Diversity in the 18S SSU rRNA V4 hyper-variable region of *Theileria* spp. in Cape buffalo (*Syncerus caffer*) and cattle from southern Africa

BEN J. MANS[1,2,*†], RONEL PIENAAR[1†], ABDALLA A. LATIF[1,2] *and* FRED T. POTGIETER[1]

[1] *Parasites, Vectors and Vector-Borne Diseases, Agricultural Research Council-Onderstepoort Veterinary Institute, Onderstepoort 0110, South Africa*
[2] *The Department of Veterinary Tropical Diseases, University of Pretoria, Pretoria, South Africa*

(*Received 27 October 2010; revised 15 December 2010 and 18 January 2011; accepted 19 January 2011; first published online 25 February 2011*)

SUMMARY

Sequence variation within the 18S SSU rRNA V4 hyper-variable region can affect the accuracy of real-time hybridization probe-based diagnostics for the detection of *Theileria* spp. infections. This is relevant for assays that use non-specific primers, such as the real-time hybridization assay for *T. parva* (Sibeko *et al*. 2008). To assess the effect of sequence variation on this test, the *Theileria* 18S gene from 62 buffalo and 49 cattle samples was cloned and ~1000 clones sequenced. Twenty-six genotypes were detected which included known and novel genotypes for the *T. buffeli, T. mutans, T. taurotragi* and *T. velifera* clades. A novel genotype related to *T.* sp. (sable) was also detected in 1 bovine sample. *Theileria* genotypic diversity was higher in buffalo compared to cattle. Polymorphism within the *T. parva* hyper-variable region was confirmed by aberrant real-time melting peaks and supported by sequencing of the S5 ribosomal gene. Analysis of the S5 gene suggests that this gene can be a marker for species differentiation. *T. parva, T.* sp. (buffalo) and *T.* sp. (bougasvlei) remain the only genotypes amplified by the primer set of the hybridization assay. Therefore, the 18S sequence diversity observed does not seem to affect the current real-time hybridization assay for *T. parva*.

Key words: 18S SSU, diversity, species, *Theileria*, *Theileria parva*.

INTRODUCTION

The order Piroplasmida (Phylum Apicomplexa) is mainly composed of the genera *Babesia* and *Theileria* (Levine *et al*. 1980; Criado *et al*. 2006). Classification of this order has been historically confined to parasite morphology, host cells in which schizogony occurs, levels of parasitaemia in the schizont and piroplasm stages, manifestation and pathology of clinical disease, as well as host and vector specificity (Barrnett, 1977; Uilenberg, 2006). By these criteria, the genus *Theileria* is distinguished by its ability to infect host leukocytes and to multiply by schizogony as part of its life cycle (Neitz, 1957; Uilenberg, 2006). *Theileria* species were first discovered based on clinical disease manifestation, linked with the observation of piroplasms in red blood cells and xenodiagnosis by tick transmission, as was the case for *Theileria parva* that caused Corridor disease and East Coast fever and *T. annulata* the causative agent of Tropical Theileriosis (Koch, 1903; Theiler, 1904; Dschunkowsky and Luhs, 1904; Lawrence, 1935). Observation of morphologically distinct piroplasm forms, with or without mild disease manifestation soon led to the discovery of more benign *Theileria*, such as *T. mutans* and *T. velifera* in cattle (Theiler, 1906; Uilenberg, 1964). New species were also discovered based on their presence in specific hosts, differential staining of blood smears and xeno-diagnosis, as in the case of *T. taurotragi* found in eland (Martin and Brocklesby, 1960; Brocklesby, 1962; Young *et al*. 1977).

Serology in the form of the indirect immunofluorescent antibody test (IFAT) was employed to distinguish species infecting the same host, but with different clinical manifestations, as is the case for *T. parva* and *T. taurotragi* (Young *et al*. 1977). The use of monoclonal antibodies and clonal cell culturing methods led to the discovery of antigenically distinct parasites from Cape buffalo (*Syncerus caffer*), thought to be *T. parva* (Conrad *et al*. 1987). Sequencing of the 18S ribosomal RNA (rRNA) gene showed, however, that *T.* sp. (buffalo) was a distinct entity from *T. parva* (Allsopp *et al*. 1993). Subsequently, the systematics of the Piroplasmida has been extensively studied by analysis of the 18S rRNA gene (Allsopp *et al*. 1994; Chae *et al*. 1999; Chansiri *et al*. 1999; Criado-Fornelio *et al*. 2004; Criado *et al*. 2006; Reichard *et al*. 2005; Altay *et al*. 2007; Bhoora *et al*. 2009). Based on this, the reverse line blot is currently used for simultaneous detection of *Babesia* and *Theileria* species diversity within a

* Corresponding author: Parasites, Vectors and Vector-Borne Diseases, Agricultural Research Council-Onderstepoort Veterinary Institute, Onderstepoort 0110, South Africa. E-mail: mansb@arc.agric.za
† These authors contributed equally.

Parasitology (2011), **138**, 766–779.
doi:10.1017/S0031182011000187

vertebrate host (Gubbels *et al*. 1999). This method depends on the amplification of a ~400 bp fragment from the 18S rRNA hyper-variable region, using universal primers for *Babesia* and *Theileria*, followed by hybridization to a panel of probes specific for different *Babesia* and *Theileria* species (Gubbels *et al*. 1999). A *Babesia*/*Theileria* genus catch-all is included to detect any species not recognized by the current probe set. This approach has been used successfully to identify new *Babesia* and *Theileria* species in antelope, black rhinoceros and dogs (Nijhof *et al*. 2003, 2005; Matjila *et al*. 2008; Oosthuizen *et al*. 2008).

The drawback of the reverse-line blot in the identification of novel species or genotypes is that it relies on animals infected with a single unique species. In the case of mixed infections, the presence of a positive signal for known species leads to the conclusion that no novel genotypes are present. This could be particularly relevant for buffalo and cattle, in which multiple *Theileria* spp. occurs (Young *et al*. 1978*a*; Allsopp *et al*. 1999; Gubbels *et al*. 1999). Random sequencing of the 18S SSU rRNA gene offers a way to obtain complementary results to the RLB for detecting *Theileria* species diversity. This approach does not depend on finding animals with single infections, while sequence coverage of an adequate sample set should give a reasonable estimate of *Theileria* sequence diversity.

T. parva is the causative agent of buffalo-derived Corridor disease and cattle-derived East Coast fever and Zimbabwe theileriosis (Norval *et al*. 1992). Whereas cattle-derived *T. parva* is considered to be eradicated in South Africa, Corridor disease is a controlled disease that is endemic in the Kruger National Park (KNP), Hluhluwe-Imfolozi Park (Kwa-Zulu Natal) and regions between and surrounding these areas (Potgieter *et al*. 1988). The movement of Cape buffalo (*Syncerus caffer*) in South Africa is thus regulated using results from a real-time hybridization PCR assay. This assay depends on amplification of a fragment of the V4 hyper-variable region of the 18S SSU rRNA gene using a *T. parva*-'specific' forward and a *Theileria* genus-specific reverse primer set (Sibeko *et al*. 2008). This primer set has been shown to amplify at least 3 *Theileria* spp. to date, *T. parva*, *T.* sp. (buffalo) and *T.* sp. (bougasvlei) (Sibeko *et al*. 2008; Zweygarth *et al*. 2009). It is therefore important to determine whether other *T. parva*-related species exist in buffalo and cattle and whether this would affect the current test. We addressed this by random sequencing of the 18S SSU rRNA hyper-variable region.

MATERIALS AND METHODS

DNA extraction and real-time hybridization and conventional PCR

Genomic DNA was extracted from 200 µl of whole blood using the MagNa Pure Large Volume DNA Isolation Kit and MagNa Pure LC (Roche Diagnostics) according the manufacturer's instructions. DNA was eluted in 100 µl of elution buffer and 2·5 µl was used for each real-time hybridization assay as described (Sibeko *et al*. 2008).

Selection of buffalo and cattle samples for analysis

Forty-nine cattle from private farms or rural communities in the vicinity of game farms and 62 buffalo samples from privately owned game farms or nature reserves were chosen from samples submitted to the Parasites, Vectors and Vector-Borne Diseases laboratory for routine *T. parva* diagnostics using the real-time hybridization PCR assay (Tables 1 and 2). The current real-time hybridization assay for *T. parva* uses 2 hybridization probe sets, the first specific for *T. parva* detectable at 640 nm and the second a *Theileria* genus-specific probe detectable at 705 nm (Sibeko *et al*. 2008). Samples negative for the 640 nm, but positive for the 705 nm probe set are designated as being *T. parva* negative, but '*T.* sp. (buffalo)-like' positive (Sibeko *et al*. 2008). This latter designation will include all *Theileria* sequences that amplify with the primer set that is not detectable with the 640 nm probe. As the drive for this study was primarily to determine the extent of closely related species to *T. parva* in southern African buffalo and cattle, samples were chosen based on their *T. parva* positive or '*T.* sp. (buffalo)-like' positive status (Tables 1 and 2). Samples were also chosen if their melting profiles differed from that of the known *T. parva*-positive controls.

Reverse-line blot

Reverse-line blot analysis was performed as described and modified (Gubbels *et al*. 1999; Matjila *et al*. 2004), with the exception that the initial 10 cycles used for touchdown PCR decreased with 1 °C steps instead of 2 °C as reported (Matjila *et al*. 2004). Probes used included those used for the detection of *Theileria* species in southern Africa (Matjila *et al*. 2008). These included probes for *T. annulata*, *T. bicornis*, *T. buffeli*, *T. equi*, *T. mutans*, *T. parva*, *T. sp*. (buffalo), *T. sp*. (duiker), *T. sp*. (kudu), *T. sp*. (sable), *T. taurotragi* and *T. velifera*, as well as a *Theileria*/*Babesia* catch-all.

Cloning and sequencing of 18S hyper-variable region

A 1101 bp fragment of the 18S rRNA gene was amplified using *Theileria* genus-specific primers designated as 989 and 990 by Allsopp *et al*. (1993). Reaction conditions included the use of 25 µl GoTaq® Green Ready reaction mix (Promega), 2·5 µl of DNA template and 10pmol of each primer up to a total volume of 50 µl. Conditions for PCR

https://doi.org/10.1017/S0031182011000187 Published online by Cambridge University Press

Table 1. Cattle samples amplified, cloned and sequenced

(*Theileria* species are designated by abbreviations and includes TbB (*T. buffeli* B), TbW (*T. buffeli* Warwick), Tm (*T. mutans*), Tm1 (*T. mutans* like-1), Tm2 (*T. mutans* like-2), Tm3 (*T. mutans* like-3), TmMSD (*T. mutans* MSD), Tp (*T. parva*), Tp1-7 (*T. parva* variants), Tt (*T. taurotragi*), Tv (*T. velifera*), TvA (*T. velifera*-like A), TsSl (*T.* sp. (sable-like)). Also shown are *Babesia bovis* (Bb).)

