Martinez & Uilenberg, 2010; Du Plessis *et al.,* 1993).

Currently, there are two main serological tests in use, based on the detection of antibodies against *E. ruminantium* major antigenic protein, MAP-1 and the fragment MAP-1B: cELISA (Katz *et al.,* 1997) and indirect ELISA MAP-1B (van Vliet *et al.,* 1995). The two tests are specific, yet, they can cross-react with *E. canis* and *E. chaffeensis,* which do not infect ruminants. The indirect MAP1-B ELISA is the routine test used at the OIE international reference laboratory. ELISA tests can be used only for epidemiological studies, but not for the diagnosis of clinical cases or to evaluate imported animals because the seropositivity period lasts for only a few weeks for bovines and less than 6 months for small ruminants. Moreover, in the case of clinical suspicion, the seroconversion appears two weeks after infection (Vachiéry *et al.,* 2013).

2.2.2 DNA probes

DNA probes targeting the *E. ruminantium pCS20* gene, 16S sRNA gene and *map1* gene were developed for detection of the bacterium.

The use of DNA probes improved the sensitivity of PCR-based diagnostic, allowing for the detection of mutations and tandem repeat sequences from several organisms (Stahl & Kane, 1992). The *pCS20* probe was the first to be developed in order to improve the detection of *E. ruminantium* in ticks (Waghela *et al.,* 1991). This probe has a high specificity and sensitivity, detecting strains from South and West Africa and the Caribbean and does not cross-react with the DNA of other pathogens (Mahan *et al.,* 1992; Waghela *et al.,* 1991). In further studies, the *pCS20* probe was used to detect *E. ruminantium* in experimentally infected *Amblyomma* ticks and sick animals (Yunker *et al.,* 1993; Mahan *et al.,* 1995; Mahan *et al.,* 2000). When compared with the 16S and *map1* probe, Allsopp *et al.,* (1999) was able to demonstrate the higher sensitivity of the *pCS20* over the other tested probes. However, studies revealed a low sensitivity of the test in infected animals and in ticks with low bacterial loads, in addition to the fact that the hybridization process was found to be heavily laborious. With the development of more sensitive PCRs, DNA probes were gradually replaced by these new techniques for the detection of *E. ruminantium* (Peter *et al.,* 1995; Peter *et al.,* 2000).

2.2.3. Reverse line blotting

Reverse line blotting (RLB) is a method that combines a PCR targeting the conserved genes 16S and 18S rRNA with DNA probes for simultaneous detection of multiple pathogens.

Bekker *et al.,* (2002) developed a reverse line blotting assay based on 16S rRNA gene PCR and probes for simultaneous detection and identification of *Anaplasma* and *Ehrlichia* in blood and ticks collected from ruminants. With this assay, the detection of *E. ruminantum*, *Anaplasma ovis* and other *Ehrlichia* spp. was successful, using blood samples from ruminants collected in Mozambique and *A. variegatum* experimentally infected with *E. ruminantium*. Additionally, 13 strains of *E. ruminantum* from South, East and West Africa were detected by RLB, but the test was less sensitive to carrier state animals (low pathogen load). With the advantage of detecting multiple pathogens, the newly developed RLB was further used for the screening of several *Anaplasma*, *Ehrlichia*, *Babesia* and *Theileria* species.

Interestingly, Faburay *et al.,* (2007b) compared the efficiency of RLB with a nested *pCS20* and nested *map1* PCRs for the diagnostic of *E. ruminantium* in Gambia. RLB was less sensitive in the detection of *Ehrlichia/Anaplasma* in *Amblyomma* ticks when compared with the nested PCRs. Thus, RLB seems not to be the best epidemiological tool for the study of heartwater in areas affected by the disease. Recently, Njiiri *et al.,* (2015) and Lorusso *et al.,* (2016) applied RLB for the detection of multiple tick-borne parasites in Kenya and Nigeria, respectively. In both studies, a low prevalence of *E. ruminantium* (around 1% or less) was found in ruminants. The authors explain the low prevalence with the hypothesis that genetic variability of the strains circulating could reduce the primer and probe hybridization and thus the sensitivity of the method. The main advantage of RLB is that it reveals several pathogens at once, but its lack of sensitivity reduces its usefulness in haemoparasite molecular surveys. For specific diagnosis, PCR and qPCR targeting *E. ruminantium* are considered to be the methods of choice.

2.2.4. *E. ruminantium* PCR and nested PCR

After the development of DNA probes, an urge to improve PCR for the detection of *E. ruminantium* started to rise and PCR became the most reliable method for diagnostic.

The diagnosis of *E. ruminantium* focused mostly on three genes; *map1*, 16S rRNA and *pCS20*, by conventional, nested and qPCR (Kock *et al.,* 1995; Allsopp *et al.,* 2001).

The first PCR for *E. ruminantium* was developed by Mahan *et al.,* (1992) in order to amplify specific DNA sequences and improve the hybridization with the *pCS20* DNA probe. Afterwards, Peter *et al.,* (1995) developed and evaluated a new PCR with AB128 and AB129 primers for detection of *E. ruminantium* in carrier animals, low bacteria load samples and differently preserved tick samples, and compared them with DNA probes. The new PCR limit of detection, sensitivity and specificity were superior to that of the DNA probe and thus could replace DNA hybridization methods. Peter *et al.,* (2000) tested the new *pCS20* PCR coupled with the DNA probe in *A. hebraeum* adult and nymphs from Zimbabwe. Tick prevalence and intensity of infection in each tick was successfully determined beyond the probe detection limit, with nymphs (10^7 to 10^9 organisms/ticks) being less infected than adults (10^5 to 10^6 organisms/tick). Later, Martinez *et al.,* (2004b) improved the sensitivity of the *pCS20* PCR by developing a hemi nested PCR using external primers AB128-AB130 in a first phase to amplify DNA and an internal primer pair AB128-AB129 to amplify the DNA matrix in a second phase. This new *pCS20* nested PCR detects up to 6 copies of bacteria per sample, from field ticks, blood, brain, and lungs from infected animals. Afterwards, the range of detected strains was increased with the improvement of the nested PCR by the development of degenerated primers, which allow for all possible nucleotides in specific positions to be obtained (Molia *et al.,* 2008; Adakal *et al.,* 2009; Adakal *et al.,* 2010). This PCR is currently the OIE reference test for detection of *E. ruminantium* and since its development, it has been used for diagnostic and epidemiological studies of *E. ruminantium* (Vachiéry *et al.,* 2008; Molia *et al.,* 2008; Adakal *et al.,* 2009; Adakal *et al.,* 2010; Esemu *et al.,* 2013).

In parallel to the *pCS20* PCR, *map1* nested PCR was developed for molecular detection of *E. ruminantium* with a detection limit of 60 copies per sample. However, it was demonstrated to be less sensitive due to the high polymorphic nature of the gene (Faburay *et al.,* 2007b; Kock *et al.,* 1995; Martinez *et al.,* 2004b).

2.2.5. *E. ruminantium* qPCR

Several qPCRs have been described for the quantification and detection of *E. ruminantium* using the *map1*, *pCS20* and *groEL* genes.

A sybergreen qPCR targeting the *map1* gene was developed to quantify *E. ruminantium* in a bioreactor during the vaccine production process. This PCR was tested on four strains and ensured uniformity between vaccine batches (Peixoto *et al.,* 2005). Likewise, bacterial load

and *E. ruminantium map1-1* transcripts were quantified by qPCR in tick midguts and salivary glands as well as from *E. ruminantium*-infected endothelial cell cultures (Postigo *et al.*, 2007). Only six strains were tested with qPCR targeting *map1* and *map1-1* polymorphic genes. In this context, they cannot be used for diagnostic purposes without further optimization. Later, a *pCS20* quantitative qPCR based on a TaqMan probe was developed by Steyn *et al.* (2008) to detect *E. ruminantium* in livestock blood and ticks from the field, but a limited number of strains were tested (15 strains). Recently, Sayler *et al.* (2015) successfully developed a multiplex Taqman qPCR assay targeting the *groEL* gene to distinguish Panola Mountain Ehrlichia infection from Heartwater in the US mainland.

3. *E. ruminantium* genetic characterization

The population genetic structure and diversity of *E. ruminantium* isolates encouraged by the need to develop an effective regional vaccine, was studied through the amplification of several genes and phylogenetic analyses such as restriction enzymes (PCR-RFLP) and the use of techniques such as VNTR (MLVA) and MLST. Complete genome sequencing of the strains, Gardel and Welgevonden (Collins *et al.,* 2005; Frutos *et al.,* 2006), allowed for the genomic plasticity and the implications for vaccine production to be deciphered. However, genome sequencing can be laborious and expensive as a means of accessing the diversity of all the strains circulating in an area. The current work applied the MLST technique to strain typing.

3.1. PCR and restriction fragment length polymorphism (RFLP)

Several studies have successfully used restriction enzymes for the characterization of *E. ruminantium* isolates. Firstly, Martinez *et al.,* (2004a) validated a *map-1* PCR and RFLP for typing of *E. ruminantium* strains. *Map1* RFLP genotypes were found to be similar to the *map1* genotypes sequenced. The study found a wide genetic diversity among 12 strains isolated in Burkina Faso and subsequent studies associated this diversity with a lack of protective immunity confirmed in cross-protection studies. Faburay *et al.* (2008) also characterized *E. ruminantium map1* genotypes from blood and ticks collected in Gambia by developing a new map 1 nested PCR and RFLP. Multiple genotypes and mixed infections of *E. ruminantium* could be detected and characterized in Gambia.

Vachiéry *et al.* (2008) were able to characterize *E. ruminantium* strains by sequencing or RFLP profiles of *map-1* PCR products. Nine distinct *map-1* genotypes from Africa and Caribbean were identified, revealing a high genetic diversity. This work supported the introduction of *E. ruminantium* in the Caribbean from the African mainland and highlighted the challenge to control heartwater trough vaccination in Guadeloupe (Vachiéry *et al.,* 2008). Afterwards, Adakal *et al.,* (2010) tested the appropriate strain to use for vaccination trials in Burkina Faso. *Map1* genotyping was used to characterize the strains tested for immunization, showing that genotype distribution is affected by time and area of study, thus affecting vaccine efficiency. Recently, Teshale *et al.* (2015) developed a method based on 16S rDNA PCR and the digestion of restriction enzymes for simultaneous detection of *Ehrlichia* and *Anaplasma* species from ticks in Ethiopia. *E. ruminantium* was successfully amplified from ticks, but fragment restriction was only achieved with MspI enzyme. Additionally, the new 16S rDNA can be used for the diagnostic of *Ehrlichia* and *Anaplasma* pathogens as well as that of other infections from ticks.

3.2. Multi-locus variable numbers of tandem repeats (MLVA)

Genome sequencing of the genus *Ehrlichia* allowed for the discovery of extensive tandem repeats, associated with expansion or contraction of intergenic regions. In fact, 8.5 % of *E. ruminantium*'s genome is composed of VNTRs (Collins *et al.,* 2005, Frutos *et al.,* 2007).

Multi-locus variable numbers of tandem repeats based on mini-satellites were developed for *E. ruminantium* by Pilet *et al.,* (2012). Thirteen reference strains from West Africa, South Africa and the Caribbean were successfully typed; no differences between the VNTR profile of virulent and attenuated strains or between strains isolated in the same area for almost 20 years were demonstrated. Similarly, Nakao *et al.* (2012) developed an MLVA scheme based on 17 *E. ruminantium* reference strains and was able to differentiate VNTR profiles between strains. Three main clusters were revealed through clustering analysis. Contrary to the findings of a study by Pilet *et al.* (2012), clustering analysis and principal component analysis revealed no association between VNTR profile and geographic origin for most of the strains, except for strains originating from West Africa.

The limited number of strains used in both studies points to the need to increase sampling intensity in order to better understand the bacterial population structure in the regions of

E.ruminantium distribution and turn this technique into a simple and routine method for population genetic structure analysis.

3.3. Multilocus sequence typing (MLST)

Multilocus sequence typing (MLST) is a technique that was developed to characterize bacteria isolates based on the sequencing of several housekeeping genes. These genes are distributed throughout the chromosome and it is unlikely that they are all affected by a single recombination event (Maiden *et al.,* 1998). Usually, a panel of 5-7 housekeeping genes with approximately 450 bp is selected and recombination and point mutation is analyzed during the initial stages of the bacterial clone's diversification (Feil *et al.,* 1999). The advantages of this technique are based on its high unambiguity, portability, reproducibility, good discriminatory power to differentiate isolates and its ability to be automated (Sullivan, Diggle & Clarke, 2005; Lindstedt, 2005). Typing of bacteria using MLST has significantly improved the understanding of the molecular epidemiology and population genetics of several species and strains (Feil & Spratt, 2001; Urwin & Maiden, 2003).

Recently, an MLST analysis was developed for *E. ruminantium* (Adakal *et al.,* 2009; Adakal *et al.,* 2010; Nakao *et al.,* 2011). However, the panel of strains should be increased to allow for typing across several geographical regions, supporting population genetic studies.

Firstly, Adakal *et al.,* (2009) developed an MLST scheme using a panel of 8 genes, *gltA*, *groEL*, *lepA*, *lipA*, *lipB*, *secY*, *sodB* and *sucA* and 3 other genes such as *map1*, *clpB* and CDS 8580 to characterize isolates from Burkina Faso. The study found a different tree topology for each locus, with the exceptions of *gltA*, *lipA* and *sodB,* as well as different topologies for concatenated sequences. None of the tests applied for neutrality (Tajima's D, Fu and Li's F and D) had significant results and thus, accept the hypothesis of neutral evolution as well as linkage disequilibrium. The authors argue that although recombination was described in *E. ruminantium,* it has never been studied on a relevant set of isolates. Thus, evolution of *E. ruminantium* is complex and contradictory because of genomic stasis and recombination present in isolates circulating in Burkina Faso. However, the MLST scheme developed can discriminate both events (Adakal *et al.,* 2009).

Secondly, Adakal *et al.,* (2010) genotyped *E. ruminantium* isolates circulating in Burkina Faso, using the previously developed MLST scheme and followed up on the evolution of these

populations over two years. The MLST dendogram of 37 field genotypes that were identified demonstrated two populations (groups). All tree topologies were congruent and discriminated the same 2 populations except for gro*EL* and *lepA* genes. Genetic diversity was found to be low in both populations and varied depending on the loci used. None of the neutrality tests rejected the hypothesis of neutral evolution. Population 1 corresponds to strains in stasis and contradicts the existence of a segregation between Southern and Western Africa (ERGA, reference strain from Guadeloupe, and ERWO, reference strain from South Africa, are included), while Population 2 is expanding rapidly, following clonal emergence. The authors argue for the existence of a homogenous population throughout Africa and clonal expansion of *E. ruminantium*. In Burkina Faso, strains are expanding rapidly following clonal emergence, but more studies need to be carried out in order to understand if this is a local, geographically limited phenomenon or a general trend in *E. ruminantium* (Adakal *et al.,* 2010).

Lastly, Nakao *et al.* (2011) typed 17 *E. ruminantium* reference strains from different geographical origins and 8 strains from Uganda (from *A. variegatum* ticks) based on a panel of 8 MLST genes (*gltA, groEL, lepA, lipA, lipB, secY, sodB and sucA*). In terms of discriminatory power, *lipA* and *SodB* had the lowest diversity index. Minimum-spanning tree revealed the presence of 3 groups: 1) Southern and Eastern Africa 2) Western Africa and Caribbean 3) Western and Eastern Africa. There was no association between groups and geographical origin except for 4 genotypes from Western Africa. The authors argue that MLST could be unsuitable to trace geographical origin due to recombination. When a neighbor net was applied to examine the impact of recombination, individual loci analysis did not detect recombination. However, for concatenated sequences, there is evidence of genetic divergence and it probably reflects the effect of recombination. Recombination tests on concatenated sequences demonstrated main recombination events. Nakao *et al.* (2011) hypothesize the existence of a homogenous population of an ancestral genotype throughout the African continent. All the predicted recombinants originated from Western Africa, but more MLST data, including strains from East and Southern Africa, are needed to understand the role of recombination in the genetic diversity of *E. ruminantium*.

3.4. Importance of recombination events

Inference of bacteria population genetics and evolution can be problematic due to recombination. This natural DNA transfer occurs through three well known processes: 1)

Transformation - competent bacteria take up free DNA in the environment and integrate it into their chromosomes, 2) Conjugation - transfer of DNA by plasmids through physical contact and 3) Transduction - transfer of DNA by bacteria and viruses followed by recombination (Redfield, 2001).

Several studies demonstrated that recombination can result in an overestimation of population expansion, leading to a false detection of positive selection (Schierup & Hein, 2000; Shriner *et al.,* 2003) and can lead to false phylogeny reconstruction (Posada & Crandall, 2002; Ruths & Nakhleh, 2005). Detection of recombination in housekeeping genes can be deciphered through a lack of congruence in gene trees (Feil *et al.,* 1996; Zhou, Bowler & Spratt, 1997), mosaic-like DNA sequences (Spratt *et al.,* 1995), excess of homoplasy in maximum-parsimony trees (Smith & Smith, 1998), and a network relationship between sequences, using split decomposition (Holmes, Urwin & Maiden, 1999).

Recombination has also been described for *E. ruminantium* in studies done by Bekker *et al.,* (2005), Hughes & French (2007) and Allsopp & Allsopp (2007). Recombination could explain part of *E. ruminantium* diversity and these data can possibly have an impact on MLST results, and more generally, on phylogenetic analysis. Currently, very few *E. ruminantium* strains have been tested for recombination. Bekker *et al.,* (2005) found recombination between two *map1* paralogs in *E. ruminantium* Gardel, while studying the transcriptomics of *map1* paralogs in one vector and several non-vector tick cell lines. Secondly, Hughes & French (2007), using a statistical and evolution approach, were able to find evidence of homologous recombination in *map1* alleles. The previous works showed the importance of the *map-1* gene family as a powerful discriminatory tool, but also demonstrated the uselessness of the *map1* gene in providing information about relatedness among genomes and about the relation with geographical strain origins.

Later, Allsopp & Allsopp (2007) genotyped a panel of eight core function genes (16S rRNA, *gltA*, groEL, *ftsZ, sodB, nuoB, rnc* and *ctaG*; shared between members of a species by horizontal gene transfer) from 12 different cultured stocks, originally isolated in different areas of Africa and the Caribbean, in order to gain information on the diversity of strains circulating in the field (Allsopp & Allsopp 2007). Phylogenetic analysis demonstrated that extensive inter-genome recombination occurs among *E. ruminantium* genotypes. The study hypothesizes that the bacteria originated in the southern or eastern regions of the African continent because of the higher genetic variability found in these isolates.

II. Aims of the study

In southern Africa, the ticks, *A. hebraeum* and *A. variegatum* are the main vectors of heartwater. Heartwater, together with East Coast fever and trypanosomosis, is regarded as one of the most important vector-borne diseases of ruminants in Africa (Uilenberg, 1983; Provost & Bezuidenhout, 1987). Further, the potential introduction of the disease to the American mainland is considered an important risk by the US homeland security department (Roth, Richt & Morozov, 2013). Heartwater control methods are limited and include the use of antibiotics or acaricides. Additionally, to date, only one vaccine is commercially available and it is only used in South Africa. Moreover, some experimental vaccines have been developed, but lack efficiency, mainly due to the genetic diversity of isolates.

In this context, it is essential to improve the molecular diagnostic and epidemiological methods for further studies of *E. ruminantium* population structure. A worldwide phylogenetic study on *E. ruminantium,* including a high number of isolates has never been conducted. Before the current study, *E. ruminantium* tick prevalence was not known in Mozambique and neither was the influence of *A. variegatum* and *A. hebraeum* tick species and wild animals on the genetic structure of isolates.

In the present work, we developed new high throughput molecular methods for *E. ruminantium* screening in ticks. The new methods consist of a new pCS20 Sol1^{TM} qPCR, an automatic DNA/RNA extraction of ticks and a 16S^{SG} rDNA qPCR for DNA quality control. These methods allowed us to screen *E. ruminantium* in a high number of ticks collected in Mozambique, where both *A. variegatum* and *hebreaum* are distributed parapatrically. We evaluated *E. ruminantium* tick prevalence in these localities as well as determining genetic diversity by MLST. More widely, we characterized genetic diversity at a global scale using isolates from West, East and South Africa, Indian Ocean and the Caribbean. Simultaneously, we identified recombination as a major driving force of *E. ruminantium* genetic diversity.

In summary, the aims of the study were to:

I. Develop high throughput molecular methods to screen *E. ruminantium* in ticks for epidemiological studies
II. Characterize the genetic diversity of *E. ruminantium* isolates at a global scale, using MLST
III. Determine *E. ruminantium* tick prevalence and isolate genetic diversity in *A. hebraeum* and *A. variegatum* ticks from Mozambique

III. **Section I**

Efficient high throughput molecular method to detect *Ehrlichia ruminantium* in ticks

(Article 1: submitted to Parasites & Vectors)

Efficient high-throughput molecular method to detect *Ehrlichia ruminantium* in ticks

Nídia Cangi[a,b,c], Valérie Pinarello[a,d], Laure Bournez[a,d*], Thierry Lefrançois[a,d], Emmanuel Albina[a,d], Luís Neves[b,e], Nathalie Vachiery[a,d]#

CIRAD, UMR CMAEE, F-97170 Petit-Bourg, Guadeloupe, France[a]; Centro de Biotecnologia-UEM, Universidade Eduardo Mondlane, Maputo, Mozambique[b]; Université des Antilles, Guadeloupe, France[c]; INRA, UMR CMAEE, F-34398, Montpellier, France[d]; Department of Veterinary Tropical Diseases, University of Pretoria, Faculty of Veterinary Science, Onderstepoort, South Africa[e].

Running Head: A new high-throughput analysis to screen *E. ruminantium*

#Address correspondence to Nathalie Vachiery: nathalie.vachiery@cirad.fr.

UMR INRA-CIRAD « Contrôle des Maladies Animales Exotiques et Emergentes »

Domaine de Duclos, Prise d'Eau, 97170 Petit Bourg, Guadeloupe.

Key words: *Ehrlichia ruminantium,* ticks, *pCS20,* automatic DNA extraction, Real-time PCR

ABSTRACT

Background

Ehrlichia ruminantium is the causal agent of heartwater, a ruminant tropical fatal disease transmitted by *Amblyomma* ticks. It is present in Sub-Saharan Africa, Indian Ocean Islands and Caribbean where it has an important economic impact and it also represents a treat for American mainland. Several molecular tools are currently available but with limited high-through put capacities especially for tick screening.

Methods

In order to improve sample screening capacity of *Ehrlichia ruminantium* in ticks and *E. ruminantium* molecular diagnostic, an automatic DNA extraction method for *Amblyomma* ticks and a new qPCR targeting *E. ruminantium pCS20* region (*pCS20* Sol1 qPCR) were developed. A comparison between the new *pCS20* Sol1 qPCR, a previously published *pCS20* Cow^{TM} qPCR and the gold standard *pCS20* nested PCR was carried out.

Results

pCS20 Sol1^{TM} qPCR was found as sensitive (up to 30 copies/sample) and specific as the gold standard *pCS20* nested PCR but less prone to sample contamination and less time-consuming. In parallel, a tick 16S^{SG} rDNA qPCR was developed for DNA extraction control, showing a good reproducibility of the automatic DNA extraction with a mean Ct value of 23+/-3 (n=37 ticks). The whole method including the automatic DNA extraction and *pCS20* Sol1 qPCR demonstrated to be sensitive, specific and reproducible. It displayed the same limit of detection of the manual DNA extraction and *pCS20* nested PCR, with 60 copies/sample. Even at 6 copies/sample, there was a detection signal for both methods but with high Ct=36.8+/-1.3 for automatic extraction and Sol1 qPCR.

Conclusions

The development of a new automatic DNA extraction using DNA/RNA viral extraction kit and qPCR allows for an accurate *E. ruminantium* epidemiological studies,

improvement of diagnostic capabilities and heartwater surveillance. In addition, the validation of a high throughput DNA/RNA viral extraction kit for ticks open new opportunities for large screening of other bacteria and viruses in ticks. This method will also contribute for tick genetic characterization and co-evolution studies as DNA ticks is mainly present in extracted samples.

Key words: *Ehrlichia ruminantium*, molecular high through put method, *pCS20,* qPCR

Background

Ehrlichia ruminantium is an obligate intracellular bacterium known for causing the infectious, virulent, transmissible and non-contagious disease, heartwater (also known as cowdriosis), in ruminants [1]. Its most important vectors are the ticks *Amblyomma hebraeum* in Southern Africa and *A. variegatum,* which is the most widespread vector, transmitting the disease to the rest of sub-Saharan Africa, Indian Ocean Islands and the Caribbean [2, 3]. This rickettsial infection is one of the major obstacles to the introduction of high-producing animals, aimed at upgrading and replacing local stock in Africa [4]. Thus, it has an important economic impact which was previously estimated for the SADC region (Southern Africa Development Community), to US$ 44.7 million [5]. Additionally, heartwater belongs to the 12th most important animal transboundary diseases listed by the US Homeland Security Department for American mainland [6]. In order to control this disease, the use of effective vaccines would be a desirable solution [1]. However, experimental vaccines such as recombinant vaccine, attenuated vaccine and inactivated vaccine have not been particularly successful so far, owing to the assumed antigenic variability of the pathogen [7, 8]. Characterization of field strains is then essential to better design appropriate vaccines including regional strains.

In order to evaluate accurately *E. ruminantium* prevalence in ticks, and further characterize the genetic diversity and population structure of *E. ruminantium* from several geographical areas by MLST or other typing methods, a large number of ticks

need to be collected (~500-2000 samples) and tested [9, 10]. Moreover, control of the absence/presence of *E. ruminantium* in ticks and in ruminants from heartwater-free area with high risk of introduction such as American mainland using rapid high through-put molecular tools will be useful for surveillance programs. Various methods for manual DNA extraction of ticks are currently available and result in high DNA yields [11, 12]. However, all of these methods have a low sample processing capacity and are time consuming. A few automated DNA extraction methods have been tested for arthropods such as spiders and flies [13-15]. Specifically for ticks, Moriarity *et al.* (2005) developed a high throughput DNA extraction method for *Ixodes scapularis* and optimized a qPCR for detection of *Rickettsia rickettsii*, *R. sibirica*, *R. africae* and *R. prowazekii* using the Promega Wizard SV96 genomic DNA purification system [16]. Further, Crowder *et al.* (2010) was able to automatize a Qiagen MiniElute Virus extraction kit and to detect the presence of *Borrelia burgdorferi* and Powassan virus in *I. scapularis* ticks [17].

In order to diagnose heartwater in ruminants and screen *E. ruminantium* in ticks, several molecular methods have been developed targeting *pCS20* gene, a highly conserved and specific gene of *E. ruminantium*. The *pCS20* nested PCR is the most used and reliable test and thus useful for epidemiological studies of *E. ruminantium* [18-20]. Additionally, this nested PCR is the OIE recommended assay and has been tested on a wide range of *E. ruminantium* strains isolated from suspicious blood and organ samples or *Amblyomma* ticks.

However, the nested *pCS20* PCR is time consuming and has a high risk of contamination due to the nested PCR procedure. To solve these issues and also address the need for quantitative results, real-time qPCRs have been developed in the past for the detection and quantification of *E. ruminantium*. A SYBR Green qPCR targeting the *map-1* gene was developed to quantify *E. ruminantium* during vaccine production [21]. Likewise, *E. ruminantium map1-1 transcripts* were quantified in tick midguts and salivary glands as well as in *E. ruminantium*-infected endothelial cell cultures by SYBR Green RT-qPCR targeting the *map-1-1* gene [22]. Later, a *pCS20* quantitative real-time PCR based on a TaqMan probe, Cow™, was developed by Steyn *et al.* (2008) to detect

E. ruminantium in livestock blood and ticks from the field [23]. Apart from this, Sayler *et al.* (2015) developed and validated recently a dual-plex Taqman qPCR assay targeting the *groEL* gene of Panola Mountain *Ehrlichia* and *E. ruminantium* in host blood [24].

The aim of this study was to further improve sample processing and screening capacity of *E. ruminantium* in ticks by developing an automated DNA extraction method using a commercial kit suitable for nucleic acid extraction of bacteria and viruses, and a new qPCR with improved sensitivity targeting the reliable *pCS20* region. For this purpose, a high throughput DNA extraction method based on the use of a 96-well plate format "Viral RNA and DNA" extraction kit from "Macherey-Nagel" and a new *pCS20* Sol1 qPCR assay were optimized. A comparison between the new *pCS20* Sol1 qPCR (both, SYBR Green, Sol1^{SG} qPCR and TaqMan, Sol1^{TM} qPCR), the previously published Cow^{TM} qPCR and the gold standard *pCS20* nested PCR were carried out to evaluate the performance of the new qPCRs. In parallel, a tick 16S rDNA real-time PCR was developed for DNA extraction quality control. The whole method including automatic DNA extraction and Sol1^{TM} qPCR was then compared to reference methods (manual DNA extraction and nested PCR).

METHODS

The study was designed in three successive steps. The first objective was to set up a new qPCR for the detection of *E. ruminantium* since the test described previously by Steyn *et al.* (2008) was not found sensitive and reproducible enough in our hands [23]. A second step was to develop a qPCR targeting the tick 16S rDNA, to be run in parallel as a mean to detect nucleic acid extraction problems and potential PCR inhibitors. Finally, to increase the throughput of tick sample preparation and testing, we designed a pipeline based on tissue lyses in a 2x24 tubes format followed by nucleic acids extraction in a 96-well plate format on an automatic platform and finally, the two qPCR targeting *E. ruminantium pCS20* gene.

Development of *pCS20* Sol1^{TM} and Sol1^{SG} qPCRs

Design of pCS20 Sol1 primers and probes. For the design of Sol1 primers and probes, the most conserved region of *E. ruminantium pCS20* gene was identified through multiple alignments of nucleotide sequences from 12 strains available in GenBank: Kwanyanga (AY236063), Mara87/7 (AY236064), Sankat430 (AY236065), Senegal (AY236066), Pokoase (AY236067), Mali (AY236068), Kumm1 (AY236069), Welgevonden (AY236058), Ball3 (AY236059), Vosloo (AY236060), Gardel (AY236061) and Blaauwkrantz (AY236062). New primers Sol1F (5'-ACAAATCTGGYCCAGATCAC-3') and Sol1R (5'-CAGCTTTCTGTTCAGCTAGT-3') and Sol1 TaqMan probe for *pCS20* Sol1 qPCR were designed targeting this conserved region using the software LightCycler® Probe Design (Table 1).

Real-time PCR setup. For the optimization of the *pCS20* Sol1 qPCR conditions, appropriate *E. ruminantium* DNA dilutions from the strain Gardel passage 48 grown in bovine aorta endothelial cell culture as previously described [25] were extracted using the QiaAmp DNA minikit (Qiagen, Courtaboeuf, France) according to the manufacturer's instructions and following the protocol of [26]. *E. ruminantium* DNA was quantified using a *map-1* TaqMan qPCR [27] and using a NanoDrop 2000c spectrophotometer (Thermo Scientific, France). DNA was serially diluted, ranging from 3.10^6 to 30 copies of elementary body/sample for further *pCS20* Sol1 qPCRs. Primers and *E. ruminantium* target probe were tested at annealing temperatures ranging from 48ºC to 56ºC to identify the appropriate qPCR conditions.

The new *pCS20* Sol1 qPCR was developed with two chemistry types, one based on the DNA intercalating dye SYBR Green (^{SG}) and the second using the TaqMan (^{TM}) technology. Both qPCR master mix contain internal passive reference dye ROX™ and Uracil-N-Glycosylase which becomes active at 50°C and inactive at 95°C.

SYBR Green (^{SG}) qPCR. The *pCS20* Sol1^{SG} qPCR assay using SYBR Green was performed using the Power SYBR® Green PCR Master Mix (Life Technologies, France). Each 25 µl reaction mixture contained 250 nM of each forward and reverse primer and standard concentrations of SYBR Green Dye, ROX™, AmpliTaq Gold® DNA Polymerase LD, dNTPs with dUTP/dTTP blend, optimized buffer components, 12.5 µl of distilled water and 2.0 µl sample DNA. The thermocycling conditions were 2' at 50°C for activation of Uracil-N-Glycosylase, 10' at 95°C for inactivation of Uracil-N-Glycosylase and activation of the AmpliTaq Gold® DNA polymerase and 40 cycles of 15 s at 95°C for denaturation and 60' at 51ºC for annealing and extension. Following amplification, specificity of the PCR products was confirmed by comparison of melting curves obtained after 10' at 95°C and 1' at 60ºC.

TaqMan (™) qPCR. The *pCS20* Sol1^{TM} qPCR assay using the TaqMan probe was performed using the TaqMan® Universal PCR Master Mix (Life Technologies, France). The final reaction contained 250 nM of each forward and reverse primer, 200 nM of the probe, ROX™, AmpliTaq Gold® DNA Polymerase LD, dNTPs with dUTP/dTTP blend, optimized buffer components, 12.5 µl of distilled water and 2.0 µl sample DNA. The qPCR was run in a final volume of 25 µl of which 2 µl corresponded to the sample DNA. The thermocycling conditions were 2' at 50°Cfor, 10' at 95°C denaturation and 40 cycles of 15' at 95°C for denaturation and 60' at 55ºC for annealing and extension.

To assess the performance of the new *pCS20* Sol1 qPCRs, a third test was used as comparison: the *pCS20* Cow^{TM} qPCR performed as previously described by Steyn *et al.* (2008), [23]. The optimal running temperature for this test is 48ºC, but owing to reduced sensitivity and reproducibility in our laboratory, we also tested 56ºC, close to the theoretical annealing temperature of its probe (58ºC).

In all the runs, positive and negative standard controls consisting of *E. ruminantium* Gardel strain and water were included. Real-time PCRs were performed on 7000 and

196 7500 System thermocyclers (Applied Biosystems) and results were analyzed by 7500
197 System SDS Software (Applied Biosystems).

198

199 ***Efficiency, limit of detection and reproducibility***. In order to evaluate the efficiency of
200 the new *pCS20* Sol1^{TM} and *pCS20* Sol1^{SG} qPCRs, 10-fold serial dilutions of *E.*
201 *ruminantium* Gardel DNA passage 48 ranging from 3.10^6 to 30 copies/sample were
202 tested in triplicates. A concentration of 3 copies/sample was additionally tested in
203 triplicates for Sol1qPCR^{SG} and once for Sol1qPCR^{TM} to complete the determination of
204 the limit of detection. The amplification efficiency (E) of the reaction was calculated
205 using the formula: $E = 10^{(1/s)}$, where "s" is the slope of the linear regression line, being
206 the Ct on the x-axis and Delta Rn on the y-axis. The percentage of efficiency was
207 calculated using the formula: % efficiency = (E - 1) x 100% [28]. Standard deviations
208 (SD) of cycle threshold values were also calculated. The detection limit of the two new
209 *pCS20* qPCRs using SYBR Green and TaqMan chemistries was determined in
210 comparison with the conventional *pCS20* nested PCR [28], considered as the OIE gold
211 standard PCR for *E. ruminantium* molecular detection. For the nested PCR, only 1µl
212 of DNA was amplified instead of 2µl for qPCR, with a final concentration of $1.5 .10^6$
213 to 1.5 copies/sample.

214

215 ***Sensitivity and specificity.*** The analytical sensitivity of the *pCS20* Sol1^{TM} qPCR was
216 evaluated with DNA extracted from 16 *E. ruminantium* strains isolated in different
217 geographical areas (Sudan, Burkina Faso, Senegal, South Africa, Zambia, Ghana,
218 Cameroon, Mozambique and Guadeloupe). Analytical specificity was evaluated using
219 nine closely related pathogens (*Anaplasma marginale, Babesia bovis* and *B. bigemina*
220 from Argentina; *Anaplasma phagocytophillum*, *A. platys* (previously *E. platys*),
221 *Ehrlichia muris*, *E. canis*, *Rickettsia felis*, *R. parkeri* from USA (Table 2). In addition,
222 nine DNA samples extracted from non-infected *A. variegatum* adult ticks (details in the
223 next section), were obtained from the tick rearing stock of the CIRAD laboratory and
224 included as negative controls.

Development of tick 16S rDNA qPCR for DNA extraction and PCR control

This test was designed to detect a conserved gene of ticks in order to evaluate the efficiency of the nucleic acid extraction using the automatic platform and also the presence of inhibitors during the qPCR. The test is a SYBR Green qPCR targeting the mitochondrial 16S ribosomal DNA (rDNA) gene and named $16S^{SG}$ rDNA qPCR in what follows.

Design of 16S rDNA primers. The forward primer 16SF 5'-CTGCTCAATGATTTTTTAAATTGCTGTGG-3' was selected from a previous paper [29] but a new reverse primer 16SR2 5'-TCTTAGGGTCTTCTTGTCDTTAATTTT-3' was designed in order to obtain an optimal product size no longer than 200 bp for qPCR (Table 1). The design of the new reverse primer 16SR2 using Primer3 [30] was based on the alignment of the partially conserved region of 16S rDNA from *Rhipicephalus geigyi* strain C7M (KF569942.1), *Ixodes minor* (KF793047.1), *Amblyomma boeroi* voucher INTA 2185 (JN828797.1), *A. glauerti* (AGU95853), *A. maculatum* clone TD02-239.T16s (AY375442.1, *I. minor* (KF793047.1), *A. glauerti* (AGU95853) and *A. maculatum* clone TD02-239.T16s (AY375442.1).

Real-time PCR setup. The qPCR was optimized by testing a temperature gradient from 58ºC to 61ºC and different primer concentrations. The qPCR assays were performed using Power SYBR® Green PCR Master Mix (Life Technologies, France). Each 25 µl reaction mixture contained 250 nM of each forward and reverse primer and standard concentrations of SYBR Green Dye, ROX™, AmpliTaq Gold DNA Polymerase LD, dNTPs with dUTP/dTTP blend, optimized buffer components, 12.5 µl of distilled water and 2.0 µl sample DNA. Platforms 7000 and 7500 (Applied Biosystems) were indifferently used for DNA amplification. The thermocycling conditions were finally set at 2’ at 50°C for activation of Uracil-N-Glycosylase, 10’ at 95°C for inactivation of Uracil-N-Glycosylase and activation of the AmpliTaq Gold® DNA polymerase, and 40 cycles of 15’ at 95°C and 60’ at 59°C. The qPCR controls included a positive control

(*A. variegatum* DNA) and a negative control (water). The results of the 16S^{SG} rDNA qPCR were analyzed by 7500 System SDS Software (Applied Biosystems).

Efficiency, limit of detection and reproducibility. Ten-fold serial dilutions of DNA extracted from a single field tick (*A. variegatum*), were tested in triplicates to evaluate the analytical performance of the new 16S^{SG} rDNA qPCR. Briefly, the tick was grinded individually using a Tissue lyser II (Qiagen, France). One steel bead of 5 mm was added to the tick in a 2 ml Eppendorf tube (Eppendorf, France) and kept at -80ºC for at least 2 hours or preferably overnight. The tick was disrupted twice in the Tissue lyser II at an oscillation frequency of 30 Hz during 2 minutes. The mashed tick was then resuspended in 450 µl of sterile PBS, vortexed and centrifuged twice at 8000 rpm for 30 sec and the supernatant recovered for nucleic acid extraction. DNA was extracted with the QiaAmp DNA minikit (Qiagen, Courtaboeuf, France) according to the supplier's instructions with a slight adjustment: tick samples of 25 to 40 mg were lysed with 180 µl of buffer ATL and 20 µl of RNase A at 20 mg/ml (Sigma-Aldrich, France). The amplification efficiency (E) and percentage of efficiency were calculated as described previously [28]. An average Ct and standard deviation (±SD) was calculated on the triplicates in order to assess the reproducibility of the 16S^{SG} rDNA qPCR.

Quality control criteria. In order to set the threshold of the new 16S^{SG} rDNA qPCR, a panel of 37 field ticks *A. hebraeum* and *A. variegatum,* collected in Mozambique and South Africa, were individually extracted on the automatic platform (as described below) and tested. The mean Ct value was calculated for these 37 tests and the upper limit to validate both the automatic extraction of nucleic acids and the absence of inhibitors in the real-time qPCR, was set using the formula:

$Ct_{sample} <$ mean Ct value $_{37\ ticks} + 2$ SD.

Validation of the whole method: performance of the automatic tick DNA extraction and *pCS20* Sol1™ qPCR

The detection limit of the automatic extraction coupled with the *pCS20* Sol1™ qPCR was first compared with the gold standard consisting in a manual extraction coupled with the nested *pCS20* PCR on tick lysates spiked with *E. ruminantium* cell culture. In a second step, performances of the automatic and manual extractions were compared using indifferently *pCS20* nested PCR and Sol1™ qPCR on tick lysates spiked with *E. ruminantium* cell cultures and experimentally infected ticks. Relative sensitivity and specificity of Sol1™ qPCR compared to nested *pCS20* PCR and combined with MLST were measured using field ticks extracted both manually and automatically. The choice of taking as a reference the combined results of nested *pCS20* PCR and MLST was made to strengthen the sensitivity and specificity of the nested PCR. Multi-band PCR products which includes a band at expected size (120 bp) can be observed using nested *pCS20* PCR and are thus non-interpretable alone. The use of a second test such as MLST is then important to determine the true status of a multiband result.

Then, the relative sensitivity and specificity of the whole method were assessed in comparison with the reference method consisting of manual DNA extraction and nested *pCS20* PCR on *E. ruminantium* serial dilutions and experimentally infected ticks. Finally, the reproducibility of the whole method including automatic DNA extraction and *pCS20* Sol1™ qPCR was determined.

Limit of detection. For the assessment of the detection limit, pools of grinded ticks were prepared to homogenize the material and spiked with serial dilutions of *E. ruminantium* passage 43 from infected cell culture with a concentration ranging from 6.10^3 to 6 copies/sample. Briefly, uninfected ticks from the rearing stock of CIRAD were grinded individually in a 2x24 tubes format using a Tissue lyser II as previously described. Crushed ticks were pooled into 3 groups of 10 ticks, each group resuspended in a final volume of 2 ml PBS. To achieve this, the first tick (first tube) was resuspended in 2 ml PBS, then the whole volume was passed on the next tick and so forth up to the tenth tick, pooling most of the lysed tissue and PBS in one tube. The 2 ml tick

311 suspension in PBS was then vortex, centrifuged twice at 8000 rpm for 30 sec and only
312 the supernatant recovered for subsequent extraction. This supernatant was used directly
313 for nucleic acid extraction to generate a negative control for the PCR or spiked with
314 serial dilutions of *E. ruminantium* from infected cells to determine the detection limit.
315 One hundred fifty µl of the tick sample was mixed with 150 µl of ten-fold serial
316 dilutions of *E. ruminantium* strain Gardel in order to generate 20 test samples. Samples
317 were processed twice for each method, automatic extraction and Sol1 QPCR or manual
318 DNA extraction and nested PCR, to compare the limit of detection of both methods.

319

320 ***Performance of the manual and automatic DNA extraction***. The automatic DNA
321 extraction was performed on 150 µl of the spiked or unspiked tick supernatant, using
322 the Biomek 4000 automated liquid handling robot (Beckman Coulter) and the kit "Viral
323 RNA and DNA from "Macherey-Nagel" in a 96-well plate format, according to the
324 supplier's instructions. This kit was selected and evaluated with the objective to use a
325 single extraction procedure to obtain tick, bacterial and viral RNA/DNA. Final elution
326 of nucleic acids was done by two successive distributions of 100 and 50 µl of nuclease-
327 free water. After extraction, the plate containing the nucleic acids was stored at -20°C
328 until use. For comparison purposes, 150 µl of the same samples were extracted
329 manually in parallel, as described in the previous section. Comparison of automated
330 and manual extractions was done using either the downstream *pCS20* Sol1^{TM} qPCR or
331 the gold standard nested *pCS20* PCR. For combinatorial purposes, quantitative and
332 qualitative results of the two tests were secondarily converted as positive/detected or
333 negative/not detected. Paired extractions (manual versus automated) were generated on
334 samples prepared for the assessment of the detection limit as described earlier in this
335 section and tested by qPCR (n=17) and nested PCR (n=17) with exclusion of any
336 doubtful results both by nested PCR (multiband PCR products) or by qPCR (Ct>limit
337 of positivity). Likewise, 30 samples from *A. variegatum* adults moulted from nymphs
338 engorged on goats challenged experimentally with the *E. ruminantium* strain Bekuy
339 255 were submitted to automatic and manual DNA extractions and then to repeated
340 qPCR (n=30) and nested PCR (n=30). It is worthy to note that only a proportion of 30%

(9/30) of experimentally infected ticks were then found to be infected after experimental infection. The degree of agreement between the two extraction methods was calculated using Kappa statistics [31]. Kappa values are interpreted as following: ≥0.81 is very good agreement, from 0.61 to 0.80 is a good agreement, from 0.41 to 0.6 is moderate agreement, from 0.21 to 0.4 is fair agreement and ≤0.20 is poor agreement [31].

In addition, distributions of the Ct values generated by *pCS20* Sol1™ QPCR (n=17) on samples both extracted automatically and manually, were represented onto a 2-D dot plot [32].

Relative sensitivity and specificity of pCS20 Sol1™. Relative sensitivity, specificity and accuracy of the *pCS20* Sol1™ qPCR were determined on 60 field ticks indifferently extracted manually or automatically. Adult *A. hebraeum* and *A. variegatum* ticks were collected from cattle from several localities in Mozambique and South Africa during another epidemiological study [33]. The true status (positive/negative) of these ticks was established by the combined results of two tests, here below named as the reference method. The first test is the OIE gold standard *pCS20* nested PCR [11]. The second test is based on Multi Locus Sequence Typing (MLST) performed according to Adakal *et al*. (2009), with a small modification: only five out of 8 housekeeping genes, *lipA*, *lipB*, *secY*, *sodB* and *sucA* were amplified and sequenced [34]. There was a potential default of sensitivity and specificity of the nested PCR and it was decided to combine with MLST to improve the detection. Given that no cross-reactions with closely related pathogens were evidenced for the *pCS20* nested PCR and MLST [11, 34, 34], we consider that these tests cannot render false positive results. For some samples, only partial amplification of genes (1 out of 5 genes) from MLST occurred and were classified as positive samples. Therefore, a sample was considered as non-detected or negative when the two tests did not render any positive result. However, when multi-bands were detected for nested *pCS20* PCR and MLST was negative for any sample, it was also considered as negative. In the other cases, the sample was considered as positive or detected.

The relative sensitivity is the ability of the test *pCS20* Sol1™ qPCR to detect samples scored positive by the reference method (*pCS20* nested PCR and MLST): Se = 100*TP/ (TP+FN) %, where TP stands for true positive (e.g. positive in the two tests) and FN stands for false negative (e.g. negative with *pCS20* Sol1™ qPCR but positive in the reference method). The relative specificity is the ability of the *pCS20* Sol1™ qPCR to score as not detected or negative, samples that were not detected by the reference method: Sp = 100*TN/ (TN+FP) %, where TN stands for true negative (e.g. negative in the two tests) and FP stands for false positive (e.g. positive with *pCS20* Sol1™ QPCR and negative in the reference method), [35]. TN status was defined for samples with multiband for nested PCR and negative for MLST. Results of the Sol1™ qPCR and the *pCS20* nested PCR in combination with MLST were cross-tabulated (2 x 2 table). The relative accuracy (Ac) is the degree of agreement between the results obtained by the *pCS20* Sol1™ qPCR and the reference method: Ac=100*(TP+TN)/ (TP+TN+FP+FN) %. Additionally, the Kappa agreement between the two methods was determined as previously described.

Relative sensitivity and specificity of the whole method. Finally, the comparison of the whole method (automatic extraction + *pCS20* Sol1™ QPCR) versus the standard method (manual extraction + nested *pCS20* PCR) was done on 17 samples spiked with serial tenfold *E. ruminantium* dilutions from infected cell culture and 30 experimentally infected ticks as previously described. Any doubtful sample, multiband PCR product or Ct>limit of positivity, was excluded from the analysis. Sensitivity, specificity, accuracy and kappa agreement between the two methods were determined accordingly.

Reproducibility. In order to estimate the reproducibility of the automatic DNA extraction coupled with the *pCS20* Sol1™ qPCR, *E. ruminantium* strain Gardel passage 43 was appropriately diluted and added to tick supernatant to achieve concentrations of

60 and 6 copies/sample. Each spiked tick supernatants were extracted in triplicates in separate procedures and further tested by qPCR.

RESULTS

Development of *pCS20* Sol1^{TM} and Sol1^{SG} qPCRs

Optimization and efficiency of pCS20 Sol1 ™ and Sol1^{SG} qPCR. The *pCS20* Sol1^{TM} and Sol1^{SG} qPCR efficiencies (%) were tested at different temperatures from 50°C to 56°C using 10-fold serial dilutions of *E. ruminantium* Gardel DNA (from 3.10^6 to 30 copies/sample) in 3 separate experiments.

For *pCS20* Sol1^{SG} qPCR, the optimal annealing temperature was 51°C with an efficiency of 98.1% (+/-1.9) (data not shown). The mean expected temperature of dissociation for Sol1^{SG} qPCR using the serial dilution of the positive control *E. ruminantium* Gardel in triplicates was 74.2ºC (±0.5). Maximal PCR efficiency of Sol1^{TM} was obtained at 55°C with 94.4% (±3.6). At 56°C, an efficiency of 93.8% (±5.8) was very close to the maximal value at 55°C but with a higher variation between the tests (data not shown). Even at 54°C, the efficiency was still good with 89.1% (±6.1), (data not shown).

The optimal PCR conditions were defined for the new *pCS20* Sol1^{TM} and Sol1^{SG} qPCR, using an annealing temperature of 55°C and 51°C respectively.

For Cow^{TM} qPCR, the use of the optimal temperature of 48°C recommended by the authors, allowed the detection of 3.10^4 copies/sample in only one instance out of three independent assays. Thus it was not possible to determine PCR efficiency due to a lack of analytical sensitivity. Cow^{TM} qPCR was then tested at 56°C, a temperature closer to the melting temperature of the probe: it resulted in a low PCR efficiency of 69.2% (±3.1) estimated only for 4 dilutions (from 3.10^6 to 3.10^3 copies/sample). The Ct for 3.10^2 copies/sample was higher than 38 or undetermined (data not shown). Thus, the limit of detection for Cow^{TM} qPCR at 56°C was at 3 10^3 copies of bacteria per sample.

Limit of detection and reproducibility of pCS20 Sol1 ™ and Sol1^{SG} qPCR. At the optimal annealing temperatures, *pCS20* Sol1^{TM} and Sol1^{SG} qPCRs were performed on *E. ruminantium* Gardel DNA at 3.10^6 to 3 copies/sample in parallel with *pCS20* nested PCR. The mean of Ct values of three independent runs and standard deviations are shown in Table 3. The limit of detection of both *pCS20* Sol1 qPCRs is similar to *pCS20* nested PCR, with detection until 30 copies/sample and 15 copies/sample respectively (Table 3). The results obtained with the *pCS20* Sol1^{TM} gave a Ct value of 34.2 (+/-0.3) for 30 copies/sample and even at 3 copies tested once, we detected a Ct of 36.68. Thus, up to Ct of 37, samples were considered as positive samples. For Sol1^{SG} qPCR, the Ct was 30.5 (+/-1.3) and 34 (+/-0.7) for 30 copies and 3 copies respectively. However, for 3 copies, the dissociation curve was bi-phasic, highlighting the presence of both primer dimer product and the *pCS20* specific PCR product. Moreover, a signal for Sol1^{SG} qPCR was detected for non-template control with a Ct of 35 (± 1.1) due to dimers of primers as evidenced by a lower temperature than the expected dissociation temperature of the target (data not shown). The positive threshold for Sol1^{SG} qPCR was established at 35 Ct but particular attention will be drawn on the associated dissociation curve to detect any unspecific PCR product for samples around this Ct of 35 cycles.

In general, at a Ct>37 and Ct>35 for Sol1^{TM} and Sol1^{SG} qPCR respectively, with an expected amplification curve shape, samples were doubtful. However, if there was an atypical amplification curve, samples were considered as negative.

For both *pCS20* QPCRs, the standard deviation of Ct was extremely low, ranged from 0 to 1.3 and demonstrating a good reproducibility of both assays (Table 3).

Sensitivity and specificity of pCS20 Sol1^{TM}. Sixteen *E. ruminantium* strains from different geographic origins (Table 2) were successfully amplified by both *pCS20* Sol1^{TM} qPCR and gold standard test *pCS20* nested PCR. There was a weak positive signal for Banankeledaga strain for *pCS20* Sol1^{TM} qPCR compared to a strong signal for *pCS20* nested PCR. Concerning the specificity of the assay, there was no detection

of *A. marginale*, *A. phagocytophilum*, *A. platys*, *B. bovis* and *bigemina*, *E. canis* and *muris*, *R. felis* and *parkeri* by both *pCS20* Sol1™ qPCR and *pCS20* nested PCR. Moreover, nine uninfected *A. variegatum* DNA samples from CIRAD rearing facilities were used as negative controls and as expected, there was no signal either with *pCS20* Sol1™ qPCR and *pCS20* nested PCR, demonstrating the specificity of Sol1™ qPCR.

