CD4 T cell responses to Theileria parva in immune cattle recognise a diverse set of parasite antigens presented on the surface of infected lymphoblasts

CD4 T cell responses to *Theileria parva* in immune cattle recognise a diverse set of parasite antigens presented on the surface of infected lymphoblasts

W Ivan Morrison *, Adriana Aguado * Ɛ, Tara A. Sheldrake *, Nicholas C. Palmateer †, Olukemi O. Ifeonu †, Kyle Tretina †, Keith Parsons ‡ κ, Emilio Fenoy ¶, Timothy Connelley *, Morten Nielsen ¶ #, and Joana C. Silva † ¥

* The Roslin Institute, Royal (Dick) School of Veterinary Studies, Easter Bush Campus, University of Edinburgh, Roslin EH25 9RG, UK

† Institute for Genome Sciences, University of Maryland School of Medicine, Baltimore, Maryland, USA

‡3 Institute for Animal Health, Compton, Berkshire UK

¶ Biotechnological Research Institute, National University of San Martin, San Martin, Aeres, Argentina

Department of Health Technology, Technical University of Denmark, 2800 Kongens Lyngby, Denmark

¥ Department of Microbiology and Immunology, University of Maryland School of Medicine, Baltimore, Maryland, USA

Ɛ Current contact address: adriana.aguado@hotmail.com

κ Current contact address: keithparsons54@btinternet.com

Corresponding author: Professor W Ivan Morrison – ivan.morrison@roslin.ed.ac.uk

Keywords: *Theileria parva*, cattle, vaccination, CD4 T cell, antigens, polymorphism

Running Title: Specificity of CD4 T cell responses to *Theileria parva*

This work was supported by a grant awarded by the Bill and Melinda Gates Foundation jointly with the UK Department for International development (DfID) [No. OPP1078791] and a Biotechnology and Biological Sciences Research Council (BBSRC) Institute ISP grant (grant number BB/J004227/1).

Abstract

Parasite-specific CD8 T cell responses play a key role in mediating immunity against *Theileria parva* in cattle (*Bos taurus*) and there is evidence that efficient induction of these responses requires CD4 T cell responses. However, information on the antigenic specificity of the CD4 T cell response is lacking. The current study used a high-throughput system for antigen identification using CD4 T cells from immune animals to screen a library of ~40,000 synthetic peptides representing 499 *T. parva* gene products. Use of CD4 T cells from 12 immune cattle, representing 12 class II MHC types, identified 26 antigens. Unlike CD8 T cell responses, which are focused on a few dominant antigens, multiple antigens were recognised by CD4 T cell responses of individual animals. The antigens had diverse properties, but included proteins encoded by two multi-member gene families – five haloacid dehalogenases and five subtelomere-encoded variable secreted proteins (SVSPs). Most antigens had predicted signal peptides and/or were encoded by abundantly transcribed genes, but neither parameter on their own was reliable for predicting antigenicity. Mapping of the epitopes confirmed presentation by DR or DQ class II alleles and comparison of available *T. parva* genome sequences demonstrated that they included both conserved and polymorphic epitopes. Immunisation of animals with vaccine vectors expressing two of the antigens demonstrated induction of CD4 T cell responses capable of recognising parasitised cells. The results of this study provide detailed insight into the CD4 T cell responses induced by *T. parva,* and identify antigens suitable for use in vaccine development.

Key points

- Multiple CD4 T cell antigens identified by screening a *T. parva* peptide library
- Antigens have diverse properties and include polymorphic and conserved proteins
- Parasite infection and viral-vector delivered antigen induce similar CD4 responses

Introduction

Intracellular protozoan parasites induce T cell-mediated immune responses that play a central role in development of immunity. However, the antigenic complexity of these parasites presents a significant challenge for identifying the antigens that mediate immunity and hence are candidates for vaccine development. One such pathogen is the tick-borne parasite *Theileria parva,* which causes one of the most economically important diseases affecting cattle in sub-Saharan Africa. The acute and frequently fatal nature of the disease in cattle results in high levels of mortality and presents major challenges for implementing control measures (1). Control relies predominantly on prevention of infestation with the tick vector and in some regions use of vaccination. The latter involves an infection and treatment regime using live parasites harvested from ticks and cryopreserved prior to use, administered along with a long-acting formulation of oxytetracycline (2). The most frequently used version of this vaccine incorporates three parasite isolates, to overcome parasite antigenic diversity (3), two of which have recently been shown to be nearly identical genome-wide (4). Although shown to be effective in the field (5), widespread application of this multi-component live vaccine has been hampered by the cumbersome processes involved in its production and quality control and the requirement of a cold chain for vaccine distribution. In some regions, concerns about potential risks of introducing the vaccine parasite strains into local tick populations have also resulted in reluctance to adopt vaccination. These shortcomings have led to efforts to develop alternative vaccines based on the use of defined parasite antigens (reviewed in 6).

Strategies adopted for development of subunit vaccines have built on knowledge of immune responses in animals immunised with live parasites. A large body of evidence indicates that immunity against *T. parva* is mediated by T cell responses against the intra-lymphocytic schizont stage of the parasite and that parasite-specific CD8 T cells play a central role in immunity (reviewed in 7). Notably, it has been possible to confer immunity against parasite challenge by adoptive transfer of purified actively responding CD8 T cells from immune to

naïve identical twin calves (8). Moreover, CD8 T cell responses in cattle immunised with a single parasite isolate are frequently strain-restricted, which has been shown to correlate with susceptibility to subsequent challenge with heterologous parasite strains (9). A striking feature of the parasite-specific CD8 T cell response is that the response in individual animals is focused on a few highly dominant antigens, which are often polymorphic (10). This immunodominance is critical in determining the strain restriction of immunity.

More recent work has successfully identified a series of *T. parva* antigens recognised by CD8 T cells from immune cattle (11, 12). However, although immunisation of cattle with some of these antigens, employing prime-boost protocols with recombinant viral vectors, successfully induced parasite-specific CD8 T cell responses, protection against parasite challenge was only achieved in a proportion of the immunised animals (11, 13).

The adoptive transfer studies referred to above utilised CD8 T cells collected during the active response of immune calves to challenge with *T. parva*. Hence, it remains unclear whether or not other components of the immune response are required for efficient activation of memory CD8 T cells following parasite challenge. Help from CD4 T cells is required for efficient induction of CD8 T cell responses against many viral infections (Reviewed in 14). Moreover, studies of responses to *Plasmodium* parasites using murine models have shown that the induction and maintenance of CD8 T cell responses to the pre-erythrocytic stage of the parasite are strongly dependent on the presence of specific CD4 T cells (15-17). *T. parva*-infected lymphoblasts express class II MHC on their surface and immune animals exhibit strong parasite-specific CD4 T cell responses against parasitized cells (7, 18). However, their role in immunity has received relatively little attention. Bovine CD8 T cells do not appear to produce IL-2 upon activation and *in vitro* studies have indicated that their ability to respond following antigen recognition is dependent on production of soluble growth factors by other cell types. Experiments using different combinations CD4 and CD8 T cells purified from genetically identical *T. parva*-immune or -naïve twin calves have demonstrated that optimal proliferation of the immune CD8 T cells required the presence of antigen-

specific CD4 T cells (19). Moreover, the results indicated that both cell types had to recognise antigen presented by the same *T. parva*-infected cells. These observations suggest that successful vaccination against *T. parva* with defined antigens requires induction of both CD4 and CD8 T cell responses of appropriate specificities and functional activities. Although a series of early studies identified crude antigenic fractions of parasitized cells recognised by *T. parva*-specific CD4 T cells (18, 20-22), the precise antigenic specificities of the responding T cells have not been defined.

Since the development of efficacious vaccines against *T. parva* is likely to require inclusion of CD4 as well as CD8 T cell antigens to provide robust protection, the current study set out to identify the parasite antigens recognised by CD4 T cell responses in cattle immunised against *T. parva* by infection and treatment. Antigens were identified by undertaking a high-throughput screen with CD4 T cells from immune cattle of diverse MHC types, using a library of peptides representing 499 parasite proteins (approximately 12.5 % of the parasite proteome). The results demonstrate that, in contrast to the narrow specificity of CD8 T cell responses, CD4 T cells from immune cattle recognise multiple antigens presented by parasite-infected cells and that such responses can readily be induced using viral vaccine vectors.

Methods

Animals and immunisation

The study utilised castrated male Holstein cattle between 12 and 24 months of age. All animals were selected by initially determining their MHC class I types using a combination of serological typing with monoclonal antibodies (23) and allele-specific PCR (24). Some of the animals, including two DR 011.01-homozygous animals, produced by father-daughter mating with sires of known genotype, were included. The class II DRB3 alleles expressed by the selected animals were identified by sequencing of cloned PCR products obtained from cDNA with DRB3-specific primers (25).

Animals used for antigen screening were immunised with the Muguga isolate of *T. parva* by infection with cryopreserved sporozoites and simultaneous treatment with a slow-release formulation of oxytetracycline, as described previously (2). Additional animals were immunised with selected antigens expressed individually in recombinant human adenovirus (hAd5) and vaccinia virus (modified Ankara strain - MVA) vectors. Animals received 2 x 10^9 IU of each adenovirus followed 8 weeks later by 5 x 10^8 PFU of each vaccinia virus, both administered intramuscularly.

T cell lines

All cultures of parasitized cells and T cells were conducted in RPMI 1640 medium supplemented with 10% foetal bovine serum, 2-mM L-glutamine, 5 x 10^{-5} M 2-mercapto-ethanol and penicillin and streptomycin. Cell lines infected with the Muguga stock of *T. parva*, from which the reference genome sequence was derived (26), were generated by infection of peripheral blood mononuclear cells (PBMC) *in vitro* with sporozoites (27). A parasite-specific CD4 T cell line was generated from each immune animal as described previously (28); briefly, PBMC were stimulated with irradiated autologous *T. parva*-infected cells three times at weekly intervals, followed by depletion of CD8 and γδ T cells by antibody-and-complement-mediated lysis. In some cases, the cultures were subjected to a

192 further positive selection step involving cell sorting of cells stained with a CD4-specific
193 monoclonal antibody (IL-A11) using a FACSaria cell sorter (Becton-Dickenson Biosciences,
194 Oxford, UK). Purified CD4 T cells (>98%) were maintained by stimulation at 7-10 day
195 intervals with equal numbers of irradiated *T. parva*-infected cells in medium containing 50
196 units/ml of recombinant human IL-2 (Chiron, Emryville, CA, USA)..

197 **Antigen screening**

198 A library of synthetic peptides representing 499 selected *T. parva* gene products was used
199 for antigen screening. The gene products included in the library are listed in Supplementary
200 Table S1. This library consisted of 40,921 18-mer peptides overlapping by 12 residues,
201 produced as a series of 819 pools each containing 50 contiguous peptides. Peptides were
202 supplied by JPT Peptide Technologies GmbH (Berlin, Germany) and were produced by their
203 Microscale system, which results in an additional glycine residue at the C-terminus of each
204 peptide. Antigen screening was conducted in 96-well round-bottomed plates by incubation of
205 each peptide pool with 2 x 10^4 CD4 T cells/well in a total volume of 150 ul of medium, each
206 peptide at a final concentration of 2ug/ml. CD4 T cells harvested 9-10 days after antigenic
207 stimulation were used for antigen screening to ensure low background levels of cytokine
208 secretion. Because of the large numbers of CD4 T cells required for each screen, they were
209 conducted in single wells. After incubation with peptide at 37^{o}C for 48 h, culture
210 supernatants collected from the wells were assayed for IFNγ, using a biological assay
211 performed as described previously (29), based on the ability of the IFNγ in culture
212 supernatants to up-regulate surface expression of MHC class II on Madin Darby Bovine
213 Kidney (MDBK) cells. The MDBK cells were harvested after 48 hours and class II MHC
214 expression assessed by flow cytometry following immunofluorescence staining with a class II
215 DR-specific monoclonal antibody (IL-A21). Results are expressed as the percentage of
216 MDBK cells expressing surface class II.

