Genetic diversity and antigenicity variation of *Babesia bovis* merozoite surface antigen-1 (MSA-1) in Thailand

Muncharee Tattiyapong[a,b], Thillaiampalam Sivakumar[a], Hitoshi Takemae[a], Pacharathon Simking[c], Sathaporn Jittapalapong[c], Ikuo Igarashi[a], Naoaki Yokoyama[a,*]

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

[b] National Institute of Animal Health, Department of Livestock Development, Bangkok 10900, Thailand

[c] Department of Parasitology, Faculty of Veterinary Medicine, Kasetsart University, Bangkok 10900, Thailand

Corresponding author: Naoaki Yokoyama

National Research Center for Protozoan Diseases, Obihiro University of Agriculture and Veterinary Medicine, Inada-cho, Obihiro, Hokkaido 080-8555, Japan.

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

E-mail address:yokoyama@obihiro.ac.jp

Abstract

Babesia bovis, an intraerythrocytic protozoan parasite, causes severe clinical disease in cattle worldwide. The genetic diversity of parasite antigens often results in different immune profiles in infected animals, hindering efforts to develop immune control methodologies against the *B*. *bovis* infection. In this study, we analyzed the genetic diversity of the merozoite surface antigen-1 (*msa-1*) gene using 162 *B*. *bovis*-positive blood DNA samples sourced from cattle populations reared in different geographical regions of Thailand. The identity scores shared among 93 *msa-1* gene sequences isolated by PCR amplification were 43.5–100%, and the similarity values among the translated amino acid sequences were 42.8–100%. Of 23 total clades detected in our phylogenetic analysis, Thai *msa-1* gene sequences occurred in 18 clades; seven among them were composed of sequences exclusively from Thailand. To investigate differential antigenicity of isolated MSA-1 proteins, we expressed and purified eight recombinant MSA-1 (rMSA-1) proteins, including an rMSA-1 from *B*. *bovis* Texas (T2Bo) strain and seven rMSA-1 proteins based on the Thai *msa-1* sequences. When these antigens were analyzed in a western blot assay, anti-T2Bo cattle serum strongly reacted with the rMSA-1 from T2Bo, as well as with three other rMSA-1 proteins that shared 54.9–68.4% sequence similarity with T2Bo MSA-1. In

contrast, no or weak reactivity was observed for the remaining rMSA-1 proteins, which shared low sequence similarity (35.0–39.7%) with T2Bo MSA-1. While demonstrating the high genetic diversity of the *B. bovis msa-1* gene in Thailand, the present findings suggest that the genetic diversity results in antigenicity variations among the MSA-1 antigens of *B. bovis* in Thailand.

1. Introduction

Babesia bovis is a tick-transmitted, highly virulent *Babesia* species that induces a fatal bovine babesiosis, leading to severe economic losses in the cattle industry worldwide (Bock et al., 2004). Control of *B. bovis* infection largely relies on tick control and immunization with live attenuated vaccines (De Vos and Bock, 2000; Smith et al., 2000). However, increasing occurrence of acaricide resistance in wild populations limits tick control strategies (Abbas et al., 2014). Broad deployment of live vaccine is limited for several reasons, including strain-specific immunity and contamination with other blood pathogens (Bock et al., 1992; Rogers et al., 1988).

Strain-specific immunity against *B. bovis* is likely associated with genetic diversity of genes that encode parasitic antigens, particularly the merozoite surface antigens (MSAs) (Berens et al., 2005; Leroith et al., 2005). Known MSAs consist of at least five antigens, including MSA-1 and four antigens of the MSA-2 family (MSA-2c, MSA-2a1, MSA-2a2, and MSA-2b) (Florin-Christensen et al., 2002; Suarez et al., 2000). However, the number of MSA-2 antigens may vary between isolates. For example, in all *B. bovis* isolates investigated in Australia, only two genes that encode MSA-2 antigens were detected (*msa-2c* and *msa-2a/b*). The sequence of *msa-2a/b* from these isolates was similar to *msa-2a1*, *msa-2a2*, and *msa-2b* of *B. bovis* strains such as

Texas (T2Bo) and Mo7 that have all four genes in their *msa-2* locus (Berens et al., 2005; Florin-Christensen et al., 2002). In addition, the individual *msa* genes of *B. bovis* isolates from several endemic countries are highly diverse (Genis et al., 2008; Molad et al., 2014; Sivakumar et al., 2013; Tattiyapong et al., 2014; Yokoyama et al., 2015). In particular, the studies that analyzed *msa-1* and *msa-2a/b* from isolates obtained in Australia found marked genetic differences between the vaccine strain and breakthrough isolates, suggesting that the diversity of *msa* genes could alter the antigenicity of *B. bovis* and result in a lack of vaccinal protection (Berens et al., 2005; Leroith et al., 2005). Therefore, genetic analyses of *msa* genes might be useful to identify antigenically different *B. bovis* isolates.