Development of tick 16S^{SG} rDNA qPCR for DNA extraction and PCR control

16S^{SG} rDNA qPCR efficiency and limit of detection. Serial dilutions of *A. variegatum* DNA from 10^{-1} to 10^{-5} were amplified in triplicates with 16S^{SG} rDNA qPCR at different temperatures from 58°C to 61°C. The range of percentage of efficiency was between 80% and 84%, depending on the temperature of hybridization with no significant difference (data not shown). However at 60ºC and 61ºC, Cts, were higher compared to 58°C and 59°C, with an increment of two to seven Ct for each dilution. Also, there was a lack of 16S detection for dilution 10^{-5} at 61ºC. The optimal temperature of hybridization for the 16S^{SG} rDNA qPCR was defined at 59°C with 84% (±5.1) of efficiency, based on five replicates. The mean expected temperature of dissociation for 16S^{SG} rDNA qPCR using the serial dilutions of the positive control was 72.1ºC (± 0.2, n=5).

Quality control of the automatic DNA extraction and reproducibility. From a panel of 37 field samples submitted to automatic DNA extraction, all samples were successfully amplified by 16S^{SG} rDNA qPCR, with a mean Ct of 23.3 (±2.8) suggesting good DNA quality extracted by the robot and no inhibitors (data not shown). The acceptable limit Ct attesting good DNA quality was calculated as the mean of Ct + 2xSD and was 28.9 Ct. Moreover, the reproducibility of the whole method, DNA extraction and 16S^{SG} qPCR was evaluated and variation was less than 4 Ct between ticks extracted at four different periods.

Validation of the whole method: performance of the automatic tick DNA extraction coupled with *pCS20* Sol1™ qPCR

Limit of detection. The automatic DNA extraction followed by *pCS20* Sol1™ qPCR allowed the detection of *E. ruminantium* from infected cell culture until 60 copies/sample with a Ct=33+/-1.4. These results were based on two independent assays on tick lysate spiked with *E. ruminantium* Gardel passage 43 (data not shown). At 6 copies/sample, there was a detection signal with a high Ct of 37.6+/- 1. In parallel, the same samples were processed by manual extraction and *pCS20* nested PCR. Detection of *E. ruminantium* was also obtained until 6 copies by manual extraction and p*CS20* nested PCR showing the same performance as automatic extraction and Sol1 qPCR.

Comparison of automatic and manual DNA extraction performances. Comparison between automatic and manual DNA extraction was globally performed on a total of 94 samples screened indifferently by *pCS20* Sol1™ qPCR or nested pCS20: 17 samples of tick lysates spiked with *E. ruminantium* serial dilutions and 30 samples from adult ticks moulted from nymphs that were experimentally engorged on infected goats. Sensitivity, specificity, relative accuracy and Kappa obtained for automatic extraction method were good with 84.1%, 88%, 86.2% and 72% (good agreement), respectively (Table 4).

Six and seven samples not detected for manual and automatic extraction respectively, were balanced between the two extractions methods, thus suggesting equal performances of the automated compared to the manual extraction (Table 4). Comparison of Ct values obtained for tick lysates spiked with *ER* serial dilutions and experimentally infected ticks extracted automatically and manually in parallel are shown in Figure 1. A good correlation was observed (R^2=85%) and the Ct values were slightly improved with the automatic DNA extraction (-1.96 Ct, $p<0.001$). In conclusion, automatic DNA extraction method had at least the same performance as

the manual extraction for subsequent detection of *E. ruminantium* by *pCS20* Sol1^{TM} QPCR and nested QPCR.

Relative sensitivity and specificity of the pCS20 Sol1^{TM} QPCR. Compared to the reference method (*pCS20* nested PCR and MLST combined), the relative sensitivity and specificity of the *pCS20* Sol1^{TM} QPCR were 75.8% and 85.2% respectively with an accuracy of 80% (Table 5, n=60 field ticks extracted manually and automatically). Within the 8 false negative samples, four were positive for both MLST and nested *pCS20* PCR and four were positive for nested PCR and negative for MLST. On the four false positive samples, 2 displayed multi-bands by nested PCR and are negative by MLST and 2 are negative both by MLST and nested PCR. On the 25 true positive samples, 76% (19 samples) are positive for nested PCR and MLST (5 with partial MLST amplification), 12% (3 samples) are negative by nested PCR and positive by MLST (with partial amplification only) and 12% (3 samples) are positive by nested PCR and negative by MLST (data not shown). On 23 true negative samples, 78% (18 samples) with multi-bands for *pCS20* nested PCR were negative for both MLST and qPCR (data not shown).

The Kappa statistics for the *pCS20* Sol1^{TM} QPCR/*pCS20* nested PCR+MLST comparisons were 60%, demonstrating a fair to good agreement between the tests.

Relative sensitivity and specificity of the whole method. The relative sensitivity and specificity of the whole method, automatic DNA extraction+*pCS20* Sol1^{TM} qPCR as compared with manual extraction+*pCS20* nested were 76.2% and 73.1%, respectively (Table 6, n= 17 *E. ruminantium* serial dilutions and 30 experimentally infected ticks), with more positive (7 samples) detected compared to manual extraction + nested PCR (5 samples). Four out of seven false positive samples had a Ct of 37, at the limit of positivity. The five false negative samples were clearly positive by nested *pCS20* PCR. The Kappa test for this analysis was 49%, demonstrating a moderate agreement between tests.

Reproducibility of the automatic DNA extraction and pCS20 Sol1 ™ QPCR. The reproducibility of the whole method (automated extraction + *pCS20* Sol1™ qPCR) on independent triplicates was high as demonstrated with the low standard deviation of Ct values: Ct=33±0.7 (CV=2.1%) and Ct=36.8±1.3 (CV=3.5%) for samples titrating 60 and 6 *E. ruminantium* copies/sample, respectively (data not shown).

DISCUSSION

We have developed a new automatic DNA extraction for *A. hebraeum*, *A. variegatum* and more widely for ticks, giving reliable DNA quality and yield as well as a specific, sensitive and efficient new qPCR assay for the detection of *E. ruminantium*.

Development of *pCS20* Sol1™ and Sol1SG qPCRs

Both qPCR *pCS20* Sol1 using SYBR Green and TaqMan probe chemistries for detection of *E. ruminantium* have an efficiency greater than 94%, but when using SYBR Green, our results showed that aspecific signal could appear at very low bacteria load and in non-template control. SYBR Green is the most commonly used dye in qPCR. It is a cost-effective method, but it is known for producing false positive results due to the non-specific binding of double-stranded DNA molecules, especially at the limit of detection. Thus, attention should be given to the dissociation curve analysis of the SYBR Green qPCR. In contrast, the use of a TaqMan probe can guarantee the specificity of the fluorescence emitted and has been well documented in other studies [36-38]. TaqMan probes are routinely used for molecular diagnostic assays. Nonetheless, the new Sol1SG qPCR developed for *E. ruminantium* detection could be a cheaper alternative for use in the field and in low income country laboratories. Furthermore, similar PCR efficiencies obtained for 55ºC and 56ºC for Sol1™ demonstrate the robustness of the assay. Even at 54°C, the efficiency was lower but still acceptable. These characteristics could allow for the development of a multiplex PCR targeting several pathogens using this range of temperatures.

571 The limit of detection of the new *pCS20* Sol1^{TM} and Sol1^{SG} is 30 *E. ruminantium* copies
572 per sample, which is similar to that of the gold-standard *pCS20* nested PCR (15 copies
573 per sample). Even at 3 copies per sample, we obtained a signal for *pCS20* Sol1^{TM}. In
574 fact, the limit of detection of the nested *pCS20* in the current study is in accordance
575 with previous work, which determined a limit of detection of 6 copies per sample [18].
576 In this study, the threshold of positivity was established at 37 and 35 cycles for Sol1^{TM}
577 and Sol1^{SG} qPCRs, respectively. Above these thresholds, samples are considered as
578 doubtful if amplification curves are correct, or as negative if not. The limit of positivity
579 at 37 cycles is confirmed by results obtained on the automatic extraction and Sol1^{TM}
580 qPCR (whole method) with a Ct of 36.8 for 6 copies per sample. The performance of
581 the *pCS20* Sol1^{TM} was also similar to a multiplex qPCR developed recently, targeting
582 both Panola Mountain *Ehrlichia* and *E. ruminantium groEL* genes with a limit of 10
583 copies per sample [24]. Such a limit of detection is important as it allows for the
584 detection of low bacteria load in a sample from field infected ticks as shown in this
585 study. Moreover, *E. ruminantium* was successfully detected in blood samples of three
586 experimentally infected goats during hyperthermia by *pCS20* Sol1^{TM} qPCR, as
587 previously shown by Martinez *et al.* (2004) with *pCS20* nested PCR (data not shown)
588 [18].

589 Thus, *E. ruminantium* screening with both *pCS20* Sol1^{TM} and Sol1^{SG} qPCR, optimized
590 using ticks and *E. ruminantium* cell culture inoculum, can be easily extended to blood
591 samples and other tissues of suspicious clinical cases. The development of the new
592 diagnostic method will allow for the implementation of a quick and efficient diagnostic
593 assay on a large number of animal samples in case of suspicion of disease outbreaks in
594 heartwater-free areas. As an example, it will be useful in the American mainland, where
595 there is a potential risk of heartwater introduction [6, 39].

596 Besides the reference gene, *pCS20*, other genes such as *Map1* and *groEL* were used to
597 develop molecular tests for the diagnosis of heartwater in ruminants and for the
598 screening of ticks. *Map1* genes (major antigenic protein 1) are highly diverse across *E.*
599 *ruminantium* strains and might not always be effective to detect isolates from the field
600 [40-42]. Thus, the use of the *Map1* gene is not actually suitable for diagnostic tests.

Conversely, the *groEL* gene codes for highly conserved and essential proteins for the survival of cells and was demonstrated to be useful for the identification and characterization of the genus Rickettsia [43, 44].Thus, it is a possible gene candidate for the diagnostic of *E. ruminantium*. However, until now, it has been applied in strain typing rather than diagnostic and has not yet been tested on a wide panel of *E. ruminantium* strains compared to *pCS20* [9, 45].

The use of *pCS20* Cow™, described by Steyn *et al*. (2008), was surprisingly inefficient throughout this study [23]. Either no efficiency or low PCR efficiency (69.2%) was obtained, both at the previously defined optimal temperature of 48ºC and at 56°C, which is close to the theoretical annealing temperature of the probe. The only modifications from Steyn *et al*. (2008) were the use of a quencher without a 3' end phosphorylated, a different thermocycler (Light cycler system instead of 7000 and 7500 Applied Biosystem) and a different PCR kit. This lack of efficiency and reproducibility can probably be attributed to a PCR inhibitor, inappropriate PCR primer and/or probe design as well as inappropriate reagents and concentrations [46]. However, as we successfully optimized Sol1 qPCR, in parallel, using the same DNA and reagent kit, it cannot be explained by the presence of inhibitors or a problem with the reagents. Thus, the reason for unsuccessful implementation of *pCS20* Cow™ remains unknown. Njiiri *et al.* (2015), after screening *E. ruminantium* in blood samples from Western Kenya using *pCS20* Cow™, did not find any positive samples, including samples previously identified as positive by RLB [47]. The authors hypothesized that the lack of amplification was caused by the genetic variability of Kenyan strains, but did not describe any difficulties with PCR optimization. In our study, Gardel strain was used as previously tested by Steyn *et al.* (2008), [23] and Cow™ primers and probe hybridized completely on the Gardel *pCS20* gene. Thus, the lack of PCR efficiency cannot be explained by genetic diversity.

In terms of specificity and sensitivity, the new *pCS20* Sol1™ qPCR does not cross react with other tick-borne pathogens and can detect 16 *E. ruminantium* strains from a wide geographic range. Also, it has the same performance as the nested *pCS20* PCR except for the strain Banankeledaga. Thi difference of signal between the PCRs could be due

631 to a nucleotide mismatch at hybridization sites of Sol1 primers and probes, as the target
632 *pCS20* fragment for Sol1 qPCR is different from *pCS20* nested PCR. Conversely,
633 Cow^{TM} qPCR [23] cross-reacted with *E. chaffeensis* and *E. canis,* reducing the
634 specificity of the test. The sequence identities with Panola Mountain *Ehrlichia pCS20*
635 are94% and 100% for reverse on 16 nucleotides and forward on 17 nucleotides Sol1
636 primers, respectively. Thus, further analysis should be performed to check if the new
637 *pCS20* Sol1 qPCR does not cross react with Panola Mountain *Ehrlichia* DNA, as
638 demonstrated by Sayler *et al*. (2015), using the *groEL* qPCR [24].

639

640 **Development of tick 16S rDNA qPCR for DNA extraction and PCR control**

641 A new qPCR targeting the 16S rDNA gene from ticks was successfully optimized and
642 most of the samples presented a good DNA quality. Given the successful amplification
643 of the 16S gene in ticks, this real-time PCR targeting a conserved gene within tick
644 species can possibly be used for several tick genera and species such as *Amblyomma*,
645 *Rhipicephalus* and *Ixodes* [48, 49]. The use of a qPCR for DNA quality control is a
646 powerful method for DNA quality, quantification and yield assessment compared to
647 classical methods using gel migration and nanodrop [50]. This method has the
648 advantage of being able to determine the amount of nucleic acids that can be amplified
649 and the presence of inhibitors in the reaction mixture [50].

650 The use of a real-time PCR for DNA quantification can also turn the sample processing
651 easier and quicker. Accordingly, assessment of DNA quality is important to validate
652 DNA extraction method and hence, ensure successful amplification of *E. ruminantium*
653 DNA. Further, the use of the 16S qPCR for DNA quality evaluation circumvents the
654 limitations of photometric and fluorometric methods for DNA quality assessment when
655 using the “Viral RNA and DNA from Macherey-Nagel" kit due to the presence of a
656 carrier, which can induce an overestimation of the amount of nucleic acids. This carrier
657 RNA enhances binding of nucleic acids to the silica membrane and reduces the risk of
658 degradation, thus improving the performance of DNA extraction methodologies [51].

659

Validation of the whole method: performance of the automatic tick DNA extraction coupled with *pCS20* Sol1™ qPCR

We observed a similar performance between the automatic DNA extraction method and the manual extraction methods. Accordingly, DNA extraction methods can be interchangeable depending on the number of samples. The extracted DNA can subsequently be screened by *pCS20* Sol1™ qPCR or nested qPCR for *E. ruminantium* detection.

The performance of *pCS20* Sol1™ qPCR was tested by comparing the detection results with those obtained with *pCS20* nested PCR and MLST.

Considering, that these tests were not perfect and may not have detected all infected samples, we chose to combine the positive results of the *pCS20* nested PCR and MLST in order to have more robust data. Our results confirmed this choice. Indeed, we obtained results that were positive for nested PCR but negative for MLST, positive for MLST and negative for nested PCR, and some of these divergent results were positive with *pCS20* Sol1™ qPCR (data not shown). Within these Sol1™ qPCR positive samples, three samples were negative by nested PCR and positive by MLST (only samples with partial amplification of at least one MLST gene were noticeable, stressing high gene variability). These results showed that some strains that could not be amplified by nested PCR, probably due to genetic polymorphism of *E. ruminantium* *pCS20* fragment targeted by nested PCR, were detected with Sol1™ qPCR. Additionally, Sol1™ qPCR was also able to detect samples that were positive by nested PCR and negative by MLST. Moreover, on true negative samples, 78% with doubtful status (multiband PCR products) by nested *pCS20* PCR and negative by MLST, were also negative by qPCR. This showed that Sol1™ qPCR is able to detect negative samples (confirmed by MLST) which are doubtful by nested *pCS20*. The specificity of Sol1™ qPCR seemed to be better than that of nested *pCS20* PCR and avoid the problem of doubtful sample status.

As we did not observe any cross-reaction with closely related pathogens, it is possible that positive results obtained by *pCS20* Sol1™ that were negative by *pCS20* nested PCR and MLST, were real infected samples. In fact, the four false positives all had a high Ct value, close to the limit of positivity, which emphasizes the better sensitivity of Sol1™ qPCR. This confirmed that the limit of detection of Sol1™ qPCR could be better than that of nested PCR (Table 3). Moreover, half of them gave multiband results by nested PCR and were negative by MLST, suggesting that these were positive samples.

Therefore, the new PCR method allowed for an improvement in specificity and sensitivity for *E. ruminantium* detection compared to nested *pCS20* PCR.

Considering the whole method, including automatic extraction and Sol1™ qPCR, the sensitivity, specificity and accuracy values were good compared to manual extraction and nested *pCS20* PCR. There was a high number of false positives. However, as there is no cross-reaction with closely related pathogens and a good analytical specificity was found for qPCR alone, these samples are probably real positives that were not detected by conventional methods (manual extraction and *pCS20* nested PCR). This hypothesis was reinforced by the fact that four out of seven false positive samples had high Ct values, close to the positivity threshold. For samples with low loads of *E. ruminantium* (close to 6 copies per sample), it seems that the use of automatic extraction and Sol1™ qPCR performed better than conventional methods.

The commercial kit from Macherey-Nagel used in our experiments has the advantage of extracting DNA and RNA from viruses as well as bacteria, and is compatible with the Biomek 4000 automated liquid handling robot (Beckman Coulter, France).

This method has also been tested for Avian Influenza virus in our laboratory and demonstrated similar performance for virus detection (data not shown). Thus, the automatic DNA extraction method can also be useful to widely screen pathogens, including viruses in ticks. Moreover, this method is a powerful tool to obtain tick DNA from numerous samples and do further accurate genetic characterization and evolution

717 studies. On the same sample, genetic characterization of both pathogen and vector can
718 be performed allowing co-evolution analysis with gain of time due to DNA/RNA
719 automatic extraction step. The manual DNA extraction process is time consuming and
720 requires extensive training to allow reproducibility. The more samples processed, the
721 higher the probability of contamination because of manipulation [52, 53]. Thus,
722 automatic DNA extraction can possibly overcome the challenges faced by manual
723 extraction for high throughput screening of *E. ruminantium* in ticks. In our study, we
724 showed a very high reproducibility of the whole method, automatic DNA extraction
725 and Sol1 qPCR™.

726 However, automated systems are more expensive and thus less accessible.
727 Furthermore, it is more economical to load a full 96-well plate and have a significant
728 number of samples to process and justify the acquisition of the equipment.
729 Additionally, it also requires a considerable amount of space, which might not be
730 available and training for robot operation is needed [52, 53]. Specifically for the kit
731 "Viral RNA and DNA from Macherey-Nagel" tested, DNA extraction of blood samples
732 is still not available due to clotting when using the kit (Albina E., personal
733 communication). Further work is needed to optimize the conditions for the use of the
734 Viral RNA and DNA kit from Macherey-Nagel, in order to automatically extract blood
735 samples.

736 Manual tick DNA extraction or smaller versions of automatic DNA extraction robots
737 and the use of SYBR Green dye for qPCR should be considered in order to reduce
738 operational costs, especially in low-income countries.

739 An automatic DNA extraction process, coupled with a qPCR targeting the *pCS20* gene,
740 provides high sensitivity and specificity as well as a low contamination risk [53] and
741 faster detection of *E. ruminantium* in ticks. The whole method is sensitive to samples
742 with low loads of *E. ruminantium* and has the main advantages of giving a true infection
743 status, avoiding the doubtful status obtained by nested *pCS20* PCR.

744

745

CONCLUSIONS

The performance of the new qPCR *pCS20* Sol1^{TM} is equivalent to the gold standard test, *pCS20* nested PCR. Independently of the extraction methods, the *pCS20* Sol1^{TM} and Sol1^{SG} qPCRs could be valuable tools for *E. ruminantium* diagnosis in clinical case samples.

Also, the manual (Qiagen kit) and automatic DNA extraction methods (Viral RNA and DNA kit from Macherey-Nagel) show similar performance levels with regard to tick DNA extraction.

This study demonstrates that the new automatic *Amblyomma* tick DNA extraction method coupled with the new *pCS20* Sol1^{TM} qPCR for *E. ruminantium* have a good sensitivity, specificity and reproducibility, making them suitable for wide use in molecular epidemiological studies.

In the context of the implementation of control measures, the increase in processing speed gained by using qPCR is important, particularly in the case of outbreaks or clinical suspicion in Heartwater-free areas.

Overall, the present work developed and validated high throughput methods based on automatic DNA extraction and qPCR that have the potential to contribute to the improvement of *E. ruminantium* diagnostic capacity.

Moreover, a high throughput DNA extraction method using RNA/DNA kit and a new DNA quality control method based on 16S^{SG} rDNA qPCR could be widely used for tick genetic characterization studies, and for other bacteria and virus screening in field *Amblyomma* ticks as well as in other tick species.

AUTHOR CONTRIBUTIONS

NC and VP optimized and generated DNA extraction and PCR results. NC, VP, LB, EA, LN and NV interpreted results and wrote the manuscript. TL, NV, and LN designed the project. All authors critically reviewed and approved the final manuscript.

ACKNOWLEDGMENTS

DNA from *E. ruminantium* closely related species were kindly provided by Ulrike Munderloh from the University of Minnesota, Department of Entomology, Minnesota. We are grateful to the Mozambican Veterinary services, South African National Parks (Kruger National Park) and Zambeze Delta Safaris (Coutada 11 and 12) for support during sampling.

FUNDING INFORMATION

This work was financially supported by CIRAD and EPIGENESIS project which received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement No 31598". FUNDO ABERTO DA UEM 2012-2013 and FUNDO NACIONAL DE INVESTIGAÇÃO Projecto Nº 133-Inv/FNI/ 2012-2013 funded the field trips and reagents in Mozambique. This study was partly developed under the project MALIN "Surveillance, diagnosis, control and impact of infectious diseases of humans, animals and plants in tropical islands" supported by the European Union in the framework of the European Regional Development Fund (ERDF) and the Regional Council of Guadeloupe.

REFERENCES

1. Allsopp BA: **Trends in the control of heartwater.** Onderstepoort J Vet Res 2009, **76**(1):81-88.

2. Walker JB, Olwage A: **The tick vectors of *Cowdria ruminantium* (Ixodoidea, Ixodidae, genus *Amblyomma*) and their distribution.** Onderstepoort J Vet Res 1987, **54**(3):353-379.

3. Stachurski F, Tortosa P, Rahajarison P, Jacquet S, Yssouf A, Huber K: **New data regarding distribution of cattle ticks in the south-western Indian Ocean islands.** Vet Res 2013, **44**:79.

4. Provost A, Bezuidenhout JD: **The historical background and global importance of heartwater.** Onderstepoort J Vet Res 1987, **54**(3):165-169.

5. Minjauw B: **The economic impact of heartwater (*Cowdria ruminantium*) infection in the SADC region, and its control through the use of new inactivated vaccines.** ILRI, UF/USAID/SADC Heartwater Research Project Report 2000.

6. Roth JA, Richt JA, Morozov IA: **Vaccines and Diagnostics for Transboundary Animal Diseases.** In *Developments in Biologicals. Volume 135.* Ames, Iowa 17-19 September 2012 edition. Edited by Roth J.A, Richt J.A., Morozov I.A. 2013.

7. Faburay B, Geysen D, Ceesay A, Marcelino I, Alves PM, Taoufik A, Postigo M, Bell-Sakyi L, Jongejan F: **Immunisation of sheep against heartwater in The Gambia using inactivated and attenuated *Ehrlichia ruminantium* vaccines.** Vaccine 2007, **25**(46):7939-7947.

8. Adakal H, Stachurski F, Konkobo M, Zoungrana S, Meyer DF, Pinarello V, Aprelon R, Marcelino I, Alves PM, Martinez D, Lefrancois T, Vachiéry N: **Efficiency of**

inactivated vaccines against heartwater in Burkina Faso: impact of *Ehrlichia ruminantium* genetic diversity. Vaccine 2010, **28**(29):4573-4580.

9. Adakal H, Gavotte L, Stachurski F, Konkobo M, Henri H, Zoungrana S, Huber K, Vachiéry N, Martinez D, Morand S, Frutos R: **Clonal origin of emerging populations of *Ehrlichia ruminantium* in Burkina Faso.** Infect Genet Evol 2010, **10**(7):903-912.

10. Esemu SN, Besong WO, Ndip RN, Ndip LM: **Prevalence of Ehrlichia ruminantium in adult *Amblyomma variegatum* collected from cattle in Cameroon.** Exp Appl Acarol 2013, **59**(3):377-387.

11. Molia S, Frebling M, Vachiéry N, Pinarello V, Petitclerc M, Rousteau A, Martinez D, Lefrancois T: ***Amblyomma variegatum* in cattle in Marie Galante, French Antilles: prevalence, control measures, and infection by *Ehrlichia ruminantium*.** Vet Parasitol 2008, **153**(3-4):338-346.

12. Ammazzalorso AD, Zolnik CP, Daniels TJ, Kolokotronis SO: **To beat or not to beat a tick: comparison of DNA extraction methods for ticks (*Ixodes scapularis*).** PeerJ 2015, **3**:e1147.

13. Allender MC, Bunick D, Dzhaman E, Burrus L, Maddox C: **Development and use of a real-time polymerase chain reaction assay for the detection of *Ophidiomyces ophiodiicola* in snakes.** J Vet Diagn Invest 2015, **27**(2):217-220.

14. Vidergar N, Toplak N, Kuntner M: **Streamlining DNA barcoding protocols: automated DNA extraction and a new cox1 primer in arachnid systematics.** PLoS One 2014, **9**(11):e113030.

15. Rodriguez-Perez MA, Gopal H, Adeleke MA, De Luna-Santillana EJ, Gurrola-Reyes JN, Guo X: **Detection of *Onchocerca volvulus* in Latin American black flies**

for pool screening PCR using high-throughput automated DNA isolation for transmission surveillance. Parasitol Res 2013, **112**(11):3925-3931.

16. Moriarity JR, Loftis AD, Dasch GA: **High-throughput molecular testing of ticks using a liquid-handling robot.** J Med Entomol 2005, **42**(6):1063-1067.

17. Crowder CD, Rounds MA, Phillipson CA, Picuri JM, Matthews HE, Halverson J, Schutzer SE, Ecker DJ, Eshoo MW: **Extraction of total nucleic acids from ticks for the detection of bacterial and viral pathogens.** J Med Entomol 2010, **47**(1):89-94.

18. Martinez D, Vachiéry N, Stachurski F, Kandassamy Y, Raliniaina M, Aprelon R, Gueye A: **Nested PCR for detection and genotyping of *Ehrlichia ruminantium*: Use in genetic diversity analysis.** Ann N Y Acad Sci 2004, **1026**:106-113.

19. Faburay B, Geysen D, Munstermann S, Taoufik A, Postigo M, Jongejan F: **Molecular detection of *Ehrlichia ruminantium* infection in *Amblyomma variegatum* ticks in The Gambia.** Exp Appl Acarol 2007, **42**(1):61-74.

20. Vachiéry N, Jeffery H, Pegram R, Aprelon R, Pinarello V, Kandassamy RL, Raliniaina M, Molia S, Savage H, Alexander R, Frebling M, Martinez D, Lefrancois T: ***Amblyomma variegatum* ticks and heartwater on three Caribbean Islands.** Ann N Y Acad Sci 2008, **1149**:191-195.

21. Peixoto CC, Marcelino I, Vachiery N, Bensaid A, Martinez D, Carrondo MJ, Alves PM: **Quantification of *Ehrlichia ruminantium* by real time PCR.** Vet Microbiol 2005, **107**(3-4):273-278.

22. Postigo M, Taoufik A, Bell-Sakyi L, de Vries E, Morrison WI, Jongejan F: **Differential transcription of the major antigenic protein 1 multigene family of *Ehrlichia ruminantium* in Amblyomma variegatum ticks.** Vet Microbiol 2007, **122**(3-4):298-305.

23. Steyn HC, Pretorius A, McCrindle CM, Steinmann CM, Van Kleef M: **A quantitative real-time PCR assay for *Ehrlichia ruminantium* using pCS20.** Vet Microbiol 2008, **131**(3-4):258-265.

24. Sayler KA, Loftis AD, Mahan SM, Barbet AF: **Development of a Quantitative PCR Assay for Differentiating the Agent of Heartwater Disease, *Ehrlichia ruminantium*, from the Panola Mountain Ehrlichia.** Transbound Emerg Dis 2015.

25. Marcelino I, Sousa MF, Verissimo C, Cunha AE, Carrondo MJ, Alves PM: **Process development for the mass production of *Ehrlichia ruminantium*.** Vaccine 2006, **24**(10):1716-1725.

26. Frutos R, Viari A, Ferraz C, Bensaid A, Morgat A, Boyer F, Coissac E, Vachiery N, Demaille J, Martinez D: **Comparative genomics of three strains of *Ehrlichia ruminantium*: a review.** Ann N Y Acad Sci 2006, **1081**:417-433.

27. Pruneau L, Emboule L, Gely P, Marcelino I, Mari B, Pinarello V, Sheikboudou C, Martinez D, Daigle F, Lefrancois T, Meyer DF, Vachiery N: **Global gene expression profiling of *Ehrlichia ruminantium* at different stages of development.** FEMS Immunol Med Microbiol 2012, **64**(1):66-73.

28. Bustin SA, Benes V, Garson JA, Hellemans J, Huggett J, Kubista M, Mueller R, Nolan T, Pfaffl MW, Shipley GL, Vandesompele J, Wittwer CT: **The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments.** Clin Chem 2009, **55**(4):611-622.

29. Norris, DE, Klompen, JSH, Keirans, JE, Black, WC.: **Population genetics of *Ixodes scapularis* (Acari: Ixodidae) based on mitochondrial 16S and 12S genes.** J Med Entomol 1996, **33**:78-89.

30. Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M., & Rozen, S. G.: **Primer-3new capabilities and interfaces.** Nucleic Acids Res 2012, **40**(15):e115.

31. Fleiss JL, Levin B, Paik MC: **The Measurement of Interrater Agreement.** In *Statistical Methods for Rates and Proportions.* Edited by Fleiss JL, Levin B, Paik MC. John Wiley & Sons, Inc.; 2004:598-626.

32. Wilcoxon F: **Individual Comparisons by Ranking Methods.** Biometrics Bulletin 1945, **1**(6):80-83.

33. Bournez L, Cangi N, Lancelot R, Pleydell DR, Stachurski F, Bouyer J, Martinez D, Lefrancois T, Neves L, Pradel J: **Parapatric distribution and sexual competition between two tick species, *Amblyomma variegatum* and *A. hebraeum* (Acari, Ixodidae), in Mozambique.** Parasit Vectors 2015, **8**:504.

34. Adakal H, Meyer DF, Carasco-Lacombe C, Pinarello V, Allegre F, Huber K, Stachurski F, Morand S, Martinez D, Lefrancois T, Vachiéry N, Frutos R: **MLST scheme of *Ehrlichia ruminantium*: genomic stasis and recombination in strains from Burkina-Faso.** Infect Genet Evol 2009, **9**(6):1320-1328.

35. Lalkhen,AG, McCluskey,A: **Clinical tests: sensitivity and specificity.** Continuing Education in Anaesthesia, Critical Care & Pain 2008, **8**:221-223.

36. Heid CA, Stevens J, Livak KJ, Williams PM: **Real time quantitative PCR.** Genome Res 1996, **6**(10):986-994.

37. Kubista M, Andrade JM, Bengtsson M, Forootan A, Jonak J, Lind K, Sindelka R, Sjoback R, Sjogreen B, Strombom L, Stahlberg A, Zoric N: **The real-time polymerase chain reaction.** Mol Aspects Med 2006, **27**(2-3):95-125.

38. Valasek MA, Repa JJ: **The power of real-time PCR.** Adv Physiol Educ 2005, **29**(3):151-159.

39. Barré N, Uilenberg G, Morel PC, Camus E: **Danger of introducing heartwater onto the American mainland: potential role of indigenous and exotic *Amblyomma* ticks.** Onderstepoort J Vet Res 1987, **54**(3):405-417.

40. Allsopp MT, Dorfling CM, Maillard JC, Bensaid A, Haydon DT, van Heerden H, Allsopp BA: ***Ehrlichia ruminantium* major antigenic protein gene (map1) variants are not geographically constrained and show no evidence of having evolved under positive selection pressure.** J Clin Microbiol 2001, **39**(11):4200-4203.

41. Yu X-, McBride JW, Walker DH: **Restriction and expansion of Ehrlichia strain diversity.** Vet Parasitol 2007, **143**(3-4):337-346.

42. Raliniaina M, Meyer DF, Pinarello V, Sheikboudou C, Emboulé L, Kandassamy Y, Adakal H, Stachurski F, Martinez D, Lefrançois T, Vachiéry N: **Mining the genetic diversity of *Ehrlichia ruminantium* using map genes family.** Vet Parasitol 2010, **167**(2-4):187-195.

43. Sumner JW, Nicholson WL, Massung RF: **PCR amplification and comparison of nucleotide sequences from the groESL heat shock operon of Ehrlichia species.** J Clin Microbiol 1997, **35**(8):2087-2092.

44. Lee JH, Park HS, Jang WJ, Koh SE, Kim JM, Shim SK, Park MY, Kim YW, Kim BJ, Kook YH, Park KH, Lee SH: **Differentiation of rickettsiae by groEL gene analysis.** J Clin Microbiol 2003, **41**(7):2952-2960.

45. Allsopp MT, Allsopp BA: **Extensive genetic recombination occurs in the field between different genotypes of *Ehrlichia ruminantium*.** Vet Microbiol 2007, **124**(1-2):58-65.

46. Svec D, Tichopad A, Novosadova V, Pfaffl MW, Kubista M: **How good is a PCR efficiency estimate: Recommendations for precise and robust qPCR efficiency assessments .** Biomol Detect Quantif 2015, **3**:9.

47. Njiiri NE, Bronsvoort BMd, Collins NE, Steyn HC, Troskie M, Vorster I, Thumbi SM, Sibeko KP, Jennings A, van Wyk IC, Mbole-Kariuki M, Kiara H, Poole EJ, Hanotte O, Coetzer K, Oosthuizen MC, Woolhouse M, Toye P: **The epidemiology of tick-borne haemoparasites as determined by the reverse line blot hybridization assay in an intensively studied cohort of calves in western Kenya.** Vet Parasitol 2015, **210**(1–2):69-76.

48. Black WC,4th, Piesman J: **Phylogeny of hard- and soft-tick taxa (Acari: Ixodida) based on mitochondrial 16S rDNA sequences.** Proc Natl Acad Sci U S A 1994, **91**(21):10034-10038.

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

50. Boesenberg-Smith, KA, Pessarakli, MM, Wolk, DM: **Assessment of DNA Yield and Purity: an Overlooked Detail of PCR Troubleshooting.** Clin Microbiol Newsl 2012, **34**(1):1-6.

51. Shaw KJ, Thain L, Docker PT, Dyer CE, Greenman J, Greenway GM, Haswell SJ: **The use of carrier RNA to enhance DNA extraction from microfluidic-based silica monoliths.** Anal Chim Acta 2009, **652**(1-2):231-233.

52. Ivanova,NV., Dewaard, JR,HEBERT, PDN: **An inexpensive, automation-friendly protocol for recovering high-quality DNA.** Mol Ecol Notes 2006, **6**(4):998-1002.

53. Espy MJ, Uhl JR, Sloan LM, Buckwalter SP, Jones MF, Vetter EA, Yao JD, Wengenack NL, Rosenblatt JE, Cockerill FR,3rd, Smith TF: **Real-time PCR in clinical microbiology: applications for routine laboratory testing.** Clin Microbiol Rev 2006, **19**(1):165-256.

54. Pilet H, Vachiéry N, Berrich M, Bouchouicha R, Durand B, Pruneau L, Pinarello V, Saldana A, Carasco-Lacombe C, Lefrancois T, Meyer DF, Martinez D, Boulouis HJ, Haddad N: **A new typing technique for the Rickettsiales *Ehrlichia ruminantium*: multiple-locus variable number tandem repeat analysis.** J Microbiol Methods 2012, **88**(2):205-211.

55. Mathew JS, Ewing SA, Barker RW, Fox JC, Dawson JE, Warner CK, Murphy GL, Kocan KM: **Attempted transmission of *Ehrlichia canis* by *Rhipicephalus sanguineus* after passage in cell culture.** Am J Vet Res 1996, **57**(11):1594-1598.

56. Lynn GE, Oliver JD, Nelson CM, Felsheim RF, Kurtti TJ, Munderloh UG: **Tissue distribution of the *Ehrlichia muris*-like agent in a tick vector.** PLoS One 2015, **10**(3):e0122007.

57. Pornwiroon W, Pourciau SS, Foil LD, Macaluso KR: ***Rickettsia felis* from cat fleas: isolation and culture in a tick-derived cell line.** Appl Environ Microbiol 2006, **72**(8):5589-5595.

58. Paddock CD, Fournier PE, Sumner JW, Goddard J, Elshenawy Y, Metcalfe MG, Loftis AD, Varela-Stokes A: **Isolation of *Rickettsia parkeri* and identification of a novel spotted fever group Rickettsia sp. from Gulf Coast ticks (*Amblyomma maculatum*) in the United States.** Appl Environ Microbiol 2010, **76**(9):2689-2696.

TABLES AND FIGURES

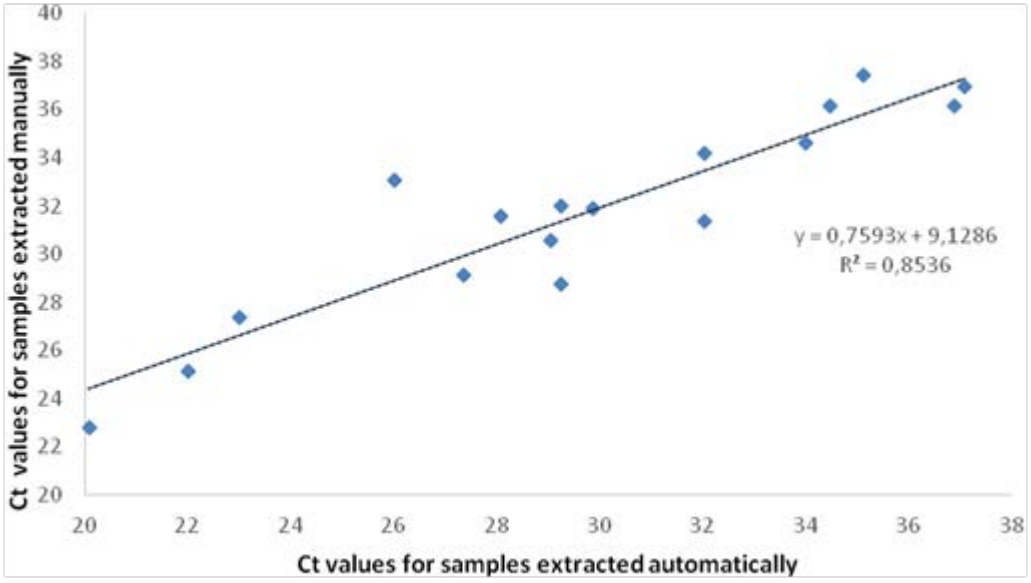


Figure 1 Comparison of Ct values obtained by Sol1^{TM} qPCR for tick lysates spiked with *E. ruminantium* serial dilutions and ticks that were experimentally engorged on infected goats extracted automatically and manually in parallel (n=17). Observed correlation was R^2=85%.

TABLE 1 Set of primer and probes for *pCS20* Sol1^{TM}, Sol1^{SG}, Cow^{TM} and 16S^{SG} qPCR

qPCR name	Primer/ Probe	Sequence (5' → 3')	Melting temp (°C)	Optimal Annealing temp (°C)	Product size (bp)	Reference
Sol1^{SG}	Sol1 F	ACA AAT CTG GYC CAG ATC AC	56.4	51	110 pb	This study
	Sol1 R	CAG CTT TCT GTT CAG CTA GT	56.5			
Sol1^{TM}	Sol1^{TM} probe	**6-FAM**-ATC AAT TCA CAT GAA ACA TTA CATG CAA CTG G-**BHQ1**	72.3	55		
Cow^{TM}	Cow F	CAA AAC TAG TAG AAA TTG CACA	56.3	48	226 bp	Adapted [23]
	Cow R	TGC ATC TTG TGG TGG TAC	57.9			
	Cow^{TM} probe	1. **FAM** -TCC TCC ATC AAG ATATATAGC ACC TAT TA-**TAM**	63.5	58		
16S^{SG}	16SF	CTGCTCAATGATTTT TTAAATTGCTGTGG	61	59		[29]
	16SR2	TCTTAGGGTCTTCTT GTCDTTAATTTT	66	59		This study

1057 **TABLE 2** Sensitivity and specificity of *pCS20* Sol1^{TM} qPCR targeting 16 *E.*
1058 *ruminantium* strains from different geographic origins and other closely related
1059 pathogens.

Name	**Origin**	**Geographical origin**	**Nested *pCS20* PCR**	**Sol1^{TM} Q PCR**	**Reference**
***E. ruminantium* strains**					
Blonde	CC p8	Guadeloupe	+	+	[42]
Gardel	CC p48		+	+	[26]
Sara 455	CC p10	Burkina Faso	+	+	[9]
Bekuy 255	CC p9		+	+	[42]
Bankouma 421	CC p15		+	+	[9]
Banankeledaga	CC p1		+	w+	[9]
Lamba 479	CC p16		+	+	[42]
Cameroun	CC p9	Cameroon	+	+	[42]
Pokoase 412	CC p10	Ghana	+	+	[42]
Senegal	CC p60	Senegal	+	+	[42]
Umbaneim	B	Sudan	+	+	[42]
Mara	CC p1	South Africa	+	+	[42]
Welgevonden	CC p12		+	+	[26]
Lutale	CC p6	Zambia	+	+	[42]
Sankat 430	CC p16	Ghana	+	+	[54]
Umpala	CC p6	Mozambique	+	+	[42]

Closely related species					
A. marginale		Argentina	-	-	
A. phagocytophilum	CC	USA	-	-	
A. platys (E. platys)	CC	USA	-	-	
B. bovis		Argentina	-	-	
B. bigemina			-	-	
E. canis	CC	USA	-	-	[55]
E. muris	CC		-	-	[56]
R. felis	CC	USA	-	-	[57]
R. parkeri	CC		-	-	[58]
9 uninfected A. variegatum	T	Guadeloupe	-	-	

Origin of samples: ticks (T), blood (B), cell culture passage (CC p), w+: weak positive

TABLE 3 Limit of detection of *pCS20* Sol1^{TM} qPCR, Sol1^{SG} QPCR and *pCS20* nested PCR

DNA *ER* strain Gardel (copies/sample)[a]	**Sol1^{TM} QPCR T=55°C Ct (± SD)[b]**	**Sol1^{SG} QPCR T=51°C Ct (± SD)[b]**	***pCS20* nested PCR signal[c]**
3.10^6	17.0 (± 0.3)	13.6 (± 0.6)	+
3.10^5	20.3 (± 0.1)	16.8 (± 0.4)	+
3.10^4	23.6 (± 0.0)	20.0 (± 0.5)	+
3.10^3	27.1 (± 0.6)	23.5 (± 0.4)	+
3.10^2	31.0 (± 0.5)	26.9 (± 1.0)	+
30	34.2 (± 0.3)	30.5 (± 1.3)	w+
3	36.68*	34 (± 0.7)	-
NTC	Undet	35.0 (± 1.1)	-

[a]: Bacteria quantity used for qPCR; for the nested PCR samples were amplified from 1µl of DNA, thus containing half quantity from 1.5 10^6 to 1.5 copies/sample; [b]: The average Ct value is indicated for each dilution (bacteria copy number) and standard deviation is from 3 replicates; [c]: conventional PCR done in duplicate; *: tested once; *ER*: *E. ruminantium*; T: *temperature of hybridization;* w+: weak positive; NTC: non-template control; Undet: Undetermined.

TABLE 4 Comparison between automatic and manual DNA extraction using *E. ruminantium* Gardel cell culture in tick lysate and experimentally infected ticks. All tick lysates spiked with *E. ruminantium* serial dilutions and ticks that were experimentally engorged on infected goats tested by nested PCR (n=17 and n=30, respectively) or *pCS20* Sol1™ qPCR (n=17 and n=30, respectively) were integrated.

		Manual extraction		Total	Se	Sp	Ac
		+	-				
Automatic extraction	+	37	6	43	84.1%	88%	86.2%
	-	7	44	51			
Total		44	50	94			

Sensibility (Se); Specificity (Sp)**;** Accuracy (Ac); The Kappa test for this analysis was 72% (good agreement).

TABLE 5 Relative sensitivity and specificity of *pCS20* Sol1™ qPCR as compared with *pCS20* nested PCR and MLST combined as the reference method on 60 field ticks.

		Nested PCR + MLST		Total	Se	Sp	Ac
		+	-				
Sol1™ qPCR	+	25	4	29	75.8%	85.2%	80%
	-	8	23	31			
Total		33	27	60			

Relative sensibility (Se); Relative specificity (Sp)**;** Accuracy (Ac). The Kappa test for this analysis was 60% (close to good agreement).

TABLE 6 Relative sensitivity and specificity of automated extraction + *pCS20* Sol1™ qPCR as compared with manual extraction + *pCS20* nested PCR (OIE gold standard) run on 17 *E. ruminantium* serial dilutions and 30 experimentally infected ticks.

		Manual extraction + *pCS20* nested PCR		Total	Se	Sp	Ac
		+	-				
Automatic extraction +Sol1™ qPCR	+	16	**7**	23	76.2%	73.1%	74.5%
	-	**5**	19	24			
Total		21	26	47			

Relative sensibility (Se); Relative specificity (Sp)**;** Accuracy (Ac); The Kappa test for this analysis was 49% (moderate agreement).

IV. **Section II**

Recombination is a major driving force of genetic diversity in the Anaplasmataceae *Ehrlichia ruminantium*

(Article 2: published in Frontiers in Cellular and Infection Microbiology)

frontiers in Cellular and Infection Microbiology

ORIGINAL RESEARCH
published: 29 September 2016
doi: 10.3389/fcimb.2016.00111

Recombination Is a Major Driving Force of Genetic Diversity in the Anaplasmataceae *Ehrlichia ruminantium*

*Nídia Cangi[1,2,3,4†], Jonathan L. Gordon[1,2†], Laure Bournez[1,2], Valérie Pinarello[1,2], Rosalie Aprelon[1,2], Karine Huber[2], Thierry Lefrançois[2], Luís Neves[3,5], Damien F. Meyer[1,2] and Nathalie Vachiéry[1,2]**

[1] CIRAD, UMR CMAEE, Petit-Bourg, Guadeloupe, France, [2] INRA, UMR1309 CMAEE, Montpellier, France, [3] Centro de Biotecnologia-UEM, Eduardo Mondlane University, Maputo, Mozambique, [4] Université des Antilles, Pointe-à-Pitre, Guadeloupe, France, [5] Department of Veterinary Tropical Diseases, Faculty of Veterinary Science, University of Pretoria, Onderstepoort, South Africa

OPEN ACCESS

Edited by:
Robert Heinzen,
National Institute of Allergy and Infectious Diseases, USA

Reviewed by:
Xue-jie Yu,
University of Texas Medical Branch, USA
Talima Pearson,
Northern Arizona University, USA
Job E. Lopez,
Baylor College of Medicine, USA

***Correspondence:**
Nathalie Vachiéry
nathalie.vachiery@cirad.fr

[†] These authors have contributed equally to this work and are co-first authors.

Received: 30 May 2016
Accepted: 09 September 2016
Published: 29 September 2016

Citation:
Cangi N, Gordon JL, Bournez L, Pinarello V, Aprelon R, Huber K, Lefrançois T, Neves L, Meyer DF and Vachiéry N (2016) Recombination Is a Major Driving Force of Genetic Diversity in the Anaplasmataceae Ehrlichia ruminantium. Front. Cell. Infect. Microbiol. 6:111. doi: 10.3389/fcimb.2016.00111

The disease, Heartwater, caused by the *Anaplasmataceae E. ruminantium*, represents a major problem for tropical livestock and wild ruminants. Up to now, no effective vaccine has been available due to a limited cross protection of vaccinal strains on field strains and a high genetic diversity of *Ehrlichia ruminantium* within geographical locations. To address this issue, we inferred the genetic diversity and population structure of 194 *E. ruminantium* isolates circulating worldwide using Multilocus Sequence Typing based on *lipA*, *lipB*, *secY*, *sodB,* and *sucA genes*. Phylogenetic trees and networks were generated using BEAST and SplitsTree, respectively, and recombination between the different genetic groups was tested using the PHI test for recombination. Our study reveals the repeated occurrence of recombination between *E. ruminantium* strains, suggesting that it may occur frequently in the genome and has likely played an important role in the maintenance of genetic diversity and the evolution of *E. ruminantium*. Despite the unclear phylogeny and phylogeography, *E. ruminantium* isolates are clustered into two main groups: Group 1 (West Africa) and a Group 2 (worldwide) which is represented by West, East, and Southern Africa, Indian Ocean, and Caribbean strains. Some sequence types are common between West Africa and Caribbean and between Southern Africa and Indian Ocean strains. These common sequence types highlight two main introduction events due to the movement of cattle: from West Africa to Caribbean and from Southern Africa to the Indian Ocean islands. Due to the long branch lengths between Group 1 and Group 2, and the propensity for recombination between these groups, it seems that the West African clusters of Subgroup 2 arrived there more recently than the original divergence of the two groups, possibly with the original waves of domesticated ruminants that spread across the African continent several thousand years ago.

Keywords: *Ehrlichia ruminantium*, MLST, recombination, genetic diversity, genetic population structure

INTRODUCTION

Ehrlichia ruminantium is an intracellular bacterium responsible for heartwater, an important and fatal tropical disease of wild and domestic ruminants (Allsopp, 2009; Moumene and Meyer, 2016). This bacterium is transmitted by *Amblyomma hebraeum* ticks in southern Africa and by *Amblyomma variegatum* ticks, the most wide spread vector through sub-Saharan Africa, Indian Ocean islands and the Caribbean (Walker and Olwage, 1987).

Heartwater occurs in almost the whole of sub-Saharan Africa (except for the very dry southwest), in São Tomé and Principe and in the Indian Ocean islands of Madagascar, Zanzibar, Mayotte, Mauritius, Reunion islands, and the Comoros (Provost and Bezuidenhout, 1987; Stachurski et al., 2013). The disease is also present in the three Caribbean islands of Guadeloupe, Marie Galante, and Antigua and it represents a threat to the American mainland (Barré et al., 1987; Roth et al., 2013).

As of yet, no efficient single vaccine against heartwater is available due to a limited cross protection between vaccinal and field strains, which is probably caused by the high genetic diversity of *E. ruminantium* in any given geographical location (Vachiéry et al., 2013).

In order to better define appropriate control strategies against heartwater, it is important to understand the diversity of *E. ruminantium* strains within regions, their evolution and origin of introduction as well as to attempt to associate *E. ruminantium* genotypes with possible protective genetic markers. Presently, the genetic diversity of *E. ruminantium* has been elucidated through polymorphic and conserved genes such as *map1, 16S* rRNA, and some housekeeping genes for a limited number of strains.

Allsopp and Allsopp (2007) studied the genetic diversity of 12 different *E. ruminantium* strains from Africa and the Caribbean using a panel of core function and housekeeping genes: *16S rRNA*, *gltA*, *groEL*, *ftsZ*, *sodB*, *nuoB*, *rnc* (*pCS20*), and *ctaG* (*pCS20*). This study highlighted inconsistent phylogenies for the different genes examined and evidence for recombination and a separation of strains into two clades, South-East Africa, and West Africa (Allsopp and Allsopp, 2007).

Using *map1* and *map1* family genes to genotype *E. ruminantium* strains, several authors showed a high genetic diversity at local (Gambia and Burkina Faso), regional and worldwide scales (Caribbean region, Africa, and Madagascar; Faburay et al., 2008; Vachiéry et al., 2008; Adakal et al., 2010b; Raliniaina et al., 2010). In localized areas of Gambia and Burkina Faso, the studies observed at least 11 sequence types and mixed infections with several *E. ruminantium* strains in ruminants and ticks. In the Caribbean, another study using the *map1* gene demonstrated a high diversity of *E. ruminantium* strains with nine different genotypes present either at the scale of locality, islands and region (Vachiéry et al., 2008). Furthermore, comparison of strains isolated in Africa, in the Caribbean and in Madagascar using the *map1* gene family, revealed a divergent evolution for this gene in *E. ruminantium* (Raliniaina et al., 2010). Divergent evolution was shown by the identification of different genotypes for each *E. ruminantium* strain depending on the *map1* paralogs used. Furthermore, important *map1* diversity was conserved independently of sampling scale (village, region, continent) and the timing of introduction (punctual for Caribbean strains vs. continuous introduction for African strains).

