Sequence diversity between class I MHC loci of African native and introduced Bos taurus cattle in Theileria parva endemic regions: in silico peptide binding prediction identifies distinct functional clusters

Isaiah Obara¹ · Morten Nielsen²,³ · Marie Jeschek⁴ · Ard Nijhof¹ · Camila J. Mazzoni⁴,⁵ · Nicholas Svitek⁶ · Lucilla Steinåa⁶ · Elias Awino⁶ · Cassandra Olds⁶ · Ahmed Jabbar¹ · Peter-Henning Clausen¹ · Richard P. Bishop⁶

Received: 18 September 2015 / Accepted: 8 January 2016 / Published online: 6 February 2016
© Springer-Verlag Berlin Heidelberg 2016

Abstract There is strong evidence that the immunity induced by live vaccination for control of the protozoan parasite Theileria parva is mediated by class I MHC-restricted CD8⁺ T cells directed against the schizont stage of the parasite that infects bovine lymphocytes. The functional competency of class I MHC genes is dependent on the presence of codons specifying certain critical amino acid residues that line the peptide binding groove. Compared with European Bos taurus in which class I MHC allelic polymorphisms have been examined extensively, published data on class I MHC transcripts in African taurines in T. parva endemic areas is very limited. We utilized the multiplexing capabilities of 454 pyrosequencing to make an initial assessment of class I MHC allelic diversity in a population of Ankole cattle. We also typed a population of exotic Holstein cattle from an African ranch for class I MHC and investigated the extent, if any, that their peptide-binding motifs overlapped with those of Ankole cattle. We report the identification of 18 novel allelic sequences in Ankole cattle and provide evidence of positive selection for sequence diversity, including in residues that predominantly interact with peptides. In silico functional analysis resulted in peptide binding specificities that were largely distinct between the two breeds. We also demonstrate that CD8⁺ T cells derived from Ankole cattle that are seropositive for T. parva do not recognize vaccine candidate antigens originally identified in Holstein and Boran (Bos indicus) cattle breeds.

Keywords Theileria parva · CD8⁺ T cell epitopes · Class I MHC · Ankole cattle · Pyrosequencing

Introduction

The protozoan parasite Theileria parva is responsible for East Coast fever (ECF), an acute lymphoproliferative disease of cattle in eastern, central and southern Africa (Norval et al. 1992). Transmitted by the tick Rhipicephalus appendiculatus, T. parva sporozoites are infective to mammalian hosts and these differentiate rapidly into intracellular schizonts that immortalize bovine lymphocytes. The infection and treatment (ITM) procedure has been used as an approach to live vaccination for control of T. parva for over 40 years (Radley et al. 1975). Despite a high efficacy against field challenge, there are a number of constraints to widespread deployment of ITM.

Electronic supplementary material The online version of this article (doi:10.1007/s00251-016-0902-5) contains supplementary material, which is available to authorized users.

Isaiah Obara
iobara@zedat.fu-berlin.de

¹ Institute for Parasitology and Tropical Veterinary Medicine, Freie Universität Berlin, Robert-von-Ostertag-Str. 7-13, 14163 Berlin, Germany
² Center for Biological Sequence Analysis (CBS), Department of Systems Biology, The Technical University of Denmark, Building 208, DK-2800 Lyngby, Denmark
³ Instituto de Investigaciones Biotecnológicas, Universidad Nacional de San Martín, San Martín B 1650 HMP, Buenos Aires, Argentina
⁴ Berlin Center for Genomics in Biodiversity Research (BeGenDiv), Königin-Luise-Str. 6-8, 14195 Berlin, Germany
⁵ Leibniz-Institute for Zoo and Wildlife Research (IZW), Evolutionary Genetics, Alfred- Kowalke-Straße 17, D-10315 Berlin, Germany
⁶ International Livestock Research Institute (ILRI), P.O. Box 30709, Nairobi 00100, Kenya
The major ones are the technical and ethical issues relating to production and subsequent delivery of live sporozoites using a cold chain combined with the perception that *T. parva* parasites that are foreign to an area may be introduced (reviewed by Di Giulio et al. 2009).

Although monoclonal antibodies against a polymorphic *T. parva* sporozoite surface molecule (Nene et al. 1999; Obara et al. 2015) have been shown to neutralize infectivity in vitro (Musoke et al. 1982), at present there is no direct evidence that ITM induces significant levels of antibody against the sporozoite stage. In contrast, available data indicate that one important mechanism of protection afforded by ITM is attributable to CD8+ T cells specific for the intra-lymphocytic schizont stage of *T. parva*. Earlier time-course studies showed an association between the kinetics of class I MHC restricted CD8+ T cell cytotoxicity and remission of infection in immune animals (Emery et al. 1981; Eugui and Emery 1981). The most compelling evidence for the role of CD8+ T cells as major mediators of immunity to *T. parva* came from a study in which transfer of CD8+ T cell-enriched lymphocyte populations from an immune animal to its naïve chimeric twin conferred protection (McKeever et al. 1994). These observations led to the screening of *T. parva* schizont cDNA libraries with CD8+ T cells from Holstein and Boran cattle immunized by ITM, with the aim of defining the antigenic specificity of the protective response (Graham et al. 2006). It was evident from the findings of the antigen screens that animals of different class I MHC genotypes recognized distinct antigens (Graham et al. 2008). A more recent study utilizing a positional scanning combinatorial peptide library (PSCPL) driven analysis corroborated the findings from the antigen screening experiments (Hansen et al. 2014). Taken together with the evidence that different cattle breeds often carry distinct class I MHC alleles (Robinson et al. 2013), the fact that recognition of *T. parva* CD8+ T cell target antigens is class I MHC restricted raises the issue of the likely impact of MHC diversity on the coverage of epitope-based vaccines.

The number of ‘breeds’ so far documented from the ECF endemic region, based on the information contained within the FAO breed database (http://dad.fao.org/), is 211 in the 11 countries where the disease currently occurs. However, such breeds cannot be directly equated to defined European breeds, such as the Holstein, since although ancestral African cattle populations typically have taurine mitochondrial haplotypes, introgression of alleles into the nuclear genome from *Bos indicus* has occurred at high frequency (MacHugh et al. 1997; Anderung et al. 2007). Although potentially serious challenges are presented by the complexity and diversity of expressed bovine MHC loci in African cattle, there are currently significant ongoing efforts to design vaccines based on CD8+ T cells targeting *T. parva*-infected lymphocytes using reverse vaccinology approaches (Nene et al. 2012; Hansen et al. 2014; Svitek et al. 2014; Graham et al. 2007). These are based on training neural network software (Hoof et al. 2009) using the CD8+ T cell target antigens and epitopes described above (Graham et al. 2006, 2008), with the predictions validated using tetramers designed to be recognized by CD8+ T cells designed from *T. parva* antigen-peptide complexes (Svitek et al. 2014).

