Received: 27 March 2019 | Revised: 28 June 2019 | Accepted: 22 July 2019

DOI: 10.1111/tbed.13311

SUPPLEMENT ARTICLE

Transboundary and Emerging Diseases WILEY

Antigen gene and variable number tandem repeat (VNTR) diversity in *Theileria parva* parasites from Ankole cattle in south-western Uganda: Evidence for conservation in antigen gene sequences combined with extensive polymorphism at VNTR loci

Anne Nanteza[1] | Isaiah Obara[2] | Paul Kasaija[1] | Elisa Mwega[3] | Fredrick Kabi[4] | Diaeldin A. Salih[5] | Moses Njahira[6] | Joyce Njuguna[6] | David Odongo[7] | Richard P. Bishop[8] | Rob A. Skilton[6] | Jabbar Ahmed[1] | Peter-Henning Clausen[2] | George W. Lubega[1]

[1]College of Veterinary Medicine, Animal Resources and Biosecurity, Makerere University, Kampala, Uganda

[2]Institute for Parasitology and Tropical Veterinary Medicine, Freie Universität Berlin, Berlin, Germany

[3]Faculty of Veterinary Medicine, Sokoine University of Agriculture, Morogoro, Tanzania

[4]College of Natural Sciences, Makerere University, Kampala, Uganda

[5]Central Veterinary Research Laboratories, Khartoum, Sudan

[6]Biosciences East and Central Africa (BecA), Nairobi, Kenya

[7]School of Biological Sciences, University of Nairobi, Nairobi, Kenya

[8]Department of Veterinary Microbiology and Pathology, Washington State University, Pullman, WA, USA

Correspondence
Anne Nanteza, College of Veterinary Medicine, Animal Resources and Biosecurity, Makerere University, Kampala, Uganda.
Email: nantezaa@covab.mak.ac.ug

Funding information
DFG; International Livestock Research Institute (ILRI)

Abstract

Theileria parva is a tick-transmitted apicomplexan protozoan parasite that infects lymphocytes of cattle and African Cape buffalo (*Syncerus caffer*), causing a frequently fatal disease of cattle in eastern, central and southern Africa. A live vaccination procedure, known as infection and treatment method (ITM), the most frequently used version of which comprises the Muguga, Serengeti-transformed and Kiambu 5 stocks of *T. parva*, delivered as a trivalent cocktail, is generally effective. However, it does not always induce 100% protection against heterologous parasite challenge. Knowledge of the genetic diversity of *T. parva* in target cattle populations is therefore important prior to extensive vaccine deployment. This study investigated the extent of genetic diversity within *T. parva* field isolates derived from Ankole (*Bos taurus*) cattle in south-western Uganda using 14 variable number tandem repeat (VNTR) satellite loci and the sequences of two antigen-encoding genes that are targets of CD8+T-cell responses induced by ITM, designated Tp1 and Tp2. The findings revealed a *T. parva* prevalence of 51% confirming endemicity of the parasite in south-western Uganda. Cattle-derived *T. parva* VNTR genotypes revealed a high degree of polymorphism. However, all of the *T. parva* Tp1 and Tp2 alleles identified in this study have been reported previously, indicating that they are widespread geographically in East Africa and highly conserved.

KEYWORDS

Ankole cattle, East Coast fever, genetic diversity, MHC1, *Theileria parva*, Tp1, Tp2, VNTR

Anne Nanteza and Isaiah Obara made equal contributions to the manuscript

1 | INTRODUCTION

East Coast fever (ECF), caused by *Theileria parva*, is an economically important cattle disease in eastern, central and southern Africa (Norval et al., 1992), due to mortality induced especially in calves including Zebu (Thumbi et al., 2013), and adult exotic *Bos taurus* cattle. At present, the only vaccination option for control of ECF involves inoculation with a potentially lethal dose of live parasites, and simultaneous oxytetracycline treatment, an approach known as the 'Infection and Treatment method' (ITM). Although several *T. parva* isolates have been evaluated for their ability to protect against a range of genetically distinct and geographically dispersed isolates, the ITM formulation that has been most widely used (Di Giulio, Lynen, Morzaria, Oura, & Bishop, 2009) is the trivalent cocktail that contains the Muguga, Kiambu 5 and Serengeti-transformed stocks (Radley et al., 1975).

Despite repeated demonstration of efficacy in the laboratory and the field, ITM vaccine delivery and uptake has not been rapid. One major concern is that when the Muguga cocktail-based ITM vaccine is first introduced into a new geographical location, *T. parva* genotypes from the vaccine that differ from the locally circulating populations and are therefore 'foreign' can be introduced into local tick populations (Bishop et al., 1992; Oura, Asiimwe, Weir, Lubega, & Tait, 2005). This concern is one reason why the ITM vaccine is yet to be deployed on a large scale in Uganda, despite the fact that ECF is arguably, the most important cattle health problem in most parts of Uganda, and that livestock frequently share grazing areas with the wildlife reservoir of *T. parva*, the African cape buffalo (*Syncerus caffer*).

The current study sought to address the knowledge gaps relating to parasite polymorphisms both at the VNTR satellite loci (Oura, Odongo, Lubega, Spooner, & Tait, 2003) and immunodominant CD8+T-cell target antigen loci (Graham et al., 2006) that might be useful for predicting the outcome of live vaccination using ITM in Uganda. The CD8+T-cell target antigens are interesting because of evidence from adoptive transfer experiments that the immunity induced in cattle by ITM is mediated by parasite-specific CD8+T cells (McKeever et al., 1994).