In general, *map1* gene appeared to be a good tool to characterize the genetic diversity among Africa, Caribbean and Madagascar, however, there was a lack of correlation between *map1* genotype and geographic origin (Raliniaina et al., 2010) and limited cross protection between strains (Adakal et al., 2010b).

In order to construct *E. ruminantium* phylogeny and elucidate genetic population structure, a multilocus sequence typing (MLST) analysis based on housekeeping genes (Adakal et al., 2009, 2010a; Nakao et al., 2011) and a multilocus variable number of tandem repeats based on mini-satellites were developed for *E. ruminantium* (Pilet et al., 2012). The apparent advantages of these techniques are based on a perceived lack of ambiguity, portability, reproducibility, good discriminatory powers to differentiate isolates and ability to be automated (Lindstedt, 2005; Sullivan et al., 2005). MLST can overcome certain challenges to phylogenetic and phylogeographic reconstruction and potentially allows the study of the diversity of the bacterium, which could aid in the design of efficient vaccines that includes appropriate local and/or regional strains (Urwin and Maiden, 2003; Lindstedt, 2005; Sullivan et al., 2005). However, inference of bacterial evolution and population genetics can be complicated by the presence of recombination.

Recombination is an exchange of genetic material between organisms or chromosomes to form new combinations of genetic material on a chromosome (Smith et al., 1993; Martin and Beiko, 2010). The unaccounted for presence of recombination can lead to an overestimation of population expansion, the false detection of positive selection (Schierup and Hein, 2000; Shriner et al., 2003) and to the reconstruction of an erroneous phylogeny (Posada and Crandall, 2002; Ruths and Nakhleh, 2005). Recombination can be detected by several means, including a lack of congruence in phylogenetic trees (Feil et al., 1996; Zhou et al., 1997), mosaic-like DNA sequences (Spratt et al., 1995), excess of homoplasy in phylogenetic trees (Smith and Smith, 1998) and a network relationship between sequences, using split decomposition (Holmes et al., 1999).

Recombination occurrence has also been described for *E. ruminantium* in studies done by Allsopp and Allsopp (2007), Bekker et al. (2005), Hughes and French (2007), and Nakao et al. (2011). Up to now, the implications of recombination for *E. ruminantium* as a major driver of its genetic diversity have been given little attention. Moreover, high genetic diversity is probably tightly associated with limited cross-protection between vaccinal and some field strains (Jongejan et al., 1991; Allsopp and Allsopp, 2007; Vachiéry et al., 2013). This omission could be important in the interpretation of the results of previous studies attempting to reconstruct the relationship and population structure between *E. ruminantium* strains.

In order to understand the genetic diversity and population structure of 194 worldwide *E. ruminantium* isolates, MLST analysis using *lipA*, *sucA*, *sodB*, *secY,* and *lipB* core function and housekeeping genes was performed. The presence of

recombination with the presence of reticulations in a split-decomposition network and tests of recombination within and between different subgroups in the network using the pairwise homoplasy index (PHI) test as well as incongruence in gene trees for these five genes was demonstrated. These results show the relevance of recombination events in the generation of *E. ruminantium* diversity and evolution. Two major genetic groups, a West African cluster and a worldwide cluster which includes West Africa, East Africa, Southern Africa, Indian Ocean, and Caribbean, could be delineated by MLST. The prevalent recombination events observed in the current study suggest a complex population structure for *E. ruminantium* strains and suggests that caution should be taken when reconstructing relationships in closely related species in the *Anaplasmataceae*, especially if using only a subset of the genome, like in MLST analyses.

MATERIALS AND METHODS

Field Collection of Ticks and Blood Samples

A. hebraeum and *A. variegatum* ticks were collected from cattle during epidemiological studies in the south and center of Mozambique from 2011 to 2013 and Indian Ocean Islands (Reunion Island, Comoros Islands, Mayotte, and Madagascar) from 2007 to 2010 and chosen randomly for each animal sampled. Blood samples of heartwater clinical cases from Guadeloupe were provided to CIRAD by the surveillance network monitoring ruminant neurological syndromes (RESPANG) with the collaboration of the veterinary services. Samples from RESPANG were collected following an ethical committee approval from 2011 to 2014. Ticks were stored in 70% ethanol at room temperature and taxonomically identified as being *A. variegatum* and *A. hebraeum*. Blood was preserved either frozen at −20°C or in 70% ethanol at room temperature until DNA extraction. All samples were tested for *E. ruminantium* positivity using *pCS20* nested PCR as described below.

pCS20 Nested PCR and Multilocus Sequence Typing

DNA was extracted from tick tissues and blood using the QiaAmp DNA minikit (Qiagen, Courtaboeuf, France) in accordance with manufacturer's instructions. Detection of *E. ruminantium* was performed using a semi-nested PCR for a fragment of the *pCS20* gene, as previously described by Molia et al. (2008). For the *pCS20* nested PCR, a first PCR phase was performed using the primer pair AB128′/AB130′ and for the second phase the primer pair used was AB129′/AB130′. Amplified products were visualized on a 1.5% agarose gel (TAE buffer) and considered positive if exhibiting a band of 280 bp. Positive samples were then typed using a modified Multilocus sequence typing scheme which includes a panel of five genes instead of eight. Five variable housekeeping genes, *lipA*, *lipB*, *secY*, *sodB,* and *sucA* were amplified using the primers and PCR conditions previously described by (Adakal et al., 2009, 2010a). PCR products were sequenced by Beckman Coulter Genomics (France).

E. ruminantium Isolates

A total of 194 isolates originating from several geographical areas were analyzed by MLST: North and Central Africa (2), West Africa (55), Eastern Africa (3), Southern Africa (64—mainly from Mozambique), Indian Ocean Islands (29), and Caribbean (41—mainly from Guadeloupe; Tables S1, S2). The geographical origin of strains and isolates, their number and name, reference and date of isolation are shown in Table S2. Previously published DNA sequences from Burkina Faso, Caribbean, and Madagascar extracted from tick tissue, blood and cell culture as well as some reference strains were also used (Adakal et al., 2010a; Raliniaina et al., 2010; Nakao et al., 2011). The number of isolates per country is shown in Table S1. The main sampling areas were Mozambique (58 isolates, 30% of sampling) for Southern Africa, Burkina Faso (44 isolates, 23% of sampling) for Western Africa, Guadeloupe (40 isolates, 21% of sampling) for the Caribbean and Madagascar (13 isolates, 7% of sampling) for the Indian Ocean.

Data Analysis

All the Figures were edited using Inkscape software (Harrington, 2004–2005). Multiple sequences were aligned and edited to produce a consensus and concatenated sequence for each strain using the software Geneious 8.1.7. R8 (Kearse et al., 2012) and AliView 1.17.1 (Larsson, 2014). Sequences were concatenated in the order *sucA-sodB-lipA-secY-lipB* resulting in a final sequence of 2314 bp length. If a set of sequences were identical for some isolates (clones), only one representative sample of each group was included in the bioinformatics analysis. Bayesian trees were generated with BEAST (Drummond et al., 2012). The dataset was partitioned into one partition per gene alignment and the trees were linked over the five partitions. The site model was set to the "BEAST Model Test," testing all reversible models and estimating the mutation rate (initial value was set to 0.001). A strict clock and an exponential population growth coalescent model were used. The Markov Chain Monte Carlo was run for 10,000,000 generations on two occasions to ensure correct mixing, and the resulting log files was reviewed in Tracer (Rambaut et al., 2014). The 95% Highest Posterior Density (HPD) for the growth rate did not contain 0, allowing the rejection of constant population growth. The Effective Sample Size (ESS) for four of the five tree likelihoods was less than 100, and examination of the traces revealed that they did not converge properly. For visualization of the trees with Densitree (Bouckaert, 2010), 25% of samples were removed from burn-in. The resulting image of 7.5 million overlaid trees and the "root canal" tree which represents the total set is shown in **Figure 2**.

The alignment of concatenated sequences from the five genes (*sucA*, *sodB*, *lipA*, *secY,* and *lipB)* was imported to SplitsTree4 program version 4.13 and a phylogenetic network was constructed using the neighbor-net algorithm (Huson and Bryant, 2006).

Maximum likelihood phylogenies were constructed using a concatenated alignment of the five genes with PhyML (Guindon et al., 2010) with and without *E. chaffeensis* as an outgroup under the GTR+G+I model of evolution with four rate categories

and aLRT (approximate Likelihood-Ratio Test)-based branch support.

Determination of genetic groups in the network was done by combined inspection of the multiple sequence alignment, split decomposition network and a heat map, which represents the degree of relatedness between strains (genetic distance). Heat maps allow the visualization of similarity and differences in the data by representing values contained in the distant matrix as colors in a graph. The heatmap was used to identify possible recombinants by finding sequence types with atypical similarity to others from outside their group. The heat map was generated using the heatmap tool, "heatmap.2" of the "gplots" (Warnes et al., 2009) package of the R software based on a distance matrix calculated using the Tamura and Nei (1993) model from the "ape" (Paradis et al., 2004) package (R Core Team, 2013). Dendograms representing hierarchical clusters of sequence types and their sequence type numbers are shown along the axes of the heatmap.

Additionally a population structure analysis was performed using STRUCTURE (Pritchard et al., 2000). The analysis was performed with an admixture model with correlated genotypes for $K = 2$ to $K = 9$, with burnin period of 10,000 MCMC generations followed by 100,000 MCMC generations. For each K, 50 iterations were performed, and the optimal K was calculated using the Evanno method (Evanno et al., 2005) on the STRUCTURE Harvester webserver (Dent and von Holdt, 2012). CLUMP (Jakobsson and Rosenberg, 2007) was used to find the optimal clustering from the multiple iterations.

The (Pairwise Homoplasy Index) PHI test (Bruen et al., 2006) was conducted to determine whether recombination events were present in the concatenated sequence alignment. A PHI test was also performed within and between determined groups and *p*-values for each comparison are presented in **Table 1**. The *p*-values were corrected for multiple testing using Benjamini-Hochberg FDR method (Benjamini and Hochberg, 1995; available online at: http://www.sdmproject.com/utilities/?show=FDR). Recombination within and between genetic groups was considered positive if the PHI test yielded a corrected $p \leq 0.05$.

ClonalFrame (Didelot and Falush, 2007) was used to estimate the ratio of recombination to mutation (r/m).

TABLE 1 | Phi test for recombination from two genetic groups (1 and 2) and five subgroups (2A, 2B, 2C, 2D, and 2E) of *E. ruminantium*.

Genetic group	Corrected *p*-values
1	1.00
2	0.00*
2A	N/A
2B	1.00
2C	0.83
2D	0.03*
2E	0.06

*N/A, not applicable because all the genotypes are recombinants; *Positive for recombination, PHI test yielded a $p \leq 0.05$.*

RESULTS

E. ruminantium Population Genetic Structure

DNA sequences were deposited in GeneBank and accession numbers are displayed in Table S4.

Concatenated sequences are based on five housekeeping genes. From 194 *E. ruminantium* isolates, only 97 representing unique sequence types are shown in the phylogenetic network and tree (**Figures 1A**, **2**, Figure S1). The remaining 97 *E. ruminantium* isolates were clonal to these sequence types for the genes examined. Several isolates, representing one sequence type from the same geographical origin, are shown by a circle with a single color, whereas clonal isolates from different geographical origins were divided into several colors within one circle (**Figures 1A**, **2**).

A split decomposition network was constructed from 97 concatenated *E. ruminantium* sequences and illustrated in **Figure 1A**. The split decomposition network resembled a network-like structure composed of two distinct and divergent groups: Group 1 (West Africa) and a Group 2 (worldwide) with West/East/Southern Africa, Indian Ocean, and Caribbean strains (**Figure 1A**). Reticulations in the graph are clearly present (especially noticeable in Group 2), and are evidence for recombination or other forms of homoplasy between different sequence types. Group 2 is composed of five subgroups, which were defined based on arrangement of sequence types in the network (**Figure 1A**), a hierarchical clustering (heat map: Figure S2), and careful examination of the multiple sequence alignment (Figure S3). Several sequence types were excluded from groups due to being recombinants identified by examination of the heat map and multiple sequence alignment (marked with blue stars in **Figure 1A**, Figure S3). In the STRUCTURE analysis (**Figure 3**) the optimal *K*-value was found to be 2, which is not unexpected given the long branch separating Group 1 and Group 2. When discounting $K = 2$, the next identified optimal grouping is $K = 5$, which is one less group than we defined using the other methods. The vast majority of sequence types are placed in equivalent groups between the two analyses, however the STRUCTURE analysis combines subgroup 2A and subgroup 2B, and clusters a few other strains in with those from different groups. In these cases, the strains tend to have a large proportion of inferred admixture (in this case equating to possible recombination), for example, ST 32 and ST 1 appear to show slightly more ancestry with the group equating to subgroup 2E instead of 2D, ST 69 from subgroup 2A clusters with subgroup 2E sequences, however both have evidence of ancestry from both subgroups.

The histogram in **Figure 1B** represents the total number of *E. ruminantium* isolates including unique sequence types and clones and only clones, per geographical origin belonging to Group 1 and Group 2. Group 1 was represented mostly by West African strains (32 isolates; **Figures 1A,B**). Twenty-five isolates were from Burkina Faso, three from Senegal, two from Ghana and one isolate from Gambia, Guadeloupe, Nigeria, Tanzania, and Mozambique (Table S3). Sequence type 97 included 17 isolates

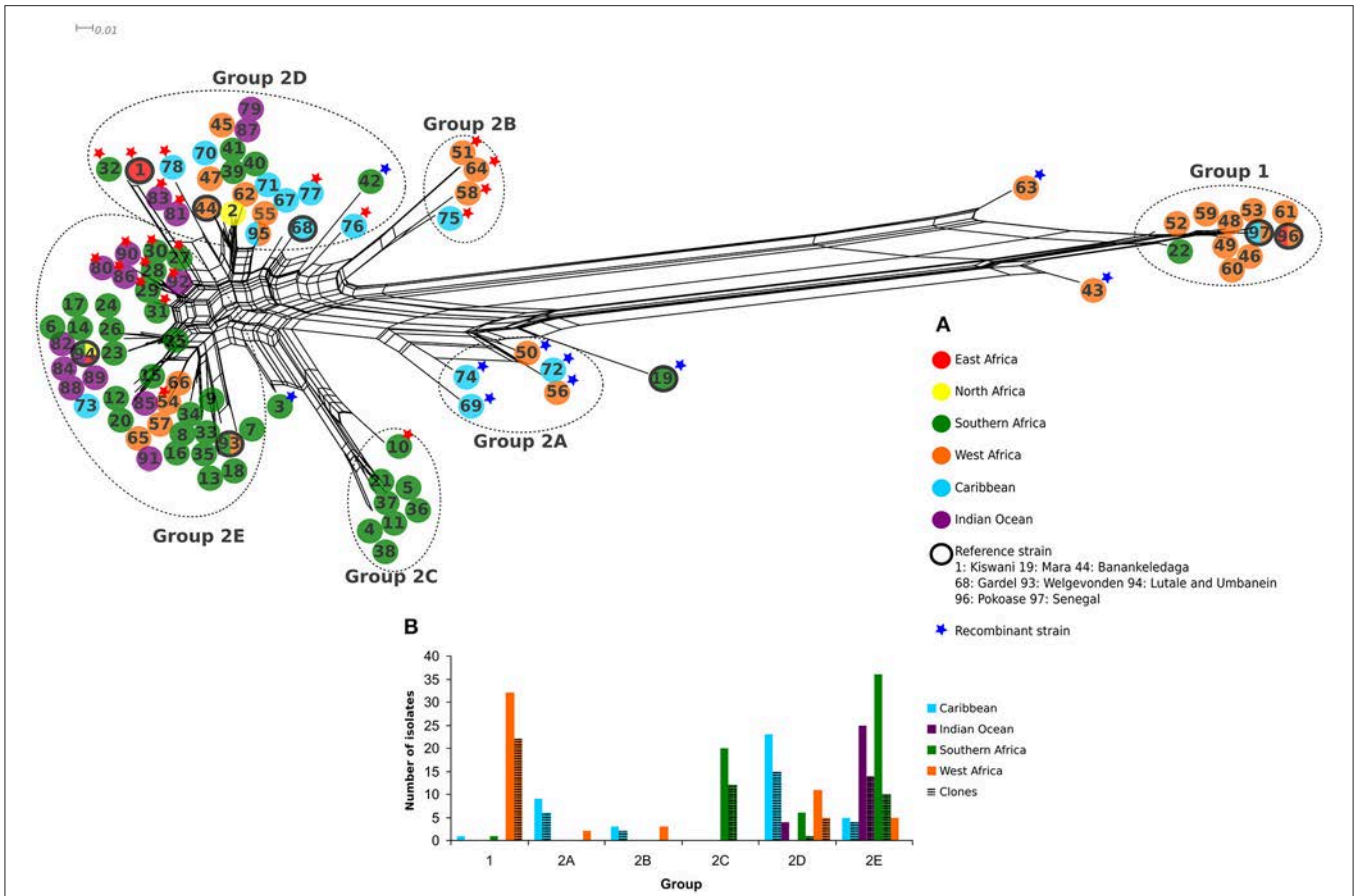


FIGURE 1 | (A) Split decomposition network of 97 *E. ruminantium* genotypes obtained from five concatenated housekeeping genes. Strains are color coded according to the geographic origin described in the legend. Reference strains are identified by black circles. Recombinant strains are tagged with a blue star, and those with less than 80% inferred ancestry from any single population in STRUCTURE with $K = 5$ are tagged with a red star. All recombinants also had less than 80% ancestry from any given population. Two main genetic groups and five subgroups are described as Group 1, 2A, 2B, 2C, 2D, and 2E. **(B)** Histogram representing the total number of *E. ruminantium* isolates sampled per geographical origin for each genetic group.

from West Africa (with the reference strain Senegal) and one from the Caribbean whereas sequence type 96 included one isolate from East Africa and the reference strain Pokoase from West Africa.

The Group 2 (worldwide) was diverse and composed of 82 unique sequence types from all the sampled geographical areas (**Figure 1A**). All of these subgroups with the exception of subgroup 2D have more frequent representation of certain geographical regions in our data. Subgroups 2A and 2B were composed of isolates from West Africa and Caribbean. Group 2A is present in Burkina Faso, Senegal, and Guadeloupe, with Guadeloupe containing most of the isolates (**Figures 1B**, **4**, Table S3). Group 2B is represented in Burkina Faso and Guadeloupe, with equal number of isolates for each group (**Figures 1B**, **4**, Table S3). Subgroup 2C was exclusively composed of Southern African isolates (20 strains), the majority being from Mozambique and one, Crystal Spring strain, from Zimbabwe (**Figures 1B**, **4**, Table S3). Subgroups 2D is the most geographically diverse subgroup at a regional scale with representatives from each large scale region considered. The majority of the strains in this subgroup are from the Caribbean (23 strains) and West Africa (11 strains; **Figure 1B**). Subgroup 2D is represented in 10 countries (Antigua, Burkina Faso, Chad, Comoros, Guadeloupe, Kenya, Madagascar, Mozambique, Nigeria, South Africa; **Figure 4**, Table S3). Sequence type 95 (subgroup 2D) includes two strains from West Africa and four from the Caribbean (**Figure 1**, Table S2). The reference strains Kiswani (Kenya), Banankeledaga (Burkina Faso), and Gardel (Guadeloupe) are part of subgroup 2D. Upon close examination of the alignment, some of the sequence types from Southern Africa (39, 40, and 41) and the Indian Ocean (79 and 87) are differentiated simply by gaps caused by lack of coverage at the end of certain gene sequences or by one residue unique to a sequence, likely increasing the representation of their regions in the group in terms of unique sequence types (Figure S3). Strains from West Africa and the Caribbean tend to be more genetically diverse within this group. Notable exceptions to the lack of diversity in Southern African sequence types in this subgroup are sequence type 42, which is recombinant as noted below, and sequence type 32, which is nearly identical to the East

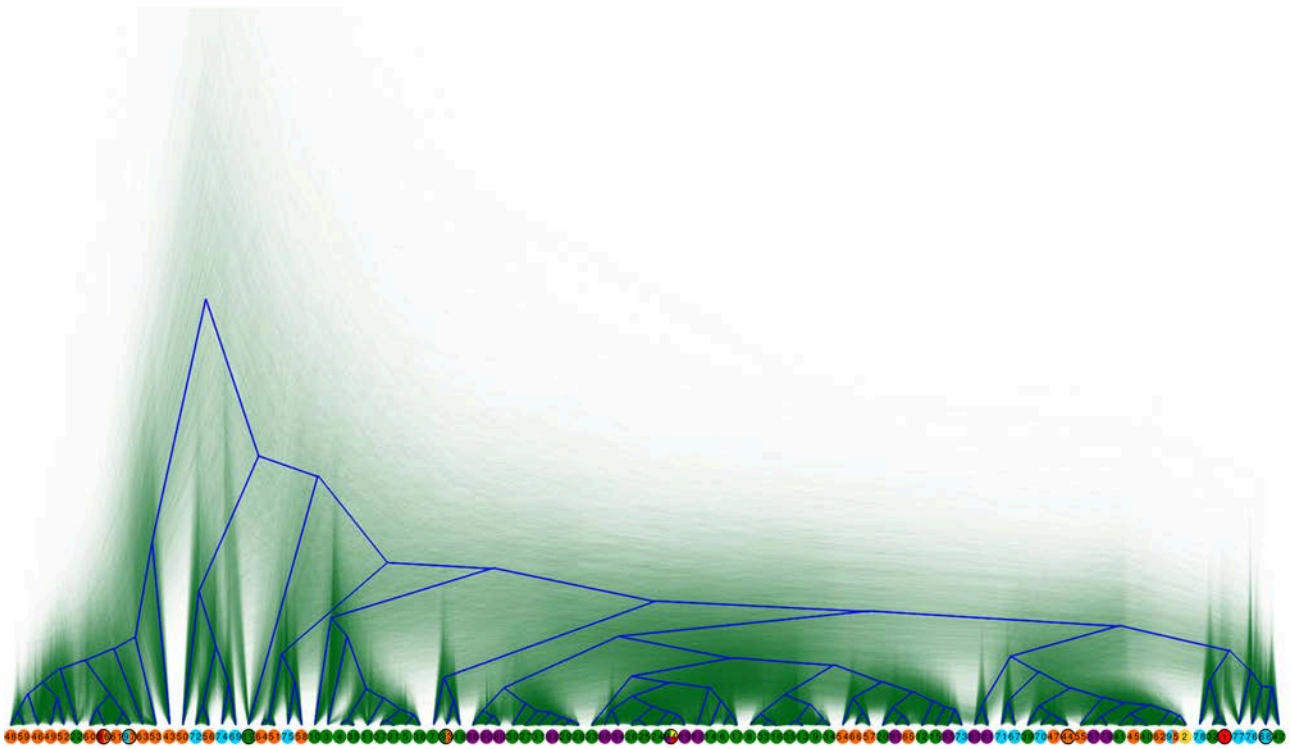


FIGURE 2 | Overlaid Bayesian consensus trees of 97 *E. ruminantium* genotypes generated using BEAST and visualized in Densitree. The root canal tree is shown in blue. Sequence types at tips are color coded according to the geographic origin: East Africa (red), North Africa (yellow), Southern Africa (green), West Africa (orange), Caribbean (blue), Indian Ocean (purple). Reference strains are identified by black circle outlines.

African sequence type 1 (Kenya), and shares its *lip*B gene with West African and Caribbean subgroup 2B types 51 and 64. They also share a *sucA* allele (that is otherwise exclusively found in subgroup 2D) with two Indian Ocean sequence types 83 and 81 in this subgroup.

Subgroup 2E is dominated by Southern African (36 strains), mainly Mozambique, and Indian Ocean isolates (25 strains) with a small number of other sequence types from West African or mixed origin. This subgroup is present in 13 countries (Burkina Faso, Cameroon, Comoros, Guadeloupe, Madagascar, Mayotte, Mozambique, Reunion, São Tome e Principe, South Africa, Sudan, Uganda, Zambia; (**Figures 1B**, **4**, Table S3). Sequence type 93 (subgroup 2E) had one isolate from West Africa and seven from Southern Africa (**Figure 1**, Table S2). Sequence type 94 (subgroup 2E) contains one isolate each from North Africa, East Africa and Southern Africa and two isolates from the Indian Ocean. The reference strains Welgevonden (South Africa), Lutale (Zambia) and Umbaneim (Sudan) are part of the subgroup 2E (**Figure 1**, Table S2).

Two sequence types from Burkina Faso (43, 63) and two sequence types from South Africa (3, 19 corresponding to reference strain Mara) were not part of any genetic group as they are recombinant strains (**Figure 1A**). Recombination was identified based on the split decomposition network location (usually at the vertices), and examination of the multiple sequences alignment and the heat map for close similarity between the given sequence type and a different subgroup (Figure S2). Several other sequence types (50, 56, 69, 72, and 74) from subgroup 2A and sequence type 42 from subgroup 2D tagged with a blue star in **Figure 1A** were also identified as recombinants based on combined inspection of the multiple sequence alignment, heat map and split decomposition network (**Figure 1A**, Figures S2, S3).

The distribution of *E. ruminantium* genetic groups in the sampled countries as well as the recent (less than 400 years ago) and ancient (more than 400 years ago) cattle movement (domestication) in Africa, the Caribbean an Indian Ocean Islands are shown in **Figure 4**, Table S3.

Evidence of Recombination Events for *E. ruminantium*

A Bayesian MCMC phylogeny created with BEAST using the five gene partitions with the 97 nucleotide sequences in each partition highlighted conflicting topologies (**Figure 2**). Reviewing the trace files revealed that the ESS of the four linked trees were all less than 100 and the trace plots reveal that they did not converge properly. This indicated conflicting signals in our data, which are also illustrated by low branch support values in the maximum likelihood trees (Figure S1). In phylogenetics, low branch support can be a quantitative indication of recombination or other forms of homoplasy in the data and is also common for short branches. In this case

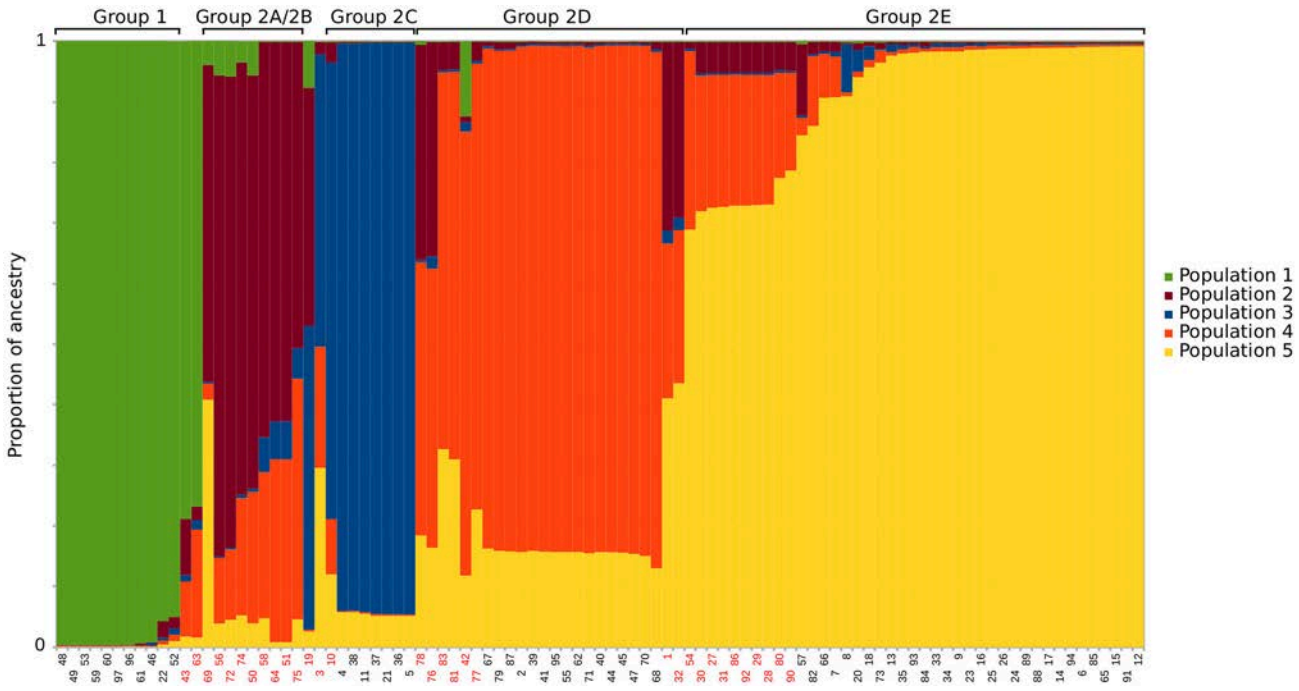


FIGURE 3 | STRUCTURE analysis for $K = 5$ using an admixture model and correlated genotypes. Each column represents the inferred ancestry for each sequence type from the five predicted populations. The groups from the hierarchical clustering and splitstree analysis are noted above the columns. Sequence types with less than 80% inferred ancestry from a single population have red labels.

the conflicting signals are most likely due to recombination events between the analyzed strains given the evidence of mosaic genes in the alignment, the reticulate structure in the split decomposition network (**Figure 1**) and the mixed ancestry exhibited in the STRUCTURE analysis (**Figure 3**). The overlaying of all the trees produced by the BEAST analysis after discarding for burn-in reveals fuzzy regions in the tree which indicate conflicts in the reconstruction. *E. ruminantium* sequence types are widespread and the tree did not show a clear phylogeographic structure between Africa, Indian Ocean islands and the Caribbean, although several isolates from the same or close geographical origin cluster together as described above (**Figures 1**, **2**).The ratio of recombination to mutation was estimated as 1.351 ± 0.0167, (any given residue difference between two sequences, is 1.351 times more likely to have been introduced by recombination than by mutation). The PHI test on 194 *E. ruminantium* isolates found statistically significant evidence for recombination ($p = 0.0$). Furthermore, recombination analysis between the six genetic groups and subgroups was mostly positive and only intragroup recombination was represented in **Table 1**. Examination of the alignment reveals clear mosaicism (blocks of multiple residues shared with a different group to the exclusion of other members of the same group) in sequence types 63, 43, 19, 42, 3 and all strains in subgroup 2A (Figure S3). These are the most obvious events because they are between regions that are divergent that contain several mutations delineating the recombination breakpoints. The strains in general do not contain many variable sites in the genes used, and there are many possible cases where recombination might transfer single SNPs. In such cases it is difficult to identify the events with certainty because the possibility of homoplasy by convergent mutation exists. The PHI test does not identify recombination breakpoints, but does allow a statistical evaluation of the likelihood of recombination between different sequences regardless of the size of the recombined blocks. This type of test is advantageous in cases where the sequence similarity is high and there are only a small number of discriminatory SNPs to infer recombination compared to methods that attempt to identify breakpoints at the edges of recombination blocks because it may detect the overall signal of recombination from a series of small events that are undetectable individually.

The STRUCTURE analysis (**Figure 3**) also hints at recombination in the groups, signified by the presence of many sequence types that are inferred to have ancestry from more than one different population in the $K = 5$ analysis. In all, 50 of the 97 sequence types have inferred ancestry of less than 90% from any one population, while 32 have less than 80% and five sequence types have less than 50% inferred ancestry from any single population. We have marked sequence types that have less than 80% inferred ancestry from a single population with red stars in **Figure 1A**, Figure S3. Although STRUCTURE doesn't provide a statistical measure for the likelihood of recombination between sequence types, the analysis does suggest widespread recombination between the five different inferred populations.

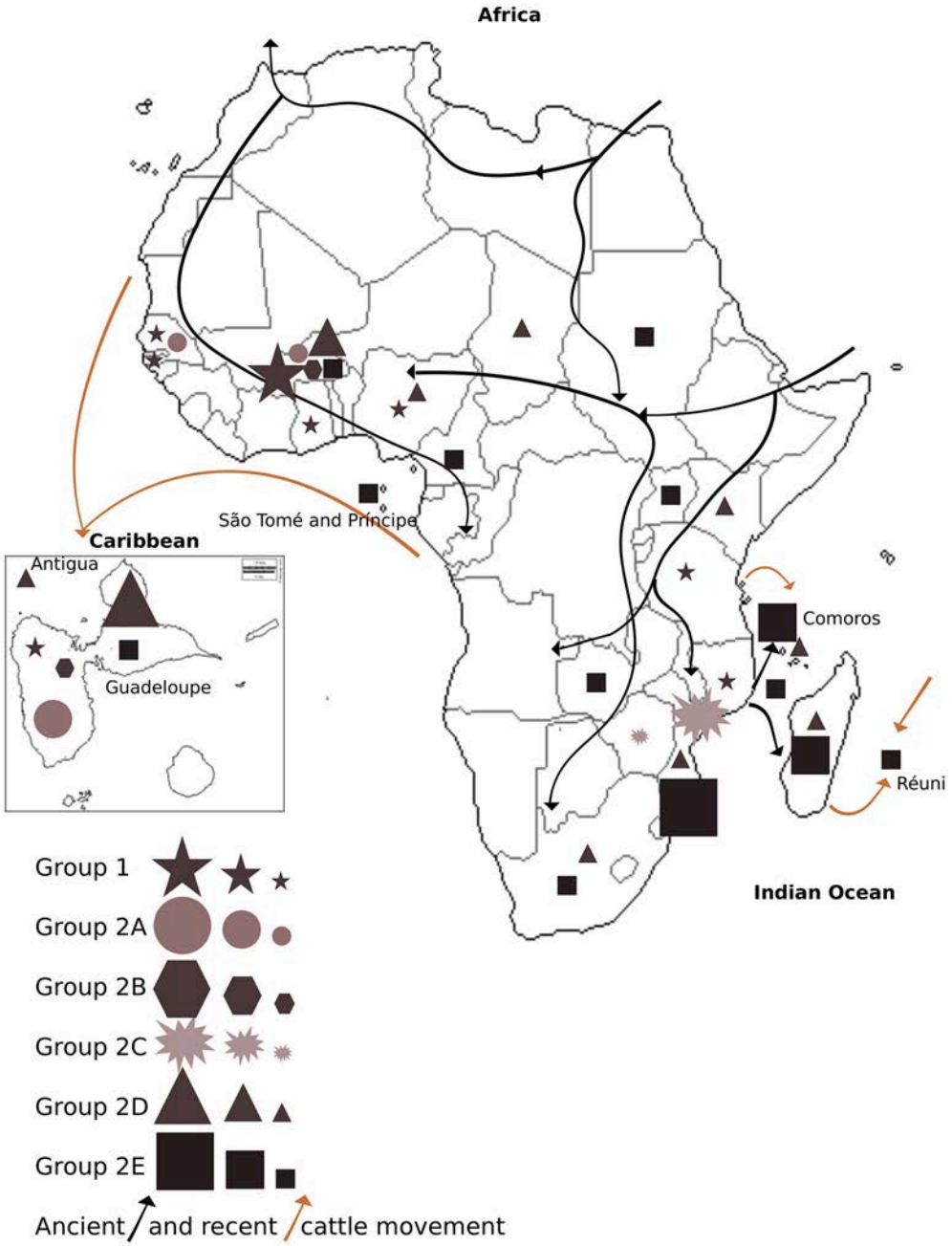


FIGURE 4 | Distribution of *E. ruminantium* genetic groups and subgroups 1, 2A, 2B, 2C, 2D, and 2E in each sampled country within Africa, Caribbean, and Indian Ocean Islands. Groups are coded by symbols according to the legend. Symbol size corresponds to sampling size defined by the following sample threshold: >15 samples (big symbol), <15 samples (medium symbol), and <5 samples (small symbol). Recent (<400 years ago; brown arrows) and ancient (>400 years ago; black arrows) movement of cattle is represented in the map.

DISCUSSION

E. ruminantium Phylogeny and Genetic Population Structure

Despite the apparent unclear population structure and phylogeography of *E. ruminantium* strains, two main distinct genetic groups were defined: Group 1 (West Africa) and a worldwide Group 2, the latter being represented by West, East and Southern Africa, Indian Ocean and Caribbean isolates. Outgrouping the tree using *E. chaffeensis* (Figure S1A) reveals that the root of the *E. ruminantium* species divides Group 1 from Group 2. The branches leading to each group are relatively long compared to the branches within the groups, allowing us to infer that the groups were probably isolated from one another for a

long time. We also infer the possible geographical locations of the ancient progenitors of the groups based on the geographic distributions of the extant strains: we propose that the Group 1 is ancestrally West African, and that Group 2 originates elsewhere. The divergence of the two main groups possibly predates the spread of livestock across the continent of Africa with the wave of domestication. The presence of several individual sequence types and all sequence types present in subgroup 2A that represent recombination between Group 1 and Group 2 in West Africa, shows the propensity of these groups to recombine. This allows us to infer a later arrival of subgroup 2D in West Africa long after the split between the two groups otherwise the clusters would be more homogeneous.

Influence of Wildlife, Ticks, and Ancient Cattle Movement on *E. ruminantium* Population Structure

Apart from *Anaplasma marginale* (de la Fuente et al., 2007; Estrada-Pena et al., 2009), no other studies have yet elucidated the influence of wildlife, cattle, and small ruminant movements on the genetic diversity of the *Anaplasmataceae* family.

We hypothesize that ancient (more than 400 years ago) and recent cattle (less than 400 years ago) movements has shaped *E. ruminantium* diversity by creating different genetic groups and subgroups as well as facilitated recombination between diverse strains. Recent and ancient cattle movement (domestication) in Africa, the Caribbean, and the Indian Ocean islands are illustrated in **Figure 4**. Group 1 and subgroups 2A, 2B, 2D, and 2E present in West Africa are also present in Guadeloupe and Antigua, reflecting the historical movement of cattle from this region to the Caribbean (Maillard et al., 1993). In addition, Indian Ocean strains from subgroup 2E, reflect the ancient movement of cattle from East and Southern Africa to Islands in the Indian Ocean. Subgroups 2D and 2E in Africa cover most of the sampled countries and might reflect the ancient movement of cattle from North and Central-East to West Africa and from East Africa to Southern Africa. Subgroup 2E is predominantly present in Southern Africa and the Indian Ocean, suggesting that the small number of strain types from this group that are present in other locations may be the result of recent cattle movement.

Given that all the current known vectors of *E. ruminantium* are African *Amblyomma* species (Walker and Olwage, 1987) and that these species have as primary hosts large African wild mammals (Voltzit and Keirans, 2003), it is reasonable to infer that well adapted *E. ruminantium* strains were circulating in Africa for thousands of years, before the introduction of the first domestic ruminants. Several studies have reported a South and East genetic break in African mammals (Pitra et al., 2002; Lorenzen et al., 2010) that can possibly help to explain the large genetic distance between Group 1 and Group 2 seen in *E. ruminantium* population. Because of limited sampling in some regions (particularly Central and East Africa) and the lack of a well-established molecular clock for *E. ruminantium* (Hughes and French, 2007), as well as the presence of recombination it is not possible to get information on the time of divergence between *E. ruminantium* strains and genetic groups.

The introduction and expansion of *Bos taurus*, small ruminants and *Bos indicus* in Africa between 8000 and 2500 years ago (Hanotte et al., 2002; Ajmone-Marsan et al., 2010) exposed naïve domesticated ruminant populations for the first time to African *Amblyomma* species and *E. ruminantium* strains. The spill over of *E. ruminantium* strains from African wildlife species to domestic ruminants was arguably associated with high mortality and triggered a long process of host-pathogen adaptation that putatively included intense recombination among *E. ruminantium* strains. We propose that this introduction and spread of domesticated animals are the main trigger for recombination between Group 2 (subgroup 2D) and Group 1.

Archaeological and genetic data on the origin and expansion of domestic ruminant populations in the African continent strongly supports the existence of migratory routes East-West, West-East (less common), North-South and more recently South-North (Hanotte et al., 2002). This domesticated host migratory pattern would support the possibility of pathogen gene flow across the above mentioned migratory routes. In the last millenium (more recently) the development of agriculture and iron technology led to a more sedentary behavior of sub-Saharan Bantu communities with the consequent progressive reduction of trans-continental migrations (Blench and MacDonald, 2000). This fact might have contributed to the gradual isolation of domestic ruminant populations and consequently a reduction in the continental gene flow of vectors and pathogens.

Cattle movement, contact with wildlife coupled with *Amblyomma* tick dispersion, have probably shaped the genetic diversity of *E. ruminantium*. However, the influence of vector species on the strain diversity is unknown. It is worth noting that the genetic diversity of the tick vector *A. variegatum* seems to follow a similar pattern that those of *E. ruminantium*.

For instance, Beati et al. (2012) studied the genetic diversity of *A. variegatum* throughout the Caribbean and Africa and found a West Africa-Caribbean clade and an Eastern Africa clade. Later, Stachurski et al. (2013) found a worldwide clade and an East Africa-Indian Ocean clade by studying the phylogeographic structure of *A. variegatum* in the Indian Ocean. *A. variegatum* genetic clusters seem to generally match *E. ruminantium* genetic groups found in the current and previous studies of Allsopp and Allsopp (2007), Vachiéry et al. (2008), and Nakao et al. (2011). Even if there are similarities between *A. variegatum* and *E. ruminantium* clusters for Caribbean/West Africa isolates and Indian Ocean/Southern Africa and South/Eastern African isolates, the phylogeography of *Amblyomma* ticks appears different from that of *E. ruminantium.*

In the Caribbean and Indian Ocean, it is very likely that *A. variegatum* and *E. ruminantium* were introduced simultaneously with cattle. In the Caribbean, studies on cattle genetic diversity found a correspondence between genetic diversity and historical domestication (importation from West Africa) of cattle from this region (Magee et al., 2002). Also, the genetic diversity of *E. ruminantium* strains (Vachiéry et al., 2008 and our results) linked the sequence types with several origins in Africa, thus suggesting that strains were introduced with *A. variegatum* in the nineteenth century.

Influence of Recent Cattle Movement on *E. ruminantium* Population Structure

In cases where a subgroup is dominated by certain geographical regions but is rarely found elsewhere, the parsimonious explanation is that the atypically located strains are the result of recent transport of a strain whose ancestry lies in the region that is predominant in the group, barring sampling biases. As we have large samples from Mozambique, Burkina Faso and Guadeloupe, the origins of subgroup 2D appear puzzling because the frequency of apparently unique sequence types from the four major geographically sampled regions are quite similar.

However, when taking into account the nearly clonal (39, 40, 41 and 79, 87) and recombinant (42 and 32) sequence types from Southern Africa and the Indian Ocean, the majority of the genetic diversity in this subgroup is found in West Africa and the Caribbean. Because livestock transport was likely unidirectional to the Caribbean from West Africa (Hanotte et al., 2002), we can consider that subgroup 2D is thus probably ancestrally West African and the presence of Southern African and Indian Ocean strains within this group may be due to recent movement of the strain types from West Africa to those regions.

The predominance of unique sequence types in subgroup 2D from West Africa/Caribbean suggests that Group 2 may have spread to that region long after the initial divergence of the two groups, but before recent transfer of *E. ruminantium* to the region. We propose that this transfer was contemporary with the ancient wave of livestock domestication.

The location of clearly recombinant strains can provide some information on their recent movement, because recombination requires direct contact between strains (although these could also be transported to new locations after recombination), which in *E. ruminantium* necessarily must take place by coinfection of either host or vector in the same location. For example, the atypical presence of the Southern African sequence type 22 in Group 1 suggests recent transport of the strain from West Africa to Southern Africa. The Southern African sequence type 19 which is recombinant between Group 1 and Group 2 also suggests movement from West Africa to Southern Africa. Most of the other clear mosaic sequences that represent recombination between Group 1 and Group 2 (subgroup 2A and strain types 63 and 43) are found in West Africa or the Caribbean, suggesting that the events occurred in West Africa. As it appears likely that subgroup 2D has its origins in West Africa, the best explanation for the presence of these recombinants is the mixing of Group 1 and Group 2 strains after the arrival of Group 2 in West Africa. Indeed the Group 2-like sequences in subgroup 2A tend to be most similar to those from subgroup 2D, especially the secY allele, supporting this assertion. Strain type 69 is an exception to this with a *secY* and *lipB* allele typically found in subgroup 2E, albeit including the West African strains present in that subgroup which are presumably recent transports. As noted above, the presence of a handful of West African strain types in subgroup 2E suggests recent movement from West Africa to Southern Africa or the Indian Ocean region.

Our results are in general concordance with previous studies which also found some inconsistent phylogeny and some separation of *E. ruminantium* strains into clusters throughout Africa, the Caribbean and the Indian Ocean for a limited number of strains. Previously, Allsopp et al. (2003) and Allsopp and Allsopp (2007) studied the phylogeny of 19 *E. ruminantium* strains from West, North, East, and Southern Africa and Caribbean region. Although inconsistent phylogeny and recombination was found, they observed segregation between West Africa and East-Southern Africa, as well as the introduction of a West African-like strain (Kumm1) to Southern Africa. Another study by Nakao et al. (2011) used an MLST scheme using *gltA*, *groEL*, *lepA*, *lipA*, *lipB*, *secY*, *sodB,* and *sucA* genes, and typed 25 *E. ruminantium* strains from West, North, East and Southern Africa and the Caribbean. They found no strict association between sequence types and geographical origin by using a minimum spanning tree (MST), except for four MLST sequence types from Western Africa, however their study also evidenced recombination amongst the strains, and a neighbor net analysis produced a network with a similar structure to the one from this study (**Figure 1A**; Nakao et al., 2011). However, a panel of five housekeeping genes was used in our study differently from Adakal et al. (2009) and Nakao et al. (2011) that used eight genes. While our data provided less than optimal resolution, given the presence of recombination and possible slow accumulation of variation, the only optimal scheme for *E. ruminantium* is whole genome analysis.

Independently of the genes used for phylogeny and typing, a similar pattern of *E. ruminantium* genetic groups appears between previous studies and our study, with the existence of *E. ruminantium* strain clusters West Africa/Caribbean and Southern/Eastern Africa. In these previous studies, no "Worldwide strain cluster" was evidenced but none of them included more than 25 strains or covered most of the *E. ruminantium* geographic distribution compared with our current study. Specifically, Caribbean and Indian Ocean strains were well represented in our sampled set and we observed clustering of Indian Ocean and Southern Africa isolates for the first time.

Although our study includes a large number of *E. ruminantium* isolates (194), sample size varied for each country. Samples mainly from Burkina Faso (West Africa), Mozambique (Southern Africa), Madagascar (Indian Ocean), and Guadeloupe (Caribbean) collected from several localities to represent sequence types circulating in each region were chosen, with 15–48 samples per country. This might have influenced the observed genetic grouping pattern; however, it is believed that the number of samples and the sampled localities in each of the countries is sufficiently high to have a good representation of the main strains circulating and to obtain a robust *E. ruminantium* genetic population structure.

In addition, strains were mostly isolated from ticks, except in Guadeloupe (Caribbean) where they came from sick ruminants. Therefore, the *E. ruminantium* strains we studied appear to represent the natural populations, including possibly adapted virulent and non-virulent clones and strains except in Guadeloupe, where virulent strains might have been over-represented (Didelot and Maiden, 2010).

E. ruminantium Genetic Population Structure Is Shaped by Recombination

While MLST has been a useful tool in understanding the relationship and population structure of many bacterial species (Sullivan et al., 2005), it only provides a limited view of the phylogenetic signals present in the genomes of the species studied. In species where the evolutionary history of the majority of the genome is the same, the use of several genes may provide a more robust estimation of the relationship of the strains or species studied, and reduce the error in the phylogeny when applied correctly (Yang and Rannala, 2012). In a species where many regions of the genome have differing evolutionary histories due to phenomena such as recombination, the reconstruction of phylogenies based on several genes (either individually or concatenated) may not provide a useful estimation of the relationship between the organisms examined. In the case of *E. ruminantium*, a small selection of the genome (only five genes from a total of roughly 950) exhibits multiple recombination events, suggesting that recombination between strains could be extensive in this species. Thus, it must be noted that the work presented here (as in any other MLST studies) represents only the uncovering of a partial image of the true underlying relationships between *E. ruminantium* strains in terms of their diversity and population structure. The addition of more genes to such an analysis will increase the resolution of the extent of the recombination, and will most likely further differentiate strains which appear to be clonal in more limited analyses.

Apart from the housekeeping genes used for MLST, recombination has also been described for *E. ruminantium* in other genomic regions such as *map1* (Bekker et al., 2005; Hughes and French, 2007). Bekker et al. (2005) found recombination between two *map1* paralogs in *E. ruminantium* Gardel strain while studying the expression of *map1* paralogs in one vector and several non-vector tick cell lines. Another study (Hughes and French, 2007) found evidence of recombination in *map1* alleles by identifying the locus as a statistical outlier in terms of the number of synonymous substitutions found between orthologs of this gene in two different strains. These previous studies using *map1* loci demonstrate the unsuitability of the *map1* gene as a typing marker to provide a phylogeographic structure of strains because of the high polymorphic nature of the locus. For results of studies using the *map1* gene for typing or genetic diversity, recombination events could explain part of *E. ruminantium* genetic diversity and unclear phylogeny.

No studies have yet elucidated inter-strain recombination mechanisms in the intracellular bacteria *E. ruminantium*. Regardless of the absence of plasmids, phages, insertion sequences, or genes for pilus assembly, *E. ruminantium* contain the genes necessary for DNA transfer and recombination by mechanisms that are not well understood (Collins et al., 2005; Thomas and Nielsen, 2005). Although obligate intracellular bacteria do not tend to have known mobile genetic elements, studies have shown the presence of plasmids, prophage, and transposon in the genome of five genera with multiple hosts such as *Wolbachia*, *Coxiella*, *Phytoplasma*, *Rickettsia*, *Chlamydia,* and *Chlamydophila* (Bordenstein and Reznikoff, 2005; Le et al., 2014). Additionally, Ogata et al. (2005) reported the presence of conjugative plasmids in the intracellular bacterium *R. felis*, from full genome sequencing, in the order Rickettsiales (to which *E. ruminantium* belongs). Moreover, repeated non-functional genes and mobile elements such as transposon were also reported in the intracellular alpha-proteobacteria *Orientia tsutsugamushi*, reflecting gene conversion and rearrangement. Furthermore, high levels of diversity and recombination are reported from these intracellular organism (Sonthayanon et al., 2010). Interestingly, these tandem repeat copy numbers are similar to *E. ruminantium* but the differences reflect distinct genetic diversification strategies (Cho et al., 2007).

The genome of *E. ruminantium* does code for gene transfer agent (GTA) genes that produce bacteriophage-like elements that package DNA into small virus-like particles that can be exchanged between cells and present a possible mechanism for genetic transfer between strains (Lang and Beatty, 2000, 2007). These GTAs could mediate exchange of genetic material in *E. ruminantium* in the absence of other mechanisms. More studies need to be done in *E. ruminantium* to identify the mechanism of gene transfer responsible for the recombination observed between strains.

Horizontal gene transfer (HGT) events from bacteria outside of the *Ehrlichia* genus have not been identified, which is probably due to its intracellular lifestyle where it is unlikely to come into contact with bacterial species other than coinfecting strains of *E. ruminantium*. HGT is a well-known mechanism of increasing bacterial genetic diversity that allows the recipients to gain new capabilities evolved in other species, and is often responsible for the spread of bacterial virulence and defense traits between strains or even divergent species (Ochman et al., 2000). The lack of any obvious interspecific events in *E. ruminantium* begs the question as to how it generates genetic diversity and manages to evolve strategies to continually evade host immune systems. We propose that coinfecting *E. ruminantium* strains readily swap genetic material, and that the frequent recombination between strains that we evidence here could play a large role in allowing *E. ruminantium* to continually increase its genetic diversity and avoid the host immune response.

The potential importance of recombination in the generation of genetic diversity in *E. ruminantium* cannot be overstated, because it allows the rapid mixing of variants generated by natural selection in the evolutionary arms race against host immune systems in a species that appears to be otherwise closed off from potential genetic diversity derived from interspecific horizontal transfer events due to its obligate intracellular lifestyle. Inter-strain recombination may allow this bacterium to survive and adapt under various environmental conditions in both vector and host species. The propensity for recombination between strains in this species is at the origin of its diversity which induces limited cross protection between vaccinal and field strains. The recombination between two previously distantly separated groups of the species has particular potential to dramatically increase local genetic diversity in the region where it occurs, outlining the importance of strategies to limit the spread of *E. ruminantium*, even between two regions where heartwater is already present.

Recombination is probably a major driver of genetic diversity in this obligate intracellular pathogen.