Available data on cattle class I MHC diversity demonstrates that major variability among allelic sequences reside in amino acid positions that determine specificity for peptide binding (Robinson et al. 2013). A listing of cattle MHC alleles and a description of the nomenclature convention is provided in the Immuno-Polymorphism database (IPD; http://www.ebi.ac.uk/ipd/mhc/bola/). By comparison to humans that have a defined complement of three classical class I loci (HLA-A, HLA-B and HLA-C), mapping and phylogenetic data (Hollms et al. 2003; Ellis 2004; Di Palma et al. 2002; Birch et al. 2006) support the grouping of cattle class I MHC sequences into six classical loci (genes 1–6). Antigen binding and presentation capacity have been demonstrated for products of all five loci (Codner et al. 2012). A further level of complexity in cattle class I MHC results from variations in gene content between haplotypes (Ellis et al. 1999; Ellis 2004). Despite recent advances in assay throughput capabilities (Harmdahl et al. 2009), the required scale imposes practical limitations that make such experimental determination of the ligand selectivity of all possible haplotypes impractical. However, the artificial neural network-based MHC binding predictor, NetMHCpan (Hoof et al. 2009; Nielsen et al. 2007), represents a significant advance in the possibility to provide accurate predictions of peptide-class I MHC interactions, even across different species. Although trained predominantly on peptide binding data covering human and nonhuman primates, the versatility of NetMHCpan has been demonstrated in the analysis of the binding specificities of porcine (Pedersen et al. 2011) and bovine (Svitek et al. 2014) class I MHC molecules. Most recently, PSCPL analysis of peptide binding for bovine class I MHC molecules known to restrict *T. parva* candidate vaccine antigens provided data that proved useful for refining the predictive performance of NetMHCpan (Hansen et al. 2014). Furthermore, procedures exist for validation of the natural presentation and immunogenicity of the predicted epitopes based on selection of cytokine-secreting or tetramer-positive T cell populations (Nene et al. 2012; Svitek et al. 2014; Hansen et al. 2014). Similar approaches have recently been applied for prediction of bovine class I MHC-restricted epitopes in foot and mouth disease (Pandya et al. 2015).

A key information gap constraining application of the expedited epitope identification procedures in strategies for *T. parva* control is the paucity of class I MHC data for cattle populations in the endemic regions of Eastern Africa. A major challenge to typing cattle for class I MHC is presented by a...
complex genomic organization that necessitates simultaneous genotyping of co-amplifying loci (reviewed by Ellis 2004). In the current study, we utilized the depth and throughput of pyrosequencing, coupled with the multiplexing capabilities of the 454 GS FLX platform to resolve variations at the structurally important peptide-binding regions of class I MHC transcripts in a population of Ankole cattle, an African taurine with demonstrated tolerance to infections with T. parva (Paling et al. 1991). For comparison, we also characterized a European Holstein population, introduced to East Africa, for the sequence of the class I MHC alleles present. To gain perspective into differential ligand selectivity of the alleles described in the present study, we derived functionally relevant allele groupings based on sequence features defining peptide binding grooves using MHCcluster (Thomsen et al. 2013). Application of this software provides evidence for major specificity groups, known as ‘supertypes’ in cattle (Pandya et al. 2015), as previously described for humans by Sette and Sidney (1999), suggesting that the constraint of class I MHC diversity for recombinant vaccine development for ECF control may not be insurmountable. In the study described herein, the use of the MHCcluster software provided evidence for minimal functional overlap between Ankole and Holstein peptide binding specificity, suggesting a requirement for additional epitopes in ECF subunit vaccines designed to induce CD8+ T cell-based protection.

Materials and methods

Taurine samples, RNA isolation and cDNA synthesis

Peripheral blood lymphocytes from 17 African taurines (Ankole) were used for preparation of pyrosequencing libraries, 14 of which were from Ol Pejeta—a ranch that integrates game conservancy with cattle production in central Kenya (www.olpejetaconservancy.org/). The other three animals, originally from a farm in Uganda, were part of a herd used for genetic studies at ILRI. For comparison, 17 Holstein cattle, which were part of a vaccine trial that combined tick salivary gland and a T. parva sporozoite antigen at ILRI (R. Bishop unpublished), were also typed by plasmid sub-cloning of class I MHC amplicons and chain termination sequencing. Ficoll-based density centrifugation was used for separation of peripheral blood mononuclear cells (PBMCs) from cattle blood drawn into heparinized vacutainers. Cellular RNA was extracted from PBMCs using Trizol (Invitrogen), quantified by measuring Abs260 with a nanodrop UV spectrophotometer (Inqaba biotech, SA) and resolved by 1.2 % agarose gel electrophoresis to determine integrity. Total RNA was used as template for first-strand cDNA synthesis using a Superscript III Reverse Transcriptase kit (Invitrogen).

Ankole class I MHC amplicon library construction and pyrosequencing

We generated PCR amplicons spanning 410 bp between conserved sequences in exons 2 and 3 of class I MHC sequences from Ankole cattle using the Phusion High-Fidelity PCR kit (Thermo scientific). For each sample, we introduced a unique ten-nucleotide multiplex identifiers (MID) index to both the sense strand primer: Bov 7 (5′-GGCTACGTGGACGACACG-3′) and the anti-sense strand primer: Bov 11 (5′-CCCTCCAGGGTAGTTCT-3′). Thermal conditions and PCR mix composition were as previously described (Birch et al. 2006). Following MinElute PCR purification (Qiagen), the amplicons were quantified using a Nanodrop UV spectrophotometer (Inqaba biotech) and normalized to equimolar concentrations. Adapter ligation, library amplification by emulsion PCR and bidirectional pyrosequencing of amplicons as one pool in 1/16 region of a 70 × 75 picotitre plate were performed at ILRI using GS FLX Titanium chemistry (454 Life Sciences).

Variant calling: putative alleles, artefacts and chimeras

Since errors inherent to pyrosequencing mean that the potential for some proportion of the reads to be artefactual cannot be excluded, we subjected the reads passing the standard GS FLX base-calling criterion to a multilevel artefact filtering workflow. Foremost, reads that failed the following initial filtering criteria were considered as putative artefacts and excluded from further analysis: (i) sequence length of at least 196 bp which could allow reads to be assigned to either exon 2 or 3; (ii) presence of complete MID sequences to allow reads to be disaggregated into the indexed individuals; (iii) absence of base pair mismatches with the PCR primers sequences; (iv) reads that could be assigned to either exon 2 or 3 on the basis of exonic nucleotide sequence alignment to the published reference sequence, BoLA-3*00101; (v) a minimum phred scaled quality score value of 20 in more than 95 % of bases; and (vi) absence of ambiguous bases (‘N’ calls). Whilst the demultiplexing steps were accomplished using the flexible barcode and adapter removal tool (FLEXBAR; Dodt et al. 2012), we used the FASTX command line tools, FASTQ/A Trimmer and FASTQ Quality Filter (http://hannonlab.cshl.edu/fastx_toolkit), for length trimming and quality filtering, respectively. Exon sequences were identified by similarity searches using the NCBI nucleotide BLAST program (Altschul et al. 1990).

The second set of analyses was intended to evaluate variant authenticity by detecting chimeric reads and base miscalls based on intra and inter-amplicon comparisons as described in Sommer et al. (2013). Briefly, identical reads from each MID indexed sample were collapsed into variants, using the FASTQ/A Collapser tool (http://hannonlab.cshl.edu/fastx_toolkit), and singletons (variants represented by single reads.
in an amplicon) were discarded. The variants, ordered by frequency, were aligned using MAFFT version 7.122 using the L-INS-i option with minor manual adjustments (Katoh and Standley 2013). The collection of aligned reads from indexed animals was stored in a SQLite database for downstream analysis. SQL queries and custom Biopython scripts were written to classify variants as putative alleles if they comprised the most frequent cluster and did not match any of the artefact definitions. The remaining variants were designated as putative alleles, ‘unclassified’ variants or artefactual sequences based on inter and intra-animal comparisons and frequency relative to artefacts within the same amplicon as described elsewhere (Sommer et al. 2013). Artefacts were excluded from downstream analysis. Finally, we used a custom Biopython script to rank, separately for each animal, all the potential combinations of the exon 2 and exon 3 variants classified as putative alleles according to their abundance in the MID demultiplexed reads from an individual. Since the exon sequences were already known at this step, we relied on partial identifying sequences to allow the inclusion of MID demultiplexed reads that were too short to comprise the complete sequences of exons 2 and 3 in this step. We further evaluated the reliability of the genotypes obtained by amplicon pyrosequencing by comparing them to partial or full-length class I MHC cDNA Sanger sequencing datasets for the three Ankole cattle from the ILRI herd.