Although they are less susceptible than exotic *Bos taurus* breeds, the African taurine Ankole form the vast majority of small and medium-sized herds in south-western Uganda where ECF is a major problem (Muhanguzi, Matovu, & Waiswa, 2010). A recent study comparing the predicted peptide-binding specificities of Ankole and Holstein class I MHC genes revealed a largely distinct peptide-binding repertoire (Obara et al., 2016). It is therefore important to document polymorphisms present within CD8+T-cell target antigens in *T. parva* isolates derived from Ankole cattle and compare with those present in the stocks that comprise the Muguga cocktail-based ITM vaccine. Additionally, the availability of the *T. parva* genome sequence has allowed identification of a genome-wide panel of satellite markers for high-resolution genotyping of *T. parva* populations (Oura et al., 2003). The number of alleles per micro- or mini-satellite locus ranges from 3 to 8, indicating a high level of genetic diversity both between and within geographically distinct *T. parva* parasite isolates (Oura et al., 2003).

In this study, we investigated *T. parva* prevalence and genotypic diversity in Ankole cattle from south-western Uganda based on 14 VNTR (satellite) and two antigen gene sequences (Tp1 and Tp2) in field parasite populations from cattle that co-graze with buffalo. These CD8+T-cell target antigens are immunodominant in cattle with specific MHC Class I alleles. The goal of this research was to provide baseline data on parasite genotypes from an ECF endemic area prior to potential future ITM deployment in this population.

2 | MATERIALS AND METHODS

2.1 | Ethics Statement

The protocols for cattle restraint and blood collection were approved by Makerere University, College of Veterinary Medicine, Animal Resources and Biosecurity (CoVAB) Institutional Animal Care and Use Ethical Committee. The approval reference number was sbls.an.2012.

2.2 | Study site and animals

Three hundred and ninety (390) Ankole cattle of variable age and sex maintained under open grazing systems and often sharing pastures with cape buffalo were sampled from Kiruhura district, south-western Uganda (Figure 1). The region is known to be endemic for *T. parva* but the majority of cattle farmers cannot afford routine use of available ECF control methods. This area is also climatically suitable for *Rhipicephalus appendiculatus* ticks that transmit *T. parva*, which can be found up to 8,000 feet above sea level in areas with an annual rainfall of over 20 inches (500 mm). The study area is warm, humid and well vegetated providing a suitable microclimate for the tick (Coetzer & Tustin, 2004). The farms or individual animals sampled were either suspected or reported to have experienced theileriosis. Sampling was performed between April and May, 2012 during the wet season which coincides with moderate to high vector tick infestations.

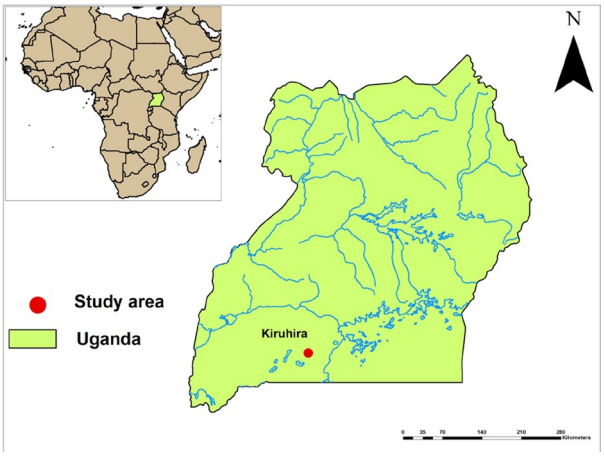

FIGURE 1 Map of Uganda showing the study site. Kiruhura district in south-western Uganda is indicated with a red circle

2.3 | Sample collection and preparation

Whole blood samples were collected from each animal into EDTA vacutainer tubes (Fisher Scientific) by jugular venipuncture. Blood was transported in iceboxes to the district veterinary laboratory, spotted onto FTA cards in duplicate (120 µl per spot), air-dried overnight at room temperature (27°C), labelled and stored individually in sealed envelopes containing desiccant at room temperature. DNA was extracted from punched dried blood spots using the PureLink™ Genomic DNA Mini kit (Invitrogen). The DNA concentration and purity were determined using UV spectrophotometry (Nanodrop).

2.4 | Screening cattle genomic DNA for *T. parva* by PCR

A *T. parva*-specific nested p104 PCR assay was used to screen for *T. parva*-positive samples from cattle blood genomic DNA, as described previously (Odongo, Sunter, Kiara, Skilton, & Bishop, 2010). The PCR amplification was performed using Bioneer tubes containing a lyophilized mastermix (Bioneer, Korea). PCR primers, water and template DNA were added to the tubes containing the PCR master mix resulting in a final reaction volume of 25 µl. The PCR amplifications were performed using a programmable thermal cycler from MJ Research (Watertown, MA, USA). The nested PCR amplicons (278 bp) were detected on 2% GelRed (Biotium, USA) pre-stained 2% agarose gels in 0.5× Tris-acetate-EDTA (TAE) buffer and viewed under UV light box. In addition to the field samples, one DNA sample from cattle-derived *T. parva* stock, Muguga from ILRI, Nairobi, Kenya, was included as a reference (positive) control sample and a negative control (distilled water) was also included. Genomic DNA from positive *T. parva* samples was selected for further molecular characterization of the different genotypes using mini- and micro-satellite markers [Table 1] (Oura et al., 2003) and Tp1 and Tp2 antigen gene sequences (Pelle et al., 2011).

2.5 | *Theileria parva* genotyping using mini- and micro-satellite markers

Fourteen VNTRs comprising, 3 micro- and 11 mini-satellites (Oura et al., 2003, 2005) that are widely distributed across all 4 chromosomes were used for *T. parva* genotyping (Table 1). The selection of the VNTRs was based on sensitivity and reproducibility (Oura et al., 2005). To increase the sensitivity of the PCR, a nested assay was performed. The forward primer of each nested primer pair was labelled at the 5′ end with one of the four standard dyes; 6-FAM (Blue), VIC (Green), NED (Yellow) and PET (Red) (Bioneer, Korea), for detection on an ABI Genetic Analyser. Two hundred *T. parva*-positive genomic DNA samples were analysed. The PCRs contained genomic DNA (20 ng), 1× PCR buffer, 200 µM of each dNTP, 0.2 µM of each forward and reverse primers and 5U of Taq polymerase (Promega) in a final volume of 10 µl. The PCR cycling parameters were as follows: denaturation at 94°C for 5 min, 30 cycles of denaturation at 94°C for 30 s, annealing at 60°C for 1 min and extension at 72°C for 1 min followed by a final extension at 72°C for 20 min. For the second round PCR, 1 µl of PCR product was used in a final volume of 10 µl and the PCR components and conditions were similar to those of the first PCR. The quality of PCR amplicons (2 µl) was assessed by agarose gel electrophoresis.

