final report

Project code: B.AHE.0213

Prepared by: Dr Cheryl Jenkins
NSW Department of Primary Industries

Date published: 16 March 2016

ISBN: 9781741919967

PUBLISHED BY
Meat and Livestock Australia Limited
Locked Bag 1961
NORTH SYDNEY NSW 2059

Bovine theileriosis: Molecular diagnosis and strain analyses

Meat & Livestock Australia acknowledges the matching funds provided by the Australian Government to support the research and development detailed in this publication.

Abstract

Bovine theileriosis is an emerging disease in Australasia caused by the red blood cell parasite, *Theileria orientalis* Ikeda. This disease has caused substantial economic losses to beef and dairy industries in Australia since 2006, estimated at $20 million per annum. The development of sensitive, rapid and cost-effective methods for diagnosis of this disease have been considered paramount to assist producers in managing their herds. Furthermore, in the absence of chemotherapeutic options to treat bovine theileriosis in Australia, a vaccine for this parasite is considered the preferred option for disease control. In this project a multiplex molecular test for the detection of *T. orientalis* and differentiation of the clinically relevant genotypes was developed and validated. This assay provides veterinarians with quantitative data on parasite levels within samples and a clinical threshold for parasite levels, delineating subclinical from clinically relevant infection, has been established. This multiplex test is the most cost-effective, sensitive and specific assay currently available for diagnosis of bovine theileriosis in Australia. A serological test for this organism has also been developed which can be applied to herd-level screening. Genomic and global proteomic analyses of *T. orientalis* genotypes conducted in this study revealed that the *T. orientalis* genotypes are genetically divergent (at least different subspecies), potentially reducing the utility of a live vaccine approach to disease control. Nonetheless the genomic data derived from this project lays the groundwork for future vaccine development work aimed at using reverse vaccinology to develop a subunit vaccine for this organism.

Executive summary

Theileria orientalis is an emerging pathogen of cattle in Asia, Australia and New Zealand, causing mortalities and ill-thrift in recent years. This organism is a vector-borne haemoprotozoan causing clinical disease characterised by anaemia, abortion and death, as well as persistent subclinical infections. Molecular methods of diagnosis are preferred due to their sensitivity and utility in differentiating between pathogenic and apathogenic genotypes. Conventional PCR (cPCR) assays for *T. orientalis* detection and typing are laborious and do not provide an estimate of parasite load. Current real-time PCR assays cannot differentiate between clinically-relevant and benign genotypes or are only semi-quantitative without a defined clinical threshold. Here, we developed and validated a hydrolysis probe quantitative PCR (qPCR) assay which universally detects and quantifies *T. orientalis* and identifies the clinically-associated Ikeda and Chitose genotypes (UIC assay). Comparison of the UIC assay with previously validated universal and genotype-specific cPCRs demonstrated that qPCR detects and differentiates *T. orientalis* with high sensitivities and specificities. A uniplex Buffeli assay was also developed for epidemiological studies on this benign genotype. The assays developed here compare favourably with a commercial multiplex tandem PCR (MT-PCR) assay in terms of sensitivity, out-perform MT-PCR in terms of genotype specificity and are considerably less expensive to run overall. Assay cost can be reduced further using an inexpensive DNA extraction method which has been developed for use with the UIC qPCR. Comparison of quantitative results from UIC qPCR with percentage parasitaemia determined via blood film analysis and packed cell volume (PCV) revealed significant positive and negative correlations respectively. One way analysis of variance (ANOVA) indicated that blood samples from animals with clinical signs of disease contained statistically higher concentrations of *T. orientalis* DNA than animals with subclinical infections. In this study we established clinical thresholds to assist in classifying high, moderate and low-level infections and describe how parasite load and the presence of the Ikeda and Chitose genotypes relates to disease.

In Australia, outbreaks of clinical theileriosis caused by *Theileria orientalis* have been largely associated with the Ikeda genotype which can occur as a sole infection, or more commonly, as a mixture of genotypes. The most prevalent genotype, Chitose, frequently co-occurs with type Ikeda, however the role of this genotype in clinical disease has not been clearly established. Furthermore, the dynamics of individual genotypes in field infection of cattle have not been examined. Quantitative PCR and genotyping methods were used to examine the role of the Chitose genotype in clinical disease and to investigate the temporal dynamics of *T. orientalis* Ikeda, Chitose and Buffeli genotypes in naïve animals introduced to a *T. orientalis*-endemic area. Analysis of the major piroplasm surface protein (MPSP) genes and internal transcribed spacers (ITS1) of Ikeda isolates revealed a high level of conservation within the population indicating that Australian populations of this parasite are relatively homogenous. In contrast, Chitose isolates revealed the presence of two distinct phylogenetic clusters, Chitose A and Chitose B. A genotyping assay aimed at determining Chitose A/B allele frequency revealed that the Chitose A phylogenetic cluster is strongly associated with clinical disease but nearly always co-occurs with the Ikeda genotype. qPCR revealed that the Chitose genotype (particularly Chitose A), undergoes temporal switching in conjunction with the Ikeda genotype and contributes substantially to the overall parasite burden. This information is of direct relevance to qPCR diagnosis, because at certain stages of infection, the Chitose genotype may dominate. The benign Buffeli genotype can also

undergo temporal switching but levels of this genotype appear to remain low relative to the Ikeda and Chitose types. Interplay between vector and host immunological factors is presumed to be critical to the population dynamics observed in this study and genotypic switching likely contributes to the persistence of *T. orientalis* in the host.

This report contains substantial new genomic and proteomic data on *T. orientalis*. The genomes of Australian strains of Ikeda, Chitose and Buffeli genotypes were sequenced using Illumina technology and were aimed at determining the level of genetic variability amongst and within the genotypes. Proteomic analyses were conducted using a global Liquid Chromatography Mass Spectrometry (LC-MS/MS) approach. Genomic data revealed extensive genetic differences between the Ikeda, Chitose and Buffeli genotypes o*f T. orientalis* which suggest that these genotypes represent at least different subspecies. The Ikeda isolate sequenced in this study closely resembled that of the Japanese Shintoku strain isolated in Japan. Illumina sequencing revealed that individual field samples of *T. orientalis* of a single MPSP genotype nonetheless contain significant population diversity (ie: each sample is composed of a large number of different haplotypes). The intra-sample diversity was significantly greater within the Chitose and Buffeli samples compared to the Ikeda sample, pointing to limited opportunity for genetic recombination (ie: genetic “inbreeding”) within *T. orientalis* Ikeda resulting from the recent introduction of this organism from a small population of parasites. Using proteomic approaches we identified the major proteins present in the red blood cell stage of *T. orientalis* Ikeda as well as those from benign *T. orientalis* genotypes. Identification of the protein components present in *T. orientalis* Ikeda can now be used to inform a) what specific features of *T. orientalis* Ikeda result in the ability to cause clinical infections b) how cattle respond to *T. orientalis* Ikeda infection immunologically and c) which proteins might represent suitable chemotherapeutic or vaccine targets for treatment or prevention of bovine theileriosis. Both the genomic and proteomic studies undertaken in this project will form the foundation for future vaccine development efforts. Prior reports have focussed on the potential to use the benign Buffeli genotype as a live vaccine against *T. orientalis* Ikeda; however this study has demonstrated that these organisms may be too divergent to facilitate the use of Buffeli as a live vaccine. Furthermore, epidemiological studies have not as yet revealed cross-protection between these genotypes. The development of subunit vaccines against Apicomplexan parasites is generally considered problematic due to extreme variability within parasite populations, allowing the organism to evade the host immune response. However, the relative homogeneity of the Ikeda genotype within Australian cattle presents a good opportunity for the development of a subunit vaccine. Without further introductions of the parasite to enable genetic recombination, the Ikeda genotype within Australia is likely to continue to be inbred.

A further tool designed and validated in this project which is relevant to future vaccine development efforts, is a serological test for *T. orientalis*. While host immunological responses to the lymphocyte transforming *Theileria* species, *T. parva* and *T. annulata*, have been well-studied, little is known about the immune response to this non-transforming species. We developed a recombinant antigen ELISA based on the major piroplasm surface protein (MPSP) of *T. orientalis* and investigated whether seroconversion to the MPSP was associated with clinical factors (anaemia), parasite burden and parasite genotype. We also examined the dynamics of seroconversion in animals acutely infected with *T. orientalis*. In cattle testing qPCR positive for *T. orientalis*, seroconversion was significantly more frequent in anaemic compared to normal cattle. The ELISA ratio (ER) was highly correlated with total

parasite burden as measured by qPCR; however when loads of individual genotypes of the parasite were examined, only the pathogenic Ikeda genotype was highly correlated with ER. Conversely, seroconversion was less frequently detected in the presence of benign *T. orientalis* genotypes. Temporal measurement of the serological response, parasite burden and packed cell volume (PCV) in acutely infected animals revealed that seroconversion to the MPSP occurs within 2-3 weeks of the initial qPCR detection of the parasite and coincides with a peak in infection intensity and a declining PCV. Whether a sustained serological response to the MPSP is immunoprotective against re-infection with the parasite or infection recrudescence requires further investigation; however the MPSP represents a promising target for a subunit vaccine given that genetic variability within the MPSP is correlated with differential pathogenicity of *T. orientalis*.

Table of Contents

Abbreviations and acronyms

CDS:	Coverage and depth of sequence
dNTPs:	deoxyribose nucleoside triphosphates.
DPK:	Detergent Proteinase K extraction method.
ELISA:	enzyme-linked immunosorbent assay
GC/µL:	gene copies per microliter.
GC/µL PE:	gene copies /µL packed erythrocytes.
ITS:	internal transcribed spacers.
MGB:	Minor groove binder.
MPSP:	Major piroplasm surface protein.
MT-PCR:	Multiplex tandem PCR.
NFQ:	non-fluorescent quencher.
SDS-PAGE:	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
UIC:	qPCR which universally detects and quantifies T orientalis and identifies Ikeda and Chitose genotypes (“Universal Ikeda Chitose PCR”)

1 Background

1.1 Development of a multiplex quantitative PCR for T. orientalis genotypes

Theileria orientalis is a vector-borne haemoprotozoan infecting cattle and buffalo and is generally spread by ticks of the *Haemaphysalis* genus (Fujisaki et al., 1994; Hammer et al., 2015; Higuchi, 1986, 1987). From a clinical perspective, *T. orientalis* can cause anaemia, lethargy, jaundice, fever, abortion and mortality in cattle (Eamens et al., 2013c). *T. orientalis* is a conditional pathogen, and while pathogenic forms are largely limited to Eastern Asia and Australasia (Eamens et al., 2013c; Islam et al., 2011; Kang et al., 2012; McFadden et al., 2016; McFadden et al., 2011; Ota et al., 2009; Pulford et al., 2016; Sivakumar et al., 2014), it is frequently detected in asymptomatic animals and these benign forms are globally spread (Jeong et al., 2010; Kakuda et al., 1998; Khukhuu et al., 2011; Sivakumar et al., 2014; Sivakumar et al., 2013; Wang et al., 1998). Analysis of the most common genotyping locus, encoding the variable MPSP, has currently identified eleven distinct *T. orientalis* genotypes (Sivakumar et al., 2014). Of these genotypes, Type 2 (Ikeda), and to a lesser extent, Type 1 (Chitose), are typically found in association with clinical disease (Eamens et al., 2013b; Eamens et al., 2013c; Kamau et al., 2011a; Lawrence et al., 2016; McFadden et al., 2011; Ota et al., 2009). The presence of pathogenic and benign forms of *T. orientalis* greatly complicates its clinical diagnosis, with standard blood film analysis unable to identify the pathogenic genotypes.

Multiple conventional PCR (cPCR) assays have been published for the identification of *T. orientalis* in blood samples (Cufos et al., 2012; Ota et al., 2009; Tanaka et al., 1993; Zakimi et al., 2006). The most commonly cited assays detect and genotype *T. orientalis* by amplifying unique regions of the MPSP gene (Zakimi et al., 2006). These assays use two universal primers to detect *T. orientalis* infection and specific forward primers to identify Ikeda, Chitose and Buffeli genotypes. While this method is highly sensitive and has been validated in prior studies (Eamens et al., 2013a; Eamens et al., 2013b; Eamens et al., 2013c), it is not multiplexed and is highly susceptible to PCR inhibitors (Eamens et al., 2013c), which are often found in nucleotide extractions sourced from blood (Al-Soud et al., 2000; Al-Soud and Radstrom, 2001). To overcome inhibition, undiluted and diluted nucleotide extracts can be examined in parallel to prevent false negative results (Eamens et al., 2013a; Eamens et al., 2013b; Eamens et al., 2013c). However, if multiple reactions per sample are required to determine infection presence, the procedure becomes both expensive and time-consuming. Furthermore, cPCR does not give an accurate representation of parasite load and therefore cannot provide an indication of the severity of *T. orientalis* infection. A recently developed multiplexed-tandem PCR (Perera et al., 2015a) aimed to discriminate between four genotypes of *T. orientalis*, but is only semi-quantitative and therefore cannot be used to define a clinical threshold. Furthermore, this assay is designed as a series of uniplex assays rather than a genuine multiplex assay, meaning that reagent costs are substantial.

Hydrolysis probe qPCR assays employ sequence-specific, fluorescently-labelled probes attached to duplex-stabilising molecules and are frequently used to accurately identify and quantify genetic sequences within clinical samples (Kim et al., 2007; Maaroufi et al., 2003; Nwakanma et al., 2009; Plain et al., 2014). These assays can be multiplexed by using

multiple sequence-specific probes attached to fluorescent molecules with unique and distinguishable emission spectra (Veron et al., 2009). While one hydrolysis-probe qPCR assay has been previously developed for *T. orientalis* detection, that assay does not identify genotypes associated with clinical disease (Jeong et al., 2003). In this study, we describe the development and validation of a multiplex hydrolysis probe qPCR which can detect *T. orientalis* infection, quantify parasite load and identify the two genotypes associated with clinical disease.

1.2 Genotyping of *T. orientalis* strains

T. orientalis infection often occurs as a mixture of genotypes. Recent outbreaks of clinical theileriosis in Australasia have been linked to infection with the Ikeda genotype (Eamens et al., 2013c; Kamau et al., 2011a; Pulford et al., 2016). In one study, this genotype was found to be present in clinical cases as a sole or mixed infection (Eamens et al., 2013c), but most commonly co-occurred with the Chitose genotype. In contrast to the Ikeda genotype, the Chitose genotype was rarely found to be associated with disease when present as a sole infection (Eamens et al., 2013c); however other studies have suggested that the Chitose genotype may directly cause clinical disease (McFadden et al., 2011). While the Buffeli genotype is considered benign and does not cause disease when present as the sole infection, it can also occur in combination with the Ikeda and/or Chitose genotypes in clinical cases (Eamens et al., 2013c).

While the Ikeda genotype is the major genotype associated with pathogenesis, infection with this genotype is often subclinical. This raises questions as to whether there are particular subpopulations of this genotype that may be responsible for disease, or whether this variation in clinical picture is driven by host, vector or environmental factors.

In this part of the study, we examined the MPSP gene and ITS region sequences of *T. orientalis* samples from across Australia to determine whether subpopuations of the parasite are more associated with clinical disease. We also used quantitative PCR methods to monitor the temporal dynamics of the Ikeda, Chitose and Buffeli genotypes of *T. orientalis* in a group of naïve animals introduced to a *Theileria*-endemic area.

1.3 Genome sequencing of Australian strains of *T. orientalis* Ikeda, Chitose and Buffeli

The genetic diversity of Apicomplexan parasites allows for rapid adaptation to selective pressures, which has significant consequences for the development of drug resistance. Moreover, the highly diverse surfaceomes of these populations allow for the avoidance of specific immune responses, challenging vaccine design and limiting pathogen clearance. Diversity in these parasites can be measured between populations of parasites in individual hosts of the same or different species (Hayashida et al., 2013). Additionally, the study of within-host diversity is central to the understanding of therapeutic and immune system avoidance mechanisms of these highly variable parasites (Manske et al., 2012). The development of inexpensive whole genome sequencing technologies that allow for direct sequencing of clinically-derived samples promises to revolutionize the study of parasitic diversity and large scale monitoring of genetic variations in field samples can provide critical information for disease surveillance (Manske et al., 2012; Miotto et al., 2013). Furthermore,

the study of diverse apicomplexan surfaceomes has the potential to improve the design of subunit vaccines, which currently have limited effectiveness (Morrison, 2009; Vaughan and Kappe, 2012).

Reduced genome apicomplexan parasites of the genus *Theileria* use genetic diversity as a mechanism to overcome host immune responses and maintain persistent infections (Sivakumar et al., 2014). Recombination through sexual reproduction is identified as a major source of genetic diversity in Apicomplexan parasites (Deitsch et al., 2009) and has been demonstrated in both *Theileria parva* and *Theileria annulata* (Henson et al., 2012; Katzer et al., 2006; Weir et al., 2007; Weir et al., 2010). In *Theileria* species that infect cattle, sexual reproduction and therefore genetic recombination, occurs within the tick vector. *T. parva* in particular has been shown to have a high rate of recombination when compared to other Apicomplexan parasites (Katzer et al., 2011). Moreover, early studies of whole genome variation in *T. parva* have shown that these parasites are highly diverse (Hayashida et al., 2013). Genetic diversity in *Theileria* spp. has also been examined in immunogenic and important surface exposed molecules such as those expressed by the TaSP, Tams-1 and MPSP genes and also the SVSP and Tpr gene families (Eamens et al., 2013a; Gubbels et al., 2001; Schnittger et al., 2002; Weir et al., 2010). As control of *Theileria* is limited to a small number of anti-theilerial therapeutics which can leave undesirable residues in meat and milk, vaccination is seen as a highly desirable option, and analysis of genomic diversity, particularly loci encoding surface exposed molecules may provide insight into the development of effective subunit vaccines.

Theileria orientalis is a tick-borne parasite that has caused outbreaks of clinical theileriosis in production cattle of Japan, Korea, China, Australia and New Zealand. Examinations of diversity in *T. orientalis* have largely focused on genes encoding immunogenic piroplasm surface proteins. Sequence variability in MPSP has been used to classify the organism into 11 distinct genotypes (Sivakumar et al., 2014). Pathogenicity is associated with Type 2 (Ikeda) while other types, such as the widespread Types 1 and 3 (Chitose and Buffeli, respectively), have been linked with benign infections (Eamens et al., 2013b; Kakuda et al., 1998; Kamau et al., 2011a; Zakimi et al., 2006). One genome of the Japanese Shintoku strain (genotype Ikeda) has been previously sequenced revealing a 9 Mb, 4 chromosome nuclear genome structure as well as mitochondrial and apicoplast genomes (Hayashida et al., 2012), which is a common karyotype within the *Theileria* genus (Gardner et al., 2005; Kappmeyer et al., 2012; Pain et al., 2005). To date, no studies have focused on the whole genomic diversity in *T. orientalis*, nor have there been studies examining within-host diversity of *T. orientalis* populations. In this part of the study we examined the genomic diversity of *T. orientalis* in single-genotype populations of Ikeda, Chitose and Buffeli genotypes and have developed a genomics-based platform for clinical surveillance of *T. orientalis* infections.

1.4 Proteomic analysis of *T. orientalis* genotypes

As outlined in Section 1.3, treatment options for *T. orientalis* in Australia are limited, with the anti-theilerial agent, buparvaquone not currently approved for use due to the detection of residues in tissues for lengthy periods post-treatment (Bailey, 2013). Furthermore, given that drug resistance is a potential issue, as demonstrated for *T. annulata* (Mhadhbi et al., 2010), a vaccine for *T. orientalis* would be a preferred method of control for this parasite. Unlike *T. parva* and *T. annulata* which exert their major pathogenic effects via transformative

oncogenic pathways which induce lymphocyte proliferation (Dobbelaere and Heussler, 1999; Heussler et al., 1999; Marsolier et al., 2015), *T. orientalis* is a non-transforming species. Instead this species induces pathogenesis via the intraerythrocytic (piroplasm) phase. Studies show that the piroplasm phase of *T. orientalis* can be transmitted mechanically either via insects (Fujisaki et al., 1993) or by direct blood transfer such as on tools used in routine animal husbandry (Hammer et al., 2016), without the need to cycle through the diploid phase in the tick vector *H. longicornis*.

A greater understanding of the antigens expressed during the piroplasm phase of this organism's lifecycle is therefore desirable for future studies of pathogenesis, to investigate potential vaccine targets and to develop serological techniques for parasite detection.

This section of the study used global proteomic techniques to identify proteins expressed during the piroplasm phase of the *T. orientalis* lifecycle, including antigens that are exposed to the cell surface.

1.5 Development of a serological test for *T. orientalis*

While host immunological responses to the lymphocyte transforming *Theileria* species, *T. parva* and *T. annulata*, have been well-studied, little is known about the immune response to this non-transforming species. Naïve cattle introduced to areas where the disease is enzootic, as well as stressed, immunocompromised, pregnant or lactating animals are most at risk of developing clinical disease (Eamens et al., 2013b; Eamens et al., 2013c; Izzo et al., 2010), while cattle in *T. orientalis*-endemic areas appear to develop a degree of resistance to disease. It is unclear whether the host develops a humoral response against the parasite prior to the intraerythrocytic phase of the parasite's lifecycle, or whether the immune response is largely cell-mediated. Prior studies on *T. annulata* and *T. parva* suggest that responses against these parasites are largely cell-mediated (Machugh et al., 2008; MacHugh et al., 2009), however these organisms represent transforming theilerias which cause a tumour-like lymphocytic proliferation (Dobbelaere and Heussler, 1999) not observed in *T. orientalis* infection.

In *T. orientalis*, the MPSP is highly expressed during both the sporozoite (Sako et al., 1999) and piroplasm (Sugimoto et al., 1991) phases of the parasite's lifecycle and is believed to mediate entry into bovine erythrocytes via interactions with heparin-like compounds on the host cell surface (Takemae et al., 2014a). Immunoblots using sera from infected animals indicate that the MPSP is strongly recognised by host IgG and that immunisation of cattle with MPSP is at least partially protective against *T. orientalis* (Onuma et al., 1997). Nonetheless, immune response to this antigen has not yet been quantified in cattle using ELISA, although this method was found to be a sensitive means of detecting *T. orientalis* infection in water buffalo (Wang et al., 2010b).

In this part of the study, we developed a recombinant MPSP ELISA to measure bovine IgG response to this major surface antigen and correlate this response with clinical disease, parasite genotype and infection intensity.

2 Project objectives

Objective 1: Develop and validate an improved molecular test for *T. orientalis* surface protein types to facilitate improved surveillance of Australian cattle

Objective 2: Genotype strains of *T. orientalis* present in Australian cattle by sequence analysis of the surface protein (MPSP) genes as well as other molecular markers

Objective 3: Sequence the genomes of representative strains of the Ikeda, Chitose and Buffeli types

Objective 4: Conduct proteomic and ultrastructural analyses on purified piroplasms from *T. orientalis* (type Ikeda)

Objective 5: Develop type-specific serological tests

3 Methodology

3.1 Development and validation of a multiplex qPCR for detection of T. orientalis and differentiation of clinically relevant genotypes

3.1.1 Samples

Samples for conventional and quantitative PCR analyses were collected by district and private veterinarians. Assay comparison was performed on 318 blood samples that were separated into sensitivity and specificity panels (Table 1). Samples for sensitivity testing were derived from animals with a clinical history consistent with *Theileria orientalis* infection or from at-risk herds sampled as part of surveillance studies (Eamens et al., 2013a; Eamens et al., 2013b; Eamens et al., 2013c). Clinical signs considered to be consistent with theileriosis were a combination of some or all of the following: lethargy, ataxia, increased respiratory rate, fever, pale and/or jaundiced mucous membranes and abortion in pregnant animals. Haematological measures were also considered and comprised anaemia, as determined by packed cell volume (PCV), a blood film positive for *Theileria* piroplasms and for erythrocytic changes consistent with regenerative anaemia (e.g. nucleated erythrocytes, poikilocytosis, polychromasia and Howell-Jolly bodies). A diverse range of breeds and herds from disparate geographical locations were represented in the sensitivity sample panel (Table 1).

Samples from herds located in areas with no clinical history of theileriosis at the time of sampling and animals of known positive status for tick fever parasites (as determined by blood smear), but negative for theileriosis, were also used in this study for specificity testing (Table 1).

Initial evaluation of the various DNA extraction methods was carried out using a subset of 53 samples selected from the sensitivity and specificity panel (Table 2).

Table 1. Samples used for multiplex qPCR test validation

Sample type	Herd type	Sample origin by State	Case type	Total no. samples
Sensitivity panel	Beef (n = 126) Angus, Angus X, Belted Galloway, Brahman, Devon, Hereford, Hereford X, Limousin, Murray Grey, Santa X, Shorthorn Dairy (n = 66) Ayrshire, Holstein-Friesian, Swiss Brown Mixed breed (n = 3) Unknown (n = 42)	QLD (n = 42) NSW (n = 164) VIC (n = 19) SA (n = 2) WA (n = 10)	Clinical cases (n = 155) Animals from *Theileria* surveillance studies (n = 82)	237
Specificity panel	Beef (n = 38) Dairy (n = 12)	SA (n = 50)	Animals found negative by cPCR and no clinical history of theileriosis on property	50
Analytical specificity panel	Unknown (n = 31) *Babesia bigemina*-infected (n = 12) *Babesia bovis*-infected (n = 13) *Anaplasma centrale*-infected (n = 6)	QLD (n = 31)	Experimentally-infected or vaccinated animals	31

3.1.2 DNA extraction methods

3.1.2.1 DNeasy Blood and Tissue Extraction Kit (Qiagen)

This extraction method was utilised and validated in prior studies on the distribution and significance of T. orientalis in Australia (Eamens, 2011; Eamens et al., 2013a; Eamens et al., 2013b) and constituted the gold standard extraction method against which the other methods were evaluated. Extractions were carried out according to the manufacturer's instructions for blood samples and were eluted in 100 µL of molecular grade water.

Table 2. Sample subset used for initial evaluation of DNA extraction methods

Sample type	Description	No. of samples	No. of herds
Positive samples	Samples previously testing positive using the gold standard DNA extraction method in at least one MPSP-specific cPCR assay	n = 47	14
Negative samples	Samples sourced from negative cattle or imported one day prior from an area with no known history of theileriosis (South Australia) to the Mid-Coast LHPA	n = 6	2

3.1.2.2 Detergent-Proteinase K (DPK method)

The DPK method constituted an inexpensive DNA extraction method, based on detergent-proteinase K cell disruption. This method was previously designed as a hair root digest method used for genotyping of cattle (Healy et al., 1995) and was adapted here for use on EDTA blood. A similar method, employing saponin as a red blood cell lysis reagent, followed by treatment with a Tris buffer containing Tween 20 and Proteinase K, has been used successfully in the past as an inexpensive extraction method for *Theileria orientalis* detection. The DPK method employs a series of centrifugation steps for the initial erythrocyte removal, rather than saponin treatment.

3.1.2.3 MagMAX viral (MMV) protocol

This protocol is a high-throughput method used routinely at EMAI for the successful isolation of viral nucleic acid (DNA and RNA) from blood. The Ambion MagMAX^{TM}-96 viral RNA isolation kit (Life Technologies Cat. No. AM1836) was used in conjunction with a Kingfisher 96 magnetic particle processor (Thermo Fisher Scientific) according to the Kingfisher MagMAX-96 viral RNA isolation protocol (Life Technologies) where the following parameters were used in deep well (200 μL) plates: A lysis binding mixture consisting of 50 µL blood, 20 µL magnetic beads and 130 µL lysis /binding solution were incubated for 5 mins (Plate A). The beads were then washed in 2 × 150 µL Wash solution 1 for 2 mins (Plates B and C), followed by 2 × 150 µL washes in Wash solution 2 (Plates D and E). The beads were then air-dried and DNA was eluted in 50 µL Elution buffer for 4-5 min at 65°C (Plate F). At the end of the program, the extracts were also diluted 1:10 in molecular grade water (Sigma) in a new plate. The extracts (both neat and diluted) were stored at -20^{o}C.

3.1.2.4 MagMAX 96 DNA Multi-Sample (MMMS) protocol

This protocol is a high-throughput method that was chosen due to its flexibility for use on diverse pathogens within a range of sample types. The MMMS extractions were performed on a Biosprint® 96 magnetic particle processor (Qiagen) using the Ambion MagMAX^{TM}-96 DNA Multi Sample kit (Life Technologies Cat. no. 4413021) according to the manufacturer's

instructions. The following parameters were used in deep well (200 μL) plates: 50 μL Proteinase K buffer/enzyme mix (Proteinase K (100 mg/mL), and 50 μL blood sample were added to plate and sealed, then incubated for 20 mins at 60-65^{o}C. 200 μL Multi-sample DNA Lysis buffer was added, shaken for 3 mins at speed 8 (Titramax 100, Heidolph), then 20 μL DNA binding bead mix (DNA binding beads [10 mg/mL] in nuclease-free water) is added to each sample. The mixture is shaken for 3 mins at speed 7. Then 240 μL 100% ethanol is added, and the plate shaken for 3 mins at speed 7. Lysis and binding then proceeds for 5 mins (Plate 1), followed by washing in 150 μL Wash buffer 1 for 1 min (Plate 2), washing in Wash buffer 2 for 2 × 1 min (Plate 3 and 4), and then followed by air drying for 2 min. DNA was then eluted at 80^{o}C for 5 minutes in Elution buffer 1 (Plate 5). A second elution step was performed in Elution buffer 2 (75 μL) for 2 mins. At the end of the program, the extracts were transferred to 0.2 mL tubes and also diluted 1:10 in molecular grade water (Sigma). The extracts (both neat and diluted) were stored at -20^{o}C.

3.1.2.5 Qiagen One-For-All Vet (OFAV) protocol

This protocol is a high-throughput method that was chosen as a potentially useful procedure designed for a variety of clinical samples/pathogenic species. The Qiagen Biosprint® 96 One-For-All Vet kit (Qiagen Cat. no. 947057) was used in conjunction with a Biosprint® 96 magnetic particle processor (Qiagen) using the One-For-All Vet protocol (Qiagen) where the following parameters were used in deep well (200 μL) plates: The lysis/binding mixture consisted of 200 μL Buffer AL, 200 μL Isopropanol, 25 μL MagAttract Suspension G, 2.7 μL Carrier RNA (1 μg/μL), 40 μL Proteinase K and 200 μL blood, and was incubated for 4 mins (Plate 1). This was followed by washing of the magnetic beads in Wash solution 1 (700 μL) for 1 min, a second wash in Wash solution 1 (500 μL) for 1 min (Plates 2 and 3), and a third wash in Buffer RPE for 2 × 1 min (Plates 4 and 5). Following the wash steps, the beads were air dried for 1 min and elution was carried out in 200 μL of Elution buffer. At the end of the program, the extracts were transferred to 0.2 mL tubes and also diluted 1:10 in molecular grade water (Sigma). The extracts (both neat and diluted) were stored at -20^{o}C.

3.1.3 PCR methods

3.1.3.1 Conventional PCR (cPCR) gold standard assay

Conventional PCR assays (cPCRs) validated in a previous project (Eamens, 2011), constituted the gold standard against which the performance of the quantitative PCRs (qPCRs) were compared. The gold standard cPCR assay was performed using DNA extracted using the DNeasy method (Section 3.1.2.1). Four sets of PCR primers, previously described (Zakimi et al., 2006) (Table 3), were used for the cPCR assays, all of which targeted the major piroplasm surface protein (MPSP) gene. The cPCRs performed were the universal (p32) screening assay for *T. orientalis* and 3 further assays specific for the Ikeda, Chitose and Buffeli subtypes. All test reactions were performed in a total volume of 25 μL including 2 μL of undiluted or 1:10 diluted template DNA. Two positive control samples, a water control (containing MilliQ water instead of template) and a PCR cocktail control (no template added) were included in every run. Each cPCR reaction contained 0.5 μM forward and reverse primers, 200 μM deoxyribose nucleoside triphosphates (dNTPs), 2 mM $MgCl_2$, 1 × Taq reaction buffer and 1 U of Taq polymerase (Roche, Basel, Switzerland). cPCR

thermal cycling was carried out on an Eppendorf Mastercycler according to the parameters described in Zakimi et al. (2006). Amplicons were assessed on 1.5% agarose-Tris borate EDTA (TBE) gels run at 120 volts for 1.5 hr. All agarose gels were stained with Gel Red™ rather than ethidium bromide as this stain was found to result in superior sensitivity and represented a less toxic option. cPCR products were visualised on a Gel Doc 2000 system (Bio-Rad Laboratories, Hercules, California).

Table 3. Primers used in the *T. orientalis* cPCR assays

cPCR assay	Forward primer (5′-3′)	Reverse primer (5′-3′)	Amplicon size (bp)
P32	CACGCTATGTTGTCCAAGAG	TGTGAGACTCAATGCGCCTA	875 bp
Ikeda	AAGGATCCGTCTCTGCTACCGCCGC	TGTGAGACTCAATGCGCCTA	826 bp
Chitose	GCGGATCCTCATCGTCTCTGCAACT	TGTGAGACTCAATGCGCCTA	831 bp
Buffeli	GCGGATCCGCTCTGCAACCGCAGAG	TGTGAGACTCAATGCGCCTA	826 bp

3.1.3.2 Multiplex quantitative PCR (qPCR)

Dual-labelled TaqMan hydrolysis probes were used for the qPCR assay. A triplex qPCR (UIC triplex) was designed which consisted of a quantitative universal (U) assay for the detection all T. orientalis types, and two type-specific semi-quantitative assays aimed at differentiating the Ikeda (I) and Chitose (C) genotypes. The assay for the non-pathogenic Buffeli type was run in singleplex. These assay formats were chosen based on limitations in the number of fluorophores that can be detected in a single reaction without bleed through of fluorophores between channels (i.e.: 4 including the baseline dye) and limitations in the number of fluorescently-labelled minor groove binder (MGB) probes that are available for simultaneous multiplexing (i.e.: 3). The assay design was selected based on prior studies demonstrating that the Ikeda and occasionally the Chitose genotype, is associated with clinical disease in Australia, and the fact that the Buffeli type has not been reported to occur in association with clinical disease (Eamens et al., 2013c; Kamau et al., 2011b). As such, the Universal-Ikeda-Chitose (UIC) multiplex was selected as the most appropriate combination of assays for clinical diagnostic purposes.

A single primer set was used for qPCR amplification, with the hydrolysis probes conferring specificity in each channel of the assay (Figure 1). This assay design precluded interactions between primer sets and allowed for accurate quantification in the universal component of the assay. The forward primer was designed to target a highly conserved sequence within the 3´ end of the MPSP gene, while the reverse primer is similar to the TsR conventional PCR primer (Zakimi et al., 2006). Each probe was 5′ labelled with a fluorophore (FAM, VIC and NED for the U, I and C components, respectively) and 3′ labelled with a non-fluorescent

quencher (NFQ). All probes contained an MGB moiety to confer additional probe specificity and were purchased from Life Technologies (Carlsbad, CA, USA). All primer and probe sequences are listed in Table 4. An in silico analysis was performed on all primers and probes by comparison with existing sequence data in GenBank. The *T. orientalis* U probe was designed as a 100% match for >95% of all available T. orientalis MPSP gene sequences. Similarly, the I and C probes were designed to maximise detection of their respective subtypes. In the case of the Chitose probe, a degenerate design was later incorporated to account for a polymorphism present in some of the Chitose sequences (Table 4). Therefore, for the pilot study, the Chitose probe initially consisted of Chitose A probe only (Table 4) but later incorporated an equal mix of the Chitose A and Chitose B probes for the full scale validation.

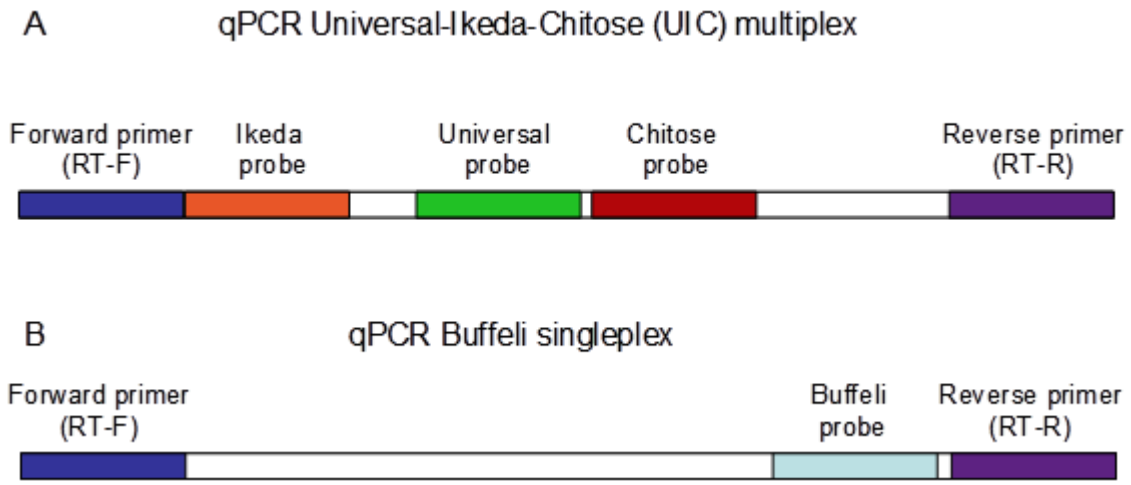


Fig. 1. qPCR assay design

The assay for the non-pathogenic Buffeli type was run in uniplex. The sequences of the primers and probes used for the Buffeli assay are listed in Table 4.

The qPCR assays were designed as hydrolysis probe (Taqman) assays with quantitation via a standard curve (see Section 3.1.3.3). All reactions were run in a 96-well plate format on an ABI 7500 Fast system (Applied Biosystems) at the standard ramp speed. Cycling parameters were as follows: Step 1: 95°C, 10 min (×1 cycle) → Step 2: 95°C, 15 sec; Step 3: 60°C, 1 min (×45 cycles).

The reagents used in each qPCR reaction are listed in Table 5.