Additionally, MSAs are possible subunit vaccine candidates, as they mediate the initial attachment of *B. bovis* merozoites to the host's red blood cells (RBCs) during their RBC invasion process, as well as induce neutralizing antibodies that can block the RBC invasion by merozoites (Mosqueda et al., 2002a,b; Suarez et al., 2000; Yokoyama et al., 2006). However, the genetic diversity of *msa* genes could prevent immunological cross-reaction among the MSAs from different isolates of *B. bovis* (Berens et al., 2005; Leroith et al., 2005). Moreover, the lack of protection with live attenuated vaccines, which contain only a single strain of *B. bovis*, against field isolates with genetically

diverse *msa* genes suggests that a cocktail of MSA antigens representing genetically diverse variants should be used if we are to use the MSAs as vaccine. Therefore, a better understanding of the genetic diversity of *msa* genes is essential to formulate effective immune control strategies against *B. bovis* infection.

An ideal target for genetic analysis is an MSA-encoding gene that is both highly diverse and shared between all *B. bovis* isolates. Among *msa-2* family genes, *msa-2b* is highly diverse, but the gene has been replaced with *msa-2a/b* in some *B. bovis* isolates, increasing the complexity of genetic analysis (Berens et al., 2005). The gene sequences of *msa-2c* form only a limited number of phylogenetic clades and therefore, the gene may not be an attractive target for genetic characterization (Altangerel et al., 2012; Sivakumar et al., 2013; Tattiyapong et al., 2014; Yokoyama et al., 2015). In contrast, *msa-1* gene sequences are highly diverse and the antigenicity properties of MSA-1 are likely associated with its genetic diversity (Leroith et al., 2005; Suarez et al., 2000). Therefore, *msa-1* is a more suitable target for genetic analyses of *B. bovis* isolates relative to *msa-2* family genes.

Thailand, a Southeast Asian country, is endemic for *B. bovis* (Simking et al., 2013). Although a previous study found that the *msa-1* gene was highly conserved in Thailand (Nagano et al., 2013), only a few *msa-1* gene sequences were investigated in

that study. In contrast to this observation, Simking et al. (2013) found that Thai *msa-2b* gene sequences were highly diverse. We speculated that the genetic diversity of *msa-1* would also be high if large numbers of sequences from Thailand were analyzed. Therefore, in the present study, we collected large numbers of *msa-1* gene sequences from *B. bovis*-positive blood DNA samples sourced from cattle in Thailand to analyze genetic diversity. Additionally, we expressed recombinant MSA-1 antigens using Thai *msa-1* gene sequences from representative phylogenetic clades and analyzed variations in antigenicity.

2. Materials and methods

2.1. DNA samples

A total of 162 *B. bovis*-positive blood DNA samples, sourced from cattle bred in the North (n = 56), Northeast (n = 9), Central (n = 96), and South (n = 1) regions of Thailand, were used in the present study (Simking et al., 2013). *B. bovis* infection was confirmed in these samples using a nested PCR assay based on the spherical body protein-2 (*sbp-2*) gene of *B. bovis*, and used to analyze the diversity of *msa-2b* gene in a previous study (Simking et al., 2013).

2.2. PCR amplification and sequencing of msa-1 gene

DNA samples were subjected to two separate PCR assays to amplify the *msa-1* gene, as described by Tattiyapong *et al*. (2014), using two previously described forward primers, MSA-1 F (5′-ATGGCTACGTTTGCTCTTTTCATTTCAGC-3′) (Altangerel et al., 2012) and MSA-1 F1 (5′-AGTACTTACCTTTTTAATGACAGCCG-3′) (Sivakumar et al., 2013), and a common reverse primer, MSA-1 R (5′-TTAAAATGCAGAGAGAACGAAGTAGC-3′) (Sivakumar et al., 2013). The PCR reactions and cycling conditions were performed as described in a previous report

(Tattiyapong et al., 2014). PCR products with the expected sizes were extracted using

NucleoSpin® Gel and PCR Clean-up kit (MACHEREY-NAGEL GmbH & Co. KG,

Düren, Germany). The extracted DNA was ligated to a PCR 2.1 plasmid vector (TOPO,

Invitrogen, Carlsbad, CA, USA), transformed into chemically competent *Escherichia*

coli cells (TOP 10, Invitrogen), and then plated on LB agar (Invitrogen). For each PCR

amplicon, two clones were sequenced with an ABI PRISM 3100 genetic analyzer

(Applied Biosystems).