**Generation of cattle class I cDNA-PCR clone libraries and Sanger sequencing**

We used enzyme Pfu DNA polymerase (Thermo scientific) to generate an amplicon that spans 632 bp between conserved sequences in exons 2 and 4 of class I MHC sequences from the 17 Holsteins and 3 Ankole cattle. The inclusion of the three Ankole samples, already genotyped by pyrosequencing, was intended to highlight the reliability of the pyrosequencing reads. We used cDNA templates and the highly conserved primer pairs, Bov 7 (5′-GGCTACGTGGACGACACG-3′) and Bov 14 (5′-GAAGGCTGGTGCTCACA-3′), for amplification as previously described (Birch et al. 2006). For the three Ankole cattle, full-length class I MHC genes were also amplified from cDNA as described in Birch et al. (2006). The amplicons were resolved on 1.2 % ethidium bromide-stained agarose gels, excised and purified with DNA gel extraction spin columns (Qiagen). PCR products were sub-cloned into the pGEM T-Easy vector (Promega) and subsequently transformed into competent JM 109 Escherichia coli cells. From each animal, 10 positive (blue/white selection) clones were picked and grown overnight in 5 ml of LB-Medium. Plasmid DNA was purified according to QIAprep Miniprep Kit Protocol (QIAGEN) and insert presence confirmed by EcoRI digestion or by colony screening PCR. Sanger sequencing of sub-cloned products was undertaken bidirectionally using T7 and SP6 primers. Reads from each clone, assembled into unidirectional contigs using CLCgenomics version 8.0.2 (http://www.clcbio.com/products/clc-genomics-workbench/), were subjected to the same artefact and chimera filter as described for the Ankole data, but with a modification of the preprocessing criterion to retain singletons after collapsing if they were carried by more than one animal. This relaxation in stringency reflects the differences in throughput between pyrosequencing and Sanger sequencing.

**In silico clustering of cattle class I MHC sequences based on predicted ligand selectivity**

We investigated overlaps, if any, in the predicted ligand selectivities of the Ankole and Holstein class I MHC products based on NetMHCpan 2.8 predictions and used MHCcluster 2.0 (Thomsen et al. 2013) to generate a specificity heat map and a functional tree constructed using 100 bootstrap calculations and 50,000 peptides. The predictions are based on sequence features defining the peptide binding grooves of the class I MHC products.

**CD8+ T cell specificities in the peripheral circulation of *T. parva*-infected Ankole cattle**

We assayed 40 field Ankole serum samples for antibodies to *T. parva* by enzyme-linked immunosorbent assay (ELISA) using the recombinant polymorphic immunodominant molecule (PIM) as described previously (Katende et al. 1998). IFN-γ responses generated following stimulation of PBMCs, as well as CD8+ T cell fractions, from serologically positive animals with complete sets of overlapping peptides representing *T. parva* candidate vaccine antigens were assayed in an ex vivo enzyme-linked immunospot (ELISpot) assay as described previously (Steinaa et al. 2012). Briefly, sterile nitrocellulose-lined 96-well microplates (Millipore, Billerica, MA, USA) were coated overnight at 4 °C with anti-bovine IFN-γ antibody (Serotec, Oxford, UK, cat.no. MCA1783) and blocked with RPMI containing 10 % heat-inactivated FBS for 2 h at 37 °C. The cryopreserved PBMCs were plated out in the coated wells in triplicate at 2 × 10^5 cells/well in 50 μl and an equal volume of complete sets of overlapping peptide pools derived from the *T. parva* vaccine candidate antigens added at 1 μg/ml final concentration (Graham et al. 2006). Negative (no peptide) control triplicate wells contained culture in supplemented RPMI medium. Cytotoxic T lymphocytes (CTLs) from BV115—an animal with demonstrated reactivity to Tp1, were used as positive assay control in triplicate wells. The cells were incubated for 36 h at 37 °C, 5 % CO₂. Spots were visualized by incubation with primary rabbit polyclonal antibovine IFN-γ antibody (Sigma-Aldrich, St. Louis, MO, USA), secondary AP-conjugated monoclonal anti-rabbit antibody (Sigma-Aldrich) and Fast substrate solution (Sigma-
A high throughput assessment of class I MHC diversity from a population of Ankole cattle

We sequenced normalized, GS FLX adaptor flanked, 410 bp amplicon libraries that extended exons 2 and 3 of cattle class I MHC from 17 Ankole cattle on a Roche 454 instrument. A total of 172,681 reads (over 56 million bases) passed the standard GS FLX base-calling filters. Majority of the bases that passed the GS FLX base-calling filters had a quality score value of more than 20 on the Phred scale, indicating a base call accuracy of more than 95 % (Supplementary Fig. 1). Furthermore, the modal sequence length was consistent with the expected product size (Supplementary Fig. 2).

Reads disaggregated on the basis of unique MID barcodes and further partitioned into individual exons were used as input for two sets of filters intended to detect and eliminate errors known to influence the reliability of pyrosequencing calls. The initial data quality check was restricted to filtering: (i) reads bearing mismatches to amplification primers; (ii) reads with ambiguous bases (‘N’ calls); (iii) reads with in-triprimer sequence lengths shorter than 196 bp; and (iv) reads containing more than 5 % low quality bases. Candidate variants that passed these initial filters were preprocessed for artefact filtering by collapsing identical reads, while maintaining read counts, and discarding singletons. The artefact filtering criterion excluded data from reads with bases that could be associated with chimerism and/or miscalls on the basis of intra and inter-amplicon comparisons as described in Sommer et al. (2013). The distribution of read counts over all MID indexed samples after successive rounds of filtering, normalized to sum to 1 (divided by the MID demultiplexed read counts) and partitioned into exons 2 and 3, is shown in Fig. 1.

After removal of PCR and sequencing associated artefacts, 27 putative alleles were obtained from a total of 966 exon 2 and 1642 exon 3 variants across the 17 Ankole samples (Table 1). The exonic variants classed as ‘putative alleles’, as well as the ‘unclassified variants’ which were carried by each of the 17 Ankole animals is shown in Supplementary Table 1. The latter includes a small number of variants whose classification as either putative allele or artefacts was inconclusive. It should be emphasized that this is a minimum estimate of alleles for these 17 animals and the stringent filtering process may have resulted in some alleles being missed. It also possible that the primers did not amplify all class I MHC loci from the Ankole, although they were designed to be as representative as possible of MHC diversity in cattle (Birch et al. 2006) and are also conserved in European Bison (Bison bonasus) (Babik et al. 2012). BLAST searches of the IPD and NCBI nucleotide databases conducted with the nucleotide sequences of the 27 unique Ankole class I MHC transcripts identified only five variants sharing 100 % identity with published cattle classical class I MHC sequences. With one exception, these previously reported sequences have been assigned to gene 2 and none has been reported in Holstein cattle. Similarly, two variants matched published non-classical class I MHC sequences with 100 % nucleotide identity. In addition to the characterized non-classical class I sequences, our analysis identified two putatively novel variants that share more than 99 % nucleotide identity with the published non-classical allele BoLA-NC2*00102 (Table 1). Furthermore, two of the sequences, BoLA-Ank_12 and BoLA-Ank_16, share in excess of 98 % nucleotide identity to BoLA-2*03201N, a null allele suffixed ‘N’ based on IPD nomenclature conventions to denote that it is transcribed but contains features that are inconsistent with expression as a functional protein.

