Establishment of a mouse–tick infection model for *Theileria orientalis* and analysis of its transcriptome

Kyoko Hayashida[a,b], Rika Umemiya-Shirafuji[a], Thillaiampalam Sivakumar[a], Junya Yamagishi[b,d], Yutaka Suzuki[c], Chihiro Sugimoto[b,d], Naoaki Yokoyama[a]

[a]*National Research Center for Protozoan Diseases, Obihiro University of Agriculture and Veterinary Medicine, Obihiro, Hokkaido 080-8555, Japan*

[b]*Division of Collaboration and Education, Hokkaido University Research Center for Zoonosis Control, Sapporo, Hokkaido 001-0020, Japan*

[c]*Department of Computational Biology and Medical Sciences, Graduate School of Frontier Sciences, the University of Tokyo, Kashiwa, Chiba, Japan*

[d]*Global Station for Zoonosis Control, GI-CoRE, Hokkaido University, Sapporo, Hokkaido 001-0020, Japan*

* Corresponding author. National Research Center for Protozoan Diseases, Obihiro University of Agriculture and Veterinary Medicine, Nishi 2-13, Inada-cho, Obihiro, Hokkaido 080-8555, Japan.

Tel.: +81-155-49-5649; Fax: +81-155-49-5643.

E-mail address: yokoyama@obihiro.ac.jp (N. Yokoyama)

Abstract

Oriental theileriosis caused by *Theileria orientalis* is an economically significant disease in cattle farming. The lack of laboratory animal models and in vitro culture systems is a major obstacle in the drive to better understand the biology of this parasite. Notably, research on the sporozoite stage of *T. orientalis* has rarely been undertaken, although such investigations are of paramount importance for vaccine development based on blocking sporozoite invasion of its host animals. In the present study, we established a mouse–tick infection model for propagating *T. orientalis* in mice and for producing the sporozoite stage in tick salivary glands. Splenectomized severe combined immunodeficient (SCID) mice transfused with bovine erythrocytes were infected with *T. orientalis*. The larval ticks of *Haemaphysalis longicornis* were then fed on the *T. orientalis*–infected mice. The piroplasm and sporozoite stages were microscopically observed in the mouse blood and nymphal salivary glands, respectively. The transcriptomics data generated from the piroplasm and sporozoite stages revealed a stage-specific expression pattern for the parasite genes. The mouse–tick infection model and the transcriptomics data it has provided will contribute to a better understanding of *T. orientalis* biology and will also provide much needed information for the design of effective control measures targeting oriental theileriosis.

Keywords: *Theileria orientalis*, Sporozoite, Infection model, Tick, SCID mouse, p67, Transcriptome

1. Introduction

Theileria orientalis, a non-transforming *Theileria* parasite, infects cattle, buffalo and yaks. In most of the countries where *T. orientalis* is endemic, it has long been considered a relatively benign parasite, unlike the malignant *Theileria* spp. *Theileria parva* and *Theileria annulata*, which infect millions of cattle in Africa and Asia (Irvin, 1987). However, under certain circumstances, *T. orientalis* infection causes anemia, jaundice, abortion, production losses and mortality in cattle, making oriental theileriosis an economically significant disease of grazing cattle in regions where this parasite is endemic. The outbreaks of hemolytic anemia that are associated with *T. orientalis* infections have been reported in southeastern Australia (Izzo et al., 2010) and New Zealand since 2006, in particular (McFadden et al., 2011). The outcomes of such infections usually depend on environmental stress factors, the level of acquired immunity in the affected cattle, and herd genetics (Terada et al., 1995). The virulence in *T. orientalis* is considered to be genotype related and is classified according to the major piroplasm surface protein (MPSP) or its 18S rRNA gene sequences. For instance, clinical theileriosis is common among cattle in countries where the *T. orientalis* type 2 (Ikeda) genotype is endemic. Thus, type 2 *T. orientalis* is considered to be more virulent than the other genotypes of this species (Hayashida et al., 2012; Sivakumar et al., 2014). *Theileria orientalis* is transmitted mainly by *Haemaphysalis longicornis* ticks, although transmission by other tick species cannot be ruled out (Yokoyama et al., 2012). Similar to other *Theileria* spp., *T. orientalis* is maintained in tick

52 vectors through transstadial persistence as the tick molts from one stage to another (larva to nymph and
53 nymph to adult), but transovarial transmission has not been reported (Higuchi, 1986).

54 Malignant theileriosis caused by *T. parva* and *T. annulata* can be prevented and controlled with
55 live vaccines. A live sporozoite vaccine, known as the Muguga cocktail vaccine, has been used in several
56 African countries to control East Coast fever (ECF) caused by *T. parva*. The vaccination methodology is
57 known as an "Infection and Treatment Method" because the Muguga cocktail, which is composed of the live
58 sporozoites derived from three different *T. parva* isolates, is administrated simultaneously with a long-acting
59 oxytetracycline formulation (Nene et al., 2016). However, the wide use of this live vaccine is limited due to
60 its time-consuming production process and concerns about the risk of introducing a "foreign" parasite
61 population into other countries, where the vaccine strains might be transmitted to unvaccinated cattle and
62 produce new parasites with genetic mosaics (McKeever, 2007; Oura et al., 2007). There is thus an urgent
63 need for a recombinant vaccine that is both suitable for mass production and addresses the drawbacks of
64 using a live vaccine. p67, a sporozoite surface glycoprotein in *T. parva*, is one of the most promising vaccine
65 candidates for ECF (Musoke et al., 1992; Nene et al., 1996; Honda et al., 1998; Bishop et al., 2003; Kaba et
66 al., 2004; Morrison and McKeever, 2006). The p67 ortholog antigen, SPAG-1, which has been identified in
67 *T. annulata*, has also been shown to be effective at inducing protective immunity against tropical theileriosis
68 in cattle (Boulter and Hall, 1999). Furthermore, a recent genome project bioinformatically identified a *p67*
69 orthologous gene (TOT_030000930) in *T. orientalis* (Hayashida et al., 2012). It would be of interest to

measure its expression profile to evaluate its potential for the development of a recombinant vaccine against *T. orientalis*.

A thorough understanding of the biology of *Theileria* in its host animals and tick vectors is vital for the development of effective control measures. However, the biology of *T. orientalis* is poorly understood, partly because an experimental laboratory system for it is lacking. There is no in vitro culture system available for any developmental stage of this parasite, and only natural hosts are known to be susceptible to it (Irvin, 1987). However, Tsuji et al. (1992) reported that *T. orientalis* could proliferate at high parasitemia in severe combined immunodeficient (SCID) mice transfused red blood cells (RBCs) from a bovine source (Bo-RBC-SCID) (Tsuji et al., 1992). Based on this study, we employed Bo-RBC-SCID mice as an infection model for *T. orientalis* and induced the sporozoite stage in *H. longicornis* ticks fed on infected mice. The aim was to establish a mouse–tick laboratory infection model, and to generate sporozoite transcriptomic data with which to gain a better understanding of parasite biology and identify new control strategies against oriental theileriosis.

2. Materials and methods

2.1. Animal ethics approval

86 The protocols for cattle blood sampling, tick feeding and experimental infections in mice were
87 approved by the Animal Care and Use Committee, Obihiro University of Agriculture and Veterinary
88 Medicine, Japan (Approval number: 25-148/ 25-78).
89
90 *2.2. Theileria orientalis–infected RBCs from cattle*
91 *Theileria orientalis*–infected cattle blood was obtained from a naturally infected cow in
92 Nakasatsunai village, Hokkaido, Japan (May 2014). The blood samples were collected from three Holstein
93 cattle, which were positive for *T. orientalis* by microscopy, using EDTA as the anticoagulant. The DNA
94 samples extracted from the blood samples using a QIAamp DNA Blood Mini kit (Qiagen, Japan) were
95 subjected to previously described type-specific PCR assays to identify *T. orientalis MPSP* genotypes 1–5,
96 which are known to be endemic in Japan (Yokoyama et al., 2011). Subsequently, a blood sample was
97 collected using citric acid as the anticoagulant from a cow that was only positive for the type 2 genotype (not
98 genotypes 1, 3, 4 or 5). The whole blood was washed three times with Vega y Marinez PBS solution (VyM
99 buffer) (Vega et al., 1985), and the purified RBCs were used for the Bo-RBC-SCID infections. Washed
100 RBCs were also prepared from a non-infected cow, after which those were stored in VyM buffer at 4°C for
101 transfusion purposes.
102
103 *2.3. Theileria orientalis infection in Bo-RBC-SCID mice*

104 The Bo-RBC-SCID mice were prepared according to the method of Tsuji et al. (1992). SCID
105 mice lack functional T cells and B cells but still retain functional macrophages that will reject xenogeneic
106 grafts. Therefore, splenectomized SCID mice were used in the present study to replace a large proportion of
107 the circulating murine erythrocytes with bovine erythrocytes by transfusion. Briefly, 10 7-week-old female
108 SCID mice (C.B-17 scid/scid, CLEA, Japan) were anesthetized with isoflurane (Intervet, Japan) and then
109 splenectomized. One week later, the splenectomized mice were infected intraperitoneally with 500 µl of
110 fresh or cryopreserved *T. orientalis*–infected bovine RBCs. Each mouse was also transfused intraperitoneally
111 with 500 µl of packed RBCs every 3–4 days, to partially replace the circulating mouse RBCs with bovine
112 RBCs. Peripheral blood smears made from the tail blood of the Bo-RBC-SCID mice every 3–4 days were
113 stained with 3% Giemsa solution (pH 7.2) (Merk, Japan) and/or acridine orange fluorescent solution (Wako,
114 Japan), followed by light microscopy or fluorescence microscopy analysis, respectively, to monitor the
115 parasitemia levels.
116 At 17–31 days p.i., whole blood was collected from the infected Bo-RBC-SCID mice by cardiac
117 puncture using heparin as the anticoagulant, and the piroplasms were purified by saponin treatment before
118 total RNA was extracted from them, as described in section 2.6.
119
120 *2.4. Theileria orientalis–infected H. longicornis ticks*

121 The parthenogenetic Okayama strain of *H. longicornis* (Fujisaki, 1978) at Obihiro University of
122 Agriculture and Veterinary Medicine has been maintained by feeding the ticks on rabbits. *H. longicornis*
123 larvae were allowed to infest the *T. orientalis*–infected Bo-RBC-SCID mice that showed various degrees of
124 parasitemia (Table 1). After 4–8 days of feeding, the engorged ticks were collected and reared in a clean
125 glass bottle at 25°C with saturated humidity in the dark. After approximately 30 days, all of the larvae
126 molted into nymphs. The newly emerged nymphs were then allowed to infest a non-infected BALB/c mouse
127 (CLEA) for 3 days to stimulate sporozoite propagation and maturation, after which they were detached from
128 the host's skin with forceps and used for further experiments. The experimental procedure is shown in Fig. 1.
129
130 *2.5. Immunostaining of T. orientalis in the salivary glands of infected ticks*
131 Immunostaining was performed on the *T. orientalis* sporozoites by probing for MPSP. Briefly,
132 each infected nymphal tick was dissected with a 27-G needle PBS, and the salivary glands were transferred
133 onto an adhesive microscope slide (MAS-coated glass slide, Matsunami Glass, Japan). The slide contents
134 were fixed with 4% paraformaldehyde in PBS at 4°C for 15 min, permeabilized with 0.2% Triton X-100 for
135 10 min at room temperature, blocked with 5% skimmed milk in PBS for 1 h at room temperature, and then
136 reacted with an anti-MPSP antibody (4G10: Type-2 genotype-specific monoclonal antibody ×250) (Iwasaki
137 et al., 1998) overnight at 4°C, and Alexa 488-conjugated anti-mouse IgG (×500, Invitrogen, Japan) was
138 added. Finally, Vectashield mounting medium containing DAPI (Vector Laboratories, Burlingame, CA,

139 USA) was added to the tick salivary gland-mounted slides and images of the glands were captured using the
140 Keyence BZ-9000 (Keyence, Japan) all-in-one microscope.