Genotyping used capillary flow fluorescent electrophoresis on an ABI 3730 Genetic Analyzer. The samples were prepared as follows: a reaction mix containing PCR product (0.5–2 µl) was added into each well of a 96-well plate containing 8 µl of a mixture of Hi-Di Formamide and Genescan 500 LIZ size standard (ABI-USA). Rapid denaturation was performed in a thermocycler at 95°C for 5 min, followed by rapid chilling of the plate on ice for 2 min. The amplicons were processed and analysed on the ABI 3730 Genetic Analyser (Applied Biosystems-USA) (SegoLip Sequencing Unit, ILRI Nairobi, Kenya). Scoring was implemented using the GeneMapper programme (GeneMapper® ID software V3.2) supplied by the manufacturer. Amplicons with maximum peak height were scored, and a predominant peak was defined as that with the largest area under the curve with the prescribed base pair range. All data generated from GeneMapper was re-sized by the Allelobin software based on the consensus sequence of the locus. A predominant allele at each locus was used to generate allele frequency data and multi-locus genotypes (MLGs) in Excel. The GenePop v 4.1.3 software was used to determine the allele frequencies of the 14 SSR markers in the 200 *T. parva* field isolates. Principal Component Analysis (PCoA) was based on pairwise Fst data (Figure 2).

2.6 | Analysis of *T. parva* Tp1 and Tp2 gene sequence variation and molecular evolution

Nested PCR amplification was performed on 200 *T. parva* p104-positive genomic DNA samples to amplify part of the gene encoding Tp1 and most of the gene encoding Tp2. A 432 bp region within Tp1 located in the centre of the gene and extending from nucleotides 523 to 954 of the reference Tp1 sequence was amplified. This region encodes 144 amino

TABLE 1 The *Theileria parva* micro-and mini-satellites used in the study and their distribution across the four Chromosomes (Oura et al., 2003)

Chromosome 1	Chromosome 2	Chromosome 3	Chromosome 4
ms 2	ms 7	MS 21	MS 33
ms 5	MS 16	MS 25	MS 34
MS 3	MS 19	MS 27	MS 40
MS 5			
MS 7			

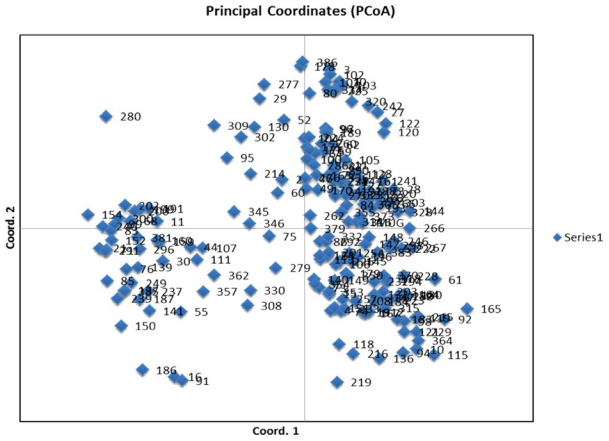

FIGURE 2 Principal Component analysis (PCoA) applied to VNTR alleles within *Theileria parva* populations present in Ankole cattle from Kiruhura district, south-western Uganda. The analysis was based on genotype data from 14 VNTR (mini- and micro-satellite) loci. Each blue square represents an individual parasite allele

acids comprising 26.5% of the 543-residue Tp1 protein (Accession number XP_762973) (Pelle et al., 2011). The full-length Tp2 gene sequence that is predicted to encode a protein of 174 amino acids was also amplified (Pelle et al., 2011). Specific forward and reverse primers were used for PCR amplification of Tp1 and Tp2 gene sequences (GenBank XM_757880.1) and (GenBank XM_760490.1), respectively. First round Tp1 primers were Tp1-F: ATGGCCACTTCAATTGCATTTGCC and Tp1-R: TTAAATGAAATATTTATGAGCTTC with a product size of 432 bp. Inner forward primers were as follows: Tp1_Forward_inner: TGCATTTGCCGCTGATCCTGGATTCTG and Tp1_Reverse_inner: TGAGCTTCGTATACACCCTCGTATTCG with a product size of 405 bp (Salih et al., 2017). The primers used to amplify the Tp2 gene were Tp2_F: ATGAAATTGGCCGCCAGATTA and Tp2_R: CTATGAAGTGCCGGAGGCTTC which produced an amplicon of 525 bp from the primary PCR. Secondary amplification primers were, Tp2_Forward_inner: ATTAGCCTTTACTTTATTATTTWCATTYTAC and Tp2_Reverse_inner: CTATGAAGCGCCGGAGGCTTCTCCT which amplify a 504 bp PCR product (Salih et al., 2017).

Primary and secondary PCR amplifications for Tp1 were performed in a total volume of 20 µl containing 10 pmol of forward and reverse primer and 2 µl of genomic DNA (20ŋg/µl) template added to a lyophilized pellet (Bioneer PCR-PreMix-Korea). The second PCR contained 1 µl of primary PCR product as template and 10pmol of each primer. The cycling conditions for Tp1 primary PCR were as follows: initial denaturation at 95°C for 5 min, 40 cycles of denaturation at 94°C for 30 s, annealing at 50°C for 1 min and extension at 72°C for 1 min, final extension at 72°C for 7 min. The cycling profile for secondary PCR was as for primary PCR except that annealing was at 55°C and the number of cycles was reduced to 30. The conditions for nested PCR for Tp2 were essentially similar to those for Tp1, except that initial denaturation at 95°C was for 11 min (hot start Taq polymerase was used) and annealing for the secondary PCR was at 58°C.

