Genotyping of *Theileria lestoquardi* from sheep and goats in Sudan to support control of Malignant Ovine Theileriosis

Ali^{1,2,*}, A.M., Salih^{1,3}, D.A., Njahira^{1,a}, M.N., Hassan³, S.K., EL Hussein³, A.M.,. Liu⁴, Z., Yin^{4,*}, H, Pelle¹, R. and Skilton^{1,a,*}, R.A.,

¹Biosciences eastern and central Africa-International Livestock Research Institute (BecA-ILRI) Hub, P.O. Box 30709-00100, Nairobi, Kenya

²University of Khartoum, Faculty of Veterinary Medicine, Khartoum, Sudan

³Central Veterinary Research Laboratory, P.O. Box 8067, Khartoum, Sudan

⁴State Key Laboratory of Veterinary Etiological Biology, Lanzhou Veterinary Research Institute, Chinese Academy of Agricultural Sciences, Xujiaping 1, Lanzhou, Gansu, 730046, China

^aInternational Centre of Insect Physiology and Ecology (*icipe*), P.O. Box 30772-00100, Nairobi, Kenya

*Corresponding author

Awadia_mosa@hotmail.com (AMA)

yinhong@caas.cn (YH)

rskilton@icipe.org (RAS)

Abstract

Theileriosis, caused by parasitic protozoa of the genus *Theileria* parasites, are among the major tick-borne diseases of ruminant livestock. The largest economic losses are attributed in particular to those caused by the leukoproliferative species of *Theileria*: *T. parva*, *T. annulata* and *T. lestoquardi*. *Theileria lestoquardi* is transmitted by *Hyalomma* ticks and causes malignant ovine theileriosis (MOT), a disease that is particularly prevalent in Sudan. The disease is considered of a high economic importance in Sudan, where export of sheep is a major component of the national economy. A live vaccine based on a Sudanese isolate of *T. lestoquardi* (Atbara strain) was previously developed for the control of MOT in Sudan, but not yet deployed in the field. The present study aims to genetically characterize and compare samples of *T. lestoquardi* circulating in Sudan as well as the live vaccine isolate in order to understand vaccine breakthroughs and failure that may occur. Sheep and goats blood samples were collected from six regions in Sudan that are known to be endemic for *T. lestoquardi* infection or have experienced outbreaks of MOT. Blood samples infected with *T. lestoquardi* were identified by PCR or RLB. Genotyping was carried out by (1) sequencing the homologues of two *T. parva* CD8⁺ T cell antigen genes, Tp1 and Tp2, and (2) using a panel of seven micro- and mini-satellite markers. A total of 100 *T. lestoquardi* positive field samples and the *T. lestoquardi* (Atbara) vaccine were genotyped. The results showed that all samples had mixed genotypes, with several alleles identified at one or more loci. The gene diversity ranged from 0.7840 (TS8) to 0.2133 (TS12) with mean values of 0.5470. PCA revealed three clusters of the parasite in Sudan; interestingly one independent cluster was clearly seen, corresponding to the vaccine isolate. The *T. lestoquardi* Tp1 homologue showed higher homology with *T. annulata* than with *T. parva* sequences included the defined single CD8⁺ T cell target epitope region. The result indicates that multiple genotypes are a common feature of *T. lestoquardi* infection in Sudan. Both genotyping and the sequencing results clearly showed that the vaccine isolate is highly distinct from the field samples. This finding raised the question whether vaccination with the prepared lived

vaccine will effectively protect animals against challenges by the field isolates of *T. lestoquardi*. The results of this work will inform on the best approach for controlling MOT in Sudan.

Keywords:

Genotyping; *Theileria lestoquardi*; malignant ovine theileriosis (MOT); Sudan; vaccination; small ruminant

1. Introduction

Small ruminant theileriosis in sheep and goats are caused by variety of *Theileria* species. A minimum of six *Theileria* species infects small ruminants. *T. seprata*, *T. ovis* and *T. recondita* are nonpathogenic, whereas *T. lestoquardi*, *Theileria* sp.1 (China) and *Theileria* sp.2 (China) are pathogenic (Uilenberg, 1997; Yin et al., 2002). *T. lestoquardi* is a highly pathogenic *Theileria* species for sheep and to lesser degree for goats, causing a disease known as malignant ovine theileriosis (MOT) and is transmitted by *Hyalomma* ticks (Levine, 1973). The disease has been reported in south-eastern Europe, North Africa, and the Near and Middle East. Control of the disease can be achieved by chemotherapy, using theilericidal drugs such as buparvaquone (El Hussein et al., 1993), as well as tick control using acaricides (Jongejan and Uilenberg, 2004). However, both methods have drawbacks. Drug treatment is expensive and requires early diagnosis of the disease (El Hussein et al., 1993) and can introduce residues in the food chain, whereas acaricide treatment rises environmental concerns and resistance to the chemical is also recorded (Foil et al., 2004). These drawbacks could be overcome by immunization of sheep with attenuated *T. lestoquardi* schizont infected ovine cells providing the animal with solid immunity (Gill et al., 1978), a method that has been applied successfully in Iraq and Iran (Hawa et al., 1981; Hashemi-Fesharki, 1997) and Southern parts of the former USSR (Dolan, 1989).

The disease causes high mortality and severely impacts on the livelihood of resource-poor farming communities in Sudan. MOT is endemic in some parts of the country, but outbreaks in new areas also occur. A live vaccine based on a Sudanese isolate of *T. lestoquardi* (Atbara strain) has been developed for the control of MOT in Sudan (Ahmed et al., 2013), but not yet deployed in the field.