Sample number	Place of origin	Species identified by sequencing	Species identified by RLB
28520	Kwa-Zulu Natal	Tp, Tv	Tm, Tp, Tv
28696	Kwa-Zulu Natal	Tp, Tp3, Tv	Tm, Tp, Tv
28701	Kwa-Zulu Natal	Tv	Bb, Tv, Tm, Tt
28845	Kwa-Zulu Natal	Tp, Tp1, Tp3, Tp5, Tv	Tm, Tp, Tt, Tv
28852	Kwa-Zulu Natal	Tp, Tp5	Tm, Tp, Tt, Tv
28914	Limpopo	TsSl	TsS
29308	Mpumalanga	TmMSD, Tv	Tm, Tt, Tv
29376	Kwa-Zulu Natal	Tp, Tp1, Tp3, Tp5, Tv	Tm, Tp, Tv
29548	Kwa-Zulu Natal	Tp	Tm, Tp, Tv
30005	Kwa-Zulu Natal	Tp, TmMSD, Tv	Tm, Tp, Tv
30006	Kwa-Zulu Natal	Tm, Tp, Tp5, Tv	Tm, Tp, Tv
30007	Kwa-Zulu Natal	Tp	Tm, Tp, Tv
30008	Kwa-Zulu Natal	TmMSD, Tp, Tt, Tv	Tm, Tp, Tt, Tv
30009	Kwa-Zulu Natal	Tp, Tp5	Tb, Tm, Tp, Tv
30065	Mpumalanga	Tp, Tp2	Tp
30066	Mpumalanga	Tp, Tv	Tm, Tp, Tv
30067	Kwa-Zulu Natal	Tp, Tt, Tv	Tm, Tp, Tt, Tv
30068	Kwa-Zulu Natal	Tp, Tv	Tp,Tt, Tv
30069	Kwa-Zulu Natal	Tp, Tp3, Tt, Tv	Tm, Tp, Tt, Tv
30078	Kwa-Zulu Natal	TbB, Tp	Tb, Tm, Tp
30079	Kwa-Zulu Natal	Tp	Tm, Tp
30082	Kwa-Zulu Natal	Tp, Tv	Tb, Tm, Tp, Tv
30166	Kwa-Zulu Natal	Tt, Tv	Tm, Tp, Tt, Tv
30175	Mpumalanga	Tm, TmMSD, Tv	Tm, Tp, Tv
30423	Kwa-Zulu Natal	Tp, Tp3	Tm, Tp
30425	Kwa-Zulu Natal	Tm, Tp3, Tp5, Tv	Tm, Tp, Tv
30463	Kwa-Zulu Natal	Tv	Tm, Tt, Tv
30467	Kwa-Zulu Natal	Tm, Tv	Tm, Tt, Tv
30492	Kwa-Zulu Natal	Tt, Tv	Tt, Tv
30521	Kwa-Zulu Natal	Tv	Tv
30525	Kwa-Zulu Natal	TmMSD, Tt	Bb, Tt
30558	Mpumalanga	TmMSD, Tv	Tm, Tv
30561	Mpumalanga	Tp, Tp3, Tv	Tm
30626	Kwa-Zulu Natal	Tp, Tv	Tm
30642	Kwa-Zulu Natal	Tv	Bb, Tm, Tt, Tv
31043	Kwa-Zulu Natal	Tt, Tv	No sample
31305	Kwa-Zulu Natal	Tp, Tv	Tm, Tp, Tv
31654	Kwa-Zulu Natal	Tm, TmMSD, Tp, Tv	Tm
34358	Kwa-Zulu Natal	Tp, Tp1, Tv	Tm, Tp, Tv
34431	Kwa-Zulu Natal	Tp3, Tp5, Tp6, Tv	Tm, Tv
34441	Kwa-Zulu Natal	Tm, Tv	Tm, Tv
34448	Kwa-Zulu Natal	Tv	Tm, Tv
35010	Kwa-Zulu Natal	Tp, Tv	Tp, Tv
35496	Kwa-Zulu Natal	Tp, Tp5, Tv	Tm, Tv
35497	Kwa-Zulu Natal	TbB, Tm, Tp, Tp4	Tb, Tm, Tp
35498	Kwa-Zulu Natal	TbB, TbW, Tv	Tb, Tm
35499	Kwa-Zulu Natal	TbB, TbW, Tp, Tp7	Tb, Tm, Tp, Tt
35500	Kwa-Zulu Natal	Tm, Tp, Tv	Tm
36478	Kwa-Zulu Natal	Tm, TmMSD	Tm, Tv

included a pre-melt of 2 min (94 °C), 40 cycles of melting at 94 °C (30s), annealing at 52 °C (30s) and extension at 72 °C (120s), with a final extension for 7 min at 72 °C. Amplified products were analysed by 1% agarose gel electrophoresis and 1 μl used for cloning into the pGEM®-T Easy cloning kit according to the manufacturer's instructions (Promega). Blue-white selection was used to identify clones with inserts, and clones were screened using the 989/990 primers. Amplified products (1101 bp) were purified for 10 positive clones from each sample using the DNA extraction kit (Fermentas). Purified fragments were sequenced using the BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) and the *Theileria* genus-specific primer 989. To confirm the novel genotypes observed, selected samples were

https://doi.org/10.1017/S0031182011000187 Published online by Cambridge University Press

Table 2. Buffalo samples amplified, cloned and sequenced

(*Theileria* species are designated by abbreviations and includes Bgvl (*T.* sp. (bougasvlei)), TbC (*T. buffeli*-like C), TbDl (*T. buffeli* type D-like), Tbic (*T. bicornis*), Tm, (*T. mutans*), Tm1 (*T. mutans* like-1), Tm2 (*T. mutans* like-2), Tm3 (*T. mutans* like-3), TmMSD (*T. mutans* MSD), Tp (*T. parva*), Tp1-5 (*T. parva* variants), TsB (*T. sp.* (buffalo)), TsB1-2 (*T. sp.* (buffalo) variants), Tv (*T. velifera*), TvA (*T. velifera*-like A), TvB (*T. velifera*-like B).)

Sample number	Place of origin	Species identified by sequencing	RLB Species identified
28698	Limpopo	TsB, Tv, TvB	Tm, TsB, Tv
31898	Mpumalanga	TsB, TsB2	Tp, TsB, Tv
31978	Limpopo	Tm1, Tp1, TsB, Tv	Tm, TsB, Tv
31989	Limpopo	Tm1, Tm2, TsB, TvB	Tm, TsB, Tv
31995	Limpopo	TbC, TbDl, Tv	Tb, TsB, Tm
31996	Limpopo	Tm, Tm1, Tm3, Tp, TsB, Tv, TvB	TsB, Tm, Tv
32006	Limpopo	Tm1, TsB, Tv, TvB	Tm, Tp, TsB, Tv
32007	Limpopo	Tm1, Tm3, TsB, Tv, TvB	Tm, Tp, TsB, Tv
32017	Limpopo	Tp, TsB, Tv, TvB	Tm, Tp, TsB, Tv
32668	Limpopo	TsBgvl, Tv, TvB	Tm, Tv
32751	Limpopo	TmMSD, Tm2, TsB, TvB	Tb, Tm, TsB, Tv
32754	Limpopo	TbC, TbDl, Tm1, Tv	Tb, Tm, TsB, Tv
32761	Limpopo	TsB, Tv	Tm, Tp, TsB, Tv
32764	Limpopo	Tm1, Tm2, Tm3, Tp, TsB, Tv, TvB	Tm, Tp, TsB, Tv
32768	Limpopo	TsB, TsBgvl, TvB	Tm, TsB, Tv
32769	Limpopo	TsB, TvB	Tm, TsB, Tv
32772	Limpopo	TsB, Tv, TvB	Tm, Tp, TsB, Tv
32780	Limpopo	Tm1, Tm2, Tm3, Tp1, TsB, TvB	Tm, Tp, TsB, Tv
32784	Limpopo	Tm1, Tp, TsB, Tv, TvB	Tm, TsB, Tv
32792	Limpopo	TmMSD, Tp TsB	Tm, TsB, Tv
32797	Mozambique	TbC, Tp1, TsB	Tb, TsB, Tv
32798	Mozambique	TbC, TsB, TsBgvl	Tb, TsB, Tv
32799	Mozambique	TbC, TsBgvl	Tb, TsB, Tv
32800	Mozambique	TbC, TsB	Tb, TsB
32806	Mozambique	TbC, TbDl, Tp1, TvA	Tb, Tm, TsB, Tv
32811	Mozambique	TbC, Tp1, TsB	Tb, TsB, Tv
32812	Mozambique	TbC, Tm, TvA	Tb, Tm, TsB, Tv
32813	Mozambique	TbC, TsB	Tb, Tm, TsB, Tv
32815	Mozambique	TbC, TsB	Tb, Tm, TsB, Tv
32817	Mozambique	TbC, Tp1, TsB, TvA	Tb, Tm, TsB, Tv
33703	Limpopo	TsBgvl, TvB	Tm, Tv
33999	Free State	Tm1, Tm3, TsB, Tv	Tm, Tp, TsB, Tv
35524	Kwa-Zulu Natal	Tp, Tp5, TsB	Tm, Tp, TsB, Tv
35608	Zimbabwe	Tm1, Tp, TsBgvl	Tm, Tp, Tv
36497	Mpumalanga	Tm, TmMSD	Tm
36780	Limpopo	Tm, TmMSD, TvB	Tm, Tv
40990	Limpopo	Tm, Tm3, TsBgvl, TvB	Tm, Tv
41001	Limpopo	Tm1, Tm3, Tv, TvB	Tm, TsB, Tv
41004	Limpopo	Tm, Tm1, TsB, Tv, TvB	Tm, TsB, Tv
41150	Kwa-Zulu Natal	TbDl, Tm2, TsB	Tb, Tbic, Tm, TsB
41168	Kwa-Zulu Natal	TbDl, TsB	Tb, Tbic, Tm, TsB
A10	Limpopo	Tm1, Tp, TsBgvl, Tv	Tv
A20	Limpopo	Tm1, Tv, TvB	Tm, Tv
C21	Limpopo	Tm1, Tm2, Tm3, TvB	Tm, Tv
E9	Limpopo	Tm1, Tm3, TsBgvl, TvB	Tm, Tv
E12	Limpopo	Tm1, Tm3, Tv	Tm, Tv
E15	Limpopo	Tm1, Tm3, TvB	Tm
E36	Limpopo	Tm1, Tm2, Tm3, Tp, TsBgvl	Tm
G14	Limpopo	Tm1, Tm3, TsBgvl, Tv, TvB	Tm, TsB
H12	Limpopo	Tm1, Tm2, Tm3, TsB, TsBgvl, Tv, TvB	Tm, TsB, Tv
R 32	Limpopo	Tm1, Tp, TsB, TsB1, Tv	Tm, Tp, TsB, Tv
R 34	Limpopo	Tm1, Tm2, Tm3, Tp, TsB, TvB	Tm, Tp, TsB, Tv
R 35	Limpopo	Tm3, TsB, TsB1	Tm, Tp, TsB, Tv
R 36	Limpopo	Tm1, Tm3, Tp, Tv	Tm, Tp, TsB, Tv
R 37	Limpopo	Tm1, Tm2, Tm3, Tp1, TsB, TvB	Tm, Tp, TsB, Tv
R 39	Limpopo	Tm1, Tm3, Tp, TsB	Tm, Tp, TsB, Tv
W1	Limpopo	Tm1, Tm2, Tp, TsB, TvB	Tm, TsB
W4	Limpopo	Tm1, Tm2, Tm3, Tp, TsB, Tv	Tm, Tp, TsB, Tv
Y19	Limpopo	Tm1, Tp, TsB, TsB2, Tv, TvB	Tm, Tp, TsB, Tv
Y29	Limpopo	Tm1, Tm3, Tp, TsB2, Tv	Tm, Tp, TsB, Tv
Y40	Limpopo	Tm1, Tm2, Tm3, TsB, Tv, TvB	Tm, Tp, TsB, Tv
Y41	Limpopo	Tm2, Tp, TsB, TsB1, TsBgvl, Tv	Tm, Tp, TsB, Tv