The PCR products for both Tp1 and Tp2 were analysed by electrophoresis in gel red-stained 1.5% agarose gels. All the positive products (10 µl from each PCR product) were purified using PCR Purification Kit (Invitrogen) and submitted to the SegoLip Sequencing Unit at BecA-ILRI hub, Nairobi-Kenya for sequencing using ABI 3730 Capillary Sequencer (Applied Biosystems-USA). Thereafter, sequences were edited and translated into protein using CLC Bio DNA Workbench Version 6.7.1 (http://www.clcbio.com). Multiple alignment of Tp1 and Tp2 sequences together with Muguga vaccine isolate (XM_760490) was performed using cluster (http://www.genome.jp/tools.bin/clusterw) (Thompson, Higgins, & Gibson, 1994). Genetic differentiation and gene flow estimates were generated using DnaSP v5 http://www.ub.edu/dnasp/ (Librado & Rozas, 2009).

2.7 | Nucleotide and codon evolution at the immunodominant Tp1 locus

An initial neighbour-joining tree was constructed based on the Jukes and Cantor model using PAUP* (phylogenetic analysis using parsimony; Swofford, 2003). We then assessed the fit of the data to 56 models of nucleotide substitution, including those that allow for a proportion of invariable sites, and recorded the following: negative log likelihood (-lnL), number of estimated parameters (K) and the information Criterion (IC). Model fit was evaluated using the program Model test version 3.7 with run settings based on standard Akaike information criterion (AIC), and all models were included in model-averaging calculations. The nucleotide substitution model that provided the best fit for the data set as selected by ModelTest was used to construct a maximum likelihood tree in PAUP* with heuristic search proceeding by rearrangements of the nearest neighbour interchange (NNI).

Likelihood settings from the best fit nucleotide substitution model were then used to investigate whether there is any evidence for positively selected codons in the Tp1 data set. In a codon by codon analysis, we compared the fit of the nearly neutral model (M1) and the model that allows for positive selection (M2) to the Tp1 data set. Both models allow for purifying selection and neutral evolution, but M2 allows for a further category of codons evolving under positive selection. Model fit was evaluated based on the AIC scores.

2.8 | Sequence logo of *T. parva* Tp1 region

We created a Kullback–Leibler logo based on a multiple alignment of the predicted amino acid sequences of the isolates genotyped based on *T. parva* Tp1 and Tp2 regions described herein using the program Seq2 Logo version 1.2 (Thomsen & Nielsen, 2012). For comparison, a separate logo was created for the same Tp1 epitope from all Tp1 sequences publicly available in GenBank.

3 | RESULTS

3.1 | *Theileria parva prevalence* and genotypic diversity assessed using mini- and micro-satellite markers

Screening of 390 cattle blood samples using the nested *T. parva* p104 gene primers (Odongo et al., 2010) revealed a total of 200

(51%) positive animals. The majority of the PCR amplicons based on the mini- and micro-satellite markers revealed multiple (2–4) bands on agarose gels (data not shown) indicating mixed infections with multiple parasite genotypes. The *T. parva* population from cattle in south-western Uganda showed a high degree of genetic diversity and were separated into two major subgroups using PCoA. However, the two subgroups were not geographically separate as all the samples came from the same district (Figure 2). Moreover, this clustering is supported by the Tp2 sequence variation data as exemplified by sample 16 in Figure 5.

3.2 | Analysis of *T. parva* Tp1 antigen gene alleles

A total of 115 sequences were generated from the 200 *T. parva* isolates and different Tp1 alleles were identified based on sequence variations across the partial gene segment (data not shown). The variant alleles in the 115 Ugandan isolates, from which a sequence was obtained, outside the mapped epitope have all been described previously (Pelle et al., 2011).

Three epitope variants described in previous studies (Pelle et al., 2011) were identified in the CD8 + T-cell epitope in the Kiruhura isolates (Figure 3). One of the Tp1 epitopes (VGYPKVKEEML) that was present in the majority of the field isolates from Uganda was identical to that present in the Muguga reference genotype and also Serengeti-transformed (Figure 3). The other two epitope variants differed only in residues 10 and 11. Analysis by neighbour-joining tree clustered Tp1 sequences into two clades with the majority of the samples (75.7%) clustering in the same clade as *T. parva* Muguga (data not shown).

We generated a Kullback–Leibler logo to provide a more useful and interpretable classification of the relative frequencies of residues at every position along the Tp1 epitope in the isolates from south-western Uganda that were genotyped (Figure 4, panel A) and compared the logo with that generated from publicly available Tp1 sequences (Figure 4 panel B). As shown in Figure 4, the level of conservation at particular epitope positions as denoted by the residue heights is identical. Similarly, the residue frequencies at particular positions as shown by the relative residue heights are also identical.

As regards molecular evolution of the Tp1 locus, AIC values showed that the Kimura 3p unequal-frequencies (K81uf+I; Kimura, 1980) provided the best fit to the Ugandan data set, with the proportion of invariable sites (I) being 0.9456. There was no difference in the likelihood score for the near-neutral model, M1 (lnL −534.550988) and the positive selection model M2 (lnL = −534.550987). However, M1 had two fewer free parameters (K) compared with M2. The near-neutral model therefore provided a marginally better fit to the current Tp1 data set based on the AIC scores.