217 For each T cell line, the screening assay was repeated with a limited set of peptide pools
218 (Supplementary Table S2) comprising all pools that were positive in the initial screen, other

pools that contained peptides from antigens represented in the positive pools and an additional 15 randomly selected pools that gave negative results with all T cell lines. This secondary screen, as well as all subsequent assays of the positive antigens were conducted in duplicate wells. Only supernatants that resulted in at least 20 percentage points above the background levels of class II expression in both the primary and secondary screens were considered positive (all were statistically significant at $p < 0.01$ in the secondary screen). The 50 peptides in each pool sometimes originated from a single gene product, but in other cases were derived from two or occasionally three gene products. In the latter cases, subcomponents of the pools were re-screened in duplicate to determine which gene product gave rise to the positive response.

Epitope localisation and MHC restriction

For each positive peptide pool, the epitope regions within the parasite protein recognised by CD4 T cells were identified by repeating the screens with individual 18-mer peptides (overlapping by 12 residues).

Two assays were used to determine whether peptides were presented by class II DR or DQ MHC proteins: A first assay used Chinese hamster ovary (CHO) cells expressing either of two defined bovine DR alleles, namely DR 010:01 and DR 011:01, known to be expressed by MHC-homozygous animals used in the study. These cells were produced by transfecting CHO cells with a construct containing cDNAs encoding both the DR α and β chains linked by an oligonucleotide from foot and mouth disease virus, which encodes a self-cleaving peptide (30). Cells transfected with this construct were stained with Mab IL-A21 and positive cells purified by cell-sorting were expanded and, after checking for DR expression, aliquots of the cells were cryopreserved prior to use. For the T cell assay, resuscitated CHO cells were distributed in 96-well u-bottomed plates at 10^4/well and duplicate wells were incubated with a range of concentrations of peptide (2ug to 100ng/ml) or medium for two hours. The cells were then washed twice with culture medium and T cells added to the wells at 10^4/well. After

incubation at 37°C for 48 hours, supernatants were harvested and assayed for IFNγ using the biological assay.

A second approach involved repeating the screening assay as described above with the individual positive peptides, with or without addition of monoclonal antibodies specific for bovine DR (Mab IL-A21) or DQ (Mab CC158). Because of the use of class II-specific antibodies in these experiments, IFNγ release could not be measured using the biological assay and therefore was measured using a sandwich ELISA as described previously (31).

Properties of gene products

Antigenic proteins need to have access to MHC-processing pathways in the host cell. As such, we identified genes whose products are targeted to the secretory pathway or are otherwise associated with membranes. TargetP 1.1 (32, 33) was used to identify proteins predicted to be targeted to the secretory pathway with high reliability (reliability classes 1). Proteins were predicted to be glycosylphosphatidylinositol (GPI)-anchored using GPI-SOM (34) and PredGPI (35). The presence of five or more transmembrane helices, which is a strong indicator of a transmembrane protein, was determined with TMHMM (36, 37). Selected proteins were all produced from genes transcribed in the schizont stage of the parasite and transcript abundance was assessed from the single published dataset on the transcriptome of *T. parva* (38).

Polymorphism

For each locus, the orthologous allele was obtained from the draft genome assembly of 17 *T. parva* strains isolated from cattle from across the range of the distribution of the parasite. The genome sequence of each strain was generated as described before (39). When alleles were mostly or completely missing from the draft assembly, the strain was removed from the respective dataset. When only a few nucleotides were missing in the 3′ or 5′ end, the sequence was kept in the dataset and the missing nucleotides were assumed to be identical to the haplotype that was otherwise identical to it. The nucleotide sequences were aligned

using Muscle (40) and manually curated in Mesquite (41). Nucleotide diversity per site averaged across each gene was estimated for silent sites (synonymous; average number of synonymous differences per synonymous site, π_S) and amino acid-changing sites (non-synonymous; average number of non-synonymous differences per non-synonymous site, π_{NS}), using DnaSP (42). In addition, π_{NS} was estimated across sliding windows of non-synonymous sites, with width and step, respectively, of 30 and 5 non-synonymous sites. Regions of low complexity and micro- and mini-satellites were identified using Tandem Repeats Finder (43). Plots of nucleotide diversity were drawn using R (44).

Statistical analysis

Data are presented as means and standard deviations (SD). Statistical analyses were performed using Student's T test. Statistically significant differences compared with the control are indicated by asterisks - p values of <0.05 = * and <0.01 = **.

Results

Selection of parasite proteins for generating the peptide library

The approach used for selection of the parasite gene products to produce the peptide library for antigen screening is illustrated in Figure 1 and the content of the library is listed in Supplementary Table S1. Selection of proteins was based predominantly on the possession of a predicted signal peptide, indicating likely secretion from the intracellular parasite (45), and/or high abundance of transcription of the genes in infected lymphoblasts. Ten proteins previously shown to be CD8 T cell antigens were included (11, 12). A re-annotated version of the *T. parva* genome was used to produce the parasite proteome sequence (46), from which 420 proteins highly likely targeted for the secretory pathway were selected; 39 of these have a likely GPI anchor and 47 have five or more transmembrane helices, with some overlap among those sets (Fig 1). An additional 69 proteins were selected based on abundance of transcription. .

Screening of the peptide library

A *T. parva*-specific CD4 T cell line from each of the 12 immunised cattle was used for antigen screening. The donor cattle represented 12 different class II DRB3 alleles and included animals that were homozygous for 5 different MHC haplotypes (Table 1). Pairs of cattle homozygous for DR 010.01 and DR 011.01 had been produced by father-daughter mating and although they were not typed for class II DQ are also highly likely to share the same DQ alleles. Screening of the peptide library (819 peptide pools representing 499 parasite proteins) with these CD4 T cell lines resulted in the identification of positive peptide pools in all animals; 32 of the pools yielded positive results. A secondary screen of all positive pools was used to confirm positive results and further screens of the sub-components of pools that contained peptides from more than one parasite protein were carried out. The results identified a total of 26 positive parasite proteins, representing 5.2% of the gene products screened. Representative results obtained with five CD4 T cell lines in the secondary screen (pools listed in Supplementary Table S2) are shown in Figure 2 and a

summary of the gene products recognised by each animal is shown in Table 1. CD4 T cell lines from individual animals recognised between 2 and 10 antigens (mean of 6.2). There was considerable variation between animals in the antigenic specificity of the detectable response. None of the antigens were recognised by all animals; some were detected by only one or two T cell lines, whereas responses to others were detected in up to 8 of the 12 lines assayed. Within the cell lines tested, two pairs of animals homozygous for DR 010:01 and DR 011:01 each shared responses to 5 antigens (5/11 = 45.5% and 5/8 = 62.5% of the antigens identified respectively), which tended to be the strongest responses detected in these animals, confirming that specificity was influenced by MHC type. Differences in detection of responses to some antigens by the same DR genotypes may relate to differences in the overall antigenic specificity of the T cell lines, resulting in failure to detect some weak specificities, and/or differences in the T cell receptor repertoire of the animals.

Characteristics of the target antigens

As shown in Table 2, which summarises the main features of the antigens, they are encoded by genes distributed across all four chromosomes and exhibit a range of properties. They include six gene products known to induce CD8 T cell responses, four of which (Tp1, Tp2, Tp7 and Tp9) were reported previously (11, 12) and the remaining two (Tp32 and Tp33) identified in recent antigen screens (N. D. MacHugh and W. I. Morrison unpublished data). Several members of two multi-member gene families, namely the haloacid dehalogenases (HAD) and the subtelomere-encoded variable secreted proteins (SVSP), are represented in the antigen panel. All five HAD proteins and five of the 14 SVSP proteins, included in the peptide library, were identified as antigens.

The panel of gene products included in the peptide library used for antigen screening was biased towards abundantly expressed genes and proteins predicted to be secreted from the parasite into the host cell cytoplasm or be otherwise exposed to the host. Reflecting this bias, a majority of the antigens recognised by CD4 T cells (19/26) have a predicted signal peptide; however, seven antigens lacked a signal peptide. Only three of the antigens

predicted to be secreted (Tp18, Tp19 and Tp36) showed evidence of a membrane anchor – the former having a predicted transmembrane domain and the latter two a predicted GPI anchor. Biological studies have confirmed that these three proteins (previously referred to as PIM, p104 and gp34) are localised to the schizont surface (47-49). Tp19 (p104) and Tp36 (gp34), as well as the *T. annulata* orthologue of another protein (Tp33), have been found to be within a complex of host and parasite proteins that associate with host microtubules on the surface of the schizont (50). The abundance of transcription of the antigen-encoding genes showed wide variation. While some of the antigens ranked very highly in their levels of transcription, including those selected on the basis of transcript abundance, others fell within the 20% genes with lowest levels of expression. In summary, no single parameter could be used reliably to predict antigenicity.

Epitope identification and MHC restriction

Experiments to identify the target epitopes were undertaken to provide reference information on the fine specificity of the response, with which to compare vaccine-induced responses. This work focused on three DR types (010:01, 011:01 and 016:01), for which the CD4 lines were derived from class II-homozygous animals. These included pairs of animals homozygous for DR 010:01 and DR 011:01. Examples of results obtained by screening 5 of the peptide pools are shown in Figure 3. The five CD4 T cell lines examined recognised 21 of the antigens identified. Screening of individual peptides for 19 of these antigens identified one or occasionally two or three epitope regions within each positive pool. Where reactions to two contiguous (18-mer) peptides were detected, testing of truncated peptides allowed each epitope to be localised to within a region of 18 amino acids or less. The results of these assays, summarised in Table 3, resulted in identification of 11-14 epitopes for each of the 3 MHC types examined. However, the HAD antigens recognised by animal 2824 contained an epitope that showed a high level of identity between all five antigens; the epitope sequences in Tp17 and Tp22 were identical and the sequences of the other three (Tp14, Tp21 and Tp22) differed at one, six and three of the 15 amino acid residues respectively (Table 3). The

same epitopes were recognised by animal 598 (data not shown), which also expresses the DR 016:01 allele. Since four of these proteins were only recognised by these two DR 016:01-positive animals, it was not possible to discern whether all or only some of the antigens were responsible for inducing the CD4 T cell responses.

As expected, MHC type influenced the epitopes that were recognised. First, where antigens were recognised by animals of identical MHC types (DR 010:01 or DR 011:01), the CD4 T cell lines recognised the same epitope; conversely, within the five antigens recognised by two or three of the DRB3 types (Tp16, Tp9, Tp20, Tp18 and Tp15) the majority of the epitopes recognised (13/17) were unique to a single DRB3 type. However, in four instances the same peptide was recognised by animals with two different class II types (Tp9 $_{280\text{-}295}$ by DR 011:01 and DR 016:01; Tp9 $_{292\text{-}310}$ by DR 010:01 and DR 016:01; Tp18 $_{97\text{-}114}$ by DR 010:01 and DR 011:01; Tp20 $_{103\text{-}120}$ by DR 010:01 and DR 016:01).

The results obtained by testing the ability of CHO cells transfected with the DR 010:01 or DR 011:01 class II MHC alleles to present peptides identified on the respective MHC backgrounds are shown in Figure 4 and summarised in Table 3. Thirteen of the 22 identified epitope specificities (seven on the DR 010:01 and six on the DR 011:01 backgrounds) were recognised when presented by CHO cells expressing the relevant DR allele. The results obtained by testing the ability of DR- and DQ-specific monoclonal antibodies to inhibit epitope recognition confirmed that those epitopes that were not recognised on DR-transfected CHO cells were inhibited by the DQ-specific antibody but not by the DR-specific antibody (Figure 5). Epitopes identified on the DR 016:01 background were tested only using the latter assay; recognition of ten of the eleven epitopes tested was inhibited by the DR-specific antibody but not by the DQ-specific antibody. The converse result was obtained with the remaining epitope (Data not shown).