https://doi.org/10.1017/S0031182011000187 Published online by Cambridge University Press

re-amplified, cloned using high fidelity *Pfu* polymerase (Fermentas) and sequenced from both ends.

Bioinformatic analysis of the 18S gene

Sequences for each sample and a master sequence set that included known *Theileria* species sequences were aligned using a consideration of the secondary structure of RNA (Q-INS-i) as implemented in MAFFT (Katoh and Toh, 2007). Alignments were manually inspected and adjusted and edges were trimmed to give a ~320 bp region that spanned the V4 hyper-variable region that was used for phylogenetic analysis. Neighbor-Joining analysis was performed using Mega4 software (Tamura *et al*. 2007). Gaps were treated as pairwise deletion and both transitions and transversions were included in the analysis. Due to the relatively limited number of phylogenetically informative positions available in the fragment analysed, pairwise-distances were calculated using the p-distance parameter, which was also used as nucleotide substitution model (Nei and Kumar, 2000). Patterns among lineages were assumed to be homogeneous and rates among sites to be uniform. Branch support was estimated using bootstrap analysis (100 000 replicates). Sequences assigned to putative species were aligned as a group and manually checked for single nucleotide polymorphisms.

Cloning and sequencing of the S5 gene

Primers for the S5 gene (S5F: ATG ACA AAC ACA GAA GTC GCC CT; S5R: ATT TCA TCC TTC TTC TTG ATT GCG T) were designed based on conserved regions between the *T. annulata* and *T. parva* S5 ribosomal genes that were published as part of their genome sequencing projects (Gardner *et al*. 2005; Pain *et al*. 2005). Conventional PCR was performed using an initial denaturation step at 94 °C (2 min), followed by 40 cycles of denaturation at 94 °C (30s), annealing at 53 °C (30s) and extension at 72 °C (60s). A final extension step at 72 °C (7 min) was performed to ensure that all PCR products were extended. Reaction conditions included the use of 25 μl GoTaq® Green Ready reaction mix, 2·5 μl of DNA template and 10pmol of each primer up to a total volume of 50 μl. Amplified products (1 μl) were used for cloning into the pGEM®-T Easy vector as described above for the 18S gene. DNA products were purified for 10 positive clones and sequenced from both sides using the BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) and the M13 forward and reverse primers. Control samples chosen were previously confirmed by sequencing of the 18S gene and included *T. taurotragi* (18565), *T.* sp. (buffalo) (25317, 26705), *T.* sp. (bougasvlei) (26531, 27071), *T. parva* (KNP102, 9288, 9446, 9468), *T. parva* variant 1 (32806), *T. parva* variant 3 (34431) and *T. lestoquardi* (Morocco). Numbers in parentheses refer to laboratory numbers of samples submitted for diagnostic analysis.

Bioinformatic analysis of the S5 gene

Sequences were aligned using ClustalX (Jeanmougin *et al*. 1996) and manually checked. Particular attention was paid to the intron-exon boundaries and open reading-frames. Neighbor-joining analysis was performed with pairwise-deletion of gapped positions. The Tamura 3-parameter was used as substitution model which included transversions, transitions, first, second and third codon positions as well as the non-coding intronic regions. Patterns among lineages were treated as heterogenous and rates among sites (787 sites) as uniform. Branch confidence was estimated using bootstrap analysis for 100 000 replicates.

RESULTS

Detection of species diversity

Cloning and sequencing of 10 clones each of the 18S gene from 49 cattle and 62 buffalo samples yielded 463 and 598 high-quality sequences, respectively. Sequences were analysed by multiple sequence alignment and phylogenetic analysis using a master sequence set of known *Theileria* genotypes (Fig. 1). Twenty-six unique genotypes were found in buffalo and cattle and grouped into 5 major evolutionary groups or clades (Fig. 1). The '*T. taurotragi*' group included *T. parva* (8 genotypes), *T.* sp. (bougasvlei), *T.* sp. (buffalo) (3 genotypes) and *T. taurotragi*. The *Theileria* antelope group included a single genotype, *T.* sp. (sable)-like found in a bovine. The '*T. mutans*' group had 5 genotypes, while the '*T. buffeli*' and '*T. velifera*' groups, possessed 4 and 3 genotypes, respectively. Of these, *T. taurotragi*, *T.* sp. (sable)-like, *T. buffeli*-Warwick and *T. buffeli*-like B was found exclusively in cattle. *T. buffeli*-like C, *T. buffeli* D-like, *T. mutans*-like 1, *T. mutans*-like 2, *T. mutans*-like 3, *T.* sp. (bougasvlei), *T.* sp. (buffalo) and its variants, as well as *T. velifera* B were found exclusively in buffalo.

Species diversity in buffalo and cattle

The *Theileria* genotypic diversity found in buffalo, both in regard to number of genotypes harboured and number of co-infecting genotypes, are more extensive than in cattle when the random sequencing data are compared (Figs 1 and 2). For the random sequencing approach, a total of 16 genotypes (9 unique) were found in cattle, while 17 genotypes (10 unique) were found in buffalo (Fig. 1). In contrast, RLB analysis of the same sample set detected only 6 genotypes in cattle (*T. buffeli*, *T. mutans*, *T. parva*, *T.* sp. (sable), *T. taurotragi*, *T. velifera*) and 5 in buffalo (*T. buffeli*,

https://doi.org/10.1017/S0031182011000187 Published online by Cambridge University Press

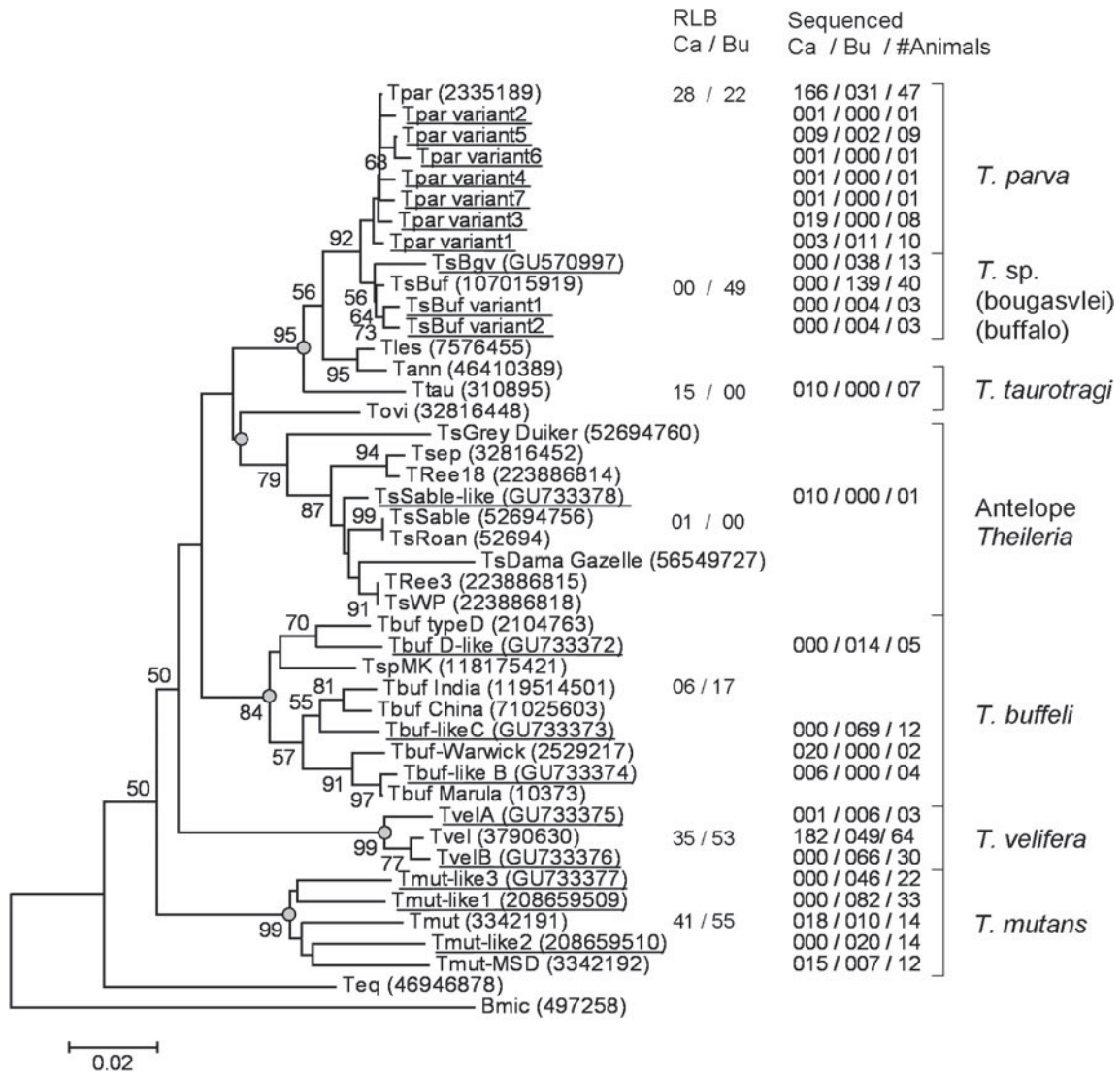

Fig. 1. Phylogenetic analysis of *Theileria* species and genotypic variants found in buffalo and cattle. Major phylogenetic groups are indicated by grey circles. The number of each species found by reverse-line blot analysis and the number of sequenced clones found in cattle (Ca) and buffalo (Bu) are indicated, as well as the number of animals (#Animals) from which these clones were derived. Unique genotypes identified in this study are underlined. Species are indicated by a 3–5 letter abbreviation, followed by their GenBank Accession code. The tree was rooted using *Babesia microti* as outgroup. Bootstrap support above 50% is indicated.