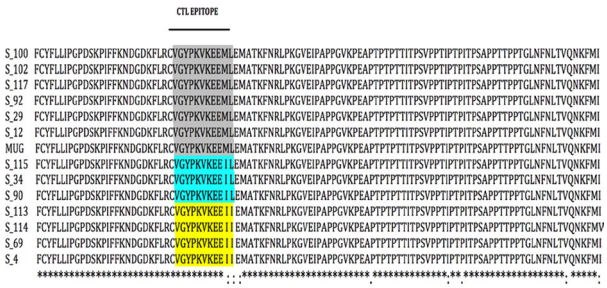

FIGURE 3 Alignment of 14 representative *Theileria parva* Tp1 amino acid sequences derived from 115 *T. parva* field isolates. S_ represents the Ugandan *T. parva* field isolate and MUG represents the *T. parva* Muguga stock reference strain. The three epitope variants are shaded in grey, blue and yellow

3.3 | Analysis of *T. parva* Tp2 antigen gene alleles

The full-length Tp2 ORF gene sequence that encodes 174 amino acids was PCR amplified and sequenced from 30 *T. parva* isolates. Amplification efficiency of Tp2 from the p104-positive cattle was relatively low, perhaps due to lack of primer conservation. All the Ugandan Tp2 variants noted both within and outside the six previously defined epitopes have been reported previously (Pelle et al., 2011). The 6 mapped CD8+epitopes in the Tp2 protein of the reference Muguga genotype were also the most commonly observed variants among the Ugandan isolates. This was similar to the data for the Tp1 epitope. Seventeen representative samples were used to construct a neighbour-joining phylogenetic tree (Figure 5). The

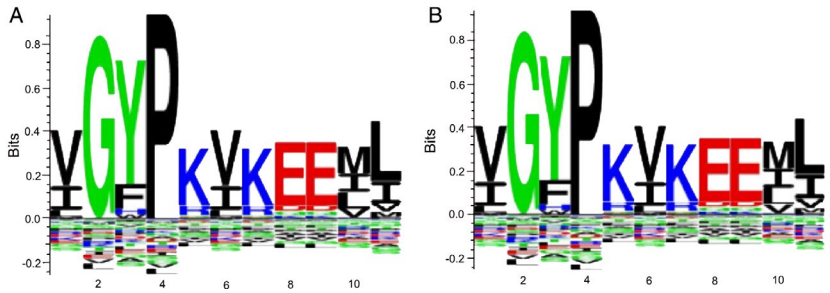

FIGURE 4 Kullback–Leibler logo for sequences corresponding to CTL epitopes within the *Theileria parva* Tp1 antigen based on a multiple alignment of the predicted amino acid (aa) sequences from the Ugandan isolates (a) and published Tp1 sequences (b). Enriched aa is shown on the positive *y*-axis and depleted underrepresented aa on the negative *y*-axis. The height of the column of aa residues represents the level of conservation at a particular position, while the relative residue height denotes its frequency. Polar aa is depicted in green, basic aa in blue, acidic aa in red and hydrophobic aa in black

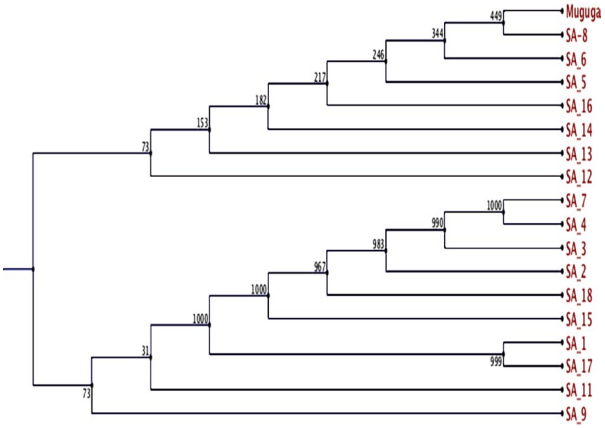

FIGURE 5 Dendrogram depicting genetic relationships between the predicted *Theileria parva* Tp2 proteins from 17 representative field isolates from South West Uganda. (SA) denotes the *T. parva* field isolates, and Muguga is the *T. parva* Muguga reference genotype. The field isolates clustered into 2 clades with the reference strain (Muguga) present in one of these clades

field samples were grouped into two clusters with *T. parva* Muguga clustering together with one group of south-western Uganda isolates (Figure 5).

4 | DISCUSSION

This study generated baseline data relating to diversity, based on polymorphic VNTR loci and CD8 + target antigen gene sequences (Tp1 and Tp2), within *T. parva* populations from an endemic region in South West Uganda. The data revealed extensive *T. parva* allelic diversity in the 14 satellite markers among 200 parasite samples. Higher levels of diversity are typically revealed when multi-locus genotypes are derived by combining data from a panel of VNTR loci, dispersed on all four chromosomes. Each VNTR was found to be highly polymorphic, as previously described (Oura et al., 2003, 2005). A high level of diversity was also observed among *T. parva* genotypes in Lira and Kayunga in northern and central Uganda respectively, with evidence for sub-structuring between and within populations in each area (Oura et al., 2005). Previous studies in other parts of East Africa, primarily Kenya, using a variety of methods including phenotypic analysis with anti-parasite monoclonal antibodies, Southern blotting using repetitive DNA probes and sequencing of the gene encoding the p67 sporozoite surface antigen (Allsopp & Allsopp, 1988; Conrad, Lams, Brown, Sohanpal, & ole-MoiYoi, 1987; Nene, Iams, Gobright, & Musoke, 1992; Odongo et al., 2006) have revealed high levels of diversity at other loci.