3.1.3.3 DNA standards

Plasmid DNA standards were generated by PCR from blood samples sourced from New South Wales and Queensland known to contain the Ikeda, Chitose and Buffeli genotypes (Accession numbers KM624619-KM624621). Type specific forward primers and a universal reverse primer targeting the MPSP genes of each genotype (Zakimi et al., 2006) were modified to facilitate cloning into the pET100/GW/D-TOPO vector (Life Technologies, Table 4). MPSP gene PCR products were cloned, as described previously, into Escherichia coli

Table 4. Primer and probe designs used in the qPCR assays

Primer/Probe	Sequence (5′-3′)	Fluorophore	Quencher
Forward primer (RT-F)	GCAAACAAGGATTTGCACGC	N/A	N/A
Reverse primer (RT-R)	TGTGAGACTCAATGCGCCTAGA	N/A	N/A
Universal probe	TCGACAAGTTCTCACCAC	5′ FAM	MGB-NFQ
Ikeda probe	CATGAACAGTGCTTGGC	5′ VIC	MGB-NFQ
Chitose A probe	TCCTCAGCGCTGTTCT	5′ NED	MGB-NFQ
Chitose B probe	TCCTCGGCGCTGTTCT	5’ NED	MGB-NFQ
Buffeli probe	CTCCTTTGCAGTATTCTTCTATCTC	5′ FAM	BHQ-1
Buffeli forward cloning primer (Ts-Bc)*	CACC-TGCTCTGCAACCGCAGAG	N/A	N/A
Chitose forward cloning primer (Ts-Cc)*	CACC-TTCCTCATCGTCTCTGCA ACT	N/A	N/A
Ikeda forward cloning primer (Ts-Ic)*	CACC-ATCGTCTCTGCTACCGCC GC	N/A	N/A
T. orientalis reverse cloning primer (Ts-Rc)ŧ	*CTA*TGTGAGACTCAATGCGCCTA	N/A	N/A
Chitose plasmid mutagenesis primer (MutF)	CTATGTGAGACTCAATGCGCCTA	N/A	N/A
Chitose plasmid mutagenesis primer (MutR)	GATGAGAACAGCGCCGAGGACGGCA AGTG	N/A	N/A
T7F primer	CACTTGCCGTCCTCGGCGCTGTTCT CATC	N/A	N/A
T7R primer	TAATACGACTCACTATAGGG	N/A	N/A

*Underscored sequences indicate overhangs incorporated to facilitate cloning

ŧUnderscored and italicised region indicates an artificial stop codon included to terminate translation

TOP10 cells (Deutscher et al., 2010). To account for the Chitose polymorphism mentioned above, a plasmid standard variant (Chitose B) was generated by site-directed mutagenesis of the original Chitose plasmid (Chitose A), as described in (Fisher and Pei, 1997) using mutagenic primers (Table 4). Plasmids were extracted from transformant colonies with the QIAprep Spin Miniprep Kit (Qiagen). All plasmids were examined with conventional dye-terminator sequencing to confirm correct sequences using T7 forward and reverse primers (Table 4). Plasmid DNA was quantified by spectrophotometry at 260 and 280 nm. Plasmid standards were prepared as serial 10-fold dilutions in (1:250) tRNA sourced from Saccharomyces cerevisiae (Sigma-Aldrich). For the multiplex assay, an equimolar mix of Ikeda MPSP gene-containing plasmid and the Chitose MPSP-containing plasmid was used. For the Buffeli assay, a Buffeli MPSP gene-containing plasmid was used. DNA standards

consisted of 10-fold serial dilutions of each plasmid from 100 picograms/µL (pg/µL) down to 100 attograms/µL (ag/µL).

Table 5. Reagents used in qPCR cocktail

Reagent	Working Concentration or Volume
Mastermix (AB Environmental mix)	1 ×
Forward primer	300 nM
Reverse primer	300 nm
Probe: Universal-FAM	100 nM
Probe: Ikeda-VIC	250 nM
Probe: Chitose A-NED	100 nM
Probe: Chitose B-NED	150 nM
Probe: Buffeli-FAM	100 nM
DNA template	2 µL
Molecular grade water	Up to 20 µL

3.1.3.4 qPCR quality acceptance criteria

Acceptance criteria for each test assay were based on analysis of the standards. The coefficient of determination (R^2) was calculated from the regression line for the standards and indicated the goodness of fit between the C_T values for each individual standard, with a value of 1 indicating a perfect fit between data points. qPCR assays with an R^2 <0.98 were considered failed runs. The slope of the standard curve is a measure of the efficiency of the assay, with the ideal efficiency of 100% resulting from a doubling in the number of copies of amplicon at each thermal cycle. An amplification efficiency of between 90 and 110% is generally considered acceptable for qPCR assays without significantly skewing the quantification calculations (Bustin et al., 2009). As such, only UIC multiplex assays or Buffeli singleplex assays with efficiencies of 90-110% were considered valid runs.

3.1.3.5 Assay scoring

A system of gradation for the qPCR results based on plasmid standards and test samples was developed for initial comparison of the qPCR and cPCR results. Given that the cPCR results could be graded as negative, very weak positive, weak positive and positive based on the strength of the gel bands following electrophoresis, a comparison of cPCR and qPCR results provided a guide to the strength of the qPCR results. Table 6 outlines the gradings that were used to arbitrarily classify the qPCR results from the pilot studies for comparison with the cPCR data.

3.1.3.6 Determination of probe specificity

Probe specificity was tested using purified single and pairwise combinations of the different plasmid standards (Ikeda, Chitose A+B and Buffeli plasmids). Buffeli plasmid was included in the standards to ensure that the Universal component of the assay efficiently detected this prevalent (albeit benign) genotype to ensure accuracy in quantification. To test for equivalent quantification of all targets in the U component of the assay, single and pairwise combinations of 10-fold serial dilutions of the different plasmid standards ranging from 3×10^1 to 3×10^7 MPSP gene copies/µL (GC/µL) were tested.

Table 6. Arbitrary gradations of qPCR results to enable comparisons between pilot qPCR and cPCR findings in multiplex UIC and the Buffeli singleplex assay

Gradation	cPCR	qPCR (C_T)	qPCR (DNA concentration in fg/µL of extract)
Positive	Clear, strong gel band consistent with target size	$C_T \leq 32$	DNA ≥ 5
Weak positive	Clear but weaker gel band consistent with target size	C_T 32.1 - 34	DNA 2 - 4.99
Very weak positive	Faint gel band consistent with target size	C_T 34.1 - 37	DNA 0.1 - 1.99
Negative	No visible gel band consistent with target size	$C_T > 37$	DNA < 0.1

3.1.3.7 Determination of the assay limit of detection (LOD)

The limit of detection (LOD) was defined as the limit where 95% of qPCRs were successful (Bustin et al., 2009) and was experimentally estimated by testing 8 replicates of an equimolar mixture of Ikeda, Chitose (A+B) and Buffeli plasmid DNA at dilutions 15000, 1500, 150, 50, 15, 5, 1.5 and 0.5 MPSP GC/µL, against each probe in singleplex and multiplex, followed by Probit analysis.

3.1.3.8 Investigation of PCR inhibition

Potential PCR inhibition from blood components was tested during UIC assay development (using a cocktail containing 900 nM primers and 250 nM U probe) by testing both undiluted and 10-fold diluted DNA extracts from the samples in the specificity panel (n = 237). In addition, PCR inhibition in the final UIC assay was tested using a 10-fold dilution series of Ikeda, Chitose (A+B) and Buffeli plasmids (1.5×10^1 to 1.5×10^7 MPSP GC/µL) spiked into a pool of 10 *Theileria*-negative DNA extracts. The negative extracts were derived from EDTA blood samples from cattle located in areas where *Theileria* was not enzootic (ie: negative samples from the specificity panel) that were further confirmed *T. orientalis*-negative by both cPCR and qPCR.

3.1.3.9 Data analysis

Data were collected and analysed using Life Technologies 7500 Software v2.3. CT was determined by calculating the cycle number when fluorescence reached ΔRn = 0.1. The Universal component of the UIC assay was quantified by comparison to a standard curve generated from a 10-fold dilution series ranging from 1.5×10^1 to 1.5×10^7 to MPSP GC/µL of an equimolar mix of Ikeda, Chitose (50% Chitose A-50% Chitose B) and Buffeli plasmids. Quality acceptance parameters for each of the standard curves were an R^2 value > 0.98 and an amplification efficiency between 90 and 110%. Nucleic acid concentrations used for analytical and diagnostic statistics (MPSP GC/µL) were obtained using plasmid molecular mass calculated from plasmid sequence as previously described (Stothard, 2000).

Any discrepant results between the gold standard cPCR assay and the qPCR assay were resolved via Sanger sequencing at the Australian Genome Research Facility (AGRF) using the relevant sequencing primers detailed in Tables 3 and 4.

3.1.3.10 Statistical analyses

McNemar's test which is recommended for assessing differences in the ability of diagnostic tests to detect disease, was used to determine whether there was a significant difference in the proportion of positives between the two tests between pairs of treatments (cPCR vs qPCR). Confidence intervals for sensitivity and specificity values were calculated using the Clopper-Pearson method (Clopper and Pearson, 1934). The units GC/µL and gene copies/µL packed erythrocytes; (GC/µL PE) were used for correlation with parasitaemia and PCV respectively. GC/µL PE values were calculated to account for erythrocyte concentrations which are considerably skewed in anaemic animals [i.e. they were calculated from GC/µL by multiplying by nucleotide extraction volume (100 µL)/PCV]. Correlations were examined using Spearman's method; two-tailed P values are listed.

3.1.3.11 Determination of a clinical threshold

To determine whether there was a significantly higher concentration of parasite DNA in clinically-affected animals, samples were divided into four groups based on their individual case histories and herd status: 1) subclinical animals from herds without clinical cases; 2) clinically-affected individuals; 3) recovering animals and 4) in contact animals. Animals were considered subclinical (n = 55) if they were positive via cPCR but belonged to herds that did not display any cases consistent with clinical theileriosis at the time of sampling (e.g. surveillance samples). Animals were considered clinically affected (n = 110) if they displayed a combination of symptoms consistent with theileriosis at the time of sampling (anaemia, jaundice, depression, ataxia, lethargy tachypnoea, tachycardia, pyrexia or late-term abortion) and tested positive for *T. orientalis* in cPCR. Cases with potentially confounding diagnoses were excluded from this group. Recovering animals (n = 25) were those that had symptoms consistent with theileriosis in the preceding weeks but had recovered at the time of sampling and were positive in cPCR. Many of these animals also had evidence of regenerative anaemia on blood smear indicating prior illness. In contact animals (n = 24) were cohorts of clinically-affected animals at the time of sampling, but were not displaying obvious symptoms of disease themselves. MPSP concentrations (GC/µL) for each group were compared with one way analysis of variance (ANOVA) using a Kruskal-Wallis test and Dunn's multiple comparison post test within Prism 4.0 (GraphPad Software, La Jolla, CA, USA).

3.2 Validation of an inexpensive DNA extraction method for *Theileria* DNA

In this part of the study, an inexpensive DNA extraction method (DPK method) which had performed well in the pilot study was validated on a large number of clinical and surveillance samples.

3.2.1 Samples

The bovine blood samples analysed in this study (n = 434) were derived from a total of 92 separate herds across five states of Australia and comprised samples PCR positive and negative for *T. orientalis.* Blood samples were collected into either EDTA (n = 425) or lithium-heparin (n = 9) and either extracted without freezing (n = 70) or frozen and extracted using both methods at a later date (n = 364).

3.2.2 DNA extraction

DNA was extracted from 100 µL of each blood sample in parallel using both the gold standard DNeasy Blood and Tissue method (Section 3.1.2.1) and the DPK method (Section 3.1.2.2). The commercial DNA extraction method has been extensively used in prior PCR studies on *T. orientalis* (Eamens et al., 2013a; Eamens et al., 2013b; Eamens et al., 2013c; Islam et al., 2011; Kang et al., 2013; Perera et al., 2015a) and was considered the gold standard method.

3.2.3 Evaluation of DNA extracts in UIC qPCR

All DNA extracts were tested using multiplex qPCR as described in Section 3.1.3.2, except that only the Chitose A probe, rather than a mix of Chitose A/B probes, was included with the Ikeda and Universal probes in the PCR reaction. All probes in the reaction were used at a concentration of 250 nM. Because extracts from the DPK method are relatively crude, both 10-fold diluted and undiluted DPK extracts were tested to determine whether inhibitors from blood had an effect on qPCR detection. Gene copy data from the quantitative (Universal) component of the assay were doubled for the DPK extraction to account for the larger elution volume used with this method and so that all concentrations could be expressed as gene copies/µL blood.

3.3 Comparison of the performance of the UIC qPCR with a commercial MT-PCR (AusDiagnostics)

During the course of this study, a commercial semi-quantitative multiplex tandem PCR (MT-PCR) was independently developed (AusDiagnostics Pty., Ltd., Australia) (Perera et al., 2015a). A small comparison trial was conducted between the commercially available MT-PCR and the UIC and Buffeli singleplex qPCR assays.

3.3.1 Samples

Samples from a wide geographic area that gave variable results for Ikeda, Chitose and Buffeli genotypes in the UIC qPCR were selected for this comparison; however only 16 samples were tested due to limitations on the run size of the AusDiagnostics platform with a

trial kit. Samples were from NSW (n=2), VIC (n=6), SA (n=3), and WA (n=2). In addition, samples derived from experimentally infected calves (n=3) which had also been subjected to genome sequencing (Section 3.5) were tested.

3.3.2 DNA extraction

All samples were extracted using the gold standard DNeasy blood and tissue method as described in Section 3.1.2.1. This method is recommended for use with the AusDiagnostics assay, although the blood volume processed and the elution volume vary slightly according to that method. The same DNA extracts were used for both assays to enable direct comparison of amplicon quantities.

3.3.3 PCR testing

DNA extracts were tested in tandem in both the UIC multiplex plus Buffeli singleplex (B) assays and the MT-PCR assay. The UIC + B assays were performed as described in Section 3.1.3.2. The AusDiagnostics MT-PCR assay was performed as described in Perera et al. (Perera et al., 2015a) as follows:

MT-PCR assays were run using the Easy-Plex platform (AusDiagnostics). The Gene-Plex CAS1212 liquid handling robot (AusDiagnostics) was set up with reagents as per the manufacturer's instructions and 10µL of genomic DNA template for each sample was added into 0.2-ml PCR strips. The strips were placed into the thermocycling block within the Gene-Plex robotic platform. The primary amplification was performed (95°C for 10 s, 60°C for 20 s and 72°C for 20 s, 15 cycles) by the program Easy-Plex Assay Setup (AusDiagnostics). The Easy-Plex ring was then sealed and transferred to the Rotor-Gene 6000 real-time PCR thermocycler (Qiagen, Germany) for secondary amplification.

3.3.4 Analysis of PCR results

For the UIC qPCR assay the concentration of target DNA for each genotype was calculated from the plasmid standard curve and was converted to gene copies/µL blood as described in Section 3.1.3.9 For the MT-PCR, the relative intensity of infection for each genotype was estimated as the DNA copy number using the AusDiagnostics software, however, for the Ikeda genotype, the relative intensity of infection was estimated by dividing the DNA copy number by 2, as there are two gene copies of the target ITS-1 region within the Ikeda genome (Perera et al., 2015a). The DNA copy number per µL was calculated by dividing the DNA copy number by 10 to account for the 10 µL template added to each reaction tube.

3.3.5 Cloning and sequencing of the Buffeli P23 gene target from MT-PCR

Discrepancies between results from the Buffeli uniplex assays and the MT-PCR assays were resolved by cloning selected amplicons from the Buffeli reaction of the MT-PCR assay. PCR amplicons were purified using the Qiaquick PCR purification kit (Qiagen) according the manufacturer's instructions. Purified PCR products were cloned into the pCR2.1 TOPO vector (Life Technologies) using the TOPO TA cloning kit as outlined by the manufacturer. Briefly, 1 µL PCR product was mixed with 1 µL of pCR2.1 TOPO vector, 1 µL salt solution and 2 µL sterile water and incubated for 5 min at room temperature. 40 µL of chemically competent *E. coli* cells (TOP10 strain) were then transformed with 2 µL of vector by

incubating together on ice for 30 min, heat shocking in a 42°C waterbath for 30 sec. The transformants were grown at 37°C for 1 hr in supplied SOC medium and then spread-plated onto Luria Bertani (LB) agar supplemented with 100 µg/mL ampicillin. The agar plates were incubated overnight and positive transformants were sequenced using the supplied vector primers at the Australian Genomic Research Facility (AGRF).

3.4 Genotyping of *T. orientalis* strains

3.4.1 Samples

All samples were received at the Elizabeth Macarthur Agricultural Institute (EMAI) as whole blood in EDTA anticoagulant. Test samples used for genotyping were collected by district and private veterinarians from animals with a clinical history consistent with *T. orientalis* infection or from herds which formed part of a prior study on the distribution and significance of MPSP types (Eamens, 2011; Eamens et al., 2013a; Eamens et al., 2013b). *Haemaphysalis longicornis* tick samples were collected from South-Eastern Victoria (Hammer et al., 2015). Samples used for MPSP genotyping and ITS-1 genotyping are shown in Tables 7 and 8 respectively.

Additional samples (n=90) were collected from a single herd of 10 naïve animals (2 year old Ayrshire heifers) that had been introduced to a property on the mid-coast of NSW with a history of clinical theileriosis cases. These samples were used for a temporal study examining genotype dynamics in *T. orientalis*-infected cattle. EDTA blood was collected from each animal immediately upon introduction to the affected property, and approximately weekly thereafter for a period of 76 days.

Table 7. Samples used for MPSP phylogeny

Sample ID	Location	Clinical signs
	Ikeda samples	
CM10/205	QLD	No
CM10/211	NSW	No
CM10/130	VIC	No
CM11/044	Kempsey, NSW	No
CM11/045	Kempsey, NSW	No
CM11/053	Monteagle, NSW	No
CM11/055	Lavington, NSW	Borderline
CM11/063	Wagga Wagga, NSW	Yes
CM11/067	Wagga Wagga, NSW	No
CM13/022	Southwest WA	Yes
CM14/054	SA, Australia	Yes
CM14/042	France, New Caledonia	Yes
CM14/003	QLD	Yes
CM14/001	Bega, NSW	Yes
	Chitose samples	
CM10/131	VIC	No
CM10/132	VIC	No
CM10/202	QLD	No
CM10/136	VIC	No
CM10/205	Mareeba, QLD	No
CM11/011	Kempsey, NSW	No
CM13/019	QLD	No
CM12/086	Barinsdale, VIC	Yes (mixed with Ikeda)
CM13/006	Lismore, NSW	No
CM10/088	Armidale, NSW	No
CM11/078	Camden, NSW	No
CM11/091	NSW	Yes (mixed with Ikeda)
CM11/028	NSW	No
CM11/033	NSW	No
CM14/054	SA	Yes
CM14/009	WA	No
	Chitose samples used in subtree	
Chitose 1	VIC	No
Chitose 2	VIC	No
Chitose 5	QLD	No
Chitose 6	VIC	No
Chitose A1	Lismore, NSW	Yes
Chitose A2	Lismore, NSW	Yes
Chitose A9	Bowral, NSW	Yes
Chitose A10	Camden, NSW	Yes
Chitose A11	Gloucester, NSW	Yes

Chitose A14	Bomaderry, NSW	Yes
Chitose A15	Willala, NSW	Yes
Chitose B1	QLD	No
Chitose B2	Kempsey, NSW	Yes
Chitose B3	QLD	No
Chitose B4	QLD	No
Chitose B5	QLD	No
Chitose B6	QLD	No
Chitose B7	Bairnsdale, VIC	Yes
Chitose B8	Lismore, NSW	Yes
Chitose B11	Armidale, NSW	Yes
Chitose B12	Armidale, NSW	Yes
Chitose AB8	Forbes, NSW	Yes
	Buffeli samples	
CM11/071	QLD	No
CM10/128	Wallangra, NSW	No
CM10/147	NSW	No
CM10/135	VIC	No
CM10/134	VIC	No
CM11/075	QLD	No
CM12/077	Singleton, NSW	No
CM12/084	Oakey, QLD	Yes
CM13/070	Lismore, NSW	No
CM11/139	Wingham, NSW	Yes

Table 8. Samples used for Ikeda ITS phylogeny

Sample ID	State/Country
13/051	NSW
11/042	NSW
11/046	NSW
11/123	VIC
14/042	New Caledonia
14/054	SA
13/031	WA
15/067	QLD
13/022	WA
14/003	QLD
12/083	QLD
15/062	NSW
14/012	VIC
14/054	SA
15/056	VIC

3.4.2 DNA extraction

DNA extraction was performed on each of the blood samples using either the DNeasy Blood and Tissue method (Qiagen; Section 3.1.2.1), or the detergent-proteinase K method (Section 3.1.2.2). The DNeasy Blood and Tissue method was used to extract DNA from tick samples. Ticks were homogenised in lysis buffer prior to extraction and the protocol recommended for tissue samples by the manufacturer was employed.

3.4.3 Amplification of the MPSP genes for sequence analysis

The MPSP genes of the test samples were amplified using conventional PCR to maximise amplicon length for downstream sequencing purposes. Amplification was carried out using the PCR primers described in Zakimi et al. (2006) (Table 3). All test reactions were performed in a total volume of 25 µL including 2 µL of undiluted or 1:10 diluted template DNA. Two positive control samples, a water control (containing MilliQ water instead of template) and a PCR cocktail control (no template added) were included in every run. Each cPCR reaction contained 0.5 µM forward and reverse primers, 200 µM dNTPs, 2 mM $MgCl_2$, 1 × Taq reaction buffer and 1 U of Taq polymerase (Roche, Basel, Switzerland). cPCR thermal cycling was carried out on an Eppendorf Mastercycler according to the parameters described in Zakimi et al. (2006). Amplicons were then assessed on 120 mL 1.5% agarose-Tris borate EDTA (TBE) gels run at 120 volts for 1.5 hr. Gels were stained with 5 µL of Gel Red dye. cPCR products were visualised on a Gel Doc 2000 system (Bio-Rad Laboratories, Hercules, California).

3.4.4 Amplification of the internal transcribed spacer (ITS) and 5.8S rDNA regions

The ITS regions ITS1 and ITS2 are located in the intergenic regions between the small subunit (SSU) rRNA and the large subunit (LSU) rRNA and are separated by the 5.8S rRNA gene (Fig. 2). These regions are commonly used for differentiating closely related strains due to the fact that intergenic spacers lack structural constraints. The ITS regions have been used previously to investigate polymorphism in *T. orientalis* genotypes from Japan and the USA (Aktas et al. 2007). These regions were investigated as alternative molecular markers for genotyping the Australian *T. orientalis* strains. Where samples contained a single *T. orientalis* subtype, the ITS1-5.8S-ITS2 regions were initially amplified via a hemi-nested PCR, using the primers and PCR conditions specified in Aktas et al. (2007). These primers were found to be relatively non-specific and therefore later studies employed the LSUR50 and ITSF primer pair from Aktas et al. (2007) in combination with newly designed primers targeting the 5.8S rRNA gene (Fig. 2, Table 9) which were more specific for the *T. orientalis* ITS regions. ITSF and 5.8S R were used for the amplification of ITS1 region and 5.8s F and LSUR50 were used for the amplification of ITS2.

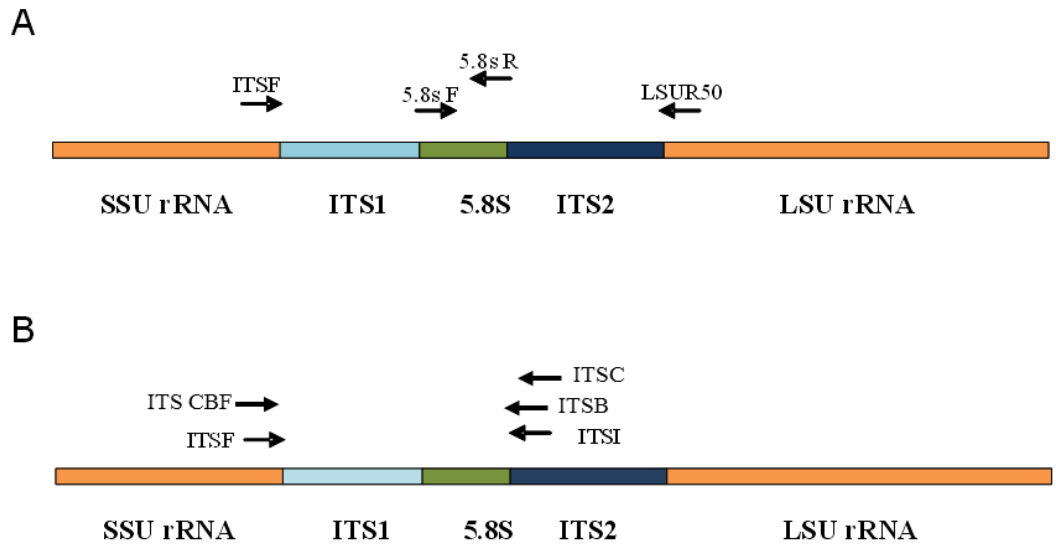


Figure 2. Structure of the ribosomal RNA operon and internal transcribed spacers.

For amplification of the ITS1-5.8S regions from mixed infections, PCRs were performed using type-specific primers designed to target the 3' end of the ITS1 region and the 5.8S rRNA gene (Table 9). The primer pairs designed to amplify each subtype were as follows: Ikeda type, ITSF and ITSI; Buffeli type, ITSF and ITSB; Chitose ITSCBF and ITSC. Each PCR reaction contained 0.5 µM each of forward and reverse primer, 200 µM dNTPs, 2 mM MgCl2, 1 × Taq reaction buffer and 1 U of Taq polymerase (Roche). Thermal cycling was carried out on an Eppendorf Mastercycler using the following cycling conditions for amplifying Buffeli and Ikeda targets: 96°C, 2 m × 1 cycle; 94°C 30 s, 60°C 30, 72°C 2m × 35 cycles; 72°C 10 m × 1 cycle. Amplicons were visualised as described in Section 3.4.3 for the MPSP genes.

Table 9. Primers used for amplification and sequencing of the ITS1-5.8S-ITS2 regions.

Primer Name	Primer Sequence (5'-3')	Reference
ITSF	GAGAAGTCGTAACAAGGTTTCCG	(Aktas et al., 2007)
LSUR50	GCTTCACTCGCCGTTACTAGG	(Aktas et al., 2007)
5.8s F	ACACTTTTAGCGGTGGATGTCTT	This study
5.8s R	AGGTTCACTGAAATGGGAGTACC	This study
ITSI	AGCCACCGGACTACGCAA	This study
ITS CBF	GGTTTCCGTAGGTGAACCTGC	This study
ITSC	GCCATTGCTCGCAACTAACAG	This study
ITSB	AASCTACCTCGCACTGCAACG	This study

3.4.5 Cloning of Ikeda genotype ITS1 regions

Cloning with the TOPO TA cloning kit involved setting up a ligation reaction containing 2 µL of PCR product, 0.5 µL of Invitrogen salt solution (1.2 M NaCl, 0.06 M $MgCl_2$) and 5 ng of pCR®2.1-TOPO® plasmid DNA (Invitrogen). The ligation reaction was incubated at room temperature for 30 min prior to transformation. Cloning was also performed with the TOPO TA cloning kit with a molar ratio of 1:3 of pCR®2.1-TOPO® plasmid DNA (Invitrogen) to insert (PCR product). For this ligation reaction, 1 x T4 DNA ligase reaction buffer, 25 ng of pCR®2.1-TOPO® plasmid and 2.5 U of ExpressLink T4 DNA Ligase was added. The ligation reaction was incubated at room temperature for 1 hr prior to transformation.

3.4.6 Sequencing of amplicons

All PCR products were purified using the Qiaquick PCR purification system (Qiagen). 30 ng of each sample was submitted along with 10 pmol sequencing primer to the Australian Genomic Sequencing Facility (AGRF) for Big Dye Terminator sequencing and capillary separation.

3.4.7 Sequence analysis

MPSP gene and ITS1-5.8S sequence contigs were assembled using the sequence editing and alignment software Geneious. Sequence alignments were assembled using the MUSCLE alignment feature in MEGA 6.0 and included representative sequences from GenBank. Sequence identities were calculated in GeneDoc. The evolutionary history was inferred using a distance method using the MEGA 6.0. The evolutionary distances were computed using the Maximum Composite Likelihood method (Tamura et al., 2004) and the Neighbor-Joining algorithm (Saitou and Nei, 1987) was used for tree construction. The ITS1 sequences were analysed separately by MPSP type due to the low levels of identity between types. For the ITS1 phylogenies, alternate MPSP types were used as outgroups. The *Theileria annulata* MPSP was used as an outgroup for the MPSP phylogeny.

3.4.8 Uniplex qPCRs for the quantification of genotypes

Because the UIC multiplex and B uniplex assays were considered quantitative for the universal assay but semi-quantitative for determining the load of individual genotypes, uniplex quantitative assays were designed such that the loads of individual genotypes could be more accurately monitored in individual animals over time. Dual-labelled TaqMan hydrolysis probes were used for all genotype-specific qPCR assays and for a Chitose subpopulation analysis. Three specific qPCR assays were designed to target the genotypes Ikeda, Chitose and Buffeli. A fourth assay was designed to determine the dominant phylogenetic subpopulation (Chitose A or Chitose B) of the Chitose genotype. Each assay consisted of a probe and a forward and reverse primer set that were all genotype-specific. Each probe was 5′ labelled with a fluorophore (Table 10) and 3′ labelled with a non-fluorescent quencher (NFQ). In the case of the Chitose-specific qPCR, a degenerate design was used to account for a polymorphism between Chitose A and Chitose B subpopulations (Table 10). All probes contained an MGB moiety to confer additional probe specificity and were purchased from Life Technologies (Carlsbad, CA, USA). All primer and probe sequences are listed in Table 10. An *in silico* analysis was performed on all primers and

probes by comparison with existing sequence data in GenBank. Each probe and primer set was designed to maximise detection of their respective subtypes.

Table 10. Primer and probe sequences used for the uniplex genotype-specific qPCRs

Oligo	Sequence (5′-3′)	Target	Working concentration
Ikeda quantitative assay			
MPSP-I-F	ATTGGTAGACGGAAAATGGAAGAAGG	Ikeda MPSP gene	900 nM
MPSP-I-R	GAGACTCAATGCGCCTAGAGATAATAGA	Ikeda MPSP gene	900 nM
Pr-I	VIC-CATGAACAGTGCTTGGC-MGB-NFQ	Ikeda MPSP gene	250 nM
Chitose quantitative assay			
MPSP-C-F	CCGGTGATGAGAGATTCAAGGAAGTA	Chitose MPSP gene	900 nM
MPSP-C-R	GACTCAATGCGCCTAGARATAGTAGG	Chitose MPSP gene	900 nM
Pr-Ca	NED-TCCTCAGCGCTGTTCT-MGB-NFQ	Chitose A MPSP gene	100 nM
Pr-Cb	NED-TCCTCGGCGCTGTTCT-MGB-NFQ	Chitose B MPSP gene	150 nM
Buffeli quantitative assay			
MPSP-B-F	AAGTATACGTAGGTACCGATGATAAGAAAGTA	Buffeli MPSP gene	900 nM
MPSP-B-R	GAGACTCAATGCGCCTAGAGATAGA	Buffeli MPSP gene	900 nM
Pr-B	FAM-AGCGCTTTCCTCATCG-MGB-NFQ	Buffeli MPSP gene	250 nM
Chitose subpopulation assay			
MPSP-C-F	TCGACAAGTTCTCACCAC	Chitose MPSP gene	900 nM
MPSP-C-R	CATGAACAGTGCTTGGC	Chitose MPSP gene	900 nM
Pr-Ca (VIC)	VIC-TCCTCAGCGCTGTTCT-MGB-NFQ	Chitose A MPSP gene	150 nM
Pr-Cb	NED-TCCTCGGCGCTGTTCT-MGB-NFQ	Chitose B MPSP gene	150 nM

3.4.9 Quantitative PCR (qPCR) and generation of plasmid standards

Plasmid DNA standards, reagents, equipment, consumables and thermal cycling parameters for qPCRs performed in this study were as described previously Section 3.1.3.2 with primers and probes for the Ikeda, Chitose and Buffeli subtype assays as listed in Table 10.

The limit of detection (LOD) of each of these assays was defined as the limit where 95% of qPCRs were successful (Bustin et al., 2009). This was determined experimentally by testing 8 replicates of an equimolar mixture of Ikeda, Chitose (A+B) and Buffeli plasmid DNA at dilutions 1500, 150, 100, 50, 15, 5, 1.5 and 0.5 MPSP GC/µL, followed by Probit analysis.

The assay used for determining subpopulation ratios of the Chitose genotype was performed using the same method as the Chitose genotyping qPCR with the exception that the Chitose A and Chitose B probes were present in different proportions and labelled with different fluorophores (Table 10).

3.4.10 qPCR specificity

Probe specificity was tested by performing genotype-specific qPCRs with pairwise combinations of each primer set, probe and plasmid (i.e. For Ikeda primers, probe/plasmid combinations Ikeda/Ikeda, Ikeda/Chitose, Ikeda/Buffeli, Chitose/Ikeda, Chitose/Chitose, Chitose/Buffeli, Buffeli/Ikeda, Buffeli/Chitose and Buffeli/Buffeli were tested). Specificity was also tested using 10 samples derived from *T. orientalis*-negative animals, or *T. orientalis*-negative animals known to be positive for *B. bovis*, *B. bigemina* or *A. centrale*. The performance of each assay in detecting its target genotype within clinical samples was also tested using 15 well-characterised (sequenced and/or tested with alternate assays) samples known to be positive for one or more of the Ikeda, Chitose and Buffeli genotypes (Eamens et al., 2013b; Eamens et al., 2013c).

3.4.11 Data Analysis

Analysis of qPCR data was performed as previously described in Section 3.1.3.9. Genotyping of selected samples were confirmed via Sanger sequencing at the Australian Genome Research Facility (AGRF). For the Chitose subpopulation assay, Chitose A/B ratios were determined by comparing to 7 duplicate ratio standards of 1.00:0.00, 0.95:0.05, 0.75:0.25, 0.50:0.50, 0.25:0.75, 0.05:0.95 and 0.00:1.00 Chitose A:Chitose B plasmid at 100 fg/µL. Subpopulation ratio determination was performed by analysis of real-time PCR curve inflexion points as described in (Chen et al., 2014). To obtain these, raw fluorescence outputs were normalized to a passive reference dye (ROX), baseline-subtracted and corrected for cross-talk using pure fluorophore calibration data. Calculated ΔRn values were then fitted to a non-linear sigmoid function using Prism 4.0 (GraphPad Software, La Jolla, CA, USA) and transformed fluorescence ratios (k') were calculated from inflexion points. Unknown Chitose A:Chitose B ratios were calculated from a standard curve of transformed fluorescence ratio (k') vs Chitose A allele frequency fitted to a non-linear function (Chen et al., 2014). A conservative upper allele frequency limit of 95% was used to define 'pure' (>95%) Chitose A or B alleles from mixed populations. Chitose A/B ratios from this analysis were used to sort Chitose positives into three groups, >95% Chitose A, mix Chitose A/B and >95% Chitose B. Correlations of Chitose A/B ratio with clinical disease and Ikeda prevalence were determined using Pearson's chi-squared test.

3.5 Genome sequencing of *T. orientalis* Ikeda, Chitose and Buffeli genotypes

3.5.1 Collection of *T. orientalis* strains

Samples were sourced from cattle testing PCR positive for a single *T. orientalis* MPSP genotype (Ikeda, Chitose B or Buffeli). The Ikeda, Chitose B and Buffeli isolates were derived from cattle in Camden, NSW, Fish Creek in Southern Victoria and East Gippsland, Victoria respectively. Confirmation that each isolate was of a single MPSP genotype was

performed with two separate PCR assays, the UIC and B assays as well as the cPCR assays described in Section 3.1.3. Two to three days following the initial bleed and subsequent PCR confirmation that the strains were of a single MPSP genotype, approximately 80 mL of blood was collected into anticoagulant (heparin or citrate) and shipped cold to the Tick Fever Centre (Wacol, QLD) laboratory cold for propagation and extraction. On arrival samples were mixed with cryopreservative and stored as stabilates at -80°C until required.

3.5.2 Calf inoculations

Individual splenectomised calves were inoculated with Ikeda, Chitose or Buffeli stabilate at the Tick Fever Centre. Blood samples were drawn from each calf at regular intervals to monitor the infection and Giemsa-stained blood smears were used to estimate the level of parasitaemia. The packed cell volume (PCV) was also monitored. The calf inoculated with *T. orientalis* Ikeda was bled twice to obtain additional material for protein analysis.

3.5.3 Enrichment of *T. orientalis* piroplasms from bovine erythrocytes

When the parasitaemia had reached an appropriate level (6-20%), approximately 3.5L of calf blood was collected into anticoagulant (CPDA1 or heparin). *T. orientalis* piroplasms were then enriched using a method adapted from Shimzu et al. (1988). The blood was transferred into 300-500mL blood bags using a dialysis pump. The blood was then passed through Terumo leukocyte filters (300 mL blood per filter) under gravity. The leukocyte-depleted blood was then transferred to centrifuge tubes, centrifuged at 2500 × *g* for 20 min and the serum and any remaining buffy coat removed. The erythrocytes were then washed with an equal volume of Dulbecco's phosphate-buffered saline (D-PBS) followed by centrifugation as described above. The erythrocytes were washed a further 3 × with D-PBS to remove excess serum. The washed erythrocytes were diluted 1:5 with D-PBS and loaded into a cell disruption vessel ("nitrogen bomb"). The vessel was infused with nitrogen gas to a pressure of 1000 psi for 1 minute and then the pressure was released. The lysed erythrocytes were collected into a clean vessel and then transferred to centrifuge tubes. The cell lysate was centrifuged at 670 × *g* for 10 min to pellet the red blood cell debris. The supernatant was then harvested and centrifuged at 2700 *x g* for 10 min to pellet the piroplasms. The piroplasms were resuspended in D-PBS and a smear prepared from the suspension, which was stained with Giemsa stain to check purity. The piroplasms were then aliquoted into microfuge tubes and centrifuged at 2700 *x g* for 10 min. The supernatant was removed and the piroplasm pellets stored at -80°C. Piroplasm preparations were transferred to EMAI on dry ice and maintained at -80°C thereafter.

3.5.4 DNA extraction

Piroplasm pellets were resuspended in 200 µL of PBS and DNA was extracted using the DNeasy Blood and Tissue kit (Qiagen, Section 3.1.2.1). This method had been shown to yield high quality sequence on prior genome sequencing (Illumina) runs. DNA concentration and quality was checked using a Qubit fluorometer.

3.5.5 Genome sequencing

Extracts were subjected to Illumina sequencing in triplicate. Isolated genomic DNA was fragmented and tagged using the NextEra system (Illumina). Appropriately sized fragments were then selected via magnetic bead size-selection. Sequencing was performed with either a MiSeq Desktop Sequencer (Illumina) using a 250 cycle kit or a HiSeq Sequencer (Illumina) using a 150 cycle kit.

3.5.6 Genome assembly

Sequence reads were assembled into contigs using a high performance computer cluster (HPC). An open source assembly pipeline, A5 MiSeq (Coil et al., 2015; Tritt et al., 2012) designed for generating assemblies from DNA sequence data generated by the Illumina platform was used for compiling the draft genomes. Genome assembly was achieved with raw reads using v2015-05-22 of the A5 pipeline (Coil et al., 2015). Levels of remaining host DNA wereas assessed using BioBloomTools (Chu et al., 2014) and the *Bos taurus* genome sequence (Zimin et al., 2009).

3.5.7 Variant calling and dN/dS analysis

For variant calling, reads were first trimmed with Trimmomatic v0.33 based on leading and trailing base quality (Q < 20) and a 4-base sliding window when average quality scores were less than 20 (Bolger et al., 2014). Read quality was assessed pre- and post-trimming with FastQC (http://www.bioinformatics.bbsrc.ac.uk/projects/fastqc/) and sequence files were manipulated using ngs-utils v0.5.7 (Breese and Liu, 2013). For between population variants, trimmed reads were mapped to the *Theileria orientalis* Shintoku reference sequence (Hayashida, et al. 2012) with NextGenMap v0.4.12 (Sedlazeck et al., 2013). Low quality mapped reads (Q < 10) were filtered using SAMtools v1.2 (Li et al., 2009) and alignments were subsequently sorted, and duplicates removed using Picard tools v1.138 (http://broadinstitute.github.io/picard/). Variant calling was performed using VarScan v2.3.8 (Koboldt et al., 2009; Koboldt et al., 2012) on Mpileup files generated by SAMtools and using a minimum coverage of 14 ×, minimum average quality of 20, minimum variant supporting reads of 4, minimum variant frequency of 0.01 and minimum p-value threshold of 0.01. A minimum frequency of 0.9 was used to define homozygous variants. For within-population variants, potential copy number variants were excluded by examining with CNVnator v0.3.2 with a bin size of 250 bp (Abyzov et al., 2011). Hard and soft clipped reads were additionally filtered from within-population variant read-alignments files with awk and small contigs (< 1000 bp) were excluded to limit potential artifacts from assembly. Variant quality was assessed by examining subsets of variant sequences with Sanger sequencing.