Because of recombination events seen in only the small fraction of its genome so far examined, MLST can only offer a limited view of the genetic diversity and population structure of *E. ruminantium* and caution should be taken when interpreting its genetic diversity and population structure, especially using non-network based methods.

Due to the progress of sequencing technologies in terms of cost and throughput, instead of studying a limited number of genes for phylogeny and phylogeography, it would be preferable to sequence the whole genomes of a large number of *E. ruminantium* strains with the sequencing of isolated strains in culture. Furthermore, the isolation of strains in cell culture would allow the measurement and association of phenotypic characteristics of strains with their sequence types, virulence and cross protection between strains. Identification of both recombination events in relevant loci as well as conserved genes that do not show signs of recombination may support the development of effective control strategies and the development of vaccines for *E. ruminantium*.

AUTHOR CONTRIBUTIONS

NC, VP, LB, KH, and RA generated sequence data. JG and NC performed analysis. NC, JG, LN, DM, and NV interpreted results and wrote the manuscript. TL, NV, and LN designed the project. All authors critically reviewed and approved the final manuscript.

FUNDING

This work was financially supported by CIRAD and EPIGENESIS project which received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement No. 31598. FUNDO ABERTO DA UEM 2012-2013 and FUNDO NACIONAL DE INVESTIGAÇÃO Projecto N° 133-Inv/FNI/ 2012-2013 funded the field trips and reagents in Mozambique. French ministry of Agriculture and "Direction de l'Alimentation, de l'Agriculture et de la Forêt de Guadeloupe" financed RESPANG work. This study was partly developed under the project MALIN "Surveillance, diagnosis, control and impact of infectious diseases of humans, animals and plants in tropical islands" supported by the European Union in the framework of the European Regional Development Fund (ERDF) and the Regional Council of Guadeloupe.

ACKNOWLEDGMENTS

We are grateful to our colleagues (Adakal et al., 2009, 2010a), and Raliniaina et al. (2010), Dr. Maxwell Opara, Mr. Asnaoui, Dr. Pablo Tortosa, Dr. Eric Cardinale and Dr. Frederic Stachurski who shared a part of strain DNAs used in this study. We thank Dr. Brigitte Marie and the French Veterinary services for providing samples from Guadeloupe through RESPANG. We are grateful to the Mozambican Veterinary services, South African National Parks and Zambeze Delta Safaris for support during sampling. We would also like to thank Dr. Adela Chavez for her precious comments and critics.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fcimb.2016.00111

REFERENCES

Adakal, H., Gavotte, L., Stachurski, F., Konkobo, M., Henri, H., Zoungrana, S., et al. (2010b). Clonal origin of emerging populations of *Ehrlichia ruminantium* in Burkina Faso. *Infect. Genet. Evol.* 10, 903–912. doi: 10.1016/j.meegid.2010.05.011

Adakal, H., Meyer, D. F., Carasco-Lacombe, C., Pinarello, V., Allegre, F., Huber, K., et al. (2009). MLST scheme of *Ehrlichia ruminantium*: genomic stasis and recombination in strains from Burkina-Faso. *Infect. Genet. Evol.* 9, 1320–1328. doi: 10.1016/j.meegid.2009.08.003

Adakal, H., Stachurski, F., Konkobo, M., Zoungrana, S., Meyer, D. F., Pinarello, V., et al. (2010a). Efficiency of inactivated vaccines against heartwater in Burkina Faso: impact of *Ehrlichia ruminantium* genetic diversity. *Vaccine* 28, 4573–4580. doi: 10.1016/j.vaccine.2010.04.087

Ajmone-Marsan, P., Garcia, J. F., and Lenstra, J. A. (2010). On the origin of cattle: how aurochs became cattle and colonized the world. *Evol. Anthropol.* 19, 148–157. doi: 10.1002/evan.20267

Allsopp, B. A. (2009). Trends in the control of heartwater. *Onderstepoort J. Vet. Res.* 76, 81–88. doi: 10.4102/ojvr.v76i1.69

Allsopp, M. T., and Allsopp, B. A. (2007). Extensive genetic recombination occurs in the field between different genotypes of *Ehrlichia ruminantium*. *Vet. Microbiol.* 124, 58–65. doi: 10.1016/j.vetmic.2007.03.012

Allsopp, M. T., Van Heerden, H., Steyn, H. C., and Allsopp, B. A. (2003). Phylogenetic relationships among *Ehrlichia ruminantium* isolates. *Ann. N.Y. Acad. Sci.* 990, 685–691. doi: 10.1111/j.1749-6632.2003.tb07444.x

Barré, N., Uilenberg, G., Morel, P. C., and Camus, E. (1987). Danger of introducing heartwater onto the American mainland: potential role of indigenous and exotic *Amblyomma* ticks. *Onderstepoort J. Vet. Res.* 54, 405–417.

Beati, L., Patel, J., Lucas-Williams, H., Adakal, H., Kanduma, E. G., Tembo-Mwase, E., et al. (2012). Phylogeography and demographic history of *Amblyomma variegatum* (Fabricius) (Acari: Ixodidae), the tropical bont tick. *Vector Borne Zoonotic Dis.* 12, 514–525. doi: 10.1089/vbz.2011.0859

Bekker, C. P., Postigo, M., Taoufik, A., Bell-Sakyi, L., Ferraz, C., Martinez, D., et al. (2005). Transcription analysis of the major antigenic protein 1 multigene family of three *in vitro*-cultured *Ehrlichia ruminantium* isolates. *J. Bacteriol.* 187, 4782–4791. doi: 10.1128/JB.187.14.4782-4791.2005

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Statist. Soc. B* 57, 289–300. doi: 10.2307/2346101

Blench, R., and MacDonald, K. (2000). *The Origins and Development of African Livestock: Archaeology, Genetics, Linguistics and Ethnography*. New York, NY: Routledge.

Bordenstein, S. R., and Reznikoff, W. S. (2005). Mobile DNA in obligate intracellular bacteria. *Nat. Rev. Microbiol.* 3, 688–699. doi: 10.1038/nrmicro1233

Bouckaert, R. R. (2010). DensiTree: making sense of sets of phylogenetic trees. *Bioinformatics* 26, 1372–1373. doi: 10.1093/bioinformatics/btq110

Bruen, T. C., Philippe, H., and Bryant, D. (2006). A simple and robust statistical test for detecting the presence of recombination. *Genetics* 172, 2665–2681. doi: 10.1534/genetics.105.048975

Cho, N. H., Kim, H. R., Lee, J. H., Kim, S. Y., Kim, J., Cha, S., et al. (2007). The *Orientia tsutsugamushi* genome reveals massive proliferation of conjugative type IV secretion system and host-cell interaction genes. *Proc. Natl. Acad. Sci. U.S.A.* 104, 7981–7986. doi: 10.1073/pnas.0611553104

Collins, N. E., Liebenberg, J., de Villiers, E. P., Brayton, K. A., Louw, E., Pretorius, A., et al. (2005). The genome of the heartwater agent *Ehrlichia ruminantium* contains multiple tandem repeats of actively variable copy number. *Proc. Natl. Acad. Sci. U.S.A.* 102, 838–843. doi: 10.1073/pnas.0406633102

de la Fuente, J., Ruybal, P., Mtshali, M. S., Naranjo, V., Shuqing, L., Mangold, A. J., et al. (2007). Analysis of world strains of *Anaplasma marginale* using major surface protein 1a repeat sequences. *Vet. Microbiol.* 119, 382–390. doi: 10.1016/j.vetmic.2006.09.015

Dent, E. A., and von Holdt, B. M. (2012). STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Res.* 4, 359–361. doi: 10.1007/s12686-011-9548-7

Didelot, X., and Falush, D. (2007). Inference of bacterial microevolution using multilocus sequence data. *Genetics* 175, 1251–1126. doi: 10.1534/genetics.106.063305

Didelot, X., and Maiden, M. C. (2010). Impact of recombination on bacterial evolution. *Trends Microbiol.* 18, 315–322. doi: 10.1016/j.tim.2010.04.002

Drummond, A. J., Suchard, M. A., Xie, D., and Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol. Biol. Evol.* 29, 1969–1973. doi: 10.1093/molbev/mss075

Estrada-Pena, A., Naranjo, V., Acevedo-Whitehouse, K., Mangold, A. J., Kocan, K. M., and de la Fuente, J. (2009). Phylogeographic analysis reveals association of tick-borne pathogen, Anaplasma marginale, MSP1a sequences with ecological traits affecting tick vector performance. *BMC Biol.* 7:57. doi: 10.1186/1741-7007-7-57

Evanno, G., Regnaut, S., and Goudet, J. (2005). Detecting the number of clusters of individuals using the software structure: a simulation study. *Mol. Ecol.* 14, 2611–2620. doi: 10.1111/j.1365-294X.2005.02553.x

Faburay, B., Jongejan, F., Taoufik, A., Ceesay, A., and Geysen, D. (2008). Genetic diversity of *Ehrlichia ruminantium* in *Amblyomma variegatum* ticks and small ruminants in The Gambia determined by restriction fragment profile analysis. *Vet. Microbiol.* 126, 189–199. doi: 10.1016/j.vetmic.2007.06.010

Feil, E., Zhou, J., Maynard Smith, J., and Spratt, B. G. (1996). A comparison of the nucleotide sequences of the adk and recA genes of pathogenic and commensal *Neisseria* species: evidence for extensive interspecies recombination within adk. *J. Mol. Evol.* 43, 631–640. doi: 10.1007/BF02202111

Guindon, S., Dufayard, J. F., Lefort, V., Anisimova, M., Hordijk, W., and Gascuel, O. (2010). New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* 59, 307–321. doi: 10.1093/sysbio/syq010

Hanotte, O., Bradley, D. G., Ochieng, J. W., Verjee, Y., Hill, E. W., and Rege, J. E. (2002). African pastoralism: genetic imprints of origins and migrations. *Science* 296, 336–339. doi: 10.1126/science.1069878

Harrington, B. (2004–2005). *Inkscape*. Available online at: http://www.inkscape.org/

Holmes, E. C., Urwin, R., and Maiden, M. C. (1999). The influence of recombination on the population structure and evolution of the human pathogen *Neisseria meningitidis*. *Mol. Biol. Evol.* 16, 741–749. doi: 10.1093/oxfordjournals.molbev.a026159

Hughes, A. L., and French, J. O. (2007). Homologous recombination and the pattern of nucleotide substitution in *Ehrlichia ruminantium*. *Gene* 387, 31–37. doi: 10.1016/j.gene.2006.08.003

Huson, D. H., and Bryant, D. (2006). Application of phylogenetic networks in evolutionary studies. *Mol. Biol. Evol.* 23, 254–267. doi: 10.1093/molbev/msj030

Jakobsson, M., and Rosenberg, N. A. (2007). CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23, 1801–1806. doi: 10.1093/bioinformatics/btm233

Jongejan, F., Thielemans, M. J., Briere, C., and Uilenberg, G. (1991). Antigenic diversity of *Cowdria ruminantium* isolates determined by cross-immunity. *Res. Vet. Sci.* 51, 24–28. doi: 10.1016/0034-5288(91)90025-J

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., et al. (2012). Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28, 1647–1649. doi: 10.1093/bioinformatics/bts199

Lang, A. S., and Beatty, J. T. (2000). Genetic analysis of a bacterial genetic exchange element: the gene transfer agent of *Rhodobacter capsulatus*. *Proc. Natl. Acad. Sci. U.S.A.* 97, 859–864. doi: 10.1073/pnas.97.2.859

Lang, A. S., and Beatty, J. T. (2007). Importance of widespread gene transfer agent genes in alpha-proteobacteria. *Trends Microbiol.* 15, 54–62. doi: 10.1016/j.tim.2006.12.001

Larsson, A. (2014). AliView: a fast and lightweight alignment viewer and editor for large datasets. *Bioinformatics* 30, 3276–3278. doi: 10.1093/bioinformatics/btu531

Le, P. T., Pontarotti, P., and Raoult, D. (2014). Alphaproteobacteria species as a source and target of lateral sequence transfers. *Trends Microbiol.* 22, 147–156. doi: 10.1016/j.tim.2013.12.006

Lindstedt, B. A. (2005). Multiple-locus variable number tandem repeats analysis for genetic fingerprinting of pathogenic bacteria. *Electrophoresis* 26, 2567–2582. doi: 10.1002/elps.200500096

Lorenzen, E. D., Masembe, C., Arctander, P., and Siegismund, H. R. (2010). A long-standing Pleistocene refugium in southern Africa and a mosaic of refugia in east Africa: insights from mtDNA and the common eland antelope. *J. Biogeogr.* 37, 571–581. doi: 10.1111/j.1365-2699.2009.02207.x

Magee, D. A., Meghen, C., Harrison, S., Troy, C. S., Cymbron, T., Gaillard, C., et al. (2002). A partial african ancestry for the creole cattle populations of the Caribbean. *J. Hered.* 93, 429–432. doi: 10.1093/jhered/93.6.429

Maillard, J. C., Kemp, S. J., Naves, M., Palin, C., Demangel, C., Accipe, A., et al. (1993). An attempt to correlate cattle breed origins and diseases associated with or transmitted by the tick *Amblyomma variegatum* in the French West Indies. *Revue. Élev. Méd. Vét. Pays Trop.* 46, 283–290.

Martin, D. P., and Beiko, R. G. (2010). "Genetic recombination and bacterial population structure," in *Bacterial Population Genetics in Infectious Disease*, eds D. A. Robinson, D. Falush, and E. J. Feil (Hoboken, NJ: John Wiley & Sons, Inc.), 61–85.

Molia, S., Frebling, M., Vachiéry, N., Pinarello, V., Petitclerc, M., Rousteau, A., et al. (2008). *Amblyomma variegatum* in cattle in Marie Galante, French Antilles: prevalence, control measures, and infection by *Ehrlichia ruminantium*. *Vet. Parasitol.* 153, 338–346. doi: 10.1016/j.vetpar.2008.01.046

Moumene, A., and Meyer, D. F. (2016). Ehrlichia's molecular tricks to manipulate their host cells. *Microbes Infect.* 18, 172–179. doi: 10.1016/j.micinf.2015.11.001

Nakao, R., Magona, J. W., Zhou, L., Jongejan, F., and Sugimoto, C. (2011). Multilocus sequence typing of *Ehrlichia ruminantium* strains from geographically diverse origins and collected in *Amblyomma variegatum* from Uganda. *Parasit. Vectors* 4:137. doi: 10.1186/1756-3305-4-137

Ochman, H., Lawrence, J. G., and Groisman, E. A. (2000). Lateral gene transfer and the nature of bacterial innovation. *Nature* 405, 299–304. doi: 10.1038/35012500

Ogata, H., Renesto, P., Audic, S., Robert, C., Blanc, G., Fournier, P. E., et al. (2005). The genome sequence of *Rickettsia felis* identifies the first putative conjugative plasmid in an obligate intracellular parasite. *PLoS Biol.* 3:e248. doi: 10.1371/journal.pbio.0030248

Paradis, E., Claude, J., and Strimmer, K. (2004). APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* 20, 289–290. doi: 10.1093/bioinformatics/btg412

Pilet, H., Vachiéry, N., Berrich, M., Bouchouicha, R., Durand, B., Pruneau, L., et al. (2012). A new typing technique for the Rickettsiales *Ehrlichia ruminantium*: multiple-locus variable number tandem repeat analysis. *J. Microbiol. Methods* 88, 205–211. doi: 10.1016/j.mimet.2011.11.011

Pitra, C., Hansen, A. J., Lieckfeldt, D., and Arctander, P. (2002). An exceptional case of historical outbreeding in African sable antelope populations. *Mol. Ecol.* 11, 1197–1208. doi: 10.1046/j.1365-294X.2002.01516.x

Posada, D., and Crandall, K. A. (2002). The effect of recombination on the accuracy of phylogeny estimation. *J. Mol. Evol.* 54, 396–402. doi: 10.1007/s00239-001-0034-9

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

Provost, A., and Bezuidenhout, J. D. (1987). The historical background and global importance of heartwater. *Onderstepoort J. Vet. Res.* 54, 165–169.

Raliniaina, M., Meyer, D. F., Pinarello, V., Sheikboudou, C., Emboule, L., Kandassamy, Y., et al. (2010). Mining the genetic diversity of *Ehrlichia*

ruminantium using map genes family. *Vet. Parasitol.* 167, 187–195. doi: 10.1016/j.vetpar.2009.09.020

Rambaut, A., Suchard, M. A., Xie, D., and Drummond, A. J. (2014). *Tracer v1.6.* Available online at: http://beast.bio.ed.ac.uk/Tracer

R Core Team (2013). *R: A Language and Environment for Statistical Computing*. Available online at: http://www.R-project.org/

Roth, J. A., Richt, J. A., and Morozov, I. A. (eds.) (2013). Vaccines and diagnostics for transboundary animal diseases. *Dev. Biol* 135, 191–200. doi: 10.1159/isbn.978-3-318-02366-4

Ruths, D., and Nakhleh, L. (2005). Recombination and phylogeny: effects and detection. *Int. J. Bioinform Res. Appl.* 1, 202–212. doi: 10.1504/IJBRA.2005.007578

Schierup, M. H., and Hein, J. (2000). Consequences of recombination on traditional phylogenetic analysis. *Genetics* 156, 879–891.

Shriner, D., Nickle, D. C., Jensen, M. A., and Mullins, J. I. (2003). Potential impact of recombination on sitewise approaches for detecting positive natural selection. *Genet. Res.* 81, 115–121. doi: 10.1017/S0016672303006128

Smith, J. M., and Smith, N. H. (1998). Detecting recombination from gene trees. *Mol. Biol. Evol.* 15, 590–599. doi: 10.1093/oxfordjournals.molbev.a025960

Smith, J. M., Smith, N. H., O'Rourke, M., and Spratt, B. G. (1993). How clonal are bacteria? *Proc. Natl. Acad. Sci. U.S.A.* 90, 4384–4388. doi: 10.1073/pnas.90.10.4384

Sonthayanon, P., Peacock, S. J., Chierakul, W., Wuthiekanun, V., Blacksell, S. D., Holden, M. T., et al. (2010). High rates of homologous recombination in the mite endosymbiont and opportunistic human pathogen *Orientia tsutsugamushi*. *PLoS Negl. Trop. Dis.* 4:e752. doi: 10.1371/journal.pntd.0000752

Spratt, B. G., Smith, N. H., Zhou, J. J., O'Rourke, M., and Feil, E. (1995). "The population genetics of the pathogenic *Neisseria*," in *Population Genetics of Bacteria*, eds S. Baumberg, J. P. W. Young, E. M. H. Wellington, and J. R. Saunders (Cambridge: Cambridge University Press), 143–160.

Stachurski, F., Tortosa, P., Rahajarison, P., Jacquet, S., Yssouf, A., and Huber, K. (2013). New data regarding distribution of cattle ticks in the south-western Indian Ocean islands. *Vet. Res.* 44:79. doi: 10.1186/1297-9716-44-79

Sullivan, C. B., Diggle, M. A., and Clarke, S. C. (2005). Multilocus sequence typing: data analysis in clinical microbiology and public health. *Mol. Biotechnol.* 29, 245–254. doi: 10.1385/MB:29:3:245

Tamura, K., and Nei, M. (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* 10, 512–526.

Thomas, C. M., and Nielsen, K. M. (2005). Mechanisms of, and barriers to, horizontal gene transfer between bacteria. *Nat. Rev. Microbiol.* 3, 711–721. doi: 10.1038/nrmicro1234

Urwin, R., and Maiden, M. C. (2003). Multi-locus sequence typing: a tool for global epidemiology. *Trends Microbiol.* 11, 479–487. doi: 10.1016/j.tim.2003.08.006

Vachiéry, N., Jeffery, H., Pegram, R., Aprelon, R., Pinarello, V., Kandassamy, R. L., et al. (2008). *Amblyomma variegatum* ticks and heartwater on three Caribbean Islands. *Ann. N.Y. Acad. Sci.* 1149, 191–195. doi: 10.1196/annals.1428.081

Vachiéry, N., Marcelino, I., Martinez, D., and Lefrancois, T. (2013). Opportunities in diagnostic and vaccine approaches to mitigate potential heartwater spreading and impact on the American mainland. *Dev. Biol. (Basel.)* 135, 191–200. doi: 10.1159/000190050

Voltzit, O. V., and Keirans, J. E. (2003). A review of African *Amblyomma* species (Acari, Ixodida, Ixodidae). *Acarina* 11, 135–214.

Walker, J. B., and Olwage, A. (1987). The tick vectors of *Cowdria ruminantium* (Ixodoidea, Ixodidae, genus *Amblyomma*) and their distribution. *Onderstepoort J. Vet. Res.* 54, 353–379.

Warnes, G. R., Bolker, B., Bonebakker, L., Gentleman, R., Liaw, W. H. A., Lumley, T., et al. (2009). *gplots: Various R Programming Tools for Plotting Data*. Available online at: http://cran.r-project.org/package=gplots

Yang, Z., and Rannala, B. (2012). Molecular phylogenetics: principles and practice. *Nat. Rev. Genet.* 13, 303–314. doi: 10.1038/nrg3186

Zhou, J., Bowler, L. D., and Spratt, B. G. (1997). Interspecies recombination, and phylogenetic distortions, within the glutamine synthetase and shikimate dehydrogenase genes of *Neisseria meningitidis* and commensal *Neisseria* species. *Mol. Microbiol.* 23, 799–812. doi: 10.1046/j.1365-2958.1997.2681633.x

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Cangi, Gordon, Bournez, Pinarello, Aprelon, Huber, Lefrançois, Neves, Meyer and Vachiéry. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

SUPPLEMENTARY MATERIAL

Figure S1 Maximum likelihood phylogeny constructed with PhyML under a GTR+G+I model of evolution with (a) and without (b) *E. chaffeensis* as an outgroup. Branch support was calculated using the aLRT method, and low support values are possibly indicative of short branches or mixed phylogenetic signals in the data, potentially introduced by recombination or other forms of homoplasy. The outgroup branch is not to scale to allow legibility of the figure.

Figure S2 Heat map of similarity and differences of 5 concatenated housekeeping genes among 97 unique *E. ruminantium* sequences. Hypothetical groups and subgroups 1, 2A, G2B, 2C, 2D and 2D are marked by dashes. Degree of relatedness is indicated by colours from white (different) to red (similar). The name of each isolate was labelled on the right side of the graphs and corresponds to the same strains on the bottom of the graph. Dendograms representing the clusters were placed on the left side and on top of the graphs. Strains 3, 19, 42, 43, 50, 56, 63, 69, 72 and 74 are marked with a square and represent recombinants.

Figure S3 Multiple sequence alignment of the different genotypes containing only variable positions. The subgroups clusters are separated by horizontal gaps, and the five different genes are boxed and separated by vertical gaps. Coloured residues represent non-consensus characters over the whole alignment. Blue stars indicate recombinant genotypes and red stars represent sequence types that are inferred to have less than 80% ancestry from a single population in STRUCTURE.

Table S1 Number of *E. ruminantium* isolates/strains per geographic region and country

Geographic region	Country	Number of isolates	Total
North and Central Africa	Chad	1	
	Sudan	1	2
West Africa	Burkina Faso	**44**	
	Cameroon	1	
	Gambia	1	
	Ghana	2	
	Nigeria	2	
	São Tome and Principe	1	
	Senegal	4	55
East Africa	Kenya	1	
	Tanzania	1	
	Uganda	1	3
Southern Africa	Mozambique	**58**	
	South Africa	4	
	Zambia	1	
	Zimbabwe	1	64
Indian Ocean	Comoros	9	
	Madagascar	**13**	
	Mayotte	5	
	Reunion	2	29
Caribbean	Antigua	1	
	Guadeloupe	**40**	41
			194

Table S2 Description of *E. ruminantium* isolates/strains based on the sequence type number, genetic group, geographic origin, country of isolation, isolate name, DNA origin, date of isolation and reference.

Sequence type number	Genetic group	Geographic origin	Country	Isolate/Strain name	DNA origin	Date of isolation	Isolate Reference
1	G2D	East Africa	Kenya	**Kiswani**	B	1985	Raliniaina et al. (2010)
2	G2D	Central Africa	Chad	TCH6	T	2008	Our study
3	No group	Southern Africa	South Africa	SAZeerust	CC	1979	Nakao et al. (2011)
4	G2C	Southern Africa	Mozambique	GAH1MH2	T	2012	Our study
4*	G2C			GAH4MH1	T	2012	Our study
5	G2C			CIT9MH1	T	2012	Our study
5*	G2C			GAH7MH1	T	2012	Our study
5*	G2C			MWAB11FH1	T	2012	Our study
5*	G2C			CIT28MH1	T	2012	Our study
5*	G2C			Umpala	B	1995	Raliniaina et al. (2010)
5*	G2C			MWAB11MH1	T	2012	Our study
5*	G2C			MAT14MH2	T	2012	Our study
5*	G2C			MAS13MH1	T	2012	Our study
6	G2E			CHIPO29MH1	T	2012	Our study
6*	G2E			ZIM15MH1	T	2012	Our study
7	G2E			CHIPO17MH1	T	2012	Our study
8	G2E			MAS27MH2	T	2012	Our study
8*	G2E			ZIM31MH1	T	2013	Our study
9	G2E			CHIPO12MH1	T	2012	Our study
10	G2C			MAH12MH1	T	2012	Our study
10*	G2C			VUL29MH1	T	2012	Our study
10*	G2C			MWAB3MH1	T	2012	Our study
10*	G2C			ZIM28MH1	T	2012	Our study
10*	G2C			Crystal Springs	CC	1990	Nakao et al. (2011)
11	G2C			MAT13MH2	T	2012	Our study
12	G2E			ZIM16MH2	T	2012	Our study
13	G2E			VUL17MH1	T	2012	Our study
14	G2E			CHIPO22MH1	T	2012	Our study
15	G2E			GAH5MH2	T	2012	Our study
16	G2E			CHIPO2MH1	T	2012	Our study
16*	G2E			CHIPA3MH1	T	2012	Our study
17	G2E			ZIM2MH1	T	2012	Our study
17*	G2E			ZIM4MH1	T	2012	Our study
18	G2E			GAH9MH1	T	2012	Our study
18*	G2E	Southern Africa	South Africa	SABall3	CC	1952	Nakao et al. (2011)
19	No group	Southern Africa	South Africa	**Mara**	CC	1998	Raliniaina et al. (2010)
20	G2E	Southern Africa	Mozambique	MAH6MH1	T	2012	Our study
21	G2C			MAS14MH1	T	2012	Our study
22	G1			303-GOV1-MV8	T	2012	Our study
23	G2E			335-MAMMV13	T	2013	Our study
24	G2E			445-MAP32MH1	T	2013	Our study
25	G2E			550-NHAMV12	T	2014	Our study
26	G2E			559-NHAMV21	T	2014	Our study
27	G2E			330-MAMMV8	T	2013	Our study
28	G2E			347-MAMMV22	T	2013	Our study

29	G2E			319-MAMMV2	T	2013	Our study
30	G2E			336-MAMFV6	T	2013	Our study
31	G2E			342-MAMMV17	T	2013	Our study
32	G2D			709-DAR5MV1	T	2014	Our study
33	G2E			765-ESP2-4MV1	T	2014	Our study
34	G2E			777-DAC8MH1	T	2014	Our study
35	G2E			823-MAG7MH1	T	2014	Our study
36	G2C			431-MAP21MH1	T	2013	Our study
37	G2C			832-MAG12MH1	T	2014	Our study
38	G2C			801-MUE3FH1	T	2014	Our study
39	G2D			488-FINMV13	T	2014	Our study
40	G2D			595-MANINMV19	T	2014	Our study
41	G2D			385-MUX3MV1	T	2013	Our study
42	G2D			394-MUX9MV1	T	2013	Our study
43	No group	West Africa	Burkina Faso	Sara401	T	2002	Raliniaina et al. (2010)
44	G2D			Lamba194	B	2003	Raliniaina et al. (2010)
44*	G2D			**Banankeledaga**	CC	1998	Raliniaina et al. (2010)
44*	G2D			BF629	T	2009	Adakal et al. (2010)
44*	G2D			BF630	T	2009	Adakal et al. (2010)
44*	G2D			BF635	T	2009	Adakal et al. (2010)
45	G2D			Banan455	B	2003	Raliniaina et al. (2010)
46	G1	West Africa	Senegal	M310	T	2002	Our study
47	G2D	West Africa	Burkina Faso	Banan033F1	T	2002	Raliniaina et al. (2010)
48	G1	West Africa	Ghana	Sankat430	CC	1996	Nakao et al. (2011)
48*	G1	West Africa	Burkina Faso	BF1210	T	2007	Adakal et al. (2010)
48*	G1			BF1795	T	2007	Adakal et al. (2010)
48*	G1			BF1796	T	2007	Adakal et al. (2010)
48*	G1			BF1798	T	2007	Adakal et al. (2010)
48*	G1			BF19	T	2007	Adakal et al. (2010)
49	G1	West Africa	Gambia	Kerr Seringe	CC	2001	Nakao et al. (2011)
49*	G1	West Africa	Senegal	M10T	T	2002	Raliniaina et al. (2010)
50	G2A			M16T	T	2002	Raliniaina et al. (2010)
51	G2B	West Africa	Burkina Faso	BF623	T	2007	Adakal et al. (2010)
52	G1	West Africa	Nigeria	SK43M1	T	2010	Our study
53	G1	West Africa	Burkina Faso	BF331	T	2007	Adakal et al. (2010)
54	G2E	West Africa	São Tome and Principe	São Tome	CC	1981	Nakao et al. (2011)
55	G2D	West Africa	Burkina Faso	BF1042	T	2007	Adakal et al. (2010)
56	G2A			BF1062	T	2007	Adakal et al. (2010)
57	G2E			BF1232	T	2007	Adakal et al. (2010)
58	G2B			BF1267	T	2007	Adakal et al. (2010)
59	G1			BF1799	T	2007	Adakal et al. (2010)
60	G1			BF1905	T	2007	Adakal et al. (2010)
61	G1			BF1948	T	2007	Adakal et al. (2010)
62	G2D			BF1951	T	2007	Adakal et al. (2010)
63	No group			BF2185	T	2007	Adakal et al. (2010)
64	G2B			BF631	T	2007	Adakal et al. (2010)
65	G2E			BF668	T	2007	Adakal et al. (2010)
66	G2E			BF708	T	2007	Adakal et al. (2010)
67	G2D	Caribbean	Antigua	GeorgesM3	T	2005	Raliniaina et al. (2010)
68	G2D	Caribbean	Guadeloupe	34-0205CM01	B	2011	Our study
68*	G2D			**Gardel**	CC	1962	Raliniaina et al. (2010)

68*	G2D			27-2103JMR03	B	2011	Our study
69	G2A			11-1711BP02	B	2010	Our study
70	G2D			25-2103JMR01	B	2011	Our study
70*	G2D			49-250112VL01	B	2012	Our study
70*	G2D			38-0507AS01	B	2011	Our study
70*	G2D			46-061211JCA01	B	2011	Our study
70*	G2D			SUI22M1B1	B	2008	Our study
70*	G2D			n6631	B	2009	Our study
71	G2D			35-0805JMR01	B	2011	Our study
71*	G2D			n5697	B	2005	Our study
71*	G2D			n5097	B	2002	Our study
72	G2A			13-2112JCA01	B	2011	Our study
72*	G2A			SUI24JM	B	2008	Our study
72*	G2A			14-2112JCA02	B	2011	Our study
72*	G2A			6-2709EH03	B	2010	Our study
72*	G2A			48-291111FB01	B	2011	Our study
72*	G2A			n971128610M2	B	2005	Our study
72*	G2A			15-2112JCA03	B	2011	Our study
73	G2E			33-2704AS01	B	2011	Our study
73*	G2E			42-1509JE01	B	2011	Our study
73*	G2E			39-2008FB01	B	2011	Our study
73*	G2E			32-1104FB01	B	2011	Our study
73*	G2E			40-2408FB02	B	2011	Our study
74	G2A			19-0202BP01	B	2011	Our study
75	G2B			44-2110JE01	B	2011	Our study
75*	G2B			26-2103JMR02	B	2011	Our study
75*	G2B			43-2610VL01	B	2011	Our study
76	G2D			36-1405BP01	B	2011	Our study
76*	G2D			30-1304MC01	B	2011	Our study
77	G2D			45-161111MM01	B	2011	Our study
77*	G2D			n6001	B	2000	Our study
77*	G2D			29-2903JMR01	B	2011	Our study
78	G2D			n6653/1-2313	B	2002	Our study
79	G2D	Indian Ocean	Comoros	AY0024	T	2010	Our study
80	G2E	Indian Ocean	Reunion	APLSM1	T	2010	Our study
81	G2D	Indian Ocean	Madagascar	Madaman1	T	2008	Our study
82	G2E	Indian Ocean	Comoros	n1690	T	2007	Our study
83	G2D	Indian Ocean	Madagascar	n13BM1	T	2010	Our study
84	G2E	Indian Ocean	Mayotte	n164B2458	T	2010	Our study
85	G2E	Indian Ocean	Madagascar	n8EM3	T	2010	Our study
86	G2E	Indian Ocean	Comoros	AY0015	T	2010	Our study
87	G2D			AY0041	T	2010	Our study
88	G2E	Indian Ocean	Mayotte	TiquesM3	T	2010	Our study
88*	G2E	Indian Ocean	Madagascar	n8DF3	T	2010	Our study
88*	G2E	Indian Ocean	Mayotte	n206B	T	2010	Our study
88*	G2E			YTBARA8M1	T	2009	Our study
89	G2E	Indian Ocean	Madagascar	n2CM4	T	2009	Our study
90	G2E			n14AF1	T	2010	Our study
91	G2E	Indian Ocean	Comoros	AY0091	T	2010	Our study
91*	G2E			AY0087	T	2010	Our study
92	G2E	Indian Ocean	Madagascar	KJSF	T	2001	Our study
92*	G2E			Madaman3	T	2008	Our study

92*	G2E			n14CM1	T	2010	Our study
92*	G2E			Madaman13	T	2008	Our study
92*	G2E	Indian Ocean	Comoros	n3700	T	2007	Our study
92*	G2E			n3683	T	2007	Our study
92*	G2E	Indian Ocean	Reunion	BDLSM3	T	2010	Our study
92*	G2E	Indian Ocean	Madagascar	Madaman4	T	2008	Our study
92*	G2E	Indian Ocean	Comoros	n3655	T	2007	Our study
92*	G2E	Indian Ocean	Madagascar	RZF	T	2001	Raliniaina et al. (2010)
93	G2E	Southern Africa	Mozambique	CHIPO26MH1	T	2012	Our study
93*	G2E			CHIPA2MH1	T	2012	Our study
93*	G2E	West Africa	Cameroon	Cameroun	CC	1994	Raliniaina et al. (2010)
93*	G2E	Southern Africa	Mozambique	MAS1MH1	T	2012	Our study
93*	G2E	Southern Africa	South Africa	**Welgevonden**	CC	1985	Raliniaina et al. (2010)
93*	G2E	Southern Africa	Mozambique	CHIPO24MH2	T	2012	Our study
93*	G2E			ZIM1MH1	T	2012	Our study
94	G2E	Southern Africa	Zambia	**Lutale**	CC	1988	Raliniaina et al. (2010)
94*	G2E	North Africa	Sudan	**Umbanein**	CC	1981	Raliniaina et al. (2010)
94*	G2E	Indian Ocean	Madagascar	Madamora3	T	2008	Our study
94*	G2E	East Africa	Uganda	KBL4M	T	1999	Our study
94*	G2E	Indian Ocean	Mayotte	YTAVI001	T	2009	Our study
95	G2D	West Africa	Nigeria	NigeriaIfe	B	1983	Nakao et al. (2011)
95*	G2D	Caribbean	Guadeloupe	4-2007AS02	T	2010	Our study
95*	G2D	West Africa	Burkina Faso	BF2	T	2007	Adakal et al. (2010)
95*	G2D	Caribbean	Guadeloupe	37-0806FB01	T	2011	Our study
95*	G2D			21-2702VL01	T	2011	Our study
95*	G2D			41-0309FB01	T	2011	Our study
96	G1	East Africa	Tanzania	AB014TAN	T	2010	Our study
96*	G1	West Africa	Ghana	**Pokoase**	CC	1996	Raliniaina et al. (2010)
97	G1	West Africa	Burkina Faso	lamba479	T	2001	Raliniaina et al. (2010)
97*	G1	Caribbean	Guadeloupe	17-2701GM01	T	2011	Our study
97*	G1	West Africa	Burkina Faso	bankouma421	T	2001	Raliniaina et al. (2010)
97*	G1	West Africa	Senegal	**Senegal**	CC	1994	Raliniaina et al. (2010)
97*	G1	West Africa	Burkina Faso	Sara292	T	2001	Raliniaina et al. (2010)
97*	G1			Lamba107	T	2002	Raliniaina et al. (2010)
97*	G1			Bekuy255	CC	2001	Raliniaina et al. (2010)
97*	G1			BF395	T	2007	Adakal et al. (2010)
97*	G1			BF1114	T	2007	Adakal et al. (2010)
97*	G1			BF1946	T	2007	Adakal et al. (2010)
97*	G1			BF2165	T	2007	Adakal et al. (2010)
97*	G1			BF461	T	2007	Adakal et al. (2010)
97*	G1			BF463	T	2007	Adakal et al. (2010)
97*	G1			BF466	T	2007	Adakal et al. (2010)
97*	G1			BF469	T	2007	Adakal et al. (2010)
97*	G1			BF474	T	2008	Adakal et al. (2010)
97*	G1			BF476	T	2007	Adakal et al. (2010)
97*	G1			BF810	T	2007	Adakal et al. (2010)

*Identical DNA sequence (clone). Reference strains are highlighted in bold. CC: Cell culture; T: Tick, B: Blood.

Table S3 Number of *E. ruminantium* isolates per genetic group and country

Group	Country	Number of samples	Total
1	Burkina Faso	**25**	
	Gambia	1	
	Ghana	2	
	Guadeloupe	1	
	Mozambique	1	
	Nigeria	1	
	Senegal	3	
	Tanzania	1	35
2A	Burkina Faso	1	
	Guadeloupe	**9**	
	Senegal	1	11
2B	Burkina Faso	3	
	Guadeloupe	3	6
2C	Mozambique	**19**	
	Zimbabwe	1	20
2D	Antigua	1	
	Burkina Faso	10	
	Chad	1	
	Comoros	2	
	Guadeloupe	**22**	
	Kenya	1	
	Madagascar	2	
	Mozambique	5	
	Nigeria	1	
	South Africa	1	46
2E	Burkina Faso	3	
	Cameroon	1	
	Comoros	7	
	Guadeloupe	5	
	Madagascar	11	
	Mayotte	5	
	Mozambique	**33**	
	Reunion	2	
	São Tome e Principe	1	
	South Africa	2	
	Sudan	1	
	Uganda	1	
	Zambia	1	73
No group	Burkina Faso	**2**	
	South Africa	1	3
Total			**194**

Table S4 GeneBank accession number corresponding to each gene sequence (order *sucA-sodB-lipA-secY-lipB*) for 67 *E. ruminantium* sequence type

Sequence type number	Accession number				
	sucA	sodB	lipA	secY	lipB
1	KX821405	KX821339	KX889850	KX821470	KX821537
2	KX821406	KX821340	KX889851	KX821471	KX821538
3	KX821407	KX821341	KX889852	KX821472	KX821539
4	KX821408	KX821342	KX889853	KX821473	KX821540
5	KX821409	KX821343	KX889854	KX821474	KX821541
6	KX821410	KX821344	KX889855	KX821475	KX821542
7	KX821411	KX821345	KX889856	KX821476	KX821543
8	KX821412	KX821346	KX889857	KX821477	KX821544
9	KX821413	KX821347	KX889858	KX821478	KX821545
10	KX821414	KX821348	KX889859	KX821479	KX821546
11	KX821415	KX821349	KX889860	KX821480	KX821547
12	KX821416	KX821350	KX889861	KX821481	KX821548
13	KX821417	KX821351	KX889862	KX821482	KX821549
14	KX821418	KX821352	KX889863	KX821483	KX821550
15	KX821419	KX821353	KX889864	KX821484	KX821551
16	KX821420	KX821354	KX889865	KX821485	KX821552
17	KX821421	KX821355	KX889866	KX821486	KX821553
18	KX821422	KX821356	KX889867	KX821487	KX821554
19	KX821423	KX821357	KX889868	KX821488	KX821555
20	KX821424	KX821358	KX889869	KX821489	KX821556
21	KX821425	KX821359	KX889870	KX821490	KX821557
22	KX821426	KX821360	KX889871	KX821491	KX821558
23	KX821427	KX821361	KX889872	KX821492	KX821559
24	KX821428	KX821362	KX889873	KX821493	KX821560
25	KX821429	KX821363	KX889874	KX821494	KX821561
26	KX821430	KX821364	KX889875	KX821495	KX821562
27	KX821431	KX821365	KX889876	KX821496	KX821563
28	KX821432	KX821366	KX889877	KX821497	KX821564
29	KX821433	KX821367	KX889878	KX821498	KX821565
30	KX821434	KX821368	KX889879	KX821499	KX821566
31	KX821435	KX821369	KX889880	KX821500	KX821567
32	KX821436	KX821370	KX889881	KX821501	KX821568
33	KX821437	KX821371	KX889882	KX821502	KX821569
34	KX821438	KX821372	KX889883	KX821503	KX821570
35	KX821439	KX821373	KX889884	KX821504	KX821571
36	KX821440	KX821374	KX889885	KX821505	KX821572

37	KX821441	KX821375	KX889886	KX821506	KX821573
38	KX821442	KX821376	KX889887	KX821507	KX821574
39	KX821443	KX821377	KX889888	KX821508	KX821575
40	KX821444	KX821378	KX889889	KX821509	KX821576
41	KX821445	KX821379	KX889890	KX821510	KX821577
42	KX821446	KX821380	KX889891	KX821511	KX821578
43	KX821447	KX821381	KX889892	KX821512	KX821579
44	KX821448	KX821382	KX889893	KX821513	KX821580
45	KX821449	KX821383	KX889894	KX821514	KX821581
46	KX821450	KX821384	KX889895	KX821515	KX821582
47	KX821451	KX821385	KX889896	KX821516	KX821583
48	KX821452	KX821386	KX889897	KX821517	KX821584
49	KX821453	KX821387	KX889898	KX821518	KX821585
50	KX821454	KX821388	KX889899	KX821519	KX821586
51	KX821455	KX821389	KX889900	KX821520	KX821587
52	KX821456	KX821390	KX889901	KX821521	KX821588
53	KX821457	KX821391	KX889902	KX821522	KX821589
54	KX821458	KX821392	KX889903	KX821523	KX821590
55	KX821459	KX821393	KX889904	KX821524	KX821591
56	KX821460	KX821394	KX889905	KX821525	KX821592
57	KX821461	KX821395	KX889906	KX821526	KX821593
58	KX821462	KX821396	KX889907	KX821527	KX821594
59	KX821463	KX821397	KX889908	KX821528	KX821595
60	KX821464	KX821398	KX889909	KX821529	KX821596
61	KX821465	KX821399	KX889910	KX821530	KX821597
62	KX821466	KX821400	KX889911	KX821531	KX821598
63	KX821467	KX821401	KX889912	KX821532	KX821599
64	KX821468	KX821402	KX889913	KX821533	KX821600
65	KX821469	KX821403	KX889914	KX821534	KX821601
66	x[a]	KX821404	KX889915	KX821535	KX821602
67	x[a]	x[a]	KX889916	KX821536	KX821603

[a]: Not possible to have accession number for this DNA sequences

Figure S1

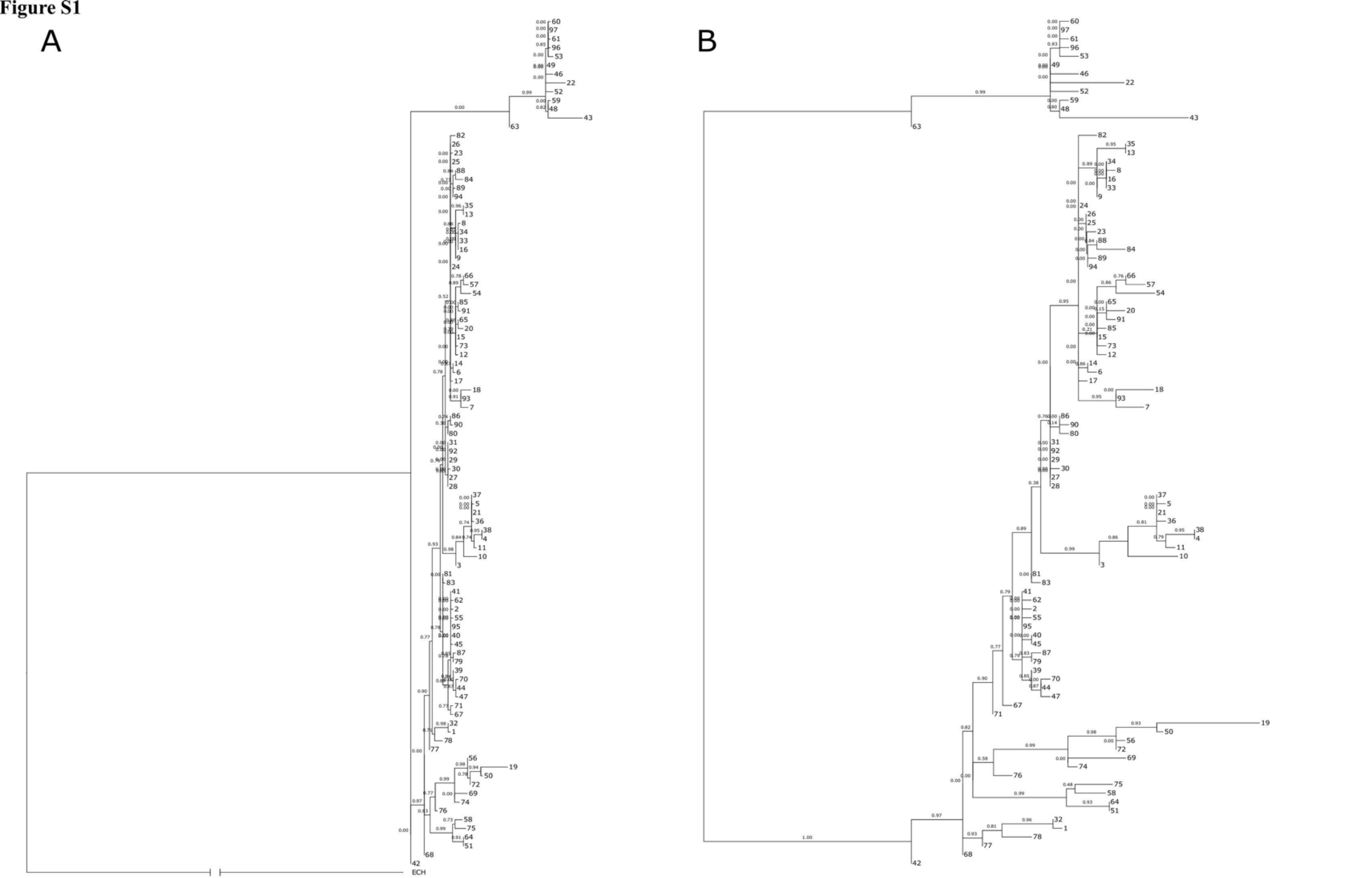


Figure S1 Maximum likelihood phylogeny constructed with PhyML under a GTR+G+I model of evolution with (a) and without (b) *E. chaffeensis* as an outgroup. Branch support was calculated using the aLRT method, and low support values are possible indicative of short branches or mixed phylogenetic signals in the data, potentially introduced by recombination or other forms of homoplasy. The outgroup branch is not to scale to allow legibility of the figure.

Figure S2

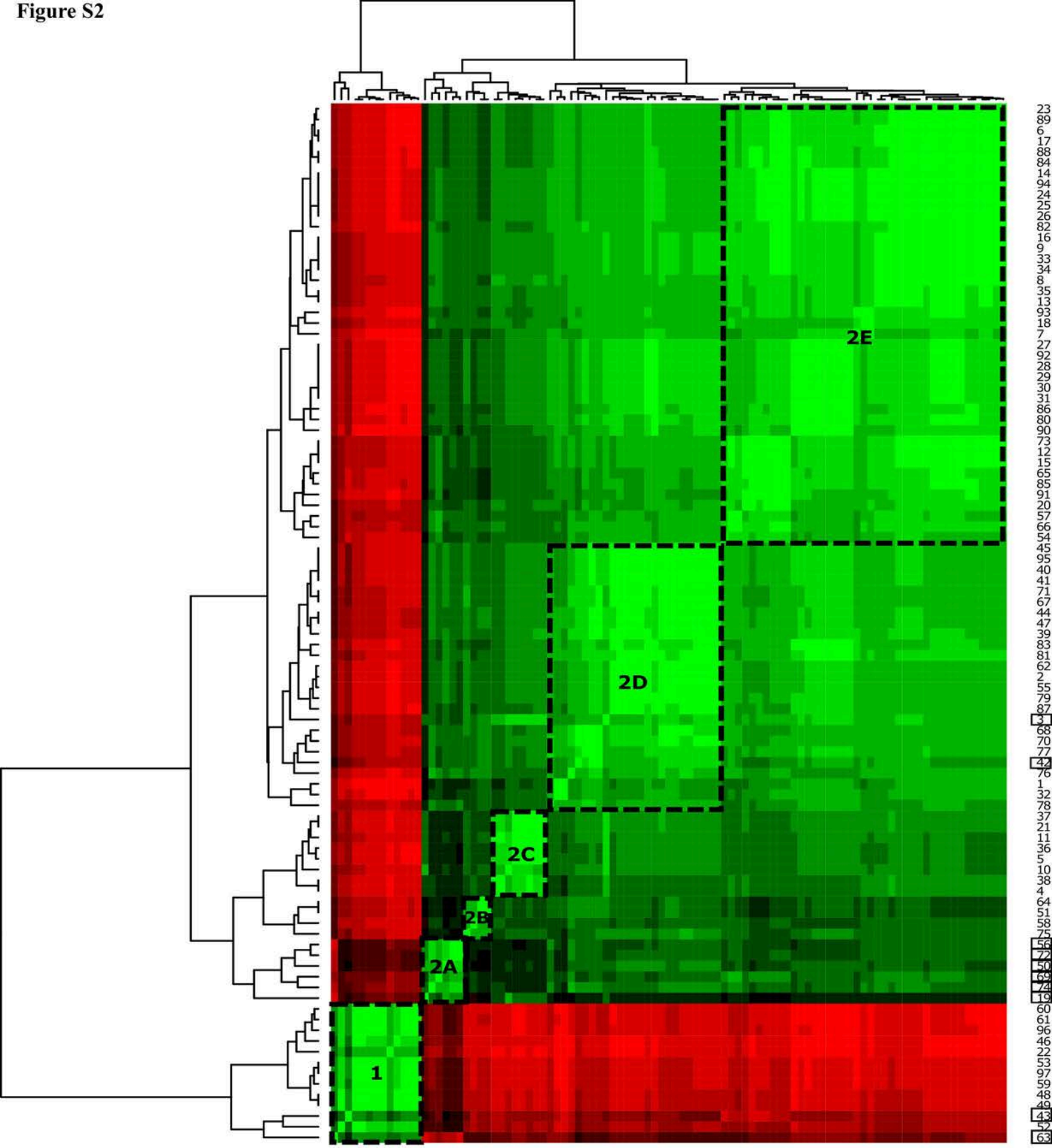


Figure S2 Heat map of similarity and differences of 5 concatenated housekeeping genes among 97 unique *E. ruminantium* sequences. Hypothetical groups and subgroups 1, 2A, G2B, 2C, 2D and 2D are marked by dashes. Degree of relatedness is indicated by colours from white (different) to red (similar). The name of each isolate was labelled on the right side of the graphs and corresponds to the same strains on the bottom of the graph. Dendograms representing the clusters were placed on the left side and on top of the graphs. Strains 3, 19, 42, 43, 50, 56, 63, 69, 72 and 74 are marked with a square and represent recombinants.