2.3. Bioinformatics analyses

The newly obtained sequences were initially analyzed using the Basic Local

Alignment Search Tool (BLAST) (http://blast.ncbi.nlm.nih.gov/Blast.cgi) to confirm

that the PCR assays amplified the *B. bovis msa-1* gene. The percent identity and

similarity scores between the nucleotide and translated amino acid sequences,

respectively, were calculated using MatGAT (Campanella et al., 2003). The amino acid

sequences translated from *msa-1* gene sequences used for the expression of recombinant

MSA-1 (rMSA-1) proteins were analyzed using SignalP 4.0 (Petersen et al., 2011)

(http://www.cbs.dtu.dk/services/SignalP/) and big-PI predictor (Eisenhaber et al., 1999)

(http://mendel.imp.ac.at/sat/gpi/gpi_server.html) to predict the signal peptide and

glycophosphatidylinositol (GPI)-anchor region, respectively. The B-cell epitopes in the

translated MSA-1 amino acid sequence from the *B. bovis* T2Bo strain were predicted, as

described by Kolaskar and Tungaonkar (1990), using an online program

(http://tools.immuneepitope.org/tools/bcell/iedb_input).

2.4. Phylogenetic analysis

The newly obtained *msa-1* gene sequences and their translated amino acid

sequences, together with those retrieved from GenBank, were aligned using MAFFT

(http://mafft.cbrc.jp/alignment/server/) (Katoh et al., 2002). The aligned nucleotide and

amino acid sequences were then used to construct maximum likelihood phylogenetic

trees, based on a General Time Reversible (Nei and Kumar, 2000) and JTT

matrix-based (Jones et al., 1992) substitution models, respectively, using MEGA

version 6.0 (Tamura et al., 2013).

2.5. Protein expression and purification

Eight recombinant proteins were expressed, based on seven Thai sequences

located in different phylogenetic clades together with an *msa-1* gene sequence from the

B. bovis T2Bo strain. PCR primers were designed to amplify the *msa-1* gene excluding

the regions that encode the signal peptide and GPI-anchor. Appropriate restriction sites

were added to each primer, as shown in Table 1, based on the corresponding restriction

sites in the pGEX-6p2 expression plasmid vector (GE Healthcare, Uppsala, Sweden)

and non-cutting restriction sites in the target gene fragments. For each PCR assay, 1 µl

of genomic DNA extracted from the *in vitro* culture of *B. bovis* (T2Bo) or the PCR 2.1

plasmids containing the target gene sequence insert, was added to a 49 µl-reaction

mixture consisting of 25 µl of 2× PCR buffer for KOD FX Neo (ToYoBo, Osaka,

Japan), 10 µl of 2 mM dNTPs (ToYoBo), 1.5 µl of 10 µM each forward and reverse

primers (Table 1), 1 µl of KOD FX Neo DNA polymerase (ToYoBo), and 10 µl of

distilled water. After an initial enzyme activation step at 94 °C for 2 min, the reaction

mixture was amplified for 25 cycles, each containing a denaturing step at 98 °C for 10 s,

an annealing step at 52 °C for 30 s, and an extension step at 68 °C for 30 s. The PCR

products were gel-extracted, digested with their respective restriction enzymes (Table 1),

and then ligated to a similarly digested pGEX-6P-2 expression plasmid vector. The

resultant plasmids with *msa-1* inserts were transformed into BL21 competent *E. coli*

cells, and then expressed as glutathione S-transferase (GST)-fusion proteins. The

GST-fusion proteins were purified by using Glutathione Sepharose 4B (GE Healthcare),

and the GST tag was then cleaved by using a PreScission Protease (GE Healthcare) to

isolate the recombinant MSA-1 (rMSA-1) proteins.

2.6. Western blot analysis

One microgram of each purified rMSA-1 protein was analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) and subsequent western blot assay, as described previously (Yokoyama et al., 2012). Briefly, after SDS-PAGE, the reactivities of rMSA-1 proteins against a serum collected from a cow experimentally infected with the T2Bo strain of *B*. *bovis* (Terkawi et al., 2013) were determined by western blot.

3. Results

From 162 *B. bovis*-positive blood DNA samples sourced from cattle in Thailand, our PCR assays amplified *msa-1* from 84 samples. Five samples were amplified with the MSA-1 F primer and 79 with the MSA-1 F1 primer. Of these 84 DNA samples, 26, 5, and 53 were from North, Northeast, and Central regions, respectively, while *msa-1* was not amplified from the only sample sourced from Southern Thailand. All of the 84 PCR amplicons were cloned, and two clones per amplicon were sequenced. Sequencing yielded identical results for both clones from 75 samples, while the remaining nine samples, all of which were initially amplified using the MSA-1 F1 primer, produced different sequences from each of their clones. The samples yielding multiple sequences from cloned PCR amplicons were distributed across multiple regions, with four, two, and three samples from the North, Northeast, and Central regions, respectively. Therefore, a total of 93 *msa-1* gene sequences were obtained in this study and were registered with GenBank (accession numbers LC099007–LC099099), as shown in Table 2. The length distributions of *msa-1* gene sequences amplified with MSA-1 F and MSA-1 F1 primers were 990–1002 bp and 823–916 bp, respectively.