T. mutans, *T. parva*, *T.* sp. (buffalo), *T. velifera*). This can be expected as the unique genotypes observed in this study are not included in the current RLB probe set.

For the sequenced samples up to 4 different *Theileria* genotypes were found per cattle sample, with the majority (94%) harbouring a maximum of 1–3 genotypes (Fig. 2). In contrast, buffalo carried up to 7 different genotypes per sample, with the majority (93%) harbouring between 2 and 6 genotypes. Up to 4 genotypes per sample were found in cattle and buffalo with the RLB. Reverse-line blot analysis of the cattle samples indicate that, in general, more genotypes are detected per sample than observed for the random sequencing approach with the majority (85%) harbouring 2–4 genotypes (Fig. 2). In the case of the buffalo, 45% of all samples harboured 4 genotypes.

T. *sp. (sable)-like sequences in cattle*

All clones sequenced for 1 bovine were identical to each other and similar (2 nucleotide differences) to sequences recently obtained from *R. evertsi evertsi* ticks collected from a gemsbok (*Oryx gazella gazella*) (Tonetti *et al*. 2009). These sequences are related to *T.* sp. (sable) and *T.* sp. (roan) derived sequences and cross-hybridize with the RLB probe for *T.* sp. (sable) (Nijhof *et al*. 2005). Differences across the fragment analysed (5 nucleotide differences) are large enough to be considered unique.

The T. buffeli *clade*

Four different genotypes were found within the *T. buffeli* group of which the complementary region to the RLB probe is more than 95% conserved

https://doi.org/10.1017/S0031182011000187 Published online by Cambridge University Press

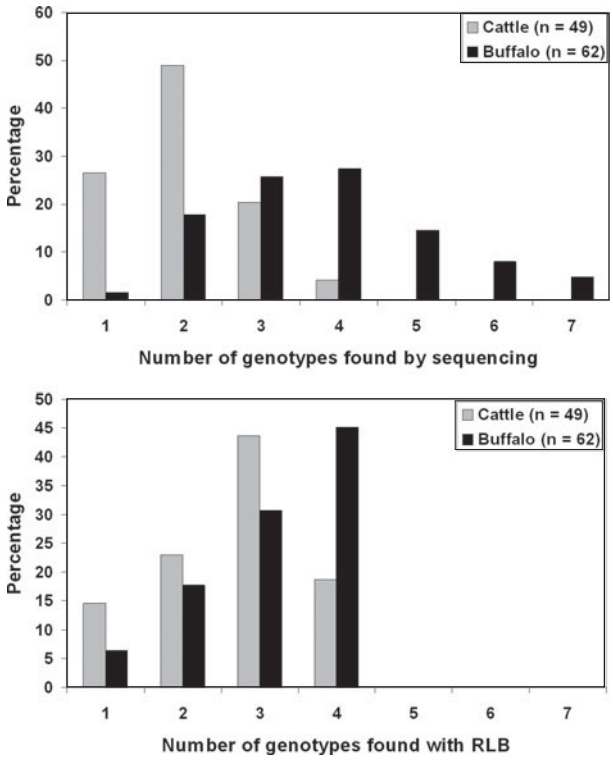

Fig. 2. Number of *Theileria* genotypes found in buffalo and cattle by random sequencing and RLB. Indicated is the number of species found per animal against the total percentage.

(Fig. 3). However, significant differences exist upstream from this region to designate all as unique genotypes. *T. buffeli* D-like were found in 3 buffalo samples and is similar to *T. buffeli* type D and *T.* sp. (Thung Song) found in cattle in Thailand, the USA and Korea (Chae *et al*. 1999; Chansiri *et al*. 1999) and a sequence designated as *T.* sp. MK from sheep and goats in Turkey (Altay *et al*. 2007), but differed by at least 8 nucleotide substitutions. Twelve buffalo samples (*T. buffeli*-like C) presented sequences similar to unpublished *T. buffeli* sequences from China and India. Four cattle possessed a genotype (*T. buffeli*-like B) similar (1 nucleotide difference) to *T. buffeli* (Marula), previously found in a bovine in Kenya (Allsopp *et al*. 1994). Two cattle samples presented sequences identical to *T. buffeli* (Warwick) originally isolated from cattle in Australia (Stewart *et al*. 1987; Chansiri *et al*. 1999).

The T. velifera *clade*

The sequence for *T. velifera*, as described from the original isolate obtained from cattle (Uilenberg and Schreuder, 1976; Gubbels *et al*. 1999), was found to be prevalent in both cattle and buffalo (64 animals). Two additional genotypes differed from the published *T. velifera* sequence with 2 (*T. velifera* B) and 7 (*T. velifera* A) nucleotides, respectively. *T. velifera* B was limited to buffalo. In all genotypes the complementary region to the RLB probe for *T. velifera* is more than 90% conserved (Fig. 3).

The T. mutans *clade*

Sequences that were identical to the Intona strain of *T. mutans* from a bovine in Kenya and *T. mutans* MSD from a bovine in South Africa (De Vos and Roos, 1981; Chae *et al*. 1999), were found in both buffalo and cattle. In addition, sequences limited to buffalo that are identical to 2 unpublished sequences (*T. mutans*-like1 and *T. mutans*-like 2) obtained from buffalo in the Kruger National Park, were found in 47 buffalo samples. A fourth unique genotype (*T. mutans*-like 3) was found in 22 buffalo samples.

The T. taurotragi *clade*

A well-supported clade is formed by the established species *T. taurotragi, T. annulata, T. lestoquardi* and *T. parva,* which have known vectors, host and disease aetiologies (Neitz, 1957; Young *et al*. 1977). Also found in this clade are the less well-characterized parasites, *T.* sp. (buffalo) and *T.* sp. (bougasvlei) known to occur in buffalo and thought to be apathogenic (Allsopp *et al*. 1993; Zweygarth *et al*. 2008). Of these, *T. parva* was found in buffalo and cattle, while *T. taurotragi* was limited to cattle. Conversely, *T.* sp. (buffalo) and *T.* sp. (bougasvlei) were only found in buffalo.

T. parva *variants*

Pairwise genetic distances based on the p–distance model of nucleotide substitution show a linear correlation with the number of pairwise nucleotide differences (Fig. 4). Genetic distances between well-described, established and closely related species show that the smallest genetic distance (0·01) found was between *T. annulata*/*T. lestoquardi* and *T. parva*/*T.* sp. (buffalo), respectively. This translates to 3 nucleotide differences between each pair, respectively for the ~300 bp fragment analysed of the SSU 18S RNA gene. In the other evolutionary related groups much higher variation is observed. Genotypes from the *T. mutans* group had pairwise distances above 0·043, which corresponds to more than 13 nucleotide differences, while the *T. buffeli* group had distances above 0·017, which corresponds to 5 or more nucleotide differences.

On the other hand, genetic distances smaller (less or equal than 3 nucleotide differences) would fall in a grey area where species discrimination would be more difficult. Pairwise distances of 0·003 (1 nucleotide difference) were obtained and corresponded with sequences that were designated as *T. parva* or *T.* sp. (buffalo) variants (Figs 1 and 3). These differences all occur within the 640 nm probe region of the hybridization probe for *T. parva* (Fig. 3).

https://doi.org/10.1017/S0031182011000187 Published online by Cambridge University Press