This study will provide background information to support the deployment of the live vaccine in Sudan. The objectives are to define the genetic and antigenic composition of the live vaccine and field isolates of *T. lestoquardi* as there are fears that the introduction of ‘foreign’ parasites by vaccination might result in parasites with novel and potentially more virulent phenotypes. Therefore, before a decision is made to use this vaccine in Sudan, there is need to characterize the isolates of *T. lestoquardi* circulating in the region. In addition, knowing the genetic and antigenic composition of the vaccine and field parasites will help to understand vaccine failure and breakthroughs that may occur. The genetic and antigenic information will also help to reveal the possible origin of the rapidly spreading of *T. lestoquardi* outbreaks in Sudan and perhaps also further west in Africa where the habitat is suitable for the parasite.

2. Materials and Methods

2.1. Collection of blood samples and DNA extraction

Sheep and goats of different age groups, different types and sex were randomly selected in six locations of Sudan (Fig. 1). Multiple blood spots (usually 2) were obtained by spotting approximately 20 µl of blood from jugular vein onto Whatman 3MM filter papers (Maidstone, UK). These blood spots were air dried overnight and stored at room temperature in plastic bags until DNA extraction. A total of 178 (91 sheep and 87 goats) DNA from field samples was extracted from blood spots using PureLink™ Genomic DNA Mini extraction kits (Invitrogen®, Germany) according to the manufacturer's protocol. Genomic DNA was isolated from *T. lestoquardi* (Atbara) cell culture vaccine (Ahmed et al., 2013) using the Gentra Puregene Cell Kit (Qiagen, Hilden, Germany) according to the manufacturer's instructions. Purity and the quality of the extracted DNA were assessed using a Nano drop spectrophotometer (Wilmington, Delaware, USA).

2.2. Identification of *T. lestoquardi*

T. lestoquardi positive samples were detected by a sensitive specific PCR assay for *T. lestoquardi* using specific primers which amplify a 785 bp fragment of the gene that codes for a 30 kD merozoite surface protein (Kirvar et al., 1998). The specificity of the PCR assay was confirmed by sequencing of selected positive amplicons. *T. lestoquardi* (Atbara) cell line (Bakheit et al., 2006b) was used as positive control.

2.3. Amplification of the mini- and micro-satellite loci

Panels of seven micro- and mini-satellite markers; microsatellite (TS 5 and 12) and minisatellite markers (TS6, 8, 15, 20 and 25), were used to genotype *T. lestoquardi* samples. These markers

were developed based on the completed genome sequence of *T. annulata* and they were used for population studies (Weir et al. 2007). The forward primer of each marker was 5'-end labeled with one of four dyes: 6-FAM (Blue), VIC (Green), NED (Yellow) and PET (Red) (Bioneer, Korea), for subsequent detection of fragment length variation on an ABI genetic analyser (Weir et al. 2007).

The amplification reaction was done in PCR tubes as follows: template DNA (20 ng), 5 U Taq polymerase (Promega), 1 × PCR buffer (10 mM Tris-HCl [pH 8.4], 50 mM KCl, 2 mM MgCl₂), 0.2 mM dNTPs, and 0.2 μM each of reverse and forward primers, and made up to a final reaction volume of 10 μl with sterile distilled water. The PCR was carried out in a Gene Amp® PCR System 9700 (ABI-USA). The cycling conditions were as follows: denaturation at 94° C for 30 s, annealing at 60 ° C for 1 min and extension at 72 ° C for 1 min and final extension was at 72 ° C for 7 min before cooling to 4 ° C. The reaction was run for 35 cycles.

For fragment length analysis, (0.5–2 μl) of PCR product 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). Then denaturation was performed in a thermocycler at 95°C for 5 min followed by chilling of the plate on ice. The amplicons were processed and analysed on the ABI 3730 genetic analyser (Applied Biosystems-USA). Scoring was carried out using Gene Mapper programme (Applied Biosystems-USA). Amplicons with maximum peak height were scored, and a predominant peak was defined as that with the largest area under the curve. All data generated from GeneMapper were resized using the Allelobin software (Idury and Cardon 1997) based on consensus sequence repeats of the marker. A predominant allele at each locus was used to generate allele frequency data and multilocus genotypes (MLGs).

2.4. Population genetic analyses

Summary statistics were calculated using the genetic analysis package Power Marker V. 3.25 (Liu and Muse, 2005). Major allele frequency, gene diversity (GD) and Polymorphism information content (PIC) were used to determine diversity at each microsatellite locus. The maximum numbers of alleles identified for each marker in each population was constructed using GenAlEx V. 6.5 (Peakall and Smouse, 2006) plugged in Microsoft Excel. Similarly, the software was used to investigate patterns of distribution of each genotype in a population. In addition, principal component analysis (PCA) was constructed after calculating the genetic distance (GD) using sets of MLGs data. Analysis of molecular variance (AMOVA) was performed using the same software to assess genotypic variations across all the populations' studied.

2.5. Detection of Tp1 and Tp2 homologues of *T. lestoquardi*

Primers for nested PCR amplification of partial sequences of the Tp1 and Tp2 homologous genes of *T. lestoquardi* (TL1 and TL2, respectively) were based on the *T. annulata* homologous sequences (NCBI Reference Sequences: NC_011100.1 and NC_011129.1 respectively). Primers were designed using Primer 3 software (Table 1). The expected sizes of the nested PCR products for the *T. lestoquardi* Tp1 and Tp2 homologous sequences are 131 bp and 526 bp, based on the *T. annulata* NCBI Reference Sequences NC_011100.1 and NC_011129.1, respectively (Table 1).