Hence, overall twenty-two of the unique epitopes analysed were found to be DR-restricted and eight were DQ-restricted. Of the four epitopes that were recognised by two different class II types, two ($Tp9_{280\text{-}295}$ and $Tp20_{103\text{-}120}$) were found to be DR-restricted and two ($Tp9_{292\text{-}}$

$_{310}$ and Tp18$_{97\text{-}114}$) DQ-restricted. In separate studies, the two DR-restricted epitopes were found to have a high predictive score for binding to both DR alleles (M. Nielsen, unpublished data).

Polymorphism of the identified antigens

The identification of epitopes recognised by the CD4 T cell response in individual animals, coupled with recently generated genome sequences from 17 parasite isolates obtained from cattle, allowed us to estimate the level of polymorphism across the length of the target antigens, including regions encoding the CD4 epitopes. The target antigens ranged from some of the most polymorphic in the *T. parva* genome (39), such as Tp2 and Tp9, to those that are completely conserved (Table 2). They vary greatly in length and in the presence of low complexity regions or repeat regions (Figure 6). Some of the identified epitopes were located in antigen regions with intermediate to high variability in the parasite population (e.g. Tp2, Tp9, Tp15 – Figure 6), whereas others were in conserved regions of antigens (e.g. Tp1, Tp15, Tp16, Tp30 – Figure 6). In a few cases, epitopes were found in regions with length variation (e.g. Tp9 – Figure 6), such that they are absent in some strains, or those adjacent to repeats (e.g. Tp9, Tp15 – Figure 6), suggesting that protein conformation nearby may alter access to the epitopes. Overall, the CD4 T cell responses of all animals were found to recognise both conserved and polymorphic epitopes.

Specificity of CD4 T cell responses induced by prime-boost immunisation with selected antigens

The context in which antigens are delivered can sometimes alter the selection of epitopes (51), with obvious consequences for the efficacy of vaccination. To determine whether *T. parva* antigens delivered in vaccine vectors induce CD4 T cell responses of similar specificities to those elicited by live parasites, four animals expressing DRB3 010:01 and/or DRB3 011:01 were immunised using a prime-boost protocol, with the Tp9 and Tp15 antigens expressed in recombinant adenovirus (Ad) and vaccinia (MVA) vectors. CD4 T cell lines generated from all four animals, by stimulation of T cells with autologous *T. parva*-infected

cells, recognised Tp9 and Tp15 (Table 4). Eight epitopes had been identified in these two antigens during antigen screening with T cell lines expressing DRB3 010:01 or DRB3 011:01. Epitope mapping of the CD4 response in one of the viral vector-immunised animals, which expressed both DRB3 010:01 and DRB3 011:01, revealed significant responses of variable magnitude to 6 epitopes, four in Tp9 and two in Tp15. These epitopes were identical to those recognised by *T. parva*-immune animals (Table 5), confirming the capacity of the vectored antigens to induce CD4 T cell responses to epitope specificities displayed on infected cells.

The same four animals were concurrently immunised with recombinant Ad/MVA viruses expressing a third antigen (Tp10) (12) known to induce CD8 T cell responses but not identified as a CD4 T cell target antigen by screening in the current study. All 4 four animals generated a CD4 T cell response to Tp10, detectable by stimulation *in vitro* with *T. parva*-infected cells (Table 4). Hence, this antigen is presented by parasitised cells and can be immunogenic (ie. outwith the context of natural infection), even though it was not identified as a CD4 T cell antigen in the peptide screens.

Data on CD8 responses to these antigens were also obtained from 3 of the animals. Significant responses were detected in all 3 animals for Tp9 and in two animals for Tp10, but none of the animals responded to Tp15 (Table 4).

Discussion

The evidence that CD8 T cells play a key role in immunity to *T. parva* has led to efforts to develop a vaccine using CD8 T cell target antigens. CD4 T cell responses are also likely to be required to provide help for efficient induction and recall of CD8 T cell responses. Although infection with *T. parva* induces vigorous CD4 T cell responses against infected cells, there is a paucity of information on their specificity. The results of the current study demonstrate that, in contrast to the profound immunodominance of CD8 T cell responses to *T. parva* in immune cattle, whereby a large proportion of the response in individual animals is focused on one or two antigens (10), CD4 T cell responses recognise multiple antigens, including both polymorphic and conserved antigens. Antigen screens using CD4 T cells from twelve immune cattle of different MHC class II types identified 26 antigens, with up to eleven antigenic specificities detected in a single animal. As would be expected, there was variation between animals in the antigens that were recognised, with no single antigenic specificity detected in all animals.

Despite the multiple antigens identified in this study, the findings almost certainly underestimate the diversity of antigenic specificities presented by *T. parva*-infected cells. Since the peptide library represented only about 12.5% of *T. parva* genes, almost all of which are transcribed by the schizont stage of the parasite (46), the remaining proteome is likely to include additional unidentified antigens. Moreover, the assay used for antigen screening may well fail to detect T cell specificities present at low frequency. The finding that immunisation of calves with viral vectors expressing the Tp10 CD8 antigen, which was not detected by the antigen screens, induced CD4 responses reactive with *T. parva*-infected cells, clearly illustrates the presence of additional immunogenic antigenic specificities on infected cells.

The strategy employed in the current study for antigen screening was based on knowledge of the biology of the parasite and how parasite antigens are likely to be recognised by CD4 T cells *in vivo*. Unlike many other apicomplexan parasites, *T. parva* is found free within the

host cell cytoplasm, having escaped from the endocytic vacuole shortly after initial cell invasion by the sporozoite (52). The parasite establishes an intimate relationship with the host cell, which results in activation and transformation of infected host cells and inhibition of apoptosis (53). Synchronous division of the parasite and host cell (54, 55) results in clonal expansion of the infected cell population, thus ensuring that the parasites are retained intracellularly throughout this stage of development. The transformed state of the infected cells also results in expression of surface DR and DQ class II MHC proteins, as well as co-stimulatory proteins such as CD80, CD86 and CD40 (W I Morrison, unpublished data). These properties of infected cells favour endogenous routes of processing of parasite antigens and allow direct recognition of infected cells by parasite-specific CD4 T cells, in the absence of added antigen-presenting cells. Hence, the current study used CD4 T cell lines generated from immune cattle by stimulation *in vitro* with intact parasitized cells for antigen screening. Holstein cattle, which are highly susceptible to *T. parva* and are extensively used for milk production in endemic areas of Africa, were used in the study. The class II MHC types of the animals included the most frequently expressed DR alleles in Holsteins; a recent study of 331 Holstein cattle identified 15 DR alleles, 10 of which were represented in the current study and included the 7 most frequently detected alleles (56).

T. parva schizont antigens targeted by CD4 T cells on infected cells must access the cytoplasm of the infected cells, prior to entering class II MHC processing pathways. Consequently, proteins secreted from the parasite or released from the parasite by other mechanisms represent likely candidate antigens. Studies of antigen processing of a number of human and murine intracellular pathogens indicate that cytoplasmic proteins can be transferred into the endosomal pathway for class II processing either following autophagy within the cytoplasm or as peptides generated by degradation of parasite proteins within the proteasome (57, 58). The involvement of the latter pathway has been demonstrated for some CD4 T cell responses to influenza virus-infected cells (59-61). In preliminary studies, we have found that recognition of the Tp9 antigen on infected cells by specific CD4 T cell

clones is inhibited by pre-incubation of the infected cells with the proteasome inhibitor Epoxomicin (A. Aguada and W. I. Morrison, unpublished data), suggesting that this endogenous pathway is involved in processing *T. parva* antigens.

A high-throughput antigen screen using a peptide library representing 499 parasite proteins was used for antigen identification. As discussed above, given the requirement of antigens to access the host cell cytoplasm, selection of parasite proteins for generating the peptides was biased, firstly towards proteins predicted to have a signal peptide and therefore likely to be secreted from the parasite and secondly towards genes that are abundantly transcribed in the schizont stage of the parasite and whose products may access the cytosol by other undefined routes. The selected parasite proteins also included those known to elicit antibody or CD8 T cell responses against this stage of the parasite in immune cattle (11, 47, 62, 63).

Only 6 of the 14 CD8 antigens identified to date (Tp1, Tp2, Tp7, Tp9, Tp32 and Tp33), were also identified as CD4 T cell antigens. Previous studies have shown that Tp1 and Tp2 are highly dominant CD8 antigens in class I A18 and A10 animals respectively, yet these antigens were not recognised, to a detectable level, by CD4 T cells from animals of these class I genotypes. Thus, the CD4 and CD8 T cell responses in these animals have different dominant antigenic specificities. A recent study of prime-boost immunisation of class I A18+ animals with Tp1 expressed in adeno and vaccinia virus vectors observed Tp1-specific CD4 responses during the active response to the vaccinia boost. However, since the T cell responses were monitored using Tp1 antigen added to antigen-presenting cells, the findings did not confirm recognition of the antigen presented by infected cells. Nevertheless, the findings suggest that, as with our results with vector-induced responses to Tp10, vaccination with defined antigens may elicit specificities that are not readily detected during responses to infection.

All three proteins previously shown to induce antibody responses to the schizont stage of the parasite in infected animals (PIM, p104 and p150 – now respectively, Tp18, Tp19 and Tp20) (47, 62, 63) were identified as CD4 antigens, but only in a subset of the animals. Notably,

CD4 T cell responses to the Tp18 antigen (PIM), which is an abundant schizont protein and induces highly dominant antibody responses in all immune animals (47), were detected in only two of the animals examined. This implies that additional exogenous routes of processing of some antigens released from infected cells may be involved *in vivo* in generating CD4 T cell responses, due to more efficient display of antigen on antigen-presenting cells. Although a majority of the identified antigens (19/26) had predicted signal peptides, this appeared to reflect the composition of the peptide library, 84% of which was generated from proteins with a predicted peptide sequence. Abundance of transcription of the antigen-encoding genes was highly variable; while half of the antigens (13/26) were encoded by the top 10% most abundantly transcribed genes, others were in the lower 50%. Because transcript abundance was used to select proteins that lacked a signal sequence, the antigens identified within this subset were encoded by some of the most abundantly transcribed genes (10/11 ranked in the top 100). Although neither of these parameters on their own appeared reliable for predicting CD4 cell target antigens, the large number of antigens detected in the current study suggest that both properties may have been useful for antigen identification. However, screening of a large unbiased panel of parasite gene products would be required to address this question.

While the identified CD4 antigens included proteins with diverse properties, several members of the haloacid dehalogenase (HAD) and the subtelomere-encoded variable secreted protein (SVSP) families of proteins were identified. All five HAD proteins included in the peptide library were identified as antigens. However, epitope mapping revealed a CD4 T cell epitope that was highly conserved between the five HAD proteins and, since this was the only epitope identified in four of the antigens, it was unclear whether only some or all of these antigens were responsible for inducing the T cell responses. The HAD family of proteins are phosphatases found in a wide range of prokaryotic and eukaryotic organisms, identifiable by the presence of conserved sequence motifs (64). *Plasmodium falciparum* has 23 HAD-encoding genes, two of which have been implicated as targets for the experimental

562 therapeutic compound forsidomycin, which targets the methyl-erythrotol phosphate pathway
563 of the apicomplast organelle (65). However, unlike the *T. parva* HAD proteins identified as
564 antigens, these *Plasmodium* proteins do not possess signal peptides and hence the former
565 are likely to have different biological functions, potentially involving interaction with host
566 proteins. The SVSP family, which has only been described in *Theileria*, consists of 85
567 members encoded by genes located in the telomeric regions of all four nuclear
568 chromosomes (66, 67), 14 of which were included in the peptide library. They encode a
569 short, conserved N-terminal region including a signal peptide, followed by a QP-rich central
570 region and a conserved C-terminus. Some contain a functional nuclear localisation signal
571 and some are expressed by only a proportion of cells in cultures of infected cells. However,
572 as with the HAD proteins, the functions of SVSPs remain unclear. Further studies are
573 required to explore whether additional members of the SVSP family, not included in the
574 peptide library, are recognised by parasite-specific T cells.