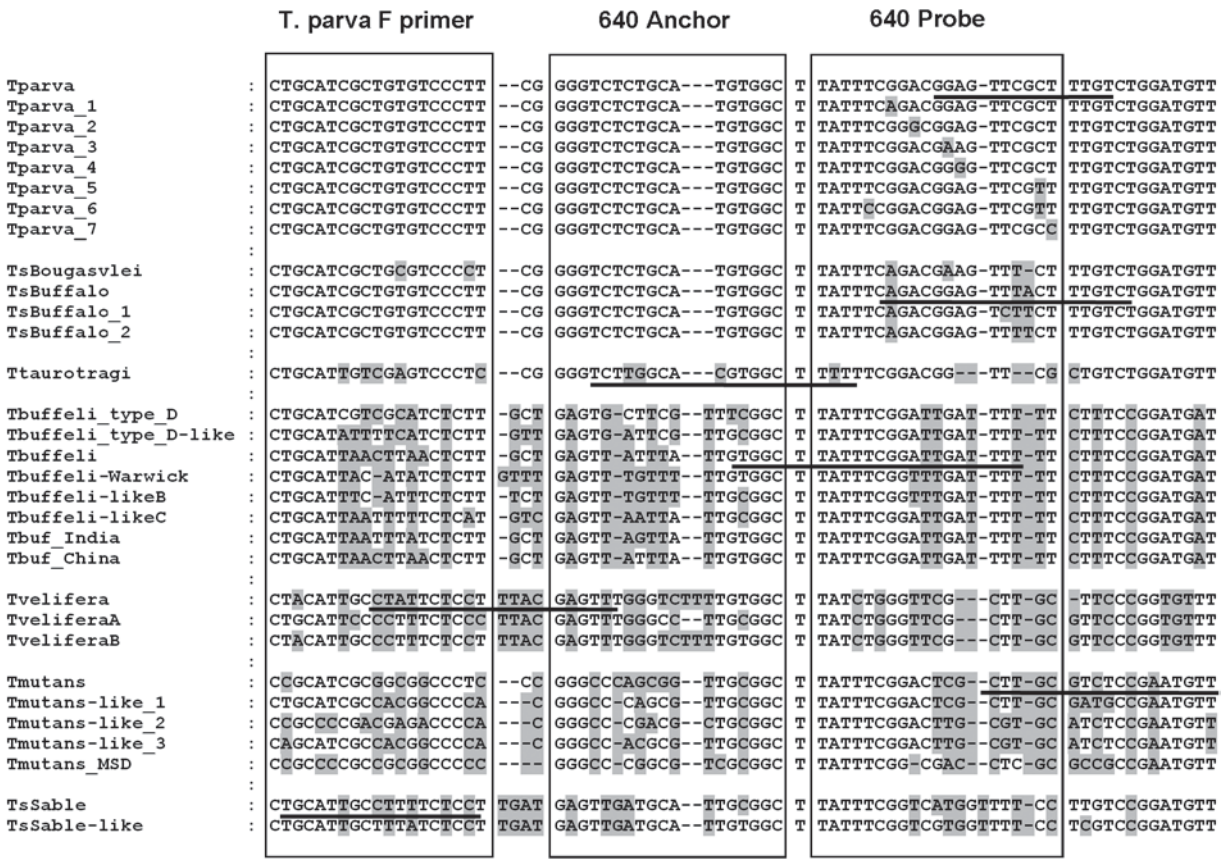

Fig. 3. Alignment of the 18S SSU RNA V4 hyper-variable region of the genotypes found in buffalo and cattle. Indicated are the regions for the *Theileria parva* 'specific' forward primer and the 640 anchor-probe pair for the real-time hybridization test for *T. parva* (Sibeko *et al*. 2008). Reverse-line blot probe regions are underlined and grey shaded areas indicate positions that differ from the *T. parva* sequence. Accession codes for sequences can be found in Fig. 1.

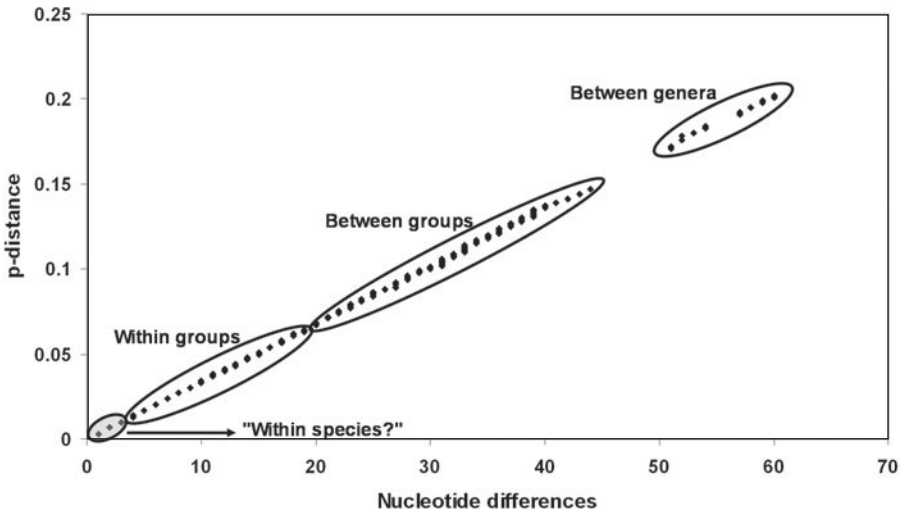

Fig. 4. Genetic and nucleotide differences of the 18S SSU V4 hypervariable region. Indicated are the ranges within and between the major groups and genera, as well as what is considered to be variation within species. Genetic distances are based on the p-distance measure and pairwise nucleotide differences were calculated with all pairwise gaps deleted.

In order to find further confirmation that the *T. parva* variants do exist, field samples for which variant sequences were obtained were analysed by real-time PCR (Fig. 5). The positive *T. parva* control sample (KNP102) gave 2 peaks as previously described (Sibeko *et al*. 2008). This includes a main peak at ~64 °C as well as a shoulder peak at ~52 °C. The samples that possess variant sequences have profiles that differ from that of the control, in that the Tm of the shoulder peak shifts to ~57 °C. Such

https://doi.org/10.1017/S0031182011000187 Published online by Cambridge University Press

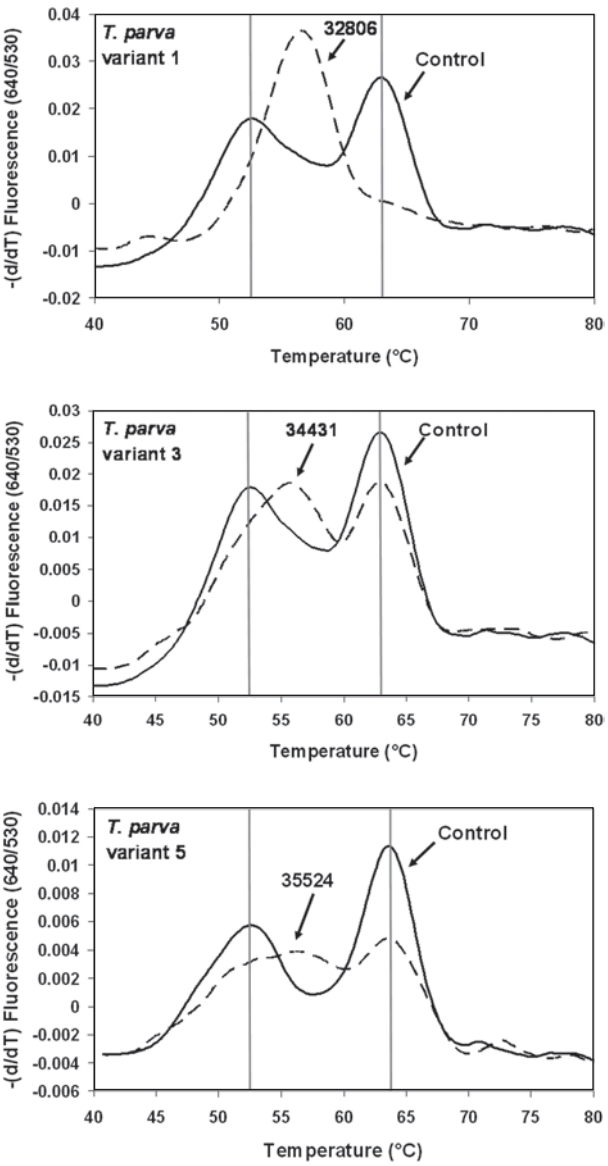

Fig. 5. Real-time PCR melting profiles of *Theileria parva* variant samples. Indicated are melting profiles obtained with the 640 nm anchor-probe pair. Vertical grey lines indicate the shoulder (51 °C) and main (64 °C) melting peaks for positive control KNP102.

profiles were observed in the case of variant 1 for samples 32797 and 32806. In both of these samples, no canonical *T. parva* sequences were obtained and in the case of 32806, 5 clones from the 10 sequenced, represented the variant 1 sequence (Table 1). For variant 3, profiles that differed from the norm were obtained for samples 30069, 30423 and 34431 and in the case of 34431 six out of 10 clones gave the sequence for variant 3 and no canonical *T. parva* clones were found. In the case of *T. parva* variant 5, all animals that possessed sequences for this variant also possessed canonical *T. parva* sequences. One sample, 35524, gave a melting profile that differed from that of the standard *T. parva* (Fig. 5), suggestive of the presence of variants in this animal.

Species differentiation using the S5 ribosomal gene

To confirm that the variants observed is indeed *T. parva* and not another closely-related species, a gene (S5 ribosomal protein), with more phylogenetic informative characters than the 18S gene was identified. The primers designed exclusively amplified a 779 bp fragment from members of the '*T. taurotragi*' clade that is composed of a 554 bp stretch of coding region, with 3 interspersed introns (86 bp, 113 bp and 30 bp, respectively). As such, the S5 fragment should be useful to investigate the genetic relationship that exists between *T. parva*, *T.* sp. (buffalo) and *T.* sp. (bougasvlei) and their variants in more detail. The primer set should also allow for the species-specific amplification of this gene from animal samples that harbour parasites from the other major groups. Phylogenetic analysis of the S5 gene shows that *T. parva*, *T.* sp. (buffalo) and *T.* sp. (bougasvlei) group into distinct and well-supported clades (Fig. 6). Furthermore, the *T. parva* variant 1 and variant 3 sequences group within the *T. parva* clade suggesting that these 18S SSU rRNA variants are indeed *T. parva*. A linear correlation can be observed when nucleotide and genetic distances are compared and correlate with the same plot for the 18S gene (Figs 4 and 6). Demarcation between genotypes are, however, much more distinct than observed for the 18S gene, with within-species variation grouping in the lower left quadrant of the plot with pairwise genetic distances less than 0·02 (<15 nucleotide differences) observed. Pairwise genetic-distances for between-species variation are 0·039 for *T. annulata* and *T. lestoquardi* (30 nucleotide differences) and above 0·067 (>50 nucleotide differences) for the rest.

DISCUSSION

The discovery of species diversity in the *Theileria* originated from direct observation of clinical disease development and parasite morphology linked to tick transmission as encapsulated in the classical paradigm of Koch's postulates. This progressed to observation of benign *Theileria* forms and the confusion engendered by morphological similarities in the absence of clinical disease progression, to the current state of species definition based on the uniqueness of their 18S SSU rRNA molecular sequences.