We next aligned these sequences and used the alignment to investigate the

phylogenetic relationships between *msa-1* sequences. Our phylogenetic analysis divided

the *msa-1* gene sequences as well as their translated amino acid sequences into a total of

twenty-three clades (designated as clades 1–23), which occurred within two major

clades, MSA-1 I (clades 1-14) and MSA-1 II (clades 15-23) (Fig. 1). The newly

obtained sequences were distributed across seventeen clades (clades 1–14, 16, 17, and

23). Together with clade 15 that contains the previously reported sequences from

Thailand (Nagano et al., 2013), there are at least 18 *msa-1* genotypes in Thailand. Of

these 18 clades, seven (clades 2, 4, 7, 8, 11, 13, and 16) were composed exclusively of

Thai sequences, while the rest of the clades contained a combination of Thai, Asian,

Australian, American, and African *msa-1* gene sequences (Fig. 1). The overall identity

scores among the newly generated *msa-1* gene sequences ranged from 43.5% to 100%,

while the similarity scores among the translated amino acid sequences ranged between

42.8% and 100% (Fig. 2).

Subsequently, eight rMSA-1 proteins, including one protein based on the T2Bo

msa-1 gene sequence (AF275911, A) and seven (B–H) proteins based on the newly

generated gene sequences assigned to phylogenetic clades 16 (LC099099), 17

(LC099098), 23 (LC099097), 8 (LC099044), 5 (LC099058), 14 (LC099088), and 6

(LC099012), respectively, were expressed in *E. coli* and purified. The rMSA-1 proteins

A, B, C, D, E, F, G, and H were composed of 276, 286, 301, 296, 263, 248, 234, and 259 amino acids, and had the expected molecular weights of 30.35, 31.69, 32.96, 33.09, 28.82, 26.71, 25.33, and 28.15 kDa, respectively. An equal amount of each recombinant protein was resolved by SDS-PAGE, and then analyzed by western blot. The apparent molecular weights of all of the rMSA-1 proteins as determined by the SDS-PAGE were higher than the expected values. The anti-*B. bovis* bovine serum, which had been prepared from a cow experimentally infected with the T2Bo strain of *B. bovis,* strongly reacted with the rMSA-1 of T2Bo (rMSA-1 A), as well as with rMSA-1 B–D. In contrast, no reactivity was observed with rMSA-1 E, G, or H (Fig. 3). Slight reactivity was detected with rMSA-1 F. The translated amino acid sequence of the T2Bo *msa-1* gene fragment, which was used to express rMSA-1 A, had 68.4, 59.9, 54.9, 39.7, 38.9, 36.9, and 35.0% similarity with those of the gene fragments used to express rMSA-1 B–H, respectively.

T2Bo MSA-1 contains multiple predicted B-cell epitopes, which have increased, though not complete, sequence similarity with aligned regions of rMSA-1 B, C, and D relative to other rMSA-1 sequences (Table 3). Of ten predicted epitopes of T2Bo MSA-1, three (YGAVGSF, QTKLYSVLSANF, and NLYKVNAL) shared half or more of amino acid residues with rMSA-1 B, C, and D. In addition, 4 and 5 of 8 amino acids

in epitope TSIVLPEG were conserved in rMSA-1 B and C, respectively, while 5 of 9

amino acids in epitope FNATIVSGF were shared with rMSA-1 B and D. Similarly, 5 of

7 and 5 of 9 amino acid residues in epitopes TEVPAPG and ASGVQQPPA were shared

with rMSA-1 B and C, respectively, while of 37 amino acids in epitope

MAGVCKEFLSPASDFYKLVESFDAFAHAKVHAQVGNF, 21 were conserved in

rMSA-1 B. In contrast, none of the epitopes of T2Bo MSA-1 shared high similarity

with the amino acid residues in rMSA E, F, G, and H (Table 3).