575 The results of the present study provide new insight into the antigenic basis of protective
576 immune responses against *T. parva* and have important implications for efforts to develop a
577 subunit vaccine. Early studies of *T. parva*-specific CD4 T cell clones revealed that some
578 clones were parasite strain-restricted (68), indicating polymorphism of the target antigens.
579 Moreover, some of the antigens identified in the current study, most notably Tp2 and Tp9,
580 are known to be highly polymorphic and result in strain-specificity of CD8 T cell responses
581 (12, 69, 70). However, many of the CD4 T cell antigens showed no or only limited
582 polymorphism. Hence, given the broad antigen specificity of CD4 T cells detected in
583 individual animals, it is unlikely that the CD4 T cell response contributes to the observed
584 parasite strain restriction of immunity to *T. parva*. Even in situations where polymorphic
585 antigens are responsible for a substantial component of the response, the panmictic
586 population structure of *T. parva* (71) arising from frequent sexual recombination during tick
587 passage (reviewed in 7), which allows independent segregation of alleles of different
588 antigens, favours cross-reactivity of the CD4 T cell responses to different parasite isolates.

With regard to vaccination, the findings have provided valuable information on which to base selection of antigens suitable for vaccine development. As discussed earlier, there is evidence that parasite-specific CD4 T cell responses are required for efficient activation of memory CD8 T cells and *in vitro* studies suggest that such help requires antigens to be displayed on the same antigen-presenting cells for recognition by CD4 and CD8 T cells (19). This is clearly the case for *T. parva*-infected cells, which express both class I and class II MHC proteins on their cell surface allowing activation of both CD4 and CD8 T cell responses following infection. The evidence that both CD8 and CD4 antigens undergo intracellular (endogenous) routes of antigen processing in infected cells implies that the same antigen delivery systems can be used in a subunit vaccine to generate both CD4 and CD8 T cell responses. In this regard, it is significant that immunisation of calves with recombinant adeno and vaccinia viruses vectors expressing two of the identified antigens were found to induce both CD4 and CD8 T cell responses reactive with parasitized cells and that the epitopes recognised by the CD4 T cells were identical to those induced by infection with *T. parva.* While induction of CD8 T cell responses by antigens expressed in viral vectors could theoretically utilise help provided by CD4 T cells specific for the vector viruses, efficient recall of such responses following parasite challenge will require vaccine-induced CD4 T cells responses against epitopes presented by *T. parva*-infected cells. The present study has shown that *T. parva*-infected cells display CD4 epitopes from a range of different parasite proteins. Data on polymorphism will permit selection of antigens with conserved amino acid sequences, to avoid strain specificity of CD4 T cell responses. Further studies are required to determine the capacity of candidate antigens to induce CD4 T cell responses in animals of different MHC genotypes when delivered in suitable vaccine vectors. Although responses of immune animals to most of the antigens in the present study were only detected in certain MHC genotypes, it is possible that immunisation with vaccine vectors will reveal immunogenicity in a wider range of genotypes. Similar investigations are ongoing to identify CD8 antigens with these properties. If different antigens prove optimal for induction of CD4

and CD8 T cell responses, it will be important to determine whether they can be delivered separately or need to be incorporated into the same vaccine vector for optimal results.

In conclusion, this study has clearly demonstrated that CD4 T cell responses of cattle to *T. parva* recognise multiple antigens presented on the surface of parasitized cells. Coupled with information on sequence polymorphism of the antigens and evidence that viral vaccine vectors expressing the antigens are capable of generating responses with the same fine antigenic specificity as those induced in infected cattle, the results provide a valuable resource for further studies to develop a subunit vaccine.

Acknowledgements: The authors wish to thank Professor Sarah Gilbert, The Edward Jenner Institute, Oxford University for producing recombinant vaccinia and adeno viruses, and JPT Peptide Technologies GmbH for advice on the peptide library. We are also grateful to Christina Vrettou for expert technical assistance. We wish to dedicate this publication to Dr Niall MacHugh, who initiated some of the early experiments reported in this study but unfortunately passed away in 2017.

References

1. Irvin, A. D. and W. I. Morrison. 1987. Immunopathology, immunology and immunoprophylaxis of *Theileria* infections. In: *Immune Responses in Parasitic Infections: Immunology, Immunopathology and Immunoprophylaxis,* Vol. III. Protozoa. L. Soulsby, ed. CRC Press Inc, Baton Rouge, Florida, *pp. 223-274*.
2. Radley, D. E., C. G. D. Brown, M. J. Burridge, M. P. Cunningham, I. M. Kirimi, R. E. Purnell and A. S. Young. 1975. East Coast fever: 1. Chemoprophylactic immunization of cattle against *Theileria parva* (Muguga) and five theilerial strains. *Vet. Parasitol. 1: 35-41.*
3. Radley, D. E., C. G. D. Brown, M. P. Cunningham, C. D. Kimber, F. L. Musisi, R. C. Payne, R. E. Purnell, S. M. Stagg and A. S. Young. 1975. East Coast fever: 3. Chemoprophylactic immunization of cattle using oxytetracycline and a combination of theilerial strains. *Vet. Parasitol. 1: 51-60*
4. Norling, M., R. P. Bishop, R. Pelle, W. Qi, S. Henson, E. F. Drábek, K. Tretina, D. Odongo, S. Mwaura, T. Njoroge, E. Bongcam-Rudloff, C. A. Daubenberger and J. C. Silva. 2015. The genomes of three stocks comprising the most widely utilized live sporozoite *Theileria parva* vaccine exhibit very different degrees and patterns of sequence divergence. *BMC Genomics 16: 729*..
5. Di Giulio, G., G. Lynen, S. Morzaria, C. Oura and R. Bishop. 2009. Live immunization against East Coast fever–current status. *Trends Parasitol. 25: 85-92*.
6. Nene, V. and W. I. Morrison. 2016. Approaches to vaccination against *Theileria parva* and *Theileria annulata. Parasite Immunol. 38: 724-734.*
7. Morrison, W. I., T. Connelley, J. D. Hemmink and N. D. MacHugh. 2015. Understanding the basis of parasite strain-restricted immunity to *Theileria parva*. *Annu. Rev. Anim. Biosci. 3: 397-418.*
8. McKeever, D. J., E. L. Taracha, E. L. Innes, N. D. MacHugh, E. Awino, B. M. Goddeeris and W. I. Morrison. 1994. Adoptive transfer of immunity to *Theileria parva* in the CD8+ fraction of responding efferent lymph. *Proc. Natl Acad. Sci USA 91: 1959-1963.*
9. Taracha, E. L., B. M. Goddeeris, A. J. Teale, S. J. Kemp and W. I. Morrison. 1995. Parasite strain specificity of bovine cytotoxic T cell responses to *Theileria parva* is determined primarily by immunodominance. *J. Immunol. 155: 4854-4860*.
10. MacHugh, N. D., T. Connelley, S. P. Graham, R. Pelle, P. Formisano, E. L. Taracha, S. A. Ellis, D. J. McKeever, A. Burrells and W. I. Morrison. 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*.

11. Graham, S. P., R. Pelle, Y. Honda, D. M. Mwangi, N. J. Tonukari, M. Yamage, E. J. Glew, E. P. de Villiers, T. Shah, R. Bishop, E. Abuya, E. Awino, J. Gachanja, A. E. Luyai, F. Mbwika, A. M. Muthiani, D. M. Ndegwa, M. Njahira, J. K. Nyanjui, F. O. Onono, J. Osaso, R. M. Saya, C. Wildmann, C. M. Fraser, I. Maudlin, M. J. Gardner, S. P. Morzaria, S. Loosmore, S. C. Gilbert, J. C. Audonnet, J.C., P. van der Bruggen, V. Nene and E. L. Taracha, E.L. 2006. *Theileria parva* candidate vaccine antigens recognized by immune bovine cytotoxic T lymphocytes. *Proc. Natl Acad. Sci. U S A. 103: 3286-3291.*

12. Hemmink, J. D., W. Weir, N. D. MacHugh, S. P. Graham, E. Patel, E. Paxton, B. Shiels, P. G. Toye, W. I. Morrison and R. Pelle. 2016. Limited genetic and antigenic diversity within parasite isolates used in a live vaccine against *Theileria parva*. *Int. J. Parasitol. 46: 495-506*.

13. Svitek, N., R. Saya, E. Awino, S. Munyao, R. Muriuki, T. Njoroge, R. Pellé, N. Ndiwa, J. Poole, S. Gilbert, V. Nene and L. Steinaa. 2018. An Ad/MVA vectored *Theileria parva* antigen induces schizont-specific $CD8^+$ central memory T cells and confers partial protection against a lethal challenge**.** *NPJ Vaccines. 3: 35.*

14. Swain, S. L., K. K. McKinstry and T. M. Strutt. 2013. Expanding roles for CD4+ T cells in immunity to viruses. *Nat. Rev. Immunol. 12: 136-148.*
15. Tsuji, M. and F. Zavala. 2003. T cells as mediators of protective immunity against liver stages of *Plasmodium. Trends Parasitol. 19: 88-93.*
16. Tse, S-W., A. J. Radtke and F. Zavala. 2011. Induction and maintenance of protective CD8+ T cells against malaria liver stages: Implications for vaccine development. *Mem. Inst. Oswaldo Cruz 106: 172-178.*
17. Crispe I. N. 2014. APC licensing and CD4+T cell help in liver-stage malaria. *Front. Microbiol.5: 617.*

18. Baldwin, C. L., K. P. Iams, W. C. Brown and D. J. Grab. 1992. *Theileria parva*: Helper and cytotoxic T-cell clones react with a schizont-derived antigen-associated with the surface of *Theileria parv*a-infected lymphoblasts. *Exp. Parasitol. 75: 19-30*.

19. Taracha, E. L., E. Awino and D. J. McKeever. 1997. T helper requirements in *Theileria parva*-immune and –naïve CTL precursors. *J. Immunol. 159: 4539-4545.*

20. Brown, W. C., C. Sugimoto and D. J. Grab. 1989. *Theileria parva*: Bovine helper T cell clones specific for both infected lymphocytes and schizont membrane antigens. *Exp. Parasitol. 69: 234-248.*

21. Brown, W. C., J. D. Lonsdale-Eccles, J. C. DeMartini and D. J. Grab. 1990. Recognition of soluble *Theileria parva* antigen by bovine helper T cell clones: Characterisation and partial purification of the antigen. *J. Immunol. 144: 271-277.*

22. Grab, D. J., C. L. Baldwin, W. C. Brown, E. A. Innes, J. D. Lonsdale-Eccles and Y. Verjee. 1992. Immune CD4+ T cells specific for *Theileria parva*-infected lymphocytes recognise a 24 kilodalton protein. *Inf. Immun. 60: 3892-3896.*

23. Ellis, S. A., W. I. Morrison, N. D. MacHugh, J. Birch, A. Burrells and M. J. Stear. 2005. Serological and molecular diversity in the cattle MHC class I region. *Immunogenetics 57: 601-606.*

24. Ellis, S. A., F. A. Staines, M. J. Stear, E. J. Hensen and w. I, Morrison. 1998. DNA typing for BoLA class I using sequence-specific primers (PCR-SSP). Eur. J. Immunogenet. 25: 365-370.

25. Baxter, R., N. Hastings, A. Law and E. J. Glass. 2008. A rapid and robust sequence based genotyping method for BoLA-DRB3 alleles in large numbers of heterozygous cattle. *Anim. Genet. 39: 561-563*.

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

27. Brown, C. G., D. A. Stagg, R. E. Purnell, G. K. Kanhai and R. C. Payne. 1973. Infection and transformation of bovine lymphoid cells i*n vitro* by infective particles of *Theileria parva. Nature 245: 101-103*.