For dN/dS analysis, SNP-containing genomes were constructed using homogeneous SNPs from Chitose and Buffeli variant analysis with the FastaAlternateReferenceMaker tool from the Genome Analysis Toolkit v3.2.2 (McKenna et al., 2010). Coding sequence alignments of Reference and SNP-containing genomes were generated for gene regions of 14 × or greater sequencing coverage using customised Python scripts and Biopython (Cock et al., 2009). Coding sequences were excluded where sequencing coverage at > 14 × made up less than 50% of the CDS. dN/dS ratios were calculated for each coding sequence using KaKs Calculator v2.0 (Wang et al., 2010a). Predicted surface proteins were identified using

SignalP 4.0 (Petersen et al., 2011) and TMHMM (**T**rans**m**embrane **H**idden **M**arkov **M**odel) (Krogh et al., 2001).

3.5.8 Piroplasmida phylogenomic analysis

Piroplasmida phylogeny was examined using 10 whole genome sequences from species representing the Piroplasmida order and one outgroup species *Plasmodium falciparum* strain 3D7 (Brayton et al., 2007; Cornillot et al., 2012; Gardner et al., 2005; Gardner et al., 2002; Hayashida et al., 2012; Jackson et al., 2014; Kappmeyer et al., 2012; Pain et al., 2005). The 10 genomes include the three *Theileria orientalis* isolates examined in this study and previously published and annotated whole genome sequences of *Theileria orientalis*¸ *Theileria parva*, *Theileria annulata*, *Theileria* (formerly *Babesia*) *equi*, *Babesia bovis*, *Babesia bigemina* and *Babesia microti*. Annotations for Ikeda, Chitose and Buffeli isolates were transferred from the *Theileria orientalis* Shintoku sequence to assemblies using RATT (Otto et al., 2011). A total of 261 orthologous groups with single copy genes found in all 11 genomes were identified using Orthofinder v0.3.0 (Emms and Kelly, 2015) with initial search results filtered to retain only groups with a BLAST e-value < 1e-30. Sequences from each of the 261 orthologous groups were aligned, trimmed and phylogeny inferred using MUSCLE v3.8.31, trimAl v1.3 and RAxML v8.2.4 (Capella-Gutierrez et al., 2009; Edgar, 2004; Stamatakis, 2006) as described previously (Cornillot et al., 2012) with the following alterations. Sequence alignments were trimmed using the nogaps automated option and two orthologous groups were removed from the final analysis after trimming due to poor alignment and high trimming. A supermatrix al ignment (63864 residues) was generated by concatenating individual gene alignments using a customised Python script. The supermatrix tree was inferred with RAxML as previously described (Cornillot et al., 2012). Average nucleotide identity comparisons were calculated using JSpecies v1.2.1 (Richter and Rossello-Mora, 2009).

3.6 Proteomic analysis of the piroplasm phase of *T. orientalis*

3.6.1 Piroplasm purification

Piroplasms were prepared from the Ikeda, Chitose and Buffeli genotypes of *T. orientalis* using blood collected from experimentally infected splenectomised calves as described in Sections 3.5.1-3.5.3 and stored at -80°C until required.

3.6.2 Protein purification

Proteins were purified from frozen piroplasm preparations using two different methods of fractionation aimed at selecting for aqueous and detergent phase proteins.

3.6.2.1 Mem-PER protein extraction method

Proteins were purified from *T. orientalis* Ikeda, Chitose and Buffeli piroplasms using the Mem-PER protein extraction kit (ThermoFisher Scientific, Waltham, MA, USA). The *T. orientalis* proteins were separated into aqueous and detergent-soluble fractions and each fraction was passed through a 300 kilodalton filter (Pall Life Sciences, Port Washington, NY,

USA) followed by a desalting column (Micro BioSpin P6 gel columns, BioRad Laboratories, Hercules, CA, USA) according to the manufacturer's instructions.

3.6.2.2 Triton X-114 protein fractionation

Approximately 0.1g of pelleted piroplasms were thawed on ice and resuspended in 1mL cold 1% triton buffer (1% TX-114, 10mM Tris pH8.0, 150mM NaCl, 1mM EDTA) by pipetting. Samples were extracted overnight on a rotary shaker at 4°C. The following morning pellet the samples were centrifuged at 4°C for 15 min at 13 000 rpm. The supernatant was removed to a new tube and incubated at 37°C for 10 min. The remaining pellet was retained and represents the TX-114 insoluble pellet. Phase separation of the supernatant was induced by centrifuging at 10 000 rpm for 5 minutes at room temperature. The top phase represented the aqueous phase and this was removed to a fresh tube. 20µL of concentrated TX114 was added to the aqueous phase (2% TX-114) and 1ml of 1% Triton buffer was added to the detergent phase (2% TX-114) and the samples were re-extracted at 4°C for 4 h on a rotary shaker. Phase separation was repeated by incubating at 37°C for 10 minutes then centrifuging at 10 000 rpm for 5 min at room temperature. The aqueous (top) and detergent (bottom) phases were removed to separate 50ml centrifuge tubes and the samples mixed with 10 volumes of ice cold acetone. Proteins were precipitated overnight at -20°C. The precipitated proteins were pelleted by centrifuging at 13 000 rpm for 30 min at 4°C. The acetone was poured off and the pellet air dried for 30 min. The aqueous pellet was resuspended in 0.5ml of SSS buffer (8M Urea, 100mM DTT, 4% chaps, 0.8% ampholytes, 40mM Tris). The detergent pellet was resuspended in 0.25ml MSS buffer (5M urea, 65mM DTT, 2% chaps, 0.8% ampholytes, 40mM tris, 2M thiourea, 2% sulfobetaine). Protein pellets were allowed to resuspend at room temperature for 30 min and then alternately sonicated (30 s) and vortexed (15 s) 4 times. Samples were centrifuged for 15 minutes at maximum speed at room temperature and the supernatants stored in fresh tubes at -20°C.

3.6.3 Protein electrophoresis

Aliquots of each fraction were electrophoresed on Criterion 12% TGX precast gels (BioRad Laboratories) at 200 V for 1 hr. Gels were washed 3 × 10 min with MilliQ water, stained with GelCode Blue stain reagent (ThermoFisher Scientific) and destained with MilliQ water.

3.6.4 Trypsin in-gel digestion ('slice and dice') protocol

Bands from stained gels were excised with a scalpel blade and chopped into small ($1mm^2$) cubes. Gel slices were kept moist with deionised water. Gel pieces from each slice were transferred to a polypropylene 0.6 mL microfuge tube and the excess water was removed. The gel was washed briefly with ammonium bicarbonate solution (100 mM NH_4HCO_3) to equilibrate the pH. To destain the gel, 200 µL of 50% acetonitrile-50 mM NH_4HCO_3 was added to each tube; samples were vortexed and incubated at room temperature for 10 min. Excess liquid was removed and 200 µl 100% acetonitrile was added to each tube. Samples were vortexed and incubated at room temperature for 10 min to allow the samples to dehydrate. Next the proteins were reduced and alkylated by covering the gel pieces with 50 µL of 5 mM tributylphosphine/20mM acrylamide in NH_4HCO_3. Samples were incubated at room temperature for 90 min. The solution was removed and gel pieces were washed with 100 mM NH_4HCO_3 for 5 min, followed by 50% acetonitrile-50 mM NH_4HCO_3 for 5 min.

Samples were then dehydrated with 200 µL of acetonitrile as above. Remaining liquid was removed and then the gel pieces were rehydrated with 12.5 ng/µL of trypsin in NH_4HCO_3 at 4°C for 30 min. More NH_4HCO_3 was added following re-swelling of the gel in order to keep the pieces covered with liquid. The proteins were then trypsin digested overnight at 37°C.

3.6.5 Mass spectrometry

Peptides were extracted by sonicating the samples in a waterbath for 10 min, centrifuging and transferring the digest solution to a clean tube. 30µL of 50% acetonitrile-2% formic acid was added to the gel pieces, followed by further sonication as above. The supernatant was removed and added to the previous supernatant sample in the clean tube. This step was repeated one further time to yield a peptide extract volume of >60 µL. The samples were then concentrated in a SpeedVac vacuum concentrator to a volume of 15 µL. Microparticulates were removed by centrifuging the samples at 14 000 rpm for 10 mins. The samples were then loaded into an autosampler tube and into the LC-MS/MS (Core Proteomics Facility, UTS) for analysis. Peptide mass fingerprints were mapped to the reference sequence (*Theileria orientalis*) and scores were assigned to protein matches using the Mascot server.

3.6.6 *In silico* analysis of the *T. orientalis* surfaceome

The predicted *T. orientalis* Ikeda proteome, based on open reading frames identified in the Japanese strain of *T. orientalis* Ikeda (http://totdb.czc.hokudai.ac.jp/), was subjected to an *in silico* analysis to identify probable surface antigens. The program TMHMM predicts which regions of a protein are likely to form a transmembrane helix based on the amino acid composition. Proteins with transmembrane helices frequently have epitopes exposed to the cell surface. All 3995 predicted open reading frames in the *T. orientalis* Ikeda genome, including all 4 chromosomes, as well as the apicoplast (an organelle unique to apicomplexan parasites) and mitochondrial genomes were subjected to TMHMM analysis using the online server at: http://www.cbs.dtu.dk/services/TMHMM/.

3.7 Development of a serological test for *T. orientalis*

3.7.1 Samples

A total of 430 EDTA blood samples and their matching sera were used to develop and validate the serological test. Of these, 280 pairs of samples were collected from 21 cattle herds from NSW and QLD. These samples were collected as part of routine clinical investigations into the significance of the *Theileria orientalis* genotypes in Australian cattle (Eamens et al., 2013b) or were submitted to the Elizabeth Macarthur Agricultural Institute as clinical samples from suspect theileriosis cases (Eamens et al., 2013c). A further 60 pairs of samples served as negative controls; 50 of these were collected from cattle herds located in regions known to be free of *T. orientalis* at the time of sampling (South Australia, 2012) and a further 10 pairs of samples were derived from cattle infected with *Babesia bigemina* (n=3), *Babesia bovis* (n=4) or *Anaplasma centrale* (n=3). The remaining pairs of samples (n=90) were collected from a single herd of 10 naïve animals (2 year old Ayrshire heifers) that had been introduced to a property on the mid-coast of NSW with a history of clinical theileriosis cases (Section 3.4.1). These samples were used for a temporal study examining

seroconversion to the *T. orientalis* MPSP. Briefly, EDTA blood was collected from each animal immediately upon introduction to the affected property, and approximately weekly thereafter for a period of 76 days. Sera were prepared from clotted blood or where clotted blood was unavailable; plasma was prepared from blood containing anti-coagulant.

3.7.2 Determination of packed cell volume (PCV)

Of the EDTA blood samples collected, 256 were examined for packed cell volume and were classed as either normal (PCV ≥ 24%) or anaemic (PCV < 24%).

3.7.3 Blood film examination

Blood films were prepared from a total of 376 EDTA bloods and were stained using Diff-Qik (Australian Biostain, Traralgon, Australia). Blood films were examined under 1000× magnification and were scored as positive or negative for intraerythrocytic *Theileria* piroplasms.

3.7.4 Cloning and expression of the *T. orientalis* MPSP genes

MPSP genes amplified from each of the *T. orientalis* Ikeda, Chitose and Buffeli MPSP types were cloned into the Champion pET100 D-TOPO expression vector (Invitrogen, Carlsbad, Calif., USA) according to the manufacturer's instructions. As described in Section 3.1.3.3, the primers used for PCR amplification of the MPSP gene fragments (Zakimi et al., 2006) were modified to ensure in-frame expression and to facilitate cloning into the pET100 vector. Primer sequences were as described in Table 4. PCR reactions were performed in a total volume of 50 µL containing 1 × Pwo polymerase reaction buffer, 100 µM dNTPs, 500 µM of each primer and 1U Pwo polymerase (Roche, Basel, Switzerland). Purified DNA (2 µL) from blood samples known to harbour the Buffeli, Chitose or Ikeda were used as template for the reactions. The Ts-Bc, Ts-Cc and Ts-Ic (forward) primers were each paired with the Ts-Rc (reverse) primer in amplification of the Ikeda, Chitose and Buffeli MPSP genes respectively. The integrity of each amplified product was assessed on a 1%-Tris borate EDTA (TBE) agarose gel and individual bands were excised and purified from the gel using the Wizard® SV PCR and Gel Clean-up System (Promega, Madison, WI, USA). The MPSP gene fragments were ligated into the pET100D-TOPO vector and transformed into *Escherichia coli* strain TOP10 One Shot® chemically competent cells (Invitrogen) according to the manufacturer's instructions. Transformants were selected on Luria-Bertani (LB) agar supplemented with 100 µg/mL ampicillin and colonies were patched onto fresh LB-ampicillin plates for subsequent PCR screening.

PCR reactions for screening transformants contained 1 × BioTaq™ buffer, 100 µM dNTPs and 500 µM of each primer. The forward primers used for PCR screening were the MPSP gene-specific primers Ts-Bc, Ts-Cc or Ts-Ic (for Buffeli, Chitose and Ikeda MPSP clones respectively), while the vector-specific primer T7R (Table 4) was used as the reverse primer. Clones yielding amplicons were considered to contain correctly oriented MPSP gene inserts and purified plasmid DNA (extracted using the QIAprep Miniprep kit (Qiagen, Alameda, Calif., USA) was then sequenced using the vector-specific T7F and T7R primers (Table 4). Sequencing reactions were performed by the Australian Genome Research Facility (AGRF, Sydney, Australia). Sequences were analysed to ensure that the vector insert sequences were free of errors and in the correct open reading frame for protein expression.

3.7.5 Recombinant protein expression and purification

Purified plasmid DNA containing Buffeli, Chitose or Ikeda MPSP gene inserts was transformed into *E. coli* expression strain BL21 Star (DE3) One Shot® chemically competent cells (Invitrogen) according to the manufacturer's instructions. Transformed cells were cultured on an orbital shaker (250 rpm) at 37°C overnight in 25 mL Luria Bertani broth (LB) each supplemented with 100 µg/mL ampicillin. The following day, 300 mL of pre-warmed LB supplemented with 100 µg/mL ampicillin were inoculated with 10 mL of overnight culture. Cultures were incubated at 37°C for 1.5 h. Protein expression was induced in each culture by the addition of isopropyl β-D-1-thiogalactopyranoside (IPTG) to a final concentration of 1 mM. Cultures were incubated for a further 2.5 h and cells were then harvested by centrifugation at 10,000 rpm in a Sorvall RC5C centrifuge. Recombinant MPSP antigens were purified by nickel affinity chromatography and dialysed as previously described (Jenkins et al., 2006). The concentration of the purified antigens was estimated using the bichinchoninic acid (BCA) assay (Thermo Scientific, Rockford, IL, USA).

3.7.6 Western blotting

Western blots of recombinant MPSP antigen derived from three genotypes of *T. orientalis* (Ikeda, Chitose and Buffeli; predicted sizes 34-35kDa) were screened with antisera from cattle previously confirmed PCR positive for a single MPSP genotype (Eamens et al., 2013b; Eamens et al., 2013c). The blots were performed as described previously using sodium dodecyl sulfate gel electrophoresis (SDS-PAGE) on a 12% polyacrylamide gel (200 volts, 1 hr) followed by transfer to polyvinylidene difluoride (PVDF) membrane (Jenkins et al., 2008). One microgram of each MPSP protein was blotted against a 1:100 dilution of each primary antiserum. The secondary antibody was a 1:20 000 dilution of alkaline phosphatase-conjugated anti-bovine (Sigma Aldrich, St Louis, Missouri USA; Catalogue number A7554). Blots were developed using NBT/BCIP substrate (Sigma, Catalogue number 11697471001).

3.7.7 ELISA

Antigens from the Ikeda, Chitose and Buffeli genotypes were mixed in equal concentrations to form an antigen cocktail with which to coat the ELISA plates. ELISA plates (Linbro/ Titertek E.I.A Cat no. 76-381-04, MP Biomedicals LLC., Santa Ana, CA, USA) were coated by passive adsorption with 0.75 µg of MPSP antigen per well diluted in 100 µL carbonate coating buffer (pH 9.6). The following day, plates were washed 5 × with PBS containing 0.05% Tween 20 (PBST) using an automatic 96-well ELISA washer and then blocked with 1 × Blocking Buffer (Cat. No. B6429; Sigma, St. Louis, MO, USA) for 1 h. After washing with PBST, 100 µL of each test serum sample and control sera diluted 1:100 in 1× blocking buffer were added to each well and incubated for 1 h in a humid chamber. Serum samples were tested in duplicate while positive and negative control sera were each tested in sextuplicate per plate.

After washing in PBST, 100 µL of monoclonal anti-bovine IgG (clone BG-18 alkaline phosphatase conjugate, Sigma) diluted in 1× Blocking buffer (Sigma) was added to each well and incubated for 1 h in a humid chamber. After further washing in PBST, plates were developed with BluePhos alkaline phosphatase substrate (KPL, Gaithersburg, Maryland, USA) (100 uL/well) on a plate shaker and then read at 610 nm (A610) on an XMark plate reader (Bio-Rad), at 5 min intervals. Optical density (OD) at 610 nm was determined for

each plate when the positive control serum had an ELISA ratio over 5 and the negative control serum an OD610 < 0.15. Results were expressed as an ELISA ratio (ER: mean OD610 test serum / mean OD610 of the negative control serum). Sera with an ER < 2 were considered negative, an ER ≥ 2 as positive.

3.7.8 DNA extraction and qPCR

DNA extractions were performed using the DNeasy Blood and Tissue DNA extraction kit (Qiagen) (Section 3.1.2.1). All DNA extractions and quantitative PCRs for *Theileria orientalis* (universal) and the genotypes Ikeda and Chitose (UIC qPCR) were performed as described in Section 3.1.3.2.

3.7.9 Statistical analysis

Associations between datasets were analysed using Spearman's rank correlation (r) for non-parametric data within Prism 4 (GraphPad Software Inc., La Jolla, California, USA). Fisher's exact test (GraphPad QuickCalcs: http://www.graphpad.com/quickcalcs/) was used to determine whether there was an association between serological status and clinical presentation (anaemic vs normal). The probability that infection level was related to serological status was determined using the Freeman-Halton extension of the Fisher's exact test via the VassarStats website (http://vassarstats.net/fisher2x3.html). P values calculated for both tests are two-tailed.

4 Results

4.1 Development and validation of an improved molecular test

4.1.1 Quality acceptance criteria

The quality criteria of all qPCR assays performed in this study were analysed at the end of each run using the DNA standards. All assays used in the comparisons presented met the quality criteria as outlined in Table 11. The quality parameters were met regardless of whether the U, I and C assays were run in uniplex or in multiplex indicating that there was no interference between the assays when run in a multiplex format.

Table 11. Quality data from the UIC qPCR assays performed in both multiplex and uniplex

	Universal		Ikeda		Chitose		Buffeli
	M	U	M	U	M	U	U
R^2	0.999	0.999	0.998	0.995	0.998	0.998	0.998
Efficiency (%)	95.0 (93.7-96.4)	95.4 (94.1-96.8)	95.2 (92.8-97.8)	96.7 (93.0-100.7)	100.0 (97.5-102.7)	97.1 (95.5-99.5)	93.4 (91.1-95.9)

4.1.2 Pilot studies

Two pilot studies were initially conducted in which a subset of 53 samples were analysed in both the qPCR assay and the gold standard cPCR assay, and in which the various DNA extraction methods were compared. The pilot studies indicated that there were no statistical differences in the number of positive detections in the qPCR compared to the cPCR assays when both neat and 1:10 dilutions of DNA extract were tested. However when only neat DNA extracts were considered, a significant number (P<0.0001) of positive samples were undetected in the cPCR screening (P32) assay relative to the qPCR assay, demonstrating the insensitivity of the qPCR to the presence of residual PCR inhibitors in the DNA extracts relative to the cPCR.

The pilot study also tested DNA extracts derived from the various inexpensive or high throughput DNA extraction procedures and indicated that all methods gave statistically similar results to the gold standard extraction method in the qPCR assay with the exception of the MagMax Viral kit. This method resulted in a statistically lower detection rate in all channels of the qPCR assay (Fig. 3). The crude, inexpensive DNA extraction method that was evaluated (DPK method) yielded very promising results with a numerically higher detection rate in the Universal and Ikeda reactions relative to the gold standard extraction method, and was subjected to further validation (Sections 3.2 and 4.2).

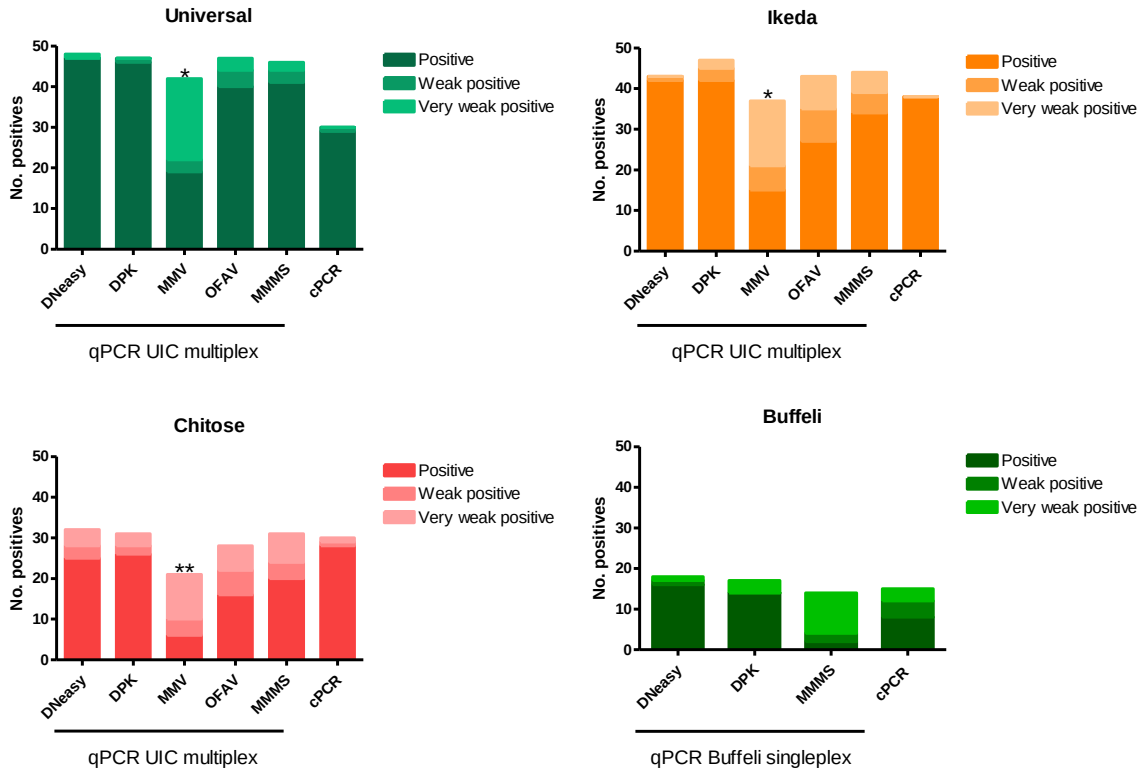


Fig. 3. Results from the pilot study examining the performance of neat DNA extracts from different extraction procedures in the qPCR assay. The results from the cPCR assay for DNeasy neat extracts (the gold standard extraction method) are also shown for comparison. Results statistically different from the DNeasy gold standard extract in the qPCR assay are indicated with an asterisk. (*P<0.05, ** P<0.01).

4.1.3 Large scale qPCR test validation

4.1.3.1 Analytical sensitivity and specificity

The linear dynamic range and analytical sensitivity statistics were determined in multiplex and uniplex for each probe used in the UIC qPCR assay as well as for the Buffeli probe in the uniplex assay. The linearity of each standard curve from reactions performed in uniplex and multiplex is demonstrated in Fig. 4A. For the U probe a linear range is observed between 3.0×10^1 and 9.0×10^7 MPSP GC/µL in both uniplex and multiplex. Similarly, the I and C components of the UIC assay exhibited a linear range between 1.5×10^1 and 1.5×10^7 MPSP GC/µL in both uniplex and multiplex. All linear ranges were observed with a coefficient of determination ≥99.5% (Table 11). Detection of single and pairwise combinations of the Ikeda, Chitose (A+B) and Buffeli plasmids within the universal (quantitative) component of the UIC assay was linear over the previously defined linear range (R^2 for each combination ranged from 0.997 – 0.999) (Fig. 4B). Furthermore, the means of each plasmid combination at each serial dilution were detected within a coefficient of variation (CV) of <20%, which is consistent with existing PCR assays that are considered quantitative (Armbruster and Pry, 2008; Buh Gasparic et al., 2008).

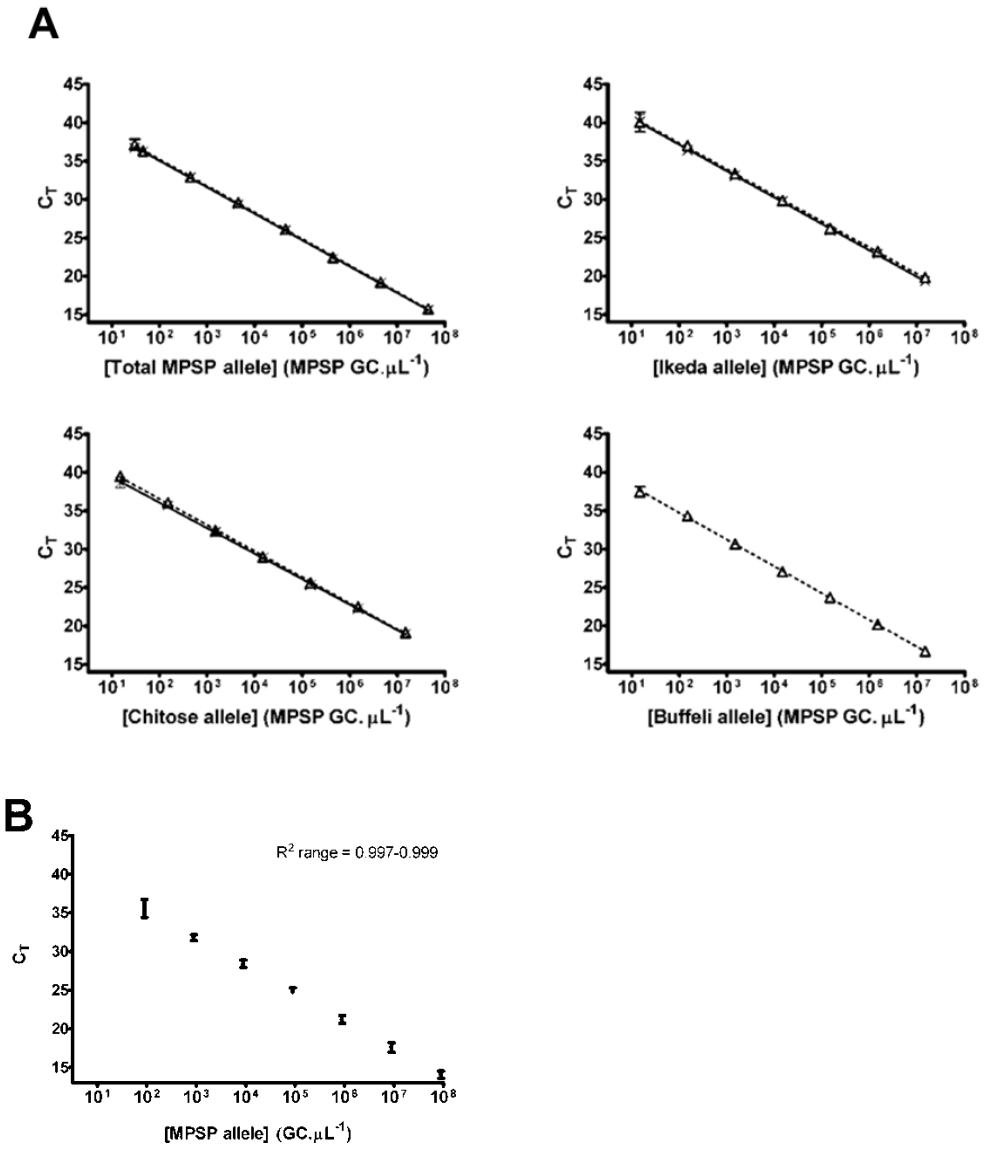


Fig. 4. The U component of the UIC assay is quantitative. A. Linear dynamic range for hydrolysis probe qPCR produced from 4-replicate assays using U (top left), I (top right), C (bottom left) and B (bottom right) probes. All qPCRs were performed using an equimolar mixture of Ikeda, Chitose (A+B) and Buffeli plasmid. Data from qPCRs performed in multiplex (×) and in uniplex (Δ) for the U, I and C components are compared. Data from the B assay in uniplex is shown. Lines-of-best-fit for multiplex and uniplex data are shown as solid and broken lines, respectively. Error bars represent standard deviation from mean. B. Means and ranges of the U component from triplicate UIC assays performed on serial dilutions of all possible combinations of Ikeda, Chitose and Buffeli plasmids (ICB, IC, IB, CB, I, C and B). R2 values for each combination ranged from 0.997 to 0.999.

The specificity of each probe for its target was confirmed using plasmids containing Ikeda, Chitose (A+B) and Buffeli MPSP gene sequences (Fig. 5). In addition, 31 samples collected from cattle cPCR-negative for *T. orientalis* that were infected with either *Anaplasma centrale*, *Babesia bigemina* or *Babesia bovis* were all negative in the UIC and B assays. Conversely, amplicons were consistently generated in the Ikeda cPCR assay when performed on *Anaplasma centrale*-infected bloods (Fig. 6) and to a lesser extent, *Babesia bovis* and *Babesia bigemina*-infected bloods. These amplicons were of a different molecular weight to

Theileria amplicons, therefore the assays could be scored as negative; however this is a factor that may confound interpretation of the cPCR Ikeda assay where animals are infected with *Anaplasma centrale* or *Babesia* spp. parasites.

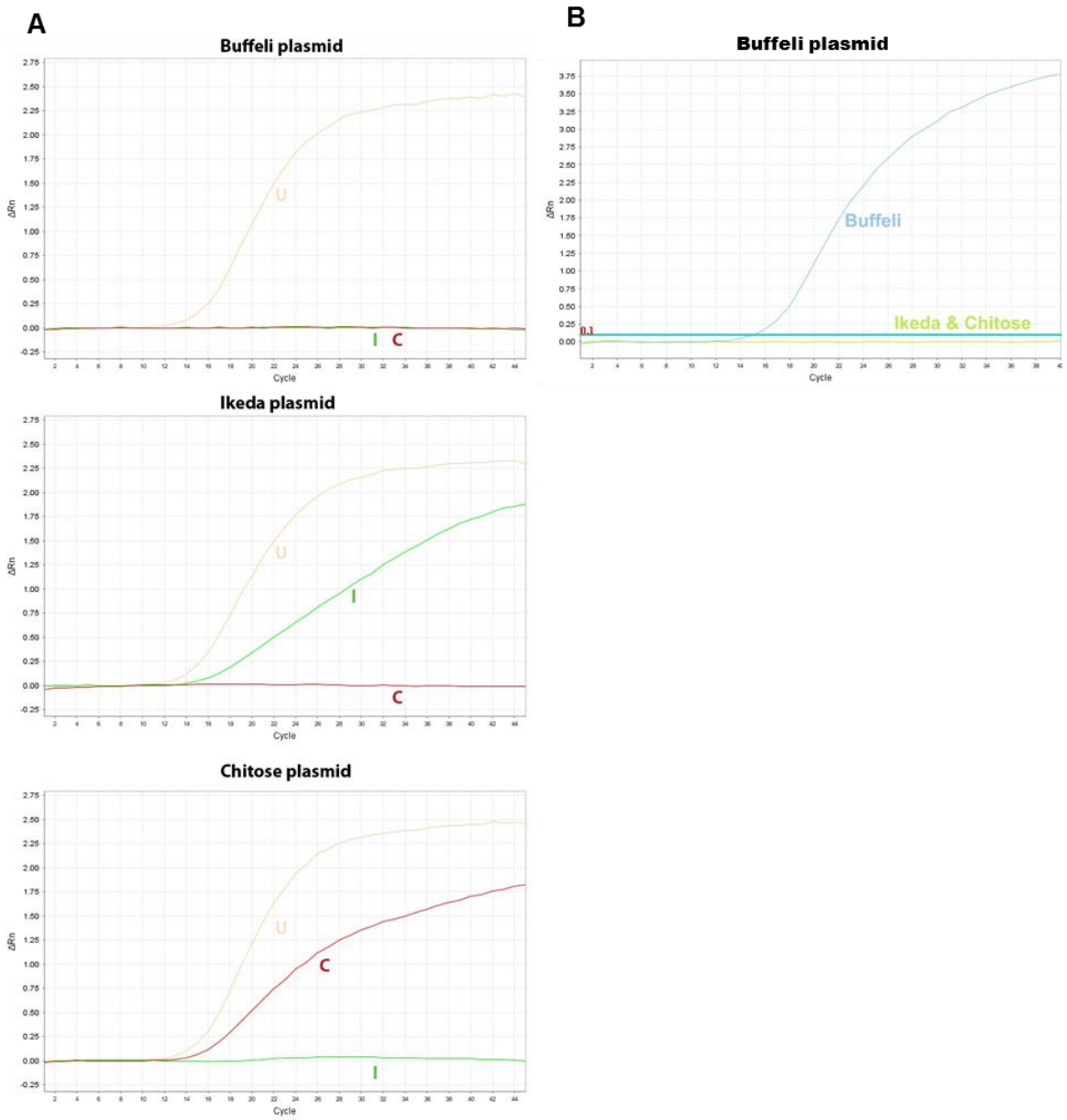


Fig. 5. Probe specificity of the UIC assay (A) and B assays (B). Results of three assays performed on plasmids (100 fg/µL) containing Buffeli, Ikeda and Chitose MPSP sequences.

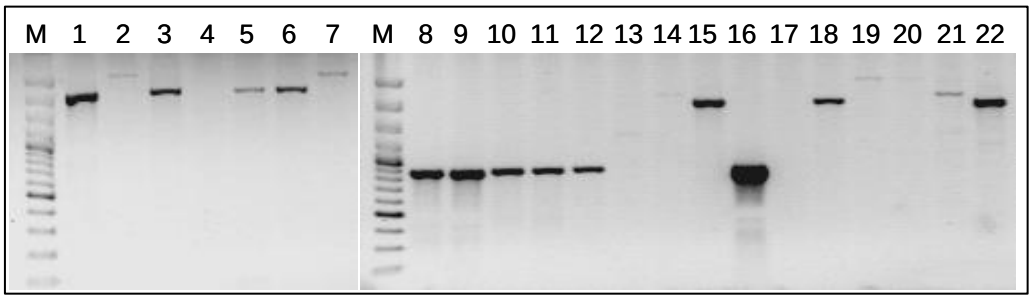


Fig. 6. Agarose gel showing results from the Ikeda cPCR assay performed on samples from the cross-reactivity panel. Amplicons were consistently observed in products derived from *Anaplasma centrale*-infected bloods (Lanes 1, 3, 5, 6, 15, 18 and 22). Faint products were also observed in products from *Babesia bovis*-infected samples (Lanes 2, 7, 19 and 20) and *Babesia bigemina*-infected samples (Lane 21). These non-specific amplicons differed in size from *Theileria orientalis* Ikeda PCR products (Lanes 8-12 and 16).

4.1.3.2 Limit of detection

The limit of detection (LOD) for each assay was defined as the point where 95% of replicate assays were positive and were determined for each probe in uniplex and for the U, I and C components when run in the multiplex format. LODs were shown to be of the same order of magnitude when the assay was run in uniplex or multiplex (Table 12).

Table 12. Limits of detection (LOD) for the each of the genotype assays in uniplex and the U, I and C components in multiplex

	Universal		Ikeda		Chitose		Buffeli
	M	U	M	U	M	U	U
LOD	17	28	27	16	20	16	20
(GC/µL)	(8-45)	(14-60)	(8-90)	(6-46)	(6-65)	(7-38)	(9-43)

4.1.3.3 Diagnostic sensitivity and specificity

The use of qPCR for the detection of infection requires the identification of a diagnostic threshold for the delineation of animals which are positive and negative for the presence of the organism. To identify this threshold the sensitivity and specificity of the U component of the UIC qPCR assay was examined against potential diagnostic thresholds (Fig. 7). A diagnostic threshold above the calculated LOD (Table 12) was selected that maximised both sensitivity and specificity of the U probe. From this analysis a diagnostic threshold of 30 MPSP GC/µL was selected.

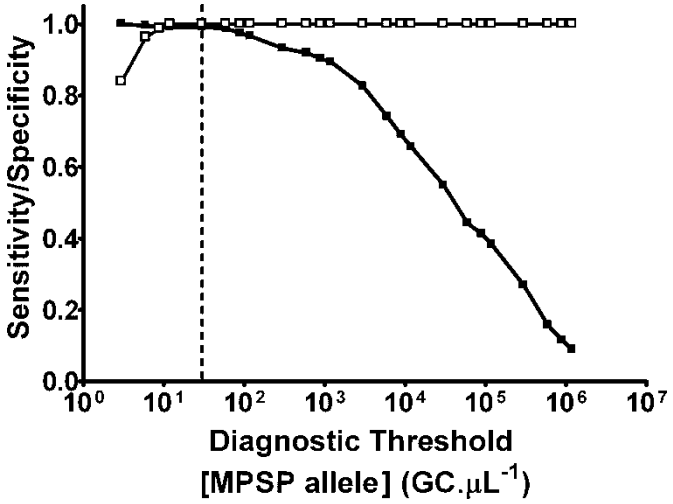


Fig. 7. *T. orientalis* detection limit using the U component. Sensitivity (■) and specificity (□) were calculated for various diagnostic thresholds (x-axis). A threshold of 30 MPSP GC/µL (dotted line) was selected as the diagnostic threshold for this assay, as it was a maximum for both sensitivity and specificity and above the calculated limit of detection.

Using this diagnostic threshold, 237 samples determined via cPCR to be *T. orientalis*-infected were analysed, and of these, 233 returned a positive result with the U component of the UIC qPCR assay (Table 13). Samples were considered qPCR positive if the concentration calculated from the U component was above the diagnostic threshold of 30 MPSP GC/µL. To rigorously test the detection limits of the qPCR assays used in this study, we examined samples that were weakly positive in cPCR. Twenty nine samples with very faint bands in cPCR were also identified as positive in qPCR (<1000 MPSP GC/µL), demonstrating that the UIC qPCR compares very well with the corresponding cPCR assays at low concentrations. An overall sensitivity of 98.3% was observed for the detection of *T. orientalis* infection and the specificity was 100.0%. No significant difference in the proportion of positives between the two tests was identified between the UIC qPCR assay and *T. orientalis* infection status as determined by cPCR using McNemar's Test (Table 13). While the sensitivity of the U probe was very high, a small number of *T. orientalis*-infected samples (n = 4) did not test positive in the UIC qPCR assay (Table 13) but were weakly positive (faint bands observed) in the p32 cPCR. Of these four samples, three demonstrated detectable amplification in the UIC qPCR but fell below the diagnostic threshold. DNA sequencing indicated that two of these three samples were of the Buffeli genotype and the remaining sample was of the Chitose genotype. The fourth sample was negative upon re-test in the p32 cPCR. Conversely, two samples included in the T. orientalis-infected group, which were detected as cPCR positive for Chitose and Buffeli but were negative when tested with the p32 cPCR, were positive when tested with the U component of the UIC qPCR assay. Both of these samples subsequently yielded positive results in the p32 cPCR assay upon re-testing and both samples were confirmed to be genuine Chitose and Buffeli positives respectively upon DNA sequencing.