141 The scutum-peeled whole bodies of the infected ticks were fixed with 4% paraformaldehyde in
142 PBS (4°C, overnight), equilibrated sequentially in 5, 10, 15 and 30% sucrose in PBS at 4°C for 4 to 12 h, and
143 then embedded and frozen in Tissue-Tek O.C.T. compound (Sakura Finetek, Japan). The frozen sections
144 were prepared at 10 μm thicknesses using a cryostat (Leica CM 3050S; Leica, Germany). The staining
145 procedures after blocking were the same as those used for mounting the whole salivary glands on slides, as
146 described above. The images were captured using a confocal laser-scanning microscope (LSM780, Zeiss
147 Microsystems, Germany).

148

149 *2.6. Theileria orientalis cDNA library preparation from T. orientalis–infected cattle, Bo-RBC-SCID mice,*
150 *tick salivary glands, and Illumina sequencing*

151 To isolate RNA from the piroplasm stage of *T. orientalis*, 25 ml and 1 ml of whole blood were
152 collected from *T. orientalis*–infected cattle and Bo-RBC-SCID mice, respectively. The samples were diluted
153 in two volumes of PBS, and the host leukocytes were removed by two rounds of filtration using a
154 Plasmodipur filter (EuroProxima, Netherlands). The filtered RBCs were lysed with ice-cold 0.05% saponin
155 for 30 min, and then pelleted down by centrifugation. The purified piroplasms were subjected to total RNA
156 extraction with TRIzol (Invitrogen, Carlsbad, CA, USA), according to the manufacturer's instructions. Four

of the *T. orientalis*–infected nymphal ticks (group #5 in Table 1) were fed on a non-infected BALB/c mouse to stimulate sporozoite maturation, removed after 48 to 72 h, and then dissected. The dissected salivary glands were collected, and total RNA was extracted from them using a Nucleospin Tissue XS kit (Macherey-Nagel, Germany), in accordance with the manufacturer's instructions, except that the carrier RNA step was omitted.

The cDNA libraries for transcriptomic analysis were prepared using SMARTer Ultra Low Input RNA from Sequencing –v3 (Clontech Laboratories, Mountain View, CA, USA) and the Nextera DNA Library Preparation Kit (Illumina K.K., Japan). The libraries were analyzed in the Hiseq 2500 36-bp Single End Rapid Mode setting, using two lanes and a multiplex index. The reads obtained were converted to a fastq file and then deposited in the DDBJ Sequence Read Archive (https://www.ddbj.nig.ac.jp/dra/, Accession number: DRA 006389).

2.7. Bioinformatics analysis

The Illumina reads were mapped against *T. orientalis* genome data (AP011946-AP011949 for Chr1–4, AB499090 for Mt, AP-11950 for Apicoplast) using the CLC genomics workbench 8.5.1 software (Qiagen Aarhus A/S, Denmark) with default parameters, and the gene expression levels obtained were normalized as reads per kilobase per million (RPKM). In total, four data sets were produced: i) sporozoite 48 h – a sample consisting of sporozoites isolated from tick salivary glands 48 h post-infestation; ii) sporozoite 72 h – a

sample collected at 72 h post-infestation; iii) Piroplasm Cattle – a sample consisting of piroplasms isolated from infected cattle blood; and iv) piroplasm SCID – a sample consisting of piroplasms isolated from infected Bo-RBC-SCID mouse blood. The data sets were then compared using the CLC Genomics Platform with the EdgeR package (Robinson et al., 2010) and *P* values were adjusted using Bonferroni correction to identify differentially expressed genes (DEGs). DEGs were selected with a false discovery rate adjusted *P* value of <0.01. The expression heat map for the extracted DEGs was generated with the heatmap 2 function of the R package gplots (Warnes et al., 2016, gplots: Various R programming tools for plotting data, (https://cran.r-project.org/web/packages/gplots/index.html)

2.8. Sequencing of ToSPAG, a p67 orthologous gene of T. orientalis, using cDNA prepared from T. orientalis sporozoites

Total RNA isolated from tick salivary glands was subjected to reverse transcription PCR (RT-PCR) to amplify *ToSPAG* (*T. orientalis* sporozoite surface antigen, TOT_030000930), which is a possible *p67* orthologous gene in *T. orientalis*, using the Qiagen One-step RT-PCR kit (Qiagen) with the following four sets of primer pairs: 5′-CGTAGAGTCACAAAGGCATAATAC-3′ and 5′-AACCGTCGTCGCCTGACTAAC-3′, 5′-TCAAAGAACAGGAGGTGGCTTAG-3′ and 5′-TACTCTTCATTCGGGTCAATC-3′, 5′-AGATAAAGGAGAAGATGGTGGCAG-3′ and 5′-CAATGGGTCCAGATGCAGATT-3′, 5′-AATCTGCATCTGGACCCATTG-3′ and 5′-CATCATCACTGGAACCACCGTG-3′.

193 The cDNA sequence was analyzed on a 3130xl Genetic Analyzer (Applied Biosystems, USA).
194 The coding sequence (CDS) was determined and submitted to the GenBank database (accession number:
195 LC342224), and the putative signal peptide cleavage sites in the CDS were determined using the SignalP 4.1
196 Server (http://www.cbs.dtu.dk/services/SignalP/).
197
198 *2.9. Quantitative RT–PCR (qRT-PCR) assay*
199 The total RNA was extracted from four tick salivary glands in group #5 (Table 1), which was a
200 different batch of ticks from that used for Illumina library preparation as described in section 2.6. *MPSP* and
201 *ToSPAG* transcription levels were determined during the different feeding phases of the tick stage using a
202 Thunderbird SYBR qPCR (TOYOBO, Japan) assay and cDNA synthesized with the TAKARA RNA PCR
203 kit (AMV) Ver.3.0 (Takara, Japan). The expression values were normalized against the value of ribosomal
204 protein *P0*, which is known to be expressed in *H. longicornis* ticks (Gong et al., 2008). Standard curves were
205 constructed using the *MPSP*, *ToSPAG,* and *P0* genes cloned into pGEM-T easy vectors (Promega, Japan).
206 The qRT-PCR assays were performed in triplicate using the ABI Prism 7900 HT Sequence Detection System
207 (Applied Biosystems). Cycling conditions were as follows: initial denaturation at 95°C for 60 s, followed by
208 40 cycles each containing a denaturation step at 95°C for 15 s and a combined step of annealing and
209 extension at 60°C for 60 s. Melting curve analysis was performed after amplification to confirm the
210 specificity of the amplified products. The primer sequences for the qRT-PCR assays were as follows: Tick

211 *PO*: 5′-CTCCATTGTCAACGGTCTCA-3′ and 5′-TCAGCCTCCTTGAAGGTGAT-3′ (Gong et al., 2008);

212 *ToSPAG*: 5′-GTAGTCAAGGTCAACAACACCAACT-3′ and

213 5′-TGTTTTCACCATCTCTCCTTACTTC-3′; *MPSP*: 5′- ATTGCTGTTCAAGAAGAAGACTGAC-3′ and

214 5′- TCTTCTCCTTCCATACAACTTCATC-3′.

215

216 2.10. Data accessibility

217 Associated data has been deposited in Mendeley Data.

218 https://data.mendeley.com/datasets/bth7xjr2vs/draft?a=3ca0c3c2-52e6-4f6e-9743-0affebea336d

219

220 **3. Results**

221 *3.1. Theileria orientalis infections in Bo-RBC-SCID mice*

222 Bo-RBC-SCID mice inoculated with *T. orientalis*–infected bovine RBCs showed high parasitemia

223 levels in their peripheral blood, as reported previously (Hagiwara et al., 1993). The morphology of the

224 piroplasms in the RBCs was pleomorphic in nature which included rod, comma, ovoid, Y and tetrad shapes

225 (Fig. 2A). At the beginning of the infection when parasitemia levels were low, rod or comma shapes were

226 predominant, whereas later in the infection, the number of oval and ovoid shapes increased, as observed in

227 naturally infected cattle (Gettinby, 1993). Parasite nuclei and cytoplasms were visualized by acridine orange

228 staining and emitted a green/yellow and orange light, respectively (Fig. 2B). The Y- and tetrad-shaped

parasites within the RBCs displayed multiple nuclei, suggesting that they were multiplying by binary fission, resulting in multiple daughter cells (Fig. 2B). A bar-like structure and a veil-like body, both of which characterize bovine theileriosis caused by *T. orientalis*, were observed in infected RBCs from the cattle and Bo-RBC-SCID mice (Fig. 2A and 2C), suggesting that the morphological changes occurring within the *T. orientalis*–infected RBCs were similar in both host species.

Immunostaining with a monoclonal antibody specific for MPSP type 2 (Zhuang et al., 1994) confirmed that the *T. orientalis* detected in the Bo-RBC-SCID mice were of the type 2 genotype that the mice were originally infected with (Fig. 2D). Microscopic monitoring of blood smears indicated a gradual increase in parasitemia in *T. orientalis*–infected Bo-RBC-SCID mice (Fig. 2E).

3.2. Theileria orientalis in the salivary glands of H. longicornis nymphs

To investigate whether *T. orientalis* in Bo-RBC-SCID mice could be ingested by ticks and could develop into sporozoite-stage parasites in their salivary glands, the larvae from *H. longicornis* were fed on the infected mice, showing different parasitemia levels at various time points (Fig. 2E, black triangle). The ticks were allowed to molt and were then fed on a non-infected BALB/c mouse to stimulate sporozoite maturation (Fig. 1). Subsequent immunostaining with the anti-MPSP antibody confirmed the presence of parasites in the acinar cells of the salivary glands from the nymphs (Fig. 3A). Furthermore, the infected acini

246 in the salivary glands appeared to have a multinucleate syncytium, indicating the presence of multiple
247 sporozoites in the cytoplasm (Fig. 3B).

248 The nymph molting rate from all available infected larvae was calculated. Additionally, 10
249 nymphs previously fed on *T. orientalis*–infected Bo-RBC-SCID mice were randomly selected to measure the
250 salivary gland infection rate by MPSP immunostaining (Table 1). The molting rate ranged from 64% to 96%
251 (average 85%). Not all of the nymphs that emerged from the larvae contained detectable sporozoites, and the
252 rate of sporozoite infection ranged from 10% to 40% (average 25%).