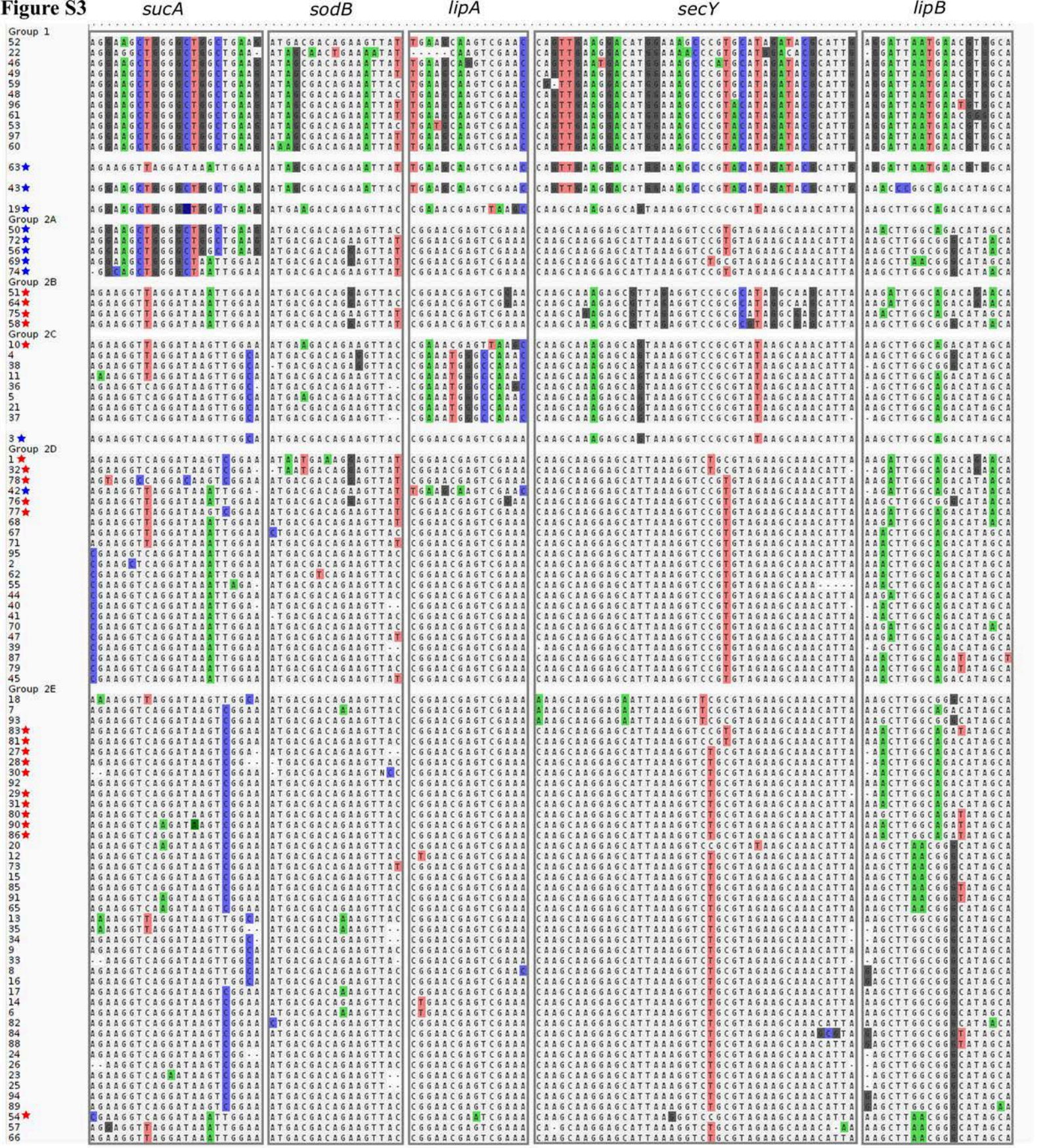


Figure S3 Multiple alignment of the different genotypes containing only variable positions. The subgroups clusters are separated by horizontal gaps, and the five different genes are boxed and separated by vertical gaps. Coloured residues represent non-consensus characters over the whole alignment. Blue stars indicate recombinant genotypes and red stars represent sequence types that are inferred to have less than 80% ancestry from a single population in STRUCTURE.

V. **Section III**

Ehrlichia ruminantium in Mozambique: a study on prevalence in ticks and isolate genetic diversity

(Draft in preparation for publication)

Ehrlichia ruminantium in Mozambique: a study on prevalence in ticks and isolate genetic diversity

Nídia Cangi[1,2,3], Laure Bournez[1,4], Jonathan Gordon[1], Rosalie Aprelon[1,4], Valérie Pinarello[1,4], Thierry Lefrançois[1,4], Luís Neves[2, 5] & Nathalie Vachiéry[1,4*]

1. CIRAD, UMR CMAEE, F-97170 Petit-Bourg, Guadeloupe, France

2. Centro de Biotecnologia-UEM, Eduardo Mondlane University, Av. de Moçambique, km 1.5, C.P. 257 Maputo, Mozambique.

3. Université des Antilles, F-97157 Pointe-à-Pitre, Guadeloupe, France

4. INRA, UMR CMAEE, F-34398 Montpellier, France

5. Department of Veterinary Tropical Diseases, University of Pretoria, Faculty of Veterinary Science P/Bag X04, Onderstepoort 0110, South Africa.

[*] Corresponding author: e-mail: nathalie.vachiery@antilles.inra.fr

Abstract

The tick species, *Amblyomma hebraeum* and *A. variegatum*, are the main vectors of Heartwater, a tropical infectious bacterial disease of ruminants caused by *Ehrlichia ruminantium*. In Mozambique, these tick species have a parapatric distribution, with *A. variegatum* present in the central and northern regions and *A. hebraeum* in the South. A narrow overlap area between the distributions of the two species occurs around parallel 22º south. In order to determine the prevalence of *E. ruminantium* in *A. hebraeum* and *A. variegatum* and to determine the genetic diversity and structure of isolates from different localities, cattle and wildlife were sampled across the south and center of Mozambique as well as in the Kruger National Park (KNP), South Africa. The prevalence of *E. rumimantium* in relation to the tick specie and locality, and correlation with tick abundance was analyzed. Afterward, Mozambican isolates were typed using Multi Locus Sequence Typing and the distribution of groups clustering genotypes was studied. In total, 722 *A. hebraeum* and 388 *A. variegatum* were collected from 31 localities and screened for *E. ruminantium*, using *pCS20* nested PCR and Sol1^{TM} qPCR. *E. ruminantium* tick prevalence in cattle varied from 0% [0-23.2 %] to 26.7% [12-45 %], with no infected ticks in 7 localities. In wildlife, prevalence was 8.2 [4-14.6 %] % in the KNP and 6.2% [0.2-30.2 %] in hunting concessions of Sofala province. However, no significant difference in prevalence was found between sampling sites and tick species, as well as no linear correlation between *E.*

ruminantium prevalence and tick abundance was observed. Most MLST genotypes from Mozambique clustered in subgroup 2C and 2E, which were present in same proportion in 5 out of 19 localities. Interestingly, MLST genotypes from group G1 and G2D were exclusively found in areas of *A. variegatum* distribution, while subgroup G2C was only detected in *A. hebraeum* areas. Moreover, genotypes from subgroup G2E were found in both *A. hebraeum* and *A. variegatum* areas. The high genetic diversity observed could be a problem to implement efficient vaccines, even if the same genotypes are widely distributed in the country. These results will contribute to a better understanding of *E. ruminantium* isolates spatial distribution in the studied regions and, therefore, to an improvement of heartwater monitoring and control strategies in Mozambique.

Key words: *Ehrlichia ruminantium, Amblyomma hebraeum, A. variegatum,* prevalence, heartwater

Introduction

Ehrlichia ruminantium is an obligate intracellular bacterium from the family Anaplasmataceae that causes a tropical infectious, virulent, transmissible and non-contagious disease named Heartwater or Cowdriosis affecting ruminants (1). The most important vectors of this disease are *Amblyomma hebraeum* in southern Africa and *A. variegatum,* a more widely distributed vector, transmitting the disease in the rest of Africa, and the Caribbean (2, 3).

Heartwater, together with East Coast fever and trypanosomosis are regarded as the most important vector-borne diseases of ruminants in Africa (4, 5). Furthermore, it is considered as a major obstacle to the introduction of improved ruminant breeds in Africa, as extremely high mortality and morbidity rates are particularly common when naïve animals are introduced in endemic areas (6). In the Southern Africa Development Community region, an annual expenditure of US$ 44.7 million was reported due to livestock production losses and the costs associated to disease control (acaricide and antibiotic treatments) (7, 8). In hyperendemic areas, only few clinical cases are generally observed (enzootic stability). This can be explained by the development of a certain immunity within the population when most animals are infected at a young age (during the period of time where they are less susceptible to the development of clinical disease), or when animals are challenged with low *E. ruminantium* infective doses (sub-lethal) and are then regularly challenged with infected ticks to maintain their immunity (9). The establishment of enzootic stability is, therefore, extremely dependent on local conditions such

as host density, tick abundance, prevalence of *E. ruminantium* in ticks, tick and host population dynamics seasonality and herd management (10, 11). Intensive acaricide usage, as performed in some commercial farms, can cause a rupture in herd immunity established over time and disease outbreaks might occur (12). Presently, no commercial vaccines are available in a wide scale, but intense research on the development of different vaccine candidates; attenuated, inactivated, and recombinant, is in progress (13). Those that were developed, showed a limited efficacy given the low levels of cross protection due to the high genetic diversity of *E. ruminantium* strains in the field (13). In this context, knowledge on the genetic diversity of *E. ruminantium* isolates at a regional scale is important to design efficient vaccines.

In Mozambique, Heartwater was first reported in 1969, but at that stage, mainly due to underdiagnoses, it was considered a non-important disease (14). The first serological surveys on heartwater were published approximately two decades later, showing that this disease was present throughout the country (15, 16). Currently, heartwater is considered to be a major cause of morbidity and mortality in ruminant production systems in Mozambique. High mortality rates are particularly common when ruminants coming from Tete province, Changara district, where *Amblyomma* ticks do not occur, are introduced into heartwater endemic areas in Gaza and Maputo provinces (16, 17). Several *E. ruminantium* serosurveys conducted in Mozambique have consistently indicated a substantial difference in seroprevalence between the central and southern regions of the country (15, 16, 18). Relatively high seroprevalence values were commonly found in the south regions of the country, varying between 56.3% (18) and 65.6% (16), contrasting with the central region, where the highest value was 10% (15).

However, serological diagnostic assays relied upon host antibody response and the length of the seropositive period in *E. ruminatium* is very short; from a few weeks to 3 months, depending on host species (13). Furthermore, a major constraint of the currently available diagnostic assays for *E. ruminantium* is their low specificity (19, 20). These limitations significantly hamper the use of serological assays as tools to investigate true epidemiological status.

Alternatively, the diagnosis of heartwater clinical cases and the accurate determination of tick infection by the use of reliable molecular tools (21, 22), should be explored and refined to be used as indicators of *E. ruminantium* intensity of circulation. Moreover, the combination of tick infection rate estimates with tick relative abundance values can be used to indirectly estimate the exposure of the animals to the pathogen and to infer herd epidemiological status regarding heartwater (23). In Mozambique, more than 80% of the domestic ruminant population is kept by small scale farmers. This is a resource constrained production system

characterized by deficient veterinary support and poor access to laboratory diagnostic facilities. Consequently, it is difficult to assess the number of clinical cases due to heartwater and hence to estimate the true economic impact of this disease in Mozambique. Additionally, due to the parapatric distribution of the main *Amblyomma* species in Mozambique, heartwater is exclusively transmitted by *A. hebraeum* to the south of Save River and *A. variegatum* in the central and northern areas of the country. *E. ruminantium* simultaneous transmission by these two tick species, occurs in a relatively small interface area located around parallel 22° south (24, 25). Taking into account the previously suggested differences in vector competence between *A. variegatum* and *A. hebraeum* (26) it would be interesting to genetically characterize *E. ruminantium* strains circulating in these parapatrically distributed tick species in Mozambique. *E. ruminantium* genetic diversity was previously studied in different regions using several approaches such as genotyping of polymorphic genes (27), MLVA (28, 29) or MLST (30, 31). A high genetic diversity was observed in the Caribbean (Guadeloupe), Burkina Faso and Gambia (27, 32, 33) with 9 to 11 *map-1* genotypes observed on restricted areas. More recently, MLST studies demonstrated *E. ruminantium* genotypic structure at the regional level (Burkina Faso), (31) and worldwide scale (34). Mozambican isolates typed by MLST were compared with isolates from West Africa, South Africa, the Caribbean and the Indian Ocean (34). High recombination events were highlighted, supporting the role of recombination in strain diversity. Many isolates appeared to be recombinant strains (34). Nevertheless, two genetic groups and 5 subgroups were defined: Group 1 clustering mainly West African isolates and Group 2 clustering worldwide isolates. Two subgroups were mainly associated with southern African isolates: subgroup 2C (Mozambique and Zimbabwe) and subgroup 2E with some isolates from the Indian Ocean and West Africa. Subgroup 2D clustered mainly West African and Caribbean isolates. However, the distribution of Mozambican isolates at locality level has not previously been shown and is detailed in the current study elucidating local genetic diversity. These data will be useful for designing further vaccines against heartwater.

In general, information on the epidemiology of heartwater and strain genetic diversity in Mozambique is scarce and requires an update. The aim of the study was to determine the prevalence of *E. ruminantium* in *A. hebraeum* and *A. variegatum* ticks and isolates genetic diversity and structure per locality, in southern and central Mozambique. This study represents the first extensive investigation on *E. ruminantium* conducted in Mozambique. The current results will contribute to the better understanding of *E. ruminantium* spatial distribution in the region and population genetic structure, which will help to implement heartwater monitoring and control strategies in Mozambique.

Material and methods

Sampling in cattle

Tick sampling from cattle was carried out between February 2012 and August 2014 in the Southern and Central regions of Mozambique, partly in association with a survey whose objectives were to determine *A. variegatum* and *A. hebraeum* tick distributions at the contact zone between these species (25). Cattle were examined at communal dip-tanks and corridors used for acaricide treatment by farmers or at commercial farms from 29 localities (Figure 1, Table 1). Two previous collections of *A. variegatum* ticks from cattle conducted in 2008, for other study purposes, in Tete province were also included in the study. At each locality, thirty ticks were randomly chosen among those collected and tested for the presence of *E. ruminantium*.

In order to have relatively comparable data, we only estimated tick abundance for places visited in February and March, which corresponds to the period of activity of adult ticks. When possible, a minimum of 50 animals within 5-10 km (approximate grazing range) from sampling locations were examined to estimate tick abundance levels. During clinical examination the total number of adults of each *Amblyomma* species on each animal was counted. Four herd infestation levels were defined according to the mean abundance of ticks (i.e. number of *Amblyomma* adults/number of animals examined): <0.1, (0.1 – 1), (1–10) and ≥ 10 adults *Amblyomma* per animal. We only included, in the analysis, animals that were treated with the acaricide product Amitraz, eight days or more prior to the visit, or those treated with pyrethroids 15 days or more prior the visit. These products were the only ones used on sampled farms. Eight days and fifteen days represent the mean duration of residual effects of Amitraz and pyrethroids on hosts respectively and the shortest-time needed for attraction and attachment of ticks on hosts (35).

The number of ticks to be tested for *E. ruminantium* was determined assuming an infinite population and an expected *E. ruminantium* prevalence of 5%, a confidence level of 95% and a precision of 8% (36). In order to test 30 ticks per locality for the presence of *E. ruminantium*, five to ten adults of *A. hebraeum* and/or *A. variegatum* ticks were collected from the first 32 infested animals that arrived at the dip-tank. In order to avoid a host effect at the adult-stage, when possible, we chose to analyse only one or two ticks per animal, randomly chosen among those collected.

Sampling in wild ruminants

In addition, ticks were sampled from wild animals from hunting concessions in Sofala Province (Coutada 11 and 12) in the central part of the country and in the Kruger National Park (KNP, South Africa) bordering Gaza and Maputo provinces in the southern part of Mozambique. The KNP was selected to be used as a proxy of the southern regions of Mozambique, given its geographical proximity, ecological similarities, wildlife density and logistic conditions. All the visible ticks were collected from hunted animals, following the procedures and ethical regulations of the hunting concessions, and were tested for the presence of *E. ruminantium*. Only known hosts of *E. ruminantium* or *Amblyomma* ticks were sampled (37). The collection was conducted during the hunting season in Coutada 11 and 12 for Mozambique, between June and November 2012 and 2014. In the KNP, where *A. hebraeum* adult stages are present throughout the year (38), collection was conducted in August 2014.

Ticks collected from each animal were preserved in 70% ethanol and stored until taxonomic confirmation and futher DNA extraction.

DNA extraction

DNA was extracted as previously described (22). Briefly, ticks were individually grinded using a Tissue lyser II (Qiagen, France). A 5 mm steel bead was added to a 2 ml Eppendorf tube (Eppendorf, France), containing a tick and kept at -80ºC for at least 2 hours or preferably overnight. The tick was disrupted twice in the Tissue lyser II at an oscillation frequency of 30 Hz for 2 minutes. The mashed tick was then resuspended in 450 µl of sterile PBS, vortexed and centrifuged twice at 8000 rpm for 30 sec. The supernatant was recovered for nucleic acid extraction.

DNA was extracted either manually or automatically following recommendations from Cangi *et al.*, (2016 submitted), (22). DNA was extracted with the QiaAmp DNA minikit (Qiagen, Courtaboeuf, France) according to the manufacturer's instructions with a slight adjustment: tick samples weighing 25 to 40 mg were lysed with 180 µl of buffer ATL and 20 µl of RNase A at 20 mg/ml (Sigma-Aldrich, France).

The automatic DNA extraction was performed using the Biomek 4000 automated liquid handling robot (Beckman Coulter) and the "Viral RNA and DNA from "Macherey-Nagel" kit in a 96-well plate format, as described previously (22). The final elution of nucleic acids was performed by sequentially applying 100 and 50 µl of nuclease-free water. After extraction, the plate containing the nucleic acids was stored at -20°C until use. The quality control of tick DNA extracted was tested using 16S rDNA qPCR for ticks.

E. ruminantium tick screening by nested and qPCR and MLST characterization

Molecular diagnostic of *E. ruminantium* in ticks was performed using both *pCS20* nested PCR and Sol1^{TM} qPCR, as previously described by Molia *et al.* (2008), (21) Furthermore, the genetic characterization of detected isolates was performed by MLST as described previously by Cangi and Gordon *et al*., (2016), (34). Mozambican isolates are described in this previous study and all MLST sequences were deposited in GeneBank and accession numbers are available (34). In the current study, Mozambican genotypes and their respective worldwide genetic groups and subgroups (group G1, G2A, G2B, G2C, G2D and G2E) were analyzed and linked with their origin in different regions of Mozambique.

Descriptive analysis

Prevalence of *E. ruminantium* in *Amblyomma* ticks was computed at site level and confidence intervals were calculated using exact binomial law. Sampling sites less than 10 km distance apart were aggregated.

The effects of tick species, type of farm and tick abundance (classified into three categories) on *E. ruminantium* presence in ticks were analysed with a generalised linear model, using a binomial probability distribution where localities were included as a random effect.

All statistical analyses were performed with R version 3.1.2 (39).

Results

E. ruminantium tick prevalence

In total, , 722 *A. hebraeum* and 388 *A. variegatum* (1110 ticks), collected from cattle and wild ruminants from 30 localities in 5 provinces from the center and south of Mozambique (Tete, Manica, Sofala, Inhambane and Maputo) and one locality in South Africa (KNP), were screened for *E. ruminantium* (Table 1, Figure 1). The sampled localities included 7 commercial farms, 21 sites with *A. hebraeum* (16 sites with more than 10 ticks tested) and 16 sites with *A. variegatum* (14 sites with more than 10 ticks tested). We found co-occurrence and collected both species at 6 sites, but there was only one site where more than 10 adult ticks for each species were collected and tested. Among the 20 localities for which tick abundance was estimated, mean infestation level was relatively low in most of them (15/20 places), with less than 10 ticks/animal, which includes five places with extremely low tick abundance (less than 1 tick/animal).

In wildlife sampling locations, 122 *A. hebraeum* ticks were collected from buffaloes (*Syncerus caffer*) in the KNP and 15 *A. variegatum* were collected from buffaloes or antelopes such as Reedbuck (*Redunca arundinum*), Bushbuck (*Tragelaphus sylvaticus*) and Sable (*Hippotragus niger*), (37) in hunting concessions of Sofala-Mozambique.

We collected and analysed at least 15 ticks for 30 localities. The prevalence of *E. ruminantium* in ticks per locality varied from 0% [0-23.2 %] to 26.7% [12-45 %] (Table 1). In the provinces with the highest sampling intensity, Manica and Inhambane, prevalence was around 0% [0-11.9 %] to 26.7% [12.3-45.9 %] and around 0% [0-23.2 %] to 20% [7.7-38.6 %], respectively. While for Tete, Sofala and Maputo the highest prevalence observed was 6.9% [0.8-22.8 %], 11.1% [2.3-29.1 %] and 8.6% [3.5-17 %], respectively.
No infected ticks were found in 7 localities, which means that the prevalence of *E. ruminantium* in these ticks was lower than 15% or 20% in those localities according to the number of ticks tested. Regarding wildlife sampling locations, the prevalence was 8.2% [4- 14.6 %] in the KNP and 6.2% [0.2-30.2 %] in hunting concessions of Sofala.

There was no significant difference of prevalence (t-test=1.57, p=0.13) between places infested by *A. variegatum* (median 6.5%, interval interquartile 25-75 % [0-10.3 %]) and those infested by *A. hebraeum* (median 9.8%, interval interquartile 25-75 % [6.7-18.3 %]). However, a slight tendency for lower prevalences in *A. variegatum* localities was observed, which was also associated to lower tick abundances (Figure 3). In the site where both species were found at equal moderate abundances, the prevalence of *E. ruminantium* was of 6.7% [0.8- 22.1 %] in *A. variegatum* adults and 19.4% [8.2-36 %] in *A. hebraeum* adults, but the difference was not statistically significant.
There was no linear correlation between *E. ruminantium* prevalence and tick abundance. However, lower prevalence values were found in places with very low abundance, i.e. inferior to one tick per animal, (n=5, median 0%, interval interquartile 25-75m% [0-6.7 %]) compared to others (n=15, median 13.6%, interval interquartile 25-75 % [10-19 %], t-test=-4.82, $p<0.001$) (Figure 3). There was no significant difference in prevalence in areas with tick abundance of 1-10 ticks per animal compared to those with a tick abundance higher than 10 ticks per animal (t-test=0.71, p=0.49).

Genetic diversity and structure of *E. ruminantium* isolates per locality

From 95 *E. ruminantium* positive samples, 57 isolates were successfully typed and 38 were not successful typed by MLST. Some positive samples could not be amplified by MLST or only partially, for one gene (*LipA*). From 57 *E. ruminantium* isolates, 39 were unique genotypes excluding clones (Table 2). As some unique genotypes were very close, with only one to 3 substitutions, only 30 were considered true different genotypes (Table 2). The distribution of genotypes per group and sub-groups is shown in Figure 4. Mozambican strains clustered in four genetic groups G1, G2C, G2D and G2E. Group G2E clustered 57% (33/57) of isolates with the presence of 25 unique genotypes and Group G2C clustered 32% (18/57) including only 8 unique genotypes (Figure 4). The remaining groups, Group G2D and Group G1 clustered only 5 and 1 isolatess, respectively. In table 2, the number of isolates and genotypes were shown per locality. The number of genotypes per locality varied from 1 (corresponding to one positive tick) to 5 when excluding recombinant genotypes (identified by *) and genetically close isolates (identified by grey shade, i.e. with only one to 3 substitutions) as shown by Cangi and Gordon *et al.*, (2016), (34), (Table 2). The geographical distribution of genetic groups and the total number of typed isolates per locality are shown in Figure 5. The two subgroups G2C and G2E are represented equally in 5 out of 19 localities with no geographical clustering (Figure 5). Surprisingly, in these 5 sites (Espungabera, Mahiza, Majuacuana, Mapinhane and Vulanjane) only 2 tick samples were typed, containing one genotypes from G2C and one from G2E, highlighting the the similar composition of subgroup populations (Table 2, Figure 5). For 10 localities, there was predominance of either subgroup G2C or G2E. In 3 localities, there was dominance of one subgroup, G2C or G2E (Figure 5). In Mambone, there were mainly 6 isolates from G2E, including 5 recombinant genetically close genotypes (Table 2). There was also detection of one genotype from G2D.

Genotype 93, identical to the reference strain Welgevonden (from G2E), genotype 10 and genotype 5 (from G2C) were widely detected in 4 localities (Table 2, figure 6). Genotypes 6, 8 and 16 were detected only in 2 localities (Table 2, figure 6). Identical genotypes were present even for distant localities like genotype 10 and 93 (300 km between 2 sites) or genotype 5 (600km between 2 sites). Interestingly, *E. ruminantium* genotypes from G2E were typed both on *A. variegatum* and *A. hebraeum*, whereas genotypes from G2D were found only in *A. variegatum* and G2C only in *A. hebraeum* (Table 2). One genotype from G1 was detected in Govuro, where only *A. variegatum* is present.

Discussion

In this study, we estimated the prevalence of *E. ruminantium* in adults of *A. hebraeum* and *A. variegatum* ticks in several localities of Mozambique and from KNP in South Africa. Our results indicate the presence of the bacterium in all regions included in the study, with evidence of relatively high circulation (with a prevalence superior to 10% in adult *Ambyomma* ticks) in several localities. The presence of *E. ruminantium* was detected in 23 out of 30 localities, which either reflects a true absence or a low prevalence level in ticks in 7 localities. It is interesting to note that two of those sites were commercial farms where acaricide treatments are probably more intensive.

Several factors can influence the level of *E. ruminantium* prevalence in adult ticks, such as host density, host species, farming practices (acaricide treatment, pastures areas and use), environmental conditions, vector species and density (11, 40, 41). It was not the objective of this study to evaluate the influence of each of these factors on heartwater prevalence. It is, however, worth noting that there was no significant difference on *E. ruminantium* estimated prevalence in adult ticks between localities where only *A. variegatum* was present compared to those with only *A. hebraeum*. This confirms the major role of both species as the main vectors of *E. ruminantium* and suggests that both species might have similar susceptibility to infection by *E. ruminantium* and potential for its transmission. However, observations from studies in southern Africa indicate differences in vector competence between *A. hebraeum* and *A. variegatum* ((26, 40, 42).

E. ruminantium tick prevalence results obtained in the current study are consistent with those described in other studies in African countries and the Caribbean, which report prevalences between 3-20 % in *A. variegatum* and *A. hebraeum* ticks (21, 43-45). Adakal *et al.* (2010) described an overall prevalence of 3.65% in *A. variegatum* and no significant difference in this parameter in three villages studied in Burkina Faso or between different tick life stages and sex in the same locations (45). Faburay *et al.* (2007) reported an overall prevalence of 16.6% in *A. variegatum* collected from Gambia, using nested *pCS20* PCR and a heartwater prevalence gradient between regions (44). The highest prevalence was described by Molia *et al.* (2008) with 19.1% of *A. variegatum* infected with *E. ruminantium* in Marie Galante (Caribbean), (21). Data on *E. ruminatium* prevalence in *A. hebraeum* is scarce and localised. Peter *et al.* (1999), using a *pCS20* probe, described a prevalence between 8.5 - 11.2 % in *A. hebraeum* collected from the lowveld and highveld of Zimbabwe (43). The majority of studies on *E. ruminantium* prevalence in ticks are cross sectional in nature. Given the biological dynamics of tick

populations and the complex interaction with their hosts, it would be of epidemiological interest to evaluate seasonal variations of *E. ruminantium* tick prevalence.

In wildlife sampled in localities with no contact with domestic animals, we found an *E. ruminantium* prevalence in adult ticks of 6.2% [0.2-30.2 %] and 8.2 % [4-14.6 %], which is similar to those previously reported by Peter *et al.* (1999) in KNP and Allsopp *et al.* (1999) who used DNA probes to measure prevalence in KNP (46, 47).

In this study, we estimated the intensity of circulation of *E. ruminantium* in host populations by calculating its prevalence in adult *Amblyomma* ticks. As only trans-stadial transmission of *E. ruminantium* was evidenced to occur in *Amblyomma* ((48), detection of the bacteria in adult ticks can reflect an infection that can have occurred during feeding on infected hosts of either one of the three life stages of the tick (larvae, nymphs and adults). Maintenance of enzootic stability, which is currently considered to be the most sustainable control method for endemic areas, seems to be driven by vertical and trans-stadial transmission, continuous feeding of males and maintenance of infectivity after infecting a host ((9, 49). Considering the current methodological constraints in measuring enzootic stability, it may be of interest to explore the combined use of prevalence in ticks and tick abundance values as means to monitor *E. ruminantium* epidemiological status.

Lower prevalence values were found in places with very low tick abundance. In fact, lower tick densities decrease the probability of tick-host contact, thus reducing heartwater tick infection and transmission. There are few published studies on the distribution of *A. variegatum* and no studies on *E. ruminantium* prevalence based on molecular detection in Mozambique. During our field sampling, we observed an equal infestation of cattle by *A. hebraeum* and *A. variegatum*, except for few *A. variegatum* localities in the central area, where cattle density was very low. Thus, in the majority of studied locations, it seems that vector density is sufficient to maintain the *E. ruminantium* life cycle.

The lack of MLST amplification for 38 *E. ruminantium* positive samples, confirmed by nested *pCS20* PCR and *pCS20* Sol1^{TM} qPCR stressed the high genetic diversity of *E. ruminantium* strains from Mozambique. It is worth emphasizing that the sensitivity and specificity of nested *pCS20* PCR and *pCS20* Sol1^{TM} qPCR, have previously been demonstrated, which validated our results (21). Several strains could not be amplified or only one gene was amplified such as *LipA* which is quite a conserved gene. This phenomenon was observed in a previous paper on the development of *pCS20* Sol1^{TM} qPCR comparing the qPCR with nested *pCS20* PCR and MLST (22). Specifically, none of the 10 positive samples collected from wildlife in KNP could

be typed, even considering that two samples gave a strong positive result (Ct value around 30). The sequence quality was poor, probably due to unspecific product or mixed isolates within one tick. . Since limited information is available on wildlife *E. ruminantium* strains, it would be interesting to try to genotype isolates using *map1*/RFLP method. Alternatively, directly sequencing the whole *E. ruminantium* genome from tick samples could be considered. However, the latter approach should take into account the constraint of *E. ruminantium* DNA limited quantity and quality.

Genotypes from G2E were strongly represented compared to other subgroups. This genetic group was previously defined as clustering of Southern Africa and Indian Ocean isolates with few genotypes from West Africa and only one genotype from the Caribbean (34). There were several recombinant genotypes, especially in Mambone_Maninga locality, which probably indicates clonal and genetic diversity expansion. Mambone has a high cattle density with records of introduction from the North, which could explain the presence of recombinant genotypes with newly introduced strains which co-infected cattle with local strains. This could also explain the presence of one genotype from G2D subgroup only found in the North of the studied region. This zone is the limit of *A. variegatum* distribution and interestingly, genotypes from G2D were only detected in *A. variegatum*. More widely, from a previous study, genotypes from G2D, clustering mainly West African and Caribbean isolates, were also observed on *A. variegatum* samples from Burkina Faso, Comoros and Madagascar (34). In Guadeloupe, genotypes from G2D were found in the blood of animals as only *A. variegatum* is present in the Caribbean. These facts can explain the presence of genotypes from G2D only in the Central regions of Mozambique where *A. variegatum* is present. The same observation applies to the genotype from Group 1 clustering mainly West Africa isolates, which was only detected once in *A. variegatum* from Mozambique. Detection of low number of genotypes from G2D and from G1 in Mozambique compared to West Africa or Caribbean reflects recent introduction in the region as hypothesized in a previous study (34). One hypothesis to explain the presence of genotypes from G1 and G2D only in *A. variegatum* is the lack of adaptation of genotypes to *A. hebraeum* due to their recent introduction in the region. Another hypothesis is the predominance of already established genotypes in *A. hebraeum* areas, compared to recently introduce ones. It could be interesting to isolate these strains from *A. variegatum in vivo* and see if they can infect *A. hebreaum*. Furthermore, it will be interesting to compare the vector competence of *A. variegatum* and *A. hebraeum* by testing their ability to transmit isolates from the South and North of Mozambique and vice-versa. Group 2C clusters mainly Mozambican isolates, including Umpala reference strain and one strain from Zimbabwe as shown by Cangi

406 and Gordon *et al.* (2016), these genotypes are not present in West Africa (34). A mixed
407 population of G2C and G2E in 5 localities was shown, highlighting the well established
408 presence of genotypes from both sub-groups originated from Southern Africa.
409 From this study, we showed that there is a high genetic diversity in Mozambique with 39 unique
410 genotypes detected. However, a high level of genetic similarity was observed. From these, 30
411 genotypes had a high proportion of recombinant strains especially in Mambone. Compared to
412 previous studies performed in the Caribbean, Burkina Faso and Gambia, mainly based on *map-*
413 *1* genotypes, which is known to be more polymorphic than MLST genes, the number of
414 different genotypes is almost three fold higher (44, 45, 50). However, there were many identical
415 or similar genotypes in distant localities. Genotype 93, identical to Welgevonden was widely
416 present. Better knowledge of the genetic diversity of isolates in Mozambique could be
417 informative for the design of control methods against heartwater, even if there is no correlation
418 between genotypes defined by MLST and cross protection between isolates.
419 It could be possible, first, to use the live attenuated vaccine Welgevonden in Mozambique, as
420 it is already present and it appeared to be protective against South African strains, such as
421 Ball3, Mara and Blaukraans reference strains (51). If further isolation of local strains *in vitro*
422 is possible, it would be interesting to prepare an inactivated vaccine including several local
423 strains i.e. genotypes which are predominant. Thus, isolation of local strains and genotypes for
424 the development of regional vaccines could be a great contribution to the control of heartwater
425 in Mozambique.
426
427
428 **Conclusion**
429 The prevalence of *E. ruminantium* in ticks per locality in Mozambique varied from 0% [0-23.2
430 %] to 26.7% [12-45 %]. There was 7 localities without infected ticks or with low prevalence.
431 For wildlife sampling locations, the prevalence was 8.2% [4-14.6 %] in the KNP and 6.2%
432 [0.2-30.2 %] in the hunting concessions of Sofala. No statistical significant difference of
433 prevalence was found between tick species and between localities. There was no linear
434 correlation between *E. ruminantium* prevalence and tick abundance. Both *Amblyomma* tick
435 species have a high bacteria prevalence, demonstrated the importance of the two tick species
436 to transmit *E. ruminantium* in Mozambique.
437 MLST genotypes from group 1 and 2D were exclusively found in areas of *A. variegatum*
438 distribution, while group 2C was only detected in *A. hebraeum* areas. Moreover, genotypes
439 from group 2E were found in both *A. hebraeum* and *A. variegatum* areas. Thirty nine unique

genotypes were found in Mozambique. Better knowledge on the genetic diversity of *E. ruminantium* in Mozambique, could be used to improve the current strategies for heartwater control.

Funding

This work was financially supported by CIRAD and EPIGENESIS project which received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement No 31598". FUNDO ABERTO DA UEM 2012-2013 and FUNDO NACIONAL DE INVESTIGAÇÃO Projecto Nº 133-Inv/FNI/ 2012-2013 funded the field trips and reagents in Mozambique.

Acknowledgments

We are grateful to our colleagues H. Mucache and N. Vaz for all the support during sampling and critical comments on the manuscript. We are thankful to all the heads of Veterinary Services and technicians for logistic support during the sampling in Maputo, Inhambane, Sofala and Manica provinces. Additionally, we thank S. Afonso for providing tick samples from Tete province. Finally, we are thankful to the Veterinary service group and SANParks staff for their help and support during tick sampling in the KNP related to the project "CANGIN 1150 Genetic diversity of *Ehrlichia ruminantium*".

References

1. **Dumler, JS, Barbet, AF, Bekker, CP, Dasch, GA, Palmer, GH, Ray, SC, Rikihisa, Y, Rurangirwa, FR.** 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 'HGE agent' as subjective synonyms of *Ehrlichia phagocytophila*. Int. J. Syst. Evol. Microbiol. **51:**2145-2165.

2. **Walker, JB, Olwage, A.** 1987. The tick vectors of *Cowdria ruminantium* (Ixodoidea, Ixodidae, genus *Amblyomma*) and their distribution. Onderstepoort J. Vet. Res. **54:**353-379.

3. **Stachurski, F, Tortosa, P, Rahajarison, P, Jacquet, S, Yssouf, A, Huber, K.** 2013. New data regarding distribution of cattle ticks in the south-western Indian Ocean islands. Vet. Res. **44:**79. doi: 10.1186/1297-9716-44-79.

4. **Uilenberg, G.** 1983. Heartwater (Cowdria ruminantium infection): current status. Adv. Vet. Sci. Comp. Med. **27:**427-480.

5. **Provost, A, Bezuidenhout, JD.** 1987. The historical background and global importance of heartwater. Onderstepoort J. Vet. Res. **54:**165-169.

6. **Uilenberg, G.** 1982. Experimental transmission of *Cowdria ruminantium* by the Gulf coast tick Amblyomma maculatum: danger of introducing heartwater and benign African theileriasis onto the American mainland. Am. J. Vet. Res. **43:**1279-1282.

7. **Minjauw, B.** 2000. The economic impact of heartwater (*Cowdria ruminantium*) infection in the SADC region, and its control through the use of new inactivated vaccines. In Proc. of the 9th Symposium of the International Society for Veterinary Epidemiology and Economics (ISVEE 9), Breckenridge, Colorado. Economics & livestock production session, 645. Available at: www.sciquest. org.nz/elibrary/edition/5415 (accessed on 26 May 2015).

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

9. **Deem, SL, Noval, R, Yonow, T, Peter, TF, Mahan, SM, Burridge, MJ.** 1996. The epidemiology of heartwater: establishment and maintenance of endemic stability. Parasitology Today. **12:**402-405. doi: doi.org/10.1016/0169-4758(96)10057-0.

10. **Tice, GA, Bryson, NR, Stewart, CG, Du Plessis, B, De Wall, DT.** 1998. The absence of clinical disease in cattle in communal grazing areas where farmers are changing from an intensive dipping programme to one of endemic stability to tick-borne diseases. Onderstepoort J. Vet. Res. **65:**169-175.

11. **Smith, RD, Evans, DE, Martins, JR, Cereser, VH, Correa, BL, Petraccia, C, Cardozo, H, Solari, MA, Nari, A.** 2000. Babesiosis (*Babesia bovis*) stability in unstable environments. Ann. N. Y. Acad. Sci. **916:**510-520. doi: 10.1111/j.1749-6632.2000.tb05330.x.

12. **Walker, AR.** 2011. Eradication and control of livestock ticks: biological, economic and social perspectives. Parasitology. **138:**945-959. doi: 10.1017/S0031182011000709.

13. **Vachiéry, N, Marcelino, I, Martinez, D, Lefrancois, T.** 2013. Opportunities in diagnostic and vaccine approaches to mitigate potential heartwater spreading and impact on the American mainland. Dev. Biol. (Basel). **135:**191-200. doi: 10.1159/000190050..

14. **Valadão, FG.** 1969. Occurrence of hyperacute heartwater in Mozambique and problems of premunition against this disease. Annls Serv. Vet. Moçambique. **12–14:**85-90.

15. **Asselbergs, M, Jongejan, F, Langa, A, Neves, L, Afonso, S.** 1993. Antibodies to *Cowdria ruminantium* in Mozambican goats and cattle detected by immunofluorescence using endothelial cell culture antigen. Trop. Anim. Health Prod. **25:**144-150.

16. **Bekker, CP, Vink, D, Lopes Pereira, CM, Wapenaar, W, Langa, A, Jongejan, F.** 2001. Heartwater (*Cowdria ruminantium* infection) as a cause of postrestocking mortality of goats in Mozambique. Clin. Diagn. Lab. Immunol. **8:**843-846. doi: 10.1128/CDLI.8.4.843-846.2001.

17. **Bila, C.J.,De Deus,N.,Fafetine,J.M.,Dimande,A., Neves,L.** 2003. Investigação preliminar sobre as causas de mortalidade de caprinos provenientes de Tete no Sul de Moçambique, Terceiro seminário de investigação, Universidade Eduardo Mondlane, p. 31-36..

18. **Atanasio, A.** 2000. Helminths, protozoa, heartwater and the effect of gastro-intestinal nematodes on productivity of goats of the family sector in Mozambique. PhD thesis, Medical University of Southern Africa, South Africa.

19. **van Vliet, AH, van der Zeijst, BA, Camus, E, Mahan, SM, Martinez, D, Jongejan, F.** 1995. Use of a specific immunogenic region on the Cowdria ruminantium MAP1 protein in a serological assay. J. Clin. Microbiol. **33:**2405-2410.

524 20. **Katz, JB, DeWald, R, Dawson, JE, Camus, E, Martinez, D, Mondry, R.** 1997.
525 Development and evaluation of a recombinant antigen, monoclonal antibody-based
526 competitive ELISA for heartwater serodiagnosis. J. Vet. Diagn. Invest. **9:**130-135.

527 21. **Molia, S, Frebling, M, Vachiéry, N, Pinarello, V, Petitclerc, M, Rousteau, A, Martinez,**
528 **D, Lefrancois, T.** 2008. *Amblyomma variegatum* in cattle in Marie Galante, French Antilles:
529 prevalence, control measures, and infection by *Ehrlichia ruminantium*. Vet. Parasitol. **153:**338-
530 346. doi: 10.1016/j.vetpar.2008.01.046.

531 22. **Cangi, N, Pinarello, V, Bournez, L, Lefrançois, T, Abina, E, Neves, L, Vachiery, N.**
532 2016 submitted. Efficient high-throughput molecular method to detect *Ehrlichia ruminantium*
533 in ticks. Manuscript submitted for publication.

534 23. **Quintao-Silva, MG, Ribeiro, MF.** 2003. Infection rate of Babesia spp. sporokinetes in
535 engorged *Boophilus microplus* from an area of enzootic stability in the State of Minas Gerais,
536 Brazil. Mem. Inst. Oswaldo Cruz. **98:**999-1002. doi: 10.1590/s0074-02762003000800003.

537 24. **Dias, JATS.** 1991. Some data concerning the ticks (Acarina-Ixodoidea) presently known
538 in Mozambique. Garcia De Orta Série De Zoologia. **18 (1-2):**27-48.

539 25. **Bournez, L, Cangi, N, Lancelot, R, Pleydell, DR, Stachurski, F, Bouyer, J, Martinez,**
540 **D, Lefrancois, T, Neves, L, Pradel, J.** 2015. Parapatric distribution and sexual competition
541 between two tick species, *Amblyomma variegatum* and *A. hebraeum* (Acari, Ixodidae), in
542 Mozambique. Parasit. Vectors. **8:**504. doi: 10.1186/s13071-015-1116-7.

543 26. **Mahan, SM, Peter, TF, Semu, SM, Simbi, BH, Norval, RA, Barbet, AF.** 1995.
544 Laboratory reared *Amblyomma hebraeum* and *Amblyomma variegatum* ticks differ in their
545 susceptibility to infection with Cowdria ruminantium. Epidemiol. Infect. **115:**345-353. doi:
546 doi.org/10.1017/s0950268800058465.

547 27. **Raliniaina, M, Meyer, DF, Pinarello, V, Sheikboudou, C, Emboulé, L, Kandassamy,**
548 **Y, Adakal, H, Stachurski, F, Martinez, D, Lefrançois, T, Vachiéry, N.** 2010. Mining the
549 genetic diversity of *Ehrlichia ruminantium* using map genes family. Vet. Parasitol. **167:**187-
550 195. doi: 10.1016/j.vetpar.2009.09.020.

28. **Pilet, H, Vachiéry, N, Berrich, M, Bouchouicha, R, Durand, B, Pruneau, L, Pinarello, V, Saldana, A, Carasco-Lacombe, C, Lefrancois, T, Meyer, DF, Martinez, D, Boulouis, HJ, Haddad, N.** 2012. A new typing technique for the Rickettsiales *Ehrlichia ruminantium*: multiple-locus variable number tandem repeat analysis. J. Microbiol. Methods. **88:**205-211. doi: 10.1016/j.mimet.2011.11.011.

29. **Nakao, R, Morrison, LJ, Zhou, L, Magona, JW, Jongejan, F, Sugimoto, C.** 2012. Development of multiple-locus variable-number tandem-repeat analysis for rapid genotyping of *Ehrlichia ruminantium* and its application to infected *Amblyomma variegatum* collected in heartwater endemic areas in Uganda. Parasitology. **139:**69-82. doi: 10.1017/S003118201100165X.

30. **Adakal, H, Meyer, DF, Carasco-Lacombe, C, Pinarello, V, Allegre, F, Huber, K, Stachurski, F, Morand, S, Martinez, D, Lefrancois, T, Vachiéry, N, Frutos, R.** 2009. MLST scheme of *Ehrlichia ruminantium*: genomic stasis and recombination in strains from Burkina-Faso. Infect. Genet. Evol. **9:**1320-1328. doi: 10.1016/j.meegid.2009.08.003.

31. **Nakao, R, Magona, JW, Zhou, L, Jongejan, F, Sugimoto, C.** 2011. Multi-locus sequence typing of *Ehrlichia ruminantium* strains from geographically diverse origins and collected in *Amblyomma variegatum* from Uganda. Parasit. Vectors. **4:**137. doi: 10.1186/1756-3305-4-137.

32. **Faburay, B, Jongejan, F, Taoufik, A, Ceesay, A, Geysen, D.** 2008. Genetic diversity of *Ehrlichia ruminantium* in *Amblyomma variegatum* ticks and small ruminants in The Gambia determined by restriction fragment profile analysis. Vet. Microbiol. **126:**189-199. doi: 10.1016/j.vetmic.2007.06.010.

33. **Adakal, H, Stachurski, F, Konkobo, M, Zoungrana, S, Meyer, DF, Pinarello, V, Aprelon, R, Marcelino, I, Alves, PM, Martinez, D, Lefrancois, T, Vachiéry, N.** 2010. Efficiency of inactivated vaccines against heartwater in Burkina Faso: impact of *Ehrlichia ruminantium* genetic diversity. Vaccine. **28:**4573-4580. doi: 10.1016/j.vaccine.2010.04.087.

34. **Cangi, N, Gordon, JL, Bournez, L, Pinarello, V, Aprelon, R, Huber, K, Lefrancois, T, Neves, L, Meyer, DF, Vachiery, N.** 2016. Recombination Is a Major Driving Force of Genetic Diversity in the Anaplasmataceae *Ehrlichia ruminantium*. Front. Cell. Infect. Microbiol. **6:**111. doi: 10.3389/fcimb.2016.00111..

35. **Barré, N, Pavis, C.** 1992. Essai d'attraction d'*Amblyomma variegatum* (Acarina: Ixodina) sur des bovins préalablement traités avec des phéromones d'agrégation-fixation et un acaricide pyréthrinoïde. Revue d'Élevage Et De Médecine Vétérinaire Des Pays Tropicaux. **45:**33-36.

36. **Cannon, RM.** 2001. Sense and sensitivity-designing surveys based on an imperfect test. Prev. Vet. Med. **49:**141-163. doi: 10.1016/s0167-5877(01)00184-2.

37. **Peter, TF, Burridge, MJ, Mahan, SM.** 2002. *Ehrlichia ruminantium* infection (heartwater) in wild animals. Trends Parasitol. **18:**214-218. doi: 10.1016/s1471-4922(02)02251-1.

38. **Horak, I, Fourie, L, Van Zyl, J.** 1995. Arthropod parasites of impalas in the Kruger National Park with particular reference to ticks. S. Afr. J. Wildl. Res. **25:**123-126. doi: doi.org/10.4102/koedoe.v38i1.306.

39. **R Core Team.** 2013. R: A language and environment for statistical computing. http://www.R-project.org/. R Foundation for Statistical Computing, Vienna, Austria.

40. **Norval, RAI.** 1983. The ticks of Zimbabwe VII. The genus *Amblyomma*. Zimbabwe Veterinary Journal. **14:**5-18.

41. **Estrada-Peña, A, de la Fuente, J.** 2014. The ecology of ticks and epidemiology of tick-borne viral diseases. Antiviral Res. **108:**104-128. doi: doi.org/10.1016/j.antiviral.2014.05.016.

42. **Karrar, G.** 1986. Epizootiological studies on heartwater in the Sudan. **Sudan J Vet Sci Anim Husb, 9:328-343:**.

43. **Peter, TF, Perry, BD, O'Callaghan, CJ, Medley, GF, Mlambo, G, Barbet, AF, Mahan, SM.** 1999. Prevalence of *Cowdria ruminantium* infection in *Amblyomma hebraeum* ticks from heartwater-endemic areas of Zimbabwe. Epidemiol. Infect. **123:**309-316. doi: 10.1017/s0950268899002861.

44. **Faburay, B, Geysen, D, Munstermann, S, Taoufik, A, Postigo, M, Jongejan, F.** 2007. Molecular detection of *Ehrlichia ruminantium* infection in *Amblyomma variegatum* ticks in The Gambia. Exp. Appl. Acarol. **42:**61-74. doi: 10.1007/s10493-007-9073-2 [doi].

606 45. **Adakal, H, Gavotte, L, Stachurski, F, Konkobo, M, Henri, H, Zoungrana, S, Huber,**
607 **K, Vachiéry, N, Martinez, D, Morand, S, Frutos, R.** 2010. Clonal origin of emerging
608 populations of *Ehrlichia ruminantium* in Burkina Faso. Infect. Genet. Evol. **10:**903-912. doi:
609 10.1016/j.meegid.2010.05.011.

610 46. **Peter, TF, Bryson, NR, Perry, BD, O'Callaghan, CJ, Medley, GF, Smith, GE,**
611 **Mlambo, G, Horak, IG, Burridge, MJ, Mahan, SM.** 1999. Cowdria ruminantium infection
612 in ticks in the Kruger National Park. Vet. Rec. **145:**304-307. doi.org/10.1136/vr.145.11.304.

613 47. **Allsopp, MT, Theron, J, Coetzee, ML, Dunsterville, MT, Allsopp, BA.** 1999. The
614 occurrence of *Theileria* and *Cowdria* parasites in African buffalo (*Syncerus caffer*) and their
615 associated Amblyomma hebraeum ticks. Onderstepoort J. Vet. Res. **66:**245-249.

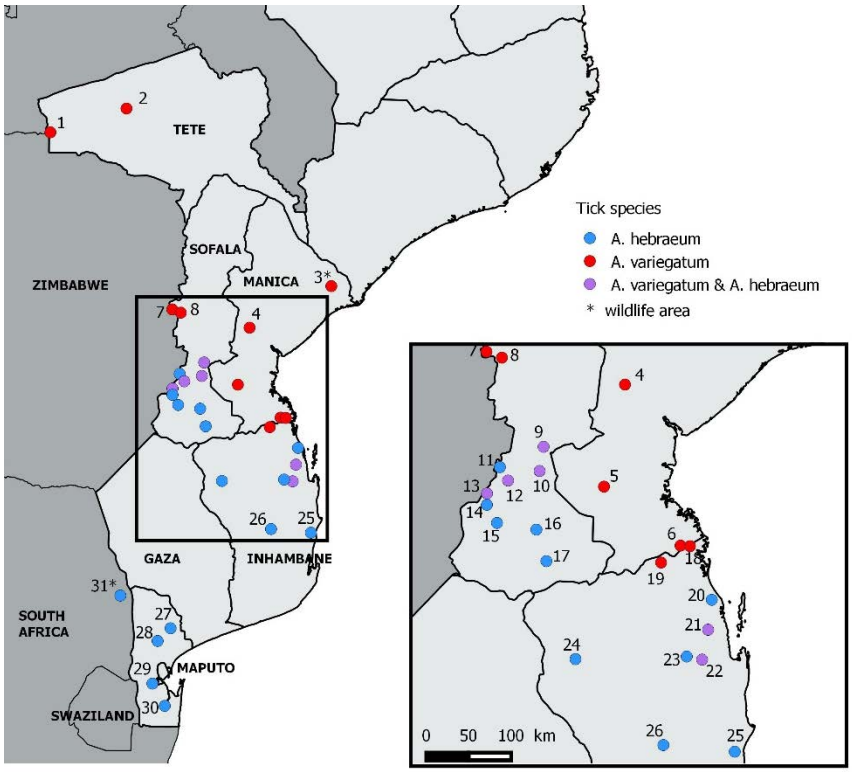
616 48. **Bezuidenhout, JD.** 1987. Natural transmission of heartwater. Onderstepoort J. Vet. Res.
617 **54:**349-351.

618 49. **Andrew, HR, Norval, RA.** 1989. The carrier status of sheep, cattle and African buffalo
619 recovered from heartwater. Vet. Parasitol. **34:**261-266.

620 50. **Vachiéry, N, Jeffery, H, Pegram, R, Aprelon, R, Pinarello, V, Kandassamy, RL,**
621 **Raliniaina, M, Molia, S, Savage, H, Alexander, R, Frebling, M, Martinez, D, Lefrancois,**
622 **T.** 2008. *Amblyomma variegatum* ticks and heartwater on three Caribbean Islands. Ann. N. Y.
623 Acad. Sci. **1149:**191-195. doi: 10.1196/annals.1428.081.