The current study showed how a random sequencing approach can be an effective discovery vehicle for detection of new genotypes and possibly new species. The question is raised whether the genotypic diversity observed in the current study is exhaustive, given the sample set and sequence depth coverage obtained. In this regard, no new genotypes were observed after ~200 clones were sequenced each for buffalo and cattle samples, respectively. This

https://doi.org/10.1017/S0031182011000187 Published online by Cambridge University Press

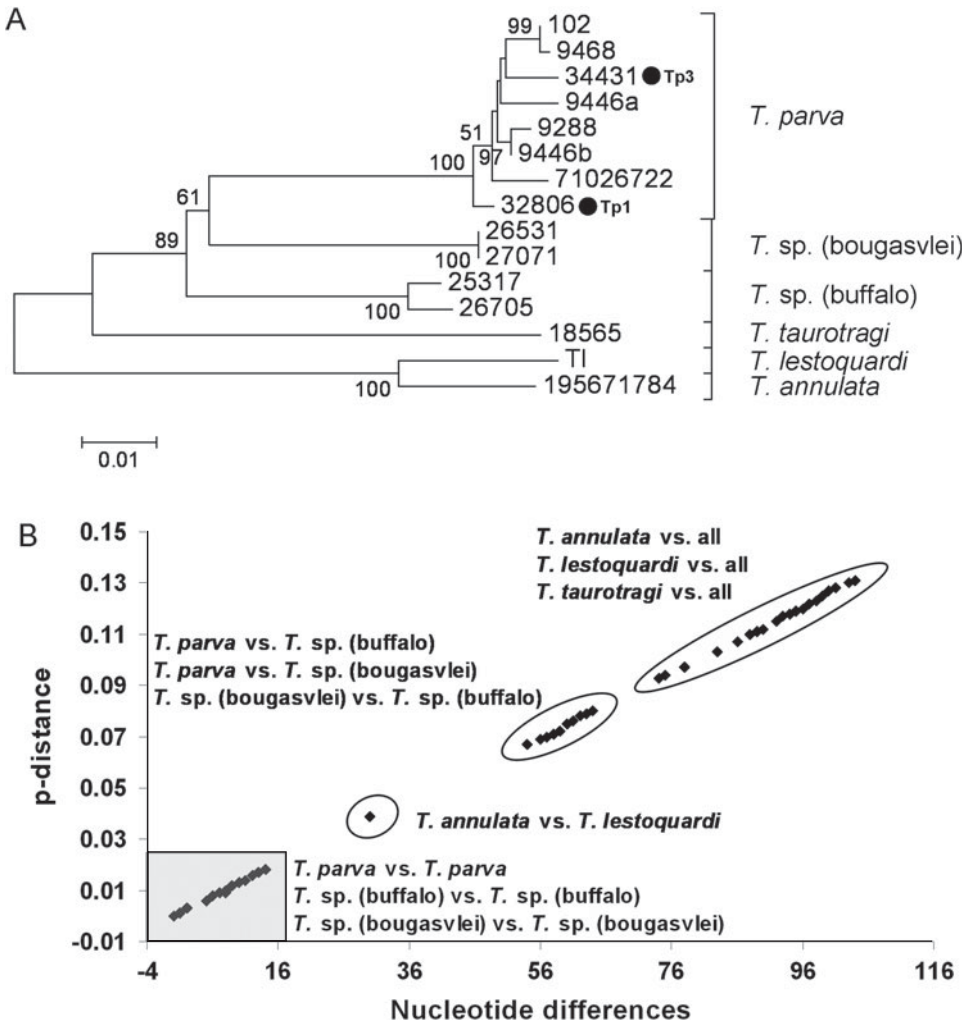

Fig. 6. Analysis of the S5 ribosomal gene. (A) Neighbor-joining analysis of the S5 ribosomal gene. Bootstrap support values (100 000 replicates) are indicated. Numbers indicate animal numbers (5 digits) or GenBank Accession codes. (B) Genetic and nucleotide differences are indicated within and between species. Genetic distances are based on the p-distance measure and pairwise nucleotide differences were calculated with all pairwise gaps deleted.

suggests that genotypic diversity was saturated with the current sample set. However, the possibility that new genotypes will be found if different geographical areas are screened cannot be excluded. This might be true for the *T. buffeli* group, of which a number of diverse genotypes have been observed (Gubbels *et al*. 2000).

The RLB analysis was more sensitive than the random sequencing approach in detecting genotypes in the case of cattle samples. This could be expected, given the fact that only 10 clones were sampled per specimen for sequencing. The larger number of genotypes found in the buffalo samples using RLB could, however, also indicate that parasitaemia in cattle compared to buffalo for any given genotype is much higher. The sampling size for cattle-derived clones may therefore affect the number of genotypes detected. Alternatively, based on the similarity of the RLB probes and the genotypes that was found in buffalo, cross-hybridization is expected to occur between members of the *T. mutans* and *T. buffeli* clades. This could explain the fact that buffalo seem to have higher numbers of genotypes per animal compared to cattle.

The discovery of numerous genotypes that may cross-hybridize with current RLB probes highlights problems that can occur with probe-based technologies. In this regard, *T.* sp. (sable) was detected previously in buffalo and cattle (Nijhof *et al*. 2005). It should be noted that the current *T.* sp. (sable) probe will also cross-hybridize with the *T. velifera* group, making previous identification of this genotype in buffalo and cattle, dubious (unpublished experimental observation).

The questions are raised whether the 26 different genotypes found in buffalo and cattle in the current study constitute different species, or mere variants or subspecies of the major groups described, or if multiple variable 18S SSU rRNA copies reside inside a single species. No consensus has yet been reached on the degree of variation within the 18S SSU rRNA or other genes that would define a species (Chae *et al*.

https://doi.org/10.1017/S0031182011000187 Published online by Cambridge University Press

1999; Gubbels *et al*. 2000; Bhoora *et al*. 2009). It has been suggested that multiple copies of the 18S SSU RNA could account for genetic variation within a species, such as the case for *Plasmodium* and *Babesia*, where differential expression of 18S rRNA occurs as a response to various host environments and life-cycle stages (Gunderson *et al*. 1987; Chae *et al*. 1999; Lau, 2009; Laughery *et al*. 2009). It has been shown that 2 copies of the 18S gene exist in the genomes of *T. annulata* and *T. parva* and is as such probably conserved in all *Theileria* spp. (Gardner *et al*. 2005; Pain *et al*. 2005). The copies are identical in the 18S hyper-variable region of the genomes sequenced, probably due to concerted evolution (Eickbush and Eickbush, 2007). Co-amplification of different copies would indicate their presence in the same genome, but in the current study, no consistent pattern of co-amplification was observed for the different genotypes. It is therefore unlikely that the high degree of variation observed within the different groups could be due to copy variation or for that matter variation within the species population.

Differences observed within the *T. buffeli*, *T. mutans* and *T. velifera* groups probably indicate unique species. Previously, a higher degree of genetic diversity has been observed within the *T. buffeli*-group for various genetic markers compared to other *Theileria* species (Gubbels *et al*. 2000). While it was acknowledged that this could be an indication of multiple species within this group, the limited knowledge at the time precluded a definitive conclusion regarding the designation of new *T. buffeli*-related species. The current study is similarly limited and can, as yet, not conclude that the novel genotypes observed are new species. However, host distributions of the various genotypes do support the idea that distinctive biological differences exist for the various genotypes. As such, *T. velifera* B, *T. buffeli*-like C and *T. buffeli*-like D were restricted to buffalo compared to *T. velifera* that was found in both buffalo and cattle and *T. buffeli*-Warwick and *T. buffeli*-like B that were limited to cattle. Similarly, *T. mutans*-like 1, *T. mutans*-like 2 and *T. mutans*-like 3 were found solely in buffalo, while *T. mutans* and *T. mutans* MSD were found in cattle and buffalo. The within-group genetic differences in the *T. buffeli*, *T. mutans* and *T. velifera* clades are also much larger than that of members of the *T. taurotragi* clade. The basal positions of the *T. buffeli*, *T. mutans* and *T. velifera* clades relative to that of the *T. taurotragi* clade suggest that they are more ancient, which could explain the larger genetic differences. This implies that much larger genetic variation can be expected across their genomes, that will likely result in genic incompatibility and hybrid sterility during recombination (Coyne and Orr, 2004).

As the tick midgut is the site of sexual recombination for *Theileria* parasites (Norval *et al*. 1992), vector specificity as a means of geographical isolation could be a viable mechanism for speciation. Even so, host specificity as some of the current data would suggest, could be a viable alternative mechanism for speciation, given the fact that transmission of *Theileria* occurs transtadially. Closely related species, such as *T.* sp. (buffalo), *T.* sp. (bougasvlei) and *T. parva* which seem to occur sympatric in their non-recombining stages would most possibly have different tick vectors. The same would hold for the '*T. buffeli*' and '*T. mutans*' groups. Several strains of *T. mutans* and *T. buffeli* were previously characterized that differed in regard to tick vector specificity (de Vos *et al*. 1981; Stewart *et al*. 1996; Young *et al*. 1978*b*). Whether this would imply that these different strains were different species remains to be determined.

The *Theileria* species diversity found in Cape buffalo relative to cattle parasites suggest that buffalo could be the ancestral reservoir host for some *Theileria* parasites found in bovids in Africa as suggested for *T. parva*, *T. annulata* and *T. mutans* (Uilenberg *et al*. 1981; Young *et al*. 1981; Conrad *et al*. 1989; Norval *et al*. 1991, 1992). In the case of genotypes limited to cattle (*T. taurotragi*, *T. sable*-like, *T. buffeli* Warwick and *T. buffeli*-like B), it is possible that antelopes could have been the ancestral hosts (Young *et al*. 1977). Alternatively, parasites such as *T. buffeli* could have been introduced from other geographical areas, where these parasites were adapted to wild ungulates other than the Cape buffalo, such as Asiatic buffalo (*Bubalus bubalis*) (Stewart *et al*. 1996; Gubbels *et al*. 2000). In contrast, some genotypes have only been found in buffalo, which suggests that these parasites have not yet adapted to cattle. Sampling of larger numbers of buffalo and cattle could show that there is no host preference. Evidence from literature indicates, however, that *T. taurotragi* was unable to transform buffalo-derived lymphocytes (Stagg *et al*. 1983), while certain *T. buffeli* strains derived from Asiatic buffalo were unable to infect cattle (Stewart *et al*. 1996). Similarly, no '*T.* sp. (buffalo)-like' profiles have been observed in ~2500 cattle samples tested with the real-time hybridization PCR assay during routine diagnostics (personal observation), suggesting that cattle are not infected by *T.* sp. (buffalo) or *T.* sp. (bougasvlei).

While some of the aims of the current study were to investigate the existence of variants, the possibility that single nucleotide polymorphism might be PCR errors should be considered. This could certainly be the case for *T. parva* variants no. 2, 4, 6 and 7, where only 1 variant sequence from a single animal was obtained, while the same samples also gave canonical or variant *T. parva* sequences. However, in the case of variants 1, 3 and 5, as well as the *T.* sp. (buffalo) variants, multiple sequences from a number of different animals were obtained, suggesting that these genotypes do occur within the local cattle and

https://doi.org/10.1017/S0031182011000187 Published online by Cambridge University Press

buffalo populations. In addition, all novel genotypes from the major groups, as well as variants 1, 3 and 5 were confirmed by re-cloning using a high fidelity polymerase. As such, the possibility that these genotypes observed is due to PCR error can be excluded. We could not confirm the sequences from variants 2, 4, 6 and 7 as we only obtained *T. parva* or variant sequences after re-cloning of the original samples and they could therefore be PCR artifacts. It is, however, likely that these sequences exist at low copy number and were outcompeted during cloning. Given the fact that variants 1, 3 and 5 have been confirmed to exist, there is no theoretical reason why other polymorphisms within the 18S hyper-variable region for *T. parva* would not exist as well, although they might not be prevalent within the population.