All PCRs were done using Bioneer's AccuPower PCR-PreMix reagents (Bioneer, Korea). PCRs with the outer primers (primary PCR) were carried out in a total volume of 20 µl containing 1 U *Taq* DNA polymerase, 250 µM of each dNTP, 10 mM Tris-HCl (pH 9.0), 30 mM KCl, 1.5 mM MgCl₂, 0.2 µM of each primers and 2 µl (20ng) genomic DNA. Reactions for the inner (nested)

primers were as for the primary PCR, except they contained 1 µl of the primary PCR product instead of genomic DNA.

Amplification cycling conditions were as follows: initial denaturation at 95 °C for 5 min followed by 40 cycles of denaturation at 94 °C for 30 s, primer annealing temperature (Table 1) for 1 min and 72 °C for 1 min, followed by a final step 72 °C for 7 min. PCR products were analysed by agarose gel electrophoresis and products from positive nested PCRs were purified using a QIAquick PCR Purification Kit (Qiagen, Germany) and sent to the Segolip unit at the BecA-ILRI Hub Nairobi-Kenya for sequencing using an ABI 3730 capillary sequencer (Applied Biosystems, USA). Compilation, editing and assembly of sequences were performed using CLC DNA workbench version 6.7.1 (<http://www.clcbio.com>). BLAST analyses of the sequences were performed on the NCBI server (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>). The sequences were analysed using bioinformatics software: <http://emboss.bioinformatics.nl> and <http://www.genome.jp/tools/clustalw>

3. Results

Out of the 178 samples examined, 36 (20.2%) was found positive for *T. lestoquardi*. These samples included 32 sheep samples out of 91 (35.2%) and 4 goat samples out of 87 (4.6%). In addition, a total of 63 positive samples used in this study were confirmed previously by reverse line blot (RLB) (Salama *et al.*, unpublished data). Also the vaccine isolate was included in the study. In total, 100 samples were included in this study. A panel of seven micro- and mini-satellite markers out of the initial 10 markers described by (Weir et al. 2007) was used to genotype these 100 positive samples of *T. lestoquardi* representing the parasite population in Sudan and *T. lestoquardi* (Atbara) cell culture vaccine. The selected seven micro- and mini-satellite markers are highly polymorphic and distributed over the four chromosomes of *T. lestoquardi*. Each sample was sequenced twice.

The number of alleles at each of the seven micro-and min-satellite loci were determined for each sample. The results showed that every sample represented mixed infection, with several alleles identified at one or more loci even in the attenuated clonal isolate of *T. lestoquardi* (Vaccine). The maximum number of alleles identified by each marker ranged from three (TS12) to 15 (TS8 and TS20), and the number of alleles for each marker is shown in Table 2. Marker TS8 and Ts20 were the most polymorphic, with 15 alleles in total across all populations whereas TS12 was the least polymorphic, with three alleles.

Allele frequencies in representative loci (for example TS 5, TS 25 and TS 8) were observed (Fig. 2). Major allele frequencies in all populations ranged from 0.3267 (TS 8) to 0.8812 (TS12) with a mean value of 0.6124. The gene diversity ranged from 0.7840 (TS8) to 0.2133 (TS12) with mean

values of 0.5470 and PIC values ranged from 0.7544 (TS8) to 0.1975 (TS12), with mean values of 0.5327 (Table 2).

The data generated by genotyping DNA from Sudanese population were used based on the predominant allele at each locus. Principal Component Analysis (PCA) for the field samples and the vaccine isolate showed that the majority formed a large cluster (1) of the sample from endemic and non-endemic area including the four goats' samples. Some samples from the endemic region clustered together separately from the majority of the samples (Cluster 2). Interestingly one independent cluster was clearly seen, corresponding to the vaccine isolate. PCA revealed that there are three genotypes of the parasite population from Sudan (Fig. 3).

Two genotypes were observed from the endemic region in Sudan, the large cluster genotype includes the genotype from non-endemic area and goat genotype. Some samples from the endemic region occupied the lower left quadrant, most likely to exhibit some level of independence from the rest of the samples. The vaccine genotype occupied the upper left quadrant separately from the other two genotypes.

In order to determine the level of antigenic diversity in Sudanese isolates, analysis of partial sequences of *T. lestoquardi* homologues of Tp1 and Tp2 CD8⁺ T-cell antigens, (TL1 and TL2, respectively) was carried out. For TL1 locus, clean sequence was obtained from the vaccine isolate only. The predicted protein sequence of this TL1 revealed that the *T. lestoquardi* vaccine isolate contains a variant of the 11-mer *T. parva* Muguga single CD8⁺ T-cell epitope (**VGYPKVKEEML**) (MacHugh et al., 2009) with the 6 residues shown in bold being conserved: **YKYPNIKQEML**. Multiple amino acid sequence alignment of the partial TL1 with the *T. annulata*

and *T. parva* homologues showed higher level of sequence similarity with *T. annulata* and quite diverse from *T. parva* (Fig. 4a).