624 51. **Zweygarth, E, Josemans, AI, Steyn, HC.** 2008. Experimental use of the attenuated
625 *Ehrlichia ruminantium* (Welgevonden) vaccine in Merino sheep and Angora goats. Vaccine.
626 **26, Supplement 6:**G34-G39. doi: http://dx.doi.org/10.1016/j.vaccine.2008.09.068.

627



Sampling site	Locality
1	Zumbo Minga
2	Finoe Mazav
3	Coutada
4	Nhamatanda
5	Muxungue
6	Machanga
7	Munene
8	Chirere
9	Dombe
10	Gunhe
11	Mahiza
12	Dacata
13	Espungabera
14	Majuacuana
15	Cita-Gaha
16	Chipambuleque
17	Chipopopo
18	Mambone
19	Govuro
20	Vulanjane
21	Pambara
22	Mapinhane
23	Mwabsa
24	Zimane
25	Massinga
26	Manhica-Inhambane
27	Chobela-Magude
28	Manhica-Maputo
29	Changalane
30	Matutune
31	Kruger National Park

Figure 1. Sampling map of *A. hebraeum* and *A. variegatum* ticks collected from cattle and wildlife in five provinces of Mozambique and one locality in South Africa. Site number and the corresponding locality is listed beside the map.*: collection of ticks on wildlife.

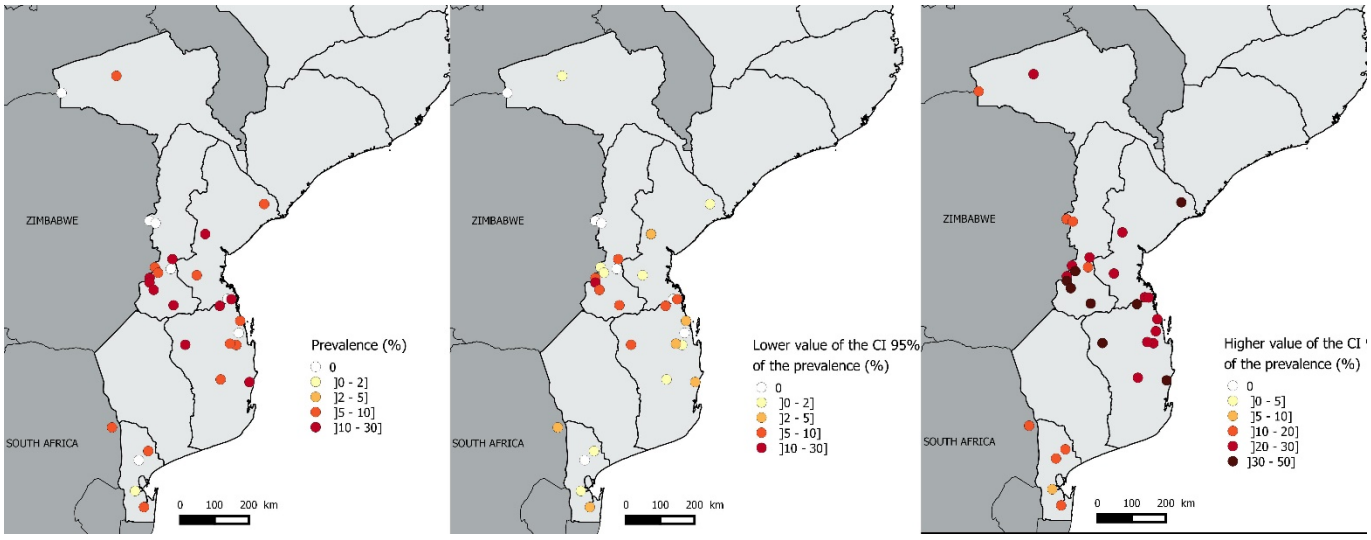
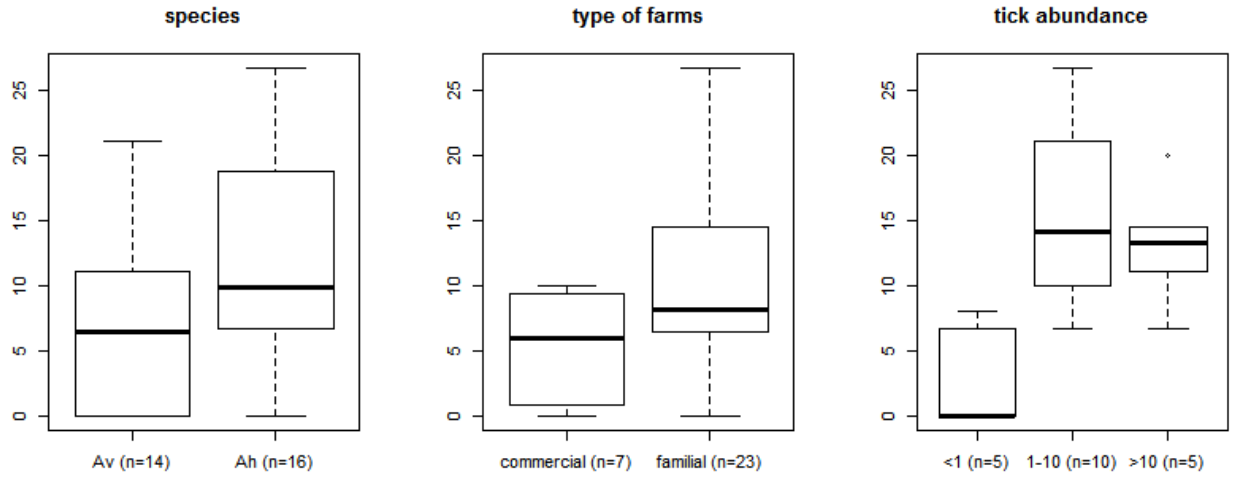


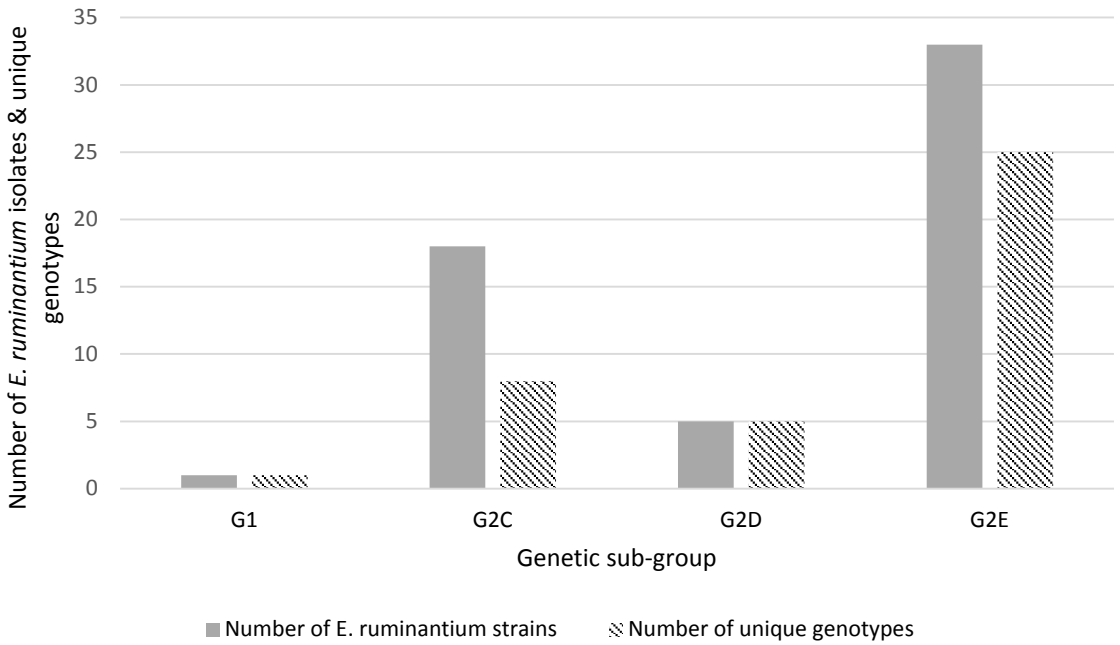
Figure 2. Prevalence of *E. ruminantium* in *Amblyomma* ticks (on the left) and lower (in the middle) and higher (on the right) values of the confidence interval 95% estimated with binomial law in sampled localities in Mozambique.



637

638 Figure 3. Boxplot of *E. ruminantium* prevalence in ticks per locality according to tick's species
639 (on the left), type of farms (in the middle) and tick abundance (on the right). In bracket: number
640 of localities, Av: *A. variegatum*, Ah: *A. hebraeum*.

641



642

643 Figure 4: *E. ruminantium* strain and unique genotype distribution per genetic group using
644 MSLT

645

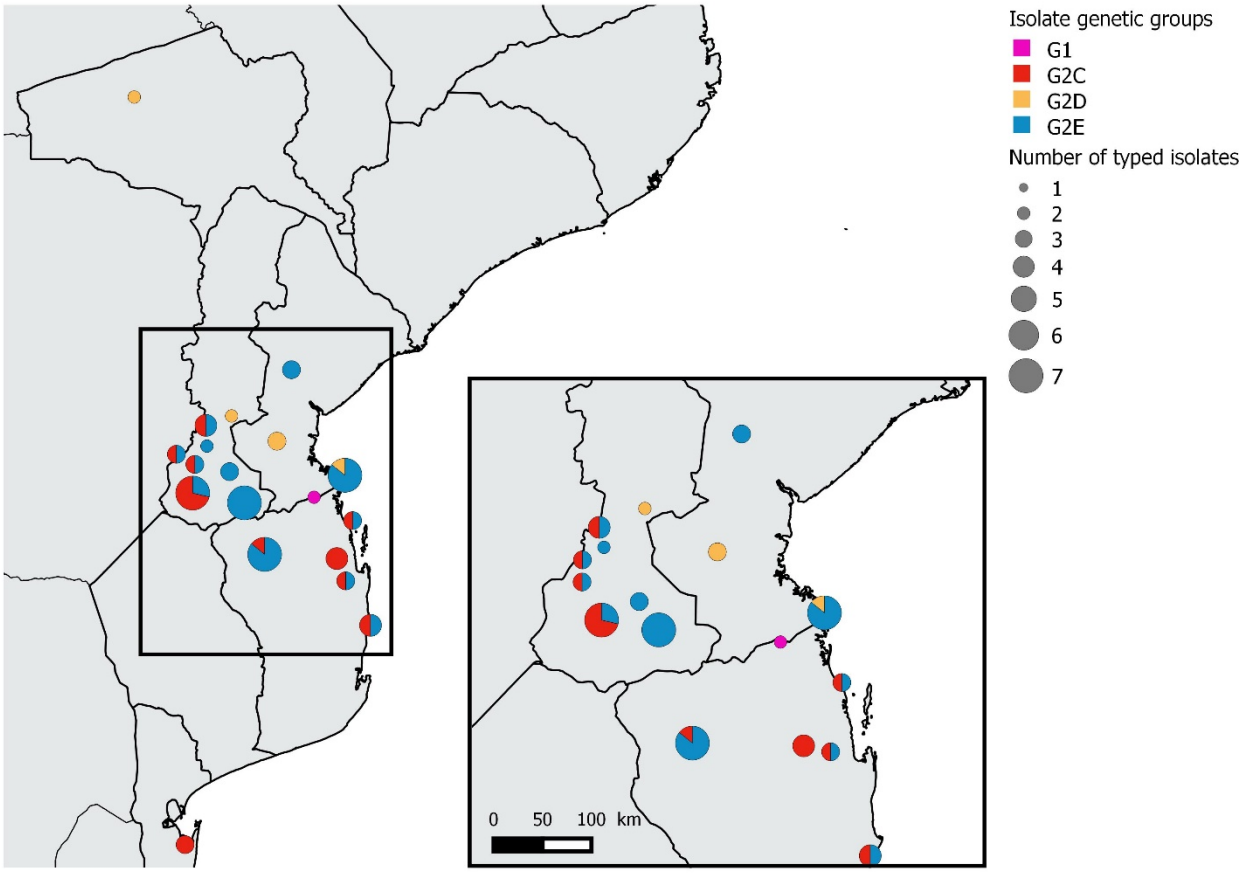


Figure 5. Distribution of *E. ruminantium* genetic groups 1, 2C, 2D and 2E per locality in Mozambique. Isolate genetic groups are color coded according to the legend. The number of isolates per locality including identical genotypes are shown and are size coded from 1 to 7 isolates, according to the legend.

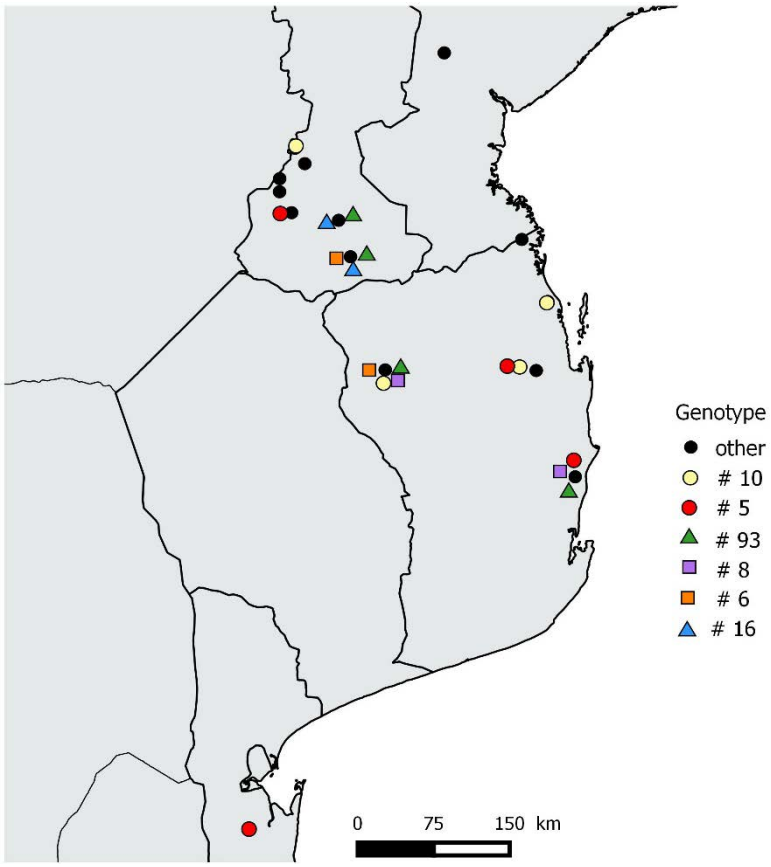


Figure 6. Map of geographical distribution of *E. ruminantium* genotypes in Mozambique.

667 Table 1. Description of sampled localities per province for collection of *A. hebraeum* and *A.*
668 *variegatum*, number of ticks collected and prevalence of *E. ruminantium* per locality.

Sampling site	Locality	Province	No. Ticks	Ah	Av	Positive to ER	Prevalence (%) (CI 95%)	Tick abundance (No. Tick/animal)
1	Zumbo Minga	Tete	27	0	27	0	0	NA
2	Finoe Mazav		29	0	29	2	6.9 (0.8 - 22.8)	NA
3	Coutada (wildlife)	Sofala	15	0	15	1	6.2 (0.2 - 30.2)	NA
4	Nhamatanda		27	0	27	3	11.1 (2.3 - 29.1)	>10
5	Muxungue		37	0	37	3	8.1 (1.7 - 21.9)	<1
6	Machanga		14	0	14	0	0 (0 - 23.2)	<1
7	Munene	Manica	20	0	20	0	0 (0 - 16.8)	NA
8	Chirere		29	0	29	0	0 (0 - 11.9)	NA
9	Dombe		48	7	41	7	14.6 (6.1 - 27.8)	1 - 10
10	Gunhe		28	4	24	0	0 (0 - 12.3)	<1
11	Mahiza		30	30	0	2	6.7 (0.8 - 22.1)	1 - 10
12	Dacata		15	4	11	1	6.7 (0.2 - 31.9)	<1
13	Espungabera		66	36	30	9	13.6 (6.4 - 24.3)	1 - 10
14	Majuacuana		30	30	0	8	26.7 (12.3 - 45.9)	1 - 10
15	Cita-Gaha		50	50	0	9	18 (8.6 - 31.4)	1 - 10
16	Chipambuleque		6	6	0	1	NA	NA
17	Chipopopo		30	30	0	7	23.3 (9.9 - 42.2)	1 - 10
18	Mambone	Inhambane	55	0	55	8	14.5 (6.5 - 26.7)	>10
19	Govuro		19	0	19	4	21 (6.1 - 45.6)	1 - 10
20	Vulanjane*		30	30	0	3	10 (2.1 - 26.5)	1 - 10
21	Pambara*		14	7	7	0	0 (0 - 23.2)	<1
22	Mapinhane*		34	31	3	3	8.8 (1.9 - 23.7)	1 - 10
23	Mwabsa*		30	30	0	3	10 (2.1 - 26.5)	1 - 10
24	Zimane		30	30	0	6	20 (7.7 - 38.6)	>10
25	Massinga		30	30	0	4	13.3 (3.8 - 30.7)	>10
26	Manhica-Inhambane		30	30	0	2	6.7 (0.8 - 22.1)	>10
27	Chobela-Magude*	Maputo	50	50	0	3	6 (1.2 - 16.5)	NA
28	Manhica-Maputo*		23	23	0	0	0 (0 - 14.8)	NA
29	Changalane*		61	61	0	1	1.6 (0.1 - 8.8)	NA
30	Matutune		81	81	0	7	8.6 (3.5 - 17)	NA
31	Kruger National Park (wildlife)	South Africa	122	122	0	10	8.2 (4 - 14.6)	NA
Total number of tested ticks			1110	722	388			

669 Ah: *A. hebraeum*, Av: *A. variegatum*, ER: *E. ruminantium*, *commercial farms.

670

671 Table 2: *E. ruminantium* genotypes per locality and tick species

Locality	Genetic Group	Isolate name	Date of isolation	Genotype number	Tick species	Number of different genotypes	Number of tested ticks
Chipambuleque	G2E	CHIPA3MH1	2012	**16**	Ah	**2**	**2**
		CHIPA2MH1	2012	**93**	Ah		
Chipopopo	G2E	CHIPO29MH1	2012	6	Ah	**4**	**7**
		CHIPO17MH1	2012	7	Ah		
		CHIPO12MH1	2012	9	Ah		
		CHIPO22MH1	2012	14	Ah		
		CHIPO2MH1	2012	**16**	Ah		
		CHIPO26MH1	2012	**93**	Ah		
		CHIPO24MH2	2012	**93**	Ah		
Cita_Gaha	G2C	CIT9MH1	2012	**5**	Ah	**4**	**7**
		GAH7MH1	2012	**5**	Ah		
		CIT28MH1	2012	**5**	Ah		
		GAH1MH2	2012	**4**	Ah		
		GAH4MH1	2012	**4**	Ah		
	G2E	GAH5MH2	2012	15	Ah		
		GAH9MH1	2012	18	Ah		
Dacata	G2E	777-DAC8MH1	2014	34	Ah	**1**	**1**
Dombe_Darue_maquina	G2D	709-DAR5MV1	2014	32*	Av	**1**	**1**
Espungabera_muedza	G2C	801-MUE3FH1	2014	38	Ah	**2**	**2**
	G2E	765-ESP2-4MV1	2014	33	Av		
Finoe Mazav	G2D	488-FINMV13	2014	39	Av	**1**	**1**
Govuro	G1	303-GOV1-MV8	2012	22	Av	**1**	**1**
Mahiza	G2C	MAH12MH1	2012	**10***	Ah	**2**	**2**
	G2E	MAH6MH1	2012	20	Ah		
Majuacuana	G2C	832-MAG12MH1	2014	37	Ah	**2**	**2**
	G2E	823-MAG7MH1	2014	35	Ah		
Mambone_Maninga2	G2E	335-MAMMV13	2013	23	Av	**3**	**10**
		330-MAMMV8	2013	27*	Av		
		347-MAMMV22	2013	28*	Av		
		319-MAMMV2	2013	29*	Av		
		336-MAMFV6	2013	30*	Av		
		342-MAMMV17	2013	31*	Av		
	G2D	595-MANINMV19	2014	40	Av		
Mapinhane_Vilankulos	G2C	431-MAP21MH1	2013	36	Ah	**2**	**2**
	G2E	445-MAP32MH1	2013	24	Ah		
Massinga	G2C	MAS14MH1	2012	21	Ah	**3**	**4**
		MAS13MH1	2012	5	Ah		
	G2E	MAS1MH1	2012	**93**	Ah		
		MAS27MH2	2012	**8**	Ah		
Matatune	G2C	MAT14MH2	2012	**5**	Ah	**2**	**2**
		MAT13MH2	2012	11	Ah		
Muxungue1	G2D	385-MUX3MV1	2013	41	Av	**2**	**2**
		394-MUX9MV1	2013	42*	Av		
Mwabsa	G2C	MWAB11FH1	2012	**5**	Ah	**2**	**3**
		MWAB11MH1	2012	**5**	Ah		
		MWAB3MH1	2012	**10***	Ah		
Nhamatanda	G2E	550-NHAMV12	2014	25	Av	**2**	**2**
		559-NHAMV21	2014	26	Av		
Vulanjane	G2C	VUL29MH1	2012	**10***	Ah	**2**	**2**
	G2E	VUL17MH1	2012	13*	Ah		
Zimane	G2C	ZIM28MH1	2012	**10***	Ah	**5**	**7**
	G2E	ZIM15MH1	2012	6	Ah		
		ZIM31MH1	2013	**8**	Ah		
		ZIM16MH2	2012	12	Ah		
		ZIM2MH1	2012	17	Ah		
		ZIM4MH1	2012	17	Ah		
		ZIM1MH1	2012	**93**	Ah		
Total number of *E ruminantium* typed isolates						57	
Number of *E ruminantium* unique genotypes						39	

*: recombinant genotype, Bold: clones, Light grey shade: genetically closed strains (one or 2 substitutions); Dark grey shade: collection on same animal

672 93: Welgevonden reference strain; Ah: A. hebraeum, Av: A. variegatum

VI. General discussion

In Section I, we developed two new qPCR, *pCS20* Sol1^{TM} and Sol1^{SG}, to screen *E. ruminantium* in *Amblyomma* ticks, which are powerful tools for: 1) heartwater epidemiological studies, 2) diagnosis in the context of heartwater clinical cases and 3) follow-up of experimental infections, both in ticks and hosts. The development of *pCS20* Sol1^{TM} qPCR coupled with an automated method for DNA/RNA extraction allowed processing of high number of tick samples collected in Mozambique that were then typed by MLST and included into a worldwide *E. ruminantium* strain genetic structure study (Section II). In section III, we focused mainly on *E. ruminantium* tick prevalence and genetic diversity and structure of Mozambican isolates from *A. variegatum* and *hebraeum* ticks collected in cattle and wildlife.

Diagnosis of tick borne pathogens such as *E. ruminantium* is important for confirmation of clinical cases (Martinez *et al*., 2004b), determination of a carrier state (Peter *et al*., 1998), determination of prevalence in ticks and spatial distribution of the pathogen strains in a given area (Molia *et al*., 2008). This will lead to effective disease management as well as monitoring of therapy and prophylactic measures (Adakal *et al*., 2010; Salih, EI Hussein & Singla, 2015; Salih, EI Hussein & Singla 2015).

The most frequently used tests for diagnosis of *E. ruminantium* are serological and molecular (Allsopp, 2015). The currently used serological tests are cELISA (Katz *et al*., 1997) and indirect ELISA MAP-1B (van Vliet *et al*., 1995). These tests have the advantage of being specific and can be used to screen a considerable number of samples for epidemiological studies and serve as an indicator of areas of infection risk. However, cross-reactions with *E. canis* and *E. chaffeensis* may occur and although these species do not infect ruminants, they can produce false-positives (van Vliet *et al*., 1995). Additionally, seropositivity as a reflection of the immune response of the host through the production of antibodies can vary with time and does not always represent a current infection. The period of seropositivity is very short, lasting from several weeks (bovine) to six months (small ruminants), (Vachiéry *et al*., 2013) Therefore, these tests alone cannot be used to confirm diagnostic in endemic areas and are ineffective for diagnosis during the acute phase of clinical cases, as IgG antibodies appear only two weeks after infection (Vachiéry *et al*., 2013; Fierz, 2004).

The most reliable test for molecular diagnostic of *E. ruminantium*, which is also the gold standard assay for the OIE reference laboratory, is the nested PCR *pCS20* (Martinez *et al*., 2004b; Molia *et al*., 2008). This PCR was tested on more than 100 *E. ruminantium* strains and the use of universal nucleotides in the primers increased strain detection without reducing the specificity. Additionally, this PCR is highly specific and sensitive, detecting approximately 6

copies of bacteria/sample. However, this method requires a high processing time and is prone to cross-contamination. Meanwhile, the development of *map1, map1-1* and *pCS20* qPCR (Peixoto *et al.*, 2005; Postigo *et al.*, 2007; Steyn *et al.*, 2008) allowed for quick diagnostic results to be obtained (within two hours) and reduction of cross-contamination risk, while maintaining the levels of specificity and sensitivity (detecting approximately 6 copies of bacteria/sample). However, *map1* and *map1-1* qPCRs could not be used for diagnostic purposes as it targets polymorphic genes and limited number of strains was tested. The current study failed to amplify through *map1* qPCR some reference strains such as Lamba 479, Sankat and Senegal (data not shown). This fact hinders the optimization of *map1* qPCR for diagnosis purposes. Moreover, in our hands, we were not able to implement the *pCS20* qPCR developed by Stein *et al.* (2008).

In order to increase sample processing capacity and contribute to diagnostic and epidemiological studies, a new automatic DNA extraction and qPCR was developed for *E. ruminantium*. Our work resulted in the optimization of an automatic DNA/RNA extraction method and a new qPCR *pCS20* Sol1^{TM} with similar performance to the manual DNA extraction kit and the gold standard test, *pCS20* nested PCR.

The present work represents a substantial contribution to the molecular diagnostic of *E. ruminantium*. The possibility of extensive epidemiological studies, through the improvement of diagnostic methods, will be important to control heartwater outbreaks and reduce infection in the future by understanding the pathogen source and circulation (Bartlett, Judge 1997).

In fact, the new qPCR could be applied as a diagnostic test to screen different samples including blood, thus allowing for the diagnosis of clinical cases. A rapid and efficient diagnostic of suspicious cases of heartwater can also be achieved in the case of outbreaks in heartwater free areas. Particularly, in regions with high risk of introduction such as the American mainland, it will be useful to be able to quickly confirm the presence of *E. ruminantium* either in host or vectors.

Interestingly, we were also able to optimize a *pCS20* Sol1 using SYBR Green detection instead of a probe, which despite having the potential to detect non-specific signal at low bacteria load, is cheaper and can be used in laboratories with resource constraints.

Besides its use for diagnostic, the *pCS20* Sol1 qPCR is also a powerful tool to follow up experimental assays performed on ticks and hosts in order to better understand the interaction between ticks, hosts and pathogens. We do not currently know the number of *E. ruminantium* copies per tick necessary to transmit the disease, so even if *E. ruminantium* tick prevalence is evaluated in the field and the infection in positive ticks corresponds to a weak signal (60

copies/µl), it does not mean that ticks can efficiently transmit the disease. To address this gap in knowledge, ticks can be experimentally infected at nymphal stage on infected goats and *E. ruminantium* load measured after moulting and after the blood meal, just before the animal reacts. It will also be possible to quantify the load in experimentally infected goats, before and during hyperthermia using Sol1 qPCR. Furthermore, the process by which *E. ruminantium* colonizes ticks is unclear. Through the use of the qPCR, it will be possible to elucidate this process by following the development of *E. ruminantium* in the midgut and salivary glands, if each organ is dissected at a different time, during blood feeding.

Because tick borne diseases are very important and represent a worldwide risk for human and veterinary public health, the first part of the method, automatic RNA/DNA extraction of ticks and DNA quality control using $16S^{SG}$ rDNA qPCR, developed within this work can have a crucial impact in tick borne pathogen detection. Nucleic acids from any tick species can be extracted, allowing any virus and bacterium other than *E. ruminantium* to be screened using multi-pathogen assays such as multiplex qPCRs or BioMark dynamic array system (Michelet *et al*., 2014). Moreover, the method can also be used for tick genetic studies or tick pathogen co-evolution studies.

The new techniques developed in the current study should be tested and implemented in other laboratories with interest in research and diagnostic of heartwater and other tick-borne pathogens.

In section II, we performed the genetic characterization of *E. ruminantium* worldwide strains, including new strains from Mozambique, Caribbean and Indian Ocean. The study was initially conducted at a global scale and later focused on a restricted area of Mozambique (Section III). Genetic characterization is the detection and differentiation of strains as a result of differences in DNA sequences, genes or modifying factors (de Vicente *et al*., 2006). Previous work on the genetic diversity of *E. ruminantium* has been based on *map1* gene sequence (Raliniaina *et al*., 2010), MLVA (Pilet *et al*., 2012; Nakao *et al*., 2012) and MLST (Adakal *et al*., 2009; Nakao *et al*., 2011), which have the advantage of giving clear results, especially if DNA is sequenced. These techniques can be reproduced in several laboratories and are sensitive enough to detect differences between isolates. However, they require enough DNA for amplification, expertise in analysis and sufficient PCR sensitivity to allow for the amplification of different strains (Sullivan, Diggle & Clarke, 2005; Lindstedt, 2005).

The different isolates from Mozambique were typed using MLST approach and compared with West, East and South Africa, Indian Ocean and Caribbean strains. The current study (Section

II) reveals the repeated occurrence of recombination between *E. ruminantium* strains and the presence of two genetic groups named West Africa (G1) and Worldwide (G2). The origin of *E. ruminantium* introduction to the Caribbean and Indian Ocean islands was also elucidated. Some hypotheses were formulated concerning the origin of *E. ruminantium* genetic diversity and the introduction of ancestral genotypes, depending on cattle migration and the role of wildlife and ticks before cattle introduction in Africa. One main hypothesis is the presence of genotypes at low frequency in atypical locations, which is probably due to their recent introduction (Section II and III). This is particularly true for group G1 and subgroup G2D, which are dominant in West Africa and Caribbean and are detected at a low frequency in areas of Mozambique where only *A. variegatum* is present. Moreover, in areas with G2D genotypes, there are a lot of recombinant genotypes, strengthening the hypothesis of recent introduction. If isolation of genotypes from subgroup 2D were to be successful, it would be interesting to infect *A. hebraeum* and *A. variegatum* ticks and compare their bacterial load and vector capacity using an automatic DNA extraction and Sol1 qPCR developed in Section I. These preliminary experiments could clarify our hypothesis of a recent introduction of subgroup 2D genotypes and their potential lack of adaptation to *A. hebraeum*. Conversely, the ability of strains from the South of Mozambique to grow in *A. variegatum* could also be tested (i.e. test isolates from G2C which were detected only in *A. hebraeum*).

Subgroup G2C clustered only Mozambican strains and coupled with G2E, are the mainly genetic groups present in Mozambique. However, it would be important to increase the sampling in other Southern African countries, such as Zimbabwe and South Africa, to confirm the specific location and/or origin of introduction of these subgroups. In the south of Tete province *Amblyomma* and *Ehrlichia* are absent. Thus, it will be important to increase the collection and screening of ticks and strain typing in the north of Tete to verify whether all the subgroups are detected or only subgroup 2D. In the north of Tete, there are heavy restrictions on exports to the South of the country due to the endemic presence of *Theileria parva* and East Coast Fever, but introduction via imported cattle could be possible.

Given that more *A. hebraeum* than *A. variegatum* ticks were tested, the differences in genotype could be due to a sampling bias. For isolates G1 and G2D present only in *A. variegatum*, the results seem to be accurate because of the high number of *A. hebraeum* ticks tested, whereas for isolates from G2C associated with *A. hebraeum* precaution should be taken, given the lower number of *A. variegatum* tested. It is essential to increase the number of *A. variegatum* ticks to be tested in order to confirm whether genotypes are associated with one tick species or the

other. Moreover, if strains could be isolated *in vivo*, differences in strain adaptation to tick species would be confirmed by experimental infections as describe above.
Unfortunately, none of the strains detected in wildlife samples could be typed by MLST in our study, making it impossible to compare them with *E. ruminantium* isolates from cattle. This problem with typing could be explained by a high genetic diversity among strains, causing a mismatch of the primers used for sequencing or amplification of non-target DNA in PCR. It could also be due to co-infection by several strains from different subgroups, which could render the sequences unreadable. For cattle, we did not have such a problem of mixture of strains when sequencing MLST products. In Burkina Faso, we previously found 18% of ticks co-infected, using *map-1* RFLP, which is much more polymorphic than MLST (Vachiéry, personal communication).

In order to complement MLST results and obtain a more informative *E. ruminantium* genetic diversity and population structure, our isolates were also typed by MLVA technique. This method includes VNTR loci, which are more polymorphic and discriminatory than the housekeeping genes used in MLST and have less associated typing costs (Vergnaud & Denoeud, 2000). To our surprise, we could not amplify most of the 8 loci from the gene panel on Mozambican isolates (data not shown). These findings could be explained by the high genetic diversity of the bacteria and strain signatures that do not allow for amplification with the current primers developed by Pilet *et al.* (2012). A possible solution could be the design of new primers and/or new primer combinations. Future studies should consider primer redesign and optimization, if a lack of amplification occurs. Thus, we continued the study, focusing only on MLST analysis.

Although it has already been reported in the literature (Bekker *et al*., 2005; Hughes & French, 2007), we were not expecting to detect such important recombination events in *E. ruminantium* conserved genes. This fact will have an implication when interpreting *E. ruminantium* genetic diversity and population structure. Moreover, the present work sheds a light on the weakness of MLST as a unique typing method for *Anaplasmataceae* due to the extensive recombination events. Thus, MLST should be complemented with other typing methods. Taking into account the progress of sequencing technologies, whole genome sequencing will possibly become the preferred method. The *map1* genotyping mentioned above also has important limitations. Using this genotyping method, no association of strains with geographic origin and evolution of ancestral founders could be determined (Raliniaina *et al.*, 2010) and the polymorphism of

this gene family probably also reflects the recombination events in these bacteria. Currently, the technical development of whole genome sequencing is contributing to cost reductions and an explosive growth of data (Land *et al*., 2015). Further, whole genome sequencing has become a tool for epidemiologists to type and track disease outbreaks in real time (Salipante *et al*., 2015). However, the main barrier to obtain whole genome sequence from field ticks is the extremely low amount of *Ehrlichia* DNA compared to the large amount of vector DNA.

With the development of new diagnostic tools (automatic DNA extraction and qPCR), the prevalence of *E. ruminantium* in southern and central Mozambique was elucidated through the screening of *A. hebraeum* and *A. variegatum* ticks collected from cattle and wildlife species (Section III). The prevalence of *E. rumimantium* in relation to the tick species and locality in Mozambique and correlation with tick abundance was analysed. *E. ruminantium* prevalence in *A. hebraeum* and *A. variegatum* from cattle varied from 0% [0-23.2 %] to 26.7% [12-45 %]. Our results support those of previous studies on *E. ruminantium* prevalence in ticks from cattle in Africa and the Caribbean (Faburay *et al*., 2007b; Molia *et al*., 2008; Adakal *et al*., 2009; Adakal *et al*., 2010; Vachiéry et al. 2008). For wildlife, few studies have attempted to determine the prevalence in ticks (Peter *et al*., 1999), tending to focus, instead, on the description of clinical cases and species' susceptibility to heartwater (Peter, Burridge & Mahan, 2002). For Mozambique, the current study is the first of its kind to molecularly determine *E. ruminantium* prevalence in ticks, collected from both cattle and wildlife species.

Prevalence estimation is important as an indicator of disease risk and can help to delineate better control strategies in a region (Rothman, 2012). Furthermore, detection of pathogens in ticks can help us to estimate the pathogen geographical distribution, while gaining some insight into the biology of the disease. The prevalence of the host and vector is probably affected by the level of herd immunity (endemic stability), (Deem *et al*., 1996). The establishment of enzootic stability requires a stable relationship between the mammal hosts, etiological agent, vector and the environment (Norval, Perry & Young, 1992). Accordingly, enzootic stability is established when the number of infected ticks are sufficient to transmit *E. ruminantium* to the mammalian host during the period of resistance to clinical disease (Deem *et al*., 1996). The maintenance of enzootic stability is currently considered to be the most sustainable approach to control heartwater in endemic areas (Allsopp, Bezuidenhout & Prozesky, 2004). However, the tools available to monitor the dynamics of enzootic stability are not efficient. In the current work, we have developed efficient molecular tools and apply them to determine *E. ruminantium* tick prevalence in a wide geographical area. The combination of prevalence and

reliable methods to determine tick density could be used to effectively monitor changes in enzootic stability.

In the current study, wildlife prevalence was 8.2 % [4-14.6 %] in the KNP and 6.2% [0.2-30.2 %] in hunting concessions of Sofala province. Wildlife have an important role in the epidemiology and spread of heartwater (Peter, Burridge & Mahan, 2002). This is the first time that *E. ruminantium* was detected in wildlife from Mozambique and these findings indicate a possible role for wild species as disease reservoirs. Nevertheless, more sampling needs to be carried out in order to isolate more strains. Ideally, *in vivo* isolation should be performed and isolates should be genotyped in order to compare them with livestock strains.

In terms of the vector species, we did not observe a significant difference for *E. ruminantium* prevalence in *A. hebraeum* ticks compared to *A. variegatum*. This lack of difference could possibly be explained by the fact that a higher number of *A. hebraeum* ticks were collected in our study compared to *A. variegatum* ticks, precluding the comparison. However, both tick species might have similar susceptibility to infection by several *E. ruminantium* strains and potential for transmitting them. This makes us question the difference in vector competence of *A. variegatum* in relation to that of *A. hebraeum* and their role in *E. ruminantium* transmission in Mozambique. Nevertheless, our sampling in Mozambique should be optimized to allow a stronger comparison between A. *hebraeum* and *A. variegatum* prevalence in the different areas, where both tick species occur.

VII. Conclusions and perspectives

Part of the study aimed to develop a high throughput molecular method to screen *E. ruminantium* in ticks including automatic acid nucleic extraction and a new *pCS20* qPCR. The study demonstrated that the new automatic DNA extraction of *Amblyomma* ticks and *pCS20* Sol1^{TM} qPCR for *E. ruminantium* detection have a good sensitivity, specificity and reproducibility. The new qPCR is a powerful tool for heartwater diagnostic but also to follow-up experimental infection both in vectors and host. Furthermore, a new tick DNA quality control method based on 16S^{SG} rDNA qPCR was optimized, meaning that the extraction method and DNA quality control could be widely used for other bacteria and virus screening in field ticks and for other tick species. The work paves the way to screen any tick-borne pathogen on a large scale given the development of a tick automatic DNA/RNA extraction.

In terms of the isolates genetic structure analysis using MLST at worldwide scale, two main genetic groups were found in the present study: West African and a worldwide group. In the study, the origin of *E. ruminantium* introduction due to cattle movement was clarified, with introductions from West Africa to the Caribbean and from Southern Africa to the Indian Ocean islands. Equally important, the results highlight that recombination is probably a major driver of genetic diversity in this obligate intracellular pathogen and the difficulty to use solely MLST to perform phylogenetic and phylogeographical studies.

Last, the study aimed to determine the prevalence of *E. ruminantium* in *A. hebraeum* and *A. variegatum* ticks and genetic diversity of isolates in Mozambique. *E. ruminantium* tick prevalence in cattle varied from 0% [0-23.2 %] to 26.7% [12-45 %]. In wildlife the overall prevalence was 8.2 % [4-14.6 %] in the KNP and 6.2% [0.2-30.2 %] in hunting concessions of Sofala province. Wildlife probably play a role as reservoirs of *E. ruminantium*.
The results did not show a linear correlation between *E. ruminantium* prevalence and tick abundance. MLST genotypes from group 1 and 2D were exclusively found in areas of *A. variegatum* distribution, while group 2C was only detected in *A. hebraeum* areas. Moreover, genotypes from group 2E were found in both *A. hebraeum* and *A. variegatum* areas.

Optimization and implementation of the new highly specific and sensitive qPCRs, Sol1^{TM} and Sol1^{SG}, should be considered in laboratories with interest in investigation of *E. ruminantium* and heartwater diagnosis.
The remaining *A. variegatum* ticks collected during this work were not screened for *E. ruminantium*. With the use of automatic DNA extraction and *pCS20* qPCR, a higher number

of *A. variegatum* ticks will be tested and *E. ruminantium* positive samples will be typed by MLST to screen for the presence of genotypes.

More sampling needs to be done in order to compare the prevalence and to confirm the adaptation of certain genotypes to the two main *Amblyomma* vectors in Mozambique. Wildlife sampling also presents a good opportunity to identify the strains circulating in the wild or wildlife-interface areas and to understand how the dynamics of heartwater transmission and epidemiology of the disease is affected. Collaboration agreements with hunting concessions and national parks should be established in Mozambique to facilitate sampling on wildlife.

The current thesis made a contribution to heartwater research by improving the molecular diagnostic capacity, by providing epidemiological data on *E. ruminantium* prevalence in southern Africa (Mozambique) and by characterizing new *E. ruminantum* genotypes and analyzing their genetic population structure. We hope that the current investigation could shed a light on reducing the impact of tick-borne diseases such as heartwater in cattle production systems.

VIII. References

Adakal, H., Gavotte, L., Stachurski, F., Konkobo, M., Henri, H., Zoungrana, S., Huber, K., Vachiéry, N., Martinez, D., Morand, S. & Frutos, R. 2010. "Clonal origin of emerging populations of *Ehrlichia ruminantium* in Burkina Faso". Infection, Genetics and Evolution, vol. 10, no. 7, pp. 903-912.

Adakal, H., Meyer, D.F., Carasco-Lacombe, C., Pinarello, V., Allegre, F., Huber, K., Stachurski, F., Morand, S., Martinez, D., Lefrancois, T., Vachiéry, N. & Frutos, R. 2009. MLST scheme of *Ehrlichia ruminantium*: genomic stasis and recombination in strains from Burkina-Faso. Infection, Genetics and Evolution, vol. 9, no. 6, pp. 1320-1328.

Adakal, H., Stachurski, F., Konkobo, M., Zoungrana, S., Meyer, D.F., Pinarello, V., Aprelon, R., Marcelino, I., Alves, P.M., Martinez, D., Lefrancois, T. & Vachiéry, N. 2010. Efficiency of inactivated vaccines against heartwater in Burkina Faso: impact of *Ehrlichia ruminantium* genetic diversity. Vaccine, vol. 28, no. 29, pp. 4573-4580.

Allender, M.C., Bunick, D., Dzhaman, E., Burrus, L. & Maddox, C. 2015. Development and use of a real-time polymerase chain reaction assay for the detection of Ophidiomyces ophiodiicola in snakes. Journal of Veterinary diagnostic investigation, vol. 27, no. 2, pp. 217-220.

Allsopp, B.A. 2015. Heartwater - *Ehrlichia ruminantium* infection. Revue scientifique et technique (International Office of Epizootics), vol. 34, no. 2, pp. 557-568.

Allsopp, B.A. 2010. Natural history of *Ehrlichia ruminantium*. Veterinary parasitology, vol. 167, no. 2-4, pp. 123-135.

Allsopp, B.A. 2009. Trends in the control of heartwater. The Onderstepoort journal of veterinary research, vol. 76, no. 1, pp. 81-88.

Allsopp, B.A., Bezuidenhout, S.M., Prozesky, L 2004. “Heartwater” in Infectious diseases of livestock, ed. Coetzer, J.A.W.,Tustin, R.C., 2nd edn, Oxford University Press, Oxford and Cape Town, pp. 507-535.

Allsopp, M.T. & Allsopp, B.A. 2007. Extensive genetic recombination occurs in the field between different genotypes of *Ehrlichia ruminantium*. Veterinary microbiology, vol. 124, no. 1-2, pp. 58-65.

Allsopp, M.T., Dorfling, C.M., Maillard, J.C., Bensaid, A., Haydon, D.T., van Heerden, H. & Allsopp, B.A. 2001. *Ehrlichia ruminantium* major antigenic protein gene (map1) variants are not geographically constrained and show no evidence of having evolved under positive selection pressure. Journal of Clinical Microbiology, vol. 39, no. 11, pp. 4200-4203.

Allsopp, M.T., Hattingh, C.M., Vogel, S.W. & Allsopp, B.A. 1999. Evaluation of 16S, map1 and pCS20 probes for detection of Cowdria and Ehrlichia species. Epidemiology and infection, vol. 122, no. 2, pp. 323-328.

Allsopp, M.T., Louw, M. & Meyer, E.C. 2005. *Ehrlichia ruminantium*: an emerging human pathogen? Annals of the New York Academy of Sciences, vol. 1063, pp. 358-360.

Allsopp, M.T., Theron, J., Coetzee, M.L., Dunsterville, M.T. & Allsopp, B.A. 1999. The occurrence of *Theileria* and *Cowdria* parasites in African buffalo (*Syncerus caffer*) and their

associated *Amblyomma hebraeum* ticks. The Onderstepoort journal of veterinary research, vol. 66, no. 3, pp. 245-249.

Ammazzalorso, A.D., Zolnik, C.P., Daniels, T.J. & Kolokotronis, S.O. 2015. To beat or not to beat a tick: comparison of DNA extraction methods for ticks (*Ixodes scapularis*). PeerJ, vol. 3, pp. e1147.

Andrew, H.R. & Norval, R.A. 1989. The carrier status of sheep, cattle and African buffalo recovered from heartwater. Veterinary parasitology, vol. 34, no. 3, pp. 261-266.

Asselbergs, M., Jongejan, F., Langa, A., Neves, L. & Afonso, S. 1993. Antibodies to *Cowdria ruminantium* in Mozambican goats and cattle detected by immunofluorescence using endothelial cell culture antigen. Tropical animal health and production, vol. 25, no. 3, pp. 144-150.

Atanasio, A. 2000, Helminths, protozoa, heartwater and the effect of gastro-intestinal nematodes on productivity of goats of the family sector in Mozambique. PhD thesis, Medical University of Southern Africa.

Barré, N. & Pavis, C. 1992. Essai d'attraction d'*Amblyomma variegatum* (Acarina: Ixodina) sur des bovins préalablement traités avec des phéromones d'agrégation-fixation et un acaricide pyréthrinoïde. Revue d'élevage et de médecine vétérinaire des pays tropicaux, vol. 45, no. 1, pp. 33-36.

Barré, N., Garris, G. & Camus, E. 1995.Propagation of the tick *Amblyomma variegatum* in the Caribbean. Revue scientifique et technique (International Office of Epizootics), vol. 14, no. 3, pp. 841-855.

Barré, N., Uilenberg, G., Morel, P.C. & Camus, E. 1987.Danger of introducing heartwater onto the American mainland: potential role of indigenous and exotic *Amblyomma* ticks. The Onderstepoort journal of veterinary research, vol. 54, no. 3, pp. 405-417.

Barré, N., Uilenberg, G., Morel, P.C. & Camus, E. 1987.Danger of introducing heartwater onto the American mainland: potential role of indigenous and exotic *Amblyomma* ticks. The Onderstepoort journal of veterinary research, vol. 54, no. 3, pp. 405-417.

Bartlett, P.C. & Judge, L.J. 1997. The role of epidemiology in public health. Revue scientifique et technique (International Office of Epizootics). vol. 16, no. 2, pp. 331-336.

Bekker, C.P., de Vos, S.F., Taoufik, A., Sparagano, O.A. & Jongejan, F. 2002. Simultaneous detection of *Anaplasma* and *Ehrlichia* species in ruminants and detection of *Ehrlichia ruminantium* in *Amblyomma variegatum* ticks by reverse line blot hybridization. Veterinary microbiology JID - 7705469, .

Bekker, C.P., Postigo, M., Taoufik, A., Bell-Sakyi, L., Ferraz, C., Martinez, D. & Jongejan, F. 2005. Transcription analysis of the major antigenic protein 1 multigene family of three in vitro-cultured *Ehrlichia ruminantium* isolates. Journal of Bacteriology, vol. 187, no. 14, pp. 4782-4791.

Bekker, C.P., Vink, D., Lopes Pereira, C.M., Wapenaar, W., Langa, A. & Jongejan, F. 2001.Heartwater (*Cowdria ruminantium* infection) as a cause of postrestocking mortality of

goats in Mozambique. Clinical and diagnostic laboratory immunology, vol. 8, no. 4, pp. 843-846.

Bezuidenhout, J.D. 1987. Natural transmission of heartwater. The Onderstepoort journal of veterinary research, vol. 54, no. 3, pp. 349-351.

Bila, C.J., De Deus, N., Fafetine, J.M., Dimande, A. & Neves, L. 2003. Investigação preliminar sobre as causas de mortalidade de caprinos provenientes de Tete no Sul de Moçambique. III Seminário de Investigaçao, Direcção Cientifica, Universidade Eduardo, pp. 31-36.

Black, W.C. 4th & Piesman, J. 1994. Phylogeny of hard- and soft-tick taxa (Acari: Ixodida) based on mitochondrial 16S rDNA sequences. Proceedings of the National Academy of Sciences of the United States of America. vol. 91, no. 21, pp. 10034-10038.

Boesenberg-Smith, KA, Pessarakli, MM, Wolk, DM 2012. Assessment of DNA Yield and Purity: an Overlooked Detail of PCR Troubleshooting. Clinical Microbiology Newsletter, vol. 34, no. 1, pp. 1-6.

Bournez, L., Cangi, N., Lancelot, R., Pleydell, D.R., Stachurski, F., Bouyer, J., Martinez, D., Lefrancois, T., Neves, L. & Pradel, J. 2015. Parapatric distribution and sexual competition between two tick species, *Amblyomma variegatum* and *A. hebraeum* (Acari, Ixodidae), in Mozambique. Parasites & Vectors, vol. 8, pp. 504.

Bustin, S.A., Benes, V., Garson, J.A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M.W., Shipley, G.L., Vandesompele, J. & Wittwer, C.T. 2009. The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. Clinical chemistry, vol. 55, no. 4, pp. 611-622.

Camus, E. & Barré, N. 1992. The role of *Amblyomma variegatum* in the transmission of heartwater with special reference to Guadeloupe. Annals of the New York Academy of Sciences, vol. 653, pp. 33-41.

Cangi, N., Pinarello, V., Bournez, L., Lefrançois, T., Abina, E., Neves, L. & Vachiéry, N. 2016 submitted. Efficient high-throughput molecular method to detect *Ehrlichia ruminantium* in ticks

Cangi, N., Gordon, J.L., Bournez, L., Pinarello, V., Aprelon, R., Huber, K., Lefrancois, T., Neves, L., Meyer, D.F. & Vachiéry, N. 2016. Recombination Is a Major Driving Force of Genetic Diversity in the Anaplasmataceae *Ehrlichia ruminantium*. Frontiers in cellular and infection microbiology, vol. 6, pp. 111.

Cannon, R.M. 2001. Sense and sensitivity-designing surveys based on an imperfect test. Preventive veterinary medicine, vol. 49, no. 3-4, pp. 141-163.

Collins, N.E., Allsopp, M.T. & Allsopp, B.A. 2002. Molecular diagnosis of theileriosis and heartwater in bovines in Africa. Transactions of the Royal Society of Tropical Medicine and Hygiene, vol. 96 Suppl 1, pp. S217-24.

Collins, N.E., Liebenberg, J., de Villiers, E.P., Brayton, K.A., Louw, E., Pretorius, A., Faber, F.E., van Heerden, H., Josemans, A., van Kleef, M., Steyn, H.C., van Strijp, M.F., Zweygarth, E., Jongejan, F., Maillard, J.C., Berthier, D., Botha, M., Joubert, F., Corton, C.H., Thomson, N.R., Allsopp, M.T. & Allsopp, B.A. 2005. The genome of the heartwater agent *Ehrlichia*

ruminantium contains multiple tandem repeats of actively variable copy number. Proceedings of the National Academy of Sciences of the United States of America, vol. 102, no. 3, pp. 838-843.

Cowdry, E.V. 1925a. Studies on the Etiology of Heartwater: I. Observation of a Rickettsia, *Rickettsia Ruminantium* (N. Sp.), in the Tissues of Infected Animals. The Journal of experimental medicine, vol. 42, no. 2, pp. 231-252.

Cowdry, E.V. 1925b. Studies on the Etiology of Heartwater: Ii. *Rickettsia Ruminantium* (N. Sp.) in the Tissues of Ticks Transmitting the Disease. The Journal of experimental medicine, vol. 42, no. 2, pp. 253-274.

Crowder, C.D., Rounds, M.A., Phillipson, C.A., Picuri, J.M., Matthews, H.E., Halverson, J., Schutzer, S.E., Ecker, D.J. & Eshoo, M.W. 2010. Extraction of total nucleic acids from ticks for the detection of bacterial and viral pathogens. Journal of medical entomology, vol. 47, no. 1, pp. 89-94.

de Vicente, M.C., Guzman, F.A., Engels, J. & Rao, V.R. 2006. 12. Genetic characterization and its use in decision-making for the conservation of crop germplasm. The role of biotechnology in exploring and protecting agricultural genetic resources, pp. 129.