Samples with variants 1, 3 and 5 sequences all gave aberrant melting profiles that differ from that of the canonical *T. parva*, thus providing phenotypic evidence that these variants do exist. A precedent for the presence of the *T. parva* variant 1 genotype exists in the literature, as the sequence for the 18S SSU rRNA gene of *T. parva* in GenBank (Accession number: 2335189) possess an R at this position (Collins and Allsopp, 1999), indicating that it could either be a G as found in the canonical *T. parva* sequence, or an A as found for variant 1. One of the gold standard positive control animals (9288) used in the development of the real-time hybridization assay (Sibeko *et al*. 2008) gave a profile similar to that of variant 3. Direct sequencing of its 18S PCR product indicated a mixed infection with similar sequence trace peak intensities for the canonical as well as variant 3 genotypes (data not shown).

Previously, both cattle-derived and buffalo-derived parasites were shown to possess identical 18S SSU rRNA sequences (Collins and Allsopp, 1999). The discovery of variant sequences in the probe region for *T. parva* is therefore of considerable interest in regard to accurate diagnosis and the question is raised whether the variants observed are variant genotypes of *T. parva*, or closely related species, or subspecies. In this regard, originally 3 separate species were distinguished for the related disease syndromes known as East Coast fever (*T. parva*), Corridor disease (*T. lawrencei*) and January disease (*T. bovis*) (Theiler, 1904; Neitz *et al*. 1955; Neitz, 1957). The latter 2 were later synonymized with *T. parva*, although a trinomial system (*T. parva bovis, T. parva lawrencei* and *T. parva parva*) was proposed to differentiate between the different forms, but was subsequently dropped (Uilenberg, 1976; Lawrence, 1979; Norval *et al*. 1992). However, studies based on monoclonal antibodies and nucleic acid sequences continue to show that antigenic and sequence variation exist, keeping the notion of subspecies within the *T. parva* group viable (Conrad *et al*. 1987; Allsopp *et al*. 1989; Collins and Allsopp, 1999).

Analysis of the S5 gene suggests that the variants observed do not comprise subspecies. All *T. parva* positive control samples (KNP102, 9288, 9446 and 9468) originally came from South African buffalo, while variant 1 (32806) derived from a buffalo from the Niassa district in Mozambique. Variant 3 originated from a bovine (34431) that presumably contracted the parasites from buffalo-cattle contact. As such, the *T. parva* clade does not show any demarcation between buffalo or cattle-derived parasites for this gene and no support for subspecies classification of the variants observed in this study. This supports the notion that at genetic level no evidence for speciation can be detected between buffalo- and cattle-derived *T. parva* (Sibeko *et al*. 2010).

The S5 ribosomal gene exists as a single copy gene in the genomes of *T. parva* and *T. annulata*, so phylogenetic analysis will not be biased by pseudogenes or multi-gene copies (Gardner *et al*. 2005; Pain *et al*. 2005). The translated amino acid sequence for the fragment analysed was 100% identical for all *Theileria* species, with the majority of coding differences due to silent substitutions at the third codon position. This codon position is known to have greater numbers of parsimony-informative characters and higher overall rates of evolution compared to the first or second codon positions making it advantageous for phylogenetic inference (Simmons *et al*. 2006). Combined with the non-coding substitutions, the S5 gene seems to be a much better marker for species differentiation than the 18S hyper-variable region and it provided supporting evidence that *T. parva, T.* sp. (buffalo) and *T.* sp. (bougasvlei) are distinct species.

Cloned PCR amplicons for the 26 genotypes found in the current study were analysed with the real-time hybridization assay. No amplification at 705 nm was observed, except for *T. parva, T.* sp. (buffalo), *T.* sp. (bougasvlei) and their variants (results not shown). This indicated that no new genotype/species that amplify with the hybridization assay primer set was identified. In the case of some of the *T. parva* variants, the absence of a melting peak at 64 °C (640 nm) is abrogated by the presence of the shoulder peak, making identification of *T. parva* possible. Therefore, while numerous new genotypes/species were found in buffalo and cattle, those which affect the hybridization assay in terms of co-amplification during mixed-infections, remain *T.* sp. (buffalo) and *T.* sp. (bougasvlei). The within-species variation in the 18S SSU RNA gene observed for *T. parva* is also limited to single or double nucleotide polymorphisms that are detected by the current hybridization test. The use of this test in regions outside southern Africa should be useful for routine screening purposes. However, determination of 18S sequence variability should be considered when used as a diagnostic assay since it is possible that additional

https://doi.org/10.1017/S0031182011000187 Published online by Cambridge University Press

genotypes that amplify with the current primer set will be discovered in the future.

ACKNOWLEDGEMENTS

We thank Dr Roy Bengis, Kruger National Park for buffalo samples. The OVI sequencing facility is thanked for their support.

FINANCIAL SUPPORT

This research was funded by OVI, Project numbers OV15/08/1P01 and OV21/03/C148.

REFERENCES

Allsopp, B. A., Baylis, H. A., Allsopp, M. T. P. E., Cavalier-Smith, T., Bishop, R. P., Carrington, D. M., Sohanpal, B. and Spooner, P. (1993). Discrimination between six species of *Theileria* using oligonucleotide probes which detect small subunit ribosomal RNA sequences. *Parasitology* **107**, 157–165.

Allsopp, B., Carrington, M., Baylis, H., Sohal, S., Dolan, T. and Iams, K. (1989). Improved characterization of *Theileria parva* isolates using the polymerase chain reaction and oligonucleotide probes. *Molecular and Biochemical Parasitology* **35**, 137–147.

Allsopp, M. T., Cavalier-Smith, T., De Waal, D. T. and Allsopp, B. A. (1994). Phylogeny and evolution of the piroplasms. *Parasitology* **108**, 147–152.

Allsopp, M. T., Theron, J., Coetzee, M. L., Dunsterville, M. T. and Allsopp, B. A. (1999). The occurrence of *Theileria* and *Cowdria* parasites in African buffalo (*Syncerus caffer*) and their associated *Amblyomma hebraeum* ticks. *Onderstepoort Journal of Veterinary Research* **66**, 245–249.

Altay, K., Dumanli, N. and Aktas, M. (2007). Molecular identification, genetic diversity and distribution of *Theileria* and *Babesia* species infecting small ruminants. *Veterinary Parasitology* **147**, 161–165.

Barnett, S. F. (1977). *Theileria*. In *Parasitic Protozoa, Vol. IV* (ed. Kreier, J. P.), pp. 77–113. Academic Press, New York, USA.

Bhoora, R., Franssen, L., Oosthuizen, M. C., Guthrie, A. J., Zweygarth, E., Penzhorn, B. L., Jongejan, F. and Collins, N. E. (2009). Sequence heterogeneity in the 18S rRNA gene within *Theileria equi* and *Babesia caballi* from horses in South Africa. *Veterinary Parasitology* **159**, 112–120.

Brocklesby, D. W. (1962). *Cytauxzoon taurotragi* Martin and Brocklesby, 1960, a piroplasm of the eland (*Taurotragus oryx pattersonianus* Lydekker, 1906). *Research in Veterinary Science* **3**, 334–344.

Chae, J. S., Allsopp, B. A., Waghela, S. D., Park, J. H., Kakuda, T., Sugimoto, C., Allsopp, M. T., Wagner, G. G. and Holman, P. J. (1999). A study of the systematics of *Theileria* spp. based upon small-subunit ribosomal RNA gene sequences. *Parasitology Research* **85**, 877–883.

Chansiri, K., Kawazu, S., Kamio, T., Terada, Y., Fujisaki, K., Philippe, H. and Sarataphan, N. (1999). Molecular phylogenetic studies on *Theileria* parasites based on small subunit ribosomal RNA gene sequences. *Veterinary Parasitology* **83**, 99–105.

Collins, N. E. and Allsopp, B. A. (1999). *Theileria parva* ribosomal internal transcribed spacer sequences exhibit extensive polymorphism and mosaic evolution: application to the characterization of parasites from cattle and buffalo. *Parasitology* **118**, 541–551.

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

Conrad, P. A., Stagg, D. A., Grootenhuis, J. G., Irvin, A. D., Newson, J., Njamunggeh, R. E., Rossiter, P. B. and Young, A. S. (1987). Isolation of *Theileria* parasites from African buffalo (*Syncerus caffer*) and characterization with anti-schizont monoclonal antibodies. *Parasitology* **94**, 413–423.

Coyne, J. A. and Orr, H. A. (2004). *Speciation*. Sinauer Associates, Inc. Sunderland, MA, USA.

Criado-Fornelio, A., Gónzalez-del-Río, M. A., Buling-Saraña, A. and Barba-Carretero, J. C. (2004). The "expanding universe" of piroplasms. *Veterinary Parasitology* **119**, 337–345.

Criado, A., Martinez, J., Buling, A., Barba, J. C., Merino, S., Jefferies, R. and Irwin, P. J. (2006). New data on epizootiology and genetics of piroplasms based on sequences of small ribosomal subunit and cytochrome b genes. *Veterinary Parasitology* **142**, 238–247.

De Vos, A. J. and Roos, J. A. (1981). Observations on the transmission of *Theileria mutans* in South Africa. *Onderstepoort Journal of Veterinary Research* **48**, 1–6.

Dschunkowsky, E. and Luhs, J. (1904). Die piroplasmosen der rinder. *Centralblatt für Bakteriologie und Parasitenkunde und Infektion und Hygiene* **35**, 486–493.

Eickbush, T. H. and Eickbush, D. G. (2007). Finely orchestrated movements: evolution of the ribosomal RNA genes. *Genetics* **175**, 477–485.

Gardner, M. J., Bishop, R., Shah, T., de Villiers, E. P., Carlton, J. M., Hall, N., Ren, Q., Paulsen, I. T., Pain, A., Berriman, M., Wilson, R. J., Sato, S., Ralph, S. A., Mann, D. J., Xiong, Z., Shallom, S. J., Weidman, J., Jiang, L., Lynn, J., Weaver, B., Shoaibi, A., Domingo, A. R., Wasawo, D., Crabtree, J., Wortman, J. R., Haas, B., Angiuoli, S. V., Creasy, T. H., Lu, C., Suh, B., Silva, J. C., Utterback, T. R., Feldblyum, T. V., Pertea, M., Allen, J., Nierman, W. C., Taracha, E. L., Salzberg, S. L., White, O. R., Fitzhugh, H. A., Morzaria, S., Venter, J. C., Fraser, C. M. and Nene, V. (2005). Genome sequence of *Theileria parva*, a bovine pathogen that transforms lymphocytes. *Science* **309**, 134–137.