Table 13. Contingency tables and descriptive statistics comparing detection of *T. orientalis* infection using the UIC qPCR assay with infection status as determined by the 6 cPCR assays (Universal, Ikeda and Chitose at two dilutions). 95% confidence limits are shown in brackets.

	T. orientalis present	*T. orientalis* absent	**Total**		
qPCR positive	233	0	**233**	Relative sensitivity	98.3% (95.7% to 99.5%)
qPCR negative	4	81	**85**	Relative specificity	100% (95.6% to 100.0%)
Total	**237**	**81**	**318**	Difference (P < 0.05)	NS*

*NS – not significant

4.1.3.4 Subtype differentiation

To detect the presence of *T. orientalis* genotypes associated with clinical disease, two probes were included in the UIC assay to detect the Ikeda and Chitose genotypes. A diagnostic threshold of 30 GC/µL was similarly selected for the I and C components of the UIC assay as this was above the limits of detection (Table 12) and maximised the sensitivity and specificity of these components. In preliminary experiments with plasmid, detection of the I and C genotypes was found to be near-identical for equimolar mixtures of template. In the case of unequal concentrations of template, the dominant genotype was detected in all cases. The minor (non-domimant) genotype was detected in the majority of reactions in which the dominant type was up to 3 orders of magnitude higher in concentration than the minor genotype. The ability of the I and C assays to detect each genotype was then assessed against the genotype specific cPCRs using the 237 *T. orientalis*-infected blood samples from the sensitivity and specificity panels. Of the 181 samples testing positive for the Ikeda type in cPCR, 180 were Ikeda positive in qPCR. The remaining sample yielded amplicon in the I component of the UIC assay, but fell below the diagnostic threshold, and was therefore scored as negative. The UIC assay detected a further two samples that tested Ikeda negative in the cPCR. These samples were later confirmed as genuine positives by DNA sequencing. The overall sensitivity of the I component of the UIC assay was 99.4% and the specificity was 96.4% (Table 14). 144/237 samples were positive when tested with the Chitose type-specific cPCR and of these 129 were positive with the C component of the UIC qPCR assay giving a sensitivity of 89.6%. The samples testing Chitose negative in the qPCR but positive in the cPCR were all from samples with mixed infections containing extremely high levels of the Ikeda genotype suggesting that competition for reagents was a factor in the lower sensitivity of the C component of the UIC assay. Despite this, the C component of the UIC assay detected a further 7 positives that tested negative in the Chitose-specific cPCR yielding a specificity of 92.6% (Table 14). These 7 samples were later confirmed to be genuine positives via DNA sequencing. Although the number of Chitose-positive samples detected by the UIC was lower overall, there was no

statistical difference in the number of samples testing positive in cPCR (n = 144) versus qPCR (n = 136) (P<0.05).

The Buffeli uniplex assay was also assessed against the Buffeli cPCR assay. A total of 98 samples were positive in qPCR compared to 101 positive by cPCR. Nonetheless there were a number of discrepant samples between tests with 8 testing qPCR positive but cPCR negative and 11 testing cPCR positive but qPCR negative. The discrepant results were all from samples yielding very weak bands in the cPCR or low gene copy numbers in the qPCR around the limit of detection of this assay. The relative sensitivity of the Buffeli uinplex assay was 89.1% and the relative specificity was 94.1%. There was no statistical difference between these tests using McNemar's Test (Table 14).

Table 14. Comparison of UIC multiplex qPCR and type-specific cPCR for the detection of *T. orientalis* subtype, 95% confidence limits shown in brackets.

Type	qPCR/cPCR				Relative Sensitivity (%)	Relative Specificity (%)	Difference (P < 0.05)
	+ / +	+ / -	- / +	- / -			
Ikeda	180	2	1	54	99.4 (96.6-100.0)	96.4 (87.9-99.6)	NS
Chitose	129	7	15	86	89.6 (83.4- 94.1)	92.6 (85.3- 97.0)	NS
Buffeli	90	8	11	128	89.1 (81.3-94.4)	94.1 (88.7- 97.4)	NS

*NS – not significant

4.1.3.5 Investigation of potential PCR inhibition

Previous studies have identified that neat DNA extracts and 10-fold dilutions of sample must be run in parallel when performing the *T. orientalis* cPCR assays to account for PCR-inhibiting contaminants present within nucleic extracts from blood (Eamens et al., 2013a; Eamens et al., 2013b; Eamens et al., 2013c). To overcome this, we used a DNA polymerase mastermix that has been demonstrated to be effective at reducing contaminant-based inhibition (TaqMan Environmental MasterMix). We asssessed the efficacy of this mastermix in reducing or eliminating the effects of PCR inhibition during development of the UIC and B assays. No additional positive reactors were detected in undiluted and 10-fold diluted extracts from the 237 samples from the sensitivity panel were examined in the UIC assay. Evidence of inhibition in the cPCR assays (i.e. no or significantly less PCR product observed in undiluted extracts when compared to 10-fold dilutions) was observed in 10 of the 237 extracts from the specificity panel. The same extracts were all positive for both undiluted and 10-fold diluted extracts when tested in UIC qPCR. Furthermore, a higher concentration (MPSP GC/µL) was observed in the quantitative U component of the UIC assay for all undiluted extracts compared to their respective 10-fold dilutions. To further examine whether PCR inhibition had an effect on assay quantification, we prepared serially-diluted

plasmid DNA containing the Ikeda, Chitose and Buffeli alleles and spiked these into nucleic acid extracts of samples from regions where *T. orientalis* was not enzootic and that were also confirmed negative by both cPCR and qPCR. Plasmids were serially diluted 10-fold to concentrations spanning 1.5×10^1 to 1.5×10^7 MPSP GC/µL. In all experiments, the qPCR predicted concentrations compared well with the concentrations determined by spectrophotometry (Fig. 8). Together, these data indicated that PCR inhibition is not significant in either the UIC multiplex or the B uniplex qPCR assays. Nonetheless, it is recommended that a positive control, (or a negative control spiked with plasmid DNA) is included in each DNA extraction run, along with a negative control sample, as a batch control measure.

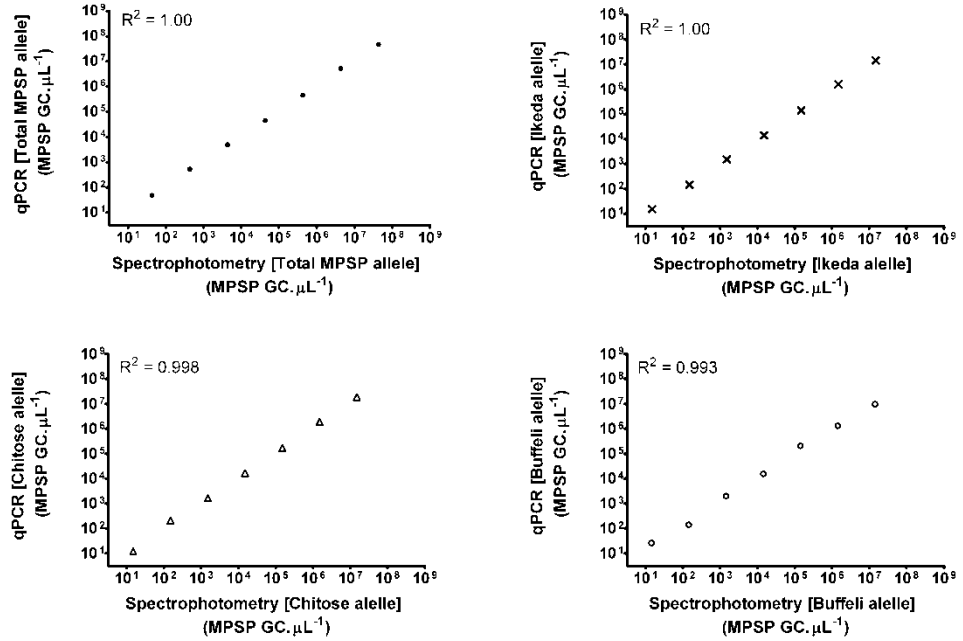


Fig. 8. Examination of the UIC multiplex and B uniplex qPCR assays on plasmids spiked into nucleotide extractions sourced from T. orientalis-negative cattle blood. U (top-left), I (top-right), C (bottom-left) and B (bottom-right) probes were used to detect spiked plasmids. Y-axis represents qPCR derived concentrations, determined with a standard plasmid series. X-axis represents spectrophotometry derived allele concentrations, determined by absorbance at 260 and 280 nm prior to dilution into T. orientalis-negative blood extractions.

4.1.3.6 Correlation with parasitaemia and PCV

Of the 237 samples from the specificity panel, 135 had corresponding data for packed cell volume (PCV) and parasitaemia and these were used for comparative analyses with total MPSP allele concentration as determined by the U component of the UIC qPCR (Fig. 9). Because blood films were scored as a percentage of erythrocytes containing piroplasms (% parasitaemia), a correction factor (as described in the Methods) was also applied to the qPCR data to account for variable erythrocyte levels, yielding an erythrocyte independent measure, gene copies per µL of packed erythrocytes (GC/µL PE). Strong and significant positive correlations ($P < 0.0001$) were observed between MPSP allele concentrations (in GC/µL and GC/µL PE) and % parasitaemia (Fig. 9A and B). A moderate negative correlation ($r = -0.482$, $P < 0.0001$) was observed when MPSP allele concentrations expressed in the erythrocyte concentration dependent measure (GC/µL) were compared with PCV values

(Fig. 9C). This correlation was stronger when MPSP allele concentrations were corrected for erythrocyte concentration (r = -0.640, P < 0.0001) (Fig. 9D).

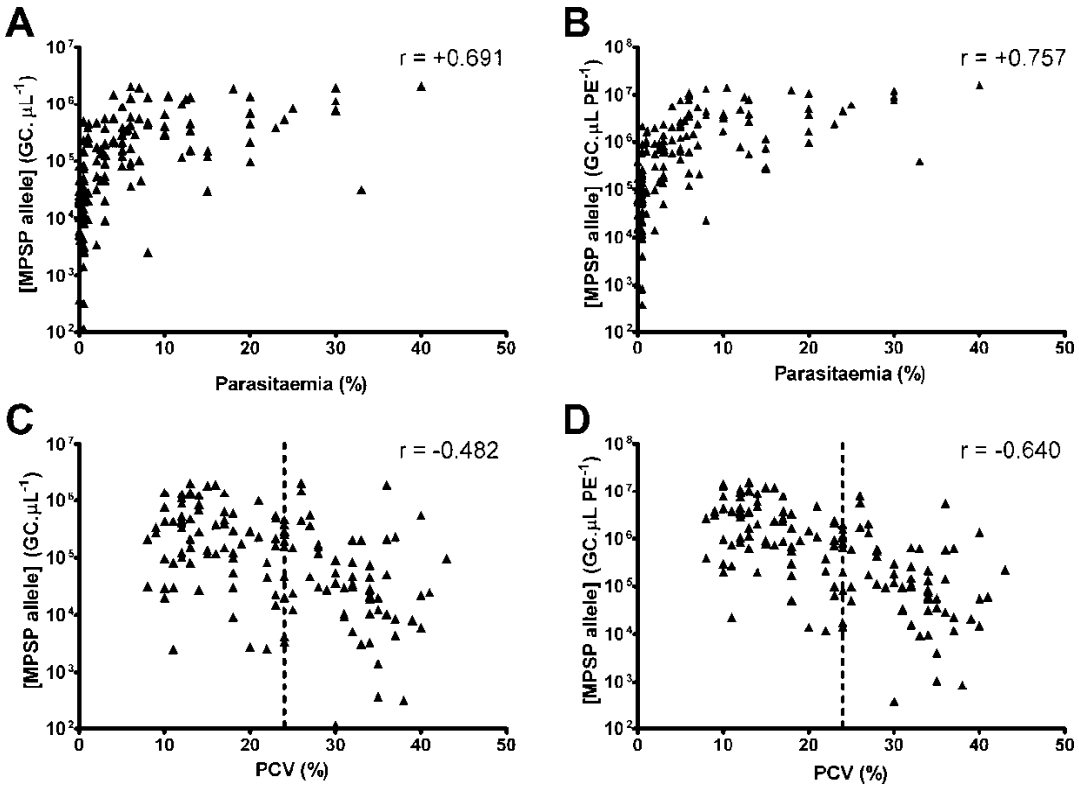


Fig. 9. MPSP allele concentrations determined by the UIC assay correlate with blood film analysis and PCV. A & B. Correlation of MPSP allele concentration (calculated from the U component of the UIC qPCR multiplex) expressed in measures (A) dependant on, and (B) independent of erythrocyte concentration with parasitaemia, as determined by blood film examination. C & D. Correlation of MPSP allele concentration (calculated from the U component of the UIC qPCR multiplex). expressed in measures (C) dependant on, and (D) independent of erythrocyte concentration with PCV. Anaemia corresponds to a PCV < 24% and is represented by vertical dotted lines.

4.1.3.7 Determination of a clinical threshold for the UIC qPCR

As total MPSP allele concentration strongly correlates with parasitaemia and PCV, we investigated whether the UIC qPCR could be used to define a clinical threshold for T. orientalis infection. Samples from the sensitivity panel were divided into four groups based on the overall clinical presentation of the donor animals as well as herd status (Fig. 10A). One way ANOVA demonstrated that animals that were clinically affected at the time of sampling (n = 110), had significantly higher mean MPSP allele concentrations (GC/µL) than recovering (n = 25) or in contact animals (n = 24) from herds with clinical cases, as well as subclinical (but *Theileria* positive) animals from herds without clinical cases (n = 55) (P<0.001). In addition, recovering and in contact animals displayed significantly higher MPSP allele concentrations than the subclinical animals from the herds without clinical cases (P<0.001). Recovering and in contact animals from herds with clinical cases did not have statistically different MPSP allele concentrations (P>0.05). Based on these data, we propose three groups of T. orientalis infection severity as determined by the UIC qPCR. The

first group describes high level *T. orientalis* infection strongly associated with clinical signs of disease and corresponding to a GC/µL > 3.0×10^5 (Fig. 10A). The second group consists of samples from animals with a moderate level *T. orientalis* infection corresponding to a GC/µL of > 1.5×10^4 but ≤ 3.0×10^5. Samples from the third group are defined by a low level of T. orientalis DNA corresponding to a GC/µL ≤ 1.5×10^4. Using these thresholds, 100% of animals (n = 55) from herds without clinical cases were classified into the low level infection group. In contrast, 91% (100/110) of animals that displayed clinical symptoms at the time of sampling were within the moderate to high level ranges. Above the upper threshold of 3×10^5 MPSP GC/µL, (high level infection) 95% of animals (42/44) were clinically affected. In addition, the majority of animals (23/25), that were recovering from clinical theileriosis fell within the low and moderate ranges, while all in contact animals (n = 24) fell within the low and moderate ranges.

4.1.3.8 Association between genotype and parasite load

The relationship between the presence of the Ikeda and Chitose genotypes and MPSP allele concentration was examined by scoring each sample as positive in each of the three channels of the UIC assay (Fig. 10B). The samples testing positive in the U assay only (n = 24) were all identified as Buffeli positives in B uniplex assay. Animals harbouring the Ikeda type only (n = 69), or the Ikeda type in combination with the Chitose type (n = 103) had significantly higher MPSP allele concentrations than animals that were Chitose only positive or samples in which no pathogenic genotype was detected ($P<0.001$). The mean MPSP allele concentrations of the C only and B only positive samples were not significantly different ($P>0.05$). Similarly, the mean allele concentrations of the I and IC positive samples were not statistically different ($P>0.05$). These data suggest that the Ikeda genotype in particular is associated with higher overall parasite loads. Indeed, 100% of samples collected from clinically-affected and recovering animals examined in this study were found to be I only positive or both I and C positive (with or without the Buffeli genotype). Semi-quantitative data from the I and C assays indicated that *T. orientalis* Ikeda was dominant genotype in 58% of mixed Ikeda-Chitose infections. Of animals that tested C positive only (n = 18), 16 were from subclinical herds, while the remaining 2 samples were derived from in contact animals within clinically-affected herds.

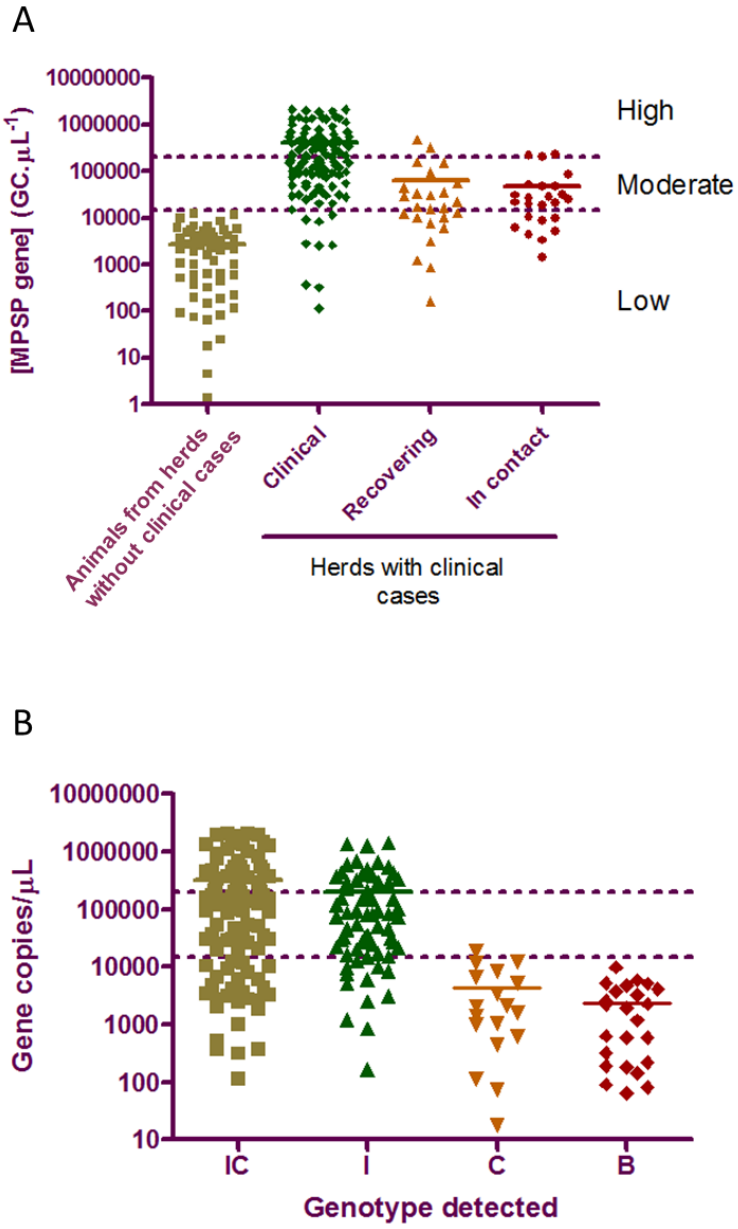


Fig. 10. A. Association of MPSP allele concentration with clinical data allows the delineation of clinical thresholds. MPSP allele concentrations (calculated from the U component of the UIC qPCR multiplex) were separated into 4 clinical groups: subclinical animals from herds without clinical cases (ie. herds with no clinical cases of theileriosis but situated in T. orientalis enzootic areas), animals with clinical symptoms consistent with theileriosis at time of sampling (Clinical), animals that had exhibited clinical symptoms prior to sampling (Recovering), and animals in contact with theileriosis-affected animals but not exhibiting clinical symptoms at time of sampling (In contact). Clinical thresholds are shown as dotted lines. B. Association of MPSP allele concentration with clinical genotype detected. MPSP allele concentrations (calculated from the U component of the UIC qPCR multiplex) were separated into 4 groups based on genotype detected by the I and C components. Clinical thresholds are shown as dotted lines.

4.2 Validation of an inexpensive DNA extraction method for *Theileria* DNA

The DNeasy and DPK methods were found to detect an equal number of positives in the Universal component of the multiplex assay, while an additional positive was detected using the DPK method in the Ikeda and Chitose A assay components of the multiplex assay as well as the Buffeli uniplex assay. There were discordant results for 2, 5, 9 and 7 samples for each of these assay components respectively (Table 15). Examination of the discordant samples that were positive when extracted with the DPK method revealed that all were very low level positives (just above the diagnostic threshold of the assay), and displayed amplification but fell just below the diagnostic threshold when extracted with the commercial method. All discordant samples that were positive when extracted with the commercial method also demonstrated below threshold amplification when extracted with the DPK method. McNemar's test demonstrated that there was no significant difference between methods for the matched pairs of qPCR results for each assay component ($P>0.05$). There were no instances where amplification was observed in the 10-fold diluted samples without corresponding amplification in the paired undiluted extract, suggesting that PCR inhibitors were not a significant issue using either the commercial or the crude DPK DNA extraction method. This may be related to the use of a qPCR mastermix designed to be relatively insensitive to PCR inhibitors.

Table 15. Number of positive and negative samples detected in *T. orientalis* multiplex qPCR using two DNA extraction methods

Assay component	qPCR outcome (Commercial method/DPK method)*				Sensitivity	Likelihood ratio	Accuracy
	+/+	+/-	-/+	-/-			
T. orientalis (Universal)	316	1	1	116	99.68%	118.6	99.5%
Ikeda	266	2	3	163	99.25%	54.83	98.8%
Chitose A	154	4	5	271	96.86%	53.51	97.9%
Buffeli	155	3	4	272	98.1%	67.69	98.3%

*$P>0.05$ for each assay component.

Regression analysis of quantitative data from the Universal component of the assay (Fig. 11) demonstrated that the DPK extraction method yielded slightly higher gene copy numbers than the commercial method (16% higher on average) and that variation was consistent over the range of DNA concentrations. This indicates that a higher yield of DNA was obtained using the DPK method relative to the commercial method, which could be due to the

difference in collection technique. DNA collection is based on precipitation in the DPK method, and presumably allows for more thorough recovery of DNA than the silica membrane binding method used by the commercial kit. Furthermore, >95% of paired (commercial and DPK), qPCR-derived, T. orientalis DNA concentrations were within an order of magnitude of each other, suggesting that DNA extracted via the DPK method results in reliable quantification of *T. orientalis* MPSP gene copies.

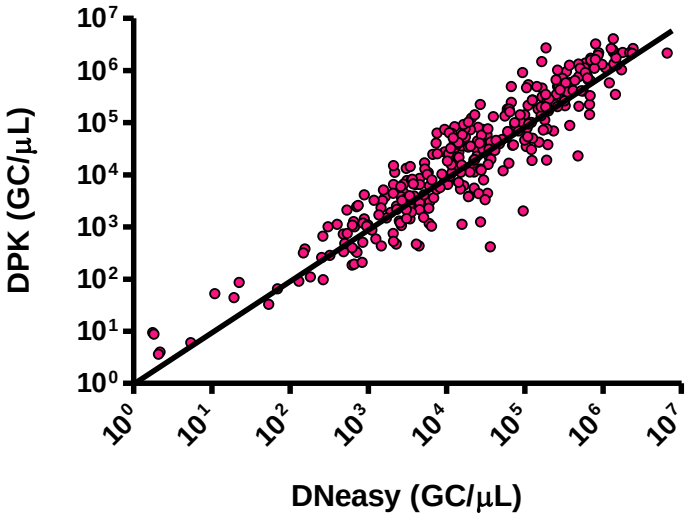


Fig. 11. Regression analysis of qPCR data from the quantitative (universal) component of the multiplex assay. Log10 MPSP gene copies determined from DPK DNA extracts are plotted against those from DNeasy extracts. Variation in results obtained from each method was consistent over a range of DNA concentrations. More data points are located above the regression line, demonstrating that of the two methods, DPK yields a greater concentration of DNA overall.

4.3 Comparison of the performance of the UIC qPCR with a commercial MT-PCR (AusDiagnostics)

A comparison of the results from the UIC plus Buffeli type-specific qPCR assays and the those from the MT-PCR assay is shown in Table 16. All 15 *T.orientalis* positive samples were detected by each assay and both assays yielded a negative result in the negative control. The UIC qPCR assay detected 12 samples positive for the Ikeda genotype, while MT-PCR detected 10 samples positive for this type with an additional 2 samples requiring manual confirmation due to deviation of the melt curves from the expected pattern. The UIC multiplex detected 5 samples positive for Chitose genotype while MT-PCR assay detected an additional 3 samples positive for the Chitose type, although 2 of these were below the limit of detection of the MT-PCR assay (20 gene copies) as determined by (Perera et al., 2015a). Three samples were positive for the Buffeli genotype according to the Buffeli uniplex qPCR; however MT-PCR detected an additional 11 samples positive for the Buffeli genotype (Table 16, underscored), many of which appeared to be at high infection intensities (see below).

MT-PCR is a semi-quantitative assay and does not contain a universal component for estimation of overall parasite load; however addition of the MT-PCR gene copy numbers for individual genotypes yielded similar results to the universal component of the UIC assay, except at high parasite loads in which case the MT-PCR overestimated the parasite load by approximately 3-5 fold. This is likely due to the fact that MT-PCR uses a single internal reference (10 000 gene copies) for calculation of gene copy number while UIC qPCR uses a standard series encompassing 6 orders of magnitude of DNA concentration. High infection intensities deviating by several orders of magnitude from the internal reference are therefore less likely to be accurate in MT-PCR.

Samples yielding positive results in MT-PCR, but negative results in the B uniplex assay were subjected to cloning and DNA sequencing (samples 1, 4, 6, 7 and 9-15). Clones and sequences were successfully obtained derived from the MT-PCR assay for all but two of these samples (ie: 1, 4, 6, and 9-14). In all cases, the PCR product was shown to correspond to the *T. orientalis* P23 gene (the target for the MT-PCR Buffeli assay); however all amplicons were of the Ikeda genotype, not the Buffeli genotype, indicating that the MT-PCR Buffeli assay cross-reacts with Ikeda genotype (Fig. 12). As shown in Table 16, the 11 additional samples for which the MT-PCR was "Buffeli positive" relative to the B uniplex assay (underscored) were all positive for Ikeda genotype further confirming cross-reactivity in MT-PCR.

Table 16. Comparison of quantitative data derived from UIC multiplex/B uniplex qPCR and MT-PCR

Sample	Parasite load determined by assay (GC/µL blood)							
	Universal		Ikeda		Chitose		Buffeli	
	MT-PCR	UIC muliplex	MT-PCR	UIC multiplex	MT-PCR	UIC multiplex	MT-PCR	B uniplex
1	N/A	141	55	161	0	0	59	0
2	N/A	321,116	0	0	142,909	317,774	0	0
3	N/A	2,534	0	0	18	0	8,506	11,792
4	N/A	79,524	58,159	96,033	0	0	37,446	0.0
5	N/A	701,267	1,415,897	551,435	151,948	235,652	227,354	3,606
6	N/A	182,601	16,480	51,954	65,903	153,783	13,487	0
7	N/A	119,687	69,671	149,285	2,146	144	46,981	0
8	N/A	471	0	0	0	0	2,510	5,750
9	N/A	157,710	*Check*[ǂ]	187,812	7	0	78,996	0
10	N/A	322,761	*Check*[ǂ]	392,952	105	0	157,735	0
11	N/A	306,747	1,315,169	359,900	0	0	150,267	0
12	N/A	129,696	75,934	162,434	0	0	57,743	0
13	N/A	180,072	60,079	110,990	30,350	89,997	35,836	0
14	N/A	146,510	94,691	178,170	0	0	60,417	0
15	N/A	174,288	100,903	207,413	0	0	74,387	0
16*	N/A	0	0	0	0	0	0	0

[ǂ]Melt curve deviated from standard

*Indicates negative control sample

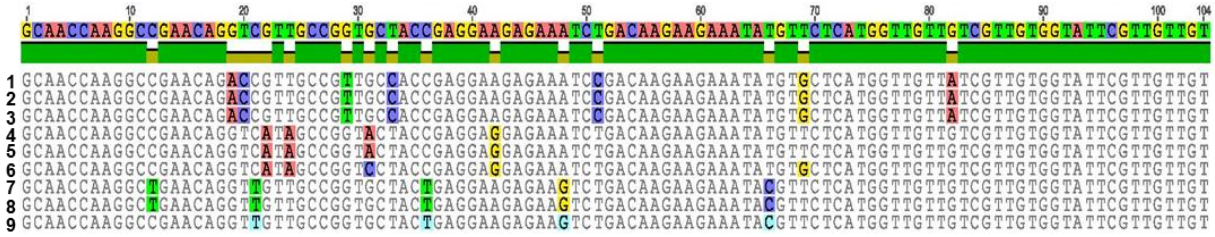


Fig. 12. Consensus sequence of MT-PCR P23 Buffeli assay (sequence 9) aligned with reference P23 sequences from Chitose (sequences 1-3), Buffeli (sequences 4-6) and Ikeda (sequences 7-8) genotype. The sequence from the MT-PCR (sequence 9) is of the Ikeda genotype rather than the Buffeli genotype indicating cross-reactivity in this component of the MT-PCR assay.

The cost of the UIC qPCR and MT-PCR assays was also compared. Perera et al. (2015a) report that MT-PCR costs $19 per sample (reagent cost only), excluding DNA extraction. Excluding DNA extraction, UIC qPCR has been costed at <$5 per sample (reagent cost only) based on 10-80 samples. The cost saving is largely a result of multiplexing the reactions in a single well with UIC qPCR, while MT-PCR consists of a series of uniplex reactions requiring more reagent.

Reagent cost for DNeasy DNA extraction (validated with both methods) costs $6.50; while the DPK method which has been validated for UIC qPCR, costs approximately 10c per reaction.

4.4 Genotyping of *T. orientalis* strains

4.4.1 MPSP gene analysis

MPSP genes from blood samples derived from animals with and without clinical signs and from *H. longicornis* ticks were collected from diverse geographical locations and sequenced to assess diversity within each MPSP type. The MPSP sequences were also compared to those from reference strains from overseas. The phylogenetic tree constructed from this dataset is shown in Fig. 13. The Ikeda sequences examined in this study were highly homologous. The strains sequenced in this study were 100% identical in terms of their MPSP gene sequence, with the exception of one sequence derived from a tick (Tick 5) which had 99.7% identity with the rest of the sequences. One MPSP sequence obtained from an infected animal in Guyra, NSW (Kamau et al., 2011a) (Sample 15, Guyra, see Fig 13) differed considerably from the sequences derived from this study, suggesting that this may represent either a chimeric sequence or perhaps a type of very low prevalence. Interestingly, the Ikeda samples examined from the southwest of Western Australia, a region geographically distant from the east coast where most theileriosis cases occur, all contained an identical MPSP gene sequence to the east coast samples. In addition, all Australian MPSP Ikeda sequences were 99.7% identical to that derived from *T. orientalis* Ikeda from Japan (*T. sergenti* D11046), suggesting a close association between the Australian Ikeda MPSP and the Japanese type (Fig. 13).

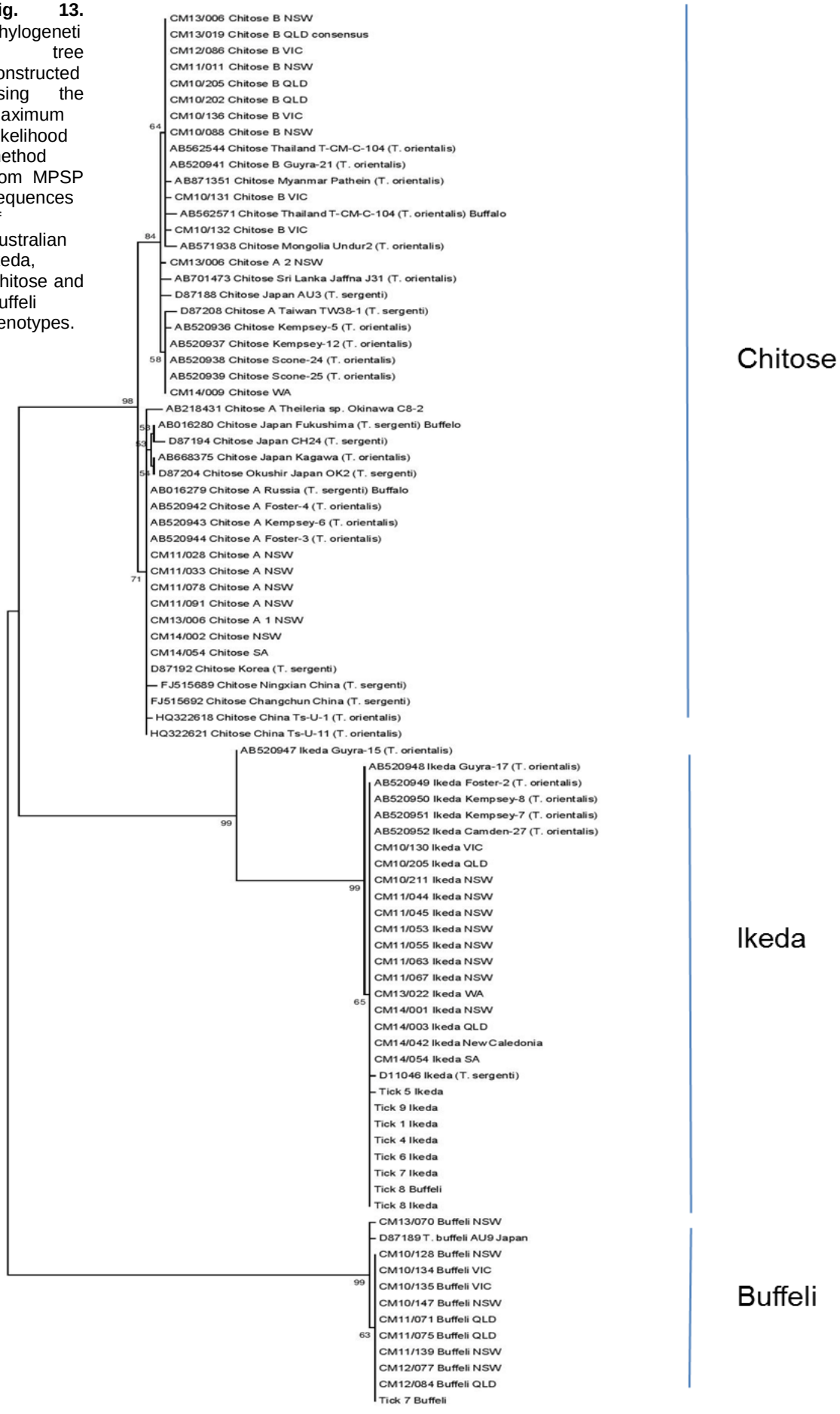


Fig. 13. Phylogenetic tree constructed using the Maximum Likelihood method from MPSP sequences of Australian Ikeda, Chitose and Buffeli genotypes.

F

The Buffeli MPSP genes sequenced in this study were also 100% homologous to each other with the exception of sample CM13/070 from Lismore, NSW which displayed 99.6% identity with the other sequences and sample Tick 7 which displayed 99.7% identity with the other sequences. The Buffeli MPSP sequences all displayed between 99.4-99.7% identity with an MPSP gene sequenced from an animal imported into Japan from Australia.

While the role of *T. orientalis* Ikeda in clinical theileriosis has been noted in several prior studies (Eamens et al., 2013c; Kamau et al., 2011a; Perera et al., 2013), the role of the Chitose genotype is less clear. Sequence identity between Chitose sequences examined in this study was also very high, ranging between 98.5-100%. Nonetheless, Chitose MPSP genes displayed a number of polymorphic sites that separated these strains into 2 groups. A phylogenetic subtree of the Chitose genotype is presented in Fig. 14. Twenty two *T. orientalis* Chitose MPSP gene sequences derived from individuals within herds displaying clinical theileriosis cases, and from herds with subclinically-affected animals were compared with reference sequences. Two distinct phylogenetic clusters of *T. orientalis* Chitose (referred to here as Chitose A and Chitose B) were consistently observed using disparate phylogenetic methods and were well-supported by the bootstrap analyses. These two phylogenetic clusters were characterised by a series of single nucleotide polymorphisms within the MPSP gene sequences. There was slightly more sequence diversity within the Chitose B cluster (98.7-100% identity) than the Chitose A cluster (99.3-100% identity). The Australian Chitose sequences were polyphyletic, with sequences from both the Chitose A and B clades represented; however the sequences analysed from Queensland, where relatively few clinical outbreaks have occurred, were all of the Chitose B type. Futhermore, sequences belonging to the Chitose A cluster were all from New South Wales where clinical outbreaks have been occurring since 2006 and showed a close relationship with Chitose sequences from Japan. Of the Australian samples sequenced in this study, it was noted that all sequences from the Chitose A phylogenetic cluster (n = 7) were derived from herds with clinical cases. In addition, 3 Australian sequences (Foster3, Foster4 and Kempsey6) from a previous study (Kamau et al., 2011a) were also derived from clinically-affected herds. All of the samples sequenced in this study from herds with subclinically infected animals (n = 9) belonged to the Chitose B cluster; however this cluster also included some sequences from herds with clinically-affected (n = 6) animals. Finally, all sequences within the Chitose A cluster had been identified in prior studies as mixed infections, all co-occurring with the Ikeda genotype (Eamens et al., 2013a), while sequences from the Chitose B cluster derived from both cases of sole Chitose infection or infections of mixed genotype.