Deem, S.L., Noval, R., Yonow, T., Peter, T.F., Mahan, S.M. & Burridge, M.J. 1996. The epidemiology of heartwater: establishment and maintenance of endemic stability. Parasitology today, vol. 12, no. 10, pp. 402-405.

Dias, J.A.T.S. 1991. Some data concerning the ticks (Acarina-Ixodoidea) presently known in Mozambique. Garcia de Orta Série de Zoologia, vol. 18 (1-2), pp. 27-48.

Du Plessis, J.L. 1970. Pathogenesis of heartwater. I. *Cowdria ruminantium* in the lymph nodes of domestic ruminants. The Onderstepoort journal of veterinary research, vol. 37, no. 2, pp. 89-95.

Du Plessis, J.L., Bezuidenhout, J.D., Brett, M.S., Camus, E., Jongejan, F., Mahan, S.M. & Martinez, D. 1993. The sero-diagnosis of heartwater: a comparison of five tests. Revue d'elevage et de medecine veterinaire des pays tropicaux, vol. 46, no. 1-2, pp. 123-129.

Du Plessis, J.L. & Malan, L. 1987. The application of the indirect fluorescent antibody test in research on heartwater. The Onderstepoort journal of veterinary research, vol. 54, no. 3, pp. 319-325.

Dumler, J.S., Barbet, A.F., Bekker, C.P., Dasch, G.A., Palmer, G.H., Ray, S.C., Rikihisa, Y. & Rurangirwa, F.R. 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 'HGE agent' as subjective synonyms of *Ehrlichia phagocytophila*. International Journal of Systematic and Evolutionary Microbiology, vol. 51, no. Pt 6, pp. 2145-2165.

Esemu, S.N., Besong, W.O., Ndip, R.N. & Ndip, L.M. 2013. Prevalence of *Ehrlichia ruminantium* in adult *Amblyomma variegatum* collected from cattle in Cameroon. Experimental & applied acarology, vol. 59, no. 3, pp. 377-387.

Espy, M.J., Uhl, J.R., Sloan, L.M., Buckwalter, S.P., Jones, M.F., Vetter, E.A., Yao, J.D., Wengenack, N.L., Rosenblatt, J.E., Cockerill, F.R. 3rd & Smith, T.F. 2006. Real-time PCR in clinical microbiology: applications for routine laboratory testing. Clinical microbiology reviews, vol. 19, no. 1, pp. 165-256.

Estrada-Peña, A. & de la Fuente, J. 2014. The ecology of ticks and epidemiology of tick-borne viral diseases. Antiviral Research, vol. 108, pp. 104-128.

Estrada-Peña, A., Pegram, R.G., Barré, N. & Venzal, J.M. 2007. Using invaded range data to model the climate suitability for *Amblyomma variegatum* (Acari: Ixodidae) in the New World. Experimental & applied acarology, vol. 41, no. 3, pp. 203-214.

Faburay, B., Geysen, D., Ceesay, A., Marcelino, I., Alves, P.M., Taoufik, A., Postigo, M., Bell-Sakyi, L. & Jongejan, F. 2007a. Immunisation of sheep against heartwater in The Gambia using inactivated and attenuated *Ehrlichia ruminantium* vaccines. Vaccine, vol. 25, no. 46, pp. 7939-7947.

Faburay, B., Geysen, D., Munstermann, S., Taoufik, A., Postigo, M. & Jongejan, F. 2007b. Molecular detection of *Ehrlichia ruminantium* infection in *Amblyomma variegatum* ticks in The Gambia. Experimental & applied acarology, vol. 42, no. 1, pp. 61-74.

Faburay, B., Jongejan, F., Taoufik, A., Ceesay, A. & Geysen, D. 2008. Genetic diversity of *Ehrlichia ruminantium* in *Amblyomma variegatum* ticks and small ruminants in The Gambia determined by restriction fragment profile analysis. Veterinary microbiology, vol. 126, no. 1-3, pp. 189-199.

Feil, E., Zhou, J., Maynard Smith, J. & Spratt, B.G. 1996. A comparison of the nucleotide sequences of the adk and recA genes of pathogenic and commensal *Neisseria* species: evidence for extensive interspecies recombination within adk. Journal of Molecular Evolution, vol. 43, no. 6, pp. 631-640.

Feil, E.J., Maiden, M.C., Achtman, M. & Spratt, B.G. 1999. The relative contributions of recombination and mutation to the divergence of clones of *Neisseria meningitidis*. Molecular biology and evolution, vol. 16, no. 11, pp. 1496-1502.

Feil, E.J. & Spratt, B.G. 2001. Recombination and the population structures of bacterial pathogens. Annual Review of Microbiology, vol. 55, pp. 561-590.

Fierz, W. 2004. Basic problems of serological laboratory diagnosis. Methods in Molecular Medicine, vol. 94, pp. 393-427.

Fleiss, J.L., Levin, B. & Paik, M.C. 2004. "The Measurement of Interrater Agreement" in Statistical Methods for Rates and Proportions, eds. J.L. Fleiss, B. Levin & M.C. Paik, John Wiley & Sons, Inc., , pp. 598-626.

Frutos, R., Viari, A., Ferraz, C., Bensaid, A., Morgat, A., Boyer, F., Coissac, E., Vachiéry, N., Demaille, J. & Martinez, D. 2006. Comparative genomics of three strains of *Ehrlichia ruminantium*: a review. Annals of the New York Academy of Sciences, vol. 1081, pp. 417-433.

Frutos, R., Viari, A., Vachiéry, N., Boyer, F. & Martinez, D. 2007. *Ehrlichia ruminantium*: genomic and evolutionary features. Trends in parasitology, vol. 23, no. 9, pp. 414-419.

Hajibabaei, M., deWaard, J.R., Ivanova, N.V., Ratnasingham, S., Dooh, R.T., Kirk, S.L., Mackie, P.M. & Hebert, P.D. 2005. Critical factors for assembling a high volume of DNA barcodes. Philosophical transactions of the Royal Society of London.Series B, Biological sciences, vol. 360, no. 1462, pp. 1959-1967.

Halos, L., Jamal, T., Vial, L., Maillard, R., Suau, A., Le Menach, A., Boulouis, H.J. & Vayssier-Taussat, M. 2004. Determination of an efficient and reliable method for DNA extraction from ticks. Veterinary research, vol. 35, no. 6, pp. 709-713.

Heid, C.A., Stevens, J., Livak, K.J. & Williams, P.M. 1996. Real time quantitative PCR. Genome research, vol. 6, no. 10, pp. 986-994.

Holmes, E.C., Urwin, R. & Maiden, M.C. 1999. The influence of recombination on the population structure and evolution of the human pathogen *Neisseria meningitidis*. Molecular biology and evolution, vol. 16, no. 6, pp. 741-749.

Horak, I., Fourie, L. & Van Zyl, J. 1995. Arthropod parasites of impalas in the Kruger National Park with particular reference to ticks. South African Journal of Wildlife Research, vol. 25, pp. 123-126.

Hughes, A.L. & French, J.O. 2007. Homologous recombination and the pattern of nucleotide substitution in *Ehrlichia ruminantium*. Gene, vol. 387, no. 1-2, pp. 31-37.

Ivanova, N.V., Dewaard, J.R., Hebert, P.D.N. 2006. An inexpensive, automation-friendly protocol for recovering high-quality DNA. Molecular Ecology Notes, vol. 6, no. 4, pp. 998-1002.

Javadi, A., Shamaei, M., Mohammadi Ziazi, L., Pourabdollah, M., Dorudinia, A., Seyedmehdi, S.M. & Karimi, S. 2014. Qualification Study of Two Genomic DNA Extraction Methods in Different Clinical Samples. Tanaffos, vol. 13, no. 4, pp. 41-47.

Jongejan, F., Thielemans, M.J., De Groot, M., van Kooten, P.J. & van der Zeijst, B.A. 1991a. Competitive enzyme-linked immunosorbent assay for heartwater using monoclonal antibodies to a *Cowdria ruminantium*-specific 32-kilodalton protein. Veterinary microbiology, vol. 28, no. 2, pp. 199-211.

Jongejan, F., Zandbergen, T.A., van de Wiel, P.A., de Groot, M. & Uilenberg, G. 1991b. The tick-borne rickettsia *Cowdria ruminantium* has a Chlamydia-like developmental cycle. The Onderstepoort journal of veterinary research, vol. 58, no. 4, pp. 227-237.

Karrar, G. 1986. Epizootiological studies on heartwater in the Sudan. The Sudan journal of veterinary science and animal husbandry, 9:328-343.

Katz, J.B., DeWald, R., Dawson, J.E., Camus, E., Martinez, D. & Mondry, R. 1997. Development and evaluation of a recombinant antigen, monoclonal antibody-based competitive ELISA for heartwater serodiagnosis. Journal of veterinary diagnostic investigation, vol. 9, no. 2, pp. 130-135.

Kock, N.D., van Vliet, A.H., Charlton, K. & Jongejan, F. 1995. Detection of *Cowdria ruminantium* in blood and bone marrow samples from clinically normal, free-ranging Zimbabwean wild ungulates. Journal of clinical microbiology, vol. 33, no. 9, pp. 2501-2504.

Kubista, M., Andrade, J.M., Bengtsson, M., Forootan, A., Jonak, J., Lind, K., Sindelka, R., Sjoback, R., Sjogreen, B., Strombom, L., Stahlberg, A. & Zoric, N. 2006. The real-time polymerase chain reaction. Molecular aspects of medicine, vol. 27, no. 2-3, pp. 95-125.

Lalkhen, A.G., McCluskey, A. 2008. Clinical tests: sensitivity and specificity. Continuing Education in Anaesthesia, Critical Care & Pain, vol. 8, pp. 221-223.

Land, M., Hauser, L., Jun, S.R., Nookaew, I., Leuze, M.R., Ahn, T.H., Karpinets, T., Lund, O., Kora, G., Wassenaar, T., Poudel, S. & Ussery, D.W. 2015. Insights from 20 years of bacterial genome sequencing. Functional & Integrative Genomics, vol. 15, no. 2, pp. 141-161.

Lee, J.H., Park, H.S., Jang, W.J., Koh, S.E., Kim, J.M., Shim, S.K., Park, M.Y., Kim, Y.W., Kim, B.J., Kook, Y.H., Park, K.H. & Lee, S.H. 2003. Differentiation of rickettsiae by groEL gene analysis. Journal of clinical microbiology, vol. 41, no. 7, pp. 2952-2960.

Lindstedt, B.A. 2005. Multiple-locus variable number tandem repeats analysis for genetic fingerprinting of pathogenic bacteria. Electrophoresis, vol. 26, no. 13, pp. 2567-2582.

Lorusso, V., Wijnveld, M., Majekodunmi, A.O., Dongkum, C., Fajinmi, A., Dogo, A.G., Thrusfield, M., Mugenyi, A., Vaumourin, E., Igweh, A.C., Jongejan, F., Welburn, S.C. & Picozzi, K. 2016. Tick-borne pathogens of zoonotic and veterinary importance in Nigerian cattle. Parasites & Vectors, vol. 9, pp. 217-016-1504-7.

Lounsbury, C.P. 1900. Tick-heartwater experiment. Agricultural journal of the Cape of Good Hope, vol. 16, pp. 682-687.

Louw, M., Allsopp, M.T. & Meyer, E.C. 2005. *Ehrlichia ruminantium*, an emerging human pathogen-a further report. South African medical journal = Suid-Afrikaanse tydskrif vir geneeskunde, vol. 95, no. 12, pp. 948, 950.

Lynn, G.E., Oliver, J.D., Nelson, C.M., Felsheim, R.F., Kurtti, T.J. & Munderloh, U.G. 2015. Tissue distribution of the *Ehrlichia muris*-like agent in a tick vector. PloS one, vol. 10, no. 3, pp. e0122007.

Mahan, S.M., Peter, T.F., Semu, S.M., Simbi, B.H., Norval, R.A. & Barbet, A.F. 1995. Laboratory reared *Amblyomma hebraeum* and *Amblyomma variegatum* ticks differ in their susceptibility to infection with *Cowdria ruminantium*. Epidemiology and infection, vol. 115, no. 2, pp. 345-353.

Mahan, S.M., Peter, T.F., Simbi, B.H., Kocan, K., Camus, E., Barbet, A.F. & Burridge, M.J. 2000. Comparison of efficacy of American and African *Amblyomma* ticks as vectors of heartwater (*Cowdria ruminantium*) infection by molecular analyses and transmission trials. The Journal of Parasitology, vol. 86, no. 1, pp. 44-49.

Mahan, S.M., Tebele, N., Mukwedeya, D., Semu, S., Nyathi, C.B., Wassink, L.A., Kelly, P.J., Peter, T. & Barbet, A.F. 1993. An immunoblotting diagnostic assay for heartwater based on the immunodominant 32-kilodalton protein of Cowdria ruminantium detects false positives in field sera. Journal of clinical microbiology, vol. 31, no. 10, pp. 2729-2737.

Mahan, S.M., Waghela, S.D., McGuire, T.C., Rurangirwa, F.R., Wassink, L.A. & Barbet, A.F. 1992. A cloned DNA probe for *Cowdria ruminantium* hybridizes with eight heartwater strains and detects infected sheep. Journal of clinical microbiology, vol. 30, no. 4, pp. 981-986.

Maiden, M.C., Bygraves, J.A., Feil, E., Morelli, G., Russell, J.E., Urwin, R., Zhang, Q., Zhou, J., Zurth, K., Caugant, D.A., Feavers, I.M., Achtman, M. & Spratt, B.G. 1998. Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. Proceedings of the National Academy of Sciences of the United States of America, vol. 95, no. 6, pp. 3140-3145.

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

Marcelino, I., de Almeida, A.M., Ventosa, M., Pruneau, L., Meyer, D.F., Martinez, D., Lefrancois, T., Vachiéry, N. & Coelho, A.V. 2012. Tick-borne diseases in cattle: applications of proteomics to develop new generation vaccines. Journal of proteomics, vol. 75, no. 14, pp. 4232-4250.

Marcelino, I., Sousa, M.F., Verissimo, C., Cunha, A.E., Carrondo, M.J. & Alves, P.M. 2006. Process development for the mass production of *Ehrlichia ruminantium*. Vaccine, vol. 24, no. 10, pp. 1716-1725.

Marcelino, I., Verissimo, C., Sousa, M.F., Carrondo, M.J. & Alves, P.M. 2005. Characterization of *Ehrlichia ruminantium* replication and release kinetics in endothelial cell cultures. Veterinary microbiology, vol. 110, no. 1-2, pp. 87-96.

Martinez, D., Uilenberg, G. 2010. “Cowdriosis (heartwater)" in Infectious and parasitic diseases of livestock. General considerations. Viral diseases, ed. Lefèvre, P.-C., Blancou, J., Chermette, R., Uilenberg, G., Paris: Lavoisier Tec et Doc, pp. 1265-1288.

Martinez, D., Vachiéry, N., Stachurski, F., Kandassamy, Y., Raliniaina, M., Aprelon, R. & Gueye, A. 2004a. Nested PCR for detection and genotyping of *Ehrlichia ruminantium*: use in genetic diversity analysis. Annals of the New York Academy of Sciences, vol. 1026, pp. 106-113.

Martinez, D., Vachiéry, N., Stachurski, F., Kandassamy, Y., Raliniaina, M., Aprelon, R. & Gueye, A. 2004b. Nested PCR for detection and genotyping of *Ehrlichia ruminantium*: Use in genetic diversity analysis. Annals of the New York Academy of Sciences, vol. 1026, pp. 106-113.

Mathew, J.S., Ewing, S.A., Barker, R.W., Fox, J.C., Dawson, J.E., Warner, C.K., Murphy, G.L. & Kocan, K.M. 1996. Attempted transmission of *Ehrlichia canis* by *Rhipicephalus sanguineus* after passage in cell culture. American Journal of Veterinary Research, vol. 57, no. 11, pp. 1594-1598.

Michelet, L., Delannoy, S., Devillers, E., Umhang, G., Aspan, A., Juremalm, M., Chirico, J., van der Wal, F.J., Sprong, H., Boye Pihl, T.P., Klitgaard, K., Bodker, R., Fach, P. & Moutailler, S. 2014. High-throughput screening of tick-borne pathogens in Europe. Frontiers in cellular and infection microbiology, vol. 4, pp. 103.

Minjauw, B. 2000. The economic impact of heartwater (*Cowdria ruminantium*) infection in the SADC region, and its control through the use of new inactivated vaccines. ILRI, UF/USAID/SADC Heartwater Research Project Report.

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

Molia, S., Frebling, M., Vachiéry, N., Pinarello, V., Petitclerc, M., Rousteau, A., Martinez, D. & Lefrancois, T. 2008. *Amblyomma variegatum* in cattle in Marie Galante, French Antilles: prevalence, control measures, and infection by *Ehrlichia ruminantium*. Veterinary parasitology, vol. 153, no. 3-4, pp. 338-346.

Moriarity, J.R., Loftis, A.D. & Dasch, G.A. 2005. High-throughput molecular testing of ticks using a liquid-handling robot. Journal of medical entomology, vol. 42, no. 6, pp. 1063-1067.

Moshkovski, S.D. 1947. Comments by readers. Science, vol. 106: 62.

Nakao, R., Magona, J.W., Zhou, L., Jongejan, F. & Sugimoto, C. 2011. Multi-locus sequence typing of *Ehrlichia ruminantium* strains from geographically diverse origins and collected in Amblyomma variegatum from Uganda. Parasites & Vectors, vol. 4, pp. 137.

Nakao, R., Morrison, L.J., Zhou, L., Magona, J.W., Jongejan, F. & Sugimoto, C. 2012. Development of multiple-locus variable-number tandem-repeat analysis for rapid genotyping of *Ehrlichia ruminantium* and its application to infected *Amblyomma variegatum* collected in heartwater endemic areas in Uganda. Parasitology, vol. 139, no. 1, pp. 69-82.

Neitz, W.O. 1967. The epidemiological pattern of viral, protophytal and protozoal zoonoses in relation to game preservation in South Africa. Journal of the South African Veterinary Association. 38, pp. 129–141.

Neitz, W.O. 1935. The blesbuck (*Damaliscus albifrons*) and the black-wildebeest (*Conochaetes gnu*) as carriers of heartwater. The Onderstepoort Journal of Veterinary Research. vol. 5, pp. 35–40.

Njiiri, N.E., Bronsvoort, B.M.d., Collins, N.E., Steyn, H.C., Troskie, M., Vorster, I., Thumbi, S.M., Sibeko, K.P., Jennings, A., van Wyk, I.C., Mbole-Kariuki, M., Kiara, H., Poole, E.J., Hanotte, O., Coetzer, K., Oosthuizen, M.C., Woolhouse, M. & Toye, P. 2015. The epidemiology of tick-borne haemoparasites as determined by the reverse line blot hybridization assay in an intensively studied cohort of calves in western Kenya. Veterinary parasitology, vol. 210, no. 1–2, pp. 69-76.

Norris, DE, Klompen, JSH, Keirans, JE, Black, WC. 1996. Population genetics of *Ixodes scapularis* (Acari: Ixodidae) based on mitochondrial 16S and 12S genes. Journal of Medical Entomology, vol. 33, pp. 78-89.

Norval, R.A.I., Perry, B.D. & Young, A. 1992. The epidemiology of theileriosis in Africa. ILRI (aka ILCA and ILRAD).

Norval, R. A., Andrew, H. R., Yunker, C. E. & Burridge, M. J. 1992. "Biological Processes in the epidemiology of Heartwater" in Tick Vector Biology: Medical and Veterinary Aspects, ed. Fivaz, B.H., Petney, T.N., Horak, I.G., Springer-Verlag, Berlin, pp. 71-86.

Norval, R.A.I. 1983. The ticks of Zimbabwe VII. The genus Amblyomma. Zimbabwe Veterinary Journal, vol. 14, pp. 5-18.

OIE 2016. OIE-Listed diseases, infections and infestations in force in 2016 [online]. Available at: http://www.oie.int/animal-health-in-the-world/oie-listed-diseases-2016/, vol. [Acessed October 2nd 2016].

Paddock, C.D., Fournier, P.E., Sumner, J.W., Goddard, J., Elshenawy, Y., Metcalfe, M.G., Loftis, A.D. & Varela-Stokes, A. 2010. Isolation of *Rickettsia parkeri* and identification of a novel spotted fever group *Rickettsia sp.* from Gulf Coast ticks (*Amblyomma maculatum*) in the United States. Applied and Environmental Microbiology, vol. 76, no. 9, pp. 2689-2696.

Peixoto, C.C., Marcelino, I., Vachiéry, N., Bensaid, A., Martinez, D., Carrondo, M.J. & Alves, P.M. 2005. Quantification of *Ehrlichia ruminantium* by real time PCR. Veterinary microbiology, vol. 107, no. 3-4, pp. 273-278.

Peter, T.F., Anderson, E.C., Burridge, M.J. & Mahan, S.M. 1998. Demonstration of a carrier state for *Cowdria ruminantium* in wild ruminants from Africa. Journal of wildlife diseases, vol. 34, no. 3, pp. 567-575.

Peter, T.F., Barbet, A.F., Alleman, A.R., Simbi, B.H., Burridge, M.J. & Mahan, S.M. 2000. Detection of the agent of heartwater, *Cowdria ruminantium*, in *Amblyomma* ticks by PCR: validation and application of the assay to field ticks. Journal of clinical microbiology, vol. 38, no. 4, pp. 1539-1544.

Peter, T.F., Bryson, N.R., Perry, B.D., O'Callaghan, C.J., Medley, G.F., Smith, G.E., Mlambo, G., Horak, I.G., Burridge, M.J. & Mahan, S.M. 1999. *Cowdria ruminantium* infection in ticks in the Kruger National Park. The Veterinary record, vol. 145, no. 11, pp. 304-307.

Peter, T.F., Burridge, M.J. & Mahan, S.M. 2002. *Ehrlichia ruminantium* infection (heartwater) in wild animals. Trends in parasitology, vol. 18, no. 5, pp. 214-218.

Peter, T.F., Deem, S.L., Barbet, A.F., Norval, R.A., Simbi, B.H., Kelly, P.J. & Mahan, S.M. 1995. Development and evaluation of PCR assay for detection of low levels of *Cowdria ruminantium* infection in *Amblyomma* ticks not detected by DNA probe. Journal of clinical microbiology, vol. 33, no. 1, pp. 166-172.

Peter, T.F., Perry, B.D., O'Callaghan, C.J., Medley, G.F., Mlambo, G., Barbet, A.F. & Mahan, S.M. 1999. Prevalence of *Cowdria ruminantium* infection in *Amblyomma hebraeum* ticks from heartwater-endemic areas of Zimbabwe. Epidemiology and infection, vol. 123, no. 2, pp. 309-316.

Pilet, H., Vachiéry, N., Berrich, M., Bouchouicha, R., Durand, B., Pruneau, L., Pinarello, V., Saldana, A., Carasco-Lacombe, C., Lefrancois, T., Meyer, D.F., Martinez, D., Boulouis, H.J. & Haddad, N. 2012. A new typing technique for the Rickettsiales *Ehrlichia ruminantium*: multiple-locus variable number tandem repeat analysis. Journal of microbiological methods, vol. 88, no. 2, pp. 205-211.

Pornwiroon, W., Pourciau, S.S., Foil, L.D. & Macaluso, K.R. 2006. *Rickettsia felis* from cat fleas: isolation and culture in a tick-derived cell line. Applied and Environmental Microbiology, vol. 72, no. 8, pp. 5589-5595.

Posada, D. & Crandall, K.A. 2002. The effect of recombination on the accuracy of phylogeny estimation. Journal of Molecular Evolution, vol. 54, no. 3, pp. 396-402.

Postigo, M., Taoufik, A., Bell-Sakyi, L., de Vries, E., Morrison, W.I. & Jongejan, F. 2007. Differential transcription of the major antigenic protein 1 multigene family of *Ehrlichia ruminantium* in *Amblyomma variegatum* ticks. Veterinary microbiology, vol. 122, no. 3-4, pp. 298-305.

Provost, A. & Bezuidenhout, J.D. 1987. The historical background and global importance of heartwater. The Onderstepoort journal of veterinary research, vol. 54, no. 3, pp. 165-169.

Prozesky, L. & Du Plessis, J.L. 1987. Heartwater. The development and life cycle of *Cowdria ruminantium* in the vertebrate host, ticks and cultured endothelial cells. The Onderstepoort journal of veterinary research, vol. 54, no. 3, pp. 193-196.

Pruneau, L., Emboule, L., Gely, P., Marcelino, I., Mari, B., Pinarello, V., Sheikboudou, C., Martinez, D., Daigle, F., Lefrancois, T., Meyer, D.F. & Vachiéry, N. 2012. Global gene expression profiling of *Ehrlichia ruminantium* at different stages of development. FEMS immunology and medical microbiology, vol. 64, no. 1, pp. 66-73.

R Core Team 2013, R: A language and environment for statistical computing. http://www.R-project.org/., R Foundation for Statistical Computing, Vienna, Austria.

Radstrom, P., Knutsson, R., Wolffs, P., Lovenklev, M. & Lofstrom, C. 2004. Pre-PCR processing: strategies to generate PCR-compatible samples. Molecular biotechnology, vol. 26, no. 2, pp. 133-146.

Raliniaina, M., Meyer, D.F., Pinarello, V., Sheikboudou, C., Emboulé, L., Kandassamy, Y., Adakal, H., Stachurski, F., Martinez, D., Lefrançois, T. & Vachiéry, N. 2010. Mining the genetic diversity of *Ehrlichia ruminantium* using map genes family. Veterinary parasitology, vol. 167, no. 2-4, pp. 187-195.

Redfield, R.J. 2001. Do bacteria have sex? Nature reviews.Genetics, vol. 2, no. 8, pp. 634-639.

Rodriguez-Perez, M.A., Gopal, H., Adeleke, M.A., De Luna-Santillana, E.J., Gurrola-Reyes, J.N. & Guo, X. 2013. Detection of Onchocerca volvulus in Latin American black flies for pool screening PCR using high-throughput automated DNA isolation for transmission surveillance. Parasitology research, vol. 112, no. 11, pp. 3925-3931.

Roth, J.A., Richt, J.A. & Morozov, I.A. 2013. “Vaccines and Diagnostics for Transboundary Animal Diseases" in Developments in Biologicals, ed. Roth, J.A, Richt, J.A., Morozov, I.A. 2012, Ames, Iowa.

Rothman, K. 2012, Epidemiology: An Introduction. 2nd ed. Oxford University Press , USA.

Ruths, D. & Nakhleh, L. 2005. Recombination and phylogeny: effects and detection. International journal of bioinformatics research and applications, vol. 1, no. 2, pp. 202-212.

Salih, D.A., EI Hussein, A.M. & Singla, L.D. 2015. Diagnostic approaches for tick-borne haemoparasitic diseases in livestock. Journal of Veterinary Medicine and Animal Health, vol. 7(2), pp. 45-56.

Salipante, S.J., SenGupta, D.J., Cummings, L.A., Land, T.A., Hoogestraat, D.R. & Cookson, B.T. 2015. Application of Whole-Genome Sequencing for Bacterial Strain Typing in Molecular Epidemiology. Journal of clinical microbiology, vol. 53, no. 4, pp. 1072-1079.

Sambrook, J, Fritschi, EF, Maniatis, T 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, New York.

Sayler, K.A., Loftis, A.D., Mahan, S.M. & Barbet, A.F. 2016. Development of a Quantitative PCR Assay for Differentiating the Agent of Heartwater Disease, *Ehrlichia ruminantium*, from the Panola Mountain Ehrlichia. Transboundary and Emerging Diseases, vol. 63, no. 6, pp. e260-e269.

Schierup, M.H. & Hein, J. 2000. Consequences of recombination on traditional phylogenetic analysis. Genetics, vol. 156, no. 2, pp. 879-891.

Shaw, K.J., Thain, L., Docker, P.T., Dyer, C.E., Greenman, J., Greenway, G.M. & Haswell, S.J. 2009. The use of carrier RNA to enhance DNA extraction from microfluidic-based silica monoliths. Analytica Chimica Acta, vol. 652, no. 1-2, pp. 231-233.

Shriner, D., Nickle, D.C., Jensen, M.A. & Mullins, J.I. 2003. Potential impact of recombination on sitewise approaches for detecting positive natural selection. Genetical research, vol. 81, no. 2, pp. 115-121.

Smith, J.M. & Smith, N.H. 1998. Detecting recombination from gene trees. Molecular biology and evolution, vol. 15, no. 5, pp. 590-599.

Smith, R.D., Evans, D.E., Martins, J.R., Cereser, V.H., Correa, B.L., Petraccia, C., Cardozo, H., Solari, M.A. & Nari, A. 2000. Babesiosis (*Babesia bovis*) stability in unstable environments. Annals of the New York Academy of Sciences, vol. 916, pp. 510-520.

Spratt, B.G., Smith, N.H., Zhou, J.J., O'Rourke, M. & Feil, E. 1995. “The population genetics of the pathogenic Neisseria” in Population genetics of bacteria, ed. Baumberg, S., Young, J.P.W., Wellington, E.M.H., Saunders, J.R., Cambridge University Press, pp. 143-160.

Stachurski, F., Tortosa, P., Rahajarison, P., Jacquet, S., Yssouf, A. & Huber, K. 2013. New data regarding distribution of cattle ticks in the south-western Indian Ocean islands. Veterinary research, vol. 44, pp. 79.

Stahl, D.A. & Kane, M.D. 1992. Methods of microbial identification, tracking and monitoring of function. Current opinion in biotechnology, vol. 3, no. 3, pp. 244-252.

Steyn, H.C., Pretorius, A., McCrindle, C.M., Steinmann, C.M. & Van Kleef, M. 2008. A quantitative real-time PCR assay for *Ehrlichia ruminantium* using pCS20.Veterinary microbiology, vol. 131, no. 3-4, pp. 258-265.

Sullivan, C.B., Diggle, M.A. & Clarke, S.C. 2005. Multilocus sequence typing: Data analysis in clinical microbiology and public health. Molecular biotechnology, vol. 29, no. 3, pp. 245-254.

Sumner, J.W., Nicholson, W.L. & Massung, R.F. 1997. PCR amplification and comparison of nucleotide sequences from the groESL heat shock operon of *Ehrlichia* species. Journal of clinical microbiology, vol. 35, no. 8, pp. 2087-2092.

Svec, D., Tichopad, A., Novosadova, V., Pfaffl, M.W. & Kubista, M. 2015. How good is a PCR efficiency estimate: Recommendations for precise and robust qPCR efficiency assessments. Biomolecular Detection and Quantification, vol. 3, pp. 9.

Teshale, S., Geysen, D., Ameni, G., Asfaw, Y. & Berkvens, D. 2015. Improved molecular detection of *Ehrlichia* and *Anaplasma* species applied to *Amblyomma* ticks collected from cattle and sheep in Ethiopia. Ticks and tick-borne diseases, vol. 6, no. 1, pp. 1-7.

Tice, G.A., Bryson, N.R., Stewart, C.G., Du Plessis, B. & De Wall, D.T. 1998. The absence of clinical disease in cattle in communal grazing areas where farmers are changing from an intensive dipping programme to one of endemic stability to tick-borne diseases. The Onderstepoort journal of veterinary research, vol. 65, no. 3, pp. 169-175.

Uilenberg, G. 1983. Heartwater (*Cowdria ruminantium* infection): current status. Advances in Veterinary Science and Comparative Medicine, vol. 27, pp. 427-480.

Uilenberg, G. 1982. Experimental transmission of *Cowdria ruminantium* by the Gulf coast tick *Amblyomma maculatum*: danger of introducing heartwater and benign African theileriasis onto the American mainland. American Journal of Veterinary Research, vol. 43, no. 7, pp. 1279-1282.

Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M., & Rozen, S. G. 2012. Primer3-new capabilities and interfaces. Nucleic Acids Research, vol. 40, no. 15, pp. e115.

Urwin, R. & Maiden, M.C. 2003. Multi-locus sequence typing: a tool for global epidemiology. Trends in microbiology, vol. 11, no. 10, pp. 479-487.

Vachiéry, N., Jeffery, H., Pegram, R., Aprelon, R., Pinarello, V., Kandassamy, R.L., Raliniaina, M., Molia, S., Savage, H., Alexander, R., Frebling, M., Martinez, D. & Lefrançois, T. 2008. *Amblyomma variegatum* ticks and heartwater on three Caribbean Islands. Annals of the New York Academy of Sciences, vol. 1149, pp. 191-195.

Vachiéry, N., Marcelino, I., Martinez, D. & Lefrancois, T. 2013. Opportunities in diagnostic and vaccine approaches to mitigate potential heartwater spreading and impact on the American mainland. Developments in biologicals, vol. 135, pp. 191-200.

Valadão, F.G. 1969. Occurrence of hyperacute heartwater in Mozambique and problems of premunition against this disease. Anais dos Serviços de Veterinaria de Moçambique, vol. 12–14, pp. 85-90.

Valasek, M.A. & Repa, J.J. 2005. The power of real-time PCR. Advances in Physiology Education, vol. 29, no. 3, pp. 151-159.

van Vliet, A.H., Jongejan, F. & van der Zeijst, B.A. 1992. Phylogenetic position of *Cowdria ruminantium* (Rickettsiales) determined by analysis of amplified 16S ribosomal DNA sequences. International Journal of Systematic Bacteriology, vol. 42, no. 3, pp. 494-498.

van Vliet, A.H., van der Zeijst, B.A., Camus, E., Mahan, S.M., Martinez, D. & Jongejan, F. 1995. Use of a specific immunogenic region on the *Cowdria ruminantium* MAP1 protein in a serological assay. Journal of clinical microbiology, vol. 33, no. 9, pp. 2405-2410.

Vergnaud, G. & Denoeud, F. 2000. Minisatellites: mutability and genome architecture. Genome research, vol. 10, no. 7, pp. 899-907.

Vidergar, N., Toplak, N. & Kuntner, M. 2014. Streamlining DNA barcoding protocols: automated DNA extraction and a new cox1 primer in arachnid systematics. PloS one, vol. 9, no. 11, pp. e113030.

Waghela, S.D., Rurangirwa, F.R., Mahan, S.M., Yunker, C.E., Crawford, T.B., Barbet, A.F., Burridge, M.J. & McGuire, T.C. 1991. A cloned DNA probe identifies *Cowdria ruminantium* in *Amblyomma variegatum* ticks. Journal of clinical microbiology, vol. 29, no. 11, pp. 2571-2577.

Walker, A.R. 2011. Eradication and control of livestock ticks: biological, economic and social perspectives. Parasitology, vol. 138, no. 8, pp. 945-959.

Walker, J.B. & Olwage, A. 1987. The tick vectors of *Cowdria ruminantium* (Ixodoidea, Ixodidae, genus Amblyomma) and their distribution. The Onderstepoort journal of veterinary research, vol. 54, no. 3, pp. 353-379.

Wesonga, F.D., Mukolwe, S.W. & Grootenhuis, J. 2001. Transmission of *Cowdria ruminantium* by *Amblyomma gemma* from infected African buffalo (*Syncerus caffer*) and eland (*Taurotragus oryx*) to sheep. Tropical animal health and production, vol. 33, no. 5, pp. 379-390.

Wilcoxon, F. 1945. Individual Comparisons by Ranking Methods. Biometrics Bulletin, vol. 1, no. 6, pp. 80-83.

Yu, X.., McBride, J.W. & Walker, D.H. 2007. Restriction and expansion of *Ehrlichia* strain diversity. Veterinary parasitology, vol. 143, no. 3-4, pp. 337-346.

Yunker, C.E., Mahan, S.M., Waghela, S.D., McGuire, T.C., Rurangirwa, F.R., Barbet, A.F. & Wassink, L.A. 1993. Detection of *Cowdria ruminantium* by means of a DNA probe, pCS20 in infected bont ticks, *Amblyomma hebraeum*, the major vector of heartwater in southern Africa. Epidemiology and infection, vol. 110, no. 1, pp. 95-104.

Zhou, J., Bowler, L.D. & Spratt, B.G. 1997. Interspecies recombination, and phylogenetic distortions, within the glutamine synthetase and shikimate dehydrogenase genes of *Neisseria meningitidis* and commensal Neisseria species. Molecular microbiology, vol. 23, no. 4, pp. 799-812.

Zweygarth, E., Josemans, A.I. & Steyn, H.C. 2008. Experimental use of the attenuated *Ehrlichia ruminantium* (Welgevonden) vaccine in Merino sheep and Angora goats. Vaccine, vol. 26, Supplement 6, pp. G34-G39.

IX. Annexe

Parapatric distribution and sexual competition between two tick species, *Amblyomma variegatum* and *A. hebraeum* (Acari, Ixodidae), in Mozambique

(Published in Parasites and Vectors)

Bournez et al. Parasites & Vectors (2015) 8:504
DOI 10.1186/s13071-015-1116-7

Parasites & Vectors

RESEARCH Open Access

Parapatric distribution and sexual competition between two tick species, *Amblyomma variegatum* and *A. hebraeum* (Acari, Ixodidae), in Mozambique

L. Bournez[1,2,3*], N. Cangi[1,2,3,4], R. Lancelot[2,5], D.R.J Pleydell[1,2], F. Stachurski[2,5], J. Bouyer[2,5,6], D. Martinez[7], T. Lefrançois[2,5], L. Neves[4,8] and J. Pradel[1,2]

Abstract

Background: *Amblyomma variegatum* and *A. hebraeum* are two ticks of veterinary and human health importance in south-east Africa. In Zimbabwe they occupy parapatric (marginally overlapping and juxtaposed) distributions. Understanding the mechanisms behind this parapatry is essential for predicting the spatio-temporal dynamics of *Amblyomma* spp. and the impacts of associated diseases. It has been hypothesized that exclusive competition between these species results from competition at the levels of male signal reception (attraction-aggregation-attachment pheromones) or sexual competition for mates. This hypothesis predicts that the parapatry described in Zimbabwe could also be present in other countries in the region.

Methods: To explore this competitive exclusion hypothesis we conducted field surveys at the two species' range limits in Mozambique to identify areas of sympatry (overlapping areas) and to study potential interactions (communicative and reproductive interference effects) in those areas. At sympatric sites, hetero-specific mating pairs were collected and inter-specific attractiveness/repellent effects acting at long and short distances were assessed by analyzing species co-occurrences on co-infested herds and co-infested hosts.

Results: Co-occurrences of both species at sampling sites were infrequent and localized in areas where both tick and host densities were low. At sympatric sites, high percentages of individuals of both species shared attachment sites on hosts and inter-specific mating rates were high. Although cross-mating rates were not significantly different for *A. variegatum* and *A. hebraeum* females, attraction towards hetero-specific males was greater for *A. hebraeum* females than for *A. variegatum* females and we observed small asymmetrical repellent effects between males at attachment sites.

Conclusions: Our observations suggest near-symmetrical reproductive interference between *A. variegatum* and *A. hebraeum*, despite between-species differences in the strength of reproductive isolation barriers acting at the aggregation, fixation and partner contact levels. Theoretical models predict that sexual competition coupled with hybrid inviability, greatly reduces the probability of one species becoming established in an otherwise suitable location when the other species is already established. This mechanism can explain why the parapatric boundary in Mozambique has formed within an area of low tick densities and relatively infrequent host-mediated dispersal events.

Keywords: Exclusive competition, Reproductive interference, Communicative interference, Parapatry, Hybrid inviability

* Correspondence: laurebournez@gmail.com
[1]CIRAD, UMR CMAEE, F-97170 Petit-Bourg, Guadeloupe, France
[2]INRA, UMR 1309 CMAEE, F-34398 Montpellier, France
Full list of author information is available at the end of the article

BioMed Central

© 2015 Bournez et al. **Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (http://creativecommons.org/publicdomain/zero/1.0/) applies to the data made available in this article, unless otherwise stated.

Background

Amblyomma variegatum and *A. hebraeum* (Acari, Ixodidae) are two tick species of veterinary and public health concern in Africa (and in the Caribbean for *A. variegatum*) [1–3]. They are the cyclical vectors of *Ehrlichia ruminantium*, the causative bacterial agent of heartwater, a fatal disease of ruminants [4]. In addition, *A. variegatum* greatly facilitates, via immunosuppression in its cattle hosts, the development of dermatophilosis, a skin disease caused by *Dermatophilus congolensis* that imposes major economic impacts [5]. Moreover, human pathogens including several species of Rickettsiae and viruses are transmitted by these ticks [1, 3]. Understanding the factors limiting their geographical distributions is a prerequisite for predicting potential changes in tick distributions and epidemiological risk of associated diseases.

Whereas *A. variegatum* is widely distributed in Africa, its southern limit (Mozambique, Zimbabwe and Botswana) corresponds to the northern limit of *A. hebraeum*'s geographical range, which also extends into South Africa and across Swaziland [6]. In Zimbabwe, the two species were allopatric (i.e. with separated and non-abutting distributions) from the 1930s onwards until civil war stopped acaricide-based tick control in 1975 [7]. An extensive field survey conducted 30 years later revealed a parapatric distribution (i.e. abutting and marginally overlapping distributions) with rare co-occurrences in the same locations [8]. Exclusive competition between the two species has been hypothesised as a mechanism explaining the parapatry seen in Zimbabwe [7, 9]. This hypothesis predicts that similar parapatric range limits could also exist in Mozambique and Botswana. However, there is currently insufficient data at the range limits to know whether or not the parapatry seen in Zimbabwe is present in those countries too. In Mozambique, tick presence data arising from occasional sampling between 1940 and 1975 was only recorded at the district level giving rise to a geographical uncertainty of some hundreds of kilometres [10, 11 & Travassos Santos Dias J,(unpublished data)]. These data suggested the contact zone between these two species was located somewhere near the Save river (19th-23rd southern parallels): in the east of the country, *A. variegatum* was found south of the river together with *A. hebraeum* (Govuro and Vilankulo districts) whereas in the west it was the only species recorded to the north of the river and was never found south of the river (Additional file 1: Figure S1). Between 2000 and 2010, only a few tick presence data records arose from opportunistic sampling in the western-part of this area (Neves, unpublished data). These data indicated the presence of *A. hebraeum* to the north-west of the Save river at a distance of 150–200 km from the nearest *A. variegatum* records (Additional file 1: Figure S1).

Sharp distribution boundaries between species may be explained either by marked environmental gradients that traverse the limits of each species' tolerance to abiotic conditions, or by biotic interactions [12, 13]. These mechanisms give rise to "ecological parapatry" and "competitive parapatry" *sensu lato* respectively [13]. The latter can arise from inter-specific competition *sensu stricto* (e.g. [14–17]), differential effects of pathogens or predators on the two species (e.g. [18, 19]) or reproductive interference (e.g. [20]) (i.e. "any kind of inter-specific interaction during the process of mate acquisition that adversely affects the fitness of at least one of the species involved and that is caused by incomplete species recognition" [21]).

Numerous physiological observations suggest that inter-specific communicative and reproductive interference can occur when these species meet. Both species are three-host ticks, i.e. larvae, nymphs and adults must quest for and feed on different hosts while between stages molting occurs in the environment after blood-fed ticks detach from their hosts. Observations made within their respective geographical ranges have shown that the two species share the same host preferences: adults feed preferentially on large ruminants and have similar feeding-site preferences on hosts [22]. In regions of south-east Africa where the annual rainfall pattern is unimodal it is principally during the rainy season (September to April) that adults (of either species) are observed on hosts, larvae are observed at the end of the rainy season and nymphs during the cold season (May to October) [22]. Experimentally, both species survive under large and overlapping ranges of humidity and temperature [23, 24]. They display similar host and mate seeking behaviour: after several days of fixation on their host, adult males (of both species) attract unfed females as well as other host-seeking males by the emission of attraction-aggregation-attachment pheromones (AAAPs) [25]. These AAAPs are composed of several volatile compounds that differently act as (i) long-range attractants facilitating host location and selection, (ii) aggregation stimulants and (iii) attachment stimulants. Some of these compounds are quite similar between the two species and have been observed to induce partial and asymmetric inter-specific attraction, aggregation and attachment responses in males and females in laboratory experiments [26, 27]. Pheromones isolated from extracts of unmated females of the two species have also been found to share common compounds [28]. Although the exact role of these later pheromones is unknown, they might contribute to aspects of the mating process such as short-distance attraction of males or partner recognition. Finally, the coupling of *A. variegatum* females with *A. hebraeum* males (and vice versa) has been observed to result in hybrid inviability in laboratory experiments [9, 29].

These elements indicate many similarities in the trophic preferences, seasonality and communication systems (pheromones) of these two species. Thus, inter-specific communicative interferences might occur between males or/and between females or/and between males and females of the two species: pheromones produced by one species might inhibit or reduce long and short distance attraction and attachment responses of individuals of the other species to their own pheromones through inhibition effects or inefficient signal reception. Such effects might induce competition for hosts (i.e. competition for free attachment sites on host or for signal reception) and lead to spatial segregation of attachment positions between the two species in sympatric areas (i.e. attachment on different hosts or on different attachment sites within the same hosts). By contrast, these pheromones might have inter-specific attractivity effects and might induce strong sexual communication errors (i.e. sexual competition) resulting in inter-specific mating. Although cross-mating has been observed between *A. variegatum* and *A. hebraeum* under experimental conditions [9, 29], its existence and importance in the field is unknown. For example, whether or not sexual competition is avoided or mitigated by spatial segregation of males of the two species at sympatric sites is unknown.

To explore the hypothesis of competitive exclusion between *A. variegatum* and *A. hebraeum* we conducted a field survey at their range limits in Mozambique (1) to analyse their spatial distribution and (2) to assess the strength of inter-specific communicative and reproductive interference effects at co-infested sites.

Methods

Distributions of *A. variegatum* and *A. hebraeum* in the Mozambican contact zone

During 2012–2013, we conducted field surveys along north–south and east–west transects crossing the known range limits of both species (Additional file 1: Figure S1) in Inhambane, Manica, and Sofala provinces (19th-23rd southern parallels, Fig. 1) in two consecutive rainy seasons (peak adult *Amblyomma* activity period). In February 20112, we sampled cattle at sites spaced 30–50 km apart to identify the location and extent of the contact zone. In February 2013, we sampled cattle from sites spaced 5–10 km apart in the south of Manica and Sofala provinces, we refer to this area as the "area of

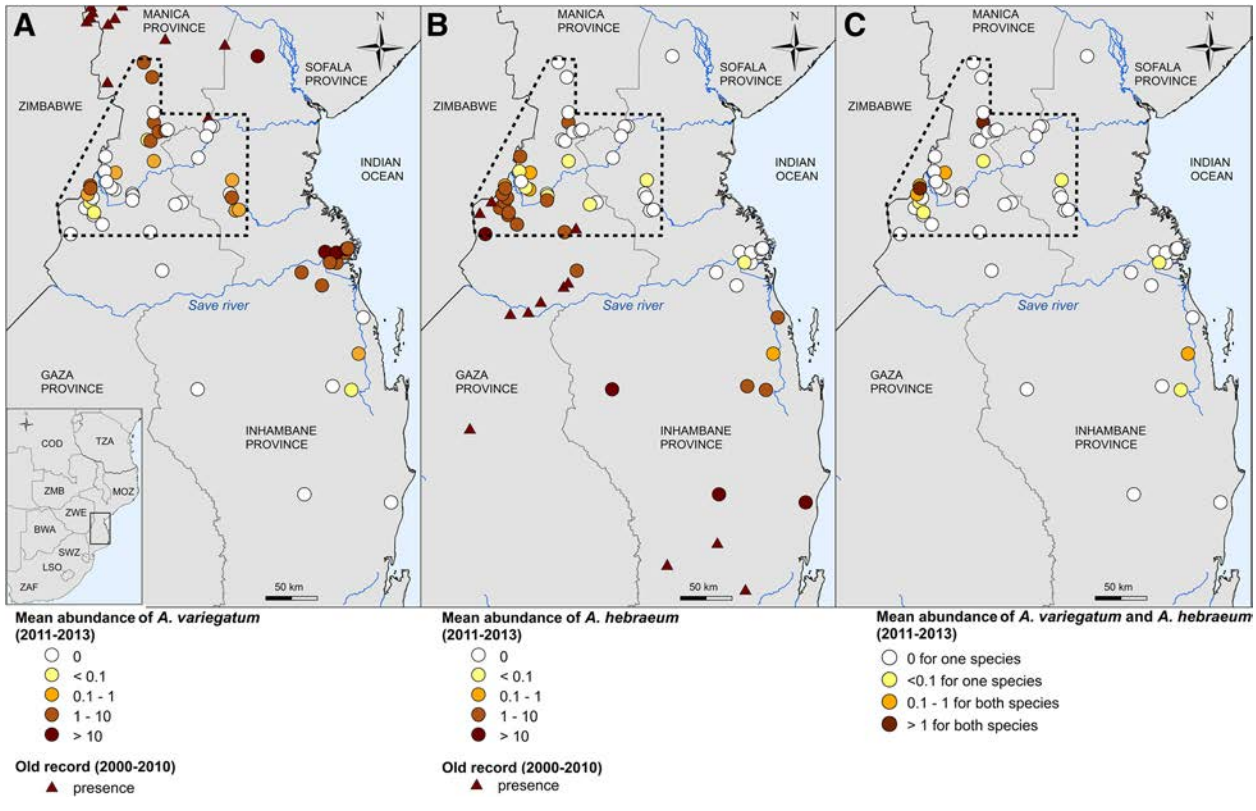


Fig. 1 Mean abundance of *Amblyomma* ticks in the Mozambican contact zone, **a** *A. variegatum*, **b** *A. hebraeum* and (**c**) co-occurrences. Dotted lines: limits of quasi-exhaustive sampling of sites with cattle. BWA = Botswana, COD = Democratic Republic of Congo, TZA = United Republic of Tanzania, LSO = Lesotho, MOZ = Mozambique, SWZ = Swaziland, ZMB = Zambia, ZWE = Zimbabwe, ZAF = South Africa

quasi-exhaustive sampling" since our survey visited a large majority of map pixels (3 × 3 km) with farms in that area. Study sites included communal dip-tanks and corridors used for acaricide treatment by farmers, plus some farms with no access to such facilities. When possible, a minimum of 50 animals within 3 km (approximate grazing range) of sampling locations, or (when less) all animals present on farms, were examined for tick presence and species identification. When feasible, 10 of these animals were laid down for detailed examination and the remaining cattle were examined in the corridor. During clinical examinations we counted the total number of adults of each *Amblyomma* species on each animal. Four herd infestation levels were defined according to mean abundance of ticks (i.e. number of *Amblyomma* adults/number of animals examined): <0.1, [0.1 – 1), [1–10) and ≥ 10 adults *Amblyomma* per animal[1]. We only included in the analysis animals that were treated with the acaricide product Amitraz eight days or more prior to the visitor those treated with pyrethroids 15 days or more prior the visit. These products were the only ones used on sampled farms. Eight days and fifteen days represent the mean duration of residual effects on hosts of Amitraz and pyrethroids respectively and the short-time needed for attraction and attachment of ticks on hosts. *Amblyomma* data were mapped to enable visualisation of the spatial distributions of the two species and their range limits.

In order to estimate the probability to detect ticks on cattle in a given area, cattle population data, encompassing census information, cattle movements and herd management practices including tick control, were collected through interviews with farmers and local veterinary service staff. At each location the probability to detect ticks in a herd, given the sample size, an assumed prevalence of 20 % and an assumed probability of tick detection on an infested host ("sensitivity") of 1 for inspections on laid-down animals and 0.6 for corridor inspections, was deemed "high" when it was estimated to be greater than 0.7 (Additional file 1: Method 1). These probabilities were also mapped.

Since the abundance of *A. variegatum* and *A. hebraeum* in a given area is conditionally related to the presence of large ruminants, we collected qualitative data on the presence of large wild ruminants and cattle. Cattle densities were mapped at a 10-km resolution to help identify potential tick infestation areas in the field and identify areas where low cattle host densities might amplify stochastic effects in the stability of *Amblyomma* populations. For this, cattle census data were available in the districts of Govuro, Chibababva, Sussudenga, Mossurize and Mabote (Additional file 1: Figure S2). For the other districts,we used the FAO's modelled data "Gridded Livestock of the World v2.01" [30]. All maps were produced using ArcMap v10 [ESRI, Redlands, California, USA].

Inter-specific communicative interference between *A. variegatum* and *A. hebraeum*

Attachment site preferences of A. variegatum and A. hebraeum with and without co-infestation

To assess potential differences in the attachment site preferences of the two species on co-infested animals, we measured overlap in adult attachment site preferences of each species using Schoener's D index [31]: $D_H = 1\text{-}1/2\ \Sigma_i\ |p_{Av,i} - p_{Ah,i}|$, where $p_{Av,i}$ and $p_{Ah,i}$ are the proportion of *A. variegatum* and *A. hebraeum* attached at site *i*on all co-infested animals. Attachment sites were divided into perinea-thigh region, inguinal region, axillary region, belly, head, legs, tail and dewlap.