Gubbels, J. M., de Vos, A. P., van der Weide, M., Viseras, J., Schouls, L. M., de Vries, E. and Jongejan, F. (1999). Simultaneous detection of bovine *Theileria* and *Babesia* species by reverse line blot hybridization. *Journal of Clinical Microbiology* **37**, 1782–1789.

Gubbels, M. J., Hong, Y., van der Weide, M., Qi, B., Nijman, I. J., Guangyuan, L. and Jongejan, F. (2000). Molecular characterisation of the *Theileria buffeli/orientalis* group. *International Journal for Parasitology* **30**, 943–952.

Gunderson, J. H., Sogin, M. L., Wollett, G., Hollingdale, M., de la Cruz, V. F., Waters, A. P. and McCutchan, T. F. (1987). Structurally distinct, stage-specific ribosomes occur in Plasmodium. *Science* **238**, 933–937.

Jeanmougin, F., Thompson, J. D., Gouy, M., Higgins, D. G. and Gibson, T. J. (1998). Multiple sequence alignment with Clustal X. *Trends in Biochemical Science* **23**, 403–405.

Katoh, K. and Toh, H. (2008). Improved accuracy of multiple ncRNA alignment by incorporating structural information into a MAFFT-based framework. *BMC Bioinformatics* **9**, 212.

Koch, R. (1903). Interim report on Rhodesian redwater or African Coast Fever. *The Journal of Comparative Pathology and Therapeutics* **16**, 273–280.

Lau, A. O. (2009). An overview of the *Babesia, Plasmodium* and *Theileria* genomes: A comparative perspective. *Molecular and Biochemical Parasitology* **164**, 1–8.

Laughery, J. M., Lau, A. O., White, S. N., Howell, J. M. and Suarez, C. E. (2009). *Babesia bovis*: Transcriptional analysis of rRNA gene unit expression. *Experimental Parasitology* **123**, 45–50.

Lawrence, D. A. (1935). Report of the Director of Veterinary Research, Southern Rhodesia, for the year 1934. Rhodesian Printing and Publishing Co., Salisbury, Rhodesia.

Lawrence, J. A. (1979). The differential diagnosis of the bovine theilerias of southern Africa. *Journal of the South African Veterinary Association* **50**, 311–313.

Levine, N. D., Corliss, J. O., Cox, F. E., Deroux, G., Grain, J., Honigberg, B. M., Leedale, G. F., Loeblich, A. R. 3rd, Lom, J., Lynn, D., Merinfeld, E. G., Page, F. C., Poljansky, G., Sprague, V., Vavra, J. and Wallace, F. G. (1980). A newly revised classification of the protozoa. *Journal of Protozoology* **27**, 37–58.

Martin, H. and Brocklesby, D. W. (1960). A new parasite of the Eland. *Veterinary Record* **72**, 331–332.

Matjila, P. T., Leisewitz, A. L., Jongejan, F., Bertschinger, H. J. and Penzhorn, B. L. (2008). Molecular detection of *Babesia rossi* and *Hepatozoon* sp. in African wild dogs (*Lycaon pictus*) in South Africa. *Veterinary Parasitology* **157**, 123–127.

Matjila, P. T., Penzhorn, B. L., Bekker, C. P., Nijhof, A. M. and Jongejan, F. (2004). Confirmation of occurrence of *Babesia canis vogeli* in domestic dogs in South Africa. *Veterinary Parasitology* **122**, 119–125.

Nei, M. and Kumar, S. (2000). *Molecular Evolution and Phylogenetics*. Oxford University Press, New York, USA.

Neitz, W. O. (1957). Theileriosis, gonderiosis and cytauxzoonoses: a review. *Onderstepoort Journal of Veterinary Research* **27**, 275–430.

Neitz, W. O., Canham, A. S. and Kluge, E. (1955). Corridor disease: A fatal form of bovine theileriosis encountered in Zululand. *Journal of the South African Veterinary Association* **26**, 79–87.

Nijhof, A. M., Penzhorn, B. L., Lynen, G., Mollel, J. O., Morkel, P., Bekker, C. P. and Jongejan, F. (2003). *Babesia bicornis* sp. nov. and *Theileria bicornis* sp. nov.: tick-borne parasites associated with mortality in the black rhinoceros (*Diceros bicornis*). *Journal of Clinical Microbiology* **41**, 2249–2254.

Nijhof, A. M., Pillay, V., Steyl, J., Prozesky, L., Stoltsz, W. H., Lawrence, J. A., Penzhorn, B. L. and Jongejan, F. (2005). Molecular characterization of *Theileria* species associated with mortality in four species of African antelopes. *Journal of Clinical Microbiology* **43**, 5907–5911.

Norval, R. A., Lawrence, J. A., Young, A. S., Perry, B. D., Dolan, T. T. and Scott, J. (1991). *Theileria parva*: influence of vector, parasite and host relationships on the epidemiology of theileriosis in southern Africa. *Parasitology* **102**, 347–356.

Norval, R. A. I., Perry, B. D., Young, A. S. (1992). *The Epidemiology of Theileriosis in Africa*. Academic Press, London, UK.

Oosthuizen, M. C., Zweygarth, E., Collins, N. E., Troskie, M. and Penzhorn, B. L. (2008). Identification of a novel *Babesia* sp. from a sable antelope (*Hippotragus niger* Harris, 1838). *Journal of Clinical Microbiology* **46**, 2247–2251.

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

Potgieter, F. T., Stoltsz, W. H., Blouin, E. F. and Roos, J. A. (1988). Corridor disease in South Africa: a review of the current status. *Journal of the South African Veterinary Association* **59**, 155–160.

Reichard, M. V., Van Den Bussche, R. A., Meinkoth, J. H., Hoover, J. P. and Kocan, A. A. (2005). A new species of *Cytauxzoon* from Pallas' cats caught in Mongolia and comments on the systematics and taxonomy of piroplasmids. *Journal of Parasitology* **91**, 420–426.

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

Sibeko, K. P., Oosthuizen, M. C., Collins, N. E., Geysen, D., Rambritch, N. E., Latif, A. A., Groeneveld, H. T., Potgieter, F. T. and Coetzer, J. A. (2008). Development and evaluation of a real-time polymerase chain reaction test for the detection of *Theileria parva* infections in Cape buffalo (*Syncerus caffer*) and cattle. *Veterinary Parasitology* **155**, 37–48.

Simmons, M. P., Zhang, L. B., Webb, C. T. and Reeves, A. (2006). How can third codon positions outperform first and second codon positions in phylogenetic inference? An empirical example from the seed plants. *Systematic Biology* **55**, 245–258.

Stagg, D. A., Young, A. S., Leitch, B. L., Grootenhuis, J. G. and Dolan, T. T. (1983). Infection of mammalian cells with *Theileria* species. *Parasitology* **86**, 243–254.

Stewart, N. P., De Vos, A. J., Shiels, I. and McGregor, W. (1987). The experimental transmission of *Theileria buffeli* of cattle in Australia by *Haemaphysalis humerosa*. *Australian Veterinary Journal* **64**, 81–83.

Stewart, N. P., Uilenberg, G. and De Vos, A. J. (1996). Review of Australian species of *Theileria*, with special reference to *Theileria buffeli* of cattle. *Tropical Animal Health Production* **28**, 81–90.

Tamura, K., Dudley, J., Nei, M. and Kumar, S. (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Molecular Biology and Evolution* **24**, 1596–1599.

Theiler, A. (1904). Rhodesian tick fever. *Transvaal Agricultural Journal* **2**, 421–438.

Theiler, A. (1906). Piroplasma mutans (n. spec) of South African cattle. *The Journal of Comparative Pathology and Therapeutics* **19**, 292–300.

Tonetti, N., Berggoetz, M., Rühle, C., Pretorius, A. M. and Gern, L. (2009). Ticks and tick-borne pathogens from wildlife in the Free State Province, South Africa. *Journal of Wildlife Diseases* **45**, 437–446.

Uilenberg, G. (1964). *Haemotoxenus veliferus*, n. g., n. sp., parasite incertae sedis du sang de bovins a Madagascar. *Revue d'élevage et de médecine vétérinaire des pays tropicaux* **17**, 655–662.

Uilenberg, G. (1976). Tick-borne livestock diseases and their vectors. 2. Epizootiology of tick-borne diseases. *World Animal Review* **17**, 8–15.

Uilenberg, G. (1981). Theilerial species of domestic livestocks. In *Advances in the Control of Theileriosis* (ed. Irvin, A. D., Cunningham, M. P. and Young, A. S.), pp. 4–37. Martinus Nijhoff Publishers, The Hague, The Netherlands.

Uilenberg, G. (2006). *Babesia* – A historical perspective. *Veterinary Parasitology* **138**, 3–10.

Uilenberg, G. and Schreuder, B. E. (1976). Studies on Theileriidae (Sporozoa) in Tanzania. I. Tick transmission of *Haematoxenus veliferus*. *Tropenmedizin und Parasitologie* **27**, 106–111.

Young, A. S. (1981). The epidemiology of theileriosis in East Africa. In *Advances in the Control of Theileriosis* (ed. Irvin, A. D., Cunningham, M. P. and Young, A. S.), pp. 38–55. Martinus Nijhoff Publishers, The Hague, The Netherlands.

Young, A. S., Brown, C. G., Burridge, M. J., Grootenhuis, J. G., Kanhai, G. K., Purnell, R. E. and Stagg, D. A. (1978*a*). The incidence of theilerial parasites in East African buffalo (*Syncerus caffer*). *Tropenmedizin und Parasitologie* **29**, 281–288.

Young, A. S., Grootenhuis, J. G., Kimber, C. D., Kanhai, G. K. and Stagg, D. A. (1977). Isolation of a *Theileria* species from Eland (*Taurotragus oryx*) infective for cattle. *Tropenmedizin und Parasitologie* **27**, 185–194.

Young, A. S., Purnell, R. E., Payne, R. C., Brown, C. G. and Kanhai, G. K. (1978*b*). Studies on the transmission and course of infection of a Kenyan strain of *Theileria mutans*. *Parasitology* **76**, 99–115.

Zweygarth, E., Koekemoer, O., Josemans, A. I., Rambritch, N., Pienaar, R., Putterill, J., Latif, A. and Potgieter, F. T. (2009). *Theileria*-infected cell line from an African buffalo (*Syncerus caffer*). *Parasitology Research* **105**, 579–581.

https://doi.org/10.1017/S0031182011000187 Published online by Cambridge University Press