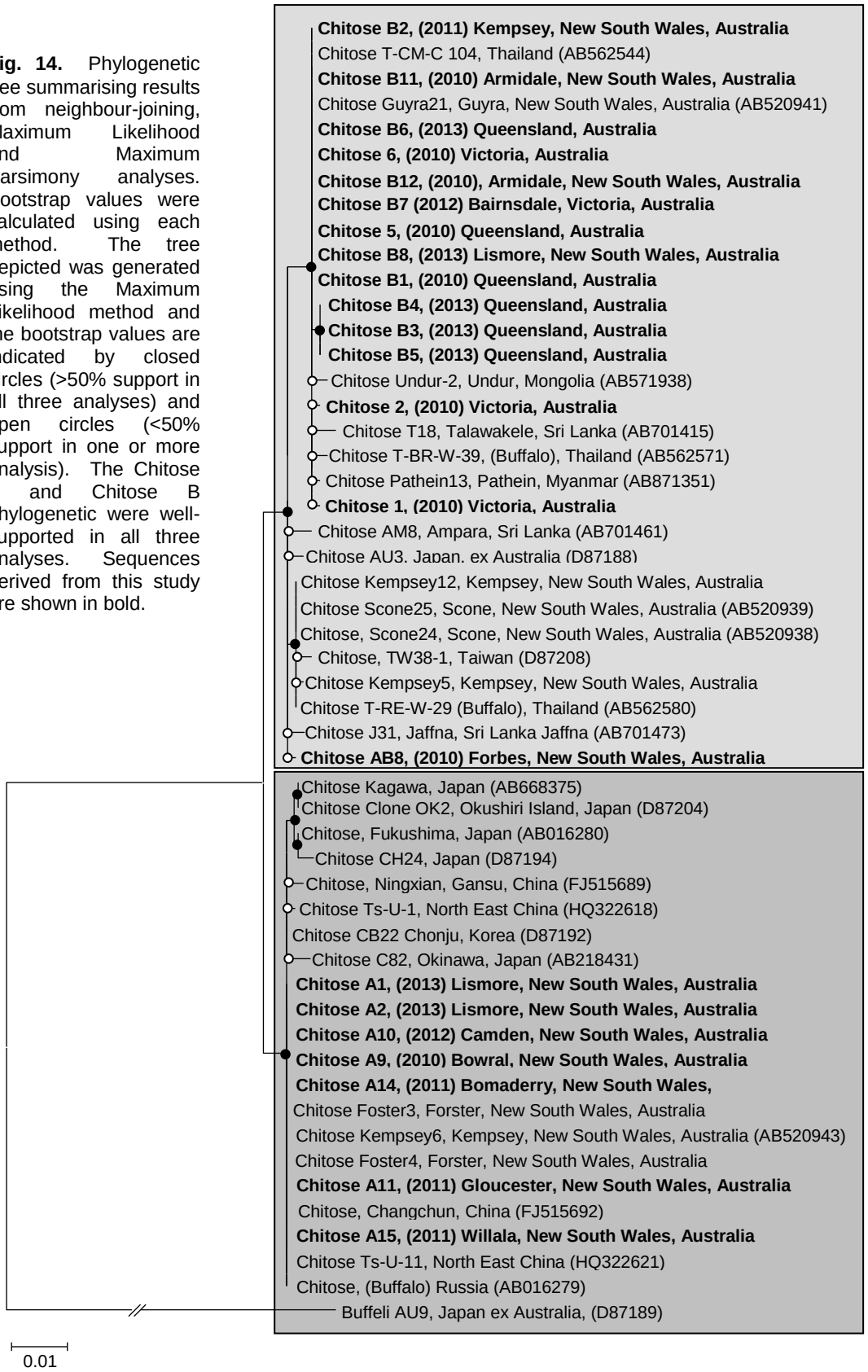


Fig. 14. Phylogenetic tree summarising results from neighbour-joining, Maximum Likelihood and Maximum Parsimony analyses. Bootstrap values were calculated using each method. The tree depicted was generated using the Maximum Likelihood method and the bootstrap values are indicated by closed circles (>50% support in all three analyses) and open circles (<50% support in one or more analysis). The Chitose A and Chitose B phylogenetic were well-supported in all three analyses. Sequences derived from this study are shown in bold.

4.4.2 Chitose A is strongly associated with both clinical disease and the Ikeda genotype

To test the association of the Chitose A phylogenetic type with clinical disease, a genotyping assay based on the method of Chen et al. (2014) aimed at discriminating Chitose A and B subpopulations, was used to examine 137 samples which tested positive for the Chitose genotype in prior studies (Eamens et al., 2013a; Eamens et al., 2013c). For this analysis the transformed fluorescence ratio (k′) and allele frequencies of the Chitose A genotype (mixed with Chitose B) were compared. Using predetermined standard plasmid ratios, transformed fluorescence ratio (k') showed a non-linear relationship with Chitose A allele frequency (R^2 = 0.998; Fig. 15). Allele frequencies determined from the blood samples are summarised in Table 17. Pearson's Chi-squared test indicated a strong association of clinical disease ($P < 0.0001$) and presence of Ikeda genotype ($P < 0.0001$) with samples that were predominantly Chitose A. Conversely, the same analysis indicated that high proportions (> 95%) of Chitose B were associated with the absence of clinical disease and negative detection of the Ikeda genotype. 100% of the samples with >95% of Chitose A allele were sourced from the state of NSW (Table 18). Additionally, the presence of Chitose A (> 95% + mixed) was found to have a significantly higher correlation with Ikeda than the presence of Chitose B ($P < 0.0001$). In addition, all samples sourced from Queensland contained >95% of the Chitose B type (Table 18).

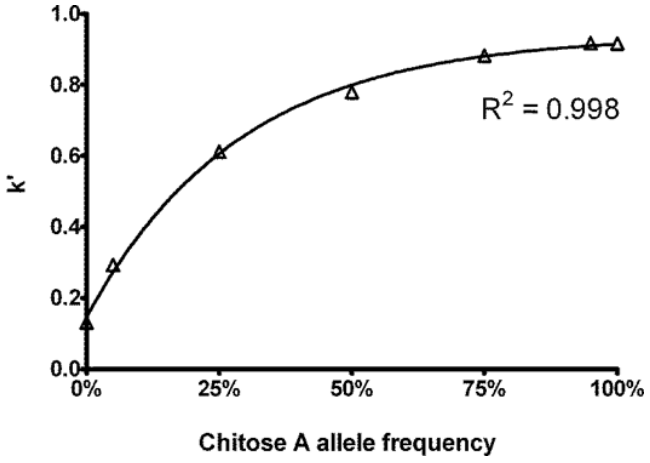


Fig. 15. Standard curve for the Chitose A/B genotyping assay generated using Chitose A/B plasmid mixes. The transformed fluorescence ratio (k′) was plotted against the allele frequencies of the Chitose A genotype and showed a non-linear relationship.

Table 17. Relative occurrence of Chitose A and B alleles in Chitose PCR positive samples and their association with clinical disease cases and the Ikeda genotype.

Allele frequency	% associated with clinical disease	% co-occurrence with Ikeda genotype
>95% Chitose A (n = 56)	100% (56)	98% (55)
Mix Chitose A/B (n = 52)	67% (35)	90.4% (47)
>95% Chitose B (n = 29)	27.6% (8)	44.8% (13)

Table 18. Relative occurrence of the Chitose A and B alleles in Chitose PCR positive samples by geographic origin

Allele frequency	Geographic origin by State			
	NSW	QLD	VIC	SA
>95% Chitose A (n = 56)	56	0	0	0
Mix Chitose A/B (n = 52)	45	0	3	2
>95% Chitose B (n = 29)	11	16	4	0
Total	112	16	7	2

4.4.3 Sequence analysis of the ITS region

The internal transcribed spacer regions ITS1 and ITS2, as well as the 5.8S rRNA gene, were all analysed initially for their potential to discriminate between *Theileria* strains within MPSP types. The ITS regions are commonly used for differentiating closely related strains due to the fact that intergenic spacers lack structural constraints and therefore evolve more rapidly. The ITS regions have been used previously to investigate polymorphism in *T. orientalis* types from Japan and the USA (Aktas et al. 2007). While sequence analysis of the ITS2 regions indicated that these sequences were discriminatory amongst MPSP types, there was no variation between ITS2 within individual MPSP types. As such, ITS2 did not provide any more information regarding genotype than already provided by the MPSP sequences and was not examined further.

Both the 5.8S rRNA gene and the ITS1 region contained sequence variation within MPSP types and specific primers were designed to amplify these regions. ITS1 was the most phylogenetically informative region and contained sufficient variation in sequence identity to discriminate between strains. Higher sequence identity was observed amongst Australian strains than between Australian and overseas strains (Table 19). In addition, the Ikeda and Chitose types were found to be considerably less variable in the ITS1 region relative to the Buffeli type which displayed identities as low as 72% within Australian strains alone (Table 19). The greater diversity in the Buffeli type ITS1 region is consistent with this strain diverging post-introduction to Australia (presumed to be >100 yr ago), while the lower level

of diversity within the Ikeda and Chitose ITS1 regions suggest that these strains have been much more recently introduced. Interestingly, some of the strains derived from Michigan, USA were identified in a prior study as Buffeli isolates (Aktas et al. 2007), however sequence comparisons show that these isolates are clearly more related to the Chitose type than the Buffeli type.

Table 19. The range in sequence identity in the ITS region observed within the Australian strains and within Australian and overseas strains.

Genotype	Range in % sequence identity within Australian strains	Range in % sequence identity within Australian and overseas strains
Buffeli	72-97%	N/A*
Chitose	98-100%	93-100%
Ikeda	95-98%	93-98%

*No sequences available in the database for comparison

The *T. orientalis* Ikeda (Shintoku) genome sequence (Hayashida et al., 2012) indicates that this genotype of the parasite contains 2 ITS1 regions, one copy on Chromosome 1 and one copy on Chromosome 3. Two copies of the ITS1 region were not found in the genome sequences of the Chitose and Buffeli genotypes sequenced in this study (Section 3.5), however because the sequence coverage for these genotypes was relatively lower, with high variability in the ITS1 region amongst genotypes, this may reflect poor assembly of these unique regions relative to the reference genome. Therefore, whether these genotypes possess two copies of the ITS1 region is uncertain. The ITS1 regions of the Ikeda genotype only, were therefore examined in more detail with respect to their chromosomal location and corresponding phylogeny. In order to separate out the disparate ITS1 genetic loci, the amplicons were cloned, sequenced and compared to the reference genome (Table 20). A phylogenetic tree was then constructed from the sequences. A phylogenetic tree showing the relationship between the Ikeda ITS1 regions derived from different samples is shown in Fig. 16. The sequences group according to their chromosomal location. Within each cluster, the Australian isolates are closely related to each other with the exception of some strains from QLD which appear to be more divergent.

Table 20. ITS1 Chromosomal loci sequenced from each sample in this study

Sample ID	Chromosome
13/051	Chromosome 1, 3
11/042	Chromosome 1 only
11/046	Chromosome 1, 3
11/123	Chromosome 1, 3
14/042	Chromosome 1, 3
14/054	Chromosome 1, 3
13/031	Chromosome 1, 3
15/067	Chromosome 1, 3
13/022	Chromosome 1, 3
14/003	Chromosome 1, 3
12/083	Chromosome 1, 3
15/062	Chromosome 3 only
14/012	Chromosome 1, 3
14/054	Chromosome 1, 3
15/056	Chromosome 1, 3

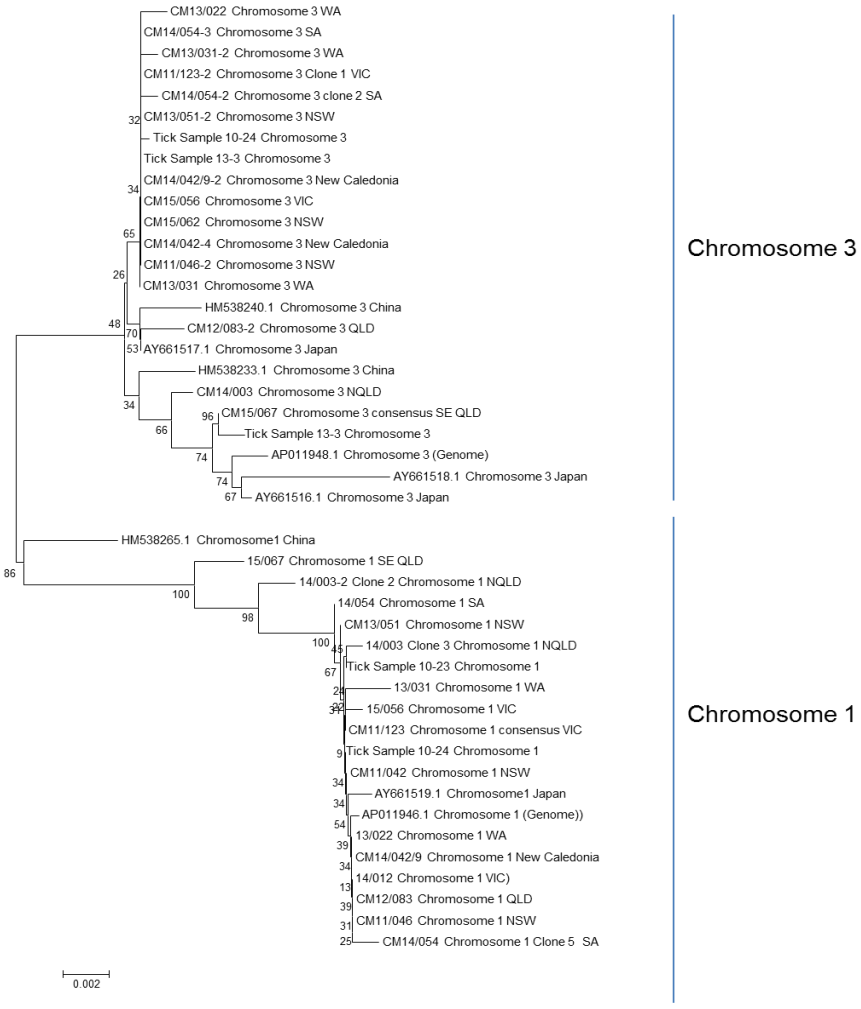


Fig 16. Phylogenetic tree of the Ikeda genotype ITS1 regions within Chromosomes 1 and 3. The sequences cluster according to their chromosomal location.

4.4.1 Temporal dynamics of genotype populations in *T. orientalis*-infected cattle

To investigate the temporal dynamics of *T. orientalis* populations in field-infected animals, we developed specific qPCR assays for the three major genotypes of this parasite found in Australia, namely the Ikeda, Chitose and Buffeli genotypes. Analytical sensitivity statistics and the linearity of standard curves from these three assays are outlined in Table 21 and Fig. 17.

Table 21. Analytical sensitivity (limit of detection, LOD) of the Ikeda- Chitose- and Buffeli-specific assays, 95% confidence limits are shown in brackets.

qPCR Parameter	Ikeda	Chitose	Buffeli
R^2	0.997	0.999	0.994
Efficiency (%)	91.2 (87.9 - 94.9)	95.1 (92.9 - 97.4)	90.7 (86.4 - 95.3)
LOD (GC/µL)	8.4 (3.1 - 23)	35 (14 - 88)	19 (8.5 - 44)

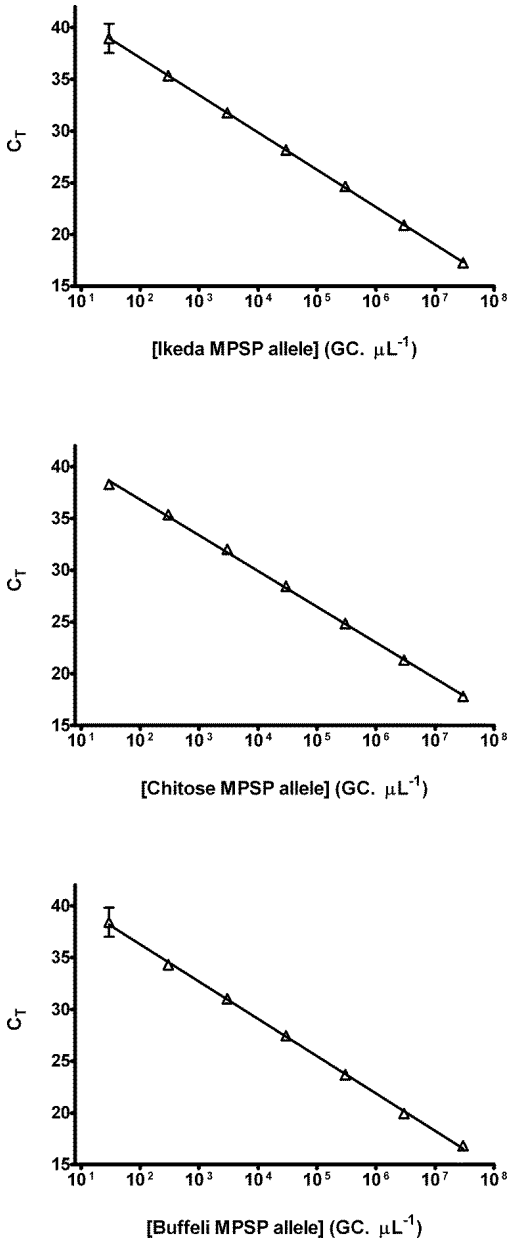


Fig. 17. Standard curves for the Ikeda (top), Chitose (middle) and Buffeli (bottom) quantitative PCR assays. Cycle threshold (C_T) is plotted against MPSP allele concentration and for all assays demonstrated a linear relationship.

The specificity of primer-probe sets for each intended target was confirmed using plasmids containing Ikeda, Chitose (A+B) and Buffeli MPSP gene sequences at 100 pg/µL (Fig. 18). In addition, 10 samples collected from cattle cPCR-negative for *T. orientalis* or that were infected with either *B. bovis*, *B. bigemina* or *A. centrale* were all negative in each genotype-specific assay. The performance of each assay was further assessed on 15 well-characterised blood samples from *T. orientalis*-positive animals and yielded comparable results to conventional PCR (Eamens et al., 2013c).

The Ikeda, Chitose and Buffeli qPCR assays were applied to DNA extracts from 10 naïve cattle that had been introduced onto a property with a recent history of clinical theileriosis. The 10 animals became infected with all three genotypes over the study period, with the Ikeda allele the first to be detected in the majority of animals (7/10 cows positive at Day 11 and the remaining three cows positive at Day 20). The Chitose allele was also detectable in Cow 8 at 11 days post-introduction, but was not detected in the other nine animals until Day 20. Buffeli allele was detected in nine animals at Day 20 and the remaining animal at Day 34. Together these data suggest that the Ikeda genotype may be either favoured during vector transmission, or out-competes the Chitose and Buffeli genotypes. Interestingly, in all animals the dominant genotype, as detected by qPCR, shifted over the course of the study, from the Ikeda genotype to the Chitose genotype, with the Ikeda genotype re-emerging as the dominant allele around Day 60 in most cases (Fig. 19). Quantitative data for each genotype revealed that the levels of Ikeda allele peaked 35-40 days post-introduction in the majority of animals (Fig. 19). In most animals, the concentration of Chitose allele peaked only after the initial peak in the Ikeda genotype, from 40-55 days post-introduction. These peaks coincided with the peak in parasitaemia as determined by blood film, which in most animals was at Day 40 (Fig. 19). The Buffeli genotype, while detected in all animals over the course of the study, was usually detected at concentrations an order of magnitude lower than the Ikeda and Chitose alleles. One exception was Cow 1, in which the Buffeli genotype reached a higher concentration than the Ikeda and Chitose genotypes. Despite this, the peak in parasite DNA in Cow 1 was 1-2 orders of magnitude lower than in the other nine animals. Furthermore, Cow 1 was the only animal that did not become anaemic during the course of the study. Eight of the remaining nine animals first became anaemic approximately eight days following the peak in concentration of the Ikeda genotype, and around the time when the Chitose genotype was peaking, the exception being Cow 10 which first became anaemic at Day 34, coinciding with the peak concentration of the Ikeda allele. These fluctuations in parasite load should also be considered in reference to the level of anaemia i.e. an increase in the GC/µL in an anaemic animal equates to an increase in the GC/ µL PE.

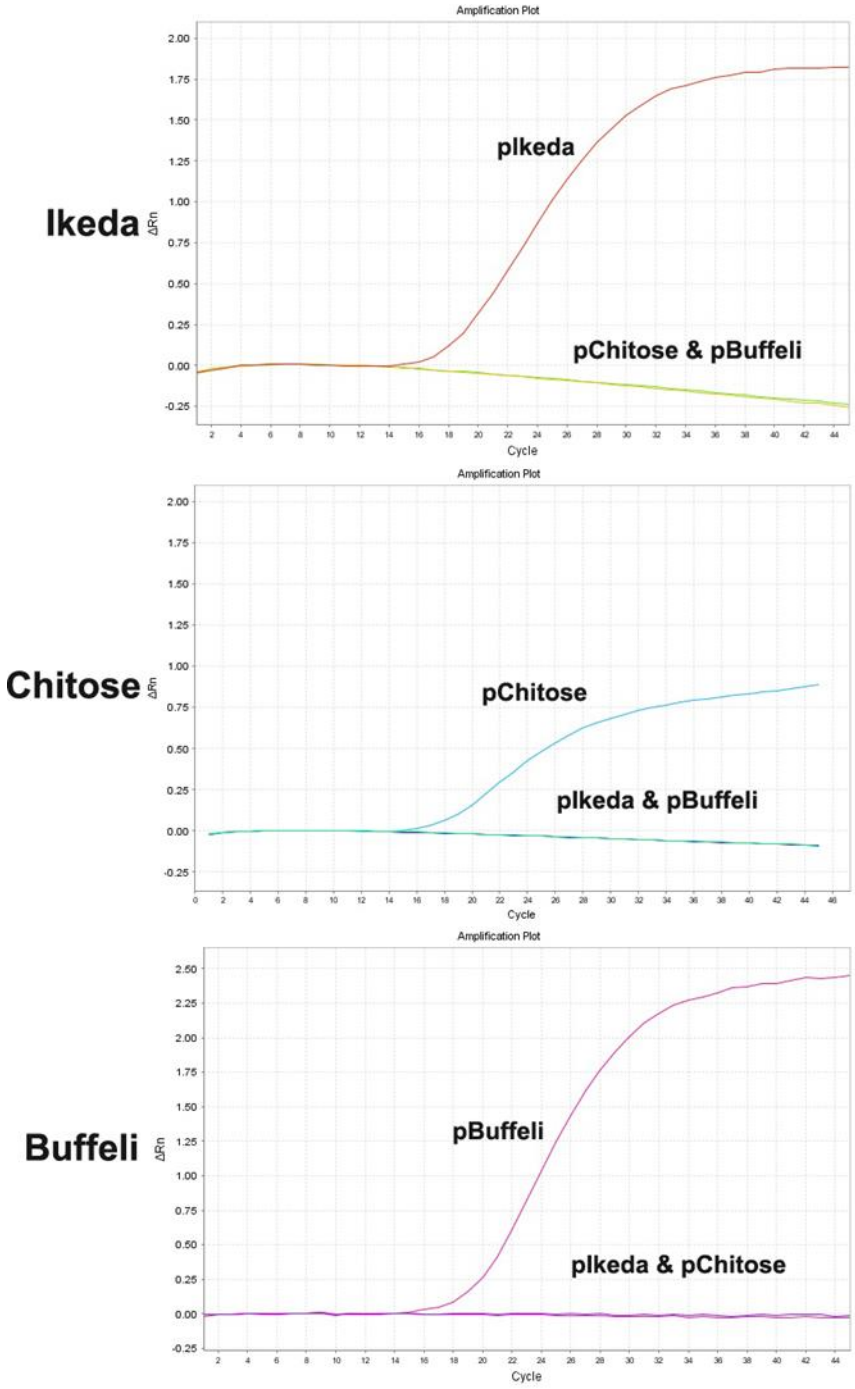


Fig. 18. Specificity of the Ikeda (top), Chitose (middle) and Buffeli (bottom) quantitative PCR assays as demonstrated using 100 pg/µL non-target plasmid combinations (ie: pChitose + pBuffeli, pIkeda + pBuffeli or pIkeda + pChitose respectively).

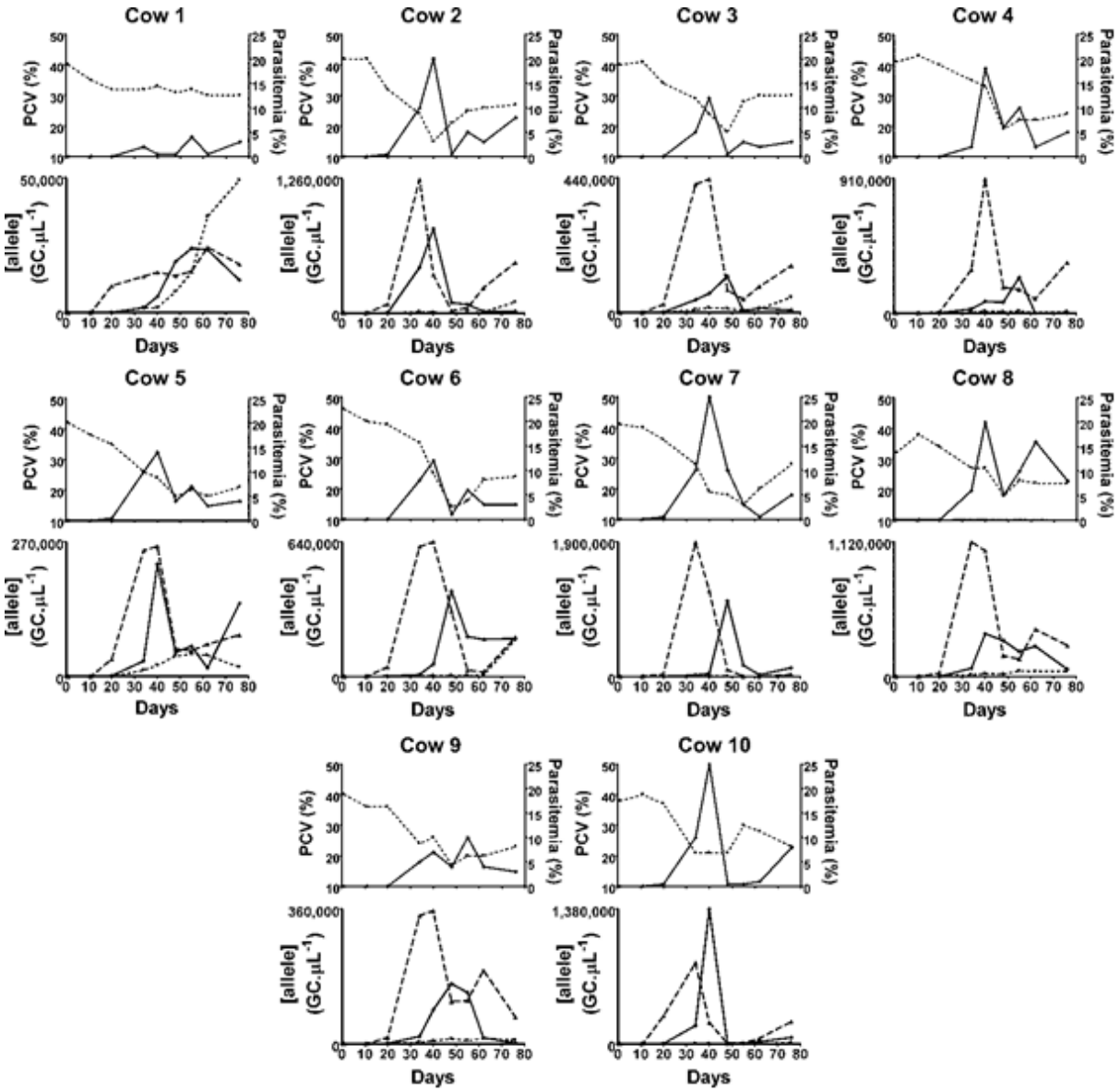


Fig. 19. Temporal population dynamics of three *T. orientalis* subtypes in 10 field-infected cattle. For each cow, Top: PCV (dotted line) and parasitaemia as determined by blood film analysis (solid line). Bottom: Allele concentrations determined by qPCR targeting the Ikeda (dashed line), Chitose (solid line) and Buffeli (dotted line) genotypes. Note that y-axes are not equivalent for each cow.

The Chitose subpopulation genotyping assay was applied to samples from the time course study where the Chitose genotype was at its peak. Of the ten samples, six were predominantly (>95%) Chitose A allele (Cows 2, 4, 5, 6, 7 and 8) and the remaining four samples were a mix of Chitose A and B. Of the samples containing a mix of Chitose A and B, Cows 9 and 10 contained a much higher proportion of Chitose A (75% and 86% respectively), while Cows 1 and 3 contained closer to an equimolar mix of the two alleles (43% and 52% Chitose A respectively). Quantitative data from the Chitose qPCR assay also demonstrated that Cows 1 and 3 had the lowest overall concentration of the Chitose allele when this genotype was at its peak.

4.5 Genome sequencing of *T. orientalis* Ikeda, Chitose and Buffeli genotypes

4.5.1 Genome sequencing of *T. orientalis* single genotype populations

For each isolate, reads from three technical replicate sequencing experiments were aligned to the Shintoku reference genome (Assembly ASM74089v1) and merged to generate a single alignment. The Shintoku reference sequence is of the Ikeda genotype and hence the Australian Ikeda genotype showed high proportions of reads mapped, reference sequence coverage and depth of sequence coverage (Table 22). Chitose and Buffeli isolates showed reduced percentages of mapped reads and reference coverage indicating a high amount of sequence variance from Ikeda, and include large insertion and deletion events (Fig. 20). Bos taurus DNA was detected at low levels and represented less than 1.35% of reads for all isolates. Assemblies of Ikeda, Chitose and Buffeli genome sequences produced varied results with the Ikeda assembly producing longer and fewer total contigs. In contrast, assemblies of the Chitose and Buffeli isolates were more fragmented (Table 22).

4.5.2 Variation analysis

Sexual reproduction causes apicomplexan parasites to exist as variable populations of individuals with the potential for a wide range of haplotypes in a single host. In this study, variant analysis of the Ikeda, Chitose and Buffeli isolates was limited to SNP mutations due to reported difficulties in the analysis of indel mutations using short read alignment methods (Koboldt et al., 2012). To assess isolate variation, analysis was performed using two methods: 1. by aligning reads to the Shintoku reference sequence and calling homozygous SNPs (i.e. between population variation; Table 23) and 2. Aligning reads to genome assemblies and calling heterozygous SNPs (i.e. within population variation; Table 24). To examine the effectiveness of these two methodologies in detecting SNP variants we examined representative sections from each genome by Sanger sequencing (Table 25). Despite difficulties identifying low allele frequency SNPs with Sanger sequencing, 86.8% (217/250) of within-population variants detected by Illumina sequencing showed evidence of dual-base peaks in Sanger chromatograms. For homozygous SNPs, very high sensitivity

Table 22: Genome assembly and resequencing statistics

	Ikeda	Chitose	Buffeli
Resequencing			
% reads mapped	96.9 %	75.2 %	78.0 %
% host DNA	0.072 %	1.34 %	0.069 %
Genome % ref. coverage (≥ 14×)	93.1 %	74.4 %	65.0 %
CDS % ref. coverage (≥ 14×)	96.5%	85.3%	78.3%
Mean coverage depth	60.5 ×	43.5 ×	47.3 ×
Assembly			
# scaffolds	639	1557	6043
N50 (median scaffold length)	54695	12243	2646
% GC	41.7 %	39.3 %	39.8 %

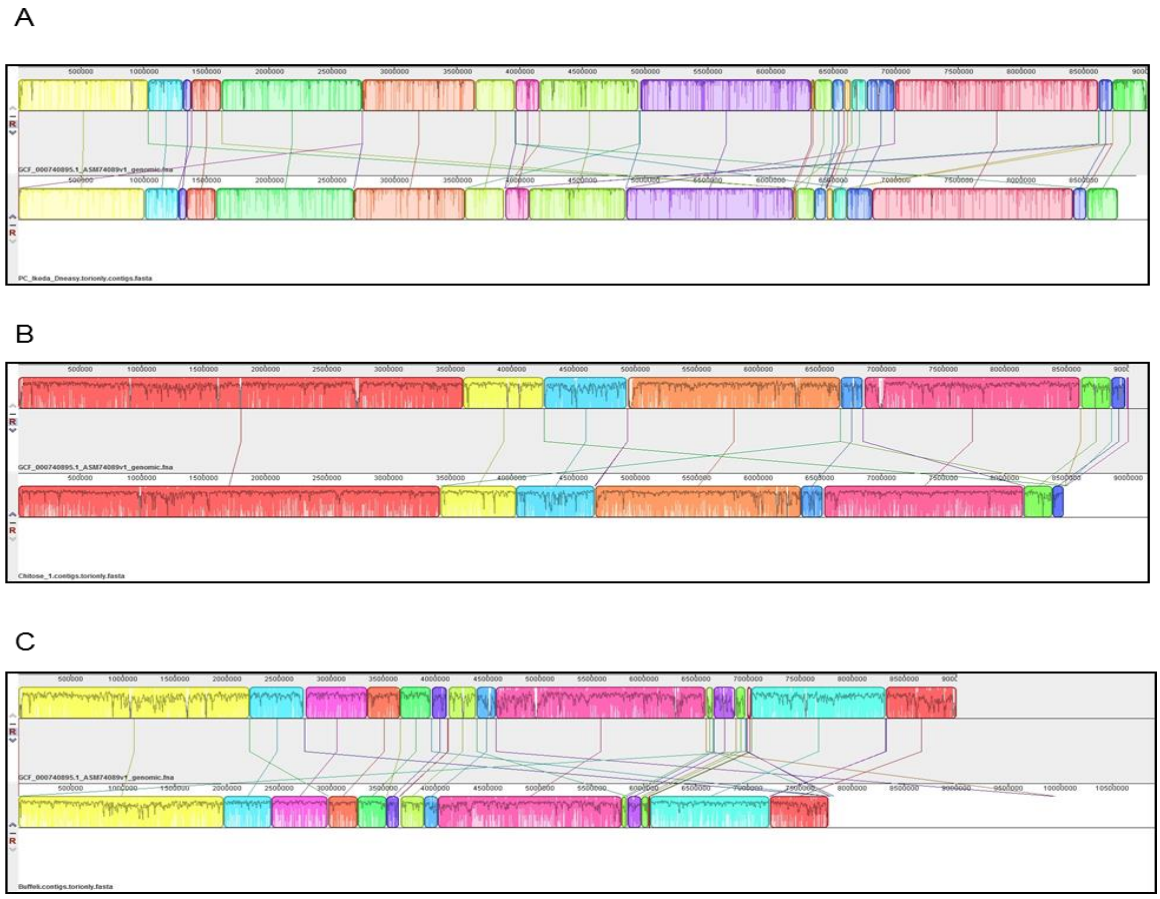


Fig. 20. Diagrammatic representation of the *T. orientalis* genomes derived from genotypes Ikeda (A), Chitose (B) and Buffeli (C) (bottom panels) aligned with the *T. orientalis* Shintoku reference genome (top panels), generated in Mauve. Coloured blocks represent “localised co-linear blocks”, which are conserved segments that appear to be internally free from genome rearrangements (i.e.: recombination events) relative to the reference genome.

Table 23. Between population SNP variance of Australian Ikeda, Chitose and Buffeli strains

Between-population SNPs	Ikeda	Chitose	Buffeli
Total	24132	788412	676284
Within CDS	16974	637070	565947

Table 24. Within population SNP variance of Australian Ikeda, Chitose and Buffeli strains

Within-population SNPs	Ikeda	Chitose	Buffeli
Total	1669	4634	27947
Within CDS	1144	2086	16661

Table 25. Validation of SNP variant calling pipeline

Strain	Next generation sequencing SNP call/Sanger SNP call				Sensitivity
	+/+	+/-	-/+	-/-	
Ikeda	116	0	1	6953	0.991
Chitose	515	0	64	3652	0.889
Buffeli	399	0	72	2852	0.847

values were observed in the Ikeda isolate. No false positive calls (Illumina positive/Sanger negative) were observed in any isolate, while false negative calls were very low in Ikeda but increased in Chitose and Buffeli isolates. When these calls were examined in depth it was found that all were closely associated with observed small insertion or deletion events which have been previously reported to be problematic in SNP calling pipelines (Koboldt et al., 2012).

When comparing between-population variation, the Ikeda isolate showed expectedly high similarity to the Shintoku reference sequence, while Chitose and Buffeli isolates showed similar levels of divergence (Table 23). Total numbers of variants were similar in the Chitose and Buffeli isolates, however, these variants were largely found in differing positions reflecting the divergence between all three *T. orientalis* isolates. A high proportion of between population variants were observed within coding sequences although sequencing

coverage (and hence variant calling ability) was significantly higher in coding regions (Table 22).

The three isolates also showed large differences in within-population variation (Table 24). The number of Ikeda within-population variant positions observed was considerably less than that of Chitose (~3 × higher) and Buffeli isolates (~17 × higher), with much of this variation occurring within predicted coding sequences. When the distribution of variant allele frequencies was examined (Fig. 21) the Ikeda isolate showed a sharp peak at 10-15%, indicating a mixture of fewer haplotypes while the Chitose and Buffeli isolates showed much broader distributions indicating more complex and varied populations.

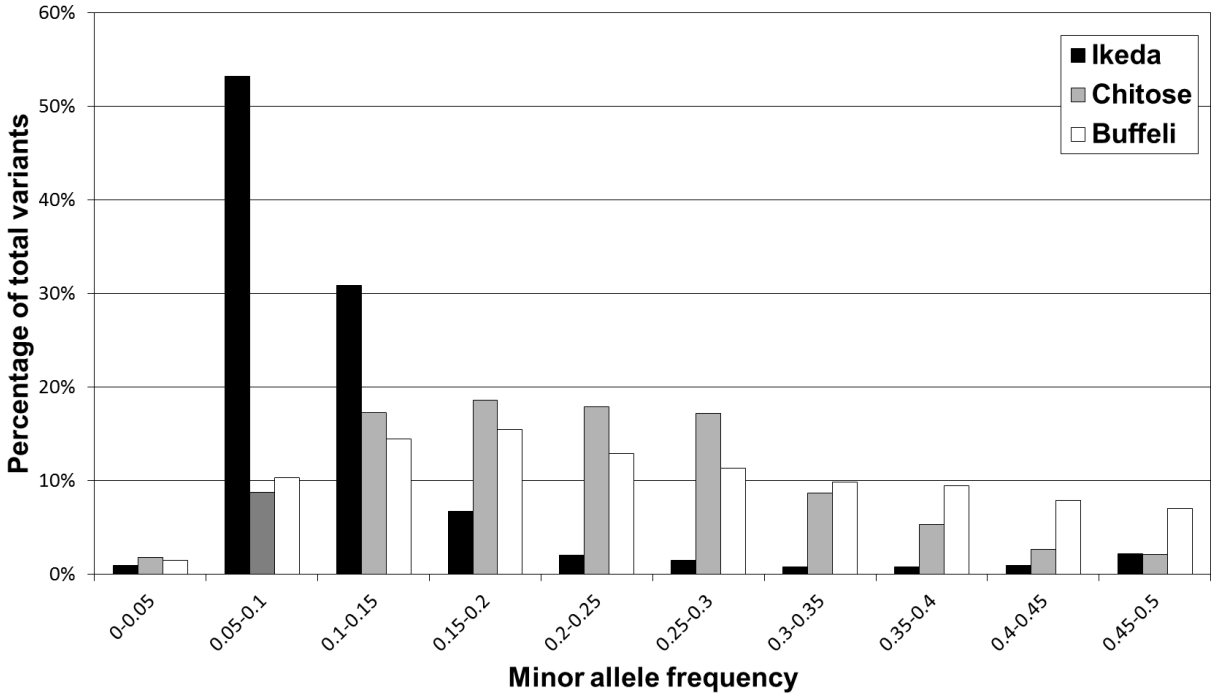


Fig. 21. Distribution of within-host variants observed in strains Ikeda (black), Chitose (grey) and Buffeli (white). Minor allele frequencies at within-population variant positions are sorted into 5% bins (x-axis) and are shown against the percentage of total variants (y-axis).

4.5.3 Mitochondrial genomes of Ikeda, Chitose and Buffeli isolates

Theileria mitochondrial genomes have been previously observed to be linear and relatively small at approximately 6 kbp in size (Hikosaka et al., 2010). In *T. orientalis* isolates that were examined in this study we were able to confirm a linear structure with inverted PCR utilizing outward facing primers at the edge each mitochondrial genome (Fig. 22). Furthermore, some *Babesia* mitochondrial genomes have been described to undergo multiple inversions (Hikosaka et al., 2012). To determine if this occurs in *T. orientalis* mitochondrial genomes we examined sequencing reads mapped to assembled Ikeda, Chitose and Buffeli mitochondrial genomes for evidence of split-reads, i.e. reads partially mapped to the contig and soft-clipped. We found no evidence of soft-clipped reads indicative of an inversion event in any of the three *T. orientalis* mitochondrial sequences. Mitochondrial genomes of Ikeda, Chitose and Buffeli isolates for each for each *T. orientalis* genotype are shown in Fig. 22.