Discussion

Genetic diversity is a survival strategy that has been adapted by several protozoan parasites to escape the immune responses of mammalian hosts (Deitsch et al., 2009). *Babesia bovis* implements this strategy by maintaining a highly diverse population of *msa* genes (Berens et al., 2005; Genis et al., 2008; Leroith et al., 2005; Molad et al., 2014; Sivakumar et al., 2013; Tattiyapong et al., 2014; Yokoyama et al., 2015), which encode surface-exposed and neutralization-sensitive antigens (Mosqueda et al., 2002a,b; Suarez et al., 2000). Although MSAs are potential vaccine candidates, the genetic diversity often results in antigenicity variations (Suarez et al., 2000). Additionally, the genetic differences of *msa* genes between the vaccine strains and field isolates were thought to be associated with the disease outbreaks among the vaccinated animals (Berens et al., 2005; Leroith et al., 2005). Therefore, the diversity of *msa* genes complicates the development of subunit vaccines and broad deployment of live attenuated vaccines (Berens et al., 2005; Bock et al., 1992; Leroith et al., 2005).

Of 162 *B. bovis*-positive DNA samples analyzed in the present study, *msa-1* gene was amplified from only 84 samples. The DNA samples were tested positive for *B. bovis* by a highly sensitive nested PCR assay (Aboulaila et al., 2010), whereas

single-step PCR assays were used to amplify the *msa-1* gene. Therefore, differences in the sensitivities of diagnostic and *msa-1* PCR assays could explain this discrepancy. A recent study that analyzed the genetic diversity of *msa-1* gene in Thailand used primer sets that are not capable of detecting *msa-1* sequences amplifiable with the primer MSA-1 F1, with which most of the *msa-1* sequences were isolated in the present investigation (Nagano et al., 2013). Therefore, Nagano et al. (2013) was able to isolate only a limited number of *msa-1* gene sequences, which were highly similar and aligned into a single phylogenetic clade. In contrast, the Thai *msa-1* gene sequences analyzed in this investigation shared low identity and similarity values. Additionally, these sequences exhibited much higher diversity, as a phylogenetic analysis assigned them to seventeen additional clades. Thus, *B. bovis* is much more diverse in Thailand than previously reported. According to a recent report, the *msa-1* sequences can be classified into Asian, American, and Australian genotypes, based on their phylogenetic relationships (Liyanagunawardena et al., 2016). However, the *msa-1* sequences identified in this study clustered not only with Asian sequences, but also with sequences from America and Australia. It is significant to note that Thailand has imported large numbers of live cattle from Australia, New Zealand, Myanmar, and Costa Rica during the period from 1989 to 2000 (Food and Agriculture Organization, 2002). Therefore, the

possibility that the *B. bovis* strains introduced by importing live cattle from these countries might have contributed to the observed genetic diversity of *msa-1* gene sequences in Thailand cannot be ruled out. Nine DNA samples analyzed in the present study were detected with 2 different *msa-1* variants. This finding is not surprising as the mixed infections with multiple *msa-1* genotypes are not uncommon in infected cattle (Lau et al., 2010; Liyanagunawardena et al., 2016).

Genetic variation of *msa-1* sequences among different *B. bovis* strains sometimes leads to antigenicity differences (Leroith et al., 2005; Suarez et al., 2000). To test this, we expressed and purified recombinant proteins based on Thai *msa-1* sequences. Four of the rMSA-1 proteins (A–D) were expressed based on T2Bo *msa-1* and 3 Thai *msa-1* sequences, which occurred within the major phylogenetic clade MSA-1 II, while the other four (E–H) were expressed based on Thai *msa-1* sequences occurred within the major clade MSA-1 I. The molecular weights of the rMSA-1 proteins as determined by the SDS-PAGE were higher than the formula values. The apparent molecular weights of proteins might sometimes be different from the formula values, as the proteins, especially the membrane proteins, may not migrate according to their molecular weights (Goetz et al., 2004; Rath et al., 2009). This could explain why the apparent molecular weights of the rMSA-1 proteins were higher than the theoretical

values. Recombinant proteins exhibited differential reactivity to a bovine anti-serum produced from cattle infected with the T2Bo strain of *B. bovis*. Reactivity correlated with the degree of sequence similarity in predicted B-cell epitope regions of T2Bo MSA-1. Recombinant MSA-1 proteins derived from wild isolates with high similarity to these regions exhibited high reactivity, while low similarity in these regions conferred no reactivity. Therefore, these partially conserved epitopes could explain the results of our western blot analysis. However, it is unclear why rMSA-1 F, which shared high similarity with neither the MSA-1 of T2Bo, nor with its epitopes, was slightly reactive with anti-T2Bo serum. The possible explanations for this observation might include the presence of cross-reactive epitopes in other antigens of T2Bo strain. Although the amino acid sequence homology between of T2Bo MSA-1 and rMSA-1 E, F, G, and H was similar, the rMSA-1 F shared only 65.8, 61.7, and 63.7% similarities with rMSA-1 E, G, and H, respectively. This could explain why the anti-T2Bo serum did not react with rMSA-1 E, G, and H, even though a weak reactivity was observed with rMSA-1 F. Nevertheless, these findings suggest that the genetic diversity of *msa-1* in Thailand influences antigenicity variations of MSA-1 among *B. bovis* populations.