253 After the ticks had detached, three of the infected mice were subjected to a second round of tick
254 feeding (group #5, group #10, and group #15; labeled 2nd in Table 1). The molting rates did not show a clear
255 correlation between the tick groups that had fed on the Bo-RBC-SCID mice at the two different time points,
256 suggesting that the timing of tick feeding, as well as piroplasma parasitemia in the infected mice, had no
257 apparent effect on tick fitness.

258 The tick infectivity could be lost upon parasite passage in mice. To address this concern, frozen
259 stocks of piroplasms sourced from the *T. orientalis*–infected Bo-RBC-SCID mice were injected into naive
260 Bo-RBC-SCID mice, which were then infested with tick larvae. The emerging nymphs (group #16 and group
261 #17 in Table 1) still contained detectable sporozoites, suggesting retention of parasite tick infectivity even
262 after piroplasm passaging in mice.

263

3.3. Expression of ToSPAG, a p67/SPAG-1 orthologous gene, in T. orientalis sporozoites

TOT_030000930, a putative gene in *T. orientalis,* is suspected to be an ortholog of *p67* and *SPAG-1* in *T. parva* and *T. annulata*, respectively, due to some similarity in amino acid composition and genomic location (Hayashida et al., 2012; Sivakumar et al., 2014). In the present study, we confirmed that TOT_030000930 was transcribed in *T. orientalis* sporozoites and designated the gene and its protein product as *ToSPAG* and ToSPAG, respectively. The cloned *ToSPAG* cDNA, which contains six exons and is 5,589 bp in length, encodes a 1,862 amino acid polypeptide with a predicted molecular mass of 194 kDa. The cloned *ToSPAG* CDS (LC342224) differs from TOT_030000930 in its exon–intron boundary, perhaps because the database genes were predicted computationally by using the *T. orientalis* expression sequence tag (EST) pair gene models from the piroplasm stage (Supplementary Fig. S1). The *ToSPAG* gene sequenced in the present study contains 13 single nucleotide polymorphism mutations that result in seven amino acid changes, compared with the genome database sequence (TOT_030000930). These differences might result from inconsistencies between the sources of *T. orientalis* used for the genome project and the present study, although the parasites used in both investigations were of the type 2 genotype.

Despite the relatively low amino acid similarities between ToSPAG and p67/SPAG-1, they share the same Pfam domain PF05642, which was identified as a sporozoite P67 surface antigen domain, and an N-terminus 17 amino acid sequence, which is predicted to be an endoplasmic reticulum (ER) signal sequence (Supplementary Fig. S1). The qRT-PCR results revealed that the *ToSPAG* gene expression level peaked at

48 h, from the beginning of nymphal feeding on non-infected BALB/c mice. ToSPAG transcription was still active at 72 h, although the level of transcription was decreased compared with that at 48 h. By contrast, *MPSP* gene transcription peaked at 72 h, as reported previously (Sako et al., 1999) (Fig. 4), suggesting a difference in peak expression between these two possible surface antigens.

3.4. Transcriptome analyses of piroplasms from cattle, Bo-RBC-SCID mice, and sporozoites from ticks

Using two lanes of the Illumina Hiseq2000 system, we obtained 33–102 million reads in total for each dataset, and 28.6 million, 31.0 million, 0.7 million, and 2.2 million reads for Piroplasm SCID, Piroplasm Cattle, Sporozoite 48 h, and Sporozoite 72 h, respectively, which uniquely mapped to the 4,001 annotated genes in the *T. orientalis* Shintoku genome (Hayashida et al., 2012). Because our samples, especially the sporozoite samples that were prepared from whole salivary glands, contained a considerable number of host reads, the mapped read numbers for Piroplasm SCID, Piroplasm Cattle, Sporozoite 48 h, and Sporozoite 72 h were 51.0%, 30.4%, 2.1% and 6.5% of the total obtained reads, respectively (Supplementary Table S1). Highly expressed genes were selected for each data set by RPKM values up to rank 50, and the top 12 are listed in Tables 2 and 3. The full list of RPKM values is provided in Supplementary Table S2. In Piroplasm SCID and Piroplasm Cattle alike, the hypothetical ER signal sequence-positive TOT_030000643 gene was expressed most abundantly. In OrthoMCL analysis, this gene was classified as a member of the PiroF0000037 gene family, which consists of seven genes in *T. orientalis*, while four and zero homologous

300 genes were detected in *T. parva* and *T. annulata*, respectively, suggesting that this gene family might have
301 expanded preferentially in the *T. orientalis* genome (Hayashida et al., 2012). Furthermore, five and three of
302 the seven genes in the PiroF0000037 family were highly expressed in the piroplasm stage of *T. orientalis* in
303 cattle and SCID mice, respectively (Table 2), and most were predicted to be surface or secretory proteins,
304 suggesting essential and species-specific roles for these genes in the erythrocytic stages of *T. orientalis*. In
305 the sporozoite stage, both at the 48 h and 72 h time points, a known conserved hypothetical protein
306 TOT_030000930, now referred to as ToSPAG, was detected as an abundantly expressed gene (Table 3). The
307 expression of *ToSPAG* and *MPSP* in the sporozoite stage was also confirmed by qRT-PCR in different
308 biological replicates, and the expression pattern correlated with the transcriptomic data (Fig. 4,
309 Supplementary Table S2).

310 We then compared the mRNA gene expression levels between the piroplasm and sporozoite
311 stages. As a result of the lack of biological replicates, we analyzed DEGs using related replicates as follows:
312 one sample consisted of Piroplasm Cattle and Piroplasm SCID, and the other sample contained Sporozoite
313 48 h and Sporozoite 72 h. In total, 23 and 30 genes were identified that were expressed specifically at the
314 sporozoite and piroplasm stages, respectively (Fig. 5 and Supplementary Table S3). Most of the genes that
315 showed stage-specific expression were categorized within the highly expressed top 50 genes (Fig. 5,
316 Supplementary Table S3), suggesting that they may play pivotal roles at each developmental stage. The

317 transcriptomic analysis also suggested that ToSPAG was a highly expressed sporozoite-specific antigen in *T.*

318 *orientalis* (Table 3, Fig. 5, Supplementary Table S3).

319

320 **4. Discussion**

321 Investigating the developmental stages of *T. orientalis* in host animals and tick vectors is vital for

322 understanding its biological behavior. Currently, there is no established technology to culture any of the

323 developmental stages of *T. orientalis*, and ruminants are the only known mammalian hosts that can be

324 infected with this pathogen. Therefore, researchers have been compelled to rely on cattle and ticks that are

325 infected with *T. orientalis*, either naturally or experimentally, to isolate the blood and tick stages of the

326 parasite. A previous study showed that *T. orientalis* actively proliferates in SCID mice transfused with

327 bovine RBCs (Tsuji et al., 1992). Here, we employed the SCID mouse model to propagate *T. orientalis,* and

328 confirmed the active proliferation of piroplasms within the RBCs as early as 7 days p.i., a finding consistent

329 with a previous observation in cattle (Kawamoto et al., 1990). The presence of multiple piroplasms within

330 individual RBCs indicated that the piroplasms had multiplied by binary fission, which is common among

331 *Babesia* spp., *T. annulata*, and *Theileria equi,* but not for *T. parva*. The dividing forms were more common

332 in the infected Bo-RBC-SCID mice than in the cattle from which the parasites were obtained for the mouse

333 infections. This suggests that *T. orientalis* proliferation was more active in the Bo-RBC-SCID mice than in

334 cattle, a result possibly caused by the absence of host immune pressure in the former. Nevertheless, our

observations indicated that the parasite morphology and extra-parasitic structures (veil and bar) within the RBCs of the Bo-RBC-SCID mice did not differ from those observed in the RBCs of *T. orientalis*–infected cattle. The crystalline structure called the "veil" and the membrane structure known as the "bar" (Uilenberg, 1981) are formed in the cytoplasm of erythrocytes infected with piroplasms of several non-transforming *Theileria* spp. including *Theileria taurotragi*, *T. ovis*, but not in *T. parva* and *T. annulata*. Because the veil contained hemoglobin, it is believed to be part of the detoxification system (Sugimoto et al., 1992).

In our experiments, we also evaluated for the first time whether ticks fed on *T. orientalis*–infected Bo-RBC-SCID mice could maintain the complete invertebrate vector life cycle of the parasite. Similar to other *Theileria* spp., when *T. orientalis* is acquired by ticks during a blood meal it undergoes several developmental stages, culminating in the colonization of the tick salivary glands by sporozoites (Sasaki et al., 1990; Takahashi et al., 1993). Following the feeding of larvae on infected mice, sporozoites were successfully retrieved from the resulting nymphal salivary glands. This indicated that the piroplasms from infected mouse RBCs were capable of undergoing the complete life cycle in ticks, as is the case during natural infections. The infection rates in molted nymphs, as determined by immunostaining of sporozoites in the salivary glands were, however, low. This may be due to the imperfect transmission of *T. orientalis* from mice to larval ticks and/or issues associated with the development of the kinete stage in the tick gut, migration of kinetes to salivary glands, and sporozoite maturation. For instance, tick immunity may be a major deterrent factor in parasite development as previously suggested (Hajdušek et al., 2013), leading to

low rates of sporozoite infection in tick salivary glands. Detailed analyses of *T. orientalis* infection in larval ticks immediately after repletion and at defined intervals might explain whether the low infection rate in nymphal salivary glands is due to imperfect transmission or factors influencing the development of *T. orientalis* in infected ticks.

Unfortunately, our attempts to infect naive Bo-RBC-SCID mice with the parasitized nymphal ticks failed (data not shown). This outcome was not entirely unexpected, since development of schizonts is contingent on the presence of nucleated bovine cells (Shaw, 2003), for which murine cells do not offer a viable alternative. Nonetheless, recent research on humanized mice has shown that *Plasmodium falciparum* sporozoite infection could be successfully established in a chimeric human liver mouse model (Vaughan et al., 2015). Adapting the same technology to "bovinize" mice with bovine monocytes could thus provide a model for the maintenance of the complete *T. orientalis* life cycle in mice.

Although the genome sequences of three *Theileria* spp. (*T. parva*, *T. annulata* and *T. orientalis*) have been completed (Gardner et al., 2005; Pain et al., 2005; Hayashida et al., 2012), only EST and massively parallel signature sequencing (MPSS) data for the piroplasm stages of *T. orientalis*, *T. parva* and *T. annulata,* and the schizont stages of *T. parva* and *T. annulata* (Bishop et al., 2005; Hayashida et al., 2012) are available. The proteomic data for *T. parva* sporozoites have also been published recently (Nyagwange et al., 2017), but a full transcriptomic analysis of the sporozoite stage of any *Theileria* sp. has not been conducted. Here, we successfully obtained transcriptomics data from both the piroplasm and sporozoite

371 stages of *T. orientalis* using our experimental model. In our data, approximately 84.4% and 52.4% of the
372 genes were detected to be moderately or highly expressed (>10 RPKM) in the piroplasm and sporozoite
373 stages, respectively (Supplementary Table S1), indicating that a larger number of genes were transcribed in
374 the piroplasm stage than in the sporozoite stage. The results for the piroplasm were consistent with
375 observations in *T. parva*, in which 83% of potentially detectable genes were identified in the schizont stage
376 by MPSS analysis (Bishop et al., 2005). By contrast, fewer genes were expressed in the sporozoite stage. The
377 biological significance of this observation is unclear, but it may indicate that genes in the sporozoite stage
378 are more tightly regulated in a stage-specific manner.