For the TL2 locus, sequence analysis from 25 samples and the vaccine isolates (vaccine isolate and the positive control were sequenced twice) revealed polymorphisms at the DNA as well as amino acid levels (Fig. 4b). Furthermore, our study showed that variants of the six *T. parva* Muguga CD8⁺ T-cell epitopes previously reported (Graham et al., 2006, 2008; MacHugh et al., 2009) were also present in the field and vaccine isolates of *T. lestoquardi* from Sudan (Fig.5). The nucleotide sequence of TL2 alleles was illustrated by the generation of phylogenetic tree, using *T. annulata* gene for putative surface protein D (TaD) (Gen bank XM_949555) as outgroup (Fig. 6). The sequence results clearly showed the difference between the vaccine and field isolates.

4. Discussion

Malignant ovine theileriosis (MOT) is an emerging disease in the Sudan. A cell line live vaccine base on schizont attenuation has been developed to be applied for control of MOT. This study considers a comprehensive diversity assessment of *T. lestoquardi* in sheep and goats. This will provide a base to rational implementation of future control strategies and baseline data that will allow insight into the origin and diversity of *T. lestoquardi* present in this region and enable predicting potential vaccine breakthroughs and monitoring of possible future spread to adjacent regions.

Analysis of population of *Theileria* parasites using micro- and mini-satellite markers have previously been identified for *T. annulata* (Weir et al., 2007), *T. parva* (Oura et al., 2003) and recently for *T. lestoquardi* (Al-Hamidhi et al., 2016). The present study represents the first broad

population genetic analysis of *T. lestoquardi* in Sudan using seven micro-and mini-satellite markers of *T. annulata* (Weir et al., 2007). Biological and molecular phylogenetic studies clearly indicated that *T. lestoquardi* is more closely related to *T. annulata* than to *T. parva* (Katzner et al., 1998; Schnittger et al., 2003, Goh et al., 2016). Both species were probably evolved from a common ancestor and this fact is supported biologically such as both parasites share the same vector tick.

The results of genotyping revealed high level of diversity and several alleles were identified at one or more locus. This indicates that there is multiplicity of infection (MOI) even in the attenuated clonal isolate of *T. lestoquardi* (vaccine) it may also suggest that events leading to gene mutation are taken place in *T. lestoquardi* isolates.

The high level of genetic diversity may be due to the intensity of the parasite transmission and multiple re-infections of animals during the tick season resulting in genetic recombination especially in endemic area where the clonal parasite (vaccine) was isolated. Principal Component Analysis (PCA) for the Sudanese isolates showed that the majority formed a large cluster of the isolates from endemic and non-endemic area including the goats isolates while some isolates from the endemic region clustered together separately from the majority of the isolates. Interestingly one independent cluster was clearly seen, corresponding to the vaccine isolate. From this finding, one would question whether vaccination with the prepared lived vaccine will effectively protect animals against challenges by the field isolates of *T. lestoquardi*.

The identification of a number of *T. parva* antigens and epitopes recognized by CD8⁺ T-cells from *T. parva*-immune cattle (Graham et al., 2006, 2008) followed by the study of immune responses to two of these antigens, Tp1 and Tp2, have revealed that they are highly dominant

targets of the CD8⁺ T-cell response in cattle expressing the A18 and A10 class I MHC haplotypes, respectively (MacHugh et al. 2009). In the present study we sequence the homologues of these antigenic genes from *T. lestoquardi* (TL1 and TL2) and identified the epitopes variants.

The sequence of *T. lestoquardi* TL1 is more closely related to *T. annulata* than to *T. parva* including the single epitope region; our result supports the conclusion that both species probably evolved from a common ancestor (Al-Hamidhi et al., 2016). Allelic sequence analysis of *T. lestoquardi* TL2 revealed polymorphism at the DNA and amino acid levels including in the epitopes regions. These data corroborated the genetic diversity results obtained. Antigenic polymorphism can result from positive selection pressures of the host immune system, which may also be the case for TL2, a strategy that *T. lestoquardi* may have developed to escape from the host immune response.

The genetic data of the field isolates indicate a panmictic population structure of the parasite in Sudan. This result was supported by the results obtained from the cell lines that also showed different alleles in a single culture at passage 100 (vaccine). This result could also explain the long duration needed to attenuate some *T. annulata* strains isolated from endemic areas, which possibly consist of different genotypes (Ali et al., 2008)

The phylogenetic tree from TL2 sequences clearly confirmed the difference between the vaccine and the field isolates and, on one hand, this may be attributed to the difference between virulence and attenuated phenotypes of the parasite isolates. On the other hand, the difference between the vaccine and the field isolates may be a result of the introduction of 'foreign' parasites by previous vaccination campaigns that might lead to a generation of parasites with novel genotypes with potentially more virulent phenotypes.

In conclusion our result indicates that multiple genotypes are a common feature of *T. lestoquardi* infection in Sudan. The genotyping and sequencing results clearly showed that the vaccine isolate is highly distinct from the field isolates. This may indicate that a cocktail of vaccine strains might be required in order to control the MOT in Sudan. Finally the results of this work will assist in deciding on the best policy for controlling MOT in Sudan.