28. Goddeeris, B. M. and W. I. Morrison. 1988. Techniques for the generation, cloning and characterisation of bovine cytotoxic T cells specific for the protozoan *Theileria parva*. *J. Tiss. Cult. Meth. 11: 101-110.*

29. Ballingall, K. T., D. M. Mwangi, N. D. MacHugh, E. L. Taracha, P. Totte and D. J. McKeever. 2000. A highly sensitive non-radioactive assay for T cell activation in cattle: applications in screening for antigens recognised by CD4+ and CD8+ T cells. *J. Immunol. Meth. 239: 85–93.*

30. Ryan, M. D. and J. Drew. 1994. Foot and mouth disease virus 2A oligopeptide mediated cleavage of an artificial polyprotein. *EMBO J. 13: 928-933.*

31. Hart, J., N. D. MacHugh, T. Sheldrake, M. Nielsen and W. I. Morrison. 2017. Identification of immediate early gene products of bovine herpes virus 1 (BHV-1) as dominant antigens recognized by CD8 T cells in immune cattle. *J. Gen. Virol. 98: 1843-1854.*
32. Nielsen H., J. Engelbrecht, S. Brunak and G. Von Heijne 1997. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng. 10: 1–6.*
33. Emanuelsson O., H. Nielsen, S. Brunak and G. von Heijne. 2000. Predicting subcellular localization of proteins based on their N-terminal amino acid sequence. *J. Mol. Biol. 300: 1005–1016*.
34. Fankhauser N. and P. Maser. 2005. Identification of GPI anchor attachment signals by a Kohonen self-organizing map. *Bioinformatics 21: 1846–1852*.
35. Pierleoni A., P. L. Martelli and R. Casadio. 2008. PredGPI: a GPI-anchor predictor. *BMC Bioinformatics, 9: 392*.
36. Sonnhammer E. L., G. von Heijne and A. Krogh. 1998. A hidden Markov model for predicting transmembrane helices in protein sequences. *Proc. Int. Conf. Intell. Syst. Mol. Biol. 6: 175–182*.
37. Krogh A., B. Larsson, G. von Heijne and E. L. Sonnhammer. 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol. 305: 567–580*.
38. Tretina, K., R. Pelle and J. C. Silva. 2016. Cis regulatory motifsand antisense transcriptional controlin the apicomplexan Theileria parva. *BMC Genomics 17: 128*.
39. Palmateer, N., K. Tretina, J. Orvis, I. Olukemi, J. Crabtree, E. Drabék, R. Pelle, E. Awino, H. T. Gotia, J. B. Munro, L. Tallon, W. I Morrison, C. A. Daubenberger, V. Nene, D. P. Knowles, R. P. Bishop and J. C. Silva. 2020. Capture-based enrichment of *Theileria parva* genomic DNA enables whole genome assembly of a buffalo-derived strain and reveals distinct, host-associated parasite subspecies. *PLOS Negl. Trop. Dis. 14: e0008781.*
40. Edgar, R. C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res. 32 1792-1797.*
41. Maddison, W. P. and D. R. Maddison. 2019. Mesquite: a modular system for evolutionary analysis. *Version 3.61 http://www.mesquiteproject.org*.
42. Rozas, J., A. Ferrer-Mata, J. C. Sánchez-DelBarrio, S. Guirao-Rico, P. Librado, S. E. Ramos-Onsins, and A. Sánchez-Gracia. 2017. DnaSP 6: DNA sequence polymorphism analysis of large datasets. *Mol. Biol. Evol. 34: 3299-3302*.

775 43. Benson, G. 1999. Tandem repeats finder: a program to analyze DNA
776 sequences. *Nucleic Acids* Res. 27: 573-580.
777 44. R Core Team (2019). R: A language and environment for statistical computing. *R*
778 *Foundation for Statistical Computing, Vienna, Austria.* https://www.R-project.org
779 45. Armenteros, J. J. A., M. Salvatore, O. Emanuelsson, O. Winther, G. von Heijne, A.
780 Elofsson and H. Nielsen. 2020. Detecting sequence signals in targeting peptides using deep
781 learning. *Life Science Alliance 2: e201900429.*
782 46. Tretina, K., R. Pelle, J. Orvis, H.T. Gotia, O.O. Ifeonu, P. Kumari, N. C. Palmateer, S. B.
783 A. Iqbal, L. M. Fry, V. M. Nene, C. A. Daubenberger, R. P. Bishop and J. C. Silva. 2020.
784 Re-annotation of the *Theileria parva* genome refines 53% of the proteome and uncovers
785 essential components of N-glycosylation, a conserved pathway in many organisms. *BMC*
786 *Genomics 21: 279.*
787 47. Toye, P. G., B. M. Goddeeris, K. Iams, K., A. J. Musoke and W. I. Morrison. 1991.
788 Characterisation of a polymorphic immunodominant molecule in sporozoites and
789 schizonts of *Theileria parva*. *Parasite Immunol. 13: 49-62*.
790 48. Xue, G., C, von Schubert, H. Hermann, M. Peyer, R. Mausgagen, J. Schmuchli-Maurer,
791 P. Butikofer, G. Langsley and D. A. E. Dobbelaere. 2010. Characterisaation of gp34, a
792 GPI-anchored protein expressed by schizonts of *Theileria parva* and *T. annulata. Mol.*
793 *Biochem. Parasitol.172: 113-120.*
794 49. Woods, K. L., R. Theiler, M. Muhlemann, A. Segiser, S. Huber, H. R. Ansari, A. Pain and
795 D. A. E. Dobbelaere. 2013. Recruitment of EB1, a master regulator of microtubule
796 dynamics, to the surface of the *Theileria annulata* schizont. *Plos Pathogens 9:*
797 *e1003346*.
798 50. Huber, S., R. Theiler, D. de Quervain, O. Wiens, T. Karangenc, V. Huessler, D. A. E.
799 Dobbelaere and K. Woods. 2017. The microtubule-stabilising protein CLASP-1
800 associates with the *Theileria annulata* schizont surface via its kinetochore-binding
801 domain. *mSphere 2: e00215-17*.
802 51. Purnima, A., V. Ganesan and L. C. Eisenlohr, 2017. The elucidation of non-classical
803 MHC class II antigen processing through the study of viral antigens. *Curr. Opin. Virol. 22:*
804 *71-76*.
805 52. Fawcett, D. W., S. Doxsey, D. A. Stagg and A. S. Young. 1982. The entry of the
806 sporozoites of *Theileria parva* into bovine lymphocytes in vitro. *Eur. J. Cell. Biol. 27: 10-*
807 *21.*
808 53. Dobbelaere, D. A. and S. Rottenberg. 2003. *Theileria*-induced leukocyte transformation.
809 *Current Opin. Microbiol. 6: 377-382*.

54. Hulliger, L., J. H. K. Wilde, C. G. D. Brown and L. Turner. 1964. Mode of multiplication of *Theileria i*n cultures of bovine lymphocytic cells. *Nature 203: 728-730*.

55. von Schubert, C., G. Xue, J. Schmuckli-Maurer, K. L. Woods, E. A. Nigg and D. A. Dobbelaere. 2010. The transforming parasite *Theileria* co-opts host cell mitotic and central spindles to persist in continuously dividing cells. *PLoS. Biol. 8: e1000499.*

56. Vasoya, D., P. S. Oliveira, L.A. Muriel, T. Tzelos, C. Vrettou, W. I. Morrison, I. K. Ferreira de Miranda Santos and T. Connelley. 2021. High throughput analysis of MHC-I and MHC-DR diversity of Brazilian cattle populations. *HLA, in press*.

57. Gannage, M. and C. Munz. 2010. MHC presentation via autophagy and how viruses escape from it. *Semin. Immunopathol. 32: 373-381*.

58. Neefjes, J., M. L. M. Jongsma, P. Paul and O. Bakke. 2011. Towards a systems understanding of MHC class I and MHC class II antigen presentation. *Nature Rev. Immunol. 11: 823-836.*

59. Tewari, M. K., G. Sinnathamby, D. Rajagopal and L. C. Eisenlohr. 2005. A cytosolic pathway for MHC class II-restricted antigen processing that is ptoteasome and TAP-dependent. *Nature Immunol. 6: 287-294*.

60. Eisenlohr, L. C., N. Luckashanak, S. Apcher, M. A. Miller and G. Sinnathamby. 2011., Beyond the classical: Influenza virus and the elucidation of alternative MHC class II-restricted antigen processing pathways. *Immunol. Res. 51: 237-248.*

61. Miller, M. A., A. P. V. Ganesan, N. Lukashenak, M. Mendonca and L. C. Eisenlohr. 2015. Endogenous antigen processing drives the primary CD4+ T cell response to influenza. *Nature Med. 21: 1216-1222*.

62. Iams, K. P., J. R. Young, V. Nene, J. Desai, P. Webster, O. K. ole MoiYoi and A. J. Musoke. 1990. Characterisation of a gene encoding a 104-kilodalton microneme-rhoptry protein of *Theileria parva*. *Mol. Biochem. Parasitol. 39: 47-60*.

63. Skilton, R. A., R. P. Bishop, C. W. Wells, P. R. Spooner, E. Gobright, C. Knonge, A. J. Musoke, M. Macklin and K. P. Iams. 1998. Cloning and characterisation of a 150 kDa microsphere antigen of *Theileria parva* that is immunologically cross-reactive with the polymorphic immunodominant molecule (PIM). *Parasitology 117: 321-330*.

64. Siefried, A., J. Schultz and A. Gohla. 2012. Human HAD phosphatases: Structure, mechanism and roles in health and disease. *FEBS J. 280: 549-571*.

65. Frasse, P. M. and A. R. Odom John. 2019. Haloacid dehalogenase proteins: Novel mediators of metabolic plasticity in *Plasmodium falciparum*. *Microbiol. Insights 12: 1-4*.

66. Bishop, R., E. Gobright, V. Nene, S. Morzaria, A. Musoke and B. Sohanpal. 2000. Polymorphic open reading frames encoding secretory proteins are locaated vless than 3 kilobases from *Theileria parva* telomeres. *Mol. Biochem. Parasitol., 110: 359-371*.

67. Schmuckli-Maurer, J., C. Casanova, S. Schmied, S. Affentranger, I. Parvanova, S. Kang'a, V. Nene, F. Katzer, D. McKeever, J. Muller, R. Bishop, A. Pain and D. A. E. Dobbelaere. 2009. Expression analysis of the *Theileria parva* subtelomere-encoded variable secreted protein gene family. *Plos One 4 (3): e4839*.

68. Baldwin, C. L., B. M. Goddeeris and W. I. Morrison. 1988. T-cell clones in immunoparasitology. *Parasitology Today 4: 40-45*.

69. Pelle, R., S. P. Graham, M. N. Njahira, J. Osaso, R. M. Saya, D. O. Odongo, P. G. Toye, P. R. Spooner, A. J. Musoke, D. M. Mwangi, E. L. Taracha, W. I. Morrison, W. Weir, J. C. Silva and R. P. Bishop. 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: e19015*.

70. Connelley, T. K., N. D. MacHugh, R. Pelle, W. Weir and W. I. Morrison. 2011. Escape from CD8+ T cell response by natural variants of an immunodominant epitope from *Theileria parva* is predominantly due to loss of TCR recognition. *J. Immunol. 187: 5910-5920*.

71. Oura, C. A. L., B. B. Asiimwe, W. Weir, G. W. Lubega and A. Tait. 2005. Population genetic analysis and sub-structuring of *Theileria parva* in Uganda. *Mol. Biochem. Parasitol. 140: 229-239*.

.

870 **Figure Legends:**

871 Figure 1: Schematic illustration of parameters used to select proteins for production of a
872 synthetic peptide library for antigen screening, based on the predicted properties of *T. parva*
873 gene products: Secreted = proteins possessing a signal peptide; TM = transmembrane
874 domain; GPI = glycosylphosphatidylinositol anchor.

875 Figure 2: Results obtained from secondary screens of a subset of the *T. parva* peptide
876 library with CD4 T cells from immune cattle: A set of 78 peptide pools were used, including
877 all those detected as positive in primary screens of 514 pools, plus pools that were negative
878 in all screens. Cells from a *T. parva*-stimulated CD4 T cell line from each animal were tested
879 in duplicate wells. Reactivity of the CD4 T cells was assayed by testing the ability of
880 supernatants of peptide-stimulated cells to upregulate expression of surface class II MHC on
881 MDBK cells, as a measure of IFN-γ release. Results are shown for CD4 T cell lines from 3
882 MHC-homozygous animals – 1011 (DRB3 010:01, 4003 (DRB3 011:01) and 2824 (DRB3
883 016:01). Position 79 in each panel represents the result obtained with supernatants from
884 CD4 T cells incubated with MDBK cells alone. The contents of each peptide pool are shown
885 in supplementary Table S1.