379 Whole genome transcriptomic analysis of the piroplasm and sporozoite stages identified
380 stage-specific gene expression patterns. Among the 53 stage-specific genes we identified, the expression of a
381 possible secretory protein (TOT_030000643) and four other genes in the same PiroF000037 family was
382 remarkably higher in the piroplasm stage (Supplementary Table S3 and Table 2). The orthologue family
383 PiroF000037 consists of seven genes that are present on Chr. 2, 3 and 4. The presence of these genes in *T.*
384 *orientalis* causes synteny breaks between *T. parva* or *T. annulata*. The abundancy of the transcripts coupled
385 with the fact that the PiroF000037 gene family showed a greater expansion in *T. orientalis* compared with
386 other *Theileria* spp. may hint at the biological importance of this gene family in blood-stage *T. orientalis*.

387 *ToSPAG*, which is the direct ortholog of the *p67* and *SPAG-1* genes, was identified as a gene
388 expressed specifically in the *T. orientalis* sporozoite stage and its transcription was experimentally confirmed.

389 The immunogenicity, function(s), genetic diversity, and vaccine potential of ToSPAG merits further
390 investigation, as its p67 and SPAG-1 orthologs are promising vaccine candidates in *T. parva* and *T. annulata,*
391 respectively (Nene et al., 2016). With the exception of *ToSPAG*, the remaining stage-specific genes were
392 annotated as encoding hypothetical proteins. Further studies to unravel the roles of these sporozoite- and
393 piroplasm-specific genes will help researchers gain a better understanding of the biology of theileriosis, as
394 well as provide potential new targets for parasite control strategies.
395 In conclusion, our mouse–tick infection model is a promising tool that can be effectively
396 employed in *Theileria* research. In addition, the whole genome transcriptomic data generated in the present
397 study will provide useful insight into the biology of this parasite as well as potentially contributing toward
398 the development of efficient control measures targeting oriental theileriosis.
399
400 **Acknowledgements**
401 The authors acknowledge Ms. Hiroko Yamamoto (National Research Center for Protozoan
402 Diseases, Obihiro University of Agriculture and Veterinary Medicine, Japan) for her excellent technical
403 assistance. We thank Dr. Axel Martinelli (Global Station for Zoonosis Control, GI-CoRE, Hokkaido
404 University) and Kate Fox, DPhil, from the Edanz Group (www.edanzediting.com/ac), for editing a draft of
405 this manuscript. This work was supported by grants from MEXT KAKENHI (26850179 to KH, 25292169
406 and 2604087 to NY, 16K18794 to RUS), from the Science and Technology Research Promotion Program for

Agriculture, Forestry, Fisheries and Food Industry, Japan, and from the project of the NARO Biotechnology Research Advancement Institution (a special project scheme for the regional development strategy), Japan, and was partially supported by the International Collaborative Research Program for Tackling NTD (Neglected Tropical Disease) Challenges in African countries (JP18jm0510001), Japan Agency for Medical Research and Development (AMED), Japan. This work was also partially supported by MEXT, Japan as a project of the Joint Usage/Research Center. The funders had no role in the study design, data collection and analysis, the decision to publish, or preparation of the manuscript.

References

Bishop, R., Nene, V., Staeyert, J., Rowlands, J., Nyanjui, J., Osaso, J., Morzaria, S., Musoke, A., 2003. Immunity to East Coast fever in cattle induced by a polypeptide fragment of the major surface coat protein of *Theileria parva* sporozoites. Vaccine 21, 1205–1212.

Bishop, R., Shah, T., Pelle, R., Hoyle, D., Pearson, T., Haines, L., Brass, A., Hulme, H., Graham, S.P., Taracha, E.L.N., Kanga, S., Lu, C., Hass, B., Wortman, J., White, O., Gardner, M.J., Nene, V., de Villiers, E.P., 2005. Analysis of the transcriptome of the protozoan *Theileria parva* using MPSS reveals that the majority of genes are transcriptionally active in the schizont stage. Nucleic Acids Res. 33, 5503–5511. https://doi.org/10.1093/nar/gki818

Boulter, N., Hall, R., 1999. Immunity and vaccine development in the bovine theilerioses. Adv. Parasitol. 44, 41–97.

Fujisaki, K., 1978. Development of acquired resistance precipitating antibody in rabbits experimentally infested with females of *Haemaphysalis longicornis* (Ixodoidea: Ixodidae). Natl. Inst. Anim. Health Q. (Tokyo). 18, 27–38.

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.M., 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.N., 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. https://doi.org/10.1126/science.1110439

Gettinby, G., 1993. The epidemiology of theileriosis in Africa. Parasitol. Today 9, 272. https://doi.org/10.1016/0169-4758(93)90075-Q

Gong, H., Liao, M., Zhou, J., Hatta, T., Huang, P., Zhang, G., Kanuka, H., Nishikawa, Y., Xuan, X., Fujisaki, K., 2008. Gene silencing of ribosomal protein P0 is lethal to the tick *Haemaphysalis longicornis*. Vet. Parasitol. 151, 268–278. https://doi.org/10.1016/j.vetpar.2007.11.015

Hagiwara, K., Tsuji, M., Ishihara, C., Tajima, M., Kurosawa, T., Iwai, H., Takahashi, K., 1993. *Theileria sergenti* infection in the Bo-RBC-SCID mouse model. Parasitol. Res. 79, 466–470.

Hajdušek, O., Síma, R., Ayllón, N., Jalovecká, M., Perner, J., de la Fuente, J., Kopáček, P., 2013. Interaction of the tick immune system with transmitted pathogens. Front. Cell. Infect. Microbiol. 3, 26. https://doi.org/10.3389/fcimb.2013.00026

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, e00204-12-e00204-12. https://doi.org/10.1128/mBio.00204-12

Honda, Y., Waithaka, M., Taracha, E.L., Duchateau, L., Musoke, A.J., McKeever, D.J., 1998. Delivery of the *Theileria parva* p67 antigen to cattle using recombinant vaccinia virus: IL-2 enhances protection. Vaccine 16, 1276–1282.

Irvin, A.D., 1987. Characterization of species and strains of *Theileria*. Adv. Parasitol. 26, 145–97.

458 Iwasaki, T., Kakuda, T., Sako, Y., Sugimoto, C., Onuma, M., 1998. Differentiation and quantification of
459 *Theileria sergenti* piroplasm types using type-specific monoclonal antibodies. J. Vet. Med. Sci. 60,
460 665–669.
461 Izzo, M.M., Poe, I., Horadagoda, N., De Vos, A.J., House, J.K., 2010. Haemolytic anaemia in cattle in NSW
462 associated with *Theileria infections*. Aust. Vet. J. 88, 45–51.
463 https://doi.org/10.1111/j.1751-0813.2009.00540.x
464 Kaba, S.A., Schaap, D., Roode, E.C., Nene, V., Musoke, A.J., Vlak, J.M., van Oers, M.M., 2004. Improved
465 immunogenicity of novel baculovirus-derived *Theileria parva* p67 subunit antigens. Vet. Parasitol. 121,
466 53–64. https://doi.org/10.1016/j.vetpar.2004.02.013
467 Kawamoto, S., Takahashi, K., Kurosawa, T., Sonoda, M., Onuma, M., 1990. Intraerythrocytic schizogony of
468 *Theileria sergenti* in cattle. Nihon Juigaku Zasshi. 52, 1251–1259.
469 McFadden, A.M.J., Rawdon, T.G., Meyer, J., Makin, J., Morley, C.M., Clough, R.R., Tham, K., Mullner, P.,
470 Geysen, D., 2011. An outbreak of haemolytic anaemia associated with infection of *Theileria orientalis*
471 in naive cattle. N. Z. Vet. J. 59, 79–85. https://doi.org/10.1080/00480169.2011.552857
472 McKeever, D.J., 2007. Live immunisation against *Theileria parva*: containing or spreading the disease?
473 Trends Parasitol. 23, 565–568. https://doi.org/10.1016/j.pt.2007.09.002
474 Morrison, W.I., McKeever, D.J., 2006. Current status of vaccine development against *Theileria* parasites.
475 Parasitology 133, S169-S187. https://doi.org/10.1017/S0031182006001867

476 Musoke, A., Morzaria, S., Nkonge, C., Jones, E., Nene, V., 1992. A recombinant sporozoite surface antigen
477 of *Theileria parva* induces protection in cattle. Proc. Natl. Acad. Sci. U. S. A. 89, 514–518.
478 Nene, V., Kiara, H., Lacasta, A., Pelle, R., Svitek, N., Steinaa, L., 2016. The biology of *Theileria parva* and
479 control of East Coast fever - Current status and future trends. Ticks Tick. Borne. Dis. 7, 549–564.
480 https://doi.org/10.1016/j.ttbdis.2016.02.001
481 Nene, V., Musoke, A., Gobright, E., Morzaria, S., 1996. Conservation of the sporozoite p67 vaccine antigen
482 in cattle-derived *Theileria parva* stocks with different cross-immunity profiles. Infect. Immun. 64,
483 2056–2061.
484 Nyagwange, J., Tijhaar, E., Ternette, N., Mobegi, F., Tretina, K., Silva, J.C., Pelle, R., Nene, V., 2018.
485 Characterization of the *Theileria parva* sporozoite proteome. Int. J. Parasitol. 48, 265–273.
486 https://doi.org/10.1016/j.ijpara.2017.09.007
487 Oura, C.A.L., Bishop, R., Ashimwe, B.B., Spooner, P., Lubega, G.W., Tait, A., 2007. *Theileria parva* live
488 vaccination: parasite transmission, persistence and heterologous challenge in the field. Parasitology 134,
489 1205–1213. https://doi.org/10.1017/S0031182007002557
490 Pain, A., Renauld, H., Berriman, M., Murphy, L., Yeats, C.A., Weir, W., Kerhornou, A., Aslett, M., Bishop,
491 R., Bouchier, C., Cochet, M., Coulson, R.M.R., Cronin, A., de Villiers, E.P., Fraser, A., Fosker, N.,
492 Gardner, M., Goble, A., Griffiths-Jones, S., Harris, D.E., Katzer, F., Larke, N., Lord, A., Maser, P.,
493 McKellar, S., Mooney, P., Morton, F., Nene, V., O'Neil, S., Price, C., Quail, M.A., Rabbinowitsch, E.,