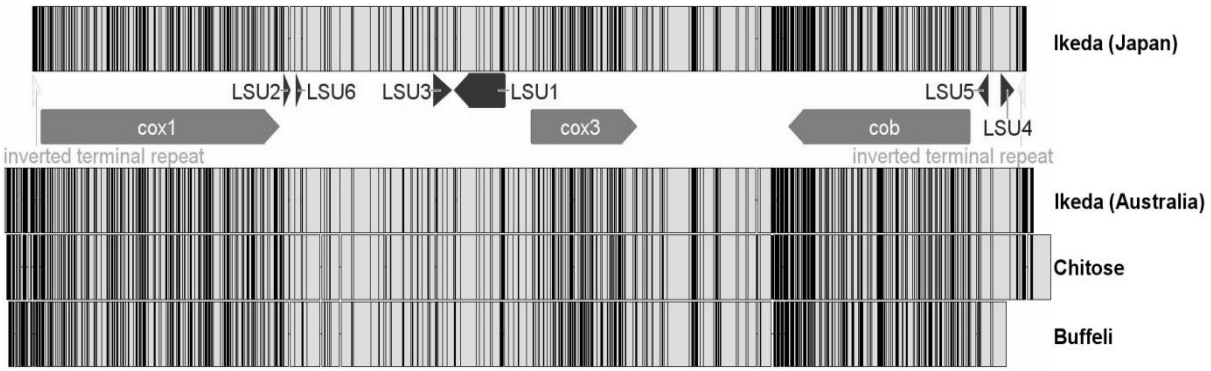


Fig. 22. The structure of the mitochondrial genomes of *T. orientalis* Ikeda, Chitose and Buffeli is linear with inverted terminal repeats at either end of the genomes. The locations of the cytochrome oxidase I (CoxI), cytochrome oxidase III (CoxIII), cytochrome oxidase b (Cob) and large subunit rRNA genes (LSU) are shown.

4.5.4 dN/dS analysis

Analysis of the dN/dS ratio has been previously used to identify surface antigens that could be potentially used as subunit vaccine components (Endo et al., 1996; Hayashida et al., 2013). In this study, we examined the dN/dS ratio calculated using reconstructed genomes generated from the Shintoku reference sequence and homozygous SNPs from the Chitose and Buffeli genome sequences. The Ikeda isolate was determined to be too closely related to the reference sequence (3377/4002 genes with < 5 mutations) and was excluded from this analysis. The dN/dS ratios of 3577/4002 and 3230/4002 genes were examined from reconstructed genomes of the Chitose and Buffeli isolates, respectively. High dN/dS predicted surface proteins (SignalP and TMHMM positive) are shown in Table 26. These dN/dS values compare well with similar studies in other *Theileria* species (Hayashida et al., 2013).

4.5.5 MLST phylogeny of Ikeda, Chitose and Buffeli and their place within the Piroplasmida

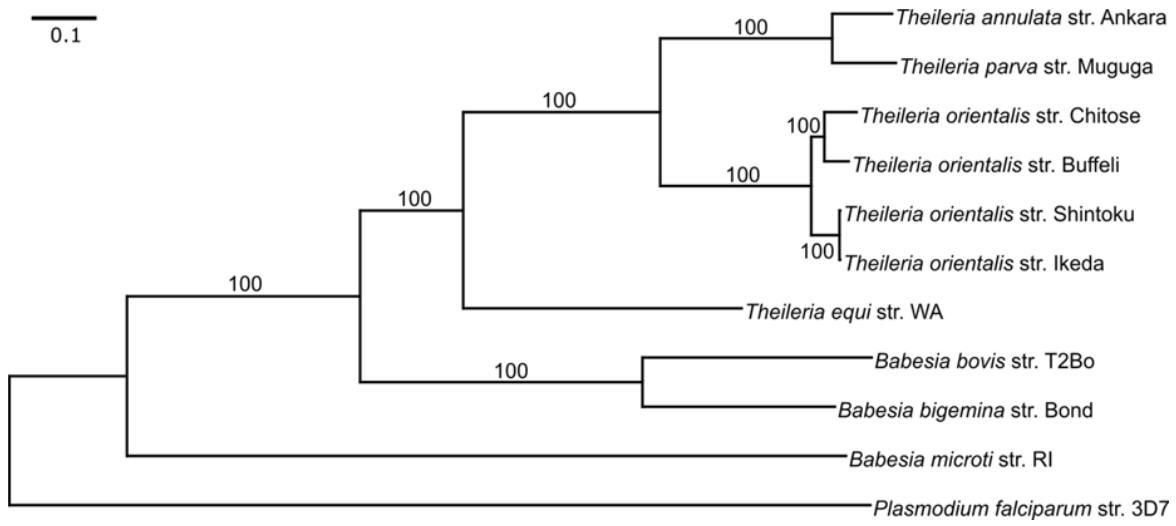
To assess if whole genome sequences could be used to further refine the taxonomy of *T. orientalis* we explored two commonly used methods of species definition, phylogeny and average nucleotide identity, using *T. orientalis* whole genome assemblies and representative whole genome sequences of the Order Piroplasmida (Fig. 23A). Both whole genome phylogeny and average nucleotide identity reveal a very close relationship between Shintoku and Ikeda genomes of *T. orientalis*, conclusively identifying that the origin of the recent Australian Ikeda outbreak was introduction from Eastern Asia, most likely Japan. In contrast, average nucleotide identities between *T. orientalis* isolates Ikeda/Shintoku, Chitose and Buffeli are low for organisms considered to be of the same species and compare well with those observed between *T. parva* and *T. annulata* (Fig. 23B).

Table 26. Predicted surface proteins with High dN/dS

Protein ID	Functional classification	Homolog in other Piroplasmida	Protein family ID (Pfam)	Mean dN/dS
XP_009690939.1	hypothetical protein	N		0.420013
XP_009691911.1	uncharacterized protein	Y	Pf04385	0.348927
XP_009690040.1	hypothetical protein	Y		0.329911
XP_009690607.1	hypothetical protein	Y		0.326084
XP_009691340.1	hypothetical protein	Y		0.325941
XP_009689372.1	hypothetical protein	Y		0.322271
XP_009689430.1	hypothetical protein	N		0.310138
XP_009690004.1	hypothetical protein	Y		0.307715
XP_009690344.1	ToLocg1 paralog	Y		0.272704
XP_009690269.1	hypothetical protein	Y		0.236663
XP_009691913.1	hypothetical protein	N		0.230252
XP_009689555.1	hypothetical protein	Y		0.228835
XP_009689383.1	hypothetical protein	N		0.225034
XP_009688884.1	CD8+ T cell target antigen Tp2	Y		0.222906
XP_009689185.1	hypothetical protein	Y		0.222693
XP_009692799.1	hypothetical protein	N	Pf04385	0.220597
XP_009692694.1	hypothetical protein	Y		0.217668
XP_009690569.1	brain protein 44-like	Y	Pf03650	0.214623
XP_009689733.1	hypothetical protein	Y		0.211825
XP_009690910.1	hypothetical protein	N		0.210374
XP_009689120.1	hypothetical protein	Y		0.198483
XP_009692522.1	thrombospondin-related anonymous protein	Y	Pf00092	0.196277
XP_009691139.1	putative protease	Y	Pf02517	0.1867
XP_009690803.1	putative apicoplast import protein	Y	Pf16166	0.176391
XP_009692438.1	hypothetical protein	Y		0.17456
XP_009692777.1	uncharacterized protein	Y	Pf04385	0.173405
XP_009692412.1	hypothetical protein	Y		0.16334
XP_009692816.1	surface protein	Y	Pf04145	0.160208
XP_009689868.1	hypothetical protein	Y		0.156548
XP_009692665.1	hypothetical protein	N		0.154705
XP_009689845.1	major piroplasm surface protein	Y	Pf02488	0.146066
XP_009690016.1	hypothetical protein	Y	Pf04385	0.136429

Protein ID	Functional classification	Homolog in other Piroplasmida	Protein family ID (Pfam)	Mean dN/dS
XP_009692660.1	putative lysophospholipase	Y	Pf12146	0.129336
XP_009692759.1	hypothetical protein	Y		0.127301
XP_009691480.1	putative ER oxidoreductin	Y	Pf04137	0.120148
XP_009689697.1	uncharacterized protein	Y		0.116339
XP_009691437.1	hypothetical protein	Y		0.113952
XP_009688850.1	hypothetical protein	Y		0.113259
XP_009690130.1	hypothetical protein	Y	Pf08320	0.1113
XP_009692225.1	uncharacterized protein	Y	Pf07691	0.107945
XP_009690580.1	p23 surface protein	Y		0.104459
XP_009691349.1	zinc transport protein	Y	Pf02535	0.10384
XP_009689509.1	hypothetical protein	Y		0.10243
XP_009689195.1	hypothetical protein	Y		0.102029
XP_009691696.1	hypothetical protein	Y		0.10029
XP_009689770.1	uncharacterized protein	Y	Pf04385	0.099851
XP_009690399.1	50S ribosomal protein L9	Y	Pf01281	0.096548
XP_009689290.1	hypothetical protein	Y	Pf05450	0.095706
XP_009689754.1	uncharacterized protein	Y		0.086991
XP_009688894.1	50S ribosomal protein L33	Y		0.0863757

A



B

T. orientalis **Ikeda**	***T. orientalis*** **Chitose**	***T. orientalis*** **Buffeli**	***T. parva*** **Muguga**	***T. annulata*** **Ankara**	***T. equi*** **WA**	
99.2%	82.6%	82.6%	66.1%	66.0%	64.9%	*T. orientalis* Shintoku
	82.8%	83.1%	66.1%	66.1%	64.5%	*T. orientalis* Ikeda
		86.5%	66.3%	66.2%	64.9%	*T. orientalis* Chitose
			67.0%	66.9%	65.5%	*T. orientalis* Buffeli
				79.5%	65.8%	*T. parva* Muguga
					65.5%	*T. annulata* Ankara

Fig. 23. Phylogenomic analysis of members of the Piroplasmida (A). *T. orientalis* Chitose and Buffeli are more closely related to each other than *T. orientalis* Ikeda. *T. orientalis* Ikeda is closely related to the Japanese Shintoku strain. Average nucleotide identities between *T. orientalis* genotypes are only slightly higher than those observed for *T. parva* and *T. annulata* which are separate species.

4.6 Proteomic analysis of the piroplasm phase of *T. orientalis*

4.6.1 Electrophoretic analysis of purified *T. orientalis* proteins

The *T. orientalis* proteins that had been separated into aqueous and detergent-soluble fractions were electrophoresed on a 1D SDS-PAGE gel to separate out individual proteins. Aqueous and detergent fractions for each *T. orientalis* genotype are shown in Fig. 24.

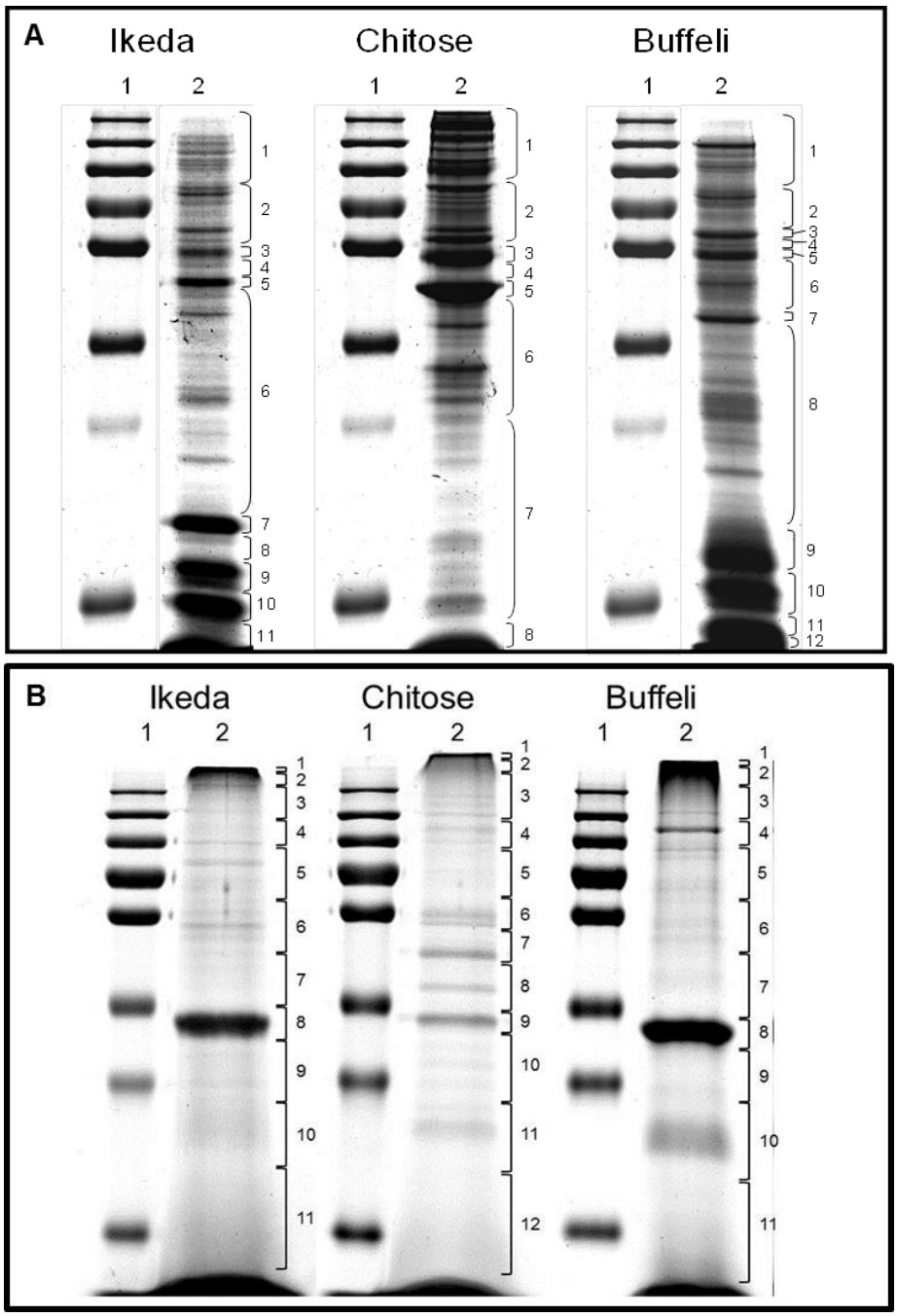


Fig. 24. One-dimensional SDS-PAGE gel showing purified proteins from aqueous (A) and detergent (B) fractions of *T. orientalis* Ikeda, Chitose and Buffeli genotypes. Lane 1: Molecular weight marker; Lane 2: purified *T. orientalis* proteins. Gel slices analysed individually by LC MS/MS are marked on the right of Lane 2 for each genotype.

4.6.2 LC-MS identification of purified *T. orientalis* proteins

From the aqueous phase, eight, eleven and twelve individual gel slices were analysed by LC-MS for the Chitose, Ikeda and Buffeli genotypes respectively. When examining the detergent phase, twelve slices for Chitose and eleven for both Ikeda and Buffeli were analysed (highlighted in Fig. 24). Spectra for individual gel slices were then converted into a peak list and peptide mass fingerprints were generated. The peptide mass fingerprints that mapped to *Theileria* proteins using Mascot are summarised in Table 27. In total, 404, 216 and 415 proteins were identified in the both aqueous and detergent fractions of the Ikeda, Chitose and Buffeli genotypes respectively. These numbers exclude contaminating bovine proteins which represented <5% of the identified proteins in each sample. A full list of proteins identified from each genotype is given in Appendices A-C.

Table 27. Proteins identified via mass spectrometry in *T. orientalis* Ikeda, Chitose and Buffeli genotypes.

Function	**No. proteins identified**		
	Ikeda	Chitose	Buffeli
Metabolism	46	20	43
DNA replication and repair	7	1	12
Cell division	15	22	10
Transcription, regulation and signalling	32	9	19
Translation	65	14	74
Protein processing and turnover	32	14	57
Intracellular transport	23	10	19
Membrane transport	4	0	4
Homeostasis	6	1	11
Virulence	17	11	8
Uncharacterised	157	114	158
Total	**404**	**216**	**415**

The number of unique protein hits identified per gel slice for each genotype are detailed in Table 28. Slice 5 from the aqueous phases of the Ikeda and Chitose genotypes, and slice 6 from the Buffeli genotype contained the ubiquitous protein, actin, which is important in cell division. Actin was also ubiquitous in the detergent fractions.

4.6.3 *In silico* analysis of *T. orientalis* proteins

Of the 3995 predicted open reading frames in the *T. orientalis* Ikeda genome, 771 proteins were predicted to contain transmembrane domains. The distribution of these proteins across the 4 chromosomes and the organellar genomes is listed in Table 29. Chromosome 1 contained the highest number of open reading frames predicted to encode transmembrane proteins. The apicoplast was not predicted to encode any transmembrane proteins. The number of predicted transmembrane helices within individual proteins ranged from 1 to 22.

Table 28. Number of *T. orientalis* proteins identified per gel slice.

Slice number	*Aqueous*			*Detergent*		
	Ikeda	Chitose	Buffeli	Ikeda	Chitose	Buffeli
1	59	35	52	27	9	12
2	54	25	58	17	8	4
3	30	31	30	17	4	2
4	26	24	29	13	2	13
5	52	42	37	17	8	7
6	N/A	34	48	28	10	1
7	34	31	25	21	6	5
8	23	18	102	12	3	13
9	29	N/A	30	29	6	10
10	31	N/A	32	18	6	6
11	24	N/A	29	12	5	3
12	N/A	N/A	24	N/A	1	N/A

4.7 Development of a serological test for *T. orientalis*

4.7.2 Immunological cross-reactivity between genotype MPSPs

While the MPSP gene is the most commonly used marker to discriminate between the various *T. orientalis* genotypes, the proteins encoded by these genes display a high degree of sequence conservation. Studies using native and recombinant MPSP proteins have shown conflicting results with respect to serological cross-reactivity between genotype

Table 29. Open reading frames from the *T. orientalis* Ikeda genome predicted to encode transmembrane proteins.

Genetic location	No. transmembrane proteins identified
Chromosome 1	215
Chromosome 2	206
Chromosome 3	177
Chromosome 4	172
Apicoplast genome	0
Mitochondrial genome	1
Total	**771**

MPSPs, with native (Kawazu et al., 1992a) but not recombinant (Iwasaki et al., 1998; Kawazu et al., 1992b) antigen being serologically distinct amongst genotypes. Nonetheless, recombinant MPSPs can be discriminated using monoclonal antibodies against specific epitopes (Iwasaki et al., 1998), while post-translational modifications are believed to be responsible for serological discrimination amongst native MPSPs (Kawazu et al., 1992b). Western blotting of recombinant MPSP antigen from the Ikeda, Chitose and Buffeli genotypes of *T. orientalis* was undertaken to determine whether there was serological cross-reactivity between the genotype MPSPs generated in this study. Sera from animals infected with individual genotypes of the parasite (Ikeda, Chitose or Buffeli as determined by qPCR) displayed cross-reactivity with all three MPSP types (representative blots are shown in Fig. 25). However, some variability in the reactivity of serum in detecting their specific targets was observed based on the intensity of the bands detected (Fig. 25); therefore the MPSP antigens were pooled for subsequent use in the ELISA

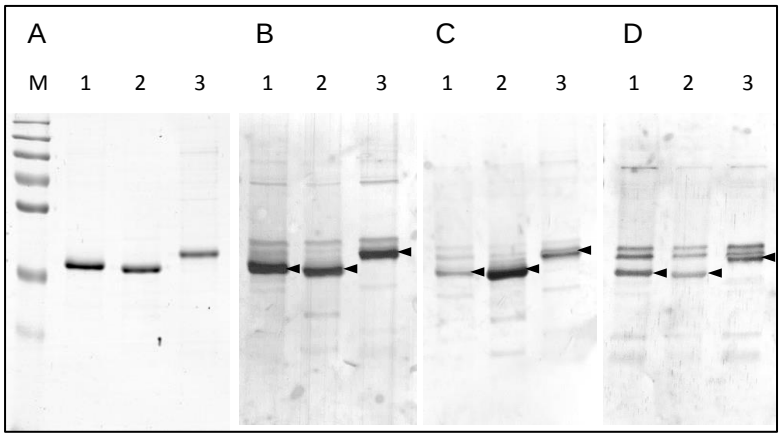


Fig. 25. Western blot analysis of recombinant MPSP antigens from the Ikeda (Lane 1), Chitose (Lane 2) and Buffeli (Lane 3) genotypes of *T. orientalis*. One microgram of each recombinant protein was run on an 1D SDS-PAGE gel and stained (A) or blotted against sera from animals infected with the Ikeda (B), Chitose (C) and Buffeli (D) genotypes of *T. orientalis*. Immunoreactive MPSP proteins are indicated with arrowheads. M = molecular weight marker (sizes 250, 130, 100, 75, 55, 35 and 25 kDa).

4.7.3 Specificity and sensitivity of the MPSP ELISA

The specificity and sensitivity of the MPSP ELISA compared to blood film examination and qPCR was assessed using 376 pairs of EDTA bloods and sera for which all three assays had been conducted. PCR assays are considered a gold standard for *T. orientalis* detection, and consistent with this, the UIC qPCR detected the highest number of positive samples (262/376). Using qPCR as a gold standard for comparison, the MPSP ELISA had a sensitivity of 68.7% and a specificity of 99.1%, while blood film examination had a sensitivity of 62.2% and a specificity of 97.4%. Animals that were PCR positive but ELISA negative either may have not yet seroconverted to the MPSP antigen or did not elicit a strong enough humoral response to enable detection. Thus, while ELISA was found to be more sensitive and specific for *T. orientalis* detection than blood film, as previously demonstrated for *T. orientalis* detection in buffalo (Wang et al., 2010b), PCR-based methods remain the most sensitive option for parasite detection.

4.7.4 MPSP seropositivity is associated with anaemia

Packed cell volume (PCV) was used as a measure of clinical status and was compared to the corresponding serological data from individual animals. Of the animals with a PCV in the normal range (PCV ≥ 24), 55% were seronegative, while the majority of animals (89%) with anaemia were MPSP seropositive (Table 30). Fisher's exact test indicated that the association between MPSP seropositivity and anaemia was significant (P<0.0001). When the antibody ELISA ratios of individual animals were plotted against the corresponding PCVs a statistically significant moderate and negative correlation (r = -0.5) was observed (Fig. 26).

Table 30. Frequency of anaemic versus normal animals in relation to MPSP serological status

	Anaemic (n=36)	**Normal (n=220)**
MPSP seropositive	32 (89%)	98 (45%)
MPSP seronegative	4 (11%)	122 (55%)

P<0.0001

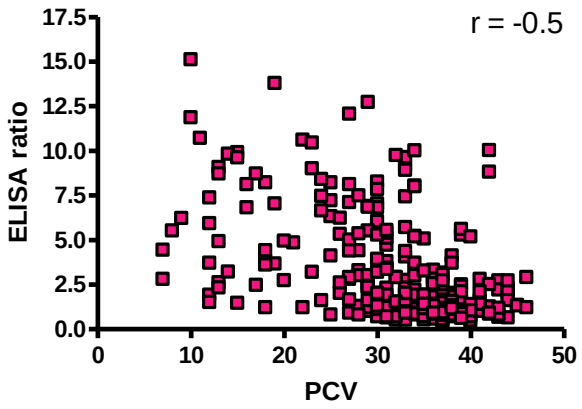


Fig. 26. The ELISA ratios (ER) of individual animal sera are negatively correlated with the packed cell volume (PCV) of whole blood (r = -0.5).

4.7.5 MPSP seropositivity is associated with total parasite load and the Ikeda genotype

The parasite load within the blood of individual animals was measured using the universal (*T. orientalis*) component of the multiplex quantitative PCR assay (Sectoin 3.1.3.2). Because *T. orientalis* infections often present as a mixture of genotypes, the relative quantities of the genotypes previously reported to be associated with clinical disease (Ikeda and Chitose genotypes) were also measured using the Ikeda and Chitose-specific components of the multiplex qPCR (Section 3.1.3.2). A Buffeli-specific qPCR was also used to measure the quantities of this genotype (Section 3.4.8 and 3.4.9). The relationship between the MPSP ELISA ratio and the total number of *T. orientalis* gene copies per µL of blood (GC/µL) or those of the individual *T. orientalis* genotypes is shown in Fig. 27. We observed a strong and significant positive correlation between the load of *T. orientalis* and the MPSP ELISA ratio ($r = 0.69$, $P<0.0001$; Fig. 27A). Comparison of individual genotype load with MPSP ELISA ratio revealed a strong positive correlation for the Ikeda genotype only ($r = 0.71$, $P<0.0001$; Fig. 27B). There was a weak but significant correlation ($r = 0.15$) between the MPSP ELISA ratio and the Chitose genotype ($p = 0.02$; Fig. 27C). No relationship was observed between the quantity of the Buffeli genotype and MPSP ELISA ratio ($r = 0.02$, $p = 0.9$; Fig. 27D).

The *T. orientalis* qPCR data were divided into three categories of infection intensity (low medium and high) based on previously established clinical thresholds (Section 4.1.3.7). Only 35% animals with infections in the low category (< 15 000 MPSP GC/µL), which typically represent subclinical carriers, were MPSP seropositive. Animals in the moderate (>15 000 but < 300 000 MPSP GC/µL) category, which usually represent recovering, in-contact or clinically-affected animals, had a much higher rate of MPSP seropositivity at 75%. In a previous study, 92% of animals in the high infection intensity range were shown to have one or more signs of clinical theileriosis; 83% of animals in this category tested seropositive (Table 31).

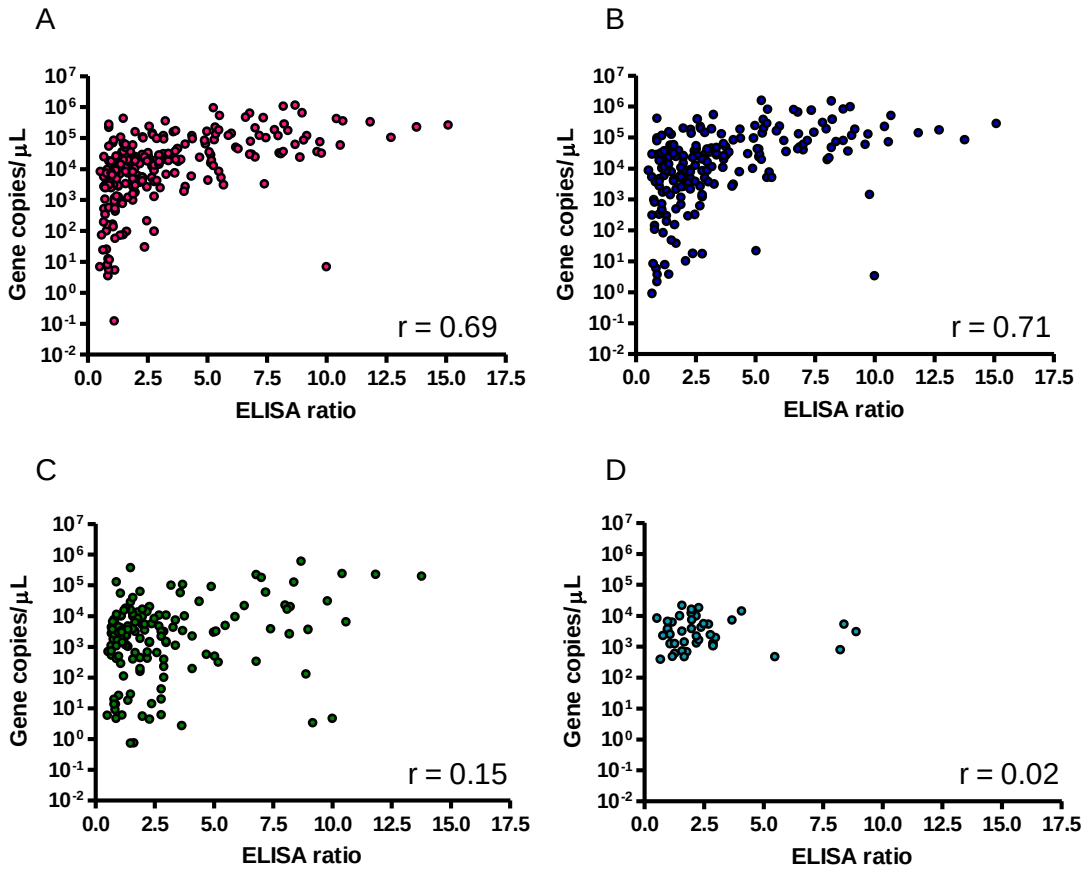


Fig. 27. (A) The total parasite load (gene copies/µL of blood) as measured by qPCR is positively correlated with the ELISA ratio (ER). Levels of the Ikeda genotype are strongly correlated with ER (B), while the levels of the Chitose genotype are only weakly correlated (C). No correlation was observed between levels of the Buffeli genotype and ER (D).

Table 31. Relationship between infection level and serological status

	Infection level*		
	High (n=12)	**Moderate (n=106)**	**Low (n=110)**
MPSP seropositive	11 (83%)	80 (75%)	39 (35%)
MPSP seronegative	1 (17%)	26 (25%)	71 (65%)

P<0.0001

*Low: <15 000 gene copies/µL blood; Moderate: >15 000 but <300 000 gene copies/µL blood; High: >300 000 gene copies/µL blood.

4.7.6 Temporal dynamics of MPSP seroconversion

We examined the temporal dynamics of seroconversion to the *T. orientalis* MPSP antigen by testing paired sera and EDTA blood from 10 naïve cattle that were introduced to a property with a known history of clinical theileriosis cases. qPCR was used to monitor the progress of infection and the serological response to the MPSP antigen was measured by MPSP ELISA. The PCV of each animal was also tested. All animals became rapidly infected with *T. orientalis* following introduction to the affected property, with 7 animals testing qPCR positive by Day 11 post-introduction and the remaining 3 animals testing positive by Day 20. The average peak in parasite load was at Day 40 post-introduction (Fig. 28A). In contrast, a positive serological response to the MPSP antigen was detected in only 1 animal at Day 20 post-introduction (this animal was qPCR positive at Day 11); however all 10 animals had seroconverted by Day 34 (Fig. 28B). The peak serological response occurred between Days 34 and 40 and preceded a sharp drop in parasite load between Days 40 and 48. The MPSP serological response declined steadily from Day 40 until the end of the sampling period (Day 76); however all animals remained in the positive range over this period. The average PCV of the cattle began to drop on Day 20, shortly after or coinciding with the onset of infection, and reaching a trough between Days 40 and 54 (Fig. 28C). All animals except one entered the anaemic range following the peak in infection intensity and the MPSP antibody response. In 9/10 cows, the decline in PCV commenced prior to the animals returning their first seropositive result; in the remaining animal, the decline in PCV coincided with the first seropositive result on Day 20; however at this time the animal was only just over the positive ER threshold of 2.

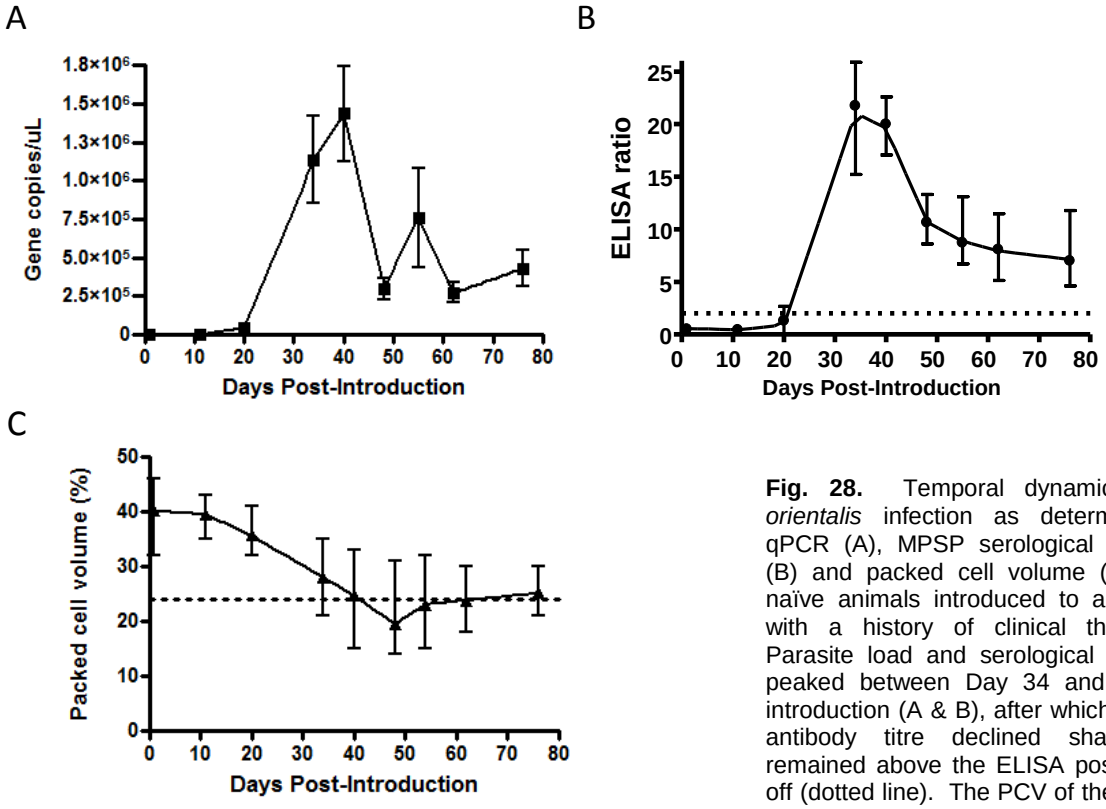


Fig. 28. Temporal dynamics of *T. orientalis* infection as determined by qPCR (A), MPSP serological response (B) and packed cell volume (C) in 10 naïve animals introduced to a property with a history of clinical theileriosis. Parasite load and serological response peaked between Day 34 and 40 post introduction (A & B), after which time the antibody titre declined sharply but remained above the ELISA positive cut-off (dotted line). The PCV of the infected animals began to decline at Day 20 post-introduction (C), with animals falling into the anaemic range between Day 40 and Day 48 post-introduction. All graphs show the mean and range of data from the 10 animals.

5 Discussion

5.1 Development and validation of an improved molecular test

Bovine theileriosis caused by *T. orientalis* is a serious problem for cattle producers in Eastern Asia and is an emerging disease in Australia and New Zealand, causing significant losses to meat and milk production (Eamens et al., 2013c; Perera et al., 2014). Disease outbreaks have been closely associated with the Ikeda type of *T. orientalis* and to a lesser extent the Chitose type (Eamens et al., 2013c; Kamau et al., 2011a; McFadden et al., 2011). The wide-spread occurrence of benign *T. orientalis* genotypes, which are morphologically and serologically indistinguishable from the clinically-associated types (Eamens et al., 2013a; Eamens et al., 2013b), necessitates the use of molecular tests for clinical diagnosis of bovine theileriosis. Furthermore, due to the high prevalence of subclinical carriers of *T. orientalis* Ikeda and Chitose, a quantitative test which can provide an accurate estimate of parasite load is essential.

The multiplex qPCR assay developed and validated in this study represents the first quantitative assay for *T. orientalis* detection that incorporates a clinical genotyping component. While real-time assays that detect *T. orientalis* have been developed previously, most have focussed on species-level detection (but not discrimination of genotypes) (Jeong et al., 2003), discrimination of different *Theileria* spp. (Chaisi et al., 2013; Yang et al., 2014), or are only semi-quantitative (Perera et al., 2015a). The UIC multiplex qPCR assay developed here was both sensitive and specific for *T. orientalis* detection compared to cPCR and reliably identified the clinically relevant Ikeda and Chitose genotypes. The additional development of a Buffeli uniplex assay, while not essential for clinical diagnostic purposes, provides additional flexibility for epidemiological studies of this benign genotype. The UIC and B assays developed here were more specific than a recently commercialised MT-PCR assay (Perera et al., 2015a), which cross-reacts between the Buffeli and Ikeda components of the assay, and provides similar sensitivity.

The U component of the UIC multiplex targets highly conserved regions of the MPSP gene in order to account for the genotypic diversity within *T. orientalis* (11 types observed currently) (Jeong et al., 2010; Khukhuu et al., 2011; Kim et al., 1998). Therefore the U component allows for additional flexibility in detecting any genotypes not typically associated with disease or for which the clinical relevance has not been fully established (Aparna et al., 2011).

Quantitative PCR has previously been used to determine parasite load within clinical blood samples (Jeong et al., 2003; Ros-Garcia et al., 2012). In this study, we observed a strong and significant correlation between MPSP allele concentration and parasitaemia (as determined by blood film), demonstrating the utility of the U component of the UIC assay in quantifying parasite load. While the correlation between total MPSP allele concentration and PCV was also significant, it was not as strong as the correlation observed between total MPSP allele concentration and parasitaemia. This observation could be due to the cyclical nature of apicomplexan infections which can recrudesce periodically (Ros-Garcia et al., 2012; Sasaki et al., 2013). Furthermore, a lag period between peak parasitaemia and minimum PCV has been observed in prior studies of both *Babesia* and *Theileria* spp. (Ros-Garcia et al., 2012; Sasaki et al., 2013). In field-based studies of *T. orientalis* infection, we

have observed that parasitaemia can peak up to 8 days before cattle become anaemic and up to 15 days before minimum PCV is reached (Section 4.4.1).

In this study, we present the first comprehensive comparison of MPSP allele concentration and genotype with clinical data from *T. orientalis*-positive animals. Indeed, the quantitative data derived from the UIC assay was found to be strongly associated with individual clinical presentation and herd status, enabling the establishment of clinical thresholds for the assay. MPSP allele concentrations of $>1.5 \times 10^4$ GC/µL (moderate infection) were strongly associated with clinical theileriosis at the herd level and within individuals. Furthermore, 95% of animals with MPSP allele concentrations in excess of 3×10^5 (high level infection) displayed symptoms of clinical theileriosis. In contrast, 100% of subclinical animals from herds without clinical cases had MPSP allele concentrations below 1.5×10^4 (low level infection). These thresholds can be applied as a guide in the differential diagnosis of theileriosis caused by *T. orientalis*, in addition to information derived from the genotyping components of the UIC assay. Based on the samples tested, there was strong evidence for the association of the Ikeda genotype, alone or in combination with Chitose and/or Buffeli genotypes, with clinical disease (100% of clinical cases examined). Where the Chitose and Buffeli genotypes occurred as sole infections, they were most often associated with herds having only subclinical cases. These findings are consistent with a prior cPCR studies (Eamens et al., 2013c; Kamau et al., 2011a) investigating the association of MPSP genotypes with clinical cases in New South Wales, Australia. Nonetheless, other studies conducted elsewhere have implicated the Chitose genotype as a direct cause of disease (Islam et al., 2011; McFadden et al., 2011), therefore the role of this genotype in clinical cases, in the presence and absence of the Ikeda genotype, warranted further investigation (Section 4.4.2).