886 Figure 3: Mapping of epitopes recognised by CD4 T cells: Results are shown for responses
887 to individual overlapping 18-mer peptides for five pools containing peptides for Tp18, Tp24,
888 Tp29, Tp30 and Tp32, using CD4 T cells from animals 2059, 3846, 2824, 2824 and 4003
889 respectively. The assays used cells from a *T. parva*-stimulated CD4 T cell line from each
890 animal, tested in duplicate wells. Responses were assayed as described in the legend for
891 figure 2. The final well in each panel (position 51) shows the response of T cells to MDBK
892 cells incubated with the respective peptide pool.

Figure 4: Class II MHC-restriction of CD4 T cells specific for *T. parva* epitopes: Presentation of individual epitopes by MHC class II DR was examined by measuring responses to Chinese hamster ovary (CHO) cells transfected with the DRB3 010:01 or DRB3 011:01 alleles, which had been pre-incubated with peptide and washed prior to incubation with CD4 T cells. CD4 T cell lines from animals homozygous for DRB3 010:01 or DRB3 011:01, were assayed in two experiments for each (i and ii). The assays used cells from a *T. parva*-stimulated CD4 T cell line from each animal, tested in duplicate wells. Reactivity of the CD4 T cells was assayed by testing the ability of supernatants of stimulated T cells to upregulate expression of surface class II MHC on MDBK cells, as a measure of IFN-γ release. Results are shown for responses to peptide added directly to CD4 T-cells and auto-presented (positive control), responses against peptide pre-incubated with CHO cells expressing an irrelevant BoLA-DR molecule (negative control) and responses against peptides pre-incubated with CHO cells expressing the relevant BoLA-DR molecule. * $p < 0.05$ compared to responses to cells expressing the irrelevant DR.

Figure 5: Class II MHC DR- or DQ-restriction of CD4 T cells specific for *T. parva* epitopes: Inhibition of epitope recognition by monoclonal antibodies specific for bovine DR (IL-A21) and DQ (CC158) was examined by measuring IFN-γ release in supernatants of peptide-stimulated CD4 T cells using a specific ELISA. Cells from a *T. parva*-stimulated CD4 T cell line from each animal were tested in duplicate wells. Results are shown for two DR-restricted ($Tp16_{103\text{-}120}$ and $Tp19_{835\text{-}852}$) and two DQ-restricted epitopes ($Tp2_{76\text{-}92}$ and $Tp15_{517\text{-}534}$). Responses to peptide in the absence of added antibody are shown in grey. For each epitope, inhibition by the respective antibody was observed compared to the non-inhibitory antibody $p < 0.01$.

Figure 6: Amino acid sequence variation across coding sequence and location of epitopes: Average pairwise difference per non-synonymous site between sequences (π_{NS}) is shown for six representative antigens. π_{NS} was calculated using a sliding window of 30 sites,

with a five-site step. The π_{NS} plot is interrupted in regions of the alignment with indels between alleles. Repeat regions are shown in grey below the plot. The location of epitopes shown in Table 3 is marked by MHC genotype in yellow (DR 10:01), green (DR 11:01) and blue (DR 16:01).

Figure 1

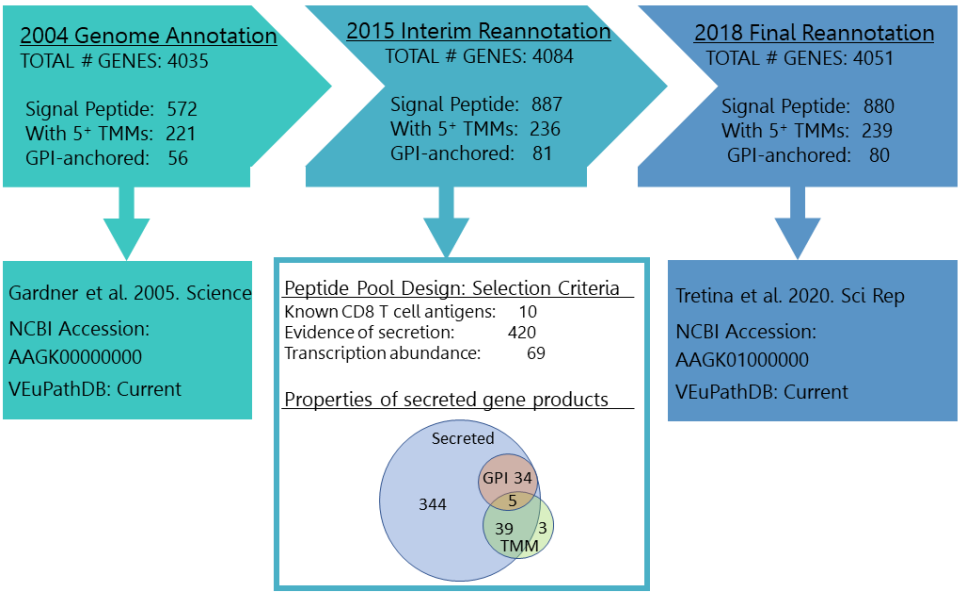
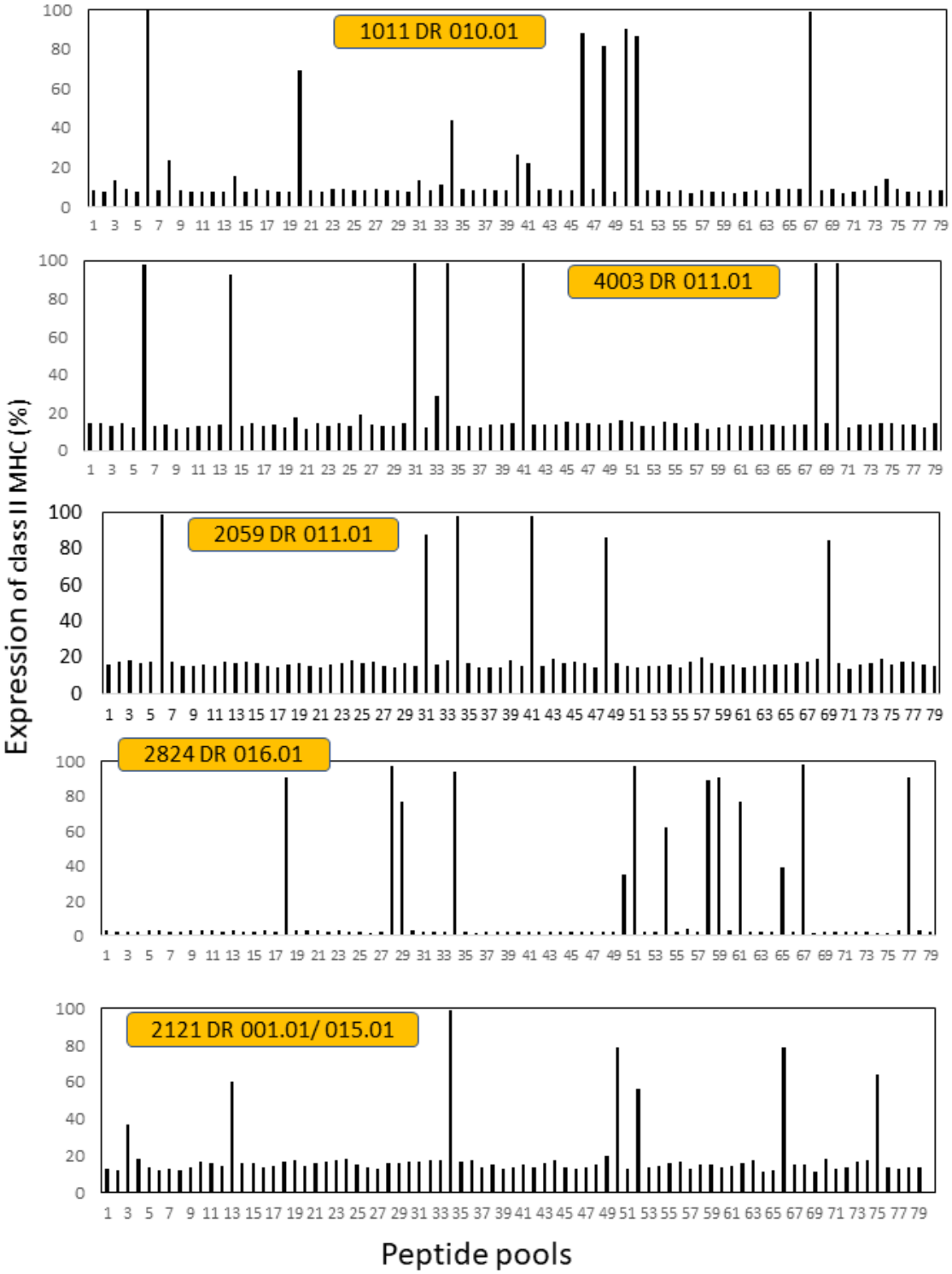