Considerable variation was observed among the remainder of the sequences with respect to the percent nucleotide identity to previously described cattle classical MHC I alleles. Interestingly, with a few exceptions, these putatively novel classical class I variants share identities with published sequences known to be expressed in breeds other than the Holstein. The diversity per site, excluding variants shown to share nucleotide identity with non-classical class I alleles, was $p=0.09436$ (or 9.436 %) $=0.00385$. The nucleotide sequences designated as putatively novel in this study have been submitted to the European nucleotide archive (LN8886781–LN8886805). In addition to the stepwise variant validation procedure, a further consideration involved in evaluating the reliability of the genotypes obtained by pyrosequencing related to comparing them to the genotypes obtained by cloning and Sanger sequencing for the three Ankole cattle from the ILRI herd. Amongst the partial (632 bp) as well as the full-length Sanger sequencing datasets, five distinct alleles were identified, namely BoLA-Ank_04, BoLA-Ank_06, BoLA-Ank_08, BoLA-Ank_09 and BoLA-Ank_25. These matched the sequences retained from the amplicon pyrosequencing data by the cascade of filters used for variant calling.

Holstein class I MHC amplification, plasmid sub-cloning and Sanger sequencing

We characterized the class I MHC transcript profiles of 17 Holstein cattle by sequencing of plasmid sub-clones containing a 632-bp amplicon spanning exons 2 to 4 of the class I MHC alpha chain. After testing for incorporation of the correct insert by PCR and EcoRI digestion, 10 positive plasmid sub-clones from each animal were Sanger sequenced using T7 and SP6 primers. We subjected the Sanger sequences to similar artefact and chimera filters as described for the Ankole
data, but retained singletons if they were carried by more than one animal.

Our analysis indicated the existence of 23 distinct class I MHC sequences in the Holstein population genotyped in this study (Table 2). As anticipated, the call set was dominated by sequences displaying 100% nucleotide identity to allelic variants present in the definitive cattle MHC repository, the IPD, available at www.ebi.ac.uk/ipd/mhc/bola. The skew in the distribution of these known gene products over the class I MHC classical loci, characterized by a prominence for products of genes 1–3, is consistent with published analyses (Codner et al. 2012). In the absence of an identical match in the IPD database, the remaining sequences within the Holstein class I MHC dataset were conservatively designated as putatively novel if their existence was independently supported by plasmid sub-clone sequences from at least two animals. Two variants, working allele names BoLA-Hol_05 and BoLA-Hol_10 (Table 2), met this criterion. Interestingly, BoLA Hol_10 exhibited only 93.64% sequence similarity to any cattle class I MHC molecule currently in the databases, suggesting that despite the relatively limited diversity within the Holstein breed and intensive class I MHC sequence analysis in recent years, the BoLA database is not yet complete. The two novel Holstein sequences have been submitted to the European nucleotide archive (LN8886781–LN8886805).

**Departure from neutrality at the class I MHC locus in Bos taurus cattle**

We used neighbor-joining (NJ) trees, calculated separately for the Ankole and Holstein class I MHC datasets according to the Jukes and Cantor (JC) algorithm, as the basis for assessing the relative fits of 56 distinct models of nucleotide substitution. For each model, likelihood scores were calculated in PAUP (Swofford 2002) and the best-fit model for each dataset was selected using the Akaike Information Criterion. We found evolution of both datasets to be consistent with the model TVM+I+G, which assumes variation in base frequencies and transversion rates, as well as presence of invariable sites and rate heterogeneity among sites.

Likelihood settings from the best fit nucleotide substitution model were used in a codon-by-codon analysis to determine the ratio of non-synonymous to synonymous evolutionary changes (dN/dS, \( \omega \)) across each dataset. Specifically, we made maximum likelihood inferences of positive selection based on fitting two sets of codon evolution models and evaluating model fit using the AIC approach—a function of the logarithm of the likelihood (InL), and the number of estimated parameters (K) for each model. The first codon evolution model set comprised the nearly neutral model (M1) and the positive selection model (M2). M1 allows for two site classes: \( p_0 \), proportion of codons undergoing purifying selection...
(0 < $\omega_0 < 1$), and $p_1$, proportion of codons undergoing neutral evolution ($\omega_1 = 1$), while M2 extends M1 by additionally allowing for $p_2$, proportion of codons undergoing positive selection ($\omega_2 > 1$). The second model set assessed $\omega$ based on a $\beta$ distribution indicated by shape parameters $p$ and $q$. We implemented both the null $\beta$ model (M7) which allows for 0 < $\omega_0 < 1$ and the positive selection plus $\beta$ model M8, an extension of M7 to include a proportion of codons undergoing positive selection ($\omega_1 > 1$). The results of the analysis are shown in Table 3.

For both datasets, M2 provided a better fit than M1 indicating that evolution at the class I MHC genes in cattle is driven primarily by positive selection. Independent support for this conclusion came from the models implementing the $\beta$ distribution, with M8 providing better fit to both datasets than M7. Since the $p_2$ parameter for M8 is not substantially different from that of M2, unless explicitly stated, the positive selection remarks hereafter refer to estimates obtained under M2. For the Ankole and Holstein datasets, 9.536 and 6.27% of the codons carried historical signatures of positive selection, respectively. Furthermore, $\omega$ values of 7.2843 and 6.5286 for Ankole and Holstein class I MHC datasets, respectively, were substantially higher than 1. Similar estimates were obtained with M8 (Table 3). For codons identified as carrying signatures of positive selection, Bayes Empirical Bayes (BEB) analysis was used to assess their posterior probability of falling into this site class. For both datasets, the positively selected codons, their corresponding posterior probabilities and their location relative to sites known to be critical for antigen recognition by cattle class I MHC (Hansen et al. 2014) are shown in Fig. 2.