494 Rawlings, N.D., Rutter, S., Saunders, D., Seeger, K., Shah, T., Squares, R., Squares, S., Tivey, A.,
495 Walker, A.R., Woodward, J., Dobbelaere, D.A.E., Langsley, G., Rajandream, M..A., McKeever, D.,
496 Shiels, B., Tait, A., Barrell, B., Hall, N., 2005. Genome of the host-cell transforming parasite *Theileria*
497 *annulata* compared with *T. parva*. Science 309, 131–133. https://doi.org/10.1126/science.1110418
498 Robinson, M.D., McCarthy, D.J., Smyth, G.K., 2010. edgeR: a Bioconductor package for differential
499 expression analysis of digital gene expression data. Bioinformatics 26, 139–140.
500 https://doi.org/10.1093/bioinformatics/btp616
501 Sako, Y., Sugimoto, C., Onuma, M., 1999. Expression of a major piroplasm surface protein of *Theileria*
502 *sergenti* in sporozoite stage. J. Vet. Med. Sci. 61, 275–277.
503 Sasaki, N., Takahashi, K., Kawamoto, S., Kurosawa, T., Sonoda, M., Kawai, S., 1990. Maturation of
504 *Theileria sergenti* in salivary glands of *Haemaphysalis longicornis* induced by temperature stimulation.
505 Nihon Juigaku Zasshi. 52, 495–501.
506 Shaw, M.K., 2003. Cell invasion by *Theileria* sporozoites. Trends Parasitol. 19, 2–6.
507 Sivakumar, T., Hayashida, K., Sugimoto, C., Yokoyama, N., 2014. Evolution and genetic diversity of
508 *Theileria*. Infect. Genet. Evol. 27, 250–263. https://doi.org/10.1016/j.meegid.2014.07.013
509 Sugimoto, C., Kawazu, S., Sato, M., Kamio, T., Fujisaki, K., 1992. Preliminary biochemical characterization
510 of "veil" structure purified from *Theileria sergenti*-, *T. buffeli*- and *T. orientalis*-infected bovine
511 erythrocytes. Parasitology 104 (Pt 2), 207–213.

512 Takahashi, K., Kawai, S., Yaehata, K., Kawamoto, S., Hagiwara, K., Kurosawa, T., Tajima, M., Sonoda, M.,
513 1993. Sporogony of *Theileria sergenti* in the salivary glands of the tick vector *Haemaphysalis*
514 *longicornis*. Parasitol. Res. 79, 1–7.
515 Terada, Y., Ishida, M., Yamanaka, H., 1995. Resistibility to *Theileria sergenti* infection in Holstein and
516 Japanese Black cattle. J. Vet. Med. Sci. 57, 1003–1006.
517 Tsuji, M., Hagiwara, K., Takahashi, K., Ishihara, C., Azuma, I., Siddiqui, W.A., 1992. *Theileria sergenti*
518 proliferates in SCID mice with bovine erythrocyte transfusion. J. Parasitol. 78, 750–752.
519 Vaughan, A.M., Pinapati, R.S., Cheeseman, I.H., Camargo, N., Fishbaugher, M., Checkley, L.A., Nair, S.,
520 Hutyra, C.A., Nosten, F.H., Anderson, T.J.C., Ferdig, M.T., Kappe, S.H.I., 2015. *Plasmodium*
521 *falciparum* genetic crosses in a humanized mouse model. Nat. Methods 12, 631–633.
522 https://doi.org/10.1038/nmeth.3432
523 Vega, C.A., Buening, G.M., Green, T.J., Carson, C.A., 1985. In vitro cultivation of *Babesia bigemina*. Am. J.
524 Vet. Res. 46, 416–420.
525 Yokoyama, N., Sivakumar, T., Ota, N., Igarashi, I., Nakamura, Y., Yamashina, H., Matsui, S., Fukumoto, N.,
526 Hata, H., Kondo, S., Oshiro, M., Zakimi, S., Kuroda, Y., Kojima, N., Matsumoto, K., Inokuma, H.,
527 2012. Genetic diversity of *Theileria orientalis* in tick vectors detected in Hokkaido and Okinawa, Japan.
528 Infect. Genet. Evol. 12, 1669–1675. https://doi.org/10.1016/j.meegid.2012.07.007

529 Yokoyama, N., Ueno, A., Mizuno, D., Kuboki, N., Khukhuu, A., Igarashi, I., Miyahara, T., Shiraishi, T.,
530 Kudo, R., Oshiro, M., Zakimi, S., Sugimoto, C., Matsumoto, K., Inokuma, H., 2011. Genotypic
531 diversity of *Theileria orientalis* detected from cattle grazing in Kumamoto and Okinawa prefectures of
532 Japan. J. Vet. Med. Sci. 73, 305–312.
533 Zhuang, W., Sugimoto, C., Matsuba, T., Niinuma, S., Murata, M., Onuma, M., 1994. Analyses of antigenic
534 and genetic diversities of *Theileria sergenti* piroplasm surface proteins. J. Vet. Med. Sci. 56, 469–473.
535
536

537 **Figure legends**

538 **Fig. 1.** Experimental procedure used to generate the mouse–tick *Theileria orientalis* infection model.
539 Splenectomized SCID mice were transfused with uninfected bovine red blood cells (RBCs)
540 (Bo-RBC-SCID), and infected with *T. orientalis.* Whole blood was collected from the infected
541 Bo-RBC-SCID mice after parasitemia developed, and the piroplasms contained within it were purified.
542 Larval ticks were fed on *T. orientalis*-infected Bo-RBC-SCID mice, engorged ticks were allowed to molt
543 into nymphs, and sporozoite-containing salivary glands were collected after a feeding stimulation was
544 conducted on non-infected BALB/c mice for 48 to 72 h.

545

546 **Fig. 2.** *Theileria orientalis* parasitization of transfused bovine red blood cell (Bo-RBC) in SCID mouse A)
547 Thin blood smears prepared from *T. orientalis*–infected Bo-RBC-SCID mice were stained with Giemsa
548 solution and then observed under a light microscope (×1,000). Note that the blood stages of the parasite are
549 pleomorphic and that clear membrane-like bar (BA) and ill-defined veil (VE) structures are visible within the
550 infected RBCs. (B) Acridine orange-stained *T. orientalis* in the RBCs from a Bo-RBC-SCID mouse. The
551 parasite nuclei and cytoplasm appear as green/yellow or orange, respectively (×1,000). (C). A thin smear
552 prepared from *T. orientalis*–infected cattle blood, with which the Bo-RBC-SCID mice were infected, was
553 stained with Giemsa solution and then observed under a light microscope (×1,000). (D) Immunofluorescence
554 (×1,000) of *T. orientalis* within two Bo-RBC-SCID mouse RBCs stained with an anti- major piroplasm

surface protein (MPSP) monoclonal antibody (green). DAPI was used for nuclear staining (blue). (E) Dynamics of *T. orientalis* parasitemia in the Bo-RBC-SCID mice. Triangles indicate the parasitemia level and days p.i. when the larval ticks were allowed to feed on the individual Bo-RBC-SCID mice.

Fig. 3. Immunofluorescence staining of *Theileria orientalis* major piroplasm surface protein (MPSP) in the salivary glands of *H. longicornis*. Whole mount (A, ×100) and frozen slide sections (B, ×1,000) from *H. longicornis* salivary glands were each stained with an anti-MPSP monoclonal antibody (green). DAPI was used to stain the nuclei (blue in A and red in B).

Fig. 4. Expression profiles of major piroplasm surface protein (*MPSP*) and *T. orientalis* sporozoite surface antigen (*ToSPAG*) mRNAs in *Theileria orientalis*–infected *Haemaphysalis longicornis*. cDNA prepared from the total RNA extracted from the salivary glands of non-infected (Tick-NI), 48 h fed (Sporozoite 48 h), and 72 h fed (Sporozoite 72 h) *T. orientalis*–infected nymphal ticks was subjected to quantitative RT-PCR (qRT-PCR) assays to calculate the mRNA levels of the *MPSP* and *ToSPAG* genes. Data were normalized by qPCR analysis of *P0* levels in the cDNA samples and are presented here as the gene expression values relative to those for the non-infected ticks. Data represent the means ±S.E. of triplicate samples.

572 **Fig. 5.** Heatmap illustrating the relative expression levels of the differentially expressed genes (DEGs) in the
573 piroplasm and sporozoite stages of *Theileria orientalis*. Twenty-three and 30 genes were expressed
574 specifically in the sporozoite and piroplasm stages, respectively. The reads per kilobase per million (RPKM)
575 values for each gene were converted to standardized values (Z-scores). The scale bar is shown with the
576 minimum expression value for each gene in red and the maximum value in green. The full DEG list is shown
577 in Supplementary Table S2. The highly expressed genes specific to each developmental stage are indicated
578 on the right-hand side. Note that PiroF000037 gene family members and the *ToSPAG* gene were among the
579 highly expressed genes in the piroplasm and sporozoite stages, respectively.
580
581 **Supplementary figure legend**
582
583 **Supplementary Fig. S1.** Gene structure of *T. orientalis* sporozoite surface antigen (*ToSPAG*), an ortholog
584 of *p67* gene in *Theileria orientalis*. (A) Shematic representation of intron-exon boundaries in previously
585 annotated TOT_030000930 and *ToSPAG* _LC342224 identified in this study. Blue indicates the same
586 intron prediction, while red is a newly identified intron. (B) Amino acid sequence of *ToSPAG* (LC342224).
587 The gene contained a N-terminal endoplasmic reticulum (ER) signal sequence as predicted by signalP 4.1
588 server with a cleavage site between positions 19 and 20 (bold). A Pfam domain (PF05642.10; Sporozoite
589 P67 surface antigen; and amino acid (a.a. 876-1056) were also found (underlined).

Appendix A. Supplementary data:

Supplementary Fig. S1. Gene structure of *T. orientalis* sporozoite surface antigen **(*ToSPAG*), an ortholog of *p67* gene in *Theileria orientalis***

Supplementary Table S1. Number of reads obtained for each transcriptome data set

Supplementary Table S2. Complete list of genes and their calculated reads per kilobase per million (RPKM) values in *Theileria orientalis* piroplasm and the sporozoite stages

Supplementary Table S3. The differentially expressed genes (DEGs) in *Theileria orientalis* piroplasm and the sporozoite stages

599 **Table 1.** *Theileria orientalis* infection in *Haemaphysalis longicornis*

Tick group ID	Parasitemia in mouse[a]		No. engorged ticks	No. molted ticks (%[b])	No. infected ticks[c] (%[d])
	Beginning of feeding	End of feeding			
#3	9%	20%	10	9 (90.0)	n.d.
#4	8%	8%	89	66 (74.2)	n.d.
#5	6%	20%	34	31 (91.2)	3 (30)
#5 (2nd[e])	20%	20%	58	54 (93.1)	4 (40)
#7	10%	12%	105	93 (88.6)	1 (10)
#8	5%	5%	64	45 (70.3)	2 (20)
#10	10%	11%	28	18 (64.3)	1 (10)
#10 (2nd[e])	11%	11%	42	40 (95.2)	n.d.
#11	3%	3%	109	85 (78.0)	n.d.
#15	5%	11%	77	74 (96.1)	4 (40)
#15 (2nd[e])	11%	27%	56	53 (94.6)	n.d.
#16 (Pass[f])	1%	4%	297	270 (90.9)	4 (40%)
#17 (Pass[f])	1%	9%	222	180 (81.1)	1 (10)
Total			1191	1018 (85.5)	20 (25)

600

601 [a]Each tick group was fed on *T. orientalis*–infected Bo-RBC-SCID mice (SCID mice that had been transfused
602 with bovine red blood cells) with different parasitemia levels.
603 [b]Expressed as the percentage of engorged ticks in each group.
604 [c]A nymphal tick was considered positive for *T. orientalis* infection if major piroplasm surface protein
605 (MPSP) antigen was observed in its salivary glands by immunofluorescent staining.
606 [d]Expressed as the percentage of 10 molted ticks examined for MPSP antigen in each tick group.
607 [e]The same Bo-RBC-SCID mouse was subjected to a second round of tick feeding after completion of the
608 first feed.
609 [f]Tick groups 16 and 17 were fed on mice that were infected by injecting cryopreserved *T.*
610 *orientalis*–containing RBCs sourced from the Bo-RBC-SCID mice.
611 n.d., Not done.