Figure 2



939

940

941

942 Figure 3

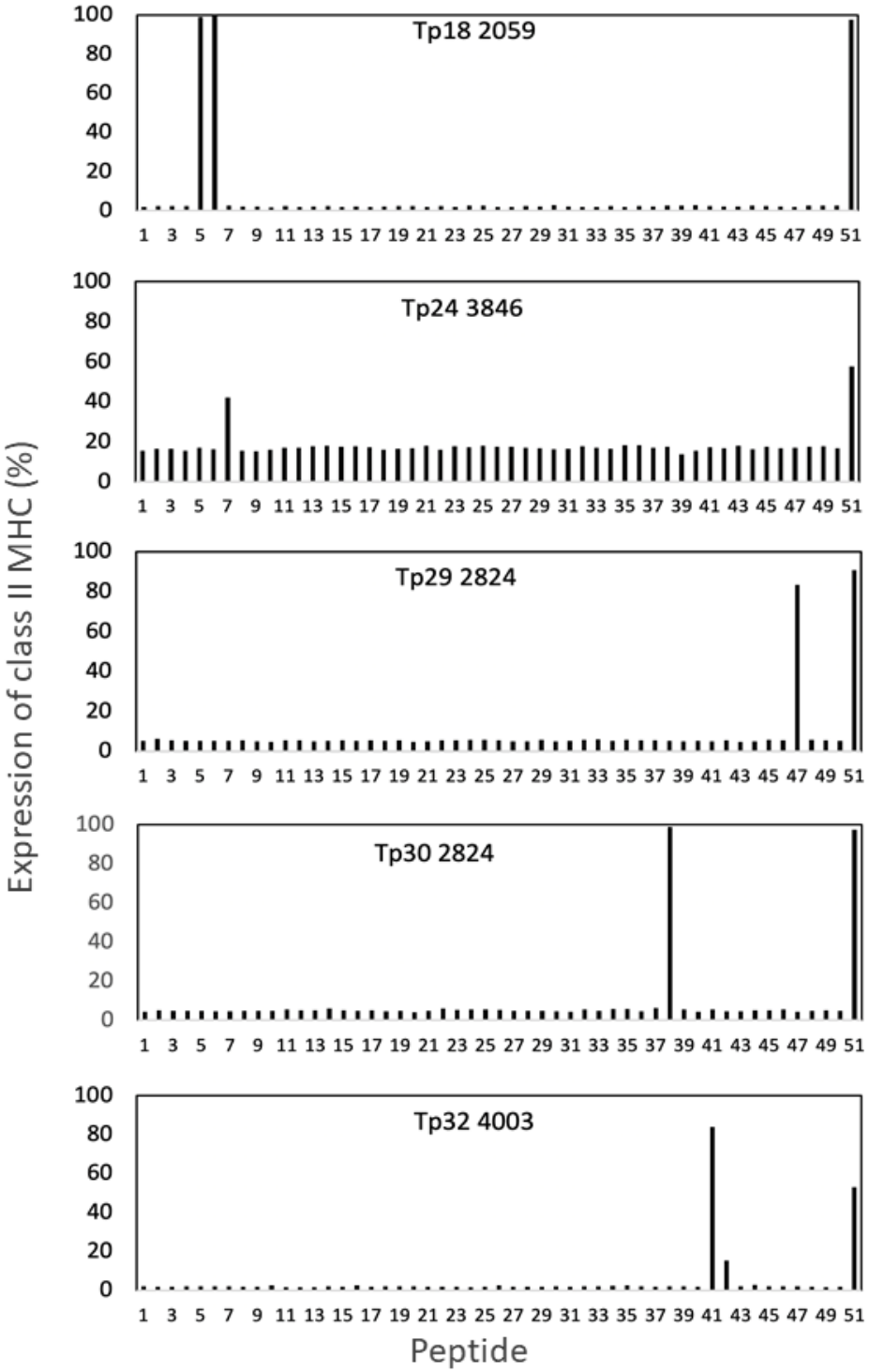
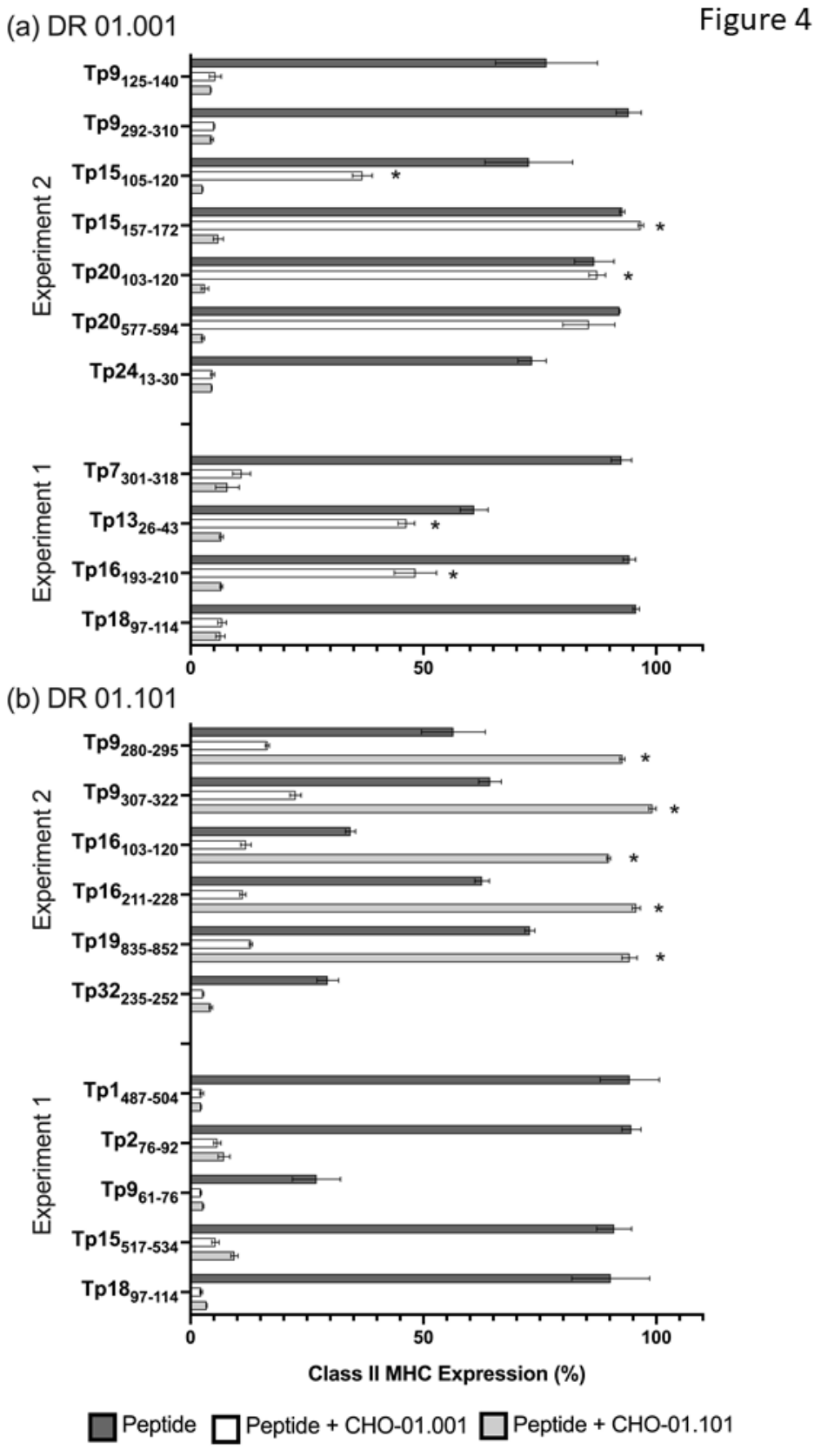


Figure 3

943

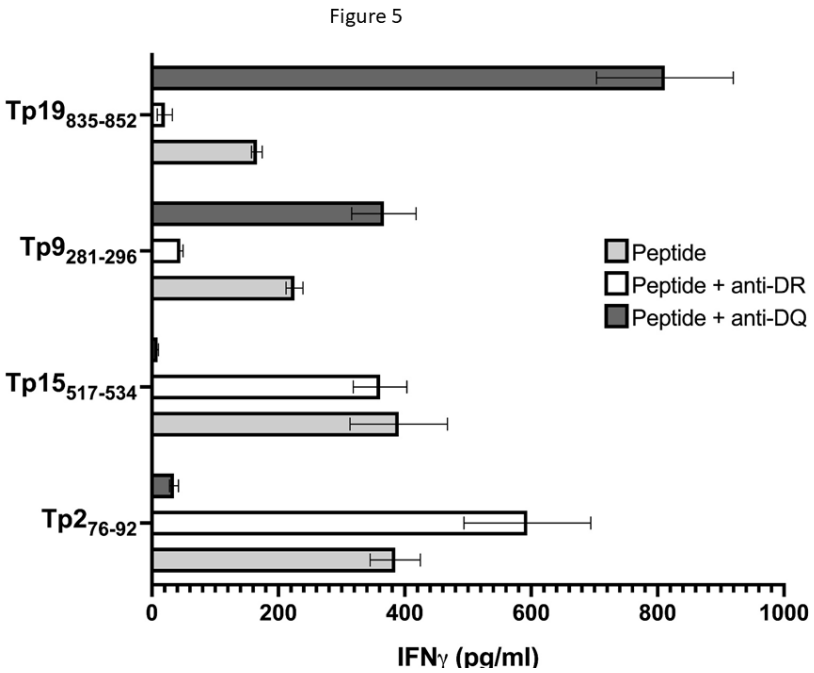
944

945 Figure 4



946

947 Figure 5

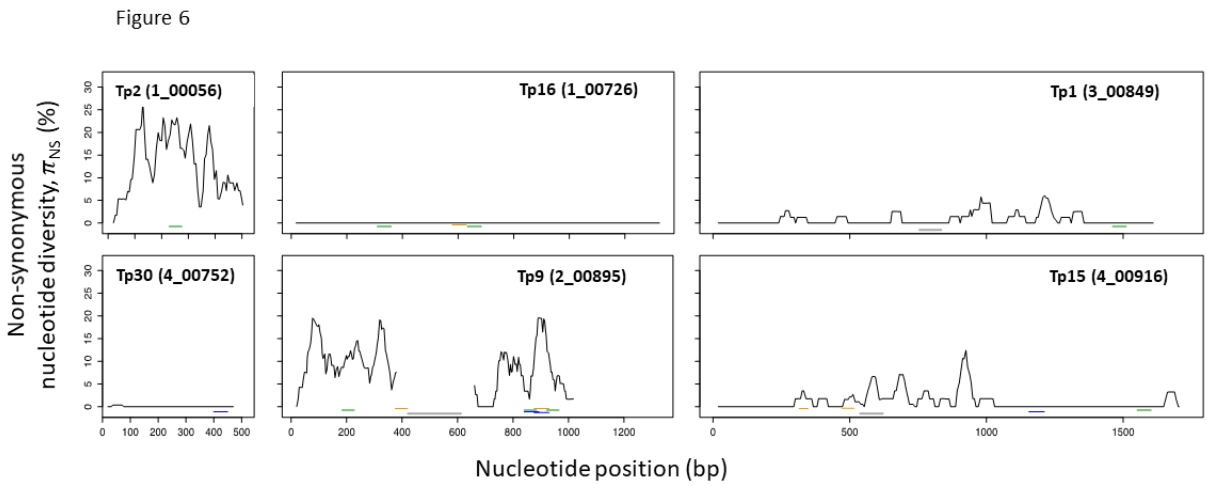


948

949

950

951 Figure 6



952

953 Table 1: Summary of antigens recognised by CD4 T cells from 12 cattle immunised against *Theileria parva*

Antigen[1]	Gene ID[2]	CD8[3] antigen	Animals and MHC types (Class I serotypes, class II DR genotypes)											
			598 A15/A17 11:01/16:01	2121 A10/A12 01:01/15:01	2786 A12/A20 12:01/15:01	2759 A13/A19 07:01/27:01	2756 A13/N5 14:01/27:03	2824 A19/A19 16:01/16:01	1011 A10/A10 10:01/10:01	3846 A10/A10 10:01/10:01	2059 A11/A15 11:01/11:01	4003 A14/A14 11:01/11:01	605 A14/A14 09:01/09:01	641 A18/A18 20:02/20:02
Tp2	1_0056	yes	+			+	+				+	+		
Tp16	1_0726		+						+	+	+	+		+
Tp36 (gp34)	1_0939									+				
Tp14	1_1074		+					+						+
Tp17	1_1077		+					+						
Tp21	1_1078		+					+						
Tp22	1_1081		+					+						
Tp23	1_1082		+					+						
Tp24	1_1182			+						+				+
Tp25	1_1225					+								+
Tp26	2_0010						+							
Tp32	2_0123	yes									+	+		
Tp27	2_0243					+								
Tp7	2_0244	yes							+	+				
Tp9	2_0895	yes	+	+	+		+	+	+	+	+	+		
Tp28	2_0958				+					+				
Tp33	3_0263	yes		+										+
Tp13	3_0655						+		+					
Tp1	3_0849	yes		+		+			+			+		
Tp20 (p150)	3_0861			+	+			+	+	+			+	+
Tp18 (PIM)	4_0051								+		+			
Tp19 (p104)	4_0437								+		+	+	+	+
Tp29	4_0683						+	+						
Tp30	4_0752							+						
Tp15	4_0916							+	+	+		+		+

Tp31	4_0917			+	+			+						

954 [1] Previously used terminology shown in parenthesis.

955 [2] Gene identifier is presented in format X_YYYY, where X is the nuclear chromosome number and YYYY represents the locus number.

956 [3] CD8 antigens Tp1, Tp2 and Tp7 reported by Graham et al., 2006 and Tp9 by Hemmink et al., 2016; Tp32 and Tp33 detected in recent antigen screens (N. D.
957 MacHugh, unpublished data.