Acknowledgments

This project was supported by the BecA-ILRI Hub through the Africa Biosciences Challenge Fund (ABCF) program. The ABCF Program is funded by the Australian Department for Foreign Affairs and Trade (DFAT) through the BecA-CSIRO partnership; the Syngenta Foundation for Sustainable Agriculture (SFSA); the Bill & Melinda Gates Foundation (BMGF); the UK Department for International Development (DFID) and; the Swedish International Development Cooperation Agency (SIDA). AWA was a recipient of an the Africa Biosciences Challenge Fund (ABCF) Fellowship

References

- Ahmed, B.M., Taha, K.M., Enan, K.A., Elfahal, A.M., El Hussein. A.M. 2013. Attenuation of *Theileria lestoquardi* infected cells and immunization of sheep against malignant ovine theileriosis. Vaccine 31: 4775-4781.
- Al-Hamidhi, S., Weir, W., Kinnaird, J., Tagledeen, M., Beja-Pereira, A., Morrison, I., Thompson, J., Tait, A., Shiels, B., Babiker, H.A., 2016. *Theileria lestoquardi* displays reduced genetic diversity relative to sympatric *Theileria annulata* in Oman. In: Infection, Genetics and Evolution, 10.1016/j.meegid.2016.05.007
- Ali, A.M., Salih, D., Bakheit, M., El Hussein, A.R.M., Hassan, S.M., Mukhtar, M.M., Ahmed, J.S., Seitzer, U., 2008. Comparison of the Molecular Structure of the TaSP Gene of *Theileria annulata* from Sudanese Isolates. Annals of the New York Academy of Sciences: 1149: 218–220.
- Bakheit, M.A., Endl, E., Ahmed, J.S., Seitzer, U., 2006. Purification of macroschizonts of a Sudanese isolate of *Theileria lestoquardi* (*T. lestoquardi* [Atbara]). Annals of the New York Academy of Sciences: 1081: 453–462.
- Dolan, T.T., 1989. Theileriosis: A comprehensive review. Rev. Sci. Tech. Off Int. Epiz., 8: 11-36.
- El Hussein, A.M., El Ghali, A.A., Mohammed, S.A., 1993. Efficacy of buparvaquone in the treatment of malignant theileriosis of sheep in Ed-Damer Province, North State, Sudan. A field trial. Sud J Vet Res, 12:51–57.

- Foil, L.D., Coleman, P., Eisler, M., Fragoso-Sanchez, H., Garcia-Vazquez, Z., Guerrero, F.D., 2004. Factors that influence the prevalence of acaricide resistance and tick-borne diseases. *Vet. Parasitol.* 125: 163–81.
- Gill, B.S., Bhattacharyulu, Y., Kaur, D., Singh, A., 1978. Chemoprophylaxis with tetracycline drugs in the immunization of cattle against *Theileria annulata* infection. *Int. J. Parasitol.* 8: 467-469.
- Goh, S., Ngugi, D., Lizundia, R., Hostettler, I., Woods, K., Ballingall, K., MacHugh, N.D., Morrison, W.I., Weir, W., Shiels, B., Werling, D., 2016. Identification of *Theileria lestoquardi* antigens recognized by CD8+ T Cells. *PLoS One*, 11(9):e0162571. doi: 10.1371/journal.pone.0162571. eCollection 2016.
- Graham, S.P., Pellé, R., Honda, Y., Mwangi, D.M., Tonukari, N.J., 2006. *Theileria parva* candidate vaccine antigens recognized by immune bovine cytotoxic T lymphocytes. *Proc Natl Acad Sci U S A* 103: 3286–3291.
- Graham, S.P., Pellé, R., Yamage, M., Mwangi, D.M., Honda, Y., 2008. Characterization of the fine specificity of bovine CD8 T-cell responses to defined antigens from the protozoan parasite *Theileria parva*. *Infect Immun* 76: 685–694.
- Hashemi-Fesharki, R., 1997. Tick-borne diseases of sheep and goats and their related vectors in Iran. *Parassitologia.* 39: 115–117.
- Hawa, N.J., Latif, B.M., Ali, S.R., 1981. Immunization of sheep against *Theileria hirci* infection with schizonts propagated in tissue culture. *Vet Parasitol.* 9: 91–97.

- Idury, R.M., Cardon, L.R. 1997. A simple method for automated allele binning in microsatellite markers. *Genome Research*, 7, 1104–1109.
- Jongejan, F., Uilenberg, G., 2004 The global importance of ticks. *Parasitology* 129 Suppl: S3–S14.
- Katzer, F., McKellar, S., Kirvar, E., Shiels, B., 1998. Phylogenetic analysis of *Theileria* and *Babesia* equi in relation to the establishment of parasite populations within novel host species and the development of diagnostic tests. *Molecular and Biochemical Parasitology*. 95: 33-44.
- Kirvar, E., Ilhan, T., Katzer, F., Wilkie, G., Hooshmand-Rad, P., Brown, D., 1998. Detection of *Theileria lestoquardi* (*hirci*) in ticks, sheep and goats using the polymerase chain reaction. *Ann. N. Y. Acad. Sci.* 849: 52-62.
- Levine, N.D., 1973. Protozoan parasites of domestic animal and of man. 2nd Edition. Burgess publishing Co, Minneapolis, USA.
- Liu, K., Muse, S.V., 2005. PowerMarker: integrated analysis environment for genetic marker data. *Bioinformatics* 21: 2128–2129.
- MacHugh, N.D., Connelley, T., Graham, S.P., Pelle, R., Formisano, P., 2009. CD8⁺ T-cell responses to *Theileria parva* are preferentially directed to a single dominant antigen: Implications for parasite strain-specific immunity. *Eur J Immunol* 39: 2459–2469.
- Oura, C.A., Odongo, D., Lubega, G., Spooner, P., Tait, A., Bishop, R., 2003. A panel of microsatellite and minisatellite markers for the characterisation of field isolates of *Theileria parva*. *International Journal of Parasitology* 33: 1641–1653.