East coast fever control strategies are only occasionally implemented in South West Uganda and the African taurine (Bos taurus) Ankole cattle frequently co-graze with cape buffalo in this area. The African cape buffalo are important wildlife reservoirs of T. parva and are typically infected with parasite populations exhibiting higher antigenic diversity than T. parva that can be transmitted between cattle by ticks (Obara et al., 2015; Pelle et al., 2011). In addition, it has recently been demonstrated that cattle immunized by (ITM) using the Muguga cocktail are susceptible to challenge with buffalo-derived parasites (Bishop et al., 2015; Sitt et al., 2015). Genome sequencing of schizonts purified from cell lines (Hayashida et al., 2013), a comprehensive study of variation within the genes encoding the Tp antigens (Pelle et al., 2011) and VNTR analyses (Oura, Tait, Asiimwe, Lubega, & Weir, 2011) have all confirmed that the cape buffalo harbours a much greater diversity of T. parva genotypes relative to the strains transmissible between cattle by ticks. It is therefore important to understand the extent of parasite antigen gene diversity at the livestock–wildlife interface since variation in the epitopes that are targets of the protective immune responses could contribute to lack of cross-immunity between vaccinated animals and local parasites. The Tp1 and Tp2 epitope conservation revealed in this study suggests that ITM vaccination in south-western Uganda may be effective against challenge with cattle-derived T. parva. However, it should be noted that it has not yet been shown that epitopes within these antigens are presented by the as yet uncharacterized Class I MHC molecules that are expressed by Ankole cattle.

In addition, the efficacy of the trivalent Muguga cocktail version of ITM against buffalo-derived parasite challenge in this region has not yet been tested. In this context, direct evaluation of Tp1 and Tp2 sequence variation in *T. parva* from co-grazing buffalo would be of interest. It is worth noting that in areas adjacent to and within Lake Mburo National park, close to the location of the current study, *T. parva* populations from co-grazing cattle and buffalo contained populations that had little or no overlap in genotypes, as indicated by VNTR analysis (Oura et al., 2011).

A high frequency of infection of cattle with mixed genotypes was also observed. This is consistent with previous studies conducted in Uganda (Oura et al., 2005) and other parts of Africa (Geysen, Bishop, Skilton, Dolan, & Morzaria, 1999; Muleya et al., 2012; Odongo et al., 2006) that revealed existence of significant numbers of multiple infections in field samples when analysing populations of *T. parva* using VNTR markers. The high *T. parva* diversity and frequency of multiple infections may be explained by heavy tick infestations resulting from irregular tick control regimes (Wambura, Gwakisa, Silayo, & Rugaimukamu, 1998). Additionally, as demonstrated by earlier studies, a high degree of genetic diversity within a population may be a result of high levels of recombination and associated gene conversion within the *R. appendiculatus* tick vector (Katzer, Lizundia, Ngugi, Blake, & McKeever, 2011; Henson et al., 2012;., Hayashida et al., 2013).

These findings from Uganda, contrast with data from Zambia (Geysen et al., 1999) and Tanzania (Mwega et al., 2015) that indicated relative homogeneity within the *T. parva* populations. These differences may partially be attributable to the fact that different polymorphic marker loci were used in Zambia, although application of VNTR markers also revealed limited diversity in *T. parva* isolates from the North East region (R. Bishop and R. Skilton unpublished observations). In Tanzania, relatively few samples were analysed and these were isolated from areas where strict tick control measures were employed

suggesting that transmission levels were low in the Tanzanian population. To improve interpretation of data in studies using VNTRs to analyse mixed infections in field samples, as observed in Uganda, cloning of infected cell lines or choosing samples with only a single allele (although this imposes a bias towards sampling younger animals, Oura et al., 2005) would ideally be desirable before generating MLGs.

The *T. parva* CD8 + target antigens Tp1 and Tp2 that are immunodominant in cattle populations expressing specific Class I MHC haplotypes were also polymorphic in the study population. The amino acid sequences of these vaccine candidate schizont antigens revealed diversity in the one and six defined CTL epitopes present in the Tp1 and Tp2 antigens respectively (Figures 3–5). Similar results were reported in a previous study indicating that six different CD8 + epitopes restricted by 5 different class 1 MHC haplotypes Tp2 protein, exhibited many variants for each epitope ranging from 18 for epitope 6 to 25 for epitope 1 (Pelle et al., 2011).

The Tp1 antigen has only a single epitope identified to date, and only three variants have been noted (Pelle et al., 2011). All the observed alleles identified in South West Uganda for both Tp1 and Tp2 have previously been documented (Pelle et al., 2011). Studies from Tanzania (Mwega et al., 2015) revealed contrasting data whereby the Tp1 fragment was a single protein sequence in all the isolates and the Tp2 epitope sequences were identical to those in the *T. parva* Muguga reference stock, except for one Tp2 epitope which was similar to the *T. parva* Kiambu, another component of the trivalent Muguga cocktail vaccine (FAO1). This deviation from present study and previous studies (Pelle et al., 2011) could be explained by the low number of samples that were investigated but is also consistent with extensive vaccination programmes in Tanzania (reviewed by Di Giulio et al., 2009) using the ITM Muguga cocktail followed by subsequent transmission to naïve ticks and cattle.

The present study failed to demonstrate evolutionary patterns consistent with positive selection at the Tp1 locus and instead showed an identical log likelihood for both the near-neutral and the positive selection model. This finding is consistent with previous analyses of selection at other *T. parva* loci that have failed to demonstrate a significant posterior probability of such selection, for example in the gene encoding p67 (Obara et al., 2015). An earlier study using multiple isolates from both cattle and buffalo failed to demonstrate preferential distribution of positively selected amino acid sites within the mapped epitopes of Tp1 and Tp2 (Pelle et al., 2011).

It is not certain that the mapped Tp1 and Tp2 epitopes are presented to the bovine immune system by Class I MHC molecules in South West Ugandan cattle, since the class I MHC loci of Ankole are distinct from the cattle used in the epitope mapping study (Obara et al., 2016). However, the conservation of the epitope and antigen gene sequences in these parasites has a practical implication, since it suggests that if Muguga cocktail genetic components are transmitted to local ticks and cattle following ITM immunization, the changes in population genetic structure of *T. parva* may serve to produce a more homogenous challenge and improved protection, at least in the short term. Thus, one major concern among veterinary authorities that has constrained large scale ITM deployment in Uganda should now be reduced.

The interpretation of phylogenetic relationships inferred from studies using a limited number of loci, rather than complete genomes, must be made carefully in the case of pathogens that acquire genetic diversity by recombination. This is due to the fact that each locus can become chimeric by crossing over between genotypes that have different evolutionary histories. Therefore, a number of independent loci should be included to estimate the 'true' relationship between isolates such as multi-locus sequence typing, or genome-wide SNP analysis. In this context, analysis of entire genomes should be more informative.