| Working allele name | Read counts | IPD BLASTn search | GenBank accession | Breed known to express allele |
|---------------------|-------------|-------------------|-------------------|------------------------------|
|                     | Exon 2 | Exon 3 | Individuals carrying allele | Allele name | Percent identity |
| BoLA-Ank_02         | 30     | 23     | 2                      | BoLA-2*00501 | 100              | LN886782         | Hereford         |
| BoLA-Ank_05         | 150    | 54     | 2                      | BoLA-2*06001 | 100              | LN886785         | Angus cross      |
| BoLA-Ank_06         | 1550   | 1199   | 3                      | BoLA-2*04701 | 100              | LN886786         | Charolais cross  |
| BoLA-Ank_08         | 133    | 70     | 2                      | BoLA-2*04801 | 100              | LN886787         | Charolais cross  |
| BoLA-Ank_015        | 376    | 281    | 2                      | BoLA-2*05301 | 100              | LN886794         | Charolais cross  |
| BoLA-Ank_01         | 72     | 37     | 1                      | BoLA-1*04901 | 93.17            | LN886781         | Charolais cross  |
| BoLA-Ank_19         | 434    | 1137   | 2                      | BoLA-1*06101 | 97.06            | LN886798         | Angus cross      |
| BoLA-Ank_21         | 315    | 245    | 1                      | BoLA-1*02101 | 96.57            | LN886799         | Holstein         |
| BoLA-Ank_03         | 49     | 37     | 1                      | BoLA-2*04801 | 94.85            | LN886783         | Charolais cross  |
| BoLA-Ank_11         | 2290   | 2139   | 3                      | BoLA-2*06901 | 99.51            | LN886790         | Angus cross      |
| BoLA-Ank_12         | 10     | 5      | 1                      | BoLA-2*03201N | 98.28          | LN886791         | Holstein         |
| BoLA-Ank_13         | 73     | 68     | 1                      | BoLA-2*07101 | 94.36            | LN886792         | Angus cross      |
| BoLA-Ank_16         | 11     | 6062   | 2                      | BoLA-2*03201N | 98.04          | LN886795         | Holstein         |
| BoLA-Ank_17         | 3204   | 3457   | 2                      | BoLA-2*04701 | 98.53            | LN886796         | Charolais cross  |
| BoLA-Ank_18         | 939    | 1007   | 2                      | BoLA-2*03001 | 93.87            | LN886797         | Holstein         |
| BoLA-Ank_27         | 11     | 15     | 2                      | BoLA-2*01602 | 94.12            | LN886803         | Charolais cross  |
| BoLA-Ank_04         | 2014   | 1707   | 2                      | BoLA-3*05002 | 98.28            | LN886784         | Angus cross      |
| BoLA-Ank_09         | 180    | 28     | 2                      | BoLA-3*01101 | 97.06            | LN886788         | Charolais cross  |
| BoLA-Ank_10         | 80     | 11     | 1                      | BoLA-3*01110 | 96.32            | LN886789         | Charolais cross  |
| BoLA-Ank_14         | 30     | 25     | 1                      | BoLA-3*05001 | 97.06            | LN886793         | Charolais cross  |
| BoLA-Ank_26         | 857    | 76     | 2                      | BoLA-3*00403 | 95.34            | LN886802         | Charolais cross  |
| BoLA-Ank_23         | 3024   | 424    | 1                      | BoLA-3*05301 | 97.3             | LN886800         | Charolais cross  |
| BoLA-Ank_25         | 81     | 78     | 2                      | BoLA-6*01501 | 96.57            | LN886801         | Holstein         |
| BoLA-Ank_07         | 83     | 20     | 3                      | BoLA-NC1*00301 | 100          | DQ140370         | Holstein         |
| BoLA-Ank_24         | 111    | 32     | 1                      | BoLA-NC1*00601 | 100          | JN792886         | Holstein         |
| BoLA-Ank_20         | 760    | 271    | 2                      | BoLA-NC2*00102 | 99.75      | Holstein         |
| BoLA-Ank_22         | 104    | 87     | 1                      | BoLA-NC2*00102 | 99.51      | Holstein         |
### Table 2  Exonic sequence diversity present within the Holstein class I MHC genes

| Working allele name  | Allele frequency | IPD BLASTn search    | GenBank accession | Breed known to express allele |
|---------------------|------------------|----------------------|-------------------|-----------------------------|
|                     | No. of plasmid subclones | Individuals carrying allele | Allele name | Percent identity |                     |
| BoLA-Hol_02         | 3                 | 1                    | BoLA-1*00902      | 100                         | DQ121149 Charolais cross |
| BoLA-Hol_19         | 2                 | 2                    | BoLA-1*01901      | 100                         | DQ140362 Holstein       |
| BoLA-Hol_21         | 6                 | 1                    | BoLA-1*02301      | 100                         | Y09208 Holstein         |
| BoLA-Hol_07         | 2                 | 2                    | BoLA-2*01201      | 100                         | DQ304655 Holstein       |
| BoLA-Hol_09         | 1                 | 1                    | BoLA-2*01802      | 100                         | DQ140361 Holstein       |
| BoLA-Hol_11         | 4                 | 2                    | BoLA-2*00802      | 100                         | L02835 Holstein         |
| BoLA-Hol_04         | 19                | 4                    | BoLA-3*00201      | 100                         | X92870 Holstein         |
| BoLA-Hol_08         | 3                 | 1                    | BoLA-3*06601      | 100                         | JN792876 Angus cross    |
| BoLA-Hol_14         | 3                 | 1                    | BoLA-3*05301      | 100                         | DQ121175 Charolais cross |
| BoLA-Hol_15         | 3                 | 2                    | BoLA-3*01701      | 100                         | Y09206 Holstein         |
| BoLA-Hol_17         | 4                 | 2                    | BoLA-3*00401      | 100                         | JN792860.1 Holstein     |
| BoLA-Hol_20         | 3                 | 1                    | BoLA-3*01101      | 100                         | M24090 Holstein         |
| BoLA-Hol_12         | 3                 | 1                    | BoLA-6*01402      | 100                         | DQ140358 Holstein       |
| BoLA-Hol_05         | 5                 | 3                    | BoLA-2*02201      | 98.04                       | LN886804                |
| BoLA-Hol_10         | 3                 | 2                    | BoLA-2*00501      | 93.64                       | LN886805                |
| BoLA-Hol_06         | 4                 | 1                    | BoLA-3*06501      | 97.56                       |                         |
| BoLA-Hol_22         | 1                 | 1                    | BoLA-1*01901      | 99.76                       |                         |
| BoLA-Hol_16         | 4                 | 4                    | BoLA-NC1*00301    | 100                         | DQ140370 Holstein       |
| BoLA-Hol_18         | 4                 | 3                    | BoLA-NC1*00501    | 100                         | JN792885 Angus cross    |
| BoLA-Hol_01         | 3                 | 3                    | BoLA-NC2*00102    | 100                         | DQ140372 Holstein       |
| BoLA-Hol_03         | 6                 | 5                    | BoLA-NC3*00101    | 100                         | DQ140373 Holstein       |
| BoLA-Hol_23         | 1                 | 1                    | BoLA-NC1*00701    | 98.58                       |                         |
| BoLA-Hol_13         | 1                 | 1                    | BoLA-NC3*00101    | 99.27                       |                         |

### Table 3  The Akaike information criterion (AIC) evaluation of the goodness of fit for different models of codon evolution

| Breed     | Model | LnL | K   | AIC   | Estimated ω | Site classes | β distribution |
|-----------|-------|-----|-----|-------|-------------|--------------|----------------|
|           |       |     |     |       | ω₀   | ω₁   | ω₂   | p₀   | p₁   | p₂   | p   | q   |
| Ankole    | M1    | -2306.153228 | 46   | 4704.306456 | 0.06109   | 1   | 7.2823 | 0.72527 | 0.27473 |
|           | M2    | -2232.131755 | 48   | 4560.26351  | 0.0605    | 1   | 7.2823 | 0.61741 | 0.28274 | 0.09536 |
|           | M7    | -2309.319323 | 46   | 4710.638646 | 0.0605    | 1   | 7.2823 | 0.14694 | 0.32814 |
|           | M8    | -2233.607141 | 48   | 4563.214282 | 0.0605    | 1   | 7.2823 | 0.14694 | 0.32814 |
| Holstein  | M1    | -2021.65299  | 34   | 4111.30598  | 0.04266   | 1   | 6.5286 | 0.07126 | 0.10506 | 0.20035 |
|           | M2    | -1976.487153 | 36   | 4024.974306 | 0.04347   | 1   | 6.5286 | 0.07126 | 0.10506 | 0.20035 |
|           | M7    | -2025.181287 | 34   | 4111.365764 | 0.04266   | 1   | 6.5286 | 0.07126 | 0.10506 | 0.20035 |
|           | M8    | -1979.071619 | 36   | 4030.143238 | 0.04266   | 1   | 6.5286 | 0.07126 | 0.10506 | 0.20035 |