- Peakall, R., Smouse, P.E., 2006. Genalex 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes*. 6(1): 288–295.
- Schnittger, L., Yin, H., Gubbels, M.J., Beyer, D., Niemann, S., Jongejan, F., Ahmed, J.S., 2003. Phylogeny of sheep and goat *Theileria* and *Babesia* parasites. *Parasitol. Res.* 91: 398-406.
- Uilenberg, G. (1997). General review of tick-borne diseases of sheep and goats world-wide. *Parassitologia* 39, 161-165.
- Weir, W., Ben-Miled, L., Karagenç, T., Katzer, F., Darghouth, M., Shiels, B., 2007. Genetic exchange and sub-structuring in *Theileria annulata* populations. *Molecular and Biochemical Parasitology* 154(2):170–180.
- Yin, H., Luo, J., Guan, G., Gao, Y., Lu, B., Zhang, Q., Ma, M., Lu, W., Lu, C. Yuan, Z., Guo, S., Wang, B., Du, H., Schnittger, L., Ahmed, J. and Jongejan, F. (2002). Transmission of an unidentified *Theileria* species to small ruminants by *Haemaphysalis qinghaiensis* ticks collected in the field. *Parasitology Research* 88, S25-S27.

Figure legends:

Fig. 1: Map of the Republic of Sudan showing Location of samples collection including the endemic region in Khartoum (central) and Atbara (north)-and other non-endemic regions in Sudan

Fig. 2: The allele frequencies of alleles in field populations of *T. lestoquardi* from different region in Sudan for representative markers Ts5, Ts25 and Ts8

Fig. 3: Principal Component Analysis (PCA) of *T.lestosquardi* populations from different field samples and the vaccine isolate. The proportion of variation in the population dataset explained by each axis is shown in parentheses. PCA was performed using multi-locus genotype data.

Fig. 4: a. Multiple amino acid sequence alignment of the deduced amino acid of TL1 from the vaccine isolate of *T. lestoquardi*, with Ta1 a partial sequence of *T. annulata* (*Theileria annulata* hypothetical protein (GenBank accession no. TA17450) and Tp1 *T. parva* (GenBank accession no. XP_762973) and Tl1 *T. lestoquardi* hypothetical protein (GenBank accession no. KT989585) The single letter amino acid code is used and the single previously identified *T. parva* CD8⁺ T-cell target epitope and Variants are bolded and underlined.

b. *T. lestoquardi* TL2 CTL epitopes variants obtained in this study from *T. lestoquardi* field samples (1—5); 6= *T. lestoquardi* Surface protein d precursor (GenBank accession no. KT989586) and the vaccine isolate (Vac).

a

TL1	PASSPTFFKNGTDFEIC	<u>YKYFNIKEML</u>	ELGVKLGRITQPTTPETTPVTK
Ta1	PASSPTFFKNGTDFEIC	<u>YKYFNIKEML</u>	ELGVKFGRITQPTTPETTPVTK
Tp1	PDSKPIFFKNDGDKFLRC	<u>VGYPKVKKEML</u>	EMATKFNRLPKGVEIPAPPGVKP
Tl1	TESRPLFFKNGDEFELC	<u>SKFENIKGEML</u>	ELAKKFGRTAPSGLPKPTPISP

b

	1	2	3	4	5	6
1	<u>SDEELDSL</u> GML	<u>PDLNK</u> PLL	<u>KTSHIL</u> GKVGKF	<u>AQSIH</u> CVANKKE	<u>TIPN</u> PCDW	
2	<u>SDEELDSL</u> GML	<u>PDLNK</u> PLL	<u>KTSHIL</u> GKVGKF	<u>AQSIH</u> CVANKKE	<u>TIPN</u> PCDW	
3	<u>SDEELDSL</u> GML	<u>PDLNK</u> PLL	<u>KTSHIL</u> GKVGKF	<u>AQSIH</u> CVANKKE	<u>SIPN</u> PCDW	
4	<u>TDEELDTL</u> GML	<u>PDLNK</u> PLL	<u>QTSHIL</u> GKVGKF	<u>AQSIH</u> CVANKKE	<u>TIPN</u> PCDW	
5	<u>TDEELDSL</u> GML	<u>PDLNK</u> PLL	<u>QTSHIL</u> GKVGKF	<u>AQSIH</u> CVANKKE	<u>TIPN</u> PCDW	
6	<u>TDEELDSL</u> GML	<u>PELNRP</u> LL	<u>QTSHIL</u> GKVGKF	<u>AQSIH</u> CVANKKE	<u>SIPN</u> PCDW	
Vac	<u>SDEELDTM</u> CML	<u>TDLNK</u> ELL	<u>QTSHIF</u> GKVGKF	<u>VQSIM</u> CVANKKE	<u>KIPN</u> PCDW	

Fig. 5: Multiple amino acid sequence alignment of 26 partial sequence of TL2 antigen variants obtained in this study. The identified six CD8⁺ T-cell target epitopes in Tp2 are coloured in red and pick.