The gene sequences used to express rMSA-1 proteins B, C, and D were each located in a major phylogenetic clade MSA-1 II. However, the finding that each of these

proteins reacts with anti-T2Bo serum at equivalent levels as rMSA-1 from T2Bo does

not necessarily mean that all of the antigens located within this major clade are

immunologically cross-reactive. For instance, LeRoith et al. (2005) found that the

serum collected from cattle immunized with K vaccine strain did not react with the

rMSA-1 from T vaccine strain in a western blot analysis, despite the fact that both

vaccine strains are members of the major clade described above. Additionally, in a

previous investigation, antibodies against rMSA-1 of *B. bovis* R1A strain reacted with

that of Mo7, but not the other way round (Suarez et al., 2000). Therefore, additional

studies using increased numbers of rMSA-1 proteins and protein-specific antibodies are

required to identity the cross-reactive antigens that belong to different *msa-1* genotypes.

Although we did not isolate *msa-1* sequences clustering with the T2Bo *msa-1*, the

sequences isolated from Thailand in a previous investigation shared 99.0 – 100%

identity scores with T2Bo *msa-1* and were found in a clade where T2Bo *msa-1* occurred

(Nagano et al., 2013). Therefore, the findings of the present study infer that the

anti-serum against an MSA-1 antigen encoded by these previously isolated gene

sequences would possibly cross react with the Thai MSA-1 antigens occurring within

the major phylogenetic clade MSA-1 II.

In summary, analyses of large numbers of *B. bovis msa-1* gene sequences

found that the gene is highly diverse in Thailand. The findings of this study also

highlight the antigenicity variation among MSA-1 antigens encoded by the genetically

diverse gene sequences.

Acknowledgements

References

Abbas, R.Z., Zaman, M.A., Colwell, D.D., Gilleard, J., Iqbal, Z., 2014. Acaricide resistance in cattle ticks and approaches to its management: the state of play. Vet. Parasitol. 203, 6–20.

Aboulaila, M., Yokoyama, N., Igarashi, I., 2010. Development and evaluation of a nested PCR based on spherical body protein 2 gene for the diagnosis of *Babesia bovis* infection. Vet. Parasitol. 169, 45–50

Altangerel, K., Sivakumar, T., Battsetseg, B., Ueno, A., Igarashi, I., Yokoyama, N., 2012. Phylogenic relationship of Mongolian *Babesia bovis* isolates based on the merozoite surface antigen (MSA)-1, MSA-2b and MSA-2c genes. Vet. Parasitol. 184, 309–316.

Berens, S.J., Brayton, K.A., Molloy, J.B., Bock, R.E., Lew, A.E., McElwain, T.F., 2005. Merozoite surface antigen 2 proteins of *Babesia bovis* vaccine breakthrough isolates contain a unique hypervariable region composed of degenerate repeats. Infect. Immun.

73, 7180–7189.

Bock, R.E., de Vos, A.J., Kingston, T.G., Shiels, I.A., Dalgliesh, R.J., 1992. Investigations of breakdowns in protection provided by living *Babesia bovis* vaccine. Vet. Parasitol. 43, 45–56.

Bock, R.E., Jackson, L., de Vos, A.J., Jorgensen, W., 2004. Babesiosis of cattle. Parasitology 129 (Suppl.), S247–S269.

Campanella, J.J., Bitincka, L., Smalley, J., 2003. MatGAT: an application that generates similarity/identity matrices using protein or DNA sequences. BMC Bioinformatics. 4, 29.

Deitsch, K.W., Lukehart, S.A., Stringer, J.R., 2009. Common strategies for antigenic variation by bacterial, fungal and protozoan pathogens. Nat. Rev. Microbiol. 7, 493–503.

De Vos, A.J., Bock, R.E., 2000. Vaccination against bovine babesiosis. Ann. N. Y.

Acad. Sci. 916, 540–545.

Eisenhaber, B., Bork, P., Eisenhaber, F., 1999. Prediction of potential GPI-modification

sites in proprotein sequences. J. Mol. Biol. 292, 741–758.

Florin-Christensen, M., Suarez, C., Hines, S.A., Palmer, G.H., Brown, W.C., McElwain,

T.F., 2002. The *Babesia bovis* merozoite surface antigen 2 locus contains four tandemly

arranged and expressed genes encoding immunologically distinct proteins. Infect.

Immun. 7, 3566–3575.

Food and Agriculture Organization., 2002. The livestock industries of Thailand.

Genis, A.D., Mosqueda, J.J., Borgonio, V.M., Falcón, A., Alvarez, A., Camacho, M., de

Lourdes Muñoz, M., Figueroa, J.V., 2008. Phylogenetic analysis of Mexican *Babesia*

bovis isolates using msa and ssrRNA gene sequences. Ann. N. Y. Acad. Sci. 1149,

121–125.