In summary, the UIC qPCR assay presented here exploits hydrolysis probe (TaqMan®) chemistry, enabling detection of multiple products in a single well, thereby circumventing laborious processes involved in running multiple screening and genotyping PCR assays. In addition, the assay can be adapted to any platform allowing multiple fluorophore-based detection. As a genuine multiplex assay, the UIC qPCR can potentially be applied to high-throughput clinical testing enabling hundreds of UIC assays to be performed in a single day. The UIC qPCR also offers a considerable advantage over other available assays which require dilution of DNA extracts to overcome PCR inhibition (Eamens et al., 2013a; Eamens et al., 2013c; Ota et al., 2009; Perera et al., 2015a). Indeed, the number of assays required per sample to achieve a result is reduced from six, via cPCR detection, to a just single assay with the UIC qPCR. This represents a considerable (approximately 75%) reduction in turn-around time. The UIC assay is also cost-effective at an estimated A$4-10 per sample, including standards and batch controls, depending on the number of samples per run (i.e: within a range of 10-80 samples). This compares favourably with MT-PCR which costs $19 per sample (Perera et al. 2015). Furthermore, the reagent cost of DNA extraction has been reduced from $6.50 per sample to approximately 10c per sample with the successful development and validation of the DPK extraction method. Finally, the quantitative data from the U assay, combined with the IC genotyping assays, provides veterinarians with a powerful tool for clinical diagnosis of bovine theileriosis caused by *T. orientalis*.

5.2 Genotyping of strains

The aim of this part of the study was to explore the genetic diversity, subpopulations and dynamics of *T. orientalis* MPSP genotypes in field-affected cattle. It has been well established that both clinical and subclinical *T. orientalis* infections frequently occur as a mixture of genotypes (Eamens et al., 2013b; Kamau et al., 2011a; Ota et al., 2009; Sivakumar et al., 2014), a phenomenon characteristic of apicomplexan parasites (Katzer et al., 2006; Nkhoma et al., 2012). Examination of many hundreds of clinical diagnostic cases over several years has demonstrated that the Ikeda MPSP type of *T. orientalis* is always associated with disease, in the presence or absence of other MPSP types (Eamens et al., 2013c). However, while many samples from surveillance and diagnostic cases test positive for *T. orientalis* Ikeda, only some animals show clinical signs of disease. The MPSP is a major surface protein on the piroplasm stage of the parasite and is also known to be immunodominant. Therefore, one objective of the MPSP genotyping study was to determine whether animals presenting with clinical signs of theileriosis were infected with a different MPSP genotype of *T. orientalis* Ikeda to those with subclinical infections. The more variable ITS1 regions within the Ikeda genotype were similarly examined. To address this objective, we tested both clinically affected and subclinical animals that were positive for *T. orientalis* Ikeda. The fact that no differences were observed in the Ikeda MPSP types from these cases suggests that the development of clinical disease, is not associated with a particular Ikeda MPSP or ITS1 sequence, although pathogenicity within the Ikeda genotype may vary according to other as yet unidentified genetic loci. Variable clinical outcomes of infections are also likely due to host factors, such as stress, (including pregnancy) or naïvety to *T. orientalis*.

While the Ikeda genotype of *T. orientalis* has been clearly associated with clinical outbreaks of disease, in Australian herds it has been demonstrated to co-occur with the Chitose genotype with high frequency (Eamens et al., 2013c) in the presence or absence of benign genotypes (Buffeli and Type 5) (Eamens et al., 2013c; Perera et al., 2013). In this study, naïve cattle introduced to a *Theileria*-endemic area rapidly became infected with a mix of MPSP genotypes (Ikeda, Chitose and Buffeli). It was not established whether these mixed infections arose from successive inoculations of distinct parasite populations from multiple tick bites, or transmission of multiple sporozoite genotypes from individual ticks. However, evidence from field studies of mosquito transmission of the malaria parasite, *Plasmodium falciparum* (Nkhoma et al., 2012), suggest that genetic recombination within the insect vector is sufficient to generate a genetically diverse population of sporozoites. Indeed, recombination within the vector is believed to be the major source of genetic diversity in *P. falciparum,* rather than “superinfection” from multiple mosquitoes. It has been demonstrated that subsequent sporozoite inoculations are in fact suppressed once blood-stage parasites reach a minimum density threshold (Portugal et al., 2011). Multiple MPSP genotypes have been detected within *T. orientalis* sporozoites harvested from *Haemaphysalis longicornis* ticks in Japan (Kubota et al., 1996), suggesting that mixed genotypes of this parasite can be transmitted in a single infective bite. *H. longicornis* is a competent vector for the transmission of the Ikeda and Chitose genotypes in Japan and is also the likely vector within Australia (Hammer et al., 2015).

The observed temporal and magnitudinal dynamics of the *T. orientalis* genotype populations observed in this study could be explained by vector and/or immunological factors. The

pathogenic Ikeda genotype was the first to be detected in the blood of all of ten cattle post-introduction to the affected property, with the levels of the Chitose and Buffeli genotypes peaking after the Ikeda genotype had reached its peak. This pattern was consistent in all 10 animals examined. Prior studies on *Theileria parva* demonstrated that different tick populations display distinct preferences for particular genotypes of the parasite (Katzer et al., 2006). Repeated early detection of the Ikeda genotype is consistent with transmission of a sporozoite population that is skewed towards this genotype, although host immunological factors could also account for this phenomenon (see Section 5.5). Experimental transmission studies conducted in the 1980s suggested that populations of *H. longicornis* from Queensland were unable to transmit the Buffeli genotype, while *H. humerosa* and *H. bancrofti* were competent vectors (Stewart et al., 1987; Stewart et al., 1989). However, of these three potential vectors, only *H. longicornis* is known to be endemic to the mid-coast of NSW where the herd examined in this study was located and therefore is the most likely candidate vector. The relatively lower levels of Buffeli genotype observed in this study is consistent with the idea that *H. longicornis* is not a competent vector for transmitting Buffeli genotypes but further research is needed to assess the epidemiological role of Haemaphysalis sp. ticks in *T. orientalis* genotype transmission dynamics. Selection of the Ikeda and Chitose (especially Chitose A) genotypes within the tick vector could also explain the common co-occurrence of these genotypes reported here and elsewhere (Eamens et al., 2013b; Eamens et al., 2013c). Preferential selection of genotypes within the tick could also explain apparent differences in the geographical prevalence of the Chitose A versus Chitose B allele. The Chitose genotype was found to be the most prevalent genotype of *T. orientalis* in Queensland herds during a recent survey (Eamens et al., 2013a); however, all of the Chitose-positive samples sourced from Queensland that were tested in this study predominantly harboured the Chitose B allele. In addition to the natural boundaries of potential vector species (Riek, 1982), local controls of cattle movements at the Queensland-NSW border, including mandatory treatment for ticks, may assist in creating genetically divergent intraspecies tick populations with disparate parasite genotype preferences.

The host immunological response likely played a key role in the epidemiology of theileriosis outbreaks in Australia and the genotype switching of *T. orientalis* observed in this study. While *T. orientalis* has been enzootic to parts of Australia for approximately 100 years, since 2006 there has been a large increase in the number of clinical cases attributed to this parasite (Eamens et al., 2013c; Izzo et al., 2010; Perera et al., 2013). This increase in clinical cases is attributed to the detection of a new genotype in Australia (Ikeda), while previous observations of *Theileria* in Australian cattle were presumed to be of the benign Buffeli genotype (Kamau et al., 2011a). In *P. falciparum* infections, immunity to one allelic variant is not necessarily cross-protective for another (Jordan et al., 2011), therefore it is reasonable to assume that lack of a cross-protective response between *T. orientalis* genotypes in Australian cattle may have contributed to the rapid spread of clinical theileriosis since 2006. Findings derived from the genome sequencing section of this study confirm the inter-genotype diversity (Section 5.3). Additionally, lack of immune suppression of type Ikeda in cattle that had not been previously exposed to this genotype would allow out-competition of genotypes to which cattle had already been exposed. Allelic diversity within the Ikeda genotype (eg. at loci other than the MPSP) may explain the apparent re-emergence of this MPSP genotype in the majority of cattle towards the end of the sampling period.

Prior research conducted in Japan using experimentally-infected splenectomised calves have suggested that the Ikeda and Chitose genotypes undergo temporal switching (Kubota et al., 1996; Matsuba et al., 1993). Temporal switching of these same genotypes was observed in the field-affected cattle monitored in this study, although switching occurred more rapidly than reported for the splenectomised animals. Kubota et al. (1996) reported that in splenectomised calves, the Ikeda genotype dominated at 38 days post-inoculation with and Ikeda-Chitose mix of sporozoites. The infection then switched to a dominant Chitose infection by Day 48 which persisted for the course of the study (120 days), only switching back to an Ikeda-dominant infection in some animals at 115 days post-inoculation. The more rapid switching observed in this study is likely related to the intact status of the cattle examined, compared to the splenectomised animals used in the prior studies, creating a greater immunological selection pressure for more rapid shifting. Indeed, prior studies on the malaria parasite, *Plasmodium falciparum* suggest that the spleen plays an important role in the sequestration of erythrocytes containing the parasite (Contamin et al., 2000), with different parasite variants inducing the expression of alternate antigens on the erythrocyte cell surface, thereby allowing the parasite to avoid splenic destruction (Borst et al., 1995). While the role of MHC in immune modulation of *T. orientalis* has not been confirmed, in T. parva and *T. annulata* infection, CD8+ T cell responses are directed against class I MHC-bound antigenic peptides and display a clear dominance hierarchy (MacHugh et al., 2009; MacHugh et al., 2011). T cell responses are restricted further by the specific MHC haplotypes of host cattle and in *T. parva* infection, the majority of these responses are focussed on a single antigenic epitope (Macdonald et al., 2010). A similarly restricted immune response to *T. orientalis* infection could explain the relative pathogenicity of the different genotypes as well as the relative susceptibility of individual animals to pathogenic forms of the parasite. Although stress (transport, pregnancy, lactation) has been widely reported as a disease trigger (Eamens et al., 2013c; Izzo et al., 2010), the inter- and intra-breed susceptibility of cattle to *T. orientalis* infection has not been extensively investigated. There is however, some suggestion that there may be differential susceptibility of *Bos taurus* versus *Bos indicus* breeds (Yang et al., 2014) and that Japanese black cattle (Wagyu) may be less susceptible to disease (Terada et al., 1995). The animals tested in our temporal study were all of a single breed (Ayrshire) and the patterns of genotype switching and disease progression were remarkably similar amongst the majority of animals. One exception was Cow 1 which contained peak parasite levels 1-2 orders of magnitude lower than its cohorts, did not become anaemic over the course of the study, and ultimately displayed infection dominated by the Buffeli genotype. These differences in the progression of infection may well be related to the nature of the immune response in Cow 1 resulting from the interplay between host and parasite genotype.

While the Ikeda genotype of *T. orientalis* has been clearly linked to clinical disease, the pathogenicity of the Chitose genotype is less clear. While this genotype frequently co-occurs with type Ikeda, it has only rarely been suggested to be the sole cause of disease (McFadden et al., 2011). Futhermore, McFadden et al., used the 18S rRNA and cytochrome oxidase III genes, rather than the MPSP gene to characterise their isolate, therefore the MPSP genotype of this isolate remains unconfirmed. Indeed, research presented in this report shows the Chitose genotype is only associated with high parasite loads when occurring in combination with Ikeda. Here, we identified a statistically higher association between clinical disease and the presence of the Chitose A compared to the Chitose B genotype; however it was noted that the Chitose A genotype almost always occurs with type

Ikeda (~95% of cases examined); therefore there is little evidence to suggest that Chitose A is a direct cause of clinical disease. While qPCR and allele frequency analysis revealed that the Chitose A genotype contributed significantly to the overall parasite burden in some animals, particularly during the acute phase of the disease where animals became anaemic, an analysis of clinical parameters in sole Ikeda versus mixed Ikeda-Chitose infections is required to determine whether this genotype exacerbates disease. A recent introduction of the Ikeda genotype to Australia, co-inciding with the increase in disease outbreaks, is believed to have occurred, with a phylogenetic relationship between Australian and Japanese Ikeda isolates proposed (Perera et al., 2015b). The phylogenetic analysis of the Chitose sequences presented here also supports a recent Japanese origin for Chitose A but not for the Chitose B genotype. Combined with the strong association observed between the Chitose A and Ikeda genotypes, we postulate that these genotypes were likely introduced to Australia at the same time. In contrast, a Chitose type B MPSP sequence was detected in a Japanese study (Kim et al., 1998) from animal imported from Australia as early as 1996, suggesting that Chitose B may have been present in Australia for some time prior to the disease outbreaks.

5.3 Genome sequencing

This study presents the first genomic analysis of Australian isolates of *T. orientalis* and the first published genomic sequences of the Chitose and Buffeli genotypes. Genomic studies allow for multi-locus strategies to define evolutionary relationships with greater confidence than those that rely on single genes. In the Piroplasmida phylum, the placement of species such as *Theileria* (formerly *Babesia*) *equi* and *Babesia microti* has been greatly clarified using these techniques (Cornillot et al., 2012; Kappmeyer et al., 2012). The taxonomy of *T. orientalis* has previously been considered controversial (Uilenberg, 2011) and originally *T. orientalis*, *T. sergenti* and *T. buffeli* were all used, in a regionally specific manner, to describe this organism. Researchers in early molecular studies that originally attempted to identify differences that could separate morphologically indistinguishable *T. orientalis, T. sergenti* and *T. buffeli* instead discovered that these “species” consisted of multiple common genotypes that often had little correlation with these definitions (Kawazu et al., 1995; Kim et al., 1998; Onuma et al., 1998). The result of these led to the general use of *T. orientalis* in the literature to describe this group with differences in MPSP used to define 11 recognized genotypes (Sivakumar et al., 2014). However, the organism is still often referred to as *T. sergenti, T. buffeli*, and the *T. orientalis/buffeli* group and hence its taxonomy requires further clarification (Omar Abdallah et al., 2016; Ziam et al., 2015). Examination of the relationship between these isolates using the multi-locus strategy in this study reveals a similar structure to previous trees that have led researchers to conclude that these organisms are one species (Kamau et al., 2011a; Kim et al., 1998). However, this analysis only considers the portion of the genome consisting of single copy orthologous genes and the differences in average nucleotide identity are only slightly less than those observed between *T. parva* and *T. annulata*. Therefore we propose that the Ikeda, Chitose and Buffeli genotypes examined in this study should be considered subspecies of *T. orientalis*, a classification that is commonly used in *Plasmodium*. Further genomic examination of globally distributed genotypes (esp. Types 4-8, N1-N3) will help to further clarify the taxonomy of *T. orientalis.*

Variation analysis of these isolates, particularly within-population variation, potentially provides insight into the history of *Theileria* genotypes in Australian cattle. The Chitose

isolate examined in this study was of the sub-genotype Chitose B, which shows most similarity with isolates from south-east Asia and first discovered in Australia in the 1990s (Kim, et al. 1998). The Buffeli genotype (a.k.a. *T. buffeli*) is believed to have been present in Australia since the early 20th century, although molecular characterization of the genotype did not occur until much later (Kawazu, et al. 1995; Kim, et al. 1998). Here we have demonstrated that Chitose and Buffeli isolates show a similarly broad distribution of within-population variant allele frequencies, although the total number of variants is much higher in the Buffeli genotype. A broad range of variant allele frequencies indicates numerous haplotypes present within a single host and suggests that variant populations of both genotypes have been introduced to Australia on multiple occasions, although study of the mutation rates of *T. orientalis* genomes would be required to confirm this. In Section 5.2, we outlined that the introduction of the Chitose isolate has occurred at least twice with the identification of two Chitose sub-genotypes in Australian cattle. Both Chitose B and Buffeli genotypes are associated with benign theileriosis infections and this, in addition to Australian quarantine procedures that apply to livestock, may explain the higher within-population diversity observed in these isolates.

The Ikeda genotype was first identified in eastern Australia in 2006 and has since become endemic to the south-eastern coastal region, closely matching the distribution of the recognized vector *Haemaphyalis longicornis* (Hammer, et al. 2015). Furthermore, this genotype shows a strong association with clinical theileriosis infections (Section 5.1, Eamens, et al. 2013b; Eamens, et al. 2013c). In contrast to Chitose and Buffeli, Ikeda shows the least number of within-population variants and a substantially tighter distribution of variant allele frequencies indicating that fewer genomic haplotypes are present within the isolate examined. This suggests that *T. orientalis* Ikeda in Australia is an inbred population and also supports the hypothesis that recombination between *T. orientalis* genotypes is unlikely (Sivakumar, et al. 2014). While some studies that have observed mosaic sequences in *T. orientalis* indicative of recombination, these may have been products of PCR chimera (Haas et al., 2011; Kamau et al., 2011a). Further understanding of *T. orientalis* recombination could be extracted from further genome sequencing of Ikeda populations.

Variation between the three isolates examined in this study was relatively high, reflecting the diversity of this parasite. In addition to SNP mutations, multiple large insertion and deletion events were identified in alignments of assembled genome sequences. Moreover, many contigs from the Chitose and Buffeli genome assemblies did not align to the Shintoku reference, or any other isolate, and appear to be unique to these organisms.

Diverse, high dN/dS ratio surface proteins identified in this study included multiple uncharacterized and hypothetical proteins, 7 of which are unique to *T. orientalis*. The FAINT domain (Pf04385) also shows high representation in these proteins as indicated previously (Hayashida, et al. 2012). Other noteworthy proteins include homologs to antigen Tp2, identified in T. parva as an immunodominant T-cell antigen (Graham et al., 2006); thrombospondin-related anonymous protein, recognized as an adhesin in several *Plasmodium* and *Babesia* species (Akhouri et al., 2004; Gaffar et al., 2004; Muller et al., 1993; Rogers et al., 1992; Zhou et al., 2006) and MPSP and P23 surface proteins which are highly expressed surface antigens in *T. orientalis* that have been shown to bind heparin and in the case of the with MPSP, also shown to bind bovine erythrocytes (Takemae et al., 2014a; Takemae et al., 2014b). Previous Japanese studies have shown partial protection and reduced clinical symptoms from subunit vaccines generated from whole or immunogenic

portions of the MPSP sequence (Onuma et al., 1998). While subunit vaccines have had limited success for apicomplexan parasites the observation of low variability in this Australian isolate may indicate increased susceptibility to such a vaccine. Trials using targets identified in this study may answer this question and inform strategies for future apicomplexan disease outbreaks.

5.4 Proteomic analysis of *T. orientalis* piroplasms

While the *T. orientalis* Ikeda genome has 3995 predicted open reading frames, these genes must encode proteins for each lifecycle stage of the parasite (sporozoite, merozoite/piroplasm, schizont, gamete and kinete stages) occurring in both tick and bovine hosts. Here we used a global proteomic approach (1D SDS-PAGE electrophoresis followed by LC-MS) to identify proteins expressed during the pathogenic piroplasm phase of the parasite's lifecycle. The results indicated that we were able to obtain relatively pure preparations (ie: <5% contaminating bovine proteins) of *T. orientalis* piroplasm phase proteins from the blood of experimentally-inoculated splenectomised calves.

Proteins identified were from the both the aqueous and detergent phase of the protein extraction procedure. The aqueous phase would be expected to contain both secreted and cytosolic proteins, while detergent soluble proteins would be expected to encompass the majority of the transmembrane proteins. An *in silico* analysis suggests that there are approximately 771 membrane-associated proteins encoded by the *T. orientalis* Ikeda genome, a proportion of which will be specific to the piroplasm lifecycle stage. A large number of proteins identified by *in silico* analysis were uncharacterised proteins; however proteins involved in hexose transport and amino acid acquisition were identified and are likely to be important for parasite nutrient uptake from the host. The major piroplasm surface protein (MPSP) of *T. orientalis*, an immunodominant cell surface protein, was predicted to contain a single transmembrane domain.

The majority of proteins identified in all three genotypes of *T. orientalis* were uncharacterised proteins, or proteins for which no function has been assigned; however large numbers of proteins involved in metabolism, translation and protein processing and turnover were also identified, particularly in the aqueous phase. A number of proteins involved in virulence were detected, including a lipoprotein, an erythrocyte membrane binding protein, a haemolysin and a protein (Tolocg 1) unique to *Theileria* spp. Interestingly, the protein Tolocg 1 was identified at concentrations 100× higher in the *T. orientalis* Ikeda protein sample compared to the Buffeli protein sample and this protein was not identified in the Chitose genotype at all. In other *Theileria* spp. this protein is believed to represent a potentially immunogenic protein. The apparently large differential expression level of this protein in the different *T. orientalis* genotypes suggests that it plays a particularly important role in the Ikeda genotype.

Global proteomic analyses such as the one conducted in this study are of central importance to future vaccine development strategies by identifying potential targets that are expressed during the pathogenic (piroplasm) phase of the protein and exposed to the cell surface. The proteomic analyses conducted here will inform future efforts aimed at using reverse vaccinology to design a subunit vaccine for *T. orientalis*.

5.5 Serological test development

The drivers of the host immunological response to *T. orientalis* infection are poorly studied but require investigation if effective vaccines are to be developed against this parasite in the future. Here we examined factors associated with stimulation of a humoral response to an MPSP of *T. orientalis* which is expressed during both the merozoite (piroplasm) and sporozoite life cycle phases of the parasite. The serological status of *T. orientalis*-infected animals is of interest given a potential role for humoral immunity in the host response to infection. While humoral responses to intracellular pathogens have traditionally been considered of less importance than cell-mediated responses, there are many recent examples where antibodies have been found to afford protection by acting in concert with the cellular immune system or by neutralising extracellular phases of the pathogen prior to gaining cell entry (Casadevall, 2003).

While the host response to infection in the transforming *Theileria* spp. (*T. parva* and *T. annulata)* appears to be largely mediated by cytotoxic T cell responses against macroschizont-infected lymphoblasts (Graham et al., 2006); *T. orientalis* lacks the ability to transform leukocytes. In contrast, the pathogenic phase of *T. orientalis* is the intraerythrocytic phase, with the major symptoms of disease resulting from erythrocyte destruction and subsequent anaemia. In this study we found a significant association between MPSP seropositivity and anaemia with 89% of anaemic animals seroconverting to the MPSP antigen. Conversely, only 45% of animals with PCVs in the normal range had seroconverted. While seroconversion to *T. parva* shows a small but significant inverse association with PCV, the association between anaemia and seroconversion to the tick fever parasite, *Babesia bigemina* is much more marked (Magona et al., 2008). Like *T. orientalis, B. bigemina* is a non-transforming member of the Piroplasmida, exerting its major pathogenic effects via its intraerythrocytic phase (Hunfeld et al., 2008). Thus, it appears that in both *B. bigemina* and *T. orientalis* infection, seroconversion is linked to erythrocyte destruction. This observation was consistent with data from our temporal study in which 10 naïve cows became acutely infected with *T. orientalis*. In these animals, seroconversion to the MPSP antigen had already occurred (Day 34) by the time the parasite load peaked (Day 40). Furthermore, while the animals did not become anaemic until approximately 10 days after parasite load and serological response had peaked, the onset of the decline in PCV commenced very early in the time course (Day 20 on average). In addition, 9/10 animals were seronegative on Day 20, despite the fact that the majority of animals tested qPCR positive for *T. orientalis* by Day 11. Together, these data suggest that seroconversion may occur as a consequence of the release of free parasites/parasite antigen from lysed erythrocytes rather than in response to the initial inoculation of sporozoites from the tick. Indeed prior studies on *Theileria* spp. indicate that the sporozoites become rapidly internalised (within 3 minutes) of cell attachment (Shaw, 1997).

In this study, we revealed a strong, positive and significant correlation between the MPSP antibody-ELISA ratio and the total parasite load as determined by qPCR. Furthermore, when the levels of the individual *T. orientalis* genotypes were examined, the correlation was strongest ($r = 0.71$) with the load of the *T. orientalis* Ikeda genotype. The Ikeda genotype has been linked to many clinical outbreaks of bovine theileriosis in the Asia-Pacific region (Eamens et al., 2013c; Kamau et al., 2011a; Perera et al., 2013; Pulford et al., 2016) and is considered the major pathogenic genotype of this parasite. In contrast the levels of the Chitose and Buffeli genotypes, which are very rarely associated with disease, showed only a

weak or no correlation with the MPSP ELISA ratio respectively. We demonstrated in this study (Section 5.1) that the Ikeda genotype of *T. orientalis* is frequently associated with high infection intensities and furthermore, that infection intensity is negatively correlated with PCV. The decline in PCV resulting from *T. orientalis* Ikeda infection reflects enhanced erythrocyte destruction in the presence of this genotype and consequently, the increased likelihood of a serological response associated with infection by this genotype.

Results from our temporal study indicated that animals acutely infected with *T. orientalis* seroconverted 2-3 wk after the parasite was detectable via qPCR (34 days post-introduction to the affected herd). This is directly comparable to a study on the tick fever parasites, *Babesia bovis* and *B. bigemina*, in which animals introduced to a property with ticks harbouring these parasites all seroconverted within 35 days, coinciding with a decline in their PCVs (Bock et al., 1999). Nonetheless, tick-borne parasites vary in their ability to induce a sustained serological response in their hosts. Animals acutely infected with *T. parva* may die prior to seroconversion; however surviving animals tend to produce a long-lasting serological response and are generally immune to re-infection thereafter (Kiara et al., 2014). While the longevity of the serological response to *Babesia* spp. is variable, seropositivity is often sustained for between 18 mos and 6 yrs (Bock et al., 1999; Homer et al., 2000; Hunfeld et al., 2008). In contrast, *Theileria mutans* induces only short-term serological responses (Kiara et al., 2014). In the case of *T. orientalis*, we demonstrated that the antibody titre declines steeply (along with the parasite burden) following acute infection; however in all animals examined, the MPSP serological response plateaued within the positive range for the remainder of the study period (76 days). Whether the serological response is sustained over the longer term is yet to be determined.

While cell-mediated responses are central to protection from re-infection with the transforming theilerias (Goddeeris et al., 1986; Machugh et al., 2008; Sivakumar et al., 2014) the pathogenesis of *T. orientalis* more closely resembles that of non-transforming piroplasmids, *B. bovis* and *B. bigemina*. *B. bovis* infection is concomitant: animals remain persistently infected with the parasite, but resist disease relapse (Brown et al., 2006). While *T. orientalis* infections are known to persist long-term (Kubota et al., 1996), it is currently unclear whether acutely infected animals become refractory to subsequent disease. In *B. bovis* infection, immune protection appears to involve a combination of innate and cell-mediated immunity and neutralising antibody (Brown et al., 2006). The spleen is central to the host response to acute disease and removes parasitised erythrocytes from circulation, while subsequent stimulation of CD4+ T lymphocytes and the production of neutralising antibody appear to be responsible for subsequent adaptive immunity to the parasite. The spleen is evidently of similar importance in responding to acute theileriosis caused by *T. orientalis*, with splenomegaly a common finding on necropsy (Izzo et al., 2010) and splenectomy inducing high parasitaemias in infected animals (Kawamoto et al., 1991). Indeed, splenic clearance of parasitised erythrocytes is most likely responsible for the declines in PCV observed in acutely affected animals in this study and the spleen may also be important in the generation of the serological response via the white pulp.

In *B. bovis* infection, antibodies are believed to be most important in neutralising extracellular phases of the parasite including the sporozoite and merozoite phases, and in capturing erythrocytes expressing parasite antigen on the cell surface (Brown et al., 2006). In some cases, immunity to *Babesia* infection only appears to occur following seroconversion (Haselbarth et al., 2007; Hunfeld et al., 2008; Krause et al., 2008). Whether the humoral

response to the *T. orientalis* MPSP protects against recrudescence or re-infection with new *T. orientalis* strains remains unclear, but is worthy of further investigation. Indeed, one study suggested that passive transfer of monoclonal antibodies raised against MPSP epitopes had an immunoprotective effect on calves challenged with *T. orientalis* (Tanaka et al., 1990). Similarly, calves vaccinated with either recombinant MPSP antigen or peptides derived from the MPSP antigen showed reduced parasitaemias and an absence of clinical signs of disease relative to control calves following experimental challenge (Onuma et al., 1997). Further investigation of the *T. orientalis* MPSP as a potential candidate for a subunit vaccine is worthwhile due to the promising results from these initial investigations. Indeed, the MPSP gene was one of the genes identified in our genomic study as having a high dN/dS ratio, suggesting that this antigen could be a suitable candidate for a subunit vaccine (Section 4.5.3).. While live vaccines are generally favoured for the prevention of apicomplexan infections due to high levels of genetic variation within these organisms, *T. orientalis* may represent a unique case given that a single MPSP genotype (Ikeda) is correlated with the majority of clinical outbreaks and our genomic study of intra-population variation suggests that this genotype is relatively inbred within Australia.

6 Conclusions/recommendations

In this project we have delivered an improved molecular diagnostic test for bovine theileriosis caused by *T. orientalis,* which accurately quantifies parasite load and determines the relative abundances of the two genotypes associated with clinical disease. The test compares favourably in terms of sensitivity and specificity to existing conventional PCR assays and a recently developed commercial assay and represents a cost-effective alternative to both of these tests. The additional development and validation of an inexpensive extraction method for *Theileria* DNA from blood further reduces the cost of testing. Furthermore, clinical thresholds have now been established for disease which allow enhanced diagnostic interpretation by veterinarians of levels of *Theileria* which represent subclinical vs clinical infection. The multiplex qPCR assay described here has been used to diagnose cases of bovine theileriosis in new regions within Australia (ie: South Australia and Western Australia) and has also been used to diagnose disease in cattle in New Caledonia. The multiplex assay has now been successfully validated in the the government lab in New Caledonia for routine diagnostic testing for *T. orientalis*.

This project has also delivered a serological test (antibody-ELISA) for *T. orientalis* detection. While PCR assays remain the gold standard, the ELISA described here would be suitable for testing at the herd level and with further investigation, may have future utility in testing animals for prior exposure to, or protection against *T. orientalis*. Investigations of factors associated with serological responses to *T. orientalis* suggest that seroconversion is more likely to occur in acutely infected animals rather than subclinical carriers and is more likely to occur in response to infection with *T. orientalis* Ikeda than non-pathogenic subtypes.

This project has laid substantial groundwork for future vaccine development against *T. orientalis* by increasing our knowledge of the genomic and proteomic features of this parasite. Limited genetic variability (inbreeding) with *T. orientalis* Ikeda in Australia, revealed by genome analysis, makes this a good candidate for subunit vaccine development relative to other apicomplexan parasites. Genomic and proteomic analyses have identified potential

target antigens for vaccine development which are surface exposed. Future work on vaccine development will use the genomic and proteomic data from this study in a reverse vaccinology pipeline (Vacceed).

Publications arising directly or indirectly from this project are listed below:

- Bogema DR, Micallef M, Liu M, Darling AE, Padula M, Djordjevic SP, Jenkins C. The genomes of Australian *Theileria orientalis* genotypes Ikeda, Chitose and Buffeli show variable within-host population diversity. (In preparation).
- Swilks E, Poynting A, Jenkins C, Krebs GL. The prevalence and effect of *Theileria orientalis* infection in young calves of the Gloucester region of New South Wales. Australian Veterinary Journal (Accepted).
- Jenkins C, Bogema DR (2016). Factors associated with seroconversion to the major piroplasm surface protein of the bovine haemoparasite, *Theileria orientalis*. Parasites and Vectors 9:106.
- Hammer JF, Jenkins C, Bogema DR, Emery D (2016). Mechanical transfer of *Theileria orientalis*: possible roles of biting arthropods, colostrum and husbandry practices in disease transmission . Parasites and Vectors 9:34.
- Proctor AKK, Ball M, Freeman P, Jenkins C and Bogema DR (2015). Prevalence of *Theileria orientalis* types in North Coast New South Wales beef cattle herds. Australian Veterinary Journal. (In Press).
- Bogema DR, Fell SA, O'Rourke BA, Collins D, Eamens GJ and Jenkins C (2015). Development and validation of an inexpensive and efficient method for extraction of *Theileria orientalis* DNA from blood. Veterinary Parasitology 212:379-381.
- Jenkins C., Micallef M, Alex SM, Collins D, Djordjevic SP and Bogema DR. (2015). Temporal dynamics and subpopulation analysis of *Theileria orientalis* genotypes in cattle. Infection, Genetics and Evolution. 32: 109-207. (doi: 10.1016/j.meegid.2015.03.017)
- Bogema DR, Deutscher AT, Fell SA, Collins D, Eamens GJ and Jenkins C. (2015). Development and validation of a multiplexed hydrolysis probe qPCR assay for the detection and quantification of *Theileria orientalis* and differentiation of clinically-relevant subtypes. Journal of Clinical Microbiology 53:941-50.
- Hammer JF, Emery D, Bogema DR, Jenkins C (2015). Detection of *Theileria orientalis* genotypes in *Haemaphysalis longicornis* ticks from Southern Australia. Parasites and Vectors. 8:229. (doi: 10.1186/s13071-015-0839-9).

7 Key messages

- An improved molecular diagnostic test for detection of T. orientalis and discrimination of the clinically relevant subtypes has been developed and validated. This test is cost effective relative to existing molecular assays and offers greater specificity than the commercially available MT-PCR.
- The qPCR assay developed in this study provides information on the parasite load in a given sample and combined with clinical thresholds established in this study, can be used to assist veterinarians in distinguishing between subclinical carriers of the parasite and animals with clinical theileriosis.

- A serological test for *T. orientalis* is now available. While less sensitive than qPCR, this test may suitable for herd-level screening or to determine whether animals have been previously exposed to *T. orientalis.* Acutely affected animals are more likely to elicit a serological response to the parasite. An investigation of the ability of a lasting serological response to *T. orientalis* to protect against disease recrudescence is warranted.
- Genotyping and genomic studies conducted have demonstrated that *T. orientalis* Ikeda is a relatively inbred population within Australia and is therefore a good candidate for vaccine development.
- Proteomic and genomic studies conducted here will form the basis of future studies employing a reverse vaccinology pipeline (Vacceed) to develop a subunit vaccine for *T. orientalis.*

8 Bibliography

Abyzov, A., Urban, A.E., Snyder, M., Gerstein, M., 2011. CNVnator: an approach to discover, genotype, and characterize typical and atypical CNVs from family and population genome sequencing. Genome research 21, 974-984.

Akhouri, R.R., Bhattacharyya, A., Pattnaik, P., Malhotra, P., Sharma, A., 2004. Structural and functional dissection of the adhesive domains of *Plasmodium falciparum* thrombospondin-related anonymous protein (TRAP). The Biochemical journal 379, 815-822.

Aktas, M., Bendele, K.G., Altay, K., Dumanli, N., Tsuji, M., Holman, P.J., 2007. Sequence polymorphism in the ribosomal DNA internal transcribed spacers differs among *Theileria* species. Vet Parasitol 147, 221-230.

Al-Soud, W.A., Jonsson, L.J., Radstrom, P., 2000. Identification and characterization of immunoglobulin G in blood as a major inhibitor of diagnostic PCR. J Clin Microbiol 38, 345-350.

Al-Soud, W.A., Radstrom, P., 2001. Purification and characterization of PCR-inhibitory components in blood cells. J Clin Microbiol 39, 485-493.

Aparna, M., Ravindran, R., Vimalkumar, M.B., Lakshmanan, B., Rameshkumar, P., Kumar, K.G., Promod, K., Ajithkumar, S., Ravishankar, C., Devada, K., Subramanian, H., George, A.J., Ghosh, S., 2011. Molecular characterization of *Theileria orientalis* causing fatal infection in crossbred adult bovines of South India. Parasitol Int 60, 524-529.

Armbruster, D.A., Pry, T., 2008. Limit of blank, limit of detection and limit of quantitation. The Clinical biochemist. Reviews / Australian Association of Clinical Biochemists 29 Suppl 1, S49-52.

Bailey, G. 2013. Meat and Livestock Australia Final Report B.AHE.0194: Buparvaquone tissue residue study.

Bock, R.E., Kingston, T.G., de Vos, A.J., 1999. Effect of breed of cattle on transmission rate and innate resistance to infection with *Babesia bovis* and *B. bigemina* transmitted by *Boophilus microplus*. Aust Vet J 77, 461-464.

Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30, 2114-2120.

Borst, P., Bitter, W., McCulloch, R., Van Leeuwen, F., Rudenko, G., 1995. Antigenic variation in malaria. Cell 82, 1-4.

Brayton, K.A., Lau, A.O., Herndon, D.R., Hannick, L., Kappmeyer, L.S., Berens, S.J., Bidwell, S.L., Brown, W.C., Crabtree, J., Fadrosh, D., Feldblum, T., Forberger, H.A., Haas, B.J., Howell, J.M., Khouri, H., Koo, H., Mann, D.J., Norimine, J., Paulsen, I.T., Radune, D., Ren, Q., Smith, R.K., Jr., Suarez, C.E., White, O., Wortman, J.R., Knowles, D.P., Jr., McElwain, T.F., Nene, V.M., 2007. Genome sequence of *Babesia bovis* and comparative analysis of apicomplexan hemoprotozoa. PLoS pathogens 3, 1401-1413.

Breese, M.R., Liu, Y., 2013. NGSUtils: a software suite for analyzing and manipulating next-generation sequencing datasets. Bioinformatics 29, 494-496.

Brown, W.C., Norimine, J., Knowles, D.P., Goff, W.L., 2006. Immune control of *Babesia bovis* infection. Vet Parasitol 138, 75-87.

Buh Gasparic, M., Cankar, K., Zel, J., Gruden, K., 2008. Comparison of different real-time PCR chemistries and their suitability for detection and quantification of genetically modified organisms. BMC biotechnology 8, 26.

Bustin, S.A., Benes, V., Garson, J.A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M.W., Shipley, G.L., Vandesompele, J., Wittwer, C.T., 2009. The MIQE Guidelines: Minimum information for publication of quantitative real-time PCR experiments. Clinical Chemistry 55, 611-622.

Capella-Gutierrez, S., Silla-Martinez, J.M., Gabaldon, T., 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics 25, 1972-1973.

Casadevall, A., 2003. Antibody-mediated immunity against intracellular pathogens: two-dimensional thinking comes full circle. Infect Immun 71, 4225-4228.

Chaisi, M.E., Janssens, M.E., Vermeiren, L., Oosthuizen, M.C., Collins, N.E., Geysen, D., 2013. Evaluation of a real-time PCR test for the detection and discrimination of theileria species in the African buffalo (*Syncerus caffer*). PloS one 8, e75827.

Chen, Y., Bogema, D.R., Barchia, I.M., Herron, G.A., 2014. Quantification of the pirimicarb resistance allele frequency in pooled cotton aphid (*Aphis gossypii* Glover) samples by TaqMan SNP genotyping assay. PloS one 9, e91104.

Chu, J., Sadeghi, S., Raymond, A., Jackman, S.D., Nip, K.M., Mar, R., Mohamadi, H., Butterfield, Y.S., Robertson, A.G., Birol, I., 2014. BioBloom tools: fast, accurate and memory-efficient host species sequence screening using bloom filters. Bioinformatics 30, 3402-3404.

Clopper, C.J., Pearson, E.S., 1934. The use of confidence or fiducial limits illustrated in the case of the binomial. Biometrika 26, 404-413.

Cock, P.J., Antao, T., Chang, J.T., Chapman, B.A., Cox, C.J., Dalke, A., Friedberg, I., Hamelryck, T., Kauff, F., Wilczynski, B., de Hoon, M.J., 2009. Biopython: freely available Python tools for computational molecular biology and bioinformatics. Bioinformatics 25, 1422-1423.

Coil, D., Jospin, G., Darling, A.E., 2015. A5-miseq: an updated pipeline to assemble microbial genomes from Illumina MiSeq data. Bioinformatics 31, 587-589.