The three site classes, purifying selection, neutral selection and positive selection, are denoted by ω₀, ω₁ and ω₂, respectively, while p₀, p₁ and p₂ indicate the proportion of codons belonging to each of the three site categories, respectively. β distribution shape parameters are denoted p and q.
In silico analysis of peptide-binding motifs for Ankole and Holstein class I MHC proteins

Using the sequences described in the current study as input for in silico analysis, we sought to compare the predicted peptide binding motifs of the Ankole and Holstein class I MHC transcripts as well as investigate the extent to which the motifs overlap. Overlaps in the NetMHCpan predicted peptide-binding motifs, conveyed in a specificity heat-map and a functional tree, were assessed using MHCcluster (Fig. 3). A prominent feature of the results of these ligand selectivity analyses was the short specificity distances among the Ankole class I MHC sequences, indicated by their juxtaposition on the heat-map. Our results suggest considerable dissimilarity in the peptide binding preferences between the Ankole and Holstein class I MHC products, as indicated by the specificity distances between them on the heat-map. The differences in the predicted peptide binding preferences are further depicted as clusters in the functional tree, while the functional repertoire of Ankole class I MHC is exemplified by sequence logo representations of predicted binding motifs (Fig. 4). Based on the height of the stack of amino acid residues, and as indicated by the cattle class I MHC allelic variants restricting known *T. parva* CD8$^+$ T cell antigens (Hansen et al. 2014), scrutiny of the predicted specificity pattern for all the Ankole sequences revealed that besides the C-terminal residues, side chains at position 2 had the second highest relative contribution to peptide binding (Fig. 4). Additional variations were predicted at other motif positions possibly resulting in additional functional divergence among the Ankole class I MHC products.

Recognition of *T. parva* candidate vaccine antigens in seropositive Ankole cattle

Studies using immune bovine CD8$^+$ T cell lines from Holstein-Friesian cattle and the Boran breed of *B. indicus* to screen expressed parasite cDNA libraries have identified *T. parva* candidate vaccine antigens, designated Tp (Graham et al. 2006). Using these antigens as stimuli, we assessed their recognition in *T. parva* seropositive Ankole cattle by measuring cytokine (IFN-γ) responses. As a pre-requisite, we first determined the titers of *T. parva*-specific antibodies in Ankole sera by quantitative antibody-detection ELISA based on the PIM antigen (Katende et al. 1998). The
result of this analysis, expressed as a percentage of the strong positive control serum, hereafter referred to as percent positivity (PP), is shown in Fig. 5. Eleven of the 40 Ankole from the field screened were found to be antibody positive for *T. parva*.

Detection and enumeration of IFN-γ secreting cells in the peripheral circulation of *T. parva* seropositive Ankole cattle were assessed ex vivo by ELISpot upon stimulation of PBMCs from the 11 seropositive cattle, with overlapping peptide pools derived from each of the defined *T. parva* candidate vaccine antigens (Graham et al. 2006). The breadth and magnitude of this cytokine response for each of the 11 positive animals are shown in Fig. 6 as SFU per million PBMCs.

The no peptide control mean numbers of IFN-γ SFU per 1.0 × 10^6 PBMCs were generally markedly less than those obtained from CD8+ T cell lines (derived from animal BV115) with known Tp1 antigenic specificity that served as a positive control for the assay. The PBMCs from animals Ankole-2 and Ankole-17 were positive in repeat experiments for peptide pools derived from antigens Tp4 and Tp5, respectively (see Fig. 6). However, when the ex vivo IFN-γ ELISpot assays were repeated using CD8+ T cell fractions purified from the PBMCs, the responses from all the Tp antigens were markedly lower than the BV115 positive control, and none attained the empirical threshold of three times mean SFU in Tp stimulated wells (Janetzki et al. 2015) compared to negative control wells (data not shown).

**Discussion**

Whilst the expressed bovine class I MHC gene repertoire has been examined extensively in European taurines in a number of studies, published analyses of class I MHC transcripts in African cattle is limited (IPD; www.ebi.ac.uk/ipd/mhc/bola). Consequently, there is still uncertainty as to the likely impacts of class I MHC diversity on the coverage of *T. parva* CD8+ T cell epitope-based vaccines in endemic regions. The current study utilized cDNA from a population of African taurine cattle (Ankole) as templates for the amplification of the domains of the class I MHC heavy chain that interact with antigenic peptides, in order to provide an initial assessment of the levels of class I MHC diversity among African taurine cattle. The amplicon libraries were typed by pyrosequencing using the Roche 454 GS-FLX chemistry, an approach that has proved useful in a wide range of situations, notably cataloguing of the complement of class I MHC transcripts in organisms with high levels of complexity at the class I MHC locus (O’Leary et al. 2009).

Whilst the catalogue of Holstein class I MHC transcripts obtained in the current study was in general agreement with
published data ([www.ebi.ac.uk/ipd/mhc/bola](http://www.ebi.ac.uk/ipd/mhc/bola)), the results of our evaluation of the Ankole transcripts pointed to minimal overlap in the ligand selectivity of the class I MHC products between the two breeds. Foremost, our data suggest a high level of inter-taurine sequence divergence, with the great majority of the Ankole class I MHC allelic sequences described in the present study being designated as putatively novel (Table 1). Notably, the nucleotide BLAST searches of the cattle section of the IPD revealed that in general the Ankole class I MHC allelic variants shared higher nucleotide identities with sequences that are not known to be expressed in the extensively studied Holstein breed. In contrast, using the

![Fig. 4](image1.png) **Comparisons of predicted peptide-binding motifs for Ankole and Holstein class I MHC products.** The MHCcluster 2.0 procedure leverages the predictions generated by NetMHCpan 2.8 to group the class I MHC molecules into a number of functional clusters. On the MHC functional distance tree, Ankole class I MHC transcripts are coloured red if they are novel or marked with an asterisk (*) if they have been described previously. Novel Holstein sequences are prefixed ‘Hol’, while previous described ones are prefixed ‘BoLA’ without an asterisk and are all coloured black. On the representative predicted binding motif sequence logos, enriched amino acids (AA) are shown on the positive y-axis and depleted (under-represented) amino acids on the negative y-axis. The height of the stack of AA residues represents the level of conservation at a particular position, while the relative residue height denotes its frequency. Polar AA are depicted in green, basic AA in blue, acidic AA in red and hydrophobic AA in black.

![Fig. 5](image2.png) **Sero-prevalence of T. parva in field Ankole cattle.** The centre line shows the medians; box limits indicate the 25th and 75th percentiles as determined by R software; whiskers extend 1.5 times the interquartile range from the 25th and 75th percentiles; the cross represents the sample mean; the grey bar indicate 95% confidence intervals of the mean; data points are plotted as open circles. Results are given as a percentage of the strong positive control serum (percent positivity, PP). Threshold PP = 20 and n = 40 Ankole cattle.
Holstein allelic sequences described herein as the query sequences, BLASTN similarity searches showed that with few exceptions, one of which was dramatically distinct in sequence, the African Holstein sequences shared 100% sequence similarity to already published cattle class I MHC sequences (Table 2).

The results of the codon-based statistical tests that we performed separately on the Ankole and Holstein class I MHC datasets revealed patterns of evolution that markedly exceeded those predicted by neutrality. Positive selection of amino acid diversity was evident from the highly contrasting levels of non-synonymous versus synonymous substitutions and the list of codons identified as carrying signatures of positive selection included those that code for residues that line the functionally important F pockets. Taken together, the evidence that the polymorphisms are generated and maintained by positive selection for amino acid diversity, and the concordance between the location of positively selected sites and known antigen binding sites are consistent with potential differences in the spectrum of T. parva epitopes recognized by CD8+ T cells.