It is important to appreciate the extent of genetic diversity in parasite populations in order to understand the potential impact of control measures, particularly vaccination, (Muleya et al., 2012). The most important finding from the current study with practical implications for deployment of the live vaccine is the fact that all of known CD8 epitope variants in Uganda are present within the trivalent Muguga cocktail version of ITM. Assuming that additional loci that represent targets of protective immune responses in cattle are similarly conserved, it is reasonable to conclude that the trivalent ITM vaccine, which has not hitherto been widely deployed in Uganda can be used, with an expectation of effectiveness.

ACKNOWLEDGEMENTS

We are grateful to the Local District Administrators (LCs), Veterinary staff and farmers in Kiruhura district in south-western Uganda for assistance during field trips, data collection and cattle sampling. The staff at the Molecular Biology Laboratory (MOBILA) at CoVAB, Makerere University for technical support in sample preparation and storage and BecA Hub, ILRI for technical support during the laboratory work and data analysis This work was supported in part by Grants from DFG (German-African Cooperation; 'Molecular epidemiology network for promotion and support of delivery of live vaccines against *Theileria parva* and *Theileria annulata* infection in Eastern and Northern Africa – CL 166/4-2') and The Africa Biosciences Challenge Fund–2012 Research Fellowships to the Biosciences eastern and central Africa (BecA) Hub at the International Livestock Research Institute (ILRI). We also acknowledge Dr Pelle (ILRI) for providing reference *T. parva* genomic DNA, guidance and useful discussions on Tp1 and Tp2 genes. We also wish to thank Dr. Selma Kamal (Central Veterinary Research Laboratory, Khartoum, Sudan) for generating the map shown in Figure 1.

COMPETING INTERESTS

The authors declare that they have no competing interests.

REFERENCES

Allsopp, B. A., & Allsopp, M. T. (1988). *Theileria parva*: Genomic DNA studies reveal intra-specific sequence diversity. *Molecular and Biochemical Parasitology*, *28*(1), 77–83. https://doi.org/10.1016/0166-6851(88)90183-1

Bishop, R. P., Hemmink, J. D., Morrison, W. I., Weir, W., Toye, P. G., Sitt, T., … Odongo, D. O. (2015). The African buffalo parasite *Theileria*. sp. (buffalo) can infect and immortalize cattle leukocytes and encodes divergent orthologues of *Theileria parva* antigen genes. *International Journal for Parasitology. Parasites and Wildlife*, *4*(3), 333–342. https://doi.org/10.1016/j.ijppaw.2015.08.006

Bishop, R., Sohanpal, B., Kariuki, D. P., Young, A. S., Nene, V., Baylis, H., … Morzaria, S. P. (1992). Detection of a carrier state in *Theileria parva*-infected cattle by the polymerase chain reaction. *Parasitology*, *104*(Pt 2), 215–232. https://doi.org/10.1017/S0031182000061655

Coetzer, J. A. W., & Tustin, R. C. (2004). *Infectious Diseases of Livestock* (2nd edn, Vol. *1* p. 461). Cape Town, South Africa: Oxford University Press.

Conrad, P. A., Lams, K., Brown, W. C., Sohanpal, B., & ole-MoiYoi, O. K. (1987). DNA probes detect genomic diversity in *Theileria parva* stocks. *Molecular and Biochemical Parasitology*, *25*(3), 213–226.

Di Giulio, G., Lynen, G., Morzaria, S., Oura, C., & Bishop, R. (2009). Live immunization against East Coast fever—status. *Trends in Parasitology*, *25*(2), 85–92. https://doi.org/10.1016/j.pt.2008.11.007

Geysen, D., Bishop, R., Skilton, R., Dolan, T. T., & Morzaria, S. (1999). Molecular epidemiology of *Theileria parva* in the field. *Tropical medicine & international health*, *4*(9), A21–A27. https://doi.org/10.1046/j.1365-3156.1999.00447.x

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

Hayashida, K., Abe, T., Weir, W., Nakao, R., Ito, K., Kajino, K., … Sugimoto, C. (2013). Whole-genome sequencing of *Theileria parva* strains provides insight into parasite migration and diversification in the African continent. *DNA research*, *20*(3), 209–220. https://doi.org/10.1093/dnares/dst003

Henson, S., Bishop, R. P., Morzaria, S., Spooner, P. R., Pelle, R., Poveda, L., … Qi, W. (2012). High-resolution genotyping and mapping of recombination and gene conversion in the protozoan *Theileria parva* using whole genome sequencing. *BMC genomics*, *13*, 503. https://doi.org/10.1186/1471-2164-13-503

Katzer, F., Lizundia, R., Ngugi, D., Blake, D., & McKeever, D. (2011). Construction of a genetic map for *Theileria parva*: Identification of hotspots of recombination. *International Journal for Parasitology*, *41*(6), 669–675. https://doi.org/10.1016/j.ijpara.2011.01.001

Kimura, M. (1980). A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution.*, *16*, 111–120. https://doi.org/10.1007/BF01731581

Librado, P., & Rozas, J. (2009). DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics*, *25*, 1451–1452. https://doi.org/10.1093/bioinformatics/btp187

McKeever, D. J., Taracha, E. L., Innes, E. L., MacHugh, N. D., Awino, E., Goddeeris, B. M., & Morrison, W. I. (1994). Adoptive transfer of immunity to Theileria parva in the CD8 + fraction of responding efferent lymph. *Proceedings of the National Academy of Sciences of the United States of America*, *91*(5), 1959–1963. https://doi.org/10.1073/pnas.91.5.1959

Muhanguzi, D., Matovu, E., & Waiswa, C. (2010). Prevalence and characterization of *Theileria* and *Babesia species* in cattle under different husbandry systems in Western Uganda. *International Journal of Animal and Veterinary Advances*, *2*, 51–58.