• Tp2 LHSFVLGG-NC SHEELKGLGMEGGGFIKDALFKS SHGAKGVKRYGLKTI PKVKVLADETLFGKHGLGGISKDC LKCFQSIHCVANAKCGACLGPCITDDCQNCDFRNC SALLE CIGKETS IPNPDWKE
 • T-an ISSSVLGANNNC KDEELDAMGLIKID LNKELLFQT SHI LKGVGKKHG IKP-GTKVEKFLTETQKLFGEHGISGVSEDC LKCFQSIHCVANAKCGPCLRGPCITEEQNCVRRNCMPALLE CIGKETS IPNPDWKE
 •
 • Pos ISSSVLGANNNC SDEELD SLGMLDKPD LNKPLLFKT SHI LKGVGKKHG IKP-GTKVEKFLTETQKLFGEHGISGVSEDC LKCFQSIHCVANAKCGPCLRGPCITEEQNCIKKNCMQJLLE CIGKET IPNPDWKE
 • 1Uk ISSSVLGANNNC SDEELD SLGMLDKPD LNKPLLFKT SHI LKGVGKKHG IKP-GTKVEKFLTETQKLFGEHGISGVSEDC LKCFQSIHCVANAKCGPCLRGPCITEEQNCIKKNCMQJLLE CIGKET IPNPDWKE
 • 22A ISSSVLGANNNC SDEELD SLGMLDKPD LNKPLLFKT SHI LKGVGKKHG IKP-GTKVEKFLTETQKLFGEHGISGVSEDC LKCFQSIHCVANAKCGPCLRGPCITEEQNCIKKNCMQJLLE CIGKET IPNPDWKE
 • 2 ISSSVLGANNNC SDEELD SLGMLDKPD LNKPLLFKT SHI LKGVGKKHG IKP-GTKVEKFLTETQKLFGEHGISGVSEDC LKCFQSIHCVANAKCGPCLRGPCITEEQNCIKKNCMQJLLE CIGKET IPNPDWKE
 • 52-T ISSSVLGANNNC SDEELD SLGMLDKPD LNKPLLFKT SHI LKGVGKKHG IKP-GTKVEKFLTETQKLFGEHGISGVSEDC LKCFQSIHCVANAKCGPCLRGPCITEEQNCIKKNCMQJLLE CIGKET IPNPDWKE
 • 15 ISSSVLGANNNC SDEELD SLGMLDKPD LNKPLLFKT SHI LKGVGKKHG IKP-GTKVEKFLTETQKLFGEHGISGVSEDC LKCFQSIHCVANAKCGPCLRGPCITEEQNCIKKNCMQJLLE CIGKET IPNPDWKE
 • K34 ISSSVLGANNNC SDEELD SLGMLDKPD LNKPLLFKT SHI LKGVGKKHG IKP-GTKVEKFLTETQKLFGEHGISGVSEDC LKCFQSIHCVANAKCGPCLRGPCITEEQNCIKKNCMQJLLE CIGKET IPNPDWKE
 • posC ISSSVLGANNNC SDEELD SLGMLDKPD LNKPLLFKT SHI LKGVGKKHG IKP-GTKVEKFLTETQKLFGEHGISGVSEDC LKCFQSIHCVANAKCGPCLRGPCITEEQNCIKKNCMQJLLE CIGKET IPNPDWKE
 • K11 ISSSVLGANNNC SDEELD SLGMLDKPD LNKPLLFKT SHI LKGVGKKHG IKP-GTKVEKFLTETQKLFGEHGISGVSEDC LKCFQSIHCVANAKCGPCLRGPCITEEQNCIKKNCMQJLLE CIGKET IPNPDWKE
 • K11.2 ISSSVLGANNNC SDEELD SLGMLDKPD LNKPLLFKT SHI LKGVGKKHG IKP-GTKVEKFLTETQKLFGEHGISGVSEDC LKCFQSIHCVANAKCGPCLRGPCITEEQNCIKKNCMQJLLE CIGKET IPNPDWKE
 • K10 ISSSVLGANNNC SDEELD SLGMLDKPD LNKPLLFKT SHI LKGVGKKHG IKP-GTKVEKFLTETQKLFGEHGISGVSEDC LKCFQSIHCVANAKCGPCLRGPCITEEQNCIKKNCMQJLLE CIGKET IPNPDWKE
 • 1alf ISSSVLGANNNC SDEELD SLGMLDKPD LNKPLLFKT SHI LKGVGKKHG IKP-GTKVEKFLTETQKLFGEHGISGVSEDC LKCFQSIHCVANAKCGPCLRGPCITEEQNCIKKNCMQJLLE CIGKET IPNPDWKE
 • A1f10 ISSSVLGANNNC SDEELD SLGMLDKPD LNKPLLFKT SHI LKGVGKKHG IKP-GTKVEKFLTETQKLFGEHGISGVSEDC LKCFQSIHCVANAKCGPCLRGPCITEEQNCIKKNCMQJLLE CIGKET IPNPDWKE
 • 13alf ISSSVLGANNNC SDEELD SLGMLDKPD LNKPLLFKT SHI LKGVGKKHG IKP-GTKVEKFLTETQKLFGEHGISGVSEDC LKCFQSIHCVANAKCGPCLRGPCITEEQNCIKKNCMQJLLE CIGKET IPNPDWKE
 • 67 ISSSVLGANNNC SDEELD SLGMLDKPD LNKPLLFKT SHI LKGVGKKHG IKP-GTKVEKFLTETQKLFGEHGISGVSEDC LKCFQSIHCVANAKCGPCLRGPCITEEQNCIKKNCMQJLLE CIGKET IPNPDWKE
 • 18 ISSSVLGANNNC SDEELD SLGMLDKPD LNKPLLFKT SHI LKGVGKKHG IKP-GTKVEKFLTETQKLFGEHGISGVSEDC LKCFQSIHCVANAKCGPCLRGPCITEEQNCIKKNCMQJLLE CIGKET IPNPDWKE
 • S9 ISSSVLGANNNC SDEELD SLGMLDKPD LNKPLLFKT SHI LKGVGKKHG IKP-GTKVEKFLTETQKLFGEHGISGVSEDC LKCFQSIHCVANAKCGPCLRGPCITEEQNCIKKNCMQJLLE CIGKET IPNPDWKE
 • 64 ISSSVLGANNNC SDEELD SLGMLDKPD LNKPLLFKT SHI LKGVGKKHG IKP-GTKVEKFLTETQKLFGEHGISGVSEDC LKCFQSIHCVANAKCGPCLRGPCITEEQNCIKKNCMQJLLE CIGKET IPNPDWKE
 • uk26 ISSSVLGANNNC SDEELD SLGMLDKPD LNKPLLFKT SHI LKGVGKKHG IKP-GTKVEKFLTETQKLFGEHGISGVSEDC LKCFQSIHCVANAKCGPCLRGPCITEEQNCIKKNCMQJLLE CIGKET IPNPDWKE
 • 74 ISSSVLGANNNC SDEELD SLGMLDKPD LNKPLLFKT SHI LKGVGKKHG IKP-GTKVEKFLTETQKLFGEHGISGVSEDC LKCFQSIHCVANAKCGPCLRGPCITEEQNCIKKNCMQJLLE CIGKET IPNPDWKE
 • ATB38 ISSSVLGANNNC SDEELD SLGMLDKPD LNKPLLFKT SHI LKGVGKKHG IKP-GTKVEKFLTETQKLFGEHGISGVSEDC LKCFQSIHCVANAKCGPCLRGPCITEEQNCIKKNCMQJLLE CIGKET IPNPDWKE
 • 1Atb ISSSVLGANNNC SDEELD SLGMLDKPD LNKPLLFKT SHI LKGVGKKHG IKP-GTKVEKFLTETQKLFGEHGISGVSEDC LKCFQSIHCVANAKCGPCLRGPCITEEQNCIKKNCMQJLLE CIGKET IPNPDWKE
 • 23Uk ISSSVLGANNNC SDEELD SLGMLDKPD LNKPLLFKT SHI LKGVGKKHG IKP-GTKVEKFLTETQKLFGEHGISGVSEDC LKCFQSIHCVANAKCGPCLRGPCITEEQNCIKKNCMQJLLE CIGKET IPNPDWKE
 • 97A ISSSVLGANNNC SDEELD SLGMLDKPD LNKPLLFKT SHI LKGVGKKHG IKP-GTKVEKFLTETQKLFGEHGISGVSEDC LKCFQSIHCVANAKCGPCLRGPCITEEQNCIKKNCMQJLLE CIGKET IPNPDWKE
 • 16 ISSSVLGANNNC TDEELD TLMGLDKPD LNKPLLFQT SHI LKGVGKKHG IKP-GTKVEKFLTETQKLFGEHGISGVSEDC LKCFQSIHCVANAKCGPCLRGPCITEEQNCIKKNCMQJLLE CIGKET IPNPDWKE
 • D3 ISSSVLGANNNC TDEELD TLMGLDKPD LNKPLLFQT SHI LKGVGKKHG IKP-GTKVEKFLTETQKLFGEHGISGVSEDC LKCFQSIHCVANAKCGPCLRGPCITEEQNCIKKNCMQJLLE CIGKET IPNPDWKE
 • vacc INSSVLGANNNC SDEELD TLMGLDKPD LNKELLFQT SHI LKGVGKKHG IKP-GTKVEKFLTETQKLFGEHGISGVSEDC LKCFQSIHCVANAKCGPCLRGPCITEEQNCIKKNCMQJLLE CIGKET IPNPDWKE
 • vac INSSVLGANNNC SDEELD TLMGLDKPD LNKELLFQT SHI LKGVGKKHG IKP-GTKVEKFLTETQKLFGEHGISGVSEDC LKCFQSIHCVANAKCGPCLRGPCITEEQNCIKKNCMQJLLE CIGKET IPNPDWKE