Goetz, H., Kuschel, M., Wulff, T., Sauber, C., Miller, C., Fisher, S., Woodward, C.,

2004. Comparison of selected analytical techniques for protein sizing, quantitation and molecular weight determination. J. Biochem. Biophys. Methods 60, 281–293.

Jones, D.T., Taylor, W.R., Thornton, J.M., 1992. The rapid generation of mutation data matrices from protein sequences. Comput. Appl. Biosci. 8, 275–282.

Katoh, K., Misawa, K., Kuma, K., Miyata, T., 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 30, 3059–3066.

Kolaskar, A.S., Tungaonkar, P.C., 1990. A semi-empirical method for prediction of antigenic determinants on protein antigens. FEBS Lett. 276, 172–174.

Lau, A.O., Cereceres, K., Palmer, G.H., Fretwell, D.L., Pedroni, M.J., Mosqueda, J., McElwain, T.F., 2010. Genotypic diversity of merozoite surface antigen 1 of *Babesia bovis* within an endemic population. Mol. Biochem. Parasitol. 172, 107–112.

LeRoith, T., Brayton, K.A., Molloy, J.B., Bock, R.E., Hines, S.A., Lew, A.E.,

McElwain, T.F., 2005. Sequence variation and immunologic cross-reactivity among *Babesia bovis* merozoite surface antigen 1 proteins from vaccine strains and vaccine breakthrough isolates. Infect. Immun. 73, 5388–5394.

Liyanagunawardena, N., Sivakumar, T., Kothalawala, H., Silva, S.S., Battsetseg, B., Lan, D.T., Inoue, N., Igarashi, I., Yokoyama, N., 2016. Type-specific PCR assays for *Babesia bovis msa-1* genotypes in Asia: Revisiting the genetic diversity in Sri Lanka, Mongolia, and Vietnam. Infect. Genet. Evol. 37, 64–69.

Molad, T., Fleiderovitz, L., Leibovich, B., Wolkomirsky, R., Erster, O., Roth, A., Mazuz, M.L., Markovics, A., Shkap, V., 2014. Genetic polymorphism of *Babesia bovis* merozoite surface antigens-2 (MSA-2) isolates from bovine blood and *Rhipicephalus annulatus* ticks in Israel. Vet. Parasitol. 205, 20–27.

Mosqueda, J., McElwain, T.F., Palmer, G.H., 2002a. *Babesia bovis* merozoite surface antigen 2 proteins are expressed on the merozoite and sporozoite surface, and specific antibodies inhibit attachment and invasion of erythrocytes. Infect. Immun. 70, 6448–6455.

28

Mosqueda, J., McElwain, T.F., Stiller, D., Palmer, G.H., 2002b. *Babesia bovis* merozoite surface antigen 1 and rhoptry-associated protein 1 are expressed in sporozoites, and specific antibodies inhibit sporozoite attachment to erythrocytes. Infect. Immun. 70, 1599–1603.

Nagano, D., Sivakumar, T., De Macedo, A.C., Inpankaew, T., Alhassan, A., Igarashi, I., Yokoyama, N., 2013. The genetic diversity of merozoite surface antigen 1 (MSA-1) among *Babesia bovis* detected from cattle populations in Thailand, Brazil and Ghana. J. Vet. Med. Sci. 75, 1463–1470.

Nei, M., Kumar, S., 2000. Molecular Evolution and Phylogenetics. Oxford University Press, New York.

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

Rath, A., Glibowicka, M., Nadeau, V.G., Chen, G., Deber, C.M., 2009. Detergent

binding explains anomalous SDS-PAGE migration of membrane proteins. Proc. Natl. Acad. Sci. U.S.A. 106, 1760–1765.

Rogers, R.J., Dimmock, C.K., de Vos, A.J., Rodwell, B.J., 1988. Bovine leucosis virus contamination of a vaccine produced *in vivo* against bovine babesiosis and anaplasmosis. Aust. Vet. J. 65, 285–287.

Simking, P., Saengow, S., bangphoomi, K., Sarataphan, N., Wongnarkpet, S., Inpankwaew, T., Jittapalapong, S., Munkhjargal, T., Sivakumar, T., Yokoyama, N., Igarashi, I., 2013. The molecular prevalence and MSA-2b gene-based genetic diversity of *Babesia bovis* in dairy cattle in Thailand. Vet. Parasitol. 197, 642-648.

Sivakumar, T., Okubo, K., Igarashi, I., de Silva, W.K., Kothalawala, H., Silva, S.S.P., Vimalkumar, S.C., Meewewa, A.S. and Yokoyama, N., 2013. Genetic diversity of merozoite surface antigens in *Babesia bovis* detected from Sri Lankan cattle. Infect. Genet. Evol. 19, 134–140.