612

Table 2. Genes identified as highly expressed in the piroplasm stage of *Theileria orientalis*

Data set[a]	Rank[b]	Gene name	Annotation	Expression Value (RPKM)	OrthoMCL[c] gene Family	ER_Signal[d]
Piroplasm Cattle	1	**TOT_030000643**	conserved hypothetical protein	22,004.47	**PiroF0000037**	**Y**
Piroplasm Cattle	2	TOT_010000613	conserved hypothetical protein	8,798.96	PiroF0000199	Y
Piroplasm Cattle	3	TOT_040000598	conserved hypothetical protein	7,111.98	**PiroF0000037**	**Y**
Piroplasm Cattle	4	TOT_030000565	Tocp1 variant tocp4	4,928.37	PiroF0100002	N
Piroplasm Cattle	5	TOT_030000105	hypothetical protein	4,150.30	PiroF0002890	N
Piroplasm Cattle	6	TOT_010000558	conserved hypothetical protein	4,099.73	PiroF0000199	Y
Piroplasm Cattle	7	TOT_040000842	conserved hypothetical protein	3,901.31	PiroF0000199	Y
Piroplasm Cattle	8	TOT_020000484	conserved hypothetical protein	3,354.05	PiroF0001462	N
Piroplasm Cattle	9	TOT_040000841	conserved hypothetical protein	3,216.60	**PiroF0000037**	**Y**
Piroplasm Cattle	10	TOT_020001109	conserved hypothetical protein	2,223.05	**PiroF0000037**	N
Piroplasm Cattle	11	TOT_020000353	conserved hypothetical protein	2,038.11	**PiroF0000037**	**Y**
Piroplasm Cattle	12	TOT_040000718	hypothetical protein	1,847.49	PiroF0003910	N
Piroplasm SCID	1	**TOT_030000643**	conserved hypothetical protein	29,153.54	**PiroF0000037**	**Y**
Piroplasm SCID	2	TOT_040000822	histone H2B variant 1	10,214.90	PiroF0000463	N
Piroplasm SCID	3	TOT_030000565	Tocp1 variant tocp4	10,119.96	PiroF0100002	N
Piroplasm SCID	4	TOT_010001167	conserved hypothetical protein	9,326.28	PiroF0000199	Y
Piroplasm SCID	5	TOT_040000710	conserved hypothetical protein	8,618.76	PiroF0000012	Y
Piroplasm SCID	6	TOT_040000598	conserved hypothetical protein	7,519.27	**PiroF0000037**	**Y**
Piroplasm SCID	7	TOT_010000613	conserved hypothetical protein	6,264.98	PiroF0000199	Y
Piroplasm SCID	8	TOT_020000484	conserved hypothetical protein	5,222.67	PiroF0001462	N
Piroplasm SCID	9	TOT_040000842	conserved hypothetical protein	5,177.38	PiroF0000199	Y
Piroplasm SCID	10	TOT_010000558	conserved hypothetical protein	4,846.78	PiroF0000199	Y
Piroplasm SCID	11	TOT_020000539	TS-Ikeda type piroplasm surface protein (p23)	4,788.67	PiroF0003021	Y
Piroplasm SCID	12	TOT_020000353	conserved hypothetical protein	3,926.88	**PiroF0000037**	**Y**

[a]Piroplasm Cattle: Transcript data set obtained from piroplasms isolated from infected cattle blood. Piroplasm SCID: Transcript data set obtained from piroplasms isolated from infected SCID mice that had been transfused with bovine red blood cells (Bo-RBC-SCID) .

617 [b]Transcripts detected from each data set were ranked by reads per kilobase per million (RPKM) value, and
618 the top 12 genes are shown.
619 [c]The orthologous gene family numbers were assigned according to Hayashida et al., 2012.
620 [d]The endoplasmic reticulum (ER) signal sequences were predicted using SignalP v4.1.

621

622 **Table 3.** Genes identified as highly expressed in the sporozoite stage of *Theileria orientalis*

Data set[a]	Rank[b]	Gene ID	annotation	Expression Value (RPKM)	OrthoMCL gene Family[c]	ER_Signal[d]
Sporozoite 48 h	1	TOT_040000822	histone H2B variant 1	2,966.53	PiroF0000463	N
Sporozoite 48 h	2	TOT_010000166	conserved hypothetical protein	2,356.16	PiroF0002919	Y
Sporozoite 48 h	3	TOT_030000549	conserved hypothetical protein	1,827.77	PiroF0003830	N
Sporozoite 48 h	4	TOT_030000930	p67 ortholog, ToSPAG	1,567.16	PiroF0003159	Y
Sporozoite 48 h	5	TOT_020000536	hypothetical protein	1,135.98	PiroF0003777	N
Sporozoite 48 h	6	TOT_030000105	hypothetical protein	1,122.61	PiroF0002890	N
Sporozoite 48 h	7	TOT_010000695	conserved hypothetical protein	900.27	PiroF0002220	N
Sporozoite 48 h	8	TOT_010000987	uncharacterized protein	728.86	PiroF0000065	N
Sporozoite 48h	9	TOT_010000624	ubiquitin	655.51	PiroF0000302	N
Sporozoite 48 h	10	TOT_020001102	actin 1	609.26	PiroF0001924	N
Sporozoite 48 h	11	TOT_020000308	conserved hypothetical protein	598.85	PiroF0003787	N
Sporozoite 48 h	12	TOT_030000283	conserved hypothetical protein	560.89	PiroF0003849	Y
Sporozoite 72 h	1	TOT_040000822	histone H2B variant 1	6,599.95	PiroF0000463	N
Sporozoite 72 h	2	TOT_040000288	uncharacterized protein	3,341.37	PiroF0002551	N
Sporozoite 72 h	3	TOT_010001072	conserved hypothetical protein	2,814.68	PiroF0002257	N
Sporozoite 72 h	4	TOT_030000035	conserved hypothetical protein	2,538.51	PiroF0002945	N
Sporozoite 72 h	5	TOT_030000105	hypothetical protein	2,468.09	PiroF0002890	N
Sporozoite 72 h	6	TOT_010000135	conserved hypothetical protein	2,288.92	PiroF0003680	N
Sporozoite 72 h	7	TOT_010000166	conserved hypothetical protein	1,936.43	PiroF0002919	Y
Sporozoite 72 h	8	TOT_030000835	hypothetical protein	1,827.83	PiroF0003809	N

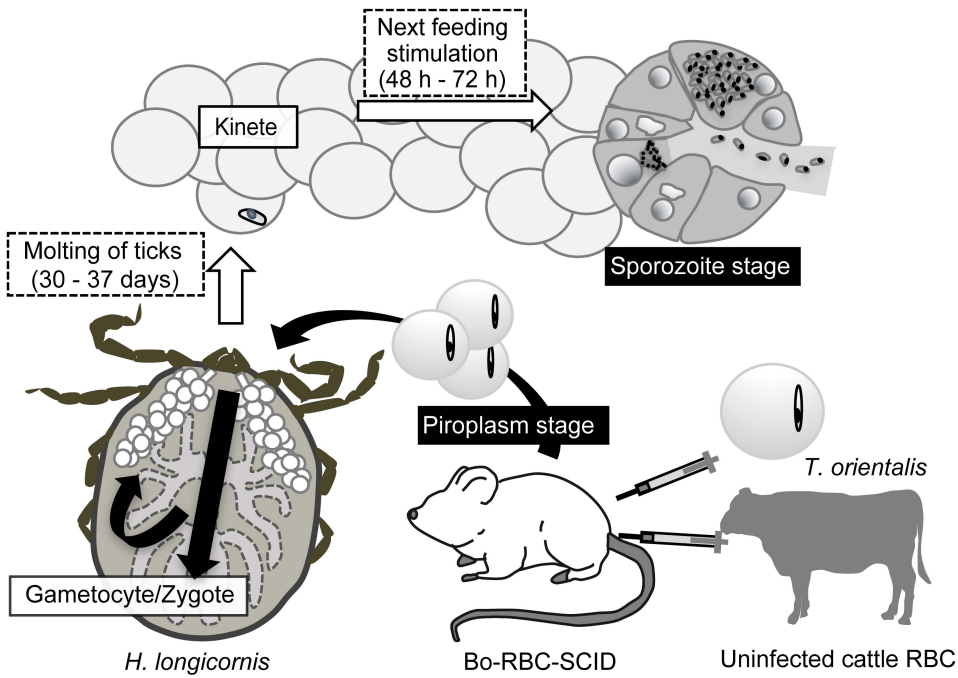
			conserved hypothetical			
Sporozoite 72 h	9	TOT_040000609	protein	1,728.56	PiroF0002753	N
Sporozoite 72 h	10	TOT_040000478	uncharacterized protein	1,706.60	PiroF0003088	Y
Sporozoite 72 h	11	TOT_030000930	p67 ortholog, ToSPAG	1,653.99	PiroF0003159	Y
Sporozoite 72 h	12	TOT_020000536	hypothetical protein	1,525.12	PiroF0003777	N