Despite the marked differences in exonic sequence diversity within the peptide binding regions of class I MHC allelic sequences expressed in Ankole and Holstein cattle, allelic polymorphism per se is not necessarily sufficient to infer functional differences. A variety of algorithms have been used for prediction of peptide binding to class I MHC and a tool called MHCcluster that functionally clusters MHC molecules based on their predicted binding specificity proved particularly
useful when applied to our analysis of Ankole and Holstein class I sequences. The classification of allelic sequences based on their predicted specificity distances, as indicated by different intensities of the colors in a heat-map (see Fig. 3), resulted in the juxtaposition of the Ankole alleles, while the Holstein alleles formed separate functional specificity clusters. Although a comprehensive definition of cattle class I MHC ‘supertypes’ is beyond the scope of the current study, it may be reasonable to conclude that the observed differences in the predicted ligand selectivity might result in differential peptide recognition by CD8+ T cells from the two breeds.

More directly relevant to ECF vaccine development is the question of whether the dissimilarities seen in in silico class I MHC ligand predictions are sustained in ex vivo tests assaying differential cytokine responses to defined T. parva antigen between the two taurine lineages. To address this, we employed pools of peptides derived from T. parva vaccine candidate antigens selected by immunoscreening of T. parva cDNA libraries using Holstein and Boran cattle immunized by ITM with the T. parva Muguga stock (Graham et al. 2008). These peptides were utilized to stimulate PBMCs and also purified CD8+ T cells isolated from T. parva serpositive Ankole cattle and assayed by gamma interferon ELISpot. Two of the 11 seropositive Ankole cattle were reproducibly positive for antigens Tp4 and Tp5, respectively. However, cytokine responses could not be demonstrated using purified Ankole CD8+ T cells. It seems possible that the Tp4 and Tp5 ELISpot responses may have been the result of CD4+ T cell responses. Whether polymorphisms within the target antigens in the Ankole population or the time point of infection in the field Ankole cattle contributed to the negative outcome of the IFN-γ ELISpot assay using purified CD8+ T cells remains to be demonstrated. However, the failure of purified CD8+ T cells from the Ankole cattle sampled in this study to recognize antigens previously shown to contain CD8+ T cell epitopes recognized in the context of other bovine haplotypes is worthy of further investigation.

Our findings, highlighting inter-taurine differences in class I MHC diversity and functional capability, are also broadly consistent with, and extend, findings from studies that examined recognition of the same T. parva candidate vaccine antigens by CD8+ T cells derived from Zebu (B. indicus) cattle (Akoolo et al. 2008). With the exception of the Tp2 antigen, this study showed that the known T. parva CD8+ T cell target candidate vaccine antigens were not recognized by CD8+ T cells from any of the animals studied. With respect to Tp2, it concluded that this T. parva candidate antigen constituted only a minor CD8+ T cell specificity in the single animal that appeared to be responding and importantly, noted that the target Tp2 epitope was distinct from those previously identified using CD8+ T cells derived from B. taurus cattle.

The data that we report herein reveals a surprising degree of functional divergence within the class I MHC of two different but relatively closely related taurine lineages and emphasizes the value of the MHCcluster software tool relative to simple comparison of sequences. Given these findings, it will be important to apply similar approaches, more comprehensively to the Ankole breed and also to a wider range of African cattle populations. This will contribute to understanding whether class I ‘supertypes’ can be defined for African cattle and the number of antigens/epitopes that may need to be incorporated in a recombinant antigen cocktail vaccine that is effective for control of ECF in the field.

Acknowledgments We are grateful to Dr. Joerg Jores of ILRI for reviewing the manuscript. We also wish to thank Giles Prettejohn of the Ol Pejeta Conservancy in Kenya whose efforts were instrumental to the completion of this study. The study was funded by the DFG German-African Cooperation Projects in Infectiology: ‘Molecular epidemiology network for promotion and support of delivery of life vaccines against Theileria parva and Theileria annulata infection in Eastern and Northern Africa (SE 862/2-1)’. Some of the work described in this paper was also supported by the CGIAR Consortium research project CRP3.7.

References

Akoolo L, Pellé R, Saya R, Awino E, Nyanjui I, Taracha EL, Kanyari P, Mwangi DM, Graham SP (2008) Evaluation of the recognition of Theileria parva vaccine candidate antigens by cytotoxic T lymphocytes from Zebu cattle. Vet Immunol Immunopathol 121(3–4):216–221

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. J Mol Biol 215(3):403–410

Anderung C, Hellborg L, Seddon J, Hanotte O, Götterström A (2007) Investigation of X- and Y-specific single nucleotide polymorphisms in taurine (Bos taurus) and indicine (Bos indicus) cattle. Anim Genet 38(6):595–600

Babik W, Kawalko A, Wojcik JM, Radwan J (2012) Low major histocompatibility complex class I variation in the European bison (Bison bonasus). J Hered 103:349–359

Birch J, Murphy L, MacHugh ND, Ellis SA (2006) Generation and maintenance of diversity in the cattle MHC class I region. Immunogenetics 58(6):670–679

Codner GF, Birch J, Hammond JA, Ellis SA (2012) Constraints on haplotype structure and variable gene frequencies suggest a functional hierarchy within cattle MHC class I. Immunogenetics 64(6):435–445

Di Giulio G, Lynen G, Morzaria S, Oura C, Bishop R (2009) Live immunization against East Coast fever—current status. Trends Parasitol 25(2):85–92

Di Palma F, Archibald SD, Young JR, Ellis SA (2002) A BAC contig of approximately 400 kb contains the classical class I major histocompatibility complex (MHC) genes of cattle. Eur J Immunogenet 29(1):65–68

Dodd M, Roehr J, Ahmed R, Dieterich C (2012) FLEXBAR—flexible barcode and adapter processing for next-generation sequencing platforms. Biology 1(3):895–905

Ellis S (2004) The cattle major histocompatibility complex: is it unique? Vet Immunol Immunopathol 102(1–2):1–8

Ellis SA, Holmes EC, Staines KA, Smith KB, Stear MJ, McKeever DJ, MacHugh ND, Morrison WJ (1999) Variation in the number of
expressed MHC genes in different cattle class I haplotypes. Immunogenetics 50(5–6):319–328

Emery DL, Eugui EM, Nelson RT, Tenywa T (1981) Cell-mediated immune responses to Theileria parva (East Coast fever) during immunization and lethal infections in cattle. Immunology 43(2):323–336

Eugui EM, Emery DL (1981) Genetically restricted cell-mediated cytotoxicity in cattle immune to Theileria parva. Nature 290(5803):251–254

Graham SP, Pellé R, Honda Y, Mwangi DM, Tonukari NJ, Janetzki S, Price L, Schroeder H, Britten CM, Welters MJ, Hoos A (2015) MacHugh DE, Shriver MD, Loftus RT, Cunningham P, Bradley DG Nene V, Buus S, Nielsen M (2014) Characterization of binding specificities of bovine leukocyte class I molecules: impacts for rational epitope discovery. Immunogenetics 66(12):705

Graham SP, Pellé R, Honda Y, Mwangi DM, Glew EJ, de Villiers EP, Shah T, Bishop R, Ayuba E, Awino E, Gachanja J, Luayai AE, Mbwika F, Muthiani AM, Ngewma DM, Njairu M, Nyanjui JK, Onono FO, Ososo J, Saya RM, Widmann C, Fraser CM, Maudlin I, Gardner MJ, Morzarizza SP, Loosmore S, Gilbert SC, Audonnet JC, van der Bruggen P, Nene V, Taracha EL. (2006) Theileria parva candidate vaccine antigens recognized by immune bovine cytotoxic T lymphocytes. Proc Natl Acad Sci U S A 103(9):3286–3291