Contamin, H., Behr, C., Mercereau-Puijalon, O., Michel, J., 2000. *Plasmodium falciparum* in the squirrel monkey (*Saimiri sciureus*): infection of non-splenectomised animals as a model for exploring clinical manifestations of malaria. Microbes and infection / Institut Pasteur 2, 945-954.

Cornillot, E., Hadj-Kaddour, K., Dassouli, A., Noel, B., Ranwez, V., Vacherie, B., Augagneur, Y., Bres, V., Duclos, A., Randazzo, S., Carcy, B., Debierre-Grockiego, F., Delbecq, S., Moubri-Menage, K., Shams-Eldin, H., Usmani-Brown, S., Bringaud, F., Wincker, P., Vivares, C.P., Schwarz, R.T., Schetters, T.P., Krause, P.J., Gorenflot, A., Berry, V., Barbe, V., Ben Mamoun, C., 2012. Sequencing of the smallest Apicomplexan genome from the human pathogen *Babesia microti*. Nucleic acids research 40, 9102-9114.

Cufos, N., Jabbar, A., de Carvalho, L.M., Gasser, R.B., 2012. Mutation scanning-based analysis of *Theileria orientalis* populations in cattle following an outbreak. Electrophoresis 33, 2036-2040.

Deitsch, K.W., Lukehart, S.A., Stringer, J.R., 2009. Common strategies for antigenic variation by bacterial, fungal and protozoan pathogens. Nature reviews. Microbiology 7, 493-503.

Deutscher, A.T., Jenkins, C., Minion, F.C., Seymour, L.M., Padula, M.P., Dixon, N.E., Walker, M.J., Djordjevic, S.P., 2010. Repeat regions R1 and R2 in the P97 paralogue Mhp271 of *Mycoplasma hyopneumoniae* bind heparin, fibronectin and porcine cilia. Molecular Microbiology 78, 444-458.

Dobbelaere, D., Heussler, V., 1999. Transformation of leukocytes by *Theileria parva* and *T. annulata*. Annual review of microbiology 53, 1-42.

Eamens, G.J. 2011. Meat and Livestock Australia final report: Bovine theileriosis - distribution and significance of major piroplasm surface protein (MPSP) types (NSW Department of Primary Industries).

Eamens, G.J., Bailey, G., Gonsalves, J.R., Jenkins, C., 2013a. Distribution and temporal prevalence of *Theileria orientalis* major piroplasm surface protein types in eastern Australian cattle herds. Aust Vet J 91, 332-340.

Eamens, G.J., Bailey, G., Jenkins, C., Gonsalves, J.R., 2013b. Significance of *Theileria orientalis* types in individual affected beef herds in New South Wales based on clinical, smear and PCR findings. Vet Parasitol 196, 96-105.

Eamens, G.J., Gonsalves, J.R., Jenkins, C., Collins, D., Bailey, G., 2013c. *Theileria orientalis* MPSP types in Australian cattle herds associated with outbreaks of clinical disease and their association with clinical pathology findings. Vet Parasitol 191, 209-217.

Edgar, R.C., 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic acids research 32, 1792-1797.

Emms, D.M., Kelly, S., 2015. OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. Genome Biol 16, 157.

Endo, T., Ikeo, K., Gojobori, T., 1996. Large-scale search for genes on which positive selection may operate. Mol Biol Evol 13, 685-690.

Fisher, C.L., Pei, G.K., 1997. Modification of a PCR-based site-directed mutagenesis method. Biotechniques 23, 570-&.

Fujisaki, K., Kamio, T., Kawazu, S., Shimizu, S., Simura, K., 1993. *Theileria sergenti*: experimental transmission by the long-nosed cattle louse, *Linognathus vituli*. Ann Trop Med Parasitol 87, 217-218.

Fujisaki, K., Kawazu, S., Kamio, T., 1994. The taxonomy of the bovine *Theileria* spp. Parasitology Today 10, 31-33.

Gaffar, F.R., Yatsuda, A.P., Franssen, F.F., de Vries, E., 2004. A *Babesia bovis* merozoite protein with a domain architecture highly similar to the thrombospondin-related anonymous protein (TRAP) present in *Plasmodium* sporozoites. Mol Biochem Parasitol 136, 25-34.

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

Gardner, M.J., Hall, N., Fung, E., White, O., Berriman, M., Hyman, R.W., Carlton, J.M., Pain, A., Nelson, K.E., Bowman, S., Paulsen, I.T., James, K., Eisen, J.A., Rutherford, K., Salzberg, S.L., Craig, A., Kyes, S., Chan, M.S., Nene, V., Shallom, S.J., Suh, B., Peterson, J., Angiuoli, S., Pertea, M., Allen, J., Selengut, J., Haft, D., Mather, M.W., Vaidya, A.B., Martin, D.M., Fairlamb, A.H., Fraunholz, M.J., Roos, D.S., Ralph, S.A., McFadden, G.I., Cummings, L.M., Subramanian, G.M., Mungall, C., Venter, J.C., Carucci, D.J., Hoffman, S.L., Newbold, C., Davis, R.W., Fraser, C.M., Barrell, B.,

2002. Genome sequence of the human malaria parasite *Plasmodium falciparum*. Nature 419, 498-511.

Goddeeris, B.M., Morrison, W.I., Teale, A.J., Bensaid, A., Baldwin, C.L., 1986. Bovine cytotoxic T-cell clones specific for cells infected with the protozoan parasite *Theileria parva*: parasite strain specificity and class I major histocompatibility complex restriction. Proc Natl Acad Sci USA 83, 5238-5242.

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

Gubbels, M.J., Katzer, F., Shiels, B.R., Jongejan, F., 2001. Study of *Theileria annulata* population structure during bovine infection and following transmission to ticks. Parasitology 123, 553-561.

Haas, B.J., Gevers, D., Earl, A.M., Feldgarden, M., Ward, D.V., Giannoukos, G., Ciulla, D., Tabbaa, D., Highlander, S.K., Sodergren, E., Methe, B., DeSantis, T.Z., Human Microbiome, C., Petrosino, J.F., Knight, R., Birren, B.W., 2011. Chimeric 16S rRNA sequence formation and detection in Sanger and 454-pyrosequenced PCR amplicons. Genome Res 21, 494-504.

Hammer, J., Emery, D., Bogema, D.R., Jenkins, C., 2015. Detection of *Theileria orientalis* genotypes in *Haemaphysalis longicornis* ticks from southern Australia. Parasit Vectors 8, 229.

Hammer, J.F., Jenkins, C., Bogema, D.R., Emery, D., 2016. Mechanical transfer of *Theileria orientalis*: possible roles of biting arthropods, colostrum and husbandry practices in disease transmission. Parasit Vectors 9, 34.

Haselbarth, K., Tenter, A.M., Brade, V., Krieger, G., Hunfeld, K.P., 2007. First case of human babesiosis in Germany - Clinical presentation and molecular characterisation of the pathogen. International journal of medical microbiology : IJMM 297, 197-204.

Hayashida, K., Abe, T., Weir, W., Nakao, R., Ito, K., Kajino, K., Suzuki, Y., Jongejan, F., Geysen, D., Sugimoto, C., 2013. Whole-genome sequencing of *Theileria parva* strains provides insight into parasite migration and diversification in the African continent. DNA research : an international journal for rapid publication of reports on genes and genomes 20, 209-220.

Hayashida, K., Hara, Y., Abe, T., Yamasaki, C., Toyoda, A., Kosuge, T., Suzuki, Y., Sato, Y., Kawashima, S., Katayama, T., Wakaguri, H., Inoue, N., Homma, K., Tada-Umezaki, M., Yagi, Y., Fujii, Y., Habara, T., Kanehisa, M., Watanabe, H., Ito, K., Gojobori, T., Sugawara, H., Imanishi, T., Weir, W., Gardner, M., Pain, A., Shiels, B., Hattori, M., Nene, V., Sugimoto, C., 2012. Comparative genome analysis of three eukaryotic parasites with differing abilities to transform leukocytes reveals key mediators of *Theileria*-induced leukocyte transformation. mBio 3, e00204-00212.

Healy, P.J., Dennis, J.A., Moule, J.F., 1995. Use of hair root as a source of DNA for the detection of heterozygotes for recessive defects in cattle. Aust Vet J 72, 392.

Henson, S., Bishop, R.P., Morzaria, S., Spooner, P.R., Pelle, R., Poveda, L., Ebeling, M., Kung, E., Certa, U., Daubenberger, C.A., Qi, W., 2012. High-resolution genotyping and mapping of recombination and gene conversion in the protozoan *Theileria parva* using whole genome sequencing. BMC Genomics 13, 503.

Heussler, V.T., Machado, J., Jr., Fernandez, P.C., Botteron, C., Chen, C.G., Pearse, M.J., Dobbelaere, D.A., 1999. The intracellular parasite *Theileria parva* protects infected T cells from apoptosis. Proc Natl Acad Sci USA 96, 7312-7317.

Higuchi, S., 1986. The development of *Theileria sergenti* in the salivary glands of the tick, *Haemaphysalis longicornis*. Nihon Juigaku Zasshi 48, 801-807.

Higuchi, S., 1987. Development of *Theileria sergenti* in the midgut of the larval tick, *Haemaphysalis longicornis*. Kitasato Arch Exp Med 60, 71-77.

Hikosaka, K., Tsuji, N., Watanabe, Y., Kishine, H., Horii, T., Igarashi, I., Kita, K., Tanabe, K., 2012. Novel type of linear mitochondrial genomes with dual flip-flop inversion system in apicomplexan parasites, *Babesia microti* and *Babesia rodhaini*. BMC Genomics 13, 622.

Hikosaka, K., Watanabe, Y., Tsuji, N., Kita, K., Kishine, H., Arisue, N., Palacpac, N.M., Kawazu, S., Sawai, H., Horii, T., Igarashi, I., Tanabe, K., 2010. Divergence of the mitochondrial genome structure in the apicomplexan parasites, *Babesia* and *Theileria*. Mol Biol Evol 27, 1107-1116.

Homer, M.J., Aguilar-Delfin, I., Telford, S.R., 3rd, Krause, P.J., Persing, D.H., 2000. Babesiosis. Clinical microbiology reviews 13, 451-469.

Hunfeld, K.P., Hildebrandt, A., Gray, J.S., 2008. Babesiosis: recent insights into an ancient disease. Int J Parasitol 38, 1219-1237.

Islam, M.K., Jabbar, A., Campbell, B.E., Cantacessi, C., Gasser, R.B., 2011. Bovine theileriosis - An emerging problem in south-eastern Australia? Infect Genet Evol 11, 2095-2097.

Iwasaki, T., Kakuda, T., Sako, Y., Sugimoto, C., Onuma, M., 1998. Differentiation and quantification of *Theileria sergent*i piroplasm types using type-specific monoclonal antibodies. J Vet Med Sci 60, 665-669.

Izzo, M., Poe, I., Horadagoda, N., De Vos, A.J., House, J.K., 2010. Haemolytic anaemia in cattle in NSW associated with *Theileria* infections. Aust Vet J 88, 45-51.

Jackson, A.P., Otto, T.D., Darby, A., Ramaprasad, A., Xia, D., Echaide, I.E., Farber, M., Gahlot, S., Gamble, J., Gupta, D., Gupta, Y., Jackson, L., Malandrin, L., Malas, T.B., Moussa, E., Nair, M., Reid, A.J., Sanders, M., Sharma, J., Tracey, A., Quail, M.A., Weir, W., Wastling, J.M., Hall, N., Willadsen, P., Lingelbach, K., Shiels, B., Tait, A., Berriman, M., Allred, D.R., Pain, A., 2014. The evolutionary dynamics of variant antigen genes in *Babesia* reveal a history of genomic innovation underlying host-parasite interaction. Nucleic Acids Res 42, 7113-7131.

Jenkins, C., Samudrala, R., Geary, S.J., Djordjevic, S.P., 2008. Structural and functional characterization of an organic hydroperoxide resistance protein from *Mycoplasma gallisepticum*. J Bact 190, 2206-2216.

Jenkins, C., Wilton, J.L., Minion, F.C., Falconer, L., Walker, M.J., Djordjevic, S.P., 2006. Two domains within the *Mycoplasma hyopneumoniae* cilium adhesin bind heparin. Infect Immun 74, 481-487.

Jeong, W., Kweon, C.H., Kang, S.W., Paik, S.G., 2003. Diagnosis and quantification of *Theileria sergenti* using TaqMan PCR. Vet Parasitol 111, 287-295.

Jeong, W., Yoon, S.H., An, D.J., Cho, S.H., Lee, K.K., Kim, J.Y., 2010. A molecular phylogeny of the benign *Theileria* parasites based on major piroplasm surface protein (MPSP) gene sequences. Parasitology 137, 241-249.

Jordan, S.J., Oliveira, A.L., Hernandez, J.N., Oster, R.A., Chattopadhyay, D., Branch, O.H., Rayner, J.C., 2011. Malaria immunoepidemiology in low transmission: correlation of

infecting genotype and immune response to domains of *Plasmodium falciparum* merozoite surface protein 3. Infect Immun 79, 2070-2078.

Kakuda, T., Shiki, M., Kubota, S., Sugimoto, C., Brown, W.C., Kosum, C., Nopporn, S., Onuma, M., 1998. Phylogeny of benign *Theileria* species from cattle in Thailand, China and the U.S.A. based on the major piroplasm surface protein and small subunit ribosomal RNA genes. Int J Parasitol 28, 1261-1267.

Kamau, J., de Vos, A.J., Playford, M., Salim, B., Kinyanjui, P., Sugimoto, C., 2011a. Emergence of new types of *Theileria orientalis* in Australian cattle and possible cause of theileriosis outbreaks. Parasit Vectors 4, 22.

Kamau, J., de Vos, A.J., Playford, M., Salim, B., Kinyanjui, P., Sugimoto, C., 2011b. Emergence of new types of *Theileria orientalis* in Australian cattle and possible cause of theileriosis outbreaks. Parasites & Vectors 4, 22.

Kang, S.W., Doan, H.T.T., Choe, S.E., Noh, J.H., Yoo, M.S., Reddy, K.E., Kim, Y.H., Kweon, C.H., Jung, S.C., Chang, K.Y., 2013. Molecular investigation of tick-borne pathogens in ticks from grazing cattle in Korea. Parasitology International 62, 276-282.

Kang, S.W., Lien, T.K.N., Noh, J.H., Reddy, K.E., Kweon, C.H., Choe, S.E., 2012. Phylogenetic analysis of benign *Theileria* species based on major piroplasm surface protein (MPSP) genes from ticks of grazing cattle in Korea. Vet Parasitol 189, 145-152.

Kappmeyer, L.S., Thiagarajan, M., Herndon, D.R., Ramsay, J.D., Caler, E., Djikeng, A., Gillespie, J.J., Lau, A.O., Roalson, E.H., Silva, J.C., Silva, M.G., Suarez, C.E., Ueti, M.W., Nene, V.M., Mealey, R.H., Knowles, D.P., Brayton, K.A., 2012. Comparative genomic analysis and phylogenetic position of *Theileria equi*. BMC Genomics 13, 603.

Katzer, F., Lizundia, R., Ngugi, D., Blake, D., McKeever, D., 2011. Construction of a genetic map for *Theileria parva*: identification of hotspots of recombination. Int J Parasitol 41, 669-675.

Katzer, F., Ngugi, D., Oura, C., Bishop, R.P., Taracha, E.L., Walker, A.R., McKeever, D.J., 2006. Extensive genotypic diversity in a recombining population of the apicomplexan parasite *Theileria parva*. Infect Immun 74, 5456-5464.

Kawamoto, S., Takahashi, K., Onuma, M., Kubota, S., Nejo, H., Kurosawa, T., Sonoda, M., 1991. Rebound phenomenon of parasitemia in splenectomized calves primarily infected with *Theileria sergenti*. J Vet Med Sci 53, 127-128.

Kawazu, S., Kamio, T., Sekizaki, T., Fujisaki, K., 1995. *Theileria sergenti* and *T. buffeli*: polymerase chain reaction-based marker system for differentiating the parasite species from infected cattle blood and infected tick salivary gland. Exp Parasitol 81, 430-435.

Kawazu, S., Sugimoto, C., Kamio, T., Fujisaki, K., 1992a. Antigenic differences between Japanese *Theileria sergenti* and other benign *Theileria* species of cattle from Australia (*T. buffeli*) and Britain (*T. orientalis*). Parasitol Res 78, 130-135.

Kawazu, S., Sugimoto, C., Kamio, T., Fujisaki, K., 1992b. Molecular cloning and immunological analysis of immunodominant piroplasm surface proteins of *Theileria sergenti* and *T. buffeli*. J Vet Med Sci 54, 305-311.

Khukhuu, A., Lan, D.T., Long, P.T., Ueno, A., Li, Y., Luo, Y., Macedo, A.C., Matsumoto, K., Inokuma, H., Kawazu, S., Igarashi, I., Xuan, X., Yokoyama, N., 2011. Molecular epidemiological survey of *Theileria orientalis* in Thua Thien Hue Province, Vietnam. J Vet Med Sci 73, 701-705.

Kiara, H., Jennings, A., Bronsvoort, B.M., Handel, I.G., Mwangi, S.T., Mbole-Kariuki, M., Van Wyk, I.C., Poole, E.J., Hanotte, O., Coetzer, J.A., Woolhouse, M.E., Toye, P.G., 2014. A longitudinal assessment of the serological response to *Theileria parva* and other tick-borne parasites from birth to one year in a cohort of indigenous calves in western Kenya. Parasitology 141, 1289-1298.

Kim, C., Iseki, H., Herbas, M.S., Yokoyama, N., Suzuki, H., Xuan, X., Fujisaki, K., Igarashi, I., 2007. Development of TaqMan-based real-time PCR assays for diagnostic detection of *Babesia bovis* and *Babesia bigemina.* Am J Trop Med Hygiene 77, 837-841.

Kim, S.J., Tsuji, M., Kubota, S., Wei, Q., Lee, J.M., Ishihara, C., Onuma, M., 1998. Sequence analysis of the major piroplasm surface protein gene of benign bovine *Theileria* parasites in east Asia. Int J Parasitol 28, 1219-1227.

Koboldt, D.C., Chen, K., Wylie, T., Larson, D.E., McLellan, M.D., Mardis, E.R., Weinstock, G.M., Wilson, R.K., Ding, L., 2009. VarScan: variant detection in massively parallel sequencing of individual and pooled samples. Bioinformatics 25, 2283-2285.

Koboldt, D.C., Zhang, Q., Larson, D.E., Shen, D., McLellan, M.D., Lin, L., Miller, C.A., Mardis, E.R., Ding, L., Wilson, R.K., 2012. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. Genome Res 22, 568-576.

Krause, P.J., Gewurz, B.E., Hill, D., Marty, F.M., Vannier, E., Foppa, I.M., Furman, R.R., Neuhaus, E., Skowron, G., Gupta, S., McCalla, C., Pesanti, E.L., Young, M., Heiman, D., Hsue, G., Gelfand, J.A., Wormser, G.P., Dickason, J., Bia, F.J., Hartman, B., Telford, S.R., 3rd, Christianson, D., Dardick, K., Coleman, M., Girotto, J.E., Spielman, A., 2008. Persistent and relapsing babesiosis in immunocompromised patients. Clin Infect Dis 46, 370-376.

Krogh, A., Larsson, B., von Heijne, G., Sonnhammer, E.L., 2001. Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol 305, 567-580.

Kubota, S., Sugimoto, C., Onuma, M., 1996. Population dynamics of *Theileria sergenti* in persistently infected cattle and vector ticks analysed by a polymerase chain reaction. Parasitology 112, 437-442.

Lawrence, K., McFadden, A., Gias, E., Pulford, D.J., Pomroy, W.E., 2016. Epidemiology of the epidemic of bovine anaemia associated with *Theileria orientalis* (Ikeda) between August 2012 and March 2014. N Z Vet J 64, 38-47.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., Genome Project Data Processing, S., 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078-2079.

Maaroufi, Y., Heymans, C., De Bruyne, J.M., Duchateau, V., Rodriguez-Villalobos, H., Aoun, M., Crokaert, F., 2003. Rapid detection of *Candida albicans* in clinical blood samples by using a TaqMan-based PCR assay. J Clin Microbiol 41, 3293-3298.

Macdonald, I.K., Harkiolaki, M., Hunt, L., Connelley, T., Carroll, A.V., MacHugh, N.D., Graham, S.P., Jones, E.Y., Morrison, W.I., Flower, D.R., Ellis, S.A., 2010. MHC class I bound to an immunodominant *Theileria parva* epitope demonstrates unconventional presentation to T cell receptors. PLoS pathogens 6, e1001149.

Machugh, N.D., Burrells, A.C., Morrison, W.I., 2008. Demonstration of strain-specific CD8 T cell responses to *Theileria annulata*. Parasite Immunol 30, 385-393.

MacHugh, N.D., Connelley, T., Graham, S.P., Pelle, R., Formisano, P., Taracha, E.L., Ellis, S.A., McKeever, D.J., Burrells, A., Morrison, W.I., 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.

MacHugh, N.D., Weir, W., Burrells, A., Lizundia, R., Graham, S.P., Taracha, E.L., Shiels, B.R., Langsley, G., Morrison, W.I., 2011. Extensive polymorphism and evidence of immune selection in a highly dominant antigen recognized by bovine CD8 T cells specific for *Theileria annulata*. Infect Immun 79, 2059-2069.

Magona, J.W., Walubengo, J., Olaho-Mukani, W., Jonsson, N.N., Welburn, S.C., Eisler, M.C., 2008. Clinical features associated with seroconversion to *Anaplasma marginale*, *Babesia bigemina* and *Theileria parva* infections in African cattle under natural tick challenge. Vet Parasitol 155, 273-280.

Manske, M., Miotto, O., Campino, S., Auburn, S., Almagro-Garcia, J., Maslen, G., O'Brien, J., Djimde, A., Doumbo, O., Zongo, I., Ouedraogo, J.B., Michon, P., Mueller, I., Siba, P., Nzila, A., Borrmann, S., Kiara, S.M., Marsh, K., Jiang, H., Su, X.Z., Amaratunga, C., Fairhurst, R., Socheat, D., Nosten, F., Imwong, M., White, N.J., Sanders, M., Anastasi, E., Alcock, D., Drury, E., Oyola, S., Quail, M.A., Turner, D.J., Ruano-Rubio, V., Jyothi, D., Amenga-Etego, L., Hubbart, C., Jeffreys, A., Rowlands, K., Sutherland, C., Roper, C., Mangano, V., Modiano, D., Tan, J.C., Ferdig, M.T., Amambua-Ngwa, A., Conway, D.J., Takala-Harrison, S., Plowe, C.V., Rayner, J.C., Rockett, K.A., Clark, T.G., Newbold, C.I., Berriman, M., MacInnis, B., Kwiatkowski, D.P., 2012. Analysis of *Plasmodium falciparum* diversity in natural infections by deep sequencing. Nature 487, 375-379.

Marsolier, J., Perichon, M., DeBarry, J.D., Villoutreix, B.O., Chluba, J., Lopez, T., Garrido, C., Zhou, X.Z., Lu, K.P., Fritsch, L., Ait-Si-Ali, S., Mhadhbi, M., Medjkane, S., Weitzman, J.B., 2015. *Theileria* parasites secrete a prolyl isomerase to maintain host leukocyte transformation. Nature 520, 378-382.

Matsuba, T., Sugimoto, C., Onoe, S., Kawakami, Y., Iwai, H., Onuma, M., 1993. Changes in the hybridization patterns of populations of *Theileria sergenti* during infection. Vet Parasitol 47, 215-223.

McFadden, A., Gias, E., Heuer, C., Stevens McFadden, F.J., Pulford, D.J., 2016. Prevalence and spatial distribution of cattle herds infected with *Theileria orientalis* in New Zealand between 2012 and 2013. N Z Vet J 64, 55-59.

McFadden, A.M., Rawdon, T.G., Meyer, J., Makin, J., Morley, C.M., Clough, R.R., Tham, K., Mullner, P., Geysen, D., 2011. An outbreak of haemolytic anaemia associated with infection of *Theileria orientalis* in naive cattle. N Z Vet J 59, 79-85.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M., DePristo, M.A., 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. Genome Res 20, 1297-1303.

Mhadhbi, M., Naouach, A., Boumiza, A., Chaabani, M.F., BenAbderazzak, S., Darghouth, M.A., 2010. In vivo evidence for the resistance of *Theileria annulata* to buparvaquone. Vet Parasitol 169, 241-247.

Miotto, O., Almagro-Garcia, J., Manske, M., Macinnis, B., Campino, S., Rockett, K.A., Amaratunga, C., Lim, P., Suon, S., Sreng, S., Anderson, J.M., Duong, S., Nguon, C., Chuor, C.M., Saunders, D., Se, Y., Lon, C., Fukuda, M.M., Amenga-Etego, L., Hodgson, A.V., Asoala, V., Imwong, M., Takala-Harrison, S., Nosten, F., Su, X.Z., Ringwald, P., Ariey, F., Dolecek, C., Hien, T.T., Boni, M.F., Thai, C.Q., Amambua-Ngwa, A., Conway, D.J., Djimde, A.A., Doumbo, O.K., Zongo, I., Ouedraogo, J.B., Alcock, D., Drury, E., Auburn, S., Koch, O., Sanders, M., Hubbart, C., Maslen, G.,

Ruano-Rubio, V., Jyothi, D., Miles, A., O'Brien, J., Gamble, C., Oyola, S.O., Rayner, J.C., Newbold, C.I., Berriman, M., Spencer, C.C., McVean, G., Day, N.P., White, N.J., Bethell, D., Dondorp, A.M., Plowe, C.V., Fairhurst, R.M., Kwiatkowski, D.P., 2013. Multiple populations of artemisinin-resistant *Plasmodium falciparum* in Cambodia. Nature Genet 45, 648-655.

Morrison, W.I., 2009. Progress towards understanding the immunobiology of *Theileria* parasites. Parasitology 136, 1415-1426.

Muller, H.M., Reckmann, I., Hollingdale, M.R., Bujard, H., Robson, K.J., Crisanti, A., 1993. Thrombospondin related anonymous protein (TRAP) of *Plasmodium falciparum* binds specifically to sulfated glycoconjugates and to HepG2 hepatoma cells suggesting a role for this molecule in sporozoite invasion of hepatocytes. EMBO J 12, 2881-2889.

Nkhoma, S.C., Nair, S., Cheeseman, I.H., Rohr-Allegrini, C., Singlam, S., Nosten, F., Anderson, T.J., 2012. Close kinship within multiple-genotype malaria parasite infections. Proc Biol Sci 279, 2589-2598.

Nwakanma, D.C., Gomez-Escobar, N., Walther, M., Crozier, S., Dubovsky, F., Malkin, E., Locke, E., Conway, D.J., 2009. Quantitative detection of *Plasmodium falciparum* DNA in saliva, blood, and urine. J Infect Dis 199, 1567-1574.

Omar Abdallah, M., Niu, Q., Yu, P., Guan, G., Yang, J., Chen, Z., Liu, G., Wei, Y., Luo, J., Yin, H., 2016. Identification of piroplasm infection in questing ticks by RLB: a broad range extension of tick-borne piroplasm in China? Parasitol Res.

Onuma, M., Kakuda, T., Sugimoto, C., 1998. *Theileria* parasite infection in East Asia and control of the disease. Comp Immunol MIcrobiol Infect Dis 21, 165-177.

Onuma, M., Kubota, S., Kakuda, T., Sako, Y., Asada, M., Kabeya, H., Sugimoto, C., 1997. Control of *Theileria sergenti* infection by vaccination. Trop Anim Health Prod 29, 119S-123S.

Ota, N., Mizuno, D., Kuboki, N., Igarashi, I., Nakamura, Y., Yamashina, H., Hanzaike, T., Fujii, K., Onoe, S., Hata, H., Kondo, S., Matsui, S., Koga, M., Matsumoto, K., Inokuma, H., Yokoyama, N., 2009. Epidemiological survey of *Theileria orientalis* infection in grazing cattle in the eastern part of Hokkaido, Japan. J Vet Med Sci 71, 937-944.

Otto, T.D., Dillon, G.P., Degrave, W.S., Berriman, M., 2011. RATT: Rapid Annotation Transfer Tool. Nucleic acids research 39, e57.

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

Perera, P.K., Gasser, R.B., Anderson, G.A., Jeffers, M., Bell, C.M., Jabbar, A., 2013. Epidemiological survey following oriental theileriosis outbreaks in Victoria, Australia, on selected cattle farms. Vet Parasitol 197, 509-521.

Perera, P.K., Gasser, R.B., Firestone, S.M., Anderson, G.A., Malmo, J., Davis, G., Beggs, D.S., Jabbar, A., 2014. Oriental theileriosis in dairy cows causes a significant milk production loss. Parasit Vect 7.

Perera, P.K., Gasser, R.B., Firestone, S.M., Smith, L., Roeber, F., Jabbar, A., 2015a. Semiquantitative multiplexed tandem PCR for detection and differentiation of four *Theileria orientalis* genotypes in cattle. J Clin Microbiol 53, 79-87.

Perera, P.K., Gasser, R.B., Jabbar, A., 2015b. Assessment of sequence variability in a p23 gene region within and among three genotypes of the *Theileria orientalis* complex from south-eastern Australia. Ticks Tick Borne Dis 6, 123-128.

Petersen, T.N., Brunak, S., von Heijne, G., Nielsen, H., 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nature Methods 8, 785-786.

Plain, K.M., Marsh, I.B., Waldron, A.M., Galea, F., Whittington, A.M., Saunders, V.F., Begg, D.J., de Silva, K., Purdie, A.C., Whittington, R.J., 2014. High-throughput direct fecal PCR assay for detection of *Mycobacterium avium* subsp. *paratuberculosis* in sheep and cattle. J Clin Microbiol 52, 745-757.

Portugal, S., Carret, C., Recker, M., Armitage, A.E., Goncalves, L.A., Epiphanio, S., Sullivan, D., Roy, C., Newbold, C.I., Drakesmith, H., Mota, M.M., 2011. Host-mediated regulation of superinfection in malaria. Nature Med 17, 732-737.

Pulford, D.J., McFadden, A., Hamilton, J.S., Donald, J., 2016. Investigation of the index case herd and identification of the genotypes of *Theileria orientalis* associated with outbreaks of bovine anaemia in New Zealand in 2012. N Z Vet J 64, 21-28.

Richter, M., Rossello-Mora, R., 2009. Shifting the genomic gold standard for the prokaryotic species definition. Proc Natl Acad Sci USA 106, 19126-19131.

Riek, R.F., 1982. Epidemiology and transmission of *Theileria* sp. of cattle in Australia. Aust Vet J 59, 89-92.

Rogers, W.O., Malik, A., Mellouk, S., Nakamura, K., Rogers, M.D., Szarfman, A., Gordon, D.M., Nussler, A.K., Aikawa, M., Hoffman, S.L., 1992. Characterization of *Plasmodium falciparum* sporozoite surface protein 2. Proc Natl Acad Sci USA 89, 9176-9180.

Ros-Garcia, A., Nicolas, A., Garcia-Perez, A.L., Juste, R.A., Hurtado, A., 2012. Development and evaluation of a real-time PCR assay for the quantitative detection of *Theileria annulata* in cattle. Parasit Vect 5, 171.

Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4, 406-425.

Sako, Y., Sugimoto, C., Onuma, M., 1999. Expression of a major piroplasm surface protein of *Theileria sergenti* in sporozoite stage. J Vet Med Sci 61, 275-277.

Sasaki, M., Fujii, Y., Iwamoto, M., Ikadai, H., 2013. Effect of sex steroids on *Babesia microti* infection in mice. Am J Trop Med Hygiene 88, 367-375.

Schnittger, L., Katzer, F., Biermann, R., Shayan, P., Boguslawski, K., McKellar, S., Beyer, D., Shiels, B.R., Ahmed, J.S., 2002. Characterization of a polymorphic *Theileria annulata* surface protein (TaSP) closely related to PIM of *Theileria parva*: implications for use in diagnostic tests and subunit vaccines. Mol Biochem Parasitol 120, 247-256.

Sedlazeck, F.J., Rescheneder, P., von Haeseler, A., 2013. NextGenMap: fast and accurate read mapping in highly polymorphic genomes. Bioinformatics 29, 2790-2791.

Shaw, M.K., 1997. The same but different: the biology of *Theileria* sporozoite entry into bovine cells. Int J Parasitol 27, 457-474.

Shimizu, S., Suzuki, K., Nakamura, K., Kadota, K., Fujisaki, K., Ito, S., Minami, T., 1988. Isolation of *Theileria sergenti* piroplasms from infected erythrocytes and development of an enzyme-linked immunosorbent assay for serodiagnosis of *T. sergenti* infections. Res Vet Sci 45, 206-212.

Sivakumar, T., Hayashida, K., Sugimoto, C., Yokoyama, N., 2014. Evolution and genetic diversity of *Theileria*. Infect Genet Evol 27, 250-263.

Sivakumar, T., Yoshinari, T., Igarashi, I., Kothalawala, H., Abeyratne, S.A.E., Vimalakumar, S.C., Meewewa, A.S., Kuleswarakumar, K., Chandrasiri, A.D.N., Yokoyama, N., 2013. Genetic diversity within *Theileria orientalis* parasites detected in Sri Lankan cattle. Ticks Tick Borne Dis 4, 235-241.

Stamatakis, A., 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22, 2688-2690.

Stewart, N.P., de Vos, A.J., McGregor, W., Shiels, I., 1987. Haemaphysalis humerosa, not H. longicornis, is the likely vector of *Theileria buffeli* in Australia. Aust Vet J 64, 280-282.

Stewart, N.P., de Vos, A.J., Shiels, I.A., Jorgensen, W.K., 1989. Transmission of *Theileria buffeli* to cattle by Haemaphysalis bancrofti fed on artificially infected mice. Vet Parasitol 34, 123-127.

Stothard, P., 2000. The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences. Biotechniques 28, 1102, 1104.

Sugimoto, C., Kawazu, S., Kamio, T., Fujisaki, K., 1991. Protein analysis of *Theileria sergenti/buffeli/orientalis* piroplasms by two-dimensional polyacrylamide gel electrophoresis. Parasitology 102 Pt 3, 341-346.

Takemae, H., Sugi, T., Kobayashi, K., Murakoshi, F., Recuenco, F.C., Ishiwa, A., Inomata, A., Horimoto, T., Yokoyama, N., Kato, K., 2014a. Analyses of the binding between *Theileria orientalis* major piroplasm surface proteins and bovine red blood cells. Vet Rec 175, 149.

Takemae, H., Sugi, T., Kobayashi, K., Murakoshi, F., Recuenco, F.C., Ishiwa, A., Inomata, A., Horimoto, T., Yokoyama, N., Kato, K., Kato, K., 2014b. Interaction between *Theileria orientalis* 23-kDa piroplasm membrane protein and heparin. Jpn J Vet Res 62, 17-24.

Tamura, K., Nei, M., Kumar, S., 2004. Prospects for inferring very large phylogenies by using the neighbor-joining method. Proc Natl Acad Sci USA 101, 11030-11035.

Tanaka, M., Ohgitani, T., Okabe, T., Kawamoto, S., Takahashi, K., Onuma, M., Kawakami, Y., Sasaki, N., 1990. Protective effect against intraerythrocytic merozoites of *Theileria sergenti* infection in calves by passive transfer of monoclonal antibody. Nihon Juigaku Zasshi 52, 631-633.

Tanaka, M., Onoe, S., Matsuba, T., Katayama, S., Yamanaka, M., Yonemichi, H., Hiramatsu, K., Baek, B.K., Sugimoto, C., Onuma, M., 1993. Detection of *Theileria sergenti* infection in cattle by polymerase chain reaction amplification of parasite-specific DNA. J Clin Microbiol 31, 2565-2569.

Terada, Y., Ishida, M., Yamanaka, H., 1995. Resistibility to *Theileria sergenti* infection in Holstein and Japanese Black cattle. J Vet Med Sci 57, 1003-1006.

Tritt, A., Eisen, J.A., Facciotti, M.T., Darling, A.E., 2012. An integrated pipeline for de novo assembly of microbial genomes. PloS One 7, e42304.

Uilenberg, G., 2011. *Theileria sergenti*. Vet Parasitol 175, 386.

Vaughan, A.M., Kappe, S.H., 2012. Malaria vaccine development: persistent challenges. Curr Opin Immunol 24, 324-331.

Veron, V., Simon, S., Carme, B., 2009. Multiplex real-time PCR detection of *P. falciparum, P. vivax* and *P. malariae* in human blood samples. Exp Parasitol 121, 346-351.

Wang, C.T., Kubota, S., Kakuda, T., Kuo, C.C., Hsu, T.L., Onuma, M., 1998. Survey of *Theileria* parasite infection in cattle in Taiwan. J Vet Med Sci 60, 253-255.

Wang, D., Zhang, Y., Zhang, Z., Zhu, J., Yu, J., 2010a. KaKs_Calculator 2.0: a toolkit incorporating gamma-series methods and sliding window strategies. Genomics, Proteomics & Bioinformatics 8, 77-80.

Wang, L.X., Zhao, J.H., He, L., Liu, Q., Zhou, D.N., Zhou, Y.Q., Zhao, J.L., 2010b. An indirect ELISA for detection of *Theileria sergenti* antibodies in water buffalo using a recombinant major piroplasm surface protein. Vet Parasitol 170, 323-326.

Weir, W., Ben-Miled, L., Karagenc, T., Katzer, F., Darghouth, M., Shiels, B., Tait, A., 2007. Genetic exchange and sub-structuring in *Theileria annulata* populations. Mol Biochem Parasitol 154, 170-180.

Weir, W., Karagenc, T., Baird, M., Tait, A., Shiels, B.R., 2010. Evolution and diversity of secretome genes in the apicomplexan parasite *Theileria annulata*. BMC genomics 11, 42.

Yang, Y., Mao, Y., Kelly, P., Yang, Z., Luan, L., Zhang, J., Li, J., El-Mahallawy, H.S., Wang, C., 2014. A pan-*Theileria* FRET-qPCR survey for *Theileria* spp. in ruminants from nine provinces of China. Parasit Vectors 7, 413.

Zakimi, S., Kim, J.Y., Oshiro, M., Hayashida, K., Fujisaki, K., Sugimoto, C., 2006. Genetic diversity of benign *Theileria* parasites of cattle in the Okinawa Prefecture. J Vet Med Sci 68, 1335-1338.

Zhou, J., Fukumoto, S., Jia, H., Yokoyama, N., Zhang, G., Fujisaki, K., Lin, J., Xuan, X., 2006. Characterization of the *Babesia gibsoni* P18 as a homologue of thrombospondin related adhesive protein. Mol Biochem Parasitol 148, 190-198.

Ziam, H., Kelanamer, R., Aissi, M., Ababou, A., Berkvens, D., Geysen, D., 2015. Prevalence of bovine theileriosis in North Central region of Algeria by real-time polymerase chain reaction with a note on its distribution. Trop Anim Health Prod 47, 787-796.

Zimin, A.V., Delcher, A.L., Florea, L., Kelley, D.R., Schatz, M.C., Puiu, D., Hanrahan, F., Pertea, G., Van Tassell, C.P., Sonstegard, T.S., Marcais, G., Roberts, M., Subramanian, P., Yorke, J.A., Salzberg, S.L., 2009. A whole-genome assembly of the domestic cow, *Bos taurus*. Genome Biol 10, R42.