To assess whether or not ticks modified their attachment site preferences in the presence of the other species, Schoener's D index was used to compare attachment site preferences of *A. variegatum* (or *A. hebraeum*) on animals infested by a single tick species with those observed on co-infested animals: $D_C = 1\text{-}1/2\ \Sigma_i\ |p_{O,i} - p_{T,i}|$, where $p_{O,i}$ and $p_{T,i}$ are the proportions of *A. variegatum* (or *A. hebraeum*) attached at site *i* for all animals infested by one (O) or two (T) *Amblyomma* species respectively. Both D_H and D_C were calculated separately for male and female ticks and for three cattle infestation levels ($n < 30$, $30 \le n < 70$, $n \ge 70$; with n the total number of ticks) using data from laid-down animals only.

Co-occurrence patterns at animal, attachment site and cluster levels

To assess the existence of between-species attraction or repulsion we analysed the spatial distribution (i.e. segregated vs. aggregated) of ticks in sympatric sites at the level of: host animal, attachment sites and cluster (a group of ticks aggregated within a 5 cm radius). These levels represent long (<4 m, i.e. average distance of attractive effects toward host-seeking ticks [32, 33]), medium (<30 cm) and short (<5 cm) range effects of pheromones respectively. The number of males and females per host, attachment site and cluster were counted on laid-down animals at sites where abundance of either species was superior to 0.1 adult ticks/host. We compared observed co-occurrences of individuals from the same or different species and/or sex to those generated by 5000 random permutations of data using the checkerboard score (C-score) [34] calculated via the function oecosimu from the R package vegan 2.2-2 [35].

Reproductive interference between *A. variegatum* and *A. hebraeum* in the field

Reproductive interference was assessed on naturally-infested cattle via the number of females mating with con- or hetero-specific males. Deviation from random mating was estimated at co-infested sites via the pair total index (PTI) and pair sexual isolation index (PSI)

[36] with the assumption of random mating of all individuals (PTI) or among individuals that actually mated (PSI). Sexual isolation was estimated via the joint isolation index (I_{PSI}) summarizing the difference in overall proportions of con- and hetero-specific pairs [36]. It ranges from −1 (fully hetero-specific mating) to 1 (complete isolation). We assessed asymmetry in hetero-specific mating by the index IA_{PSI}, the PSI ratio of hetero-specific combinations [37]. Standard errors (se) were estimated using 10,000 bootstrap replicates of the original data. Two-tail probabilities of obtaining estimates different from 0 for I_{PSI} and different from 1 for PSI, PTI and IA_{PSI} quantified deviations from random expectation (Jmating software [37] version 1.0.8).

Proportions of con-specific mating pairs were analysed using a beta-binomial logistic regression. The response variable was the frequency of con-specific couplings among all mating pairs. Fixed effects included (i) the mated-female species, (ii) the proportions of con-specific males and females centred on 0.5 (i.e., proportion - 0.5) and (iii) the interaction between these two proportions. Centring was used to facilitate the interpretation of model intercept and to decrease the correlation between fixed-effect coefficients. Over-dispersion is common in ecological data and not accounting for it can lead to spurious significance levels in statistical tests [38]. Here, over-dispersion with respect to the binomial distribution was modelled via a within-cattle correlation coefficient *r*. The statistical significance of these effects was quantified using likelihood ratio tests (LRT).

We examined the morphology of 30 ticks (or all when fewer were found) per species per study site under a stereomicroscope to search for phenotypical patterns that differed from those known for *A. variegatum* or *A. hebraeum* and may therefore represent hypothetical co-dominant or incomplete dominant hybridization. Morphological identification of *A. hebraeum* and *A. variegatum* was based on descriptions and observations documented in identifications keys [6, 39–42] of the following morphological criteria: colouration of the festoons, presence of median lateral areas of enamel, convexity of the eyes and nature of the colouration of scutal ornamentation.

Unless stated otherwise, all statistical analyses were performed with R version 3.1.2 [43].

Ethical approval

All the field work was implemented according to survey protocols approved by the Scientific Board of the Veterinary Faculty of the Eduardo Mondlane University, Maputo, Mozambique. The study permission was obtained from the Mozambican Livestock National Directorate, the Inhambane's, Sofala's and Manica's Livestock Provincial Directorate, from community leaders and from the farmers.

Results

Distributions of *A. variegatum* and *A. hebraeum* in the Mozambican contact zone

We sampled 59 sites in the Mozambican contact zone. The sampling effort was considered sufficient to enable high likelihood of species detection at 92 % (51/59) of sites (Additional file 1: Figure S3). We found *Amblyomma* ticks at 49 sites (83 %): 18 (31 %) with *A. variegatum* only, 19 (32 %) with *A. hebraeum* only and 12 (20 %) with both species (Fig. 1, Additional file 1: Table S1). Of these 12 sites: mean abundances of both species was superior to 1 tick/animal at just 2 sites (Fig. 1c, Additional file 1: Table S1), with *A. variegatum* predominant at one site and with approximately equivalent abundance of the two species at the other; mean abundance of both species was between 0.1 and 1 ticks/animal at 3 sites; one species was predominant with few individuals (<0.1 ticks/animal and <5 observed ticks in total) of the other species at 7 sites (Additional file 1: Table S1).

In this area, cattle densities were heterogeneous: absence or very low densities (<0.5 heads/km^2) of cattle was common over much of this region, otherwise cattle densities ranged between 0.5 and 25 heads/km^2 (Fig. 2). Outside protected areas (national parks, nature reserves) and hunting reserves, densities of wild animals, especially for large ruminants, were extremely low.

Areas of contact or transition between *A. variegatum* and *A. hebraeum* populations were all characterised by very low cattle densities (<0.5 heads/km^2) and mean tick infestations inferior to 10 ticks/animal. The observed zone within the quasi-exhaustive sampling area (dotted lines in Fig. 1) in which *A. variegatum* and *A. hebraeum* distributions overlapped was approximately 80–90 km wide.

Inter-specific communicative interference between *A. variegatum* and *A. hebraeum*

Attachment site preferences of A. variegatum and A. hebraeum with and without co-infestation

The great majority (>90 %) of males and females of both tick species was attached to the perineal, inguinal and axillae regions of host animals at all study sites. At sympatric sites, the preferred attachment sites of the two species were highly similar on co-infested animals (Schoener's $D_H = 0.86$ for hetero-specific males, $D_H = 0.84$ for hetero-specific females) (Additional file 1: Tables S2 and S3). However, attachment site preferences of *A. variegatum* on cattle hosting/not hosting *A. hebraeum* were less similar (Schoener's $D_C = 0.75$ for males, Dc = 0.74 for females) than attachment site preferences of *A. hebraeum* on cattle hosting/not hosting *A. variegatum* ($D_C = 0.91$ for males, Dc = 0.87 for females). Whereas *A. variegatum* attached most frequently in the inguinal region when host animals were not infested by *A. hebraeum*, on co-infested

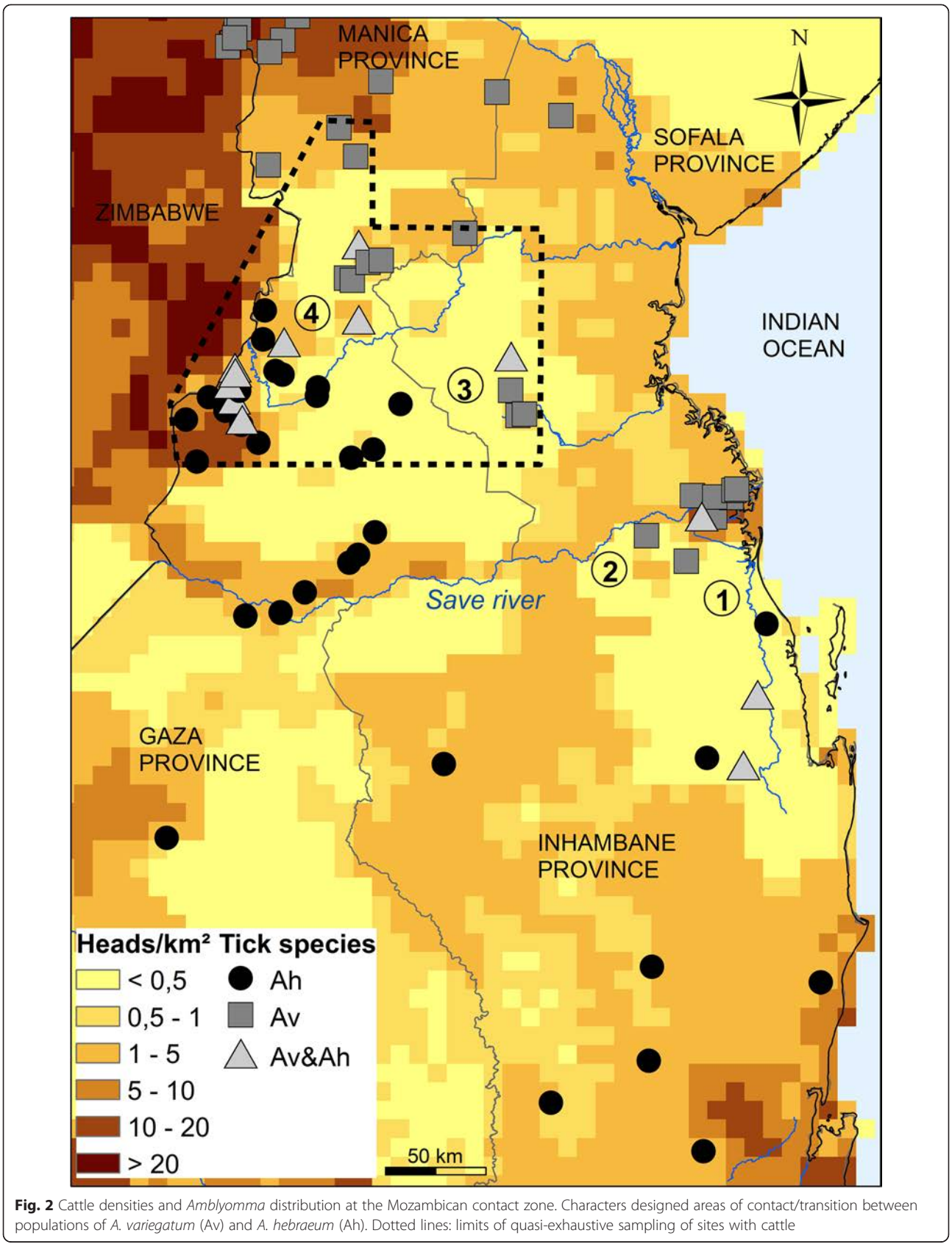


Fig. 2 Cattle densities and *Amblyomma* distribution at the Mozambican contact zone. Characters designed areas of contact/transition between populations of *A. variegatum* (Av) and *A. hebraeum* (Ah). Dotted lines: limits of quasi-exhaustive sampling of sites with cattle

animals *A. variegatum* attached preferentially in the axillae region where *A. hebraeum* was less abundant than in the inguinal region.

Co-occurrence patterns at animal, attachment site and cluster levels

At the 5 sites where both species were observed with an abundance superior to 0.1 ticks/animal (Additional file 1: Table S1), 59 % (36/61) of cattle were found to be co-infested by males of both species and 34 (21/61) and 7 % (4/61) respectively were infested only by males of either *A. variegatum* or *A. hebraeum*. The great majority (>80 %) of males and females of both species was attached on co-infested animals (Table 1).

On co-infested animals ($n = 32$ animals with tick cluster data), 55.5 % ($n = 72$) of attachment sites and 35.5 % ($n = 121$) of clusters were infested by males of both species and 33.3 and 41.3 % respectively were infested by males of *A. variegatum* but not by males of *A. hebraeum*. The majority of ticks (>50 %) of the two species were attached close to each other within the same attachment site or the same cluster (Table 1). However, males and females of *A. variegatum* attached less frequently in common attachment sites and clusters than those of *A. hebraeum* (Table 1). Animal tick infestation was low: the median [25th percentile; 75th percentile] number of males and females per co-infested animal was 4 [2.25; 9] and 2 [1; 3] for *A. variegatum* and 2 [1.25; 4] and 1 [0; 3] for *A. hebraeum* respectively.

Whereas males of the two species were independently distributed at the host and attachment site levels (C-score = 0.05, $p = 0.57$ and C-score = 0.07, $p = 0.67$), they were more segregated than expected by chance at the cluster level (C-score = 0.16, $p = 0.005$, Fig. 3) indicating that only short range aggregation was affected by potential competitive effects between hetero-specific males. *Amblyomma variegatum* males formed hetero-specific clusters less frequently than expected by chance, while the situation was reversed for *A. hebraeum* males ($\chi^2 = 13.8$, $p < 10^{-3}$). The presence of *A. variegatum* males significantly increased the probability of the presence of *A. hebraeum* females at the animal (C-score = 0, $p < 0.001$), attachment site (C-score = 0.03, $p < 0.001$) and cluster levels (C-score = 0.07, $p = 0.017$); however, the reciprocal relation was not true (C-score = 0.07, $p = 0.22$; C-score = 0.009, $p = 0.34$; C-score = 0.13, $p = 0.88$ respectively). As males of both species were independently distributed at host and attachment site levels, these effects cannot only be attributed to the presence of *A. hebraeum* males and suggest that *A. variegatum* males may also induce long-distance attraction effects to *A. hebraeum* females. Given the negative association between the distributions of males of both species at the cluster level, females of both species appeared to be locally attracted by groups of hetero-specific males, with a stronger effect observed among *A. hebraeum* females. Females of the two species were aggregated at the host level (C-score = 0.05, $p = 0.036$) but distributed independently at the attachment site (C-score = 0.08, $p = 0.67$) and cluster levels (C-score = 0.06, $p = 0.12$).

Reproductive interference between *A. variegatum* and *A. hebraeum* in the field

Cross-mating was observed at the two sites where abundance of both species was superior to 1 tick/animal (sites # 12 and 55 in Additional file 1: Table S1). At one of these sites *A. variegatum* was more abundant among observed males, whereas at the other site males of the two species were approximately equi-abundant. At the latter site, 15.5 (9/58) of *A. variegatum* females and 12.5 % (4/32) of *A. hebraeum* females mated with hetero-specific males, whereas 69 % (40/58) of *A. variegatum* females and 62.5 % (20/32) of *A. hebraeum* females mated with con-specific males and 15.5 % (9/58) of *A. variegatum* females and 25 % (8/32) of *A. hebraeum* females were attached single. *Amblyomma variegatum* and *A. hebraeum* showed substantial but incomplete sexual isolation ($I_{PSI} = 0.65$, se = 0.09, $p < 10^{-3}$; PSI and PTI values of hetero-specific pairs between *A. variegatum* males and *A. hebraeum* females were 0.27 and 0.28 respectively, and PSI and PTI of hetero-specific pairs between *A. hebraeum* males and *A. variegatum* females were 0.46 and 0.42 respectively, all PSI and PTI values were significantly different from 1, $p < 0.05$). We found no evidence of asymmetry in hetero-specific mating ($IA_{PSI} = 1.99$, se = 1.16, $p = 0.2$). At the former site (where *A. variegatum* was

Table 1 Within- and between-host co-occurrence of *Amblyomma variegatum* and *A. hebraeum*. Percentage of males and females of *A. variegatum* (Av) and *A. hebraeum* (Ah) attached on/within the same animals, attachment sites or clusters as *A. hebraeum* and *A. variegatum* males

	Av males		Ah males		Av females		Ah females	
	n	% attached in presence of Ah males	*n*	% attached in presence of Av males	*n*	% attached in presence of Ah males	*n*	% attached in presence of Av males
Animals	275	80.0	133	96.2	100	81.0	43	100.0
Attachment sites	203	80.8	119	96.6	75	74.7	39	84.6
Clusters	203	52.0	119	74.7	75	49.0	39	71.0

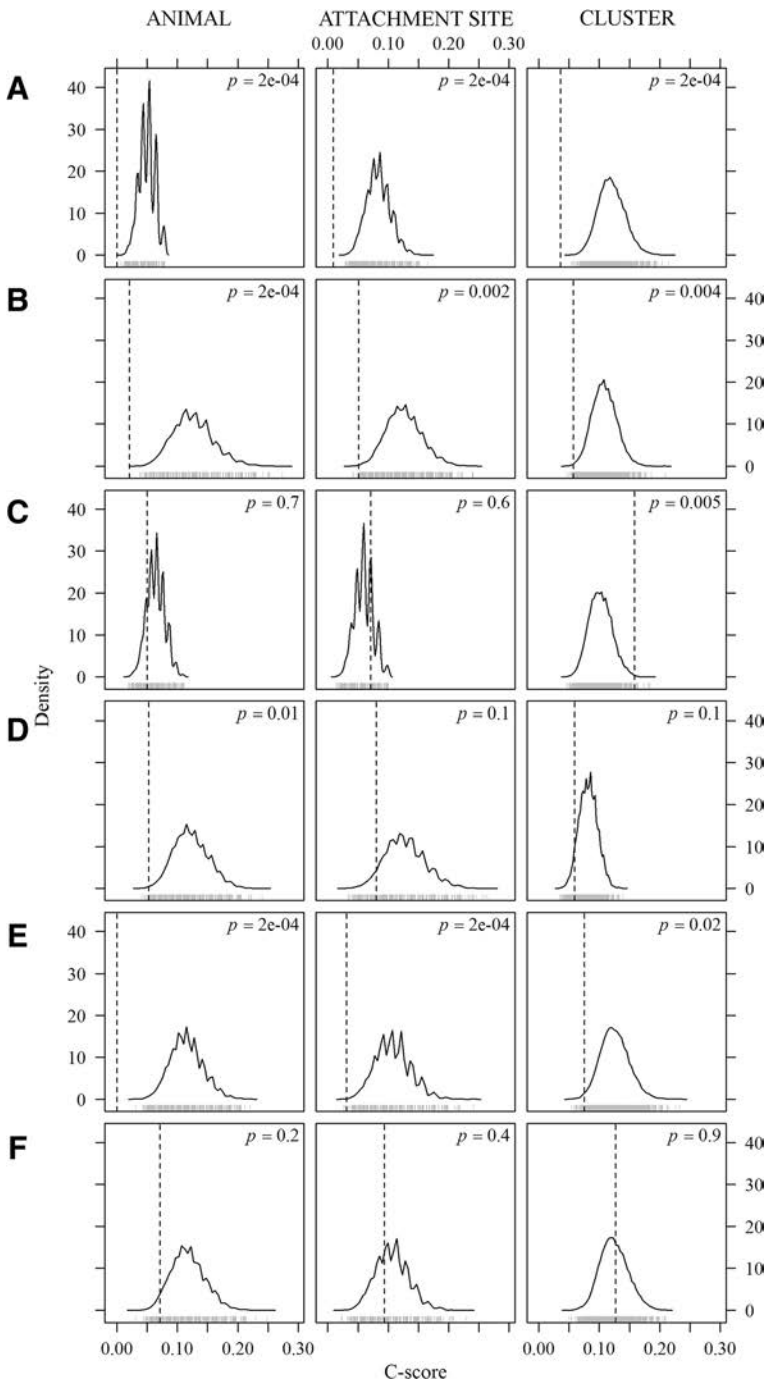


Fig. 3 C-scores of tick pairs according to species and sex at host, attachment site and cluster level (dotted line) compared to frequency distributions (solid line) generated from 10,000 Monte Carlo simulations using equivalent frequencies for each group and independent distributions (null model). Tested pairs are: **a** *A. variegatum* males vs females, **b** *A. hebraeum* males vs females, **c** *A. variegatum* males vs *A. hebraeum* males, **d** *A. variegatum* females vs *A. hebraeum* females, **e** *A. variegatum* males vs *A. hebraeum* females, **f** *A. hebraeum* males vs *A. variegatum* females

dominant), 15 (4/26) of *A. variegatum* females and 100 % (4/4) of *A. hebraeum* females mated with hetero-specific males; conversely, 77 % (20/26) of *A. variegatum* females mated with con-specific males. All the ticks examined under a stereo-microscope were identified as being either *A. variegatum* or *A. hebraeum* and no intermediate forms observed.

The beta-binomial logistic regression model of the frequency of con-specific couplings among mating pairs presented a high within-cattle correlation coefficient

($r = 0.37$, $Pr(> |r|) = 0.017$). The effect of female tick species was small and not significant (Additional file 1: Table S4), in agreement with the IA_{PSI} index. The effects of both proportions of con-specific males and females were significant (Additional file 1: Table S4) and the interaction between the proportions of con-specific males and females was strong and negative (LRT, $\chi^2 = 10.4$, $df = 1$, $p = 0.001$). Model predictions provided evidence for a preference of females for males of the same species: 67 % of the surface of Fig. 4a (magenta colour) corresponds to a majority of mating pairs being con-specific. The generally low coefficient of variation indicates that the accuracy of these predictions can be expected to be good, except when the proportions of con-specific males and females were both low (Fig. 4b). Large predicted proportions of con-specific mating pairs were obtained when the proportion of con-specific males was >30 %, except when high proportions of con-specific females (>70 %) were encountered (Fig. 4a). Even in this case, proportions of con-specific males > 50 % resulted in high predicted proportions of con-specific mating pairs. The effect of an increase in the proportion of con-specific females on the predicted proportion of con-specific pairs was different when the proportion of con-specific males was inferior (vs. superior) to 25 %: below this threshold, predicted proportions of con-specific pairs increased slightly when the proportion of con-specific females increased; inversely, above this threshold predicted proportions of con-specific pairs decreased faster as the proportion of con-specific females increased.

Discussion

Our observations confirm the existence of a parapatric boundary between the geographic distributions of *A. variegatum* and *A. hebraeum* populations in Mozambique. The adopted sampling design was based on the well documented phenology of these species – when these species are present in a given area, adults are mainly observed on cattle throughout the rainy season and it would be extremely rare not to observe *Amblyomma* during this period unless abundances were very low. Even when observations were made as late as February, a period that usually corresponds to the decreasing phase of adult infestation curves, adult infestation levels are still expected to be sufficiently high to have a large probability to observe *Amblyomma* ticks when present at moderate to large numbers. Indeed, we did observe both species at several sites during this period. To facilitate interpretation of observed absences we calculated the conditional probability to detect tick presence given our sampling at each site, an assumed 20 % of tick prevalence in the herd (which is low for these species) and an assumed sensitivity of tick detection on an infested host of 1 for inspections on laid-down animal and of 0.6 for corridor inspections. Based on our collective experiences on the field these figures are the minimum expected for scenarios of low levels of tick infestation such as 1–5 ticks per animal (Additional file 1: Method 1). These conditional probabilities were greater than 0.7 at the majority of sampled sites (median 0.98, range 0.41- 0.99). As such, the observed absence of one species at any given site most likely reflects either a true absence or a very low abundance.

In this area, we found coexistence of both species at relatively few sites. At most of the co-infested sites, densities of one or both species were low (<1 tick/animal) and often less than 5 individuals of the sub-dominant species were collected in total. Outside the quasi-exhaustive sampling area, e.g. in Inhambane province, we were able to visit only a limited number of farms due to organisational constraints. This impedes accurate estimation of the number of co-infested sites in an area 200 km wide. In the quasi-exhaustive sampling area where we visited a large majority of map pixels with cattle farms, co-infested sites

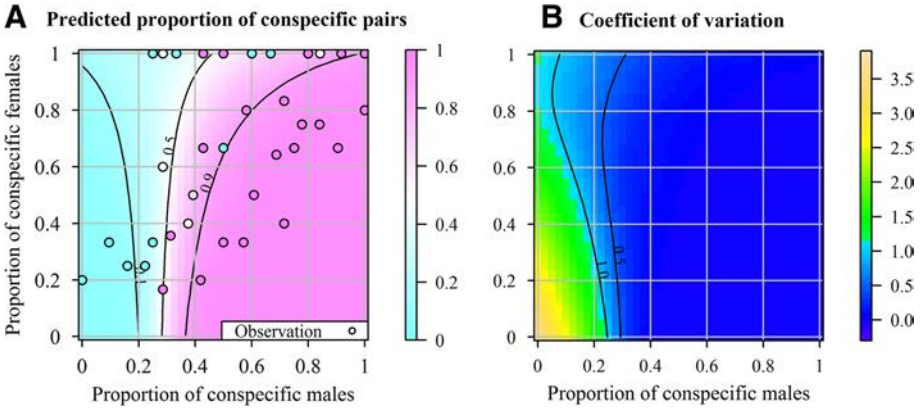


Fig. 4 Predicted (colour) and observed (circle) distribution of con-specific mating ticks according to the proportion of con-specific males and females. **a** mean proportion and **b** coefficient of variation predicted by a beta-binomial logistic regression model. Contour lines indicate the predicted probabilities to observe proportions of con-specific mating ticks of 0.1, 0.5, and 0.9

were located in a zone of 80–100 km in width. Historical tick records from 2009 and 2010 located in Manica province, northern and southern from this area, indicated the presence of a single *Amblyomma* species as well and hence may be perceived as additional hints to indicate that sympatric areas were limited to the study area. In most sites that were not sampled in this area, *Amblyomma* abundance can be expected to be very low or zero due to the absence or a very low abundance of large ruminants (cattle or large wild ruminants). Indeed, although under such conditions medium-size animals might serve as alternative hosts for adults of *A. variegatum* and *A. hebraeum* as evidenced experimentally, there is no evidence that under natural conditions these species can persist in areas where there are no large ruminants. The number of adult ticks that feed successfully on small animals is much lower than on large ruminants [44–53] and this appears to impede the tick's reproductive performance [54, 55]. Therefore, although it is possible that some co-infested sites might have escaped detection, especially where abundances of one species were low, our results do indicate that these species rarely co-exist at high or moderately high abundance levels. A similar pattern was observed in the contact zone in Zimbabwe during the last survey conducted in 1996 [8]. This pattern is consistent with a dynamical system in which sympatry, when it arises, is highly transient. More generally, *Amblyomma* densities observed in this area were low (<10 ticks/animal), compared to densities observed in the rest of our study area (>20 ticks/animal) or to previous reports of *Amblyomma* densities from other areas [56–61]. Our data do not permit accurate tick density estimates at each site, however, the between-site differences detected here during the peak adult tick activity season are sufficiently large to be highly informative regarding the location of the parapatric boundary in Mozambique within an area of low tick abundance. Further, heterogeneity in cattle distribution combined with very low densities of large wild ruminants within the contact zone probably renders tick densities highly heterogeneous across this area. As discussed previously, it can be expected that in areas with no or few large ruminants, *Amblyomma* abundance is either very low or zero.

One hypothesis to explain the parapatric distribution between *A. variegatum* and *A. hebraeum* is that some form of inter-specific competition *sensu lato* impedes the spread of either species across the contact zone. Our field observations suggest competition for sexual partners, together with the absence of strong repellent effects between species, could provide a mechanism to generate this parapatry. We observed that adults of both species preferred the same three main attachment sites on hosts (axillae, inguinal and perineal region), whether or not the other species was present. Moreover, they frequently attached to the same hosts and in the same clusters. High cross-mating rates (12.5 and 15.5 % respectively for the females of *A. hebraeum* and *A. variegatum*) were observed in the field when the two species were equally abundant. Thus, females often fail to correctly identify their sexual partners and reproductive isolation at pre-mating barriers is only partial. This phenomenon can be attributed to an inability to fully distinguish between species via the pheromones emitted for long or short distance attraction, attachment and mating.

Cross-mating rates did not differ significantly with the species of the female, thus asymmetry in reproductive interference effects was undetectably small. By contrast, asymmetry was detected in the mating preferences of females: we found the attraction-attachment effects of hetero-specific males to be stronger for *A. hebraeum* females than for *A. variegatum* females. This is consistent with experimental results of Norval et al. [26, 27] who used extracts of tick pheromones. Therefore, other mechanisms of recognition are probably at play downstream in the reproductive cycle that reduce this asymmetry. Rechav et al. [9] made similar observations: although a large number of *A. hebraeum* females attached close to *A. variegatum* males, only a small proportion actually formed hetero-specific pairs. The opposite was observed with *A. variegatum* females, which were less attracted by males of *A. hebraeum* but of which a significant proportion formed hetero-specific mating pairs.

Such "mating errors" can have a huge impact on reproductive success, the most extreme scenario being when the attraction of females towards hetero-specific males results in a complete failure to mate with conspecific males – such phenomenon greatly increase the local extinction probability of the least frequent species. We observed no major differences between the two species in the percentage of single females attached to a host when male abundances were equal, suggesting symmetry or very small asymmetry of competitive effects at that level, although in general symmetry or small asymmetry in reproductive interference effects is rare among taxa [21]. In addition, communicative interference between males could also influence the reproductive fitness of the two species. Indeed, within clusters, males of the two species were more segregated than expected under randomisation, suggesting a local and partial repellent effect between them. The repellent effect of *A. hebraeum* males towards *A. variegatum* males was more pronounced, as observed by Norval et al. [26, 27]. This may also explain the slight difference observed in *A. variegatum*'s attachment site preferences in the presence or absence of *A. hebraeum*. However, further experiments are needed to more accurately estimate the size and degree of asymmetry in deleterious effects (on each species reproductive success)

arising from communicative interference between males, mismating and hybridization.

Results of previous hetero-specific-cross experiments suggest that cross-mating between *A. variegatum* and *A. hebraeum* is most likely to generate inviable eggs [9, 29]. This implies that, if hybrids are generated, then this occurs at undetectably low frequencies. We did not observe ticks with morphologically intermediate forms that may represent hypothetical co-dominant or incomplete dominant hybridization at sympatric sites. When cross mating results in inviable offspring, a frequency-dependent mechanism, called a Satyr effect or satyrisation [62], is produced that can lead to exclusive competition and parapatry between species regardless of whether or not the exclusion generated by satyrisation is enhanced by exploitative competition (i.e. competition for resources or apparent competition through shared predators or pathogens) [62–65]. Satyrisation is predicted to decrease reproductive fitness as the relative proportion of hetero-specific individuals increases. In the Mozambican contact zone, the frequencies of hetero-specific males and (to a lesser extent) hetero-specific females both appear to influence reproductive fitness. The beta-binomial regression model of con-specific pairs at the animal level predicted that the proportion of con-specific mating pairs increases with the proportion of con-specific males, the degree of increase depending on the proportion of con-specific females. Low con-specific mating rates were predicted when the proportion of con-specific males was inferior to 0.3. Above this threshold, the proportion of con-specific pairs was generally high (>0.5) but decreased slightly as the proportion of con-specific females increased. This is coherent with our observation of high hetero-specific mating rates of *A. variegatum* females (15 %) at the sympatric site where this species was dominant. Such influences of the relative frequencies of con-specific and hetero-specific females on reproductive fitness were unexpected for this species given that males can remain on the host for several months and mate with several females and are generally more numerous than females. This may reflect that con-specific males were a limited resource for females at the animal level, which might be related to (i) the low density of ticks of any species in the study area and consequently, the low number of male ticks on infested cattle (most animals were infested by <5 male ticks), (ii) frequent acaricide treatments of cattle (usually every 15 days) reducing the number of male ticks that have been attached for long enough (at least 5–7 days) to emit AAAP pheromones and attract females; (iii) a heterogeneous proportion of the two species on animals or in the environment. Thus, sexual competition appears to decrease the population growth rate of both dominant and sub-dominant species at sympatric sites increasing the sensitivity of low populations to stochastic events. However, such effects are generally not considered in mathematical studies [62–64].

The positive frequency-dependent effect and near-symmetry of reproductive interference can be expected to induce local extinction to the least abundant species at sympatric sites [62–64]. Spatially explicit theoretical sexual competition models [62, 65, 66] predict that the dominant species at a given location at some time t is determined by (i) "initial conditions" i.e. the relative frequency of each species at some previous time t_0 at each location within the landscape, (ii) the fitness of each species given the abiotic conditions and any modifications to fitness caused by inter-specific competitive effects such as the satyrisation effect, and (iii) dispersion rates, which can be spatially heterogeneous. These models predict that the invasion of a fitter species into an area where another species is established will be unsuccessful unless the number of invading individuals exceeds a certain threshold. Otherwise, frequency dependent competition effects prevent a successful invasion and the newly introduced species typically disappears in just a few generations in the absence of continuous immigration. The influences of "initial" abundances of a resident species on the final outcome of competition with an invading species are known as "priority effects". These models also predict that when populations are initially allopatric and dispersal rates are high parapatric boundaries tend to form close to isoclines of environmental gradients that delimit equivalence in density independent fitness. But, lower dispersal rates can strengthen priority effects, accentuate the impacts of stochastic events and generate greater variation in the exact locations of parapatric boundaries.

Given the high bidirectional hetero-specific mating rates observed between *A. variegatum* and *A. hebraeum* it can be expected that strong priority effects could be generated by reproductive interference and that this process could strongly influence the limits of their geographical distributions, particularly when the spread of one species into the range of another is limited by low dispersal rates. This might occur at the contact zone in Mozambique. Tick dispersal is dependent upon host movement and the flow of cattle and wild ruminant host animals across this contact zone is thought to be low (national veterinary services, pers. comm.). Low tick and host densities as well as the patchy distribution of hosts observed at the contact zone may result in relatively infrequent tick dispersal events across the area. Some well-established populations of *A. variegatum* and *A. hebraeum* were found to be separated by "no-cattle lands" as wide as 40–60 km – such areas can be expected to function as dispersal barriers to ticks. However, even if tick dispersal events are infrequent, they are

not completely negligible as suggested by the presence of a small number of individuals of the non-dominant species at seven out of the 59 investigated sites and reported movements of cattle and small ruminants across the area (national veterinary services, pers. comm.). Thus, it is unlikely that infrequent dispersion alone provides a sufficient mechanism to prevent range expansion of both species. However, priority effects generated by sexual competition do appear to explain why the current location of the Mozambican parapatric boundary falls within a zone of apparently low tick densities and infrequent dispersal. Low densities, stochastic events and priority effects suggest that sites of sympatry (and the extent of the overlapping area) are likely to shift stochastically at relatively high frequency, although the dominant species at each site would change relatively infrequently. Existence of competitive exclusion between *A. variegatum* and *A. hebraeum* and how it interacts with abiotic factors (i.e. the relative importance of priority effects) can be explored further by studying habitat suitability for both species, studying tick dispersal across the area and by performing periodic follow up surveys at sites in and around the contact zone over a number of years. Historical tick records are insufficient to be informative about this hypothesis since these data are insufficient in quantity and quality to determine whether or not the location of parapatric boundaries have shifted with time or even which species were historically present in the area. However, it is worth noting that despite massive changes in host populations in the area between 1950 to 2014 (e.g. massive reductions of large ruminant populations (both cattle and wild large ruminants) during the civil war (1975–1992), followed by progressive reintroduction of cattle in traditional farming areas, especially from Zimbabwe), the parapatric boundaries as observed now appear to be located within the same districts as suggested by records from the 1950's. The extent of which the location of the parapatric is determined by environmental factors remains unknown (but see [67] for Zimbabwe's parapatric boundaries) and warrants further study. However, this uncertainty does not negate our main conclusion, namely that satyrisation is a key mechanism that can inhibit coexistence between these two species of tick and can inhibit the invasion of one species into an area where the other species is already abundant even when the resident species is less fit than the invader given the local environmental conditions.

When two closely related species are in regular contact, pre-mating barriers can evolve through reinforcement in response to disadvantages arising from mismating or hybridization [68, 69]. Reinforcement here is considered in its broad sense, i.e. "the evolution of mechanisms that prevent interbreeding between newly interacting incipient species, as a result of selection against inter-specific matings" [68]. Reinforcement is strongest when species co-evolve in sympatry where it typically results in pre-zygotic isolation evolving more rapidly than post-zygotic isolation [70, 71]. Here, post-mating reproductive isolation appears to be complete (inviability of hybrid eggs) yet pre-mating reproductive isolation is only partial suggesting that reinforcement between *Amblyomma* species in Mozambique has been weak. This pattern of reproductive isolation reflects a history of infrequent between-species contact [70, 71]. However, the degree of pre-mating reproductive isolation can vary spatially in response to differential competitive and sexual selection pressures [72, 73]. In Zimbabwe, tick dispersal events and between-species contacts are expected to be more frequent than in Mozambique due to higher cattle densities and more frequent host movements. It would therefore be interesting to test if female choosiness in mate choice is evolving more rapidly in Zimbabwe than in Mozambique, especially since augmentations in female choosiness could profoundly affect the future distributions of these ticks.

Conclusion

We report the first field observations of hetero-specific mating between *A. variegatum* and *A. hebraeum* from sympatric areas in Mozambique. These results, coupled with previous experimental evidence of egg inviability resulting from hybridisation, suggest that sexual competition between these species provides a key mechanism that can, at least partially, explain the spatial segregation observed in Mozambique and Zimbabwe. Sympatry within the contact zone is relatively rare, which corresponds perfectly with the transient dynamics of sympatry predicted by theoretical sexual competition models. The extent to which environmental factors determine the location of the parapatric boundary is currently unknown. Further field, laboratory and modelling work is required to quantify the extent to which sexual competition displaces the parapatric boundaries away from any potential environmentally determined lines of equi-fitness between these two species.

Endnotes

[1]In set notation, $[a,b) = \{x \in \mathbb{R};\ a \leq x < b\}$.

Additional file

Additional file 1: Supplementary information. (DOCX 287 kb)

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

LB, TL, JP, DM, LN and RL designed the study. LB and NC conducted the fieldwork and collected the data. LB, JB and RL performed the analyses. All the authors discussed the results. LB, JP, JB, DP and RL wrote the first draft of the manuscript and all authors contributed substantially to subsequent revisions. All authors read and approved the final manuscript.

Acknowledgements
We are very grateful to H. Neves Mucache, N. Nunes de Carvalho Vaz, E. Specht, the technicians and veterinarians of the national veterinary services in Mozambique, as well as to the farmers for their help and support in collecting ticks in the field. The authors acknowledge the financial support received from European project, FEDER 2007–2013, FED 1/1.4-30305, "*Gestion des Risquesen santé animale et végétale*", from the INRA "molecular biology – epidemiology" project and from the Centre of Biotechnology of the University of Maputo, Eduardo Mondlane. Laure Bournez acknowledges financial support for her PhD from the European project, FED 1/1.4-30305 and CIRAD.

Author details
[1]CIRAD, UMR CMAEE, F-97170 Petit-Bourg, Guadeloupe, France. [2]INRA, UMR 1309 CMAEE, F-34398 Montpellier, France. [3]Université des Antilles et de la Guyane, F-97159 Pointe-à-Pitre, Guadeloupe, France. [4]Centro de Biotecnologia- Eduardo Mondlane University, Av. de Moçambique, km 1,5, C.P. 257, Maputo, Mozambique. [5]CIRAD, UMR CMAEE, F-34398 Montpellier, France. [6]Institut Sénégalais de Recherches Agricoles, Laboratoire National d'Elevage et de Recherches Vétérinaires, BP 2057 Dakar – Hann, Senegal. [7]CIRAD, F-97130, Capesterre-Belle-Eau, Guadeloupe, France. [8]Department of Veterinary Tropical Diseases, Faculty of Veterinary Science, University of Pretoria, Private Bag x04, Onderstepoort 0110, South Africa.

Received: 28 April 2015 Accepted: 28 September 2015
Published online: 06 October 2015

References

1. Hoogstraal H. Viruses and ticks. Viruses Invertebr. 1973;31:349–90.
2. Jongejan F, Uilenberg G. The global importance of ticks. Parasitology. 2004;129:S3–S14.
3. Parola P, Paddock CD, Raoult D. Tick-borne rickettsioses around the world: emerging diseases challenging old concepts. ClinMicrobiol Rev. 2005;18:719–56.
4. Allsopp BA. Natural history of Ehrlichiaruminantium. Vet Parasitol. 2010;167:123–35.
5. Martinez D, Barré N, Mari B, Vidalenc T. Studies of the role of Amblyommavariegatum in the transmission of Dermatophiluscongolensis. In: Fivaz B, Petney T, Horak I, editors. Tick Vector Biology. Berlin, Germany: Springer Berlin Heidelberg; 1992. p. 87–99.
6. Walker AR, Bouattour A, Camicas JL, Estrada-Pena A, Horak IG, Latif AA, et al. Ticks of Domestic Animals in Africa: a Guide to Identification of Species, Bioscience reports Edinburgh. 2003.
7. Norval R, Perry B, Meltzer M, Kruska R, Booth T. Factors affecting the distributions of the ticks Amblyommahebraeum and A. variegatum in Zimbabwe: implications of reduced acaricide usage. ExpApplAcarol. 1994;18:383–407.
8. Peter TF, Perry BD, O'Callaghan CJ, Medley GF, Shumba W, Madzima W, et al. Distributions of the vectors of heartwater, Amblyommahebraeum and Amblyommavariegatum (Acari: Ixodidae), in Zimbabwe. ExpApplAcarol. 1998;22:725–40.
9. Rechav Y, Norval R, Oliver J. Interspecific mating of Amblyommahebraeum and Amblyommavariegatum (Acari: Ixodidae). J Med Entomol. 1982;19:139–42.
10. Travassos Santos Dias J. Lista dascarraças de Moçambique e respectivoshospedeiros. III. An Dos ServiçosVeterinária E IndústriaAnim. 1953;1954:213–87.
11. Travassos Santos Dias J. Some data concerning the ticks (Acarina-Ixodoidea) presently known in Mozambique. Garcia OrtaSérieZool. 1991;18:27–48.
12. Bridle JR, Vines TH. Limits to evolution at range margins: when and why does adaptation fail? Trends EcolEvol. 2007;22:140–7.
13. Bull C. Ecology of parapatric distributions. Annu Rev EcolSyst. 1991;22:19–36.
14. Werner P, Lötters S, Schmidt BR, Engler JO, Rödder D. The role of climate for the range limits of parapatric European land salamanders. Ecography. 2013;36:1127–37.
15. Sato Y, Sabelis MW, Mochizuki A. Asymmetry in male lethal fight between parapatric forms of a social spider mite. ExpApplAcarol. 2013;60:451–61.
16. Wisz MS, Pottier J, Kissling WD, Pellissier L, Lenoir J, Damgaard CF, et al. The role of biotic interactions in shaping distributions and realised assemblages of species: implications for species distribution modelling. Biol Rev. 2013;88:15–30.
17. Reitz SR, Trumble JT. Competitive displacement among insects and arachnids. Annu Rev Entomol. 2002;47:435–65.
18. Tompkins DM, Draycott RAH, Hudson PJ. Field evidence for apparent competition mediated via the shared parasites of two gamebird species. Ecol Lett. 2000;3:10–4.
19. Tompkins DM, White AR, Boots M. Ecological replacement of native red squirrels by invasive greys driven by disease. Ecol Lett. 2003;6:189–96.
20. Thum RA. Reproductive interference, priority effects and the maintenance of parapatry in Skistodiaptomus copepods. Oikos. 2007;116:759–68.
21. Gröning J, Hochkirch A. Reproductive interference between animal species. Q Rev Biol. 2008;83:257–82.
22. Petney TN, Horak IG, Rechav Y. The ecology of the African vectors of heartwater, with particular reference to Amblyommahebraeum and Amblyommavariegatum. Onderstepoort J Vet Res. 1987;54:381–95.
23. Yonow T. The life-cycle of Amblyommavariegatum (Acari: Ixodidae): a literature synthesis with a view to modelling. Int J Parasitol. 1995;25:1023–60.
24. Norval R. Studies on the ecology of the tick Amblyommahebraeum Koch in the Eastern Cape Province of South Africa. II. Survival and development. J Parasitol. 1977;63:740–7.
25. Sonenshine DE. Tick pheromones and their use in tick control. Annu Rev Entomol. 2006;51:557–80.
26. Norval R, Peter T, Yunker C, Sonenshine D, Burridge M. Responses of the ticks Amblyommahebraeum and A. variegatum to known or potential components of the aggregation-attachment pheromone. II. Attachment stimulation. ExpApplAcarol. 1991;13:19–26.
27. Norval R, Peter T, Sonenshine D, Burridge M. Responses of the ticks Amblyommahebraeum and A. variegatum to known or potential components of the aggregation-attachment pheromone. III. Aggregation. ExpApplAcarol. 1992;16:237–45.
28. Price Jr TL, Sonenshine DE, Norval RAI, Yunker CE, Burridge MJ. Pheromonal composition of two species of African Amblyomma ticks: similarities, differences and possible species specific components. ExpApplAcarol. 1994;18:37–50.
29. Clarke FC, Pretorius E. A comparison of geometric morphometic analyses and cross-breeding as methods to determine relatedness in three Amblyomma species (Acari: Ixodidae). Int J Acarol. 2005;31:393–405.
30. Robinson TP, Wint GW, Conchedda G, Van Boeckel TP, Ercoli V, Palamara E, et al. Mapping the global distribution of livestock. PLoS One. 2014;9:e96084.
31. Schoener TW. Nonsynchronous spatial overlap of lizards in patchy habitats. Ecology. 1970;51:408–18.
32. Barré N. Biologie et Ecologie de la Tique Amblyommavariegatum (Acarina: Ixodina) en Guadeloupe (Antilles Française), Paris-Sud University. 1989.
33. Norval R, Butler J, Yunker C. Use of carbon dioxide and natural or synthetic aggregation-attachment pheromone of the bont tick, Amblyommahebraeum, to attract and trap unfed adults in the field. ExpApplAcarol. 1989;7:171–80.
34. Stone L, Roberts A. The checkerboard score and species distributions. Oecologia. 1990;85:74–9.
35. Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB, et al. Vegan: Community Ecology Package. R Package Version 2.2-2/r2928. 2013. http://R-Forge.R-project.org/projects/vegan/.
36. Rolan-Alvarez E, Caballero A. Estimating sexual selection and sexual isolation effects from mating frequencies. Evolution. 2000;54:30–6.
37. Carvajal-Rodriguez A, Rolan-Alvarez E. JMATING: a software for the analysis of sexual selection and sexual isolation effects from mating frequency data. BMC EvolBiol. 2006;6:40.
38. Donner A. The comparison of proportions in the presence of litter effects. Prev Vet Med. 1993;18:17–26.
39. Hoogstraal H. African Ixodidea. U.S, Navy, Washington, D.C: Volume Vol. I. Ticks of the Sudan; 1956.
40. Robinson L. E: Ticks. A Monograph of the Ixodoidea. Volume Part IV. The genus Amblyomma. Great Britain: Cambridge University Press; 1926.
41. Voltzit O, Keirans J. A review of African Amblyomma species (Acari, Ixodida, Ixodidae). Acarina. 2003;11:135–214.
42. Walker JB, Olwage A. The tick vectors of Cowdriaruminantium (Ixodoidea, Ixodidae, genus Amblyomma) and their distribution. Onderstepoort J Vet Res. 1987;54:353–79.
43. R Development Core team. R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing; 2014. http://www.R-project.org/.

44. Spickett AM, Heyne IH, Williams R. Survey of the livestock ticks of the North West province, South Africa. Onderstepoort J Vet Res. 2011;78:1–12.
45. De Matos C, Sitoe C, Neves L, Nothling JO, Horak IG. The comparative prevalence of five ixodid tick species infesting cattle and goats in Maputo Province, Mozambique. Onderstepoort J Vet Res. 2009;76:201–8.
46. Nyangiwe N, Horak IG. Goats as alternative hosts of cattle ticks. Onderstepoort J Vet Res. 2007;74:1–7.
47. MacLeod J. Tick infestation patterns in the southern province of Zambia. Bull Entomol Res. 1970;60(pt. 2):253–74.
48. MacLeod J, Colbo M, Madbouly M, Mwanaumo B. Ecological studies of ixodid ticks (Acari: Ixodidae) in Zambia. III. Seasonal activity and attachment sites on cattle, with notes on other hosts. Bull Entomol Res. 1977;67:163–73.
49. Fyumagwa RD, Runyoro V, Horak IG, Hoare R. Ecology and control of ticks as disease vectors in wildlife of the Ngorongoro Crater, Tanzania. South Afr J Wildl Res. 2007;37:79–90.
50. MacLeod J, Mwanaumo B. Ecological studies of ixodid ticks (Acari: Ixodidae) in Zambia. IV. Some anomalous infestation patterns in the northern and eastern regions. Bull Entomol Res. 1978;68:409–29.
51. Horak IG, Golezardy H, Uys A. Ticks associated with the three largest wild ruminant species in southern Africa. Onderstepoort J Vet Res. 2007;74:231–42.
52. Horak IG. Parasites of domestic and wild animals in South Africa. XV. The seasonal prevalence of ectoparasites on impala and cattle in the Northern Transvaal. Onderstepoort J Vet Res. 1982;49:85–93.
53. Yeoman GH, Walker JB, Ross JPJ, Docker TM. The Ixodid Ticks of Tanzania. A Study of the Zoogeography of the Ixodidae of an East African Country. 1967.
54. Dipeolu OO, Adeyafa CAO. Studies on ticks of veterinary importance in Nigeria. VIII. Differences observed in the biology of ticks which fed on different domestic animal hosts. Folia Parasitol Praha. 1984;31:53–61.
55. Dipeolu OO, Amoo AO, Akinboade OA. Studies on ticks of veterinary importance in Nigeria: intrinsic factors influencing oviposition and egg-hatch of Amblyommavariegatum under natural conditions. Folia Parasitol (Praha). 1991;38:63–74.
56. Rechav Y. Dynamics of tick populations (Acari: Ixodidae) in the eastern Cape Province of South Africa. J Med Entomol. 1982;19:679–700.
57. Rechav Y, Kostrzewski M, Els D. Resistance of indigenous African cattle to the tick Amblyommahebraeum. ExpApplAcarol. 1991;12:229–41.
58. Meltzer M. A possible explanation of the apparent breed-related resistance in cattle to bont tick (Amblyommahebraeum) infestations. Vet Parasitol. 1996;67:275–9.
59. Kaiser M, Sutherst R, Bourne A. Relationship between ticks and zebu cattle in southern Uganda. Trop Anim Health Prod. 1982;14:63–74.
60. Knopf L, Komoin-Oka C, Betschart B, Jongejan F, Gottstein B, Zinsstag J. Seasonal epidemiology of ticks and aspects of cowdriosis in N'Dama village cattle in the Central Guinea savannah of Côte d'Ivoire. Prev Vet Med. 2002;53:21–30.
61. Pegram R, Perry B, Schels H. Seasonal dynamics of the parasitic and non-parasitic stages of cattle ticks in Zambia. In: Acarology VI. Volume 2. DA Griffiths and CE Bowman. Chichester: Ellis Horwood; 1984. p. 1183–8.
62. Ribeiro J, Spielman A. The satyr effect: a model predicting parapatry and species extinction. Am Nat. 1986;128:513–28.
63. Kuno E. Competitive exclusion through reproductive interference. Res PopulEcol. 1992;34:275–84.
64. Yoshimura J, Clark CW. Population dynamics of sexual and resource competition. Theor Popul Biol. 1994;45:121–31.
65. Case TJ, Holt RD, McPeek MA, Keitt TH. The community context of species' borders: ecological and evolutionary perspectives. Oikos. 2005;108:28–46.
66. Goldberg EE, Lande R. Species' borders and dispersal barriers. Am Nat. 2007;170:297–304.
67. Estrada-Peña A, Horak IG, Petney T. Climate changes and suitability for the ticks Amblyommahebraeum and Amblyommavariegatum (Ixodidae) in Zimbabwe (1974–1999). Vet Parasitol. 2008;151:256–67.
68. Servedio MR. The what and why of research on reinforcement. PLoSBiol. 2004;2:e420.
69. Van Doorn GS, Edelaar P, Weissing FJ. On the origin of species by natural and sexual selection. Science. 2009;326:1704–7.
70. Coyne JA, Orr HA. Patterns of speciation in Drosophila. Evolution. 1989;362–381.
71. Coyne JA, Orr HA. Speciation. Sunderland: MA. Sinauer Associates, Inc; 2004.
72. Bargielowski IE, Lounibos LP, Carrasquilla MC. Evolution of resistance to satyrization through reproductive character displacement in populations of invasive dengue vectors. Proc Natl AcadSci. 2013;110:2888–92.
73. Higgie M, Blows MW. The evolution of reproductive character displacement conflicts with how sexual selection operates within a species. Evolution. 2008;62:1192–203.

Submit your next manuscript to BioMed Central and take full advantage of:

- **Convenient online submission**
- **Thorough peer review**
- **No space constraints or color figure charges**
- **Immediate publication on acceptance**
- **Inclusion in PubMed, CAS, Scopus and Google Scholar**
- **Research which is freely available for redistribution**

Submit your manuscript at
www.biomedcentral.com/submit

BioMed Central