Graham S, Honda Y, Pelle R, Mwangi DM, Glew EJ, de Villiers EP, Shah T, Bishop R, van der Bruggen P, Nene V, Taracha ELN (2007) A novel strategy for the identification of antigens that are recognised by Bovine MHC class I restricted cytotoxic T cells in a protozoan infection using reverse vaccinomics. Immunome Res 3:2

Graham SP, Pellé R, Yamag M, Mwangi DM, Honda Y, Mwakambamba RS, de Villiers EP, Ayuba E, Awino E, Gachanja J, Mbwika F, Muthiani AM, Muriuki C, Nyanjui JK, Onono FO, Ososo J, Ritho V, Saya RM, Ellis SA, McKeever DJ, MacHugh ND, Gilbert SC, Audonnet JC, Morrison WI, van der Bruggen P, Taracha EL (2008) Characterization of the fine specificity of bovine CD8 T-cell responses to defined antigens from the protozoan parasite Theileria parva. Infect Immun 76(2):685–694

Hansen AM, Rasmussen M, Svitik N, Hamdahl M, Golte WT, Barlow J, Nene V, Buus S, Nielsen M (2014) Characterization of binding specificities of bovine leucocyte class I molecules: impacts for rational epitope discovery. Immunogenetics 66(12):705–718

Hamdahl M, Justesen S, Lambeth K, Roder G, Nielsen M, Buus S (2009) Peptide binding to HLA class I molecules: homogenous, high-throughput screening, and affinity assays. J Biomol Screen 14:173–180

Holmes EC, Roberts AF, Staines KA, Ellis SA (2003) Evolution of major histocompatibility complex class I genes in Cetartiodactyls. Immunogenetics 55(4):193–202

Hooi I, Peters B, Sidney J, Pedersen LE, Sette A, Lund O, Buus S, Nielsen M (2009) NetMHCpan, a method for MHC class I binding prediction beyond humans. Immunogenetics 61(1):1–13

Janetzki S, Price L, Schroeder H, Britten CM, Welten MJ, Hoos A (2015) Guidelines for the automated evaluation of Elistop assays. Nat Protoc 10(7):1098–1115

Katende J, Morzariza S, Toye P, Skilton R, Nene V, Nkonge C, Musoke A (1998) An enzyme-linked immunosorbent assay for detection of Theileria parva antibodies in cattle using a recombinant polymorphic immunodominant molecule. Parasitol Res 84(5):408–416

Katoh K, Standley DM (2013) MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol 30(4):772–780

MacHugh DE, Shriver MD, Loftus RT, Cunningham P, Bradley DG (1997) Microsatellite DNA variation and the evolution, domestication and phyleogeography of taurine and zebu cattle (Bos taurus and Bos indicus). Genetics 146(3):1071–1086

McKeever DJ, Taracha EL, Innes EL, MacHugh ND, Awino E, Goddeeris BM, Morrison WI (1994) Adoptive transfer of immunity to Theileria parva in the CD8+ fraction of responding effferent lymph. Proc Natl Acad Sci U S A 91(5):1959–1963

Musoke AJ, Nantulya VM, Buscher G, Masake RA, Otim B (1982) Bovine immune response to Theileria parva: neutralizing antibodies to sporozoites. Immunology 45:663–668

Nene V, Gobright E, Bishop R, Morzariza S, Musoke A (1999) Linear peptide specificity of bovine antibody responses to p67 of Theileria parva and sequence diversity of sporozoite-neutralizing epitopes: implications for a vaccine. Infect Immun 67:1261–1266

Nene V, Svitik N, Toye P, Golte WT, Barlow J, Hamdahl M, Buus S, Nielsen M (2012) Designing bovine T cell vaccines via reverse immunology. Ticks Tick Borne Dis 3(3):188–192

Nielsen M, Lundegaard C, Blicher T, Lambeth K, Hamdahl M, Justesen S, Roder G, Peters B, Sette A, Lund O, Buus S (2007) NetMHCpan, a method for quantitative predictions of peptide binding to any HLA-A and -B locus protein of known sequence. PLoS One 2(8):e796

Norval RAI, Perry BD, Young AS (1992) The epidemiology of Theileria in Africa. Academic, London

Obara I, Ulrike S, Musoke T, Spooner PR, Jabbar A, Odongo D, Kemp S, Silva JC, Bishop RP (2015) Molecular evolution of a central region containing B cell epitopes in the gene encoding the p67 sporozoite antigen within a field population of Theileria parva. Parasitol Res 114(5):1729–1737

O’Leary CE, Wiseman RW, Karl JA, Bimber BN, Lank SM, Tuscher JJ, O’Conor DH (2009) Identification of novel MHC class I sequences in pig-tailed macaques by ampiclon pyrosequencing and full-length cDNA cloning and sequencing. Immunogenetics 61(10):689–701

Paling RW, Mpangala C, Luttkhuizen B, Sibomana G (1991) Exposure of Ankole and crossbred cattle to Theileriosis in Rwanda. Trop Anim Health Prod 23(4):203–214

Pandya M, Rasmussen M, Hansen A, Nielsen M, Buus S, Golde W, Barlow J (2015) A modern approach for epitope prediction: identification of foot-and-mouth disease virus peptides binding bovine leukocyte antigen (BoLA) class I molecules. Immunogenetics 67(11):691–703

Pedersen LE, Hamdahl M, Rasmussen M, Lambeth K, Golde WT, Lund O, Nielsen M, Buus S (2011) Porcine major histocompatibility complex (MHC) class I molecules and analysis of their peptide-binding specificities. Immunogenetics 63(12):821–834

Radley DE, Brown CGD, Burridge MJ, Cunningham MP, Kirimi IM, Purnell RE, Young AS (1975) Chemophylactic immunisation of cattle against Theileria parva (Muguga) and five Theileria strains. Vet Parasitol 1:35–41

Robinson J, Halliwell JA, McWilliam H, Lopez R, Marsh SG (2013) IPD—the immuno polymorphism database. Nucleic Acids Res 41(Database issue):D1234–D1240

Sette A, Sidney J (1999) Nine major HLA class I supertypes account for the vast preponderance of HLA-A and -B polymorphism. Immunogenetics 50(4–5):201–212

Sommer S, Courtiol A, Mazzenzi CJ (2013) MHC genotyping of non-model organisms using next-generation sequencing: a new methodology to deal with artefacts and allelic dropout. BMC Genomics 14:542

Steinaa L, Saya R, Awino E, Toye P (2012) Cytotoxic T lymphocytes from cattle immunized against Theileria parva exhibit pronounced cross-reactivity among different strain-specific epitopes of the Tp1 antigen. Vet Immunol Immunopathol 145(3–4):571–581

Svitik N, Hansen AM, Steinaa L, Saya R, Awino E, Nielsen M, Buus S, Nene V (2014) Use of ‘one-pot, mix-and-read’ peptide-MHC class I tetramers and predictive algorithms to improve detection of cytotoxic T lymphocyte responses in cattle. Vet Res 45:50

Swoford DL (2002) PAUP: phylogenetic analysis using parsimony and other methods. Version 4. Sinauer Associates, Sunderland

Thomsen M, Lundegaard C, Buus S, Lund O, Nielsen M (2013) MHClcluster, a method for functional clustering of MHC molecules. Immunogenetics 65(9):655–665