Fig. 6: Phylogenetic tree showing relationships between 22 filed samples and the vaccine. These data based on the sequences of *T. lestoquardi* TL2.

Table 1: primers used in this study and their sequences

Primers	Forward (5' - 3').	Reverse (5' -3').	Annealing temperature
TaD 6	ATGAAATTGACCGCTGGATTAGTCAGGC	CTATGAACCGCCCGAGGCTTCACC	52
TaD 2	GAAATTGACCGCTGGATTAG	CCGAGGCTTCACCTTTTT	50
TI 1-1:	GACCAGCCTCTTCTCCAACATT	CAGGTTTAGTGACTGGAGTGGTC	60
TI1-2	CGGAACTGATTTTGAAATTTGCTAT	GAGTGGTCTCTGGAGTTGGAGT	60

Table 2: Major allele frequency, gene diversity, number of alleles and polymorphic information content (PIC) of the seven mini and microsatellite markers used in this study.

Marker	Major Allele Frequency	Number of Allele	Gene Diversity	PIC
TS5	0.54	8.00	0.63	0.59
TS6	0.80	6.00	0.34	0.33
TS8	0.33	15.00	0.78	0.75
TS12	0.88	3.00	0.21	0.20
TS15	0.73	9.00	0.45	0.43
TS20	0.56	15.00	0.66	0.64
TS25	0.44	12.00	0.75	0.72
Mean	0.61	9.71	0.55	0.52