Muleya, W., Namangala, B., Simuunza, M., Nakao, R., Inoue, N., Kimura, T., … Sawa, H. (2012). Population genetic analysis and sub-structuring of *Theileria parva* in the northern and eastern parts of Zambia. *Parasites and Vectors*, *5*, 255. https://doi.org/10.1186/1756-3305-5-255

Mwega, E., Salih, D. H., Njahira, M., Rukambile, E., Skilton, R., & Gwakisa, P. (2015). Genetic and antigenic diversity of *Theileria parva* in cattle in Eastern and Southern Tanzania. A study to support control of East Coast fever. *Parasitology*, *142*, 698–705. https://doi.org/10.1017/S0031182014001784

Nene, V., Iams, K. P., Gobright, E., & Musoke, A. J. (1992). Characterisation of the gene encoding a candidate vaccine antigen of *Theileria parva* sporozoites. *Molecular and Biochemical Parasitology*, *51*(1), 17–27. https://doi.org/10.1016/0166-6851(92)90196-Q

Norval, R. A. I., Perry, B. D., Young, A. S., Lawrence, J. A., Mukhebi, A. W., Bishop, R. P., & McKeever, D. (1992). *The epidemiology of Theileriosis in Africa*. London, UK: Academic Press.

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

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

Odongo, D. O., Oura, C. A., Spooner, P. R., Kiara, H., Mburu, D., Hanotte, O. H., & Bishop, R. P. (2006). Linkage disequilibrium between alleles at highly polymorphic mini- and micro-satellite loci of *Theileria parva* isolated from cattle in three regions of Kenya. *International Journal for Parasitology*, *36*, 937–946. https://doi.org/10.1016/j.ijpara.2006.01.008

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

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

Oura, C. A., Odongo, D. O., Lubega, G. W., Spooner, P. R., & Tait, A. (2003). A panel of microsatellite and minisatellite markers for the characterisation of field isolates of *Theileria parva*. *International journal for parasitology*, *33*, 1641–1653. https://doi.org/10.1016/S0020-7519(03)00280-7

Oura, C. A., Tait, A., Asiimwe, B., Lubega, G. W., & Weir, W. (2011). *Theileria parva* genetic diversity and haemoparasite prevalence in cattle and wildlife in and around Lake Mburo National Park in Uganda. *Parasitology research*, *108*(6), 1365–1374. https://doi.org/10.1007/s00436-010-2030-8

Pelle, R., Graham, S. P., Njahira, M. N., Osaso, J., Saya, R. M., Odongo, D. O., … Bishop, R. P. (2011). Two Theileria parva CD8 T cell antigen genes are more variable in buffalo than cattle parasites, but differ in pattern of sequence diversity. *PLoS ONE*, *6*(4), e19015. https://doi.org/10.1371/journal.pone.0019015

Radley, D. E., Brown, C. G. D., Cunningham, M. P., Kimber, C. D., Musisi, F. L., Payne, R. C., … Young, A. S. (1975). East Coast fever: 3.Chemoprophylactic immunization of cattle using oxytetracycline and a combination of theilerial strains. *Veterinary Parasitology*, *1*, 51–60. https://doi.org/10.1016/0304-4017(75)90007-2

Salih, D. A., Pelle, R., Mwacharo, J. M., Njahira, M. N., Marcellino, W. L., Kiara, H., … Skilton, R. A. (2017). Genes encoding two *Theileria parva* antigens recognized by CD8 + T-cells exhibit sequence diversity in South Sudanese cattle populations but the majority of alleles are similar to the Muguga component of the live vaccine cocktail. *PLoS ONE*, *12*(2), e0171426. https://doi.org/10.1371/journal.pone.0171426

Sitt, T., Poole, E. J., Ndambuki, G., Mwaura, S., Njoroge, T., Omondi, G. P., … Toye, P. (2015). Exposure of vaccinated and naive cattle to natural challenge from buffalo-derived *Theileria parva*. International journal for parasitology. *Parasites and wildlife*, *4*(2), 244–251. https://doi.org/10.1016/j.ijppaw.2015.04.006.

Swofford, D. L. (2003). *PAUP: Phylogenetic Analysis Using Parsimony*. Saunderland, MA: Sancruer Associates.

Thompson, J. D., Higgins, D. G., & Gibson, T. J. (1994). CLUSTALW: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, *22*, 4673–4680. https://doi.org/10.1093/nar/22.22.4673

Thomsen, M. C., & Nielsen, M. (2012). Seq2Logo: A method for construction and visualization of amino acid binding motifs and sequence profiles including sequence weighting, pseudo counts and two-sided representation of amino acid enrichment and depletion. *Nucleic Acids Research*, *40*, W281–W287. https://doi.org/10.1093/nar/gks469

Thumbi, S. M., Bronsvoort, M. B., Kiara, H., Toye, P. G., Poole, J., Ndila, M., … Woolhouse, M. E. (2013). Mortality in East African shorthorn zebu cattle under one year: Predictors of infectious-disease mortality. *BMC Veterinary Research*, *9*, 175. https://doi.org/10.1186/1746-6148-9-175

Wambura, P. N., Gwakisa, P. S., Silayo, R. S., & Rugaimukamu, E. A. (1998). Breed-associated resistance to tick infestation in *Bos indicus* and their crosses with *Bos taurus*. *Veterinary Parasitology*, *77*, 63–70. https://doi.org/10.1016/S0304-4017(97)00229-X

How to cite this article: Nanteza A, Obara I, Kasaija P, et al. Antigen gene and variable number tandem repeat (VNTR) diversity in *Theileria parva* parasites from Ankole cattle in south-western Uganda: Evidence for conservation in antigen gene sequences combined with extensive polymorphism at VNTR loci. *Transbound Emerg Dis*. 2020;67(Suppl. 1):99–107. https://doi.org/10.1111/tbed.13311