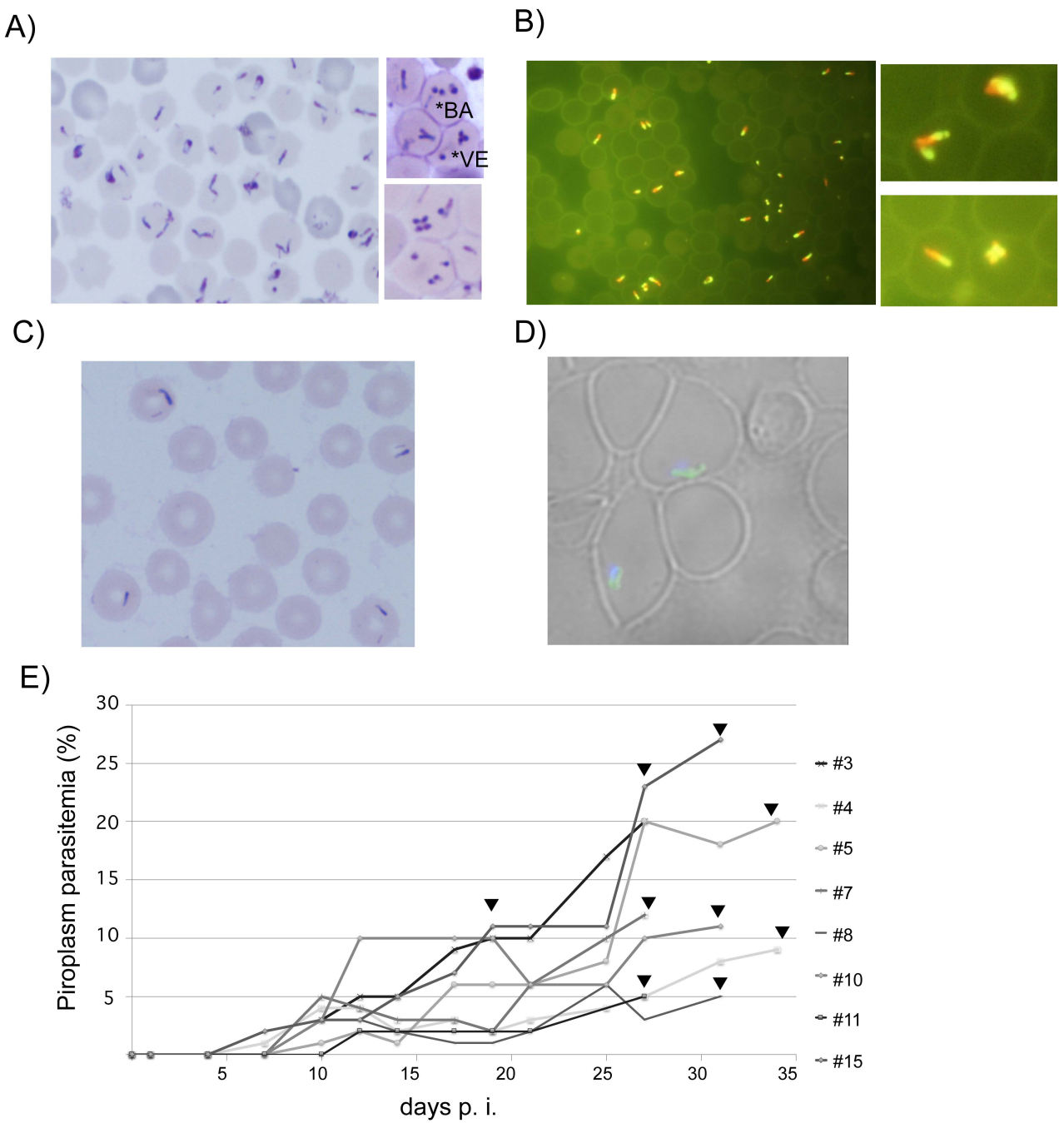
[a]Sporozoite 48 h: Transcript data set obtained from sporozoite-infected tick salivary glands collected at 48 h post-infestation. Sporozoite 72 h: Transcript data set obtained from sporozoite-infected tick salivary glands collected at 72 h post-infestation.

[b]Transcripts detected from each data set were ranked by reads per kilobase per million (RPKM) value, and the top 12 genes are shown.

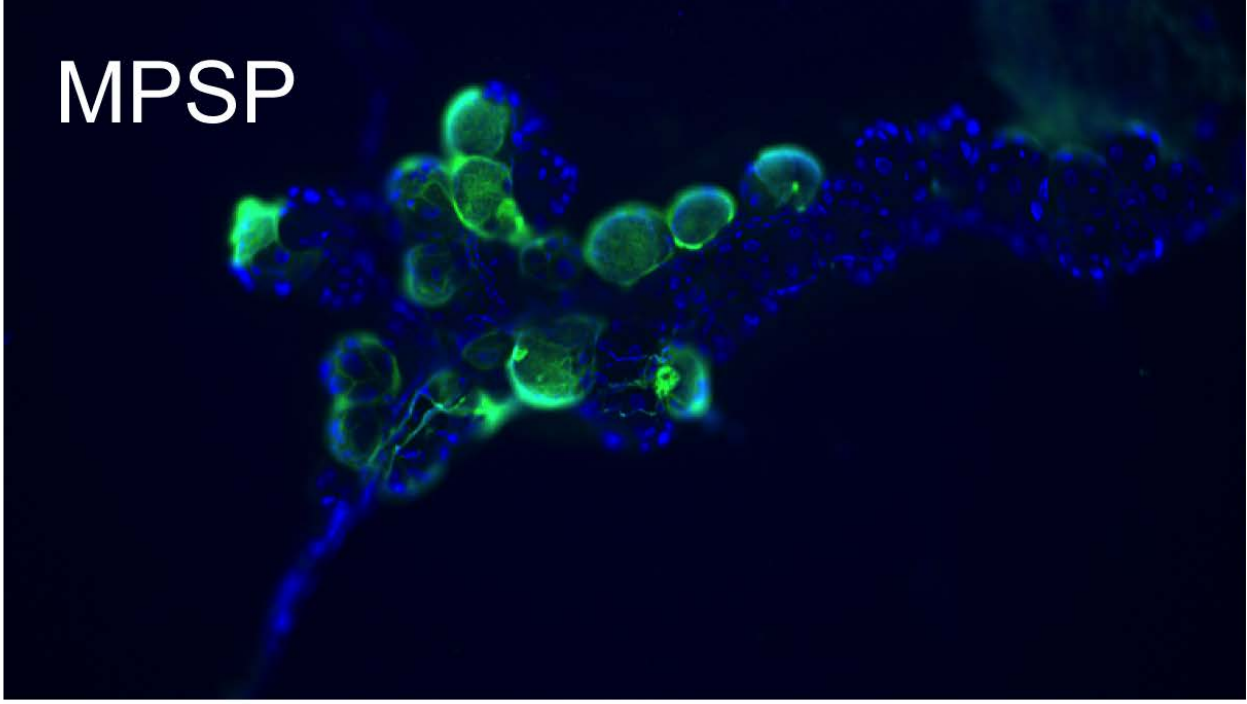
[c]The orthologous gene family numbers were assigned according to Hayashida et al., 2012.

[d] The endoplasmic reticulum (ER) signal sequences were predicted using SignalP v4.1.

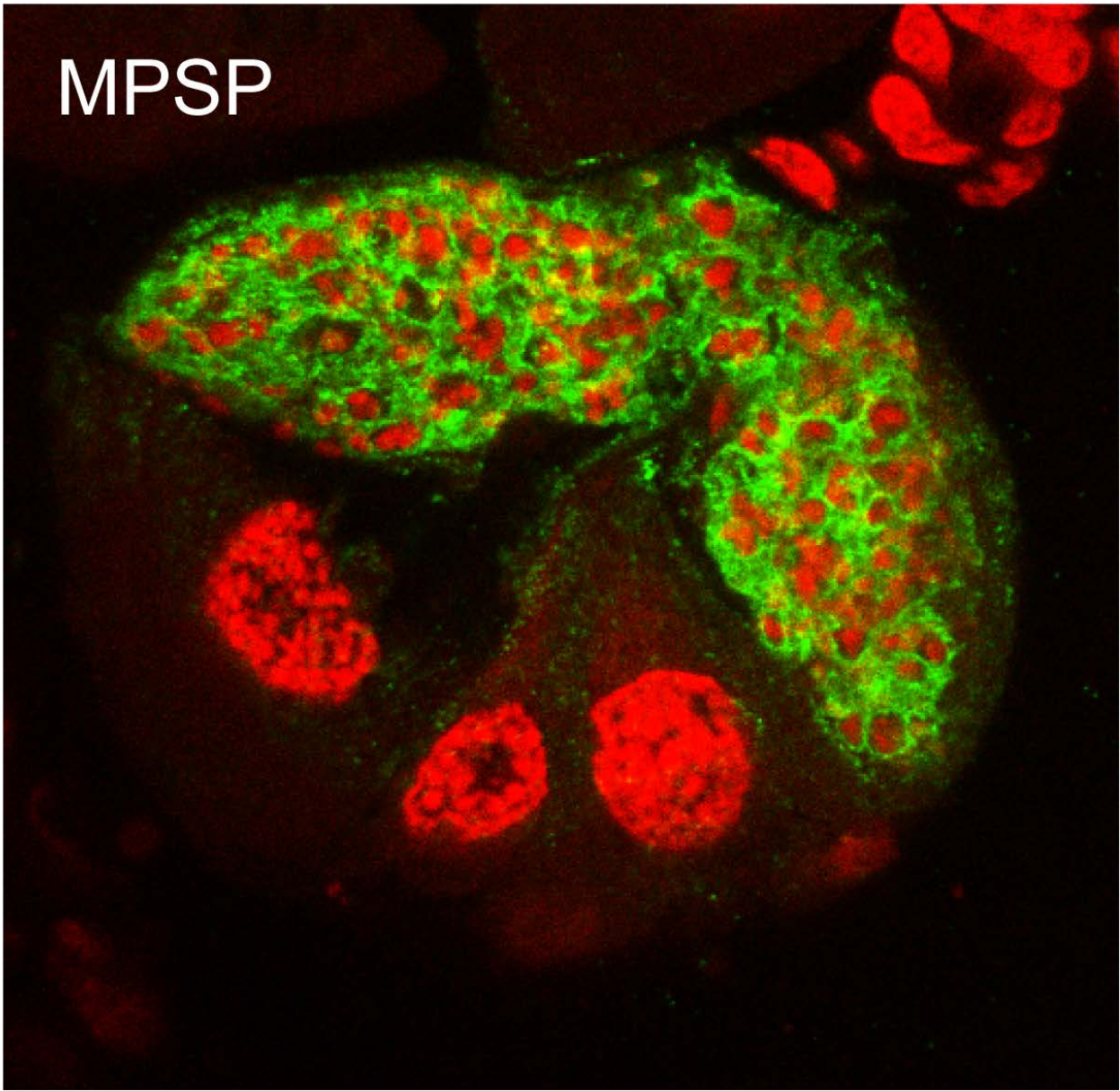

Next feeding
stimulation
(48 h - 72 h)
Kinete
Sporozoite stage
Molting of ticks
(30 - 37 days)
Piroplasm stage
T. orientalis
Gametocyte/Zygote
H. longicornis
Bo-RBC-SCID
Uninfected cattle RBC

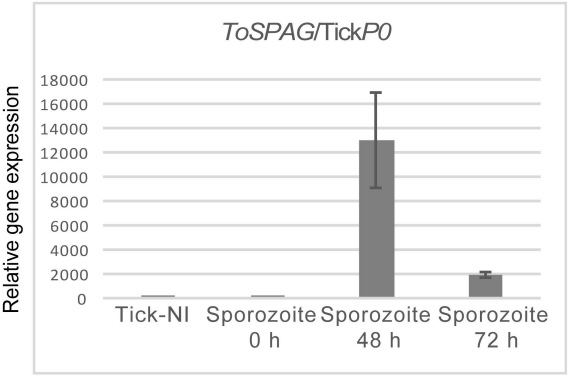

A)
*BA
*VE
B)
C)
D)
E)
Piroplasm parasitemia (%)
30
25
20
15
10
5
5
10
15
20
25
30
35
days p. i.
#3
#4
#5
#7
#8
#10
#11
#15