958

959

960

961

962

963

964

965

966

967

968

969

970

971

972

973

974

975 Table 2: Properties of antigens recognised by CD4 T cells from cattle immunised against *Theileria parva*

Ag [a]	Gene ID [b]	Protein					Gene Product[f]	Host Homologues		*N* [h]	π_{NS} (%) [i]	π_{S} (%) [j]	Expression rank [k]
		Length (aa)	Size (kD)	Sig P [c]	TM [d]	GPI [e]		Bovine	Simil. (%) [g]				
Tp2	1_0056	174	19.1	+				-	-	16	14.5	21.8	229
Tp16	1_0726	448	49.5				Elongation factor Tu GTP binding domain	DAA23235.1	72.02	15	0	1.14	4
Tp36	1_0939	307	36.4	+		+	Schizont surface protein	-	-	17	0.24	1.94	984
Tp14	1_1074	318	36.0	+			haloacid dehalogenase-like hydrolase	-	-	17	0.00	0.00	90
Tp17	1_1077	292	32.9	+			haloacid dehalogenase-like hydrolase	-	-	13	0.07	0.25	3062
Tp21	1_1078	304	34.2	+			haloacid dehalogenase-like hydrolase	-	-	16	0.02	0.00	2129
Tp22	1_1081	307	34.4	+			haloacid dehalogenase-like hydrolase	-	-	13	1.88	3.46	3331
Tp23	1_1082	303	34.5	+			haloacid dehalogenase-like hydrolase	-	-	13	0.04	0.24	3476
Tp24	1_1182	321	34.9				lactate/malate dehydrogenase	DAA13962.1	31.76	17	0.02	1.44	14
Tp25	1_1225	433	48.8	+			SVSP family protein	-	-	12	0.18	0.39	512
Tp26	2_0010	450	52.0	+			SVSP family protein	-	-	15	0.44	0.14	1851
Tp32	2_0123	400	45.5				DEAD/DEAH box helicase	DAA18212.1	65.33	17	0.00	0.00	34
Tp27	2_0243	818	93.1				Heat shock protein homolog pss1	DAA27494.1	32.48	14	0.09	1.99	520
Tp7	2_0244	721	83.7				HSP90	DAA17282.1	66.11	12	0.03	0.72	5
Tp9	2_0895	334	34.7	+				-	-	13	9.08	14.24	8
Tp28	2_0958	675	76.4	+			SVSP family protein	-	-	14	0.95	2.25	2475
Tp33	3_0263	788	90.7	+			Schizont-associated, in complex with CLASP	-	-	15	0.07	0.00	1046
Tp13	3_0655	157	17.2					-	-	17	0.04	2.00	82
Tp1	3_0849	543	61.4	+				-	-	17	0.78	0.65	622
Tp20	3_0861	1452	164.9	+			P150 - sporozoite microspheres, schizont secreted	-	-	13	0.52	1.21	287
Tp18	4_0051	480	52.4	+	+		PIM - sporozoite microspheres, schizont surface	-	-	9	2.53	1.35	3
Tp19	4_0437	924	103.6	+		+	p104 – sporozoite rhoptry, schizont surface	-	-	14	0.11	0.07	61

Tp29	4_0683	655	72.7	+			78 kDa glucose-regulated protein	DAA24281.1	65.62	14	0.04	0.83	160
Tp30	4_0752	162	18.3				Ribosomal S27a family protein	DAA24675.1	55.56	17	0.03	0.90	28
Tp15	4_0916	574	64.8	+			SVSP family protein	-	-	17	1.07	2.04	844
Tp31	4_0917	528	60.2	+			SVSP family protein	-	-	15	0.33	0.58	2298

[a] Ag: Antigen; [b] Gene: Proxy for gene identifier in the genome of the *T. parva* reference strain Muguga, where the first digit stands for chromosome number and the last four digits to the locus number in the original and the updated genome annotations (respectively, Gardner et al. 2005 and Tretina et al. 2020); [c] Sig P: Presence of a signal peptide targeting the protein to the secretory pathway with high reliability, as determined with TargetP v1.1; [d] TM: Presence of one or more trans-membrane domains; [e] GPI: Presence of a predicted GPI anchor, as determined by GPI-SOM and/or predGPI; [f] Properties relating to parasite highlighted in red; [g] Simil: Percent amino acid similarity over segment of protein aligned to host protein; [h] Number of alleles used to estimate synonymous and non-synonymous nucleotide diversity; [i] Non-synonymous nucleotide diversity: average number of non-synonymous mutations per non-synonymous site among *T. parva* strains infecting cattle, per 100 sites; [j] Synonymous nucleotide diversity: average number of synonymous mutations per synonymous site among cattle-infecting *T. parva*, per 100 sites; [k] Expression Rank: Expression of all *T. parva* genes in the schizont stage (Tretina et al 2016) was ranked, with 1 being the most expressed and 4051 the least expressed, and genes assigned the corresponding rank.

998

999 Table 3: CD4 T cell epitopes[1] identified in *Theileria parva* antigens recognised by cattle homozygous for the DR 1001, 1101 and 1601 alleles

1000

Antigen	Gene ID	DR 01.001$^+$ animals epitope	DR 01.101$^+$ animals epitope	DR 01.601$^+$ animal epitope	MHC restriction[2]	
					DR	DQ
Tp2	1_0056		$_{76}$-LETLFGKHGLGGISKDC-$_{92}$			+
Tp16	1_0726	$_{193}$-GFLGDNMIDKSDKMPWYK$_{210}$	$_{103}$-ITGTSQADVAMLVVPAES-$_{120}$		+	
					+	
			$_{211}$-GKILVEALDLMEPPKRPV-$_{228}$		+	
Tp14	1_1074			$_{55}$KYFAIDIDGTFFIKD$_{-67}$	+	
				$_{76}$-NIAAFKRLQDAGVLPFF-$_{92}$	+	
Tp17	1_1077			$_{30}$-KYFAIDIDGTFHIKD-$_{44}$	+	
Tp21	1_1078			$_{24}$-IYFGVDIDGTFYVED-$_{38}$	nt	
Tp22	1_1081			$_{34}$-KYFAIDIDGTFHIKD-$_{44}$	nt	
Tp23	1_1082			$_{40}$-KFFAIDIDGTFYIND-$_{54}$	nt	
Tp24	1_1182	$_{13}$-GSGNIGGIMGYLSQLTEL-$_{30}$				+
Tp32	2_0123		$_{235}$-ELTLEGIKQFYILIDKEY-$_{252}$		+	
Tp7	2_0244	$_{301}$-NEEYAAFYKNLTNDWEDH$_{318}$			+	
Tp9	2_0895		$_{61}$-TKQDLDAKFPGMKKSK-$_{76}$			+
		$_{125}$-GPYGQAGYVGQPGAVG-$_{140}$				+
			$_{280}$-YDGEKVWSLEVGGDYA-$_{295}$	$_{281}$-DGEKVWSLEVGGDYAV$_{-296}$	+	
		$_{292}$-GDYAVKVLVFPIGFKEKTI-$_{310}$		$_{292}$-GDYAVKVLVFPIGFKEKTI$_{-310}$		+
			$_{307}$-EKTIEITFIGGEKEIY-$_{322}$		+	
Tp13	3_0655	$_{26}$-GRVSNYVTYAKKLLSNGI-$_{43}$			+	
Tp1	3_0849		$_{487}$-SIVNVYGKNDEPLSYAPS-$_{504}$			+
Tp20	3_0861	$_{103}$-QEILYYKWEKHGFVKETY-$_{120}$		$_{103}$-QEILYYKWEKHGFVKETY$_{-120}$	+	
				$_{451}$-FNKFDMLHDGVYYSSPVP$_{-468}$	+	
		$_{577}$-PLSGYHVRYVNYGKVIMW-$_{594}$			+	
				$_{1351}$-CKANNPVVYIKAGDKTVW$_{-1368}$	+	
Tp18	4_0051	$_{97}$-QQGPDTPQPIQEPSGPVQ$_{114}$	$_{97}$-QQGPDTPQPIQEPSGPVQ$_{-114}$			+

Tp19	4_0437		$_{835}$-KSFDDLTTVELAPEPKAS-$_{852}$		+	
Tp29	4_0683			$_{499}$-QIEVTFNIDTNGILSVTA-$_{516}$	+	
Tp30	4_0752			$_{133}$-NCGRGVFMAAHNNRTYCG-$_{150}$	+	
Tp15	4_0916	$_{105}$-GTYQHYGPPVFPPQPE-$_{116}$ $_{157}$-GIQYVPYQTLQIPQPQ-$_{172}$	$_{517}$-RNQVVWIKTASEGFPSSM-$_{534}$	$_{385}$-IEMTEKEYKIIVDSRF-$_{400}$	+	
					+	
					+	
						+

1001

1002 [1] The location of the epitopes was determined by screening overlapping 18-mer peptides (overlapping by 12 residues) for the respective positive antigen,
1003 using an interferon-ɣ release assay as described by Hart et al., 2017. In some cases, further truncated peptides were tested.

1004 2 The class II restriction of T cells specific for the epitopes was determined using two assays, first by testing recognition of the peptides pre-incubated with
1005 CHO cells expressing the 1001 or 1101 class II DR alleles and, second by testing the ability of monoclonal antibodies specific for bovine DR or DQ (IL-A21 and
1006 CC158 respectively) to inhibit recognition of the peptides. Nt = not tested.

1007

1008

1009

1010

1011

1012

1013

1014

1015

1016

1017

1018

Table 4: CD4 and CD8 T cell responses of 4 animals immunised with recombinant adenovirus and vaccinia viruses expressing the Tp9, Tp10 and Tp15 *T. parva* antigens.

Antigens [1]	**T cell lines and MHC genotypes** [2]							
	2750 MHCI - A10/A14, DR - 01.001/01.101		2733 MHCI – A11/A14 DR - 01.101/01.401		2751 MHCI – A14/A15 DR - 01.101/01.401		2758 MHCI – A10/A11 DR - 01.001/01.101	
	CD4	CD8	CD4	CD8	CD4	CD8	CD4	CD8
Control	3 (0.3)	18 (1.1)	6 (1.5)	15 (0.2)	3 (1.3)	nt	6 (1.6)	3.5 (0.4)
Tp9	98 (1.5)**	95 (2.0)**	99 (0.7)**	73 (3-5)**	99 (0.2)**	nt	99 (0.2)**	97 (0.9)**
Tp10	97 (1.9)**	98 (2.8)**	33 (2.9)*	10 (0.9)	90 (4.0)**	nt	15 (2.7)*	32.6 (1.5)**
Tp15	99 (0.3)**	16 (2.0)	88 (5.4)**	14 (0.6)	29 (3.0)**	nt	85 (3.1)**	3.8 (0.7)

[1] The antigens consisted of pools of overlapping 18-mer synthetic peptides for each antigen. Controls represent responses to a peptide pool not recognised on these MHC backgrounds

[2] Responses were measured by testing the ability of supernatants from antigen-stimulated T cells to up-regulate expression of surface class II MHC on MDBK cells. Results are expressed as the % of class II MHC-positive cells, showing the mean (and standard deviation) for duplicates of each assay. * $p<0.05$, ** $p<0.01$.

nt – Not tested.

Table 5: Response to known Tp9 and Tp15 epitopes of CD4 T cells from an animal (2750) vaccinated with the Tp9 and Tp15 antigens expressed in recombinant adenovirus and vaccinia vaccine vectors

Antigen	**Responses of CD4 T cell lines (MHC DR type)** [1]		
	3846 (01.001/01.001)	**4003 (01.101/01.101)**	**2750 (01.001/01.101)**
Control [2]	9 (1.6)	5 (0.6)	6 (0.6)
$Tp9_{61-76}$		39 (3.5)*	5 (0.4)
$Tp9_{125-140}$	98 (1.3)**		94 (2.1)**
$Tp9_{280-295}$		99 (0.1)**	97 (1.1)**
$Tp9_{292-310}$	63 (8.8)*		82 (3.9)**
$Tp9_{307-322}$		99 (1.0)**	99 (0.4)**
$Tp15_{105-120}$	96 (2.7)**		5 (0.5)
$Tp15_{157-172}$	99 (0.2)**		36 (4.1)*
$Tp15_{517-534}$		90 (3.9)**	33 (5.8)*

[1] The 3846 and 4003 CD4 T cell lines were from *T. parva*-immune MHC-homozygous animals, which were used for antigen identification. Responses were measured by testing the ability supernatants from peptide-stimulated T cells to up-regulate expression of surface class II MHC on MDBK cells. Results are expressed as the % of class II MHC-positive cells, showing the mean (and standard deviation) for duplicates of each assay. * $p<0.05$, ** $p<0.01$.

[2] Controls represent responses to a peptide not recognised on either of these MHC backgrounds