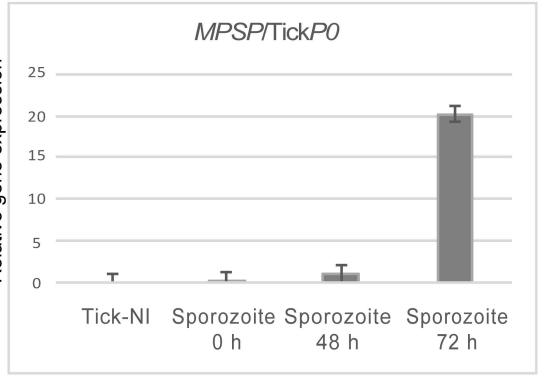
A)

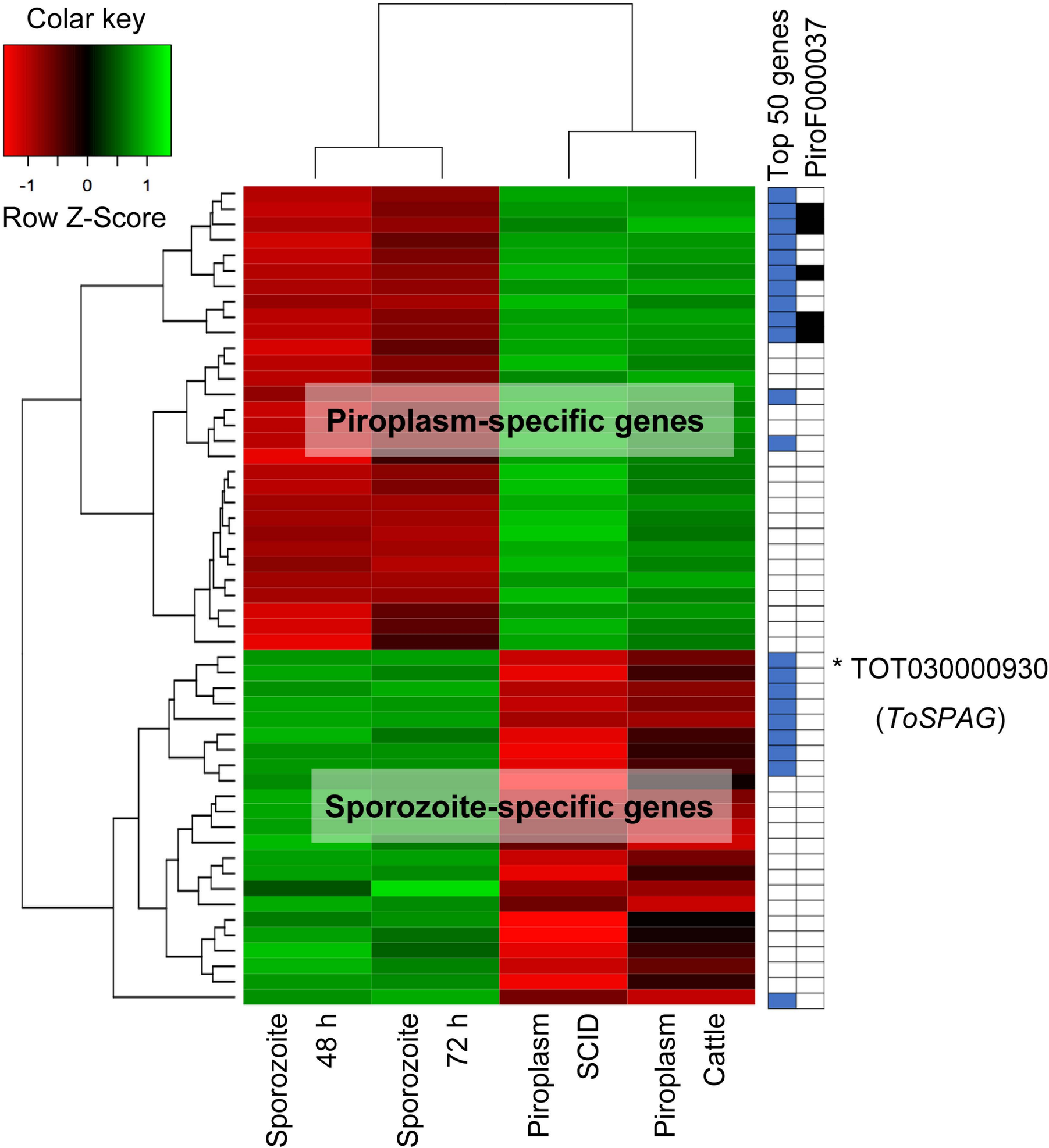


B)

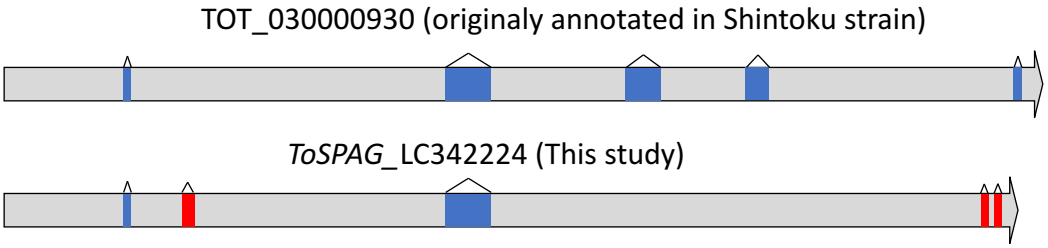



ToSPAG/Tick*P0*
Relative gene expression
18000
16000
14000
12000
10000
8000
6000
4000
2000
0
Tick-NI
Sporozoite 0 h
Sporozoite 48 h
Sporozoite 72 h


MPSP/Tick*P0*
Relative gene expression
25
20
15
10
5
0
Tick-NI
Sporozoite 0 h
Sporozoite 48 h
Sporozoite 72 h


Colar key
-1
0
1
Row Z-Score
Top 50 genes
PiroF000037
Piroplasm-specific genes
Sporozoite-specific genes
* TOT030000930
(ToSPAG)
Sporozoite
48 h
Sporozoite
72 h
Piroplasm
SCID
Piroplasm
Cattle

A)



B)

>ToSPAG_LC342224

MKLLYILLIVPVFFVTGVEEPEGTLSTGSSGRDGATDSTGLVATSESSQVVNSNVEQVQQESQVQEQHSENGHLQQEAGT
NGLPGQETDGRSEDSLTENGHLKGPQASSSLPEGGGAGAESVPRPSTGQEEQLGGARSTEDNSAVPSSSGSDTSNSHGDG
GQSTRDNQNGRQNVLSPPEEENRETTSLGTVSSDLGSQGQQHQLGASRSPADPVQGPGTSSLGTEEEVRRDGENSDSRD
PVSGEATGLAPGSPSKPGEADSRGTVVSREEGQESPTDESENSLGRSGERGPAAPGVEHRDSGDGVTLGDGGLGSAGGLR
VEPSPSGDLSSQVPQVQHTDPSGRADGLSDPAVGVPTLTQDGKQESSEEDDDDEEEDDDDEEEDDDDEDEDEVSSQGH
GDQLPGQLGEPGQDGNLGQDGRQTSQVNPGSSGGPQVTVSGVPDSTAVGGVGAVSPTDQETVRLPTGEVPSVQRTGG
GLVEQQQQESQVQERNGELTGVRGPTVGQGSPDLQQVPRTEASQGLNTVVREPGSVSDVSTEGLKGPGQEQTDPNLRT
EEEVQRVASSEAPGGDGVRDGTGNRPLEQGSEPAAVVSGPSTTQVSQATTVDTSSTTTLTTSAPTTTATTGSTETTITSAGT
TITPTSTSSTTPTVTSSPGTTPTSSLTATTTTNSPTTTTPGSTTITETPTTATVTSSAAREESTNVTTTVATSPATSVATSVGREEI
TRVTQSHLSQGQISSTQEITSQGGRRQEDQVLQTPSRVSGTQETLRQETRTQETNTGENQDQNRNTVSETTVTVSQPSPR
PNVQHTTVAGGSGQINAARRENTVTQPSGQQTSLRGSTLSTGSIKGLTMVVTRTKEKEDKIARKIKEKMVAEEE**VFDIKCF**
DYRKNDPFKLRFYMFKGIFRLWRLLQDLKFFMVVDHTLITDTFDKGVQNYLTKGLTLMNGIVVRDNGDLLAMYNGFN
KYYKAMASRLNHMKEQEEGSEIMKTIVSMSVIGYSTALRLEQEFGSWDLVEVRENEENKEGRVASYTLLGFRIAMYLTK
DIVEVIMDKFLRYTDLVGIDFGINATLARGALMEVQPEDTLVYGENEAIVTIDPNEEYKQLKAYLEYVVSEKGHSSGRGQEN
REELVRTSRKGRVTIRTSIRSGVGGEASGTVEASGTVEASGTVEASGRGEAGSGSEGRAEALEEEVRAKAWRGDPRAIEKVL
KKSMFTAVPRELGPETLEDYLRIFELSAPQEREGPAQVAEVDTEEFINEFLATAGPSYGLDVNIGAGNSKLRYSTSSTDSNVG
ARFVGGRGVESEAARGRGGSADAVARGTSAIRGATADAATAVGATKVTNGVGNEEGVARSSGFTGANAVNGQTGQTN
GDAVINGLAGNDVVINGQNNDDDSDDDDSDDDDSDSDSEDDDSVGVPGIDDNYAVDQSASGPIVEYFMKEFSTTDASM
YGVDLLIGGVTIESVFGISDTLAVDGGATNTTSGKLGVATTTTTTTTTTSETVDVSNGVVGQRTDGGNAEAYSKSVSNGVVG
QRTDGGNAESYSKSVSNGVEEECASAVRSTSVHSSSGGTGTSASTITTTTGNTIAATSTTTTISRGTGTTRNTSTGTTGTTRST
GGTRTNSVTTGIVTSSTATNSGITTRSTSTTNTTRTGTTATTVGGTKTRTGSTSTGTLIPIPSKSFRSRLWTTGAGPIRMIQSQL
GKGATASTRTEKGGLGSEGVVIRSVQTSKIPTSGLRGPAGTDRVQLTEERSKKITVMAPTNSSLTKSTSTHTITKIPTRTITSQS
LSTRSSSTVTGGQTRESSSQRSFSQGGSSSGQTRNNIGHTRSISWQDRSSSVQTGSRSAQISGKKEERRGNEKTKARLTQKR
KKKS*

Bold: ER-signal sequence
Underlined: Pfam domain (PF05642.10) Sporozoite P67 surface antigen

Supplementary Fig. S1.