Smith, R.D., Evans, D.E., Martins, J.R., Ceresér, V.H., Correa, B.L., Petraccia, C.,

Cardozo, H., Solari, M.A., Nari, A., 2000. Babesiosis (*Babesia bovis*) stability in unstable environments. Ann. N. Y. Acad. Sci. 916, 510–520.

Suarez, C.E., Florin-Christtensen, M., Hines, S.A., Palmer, G.H., Brown, W.C., McElwain, T.F., 2000. Characterization of allelic variation in the *Babesia bovis* merozoite surface antigen 1 (MSA-1) locus and identification of a cross-reactive inhibition-sensitive MSA-1 epitope. Infect. Immun. 68, 6865–6870.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: molecular evolutionary genetics analysis version 6.0. Mol. Biol. Evol. 30, 2725–2729.

Tattiyapong, M., Sivakumar, T., Ybanez, A.P., Ybanez, R. H., Perez, Z. O., Guswanto, A., Igarashi, I., Yokoyama, N., 2014. Diversity of *Babesia bovis* merozoite surface antigen genes in the Philippines. Parasitol. Int. 63, 57-63.

Terkawi, M.A., Huyen, N.X., Wibowo, P.E., Seuseu, F.J., Aboulaila, M., Ueno, A., Goo, Y.K., Yokoyama, N., Xuan, X., Igarashi, I., 2011. Spherical body protein 4 is a new serological antigen for global detection of *Babesia bovis* infection in cattle. Clin.

Vaccine Immunol. 18, 337–342.

Yokoyama, N., Okamura, M., Igarashi, I., 2006. Erythrocyte invasion by *Babesia* parasites: current advances in the elucidation of the molecular interactions between the protozoan ligands and host receptors in the invasion stage. Vet. Parasitol. 138, 22–32.

Yokoyama, N., Sivakumar, T., Ota, N., Igarashi, I., Nakamura, Y., Yamashina, H., Matsui, S., Fukumoto, N., Hata, H., Kondo, S., Oshiro, M., Zakimi, S., Kuroda, Y., Kojima, N., Matsumoto, M., Inokuma, H., 2012. Genetic diversity of *Theileria orientalis* in tick vectors detected in Hokkaido and Okinawa, Japan. Infect. Genet. Evol. 12, 1669–1675.

Yokoyama, N., Sivakumar, T., Tuvshintulga, B., Hayashida, K., Igarashi, I., Inoue, N., Long, P. H., Lan, D. T. B., 2015. Genetic variations in merozoite surface antigens genes of *Babesia bovis* detected in Vietnamese cattle and water buffaloes. Infect. Genet. Evol. 30, 288-295.

Figure legends

Figure 1. Phylogenetic trees. Maximum likelihood phylogenetic trees were constructed using *msa-1* gene sequences (panel A) and their translated amino acid sequences (panel B) based on General Time Reversible and JTT matrix-based substitution models, respectively. The sequences determined in the present study are indicated by boldface-type letters. Note that the nucleotide as well as amino acid sequences were divided into twenty-three clades within two major clades, MSA-1 I and MSA-1 II. The letters A–F shown with the *msa-1* gene sequences in the clades 15, 16, 17, 23, 8, 5, 14, and 6, respectively, indicate the phylogenetic positions of gene sequences that were used to express the rMSA-1 proteins described in Fig. 3. Note that the newly generated Thai sequences were found in seventeen clades, while previously reported sequences in Thailand are included in clade 15.

Figure 2. Identity and similarity scores among the Thai sequences within and between phylogenetic clades. The percent identity and similarity scores among the nucleotide and translated amino acid sequences, respectively, were estimated with MatGAT. The clades refer to the phylogenetic clades as shown in Fig. 1. Clades 10, 16, and 17 contain

33

only a single Thai sequence. The similarity values are highlighted in gray. Note that the

ranges of overall identity and similarity scores were 43.5–100% and 42.8–100%,

respectively.

Figure 3. Expression of rMSA-1 proteins and western blot analysis. Equal amounts of

the rMSA-1 proteins expressed in this study were analyzed by SDS-PAGE (panel A)

and western blot assay (panel B). In both panels, M indicates the molecular size (kDa)

marker. The rMSA-1 expressed from T2Bo strain is indicated by lane A, while lanes

B–F refer to rMSA-1 proteins B–F, respectively that were expressed based on Thai

msa-1 gene sequences. Panel B was treated with anti-*B. bovis* T2Bo bovine serum. Note

that the rMSA-1 proteins A–D strongly reacted with the anti-serum, while no reaction

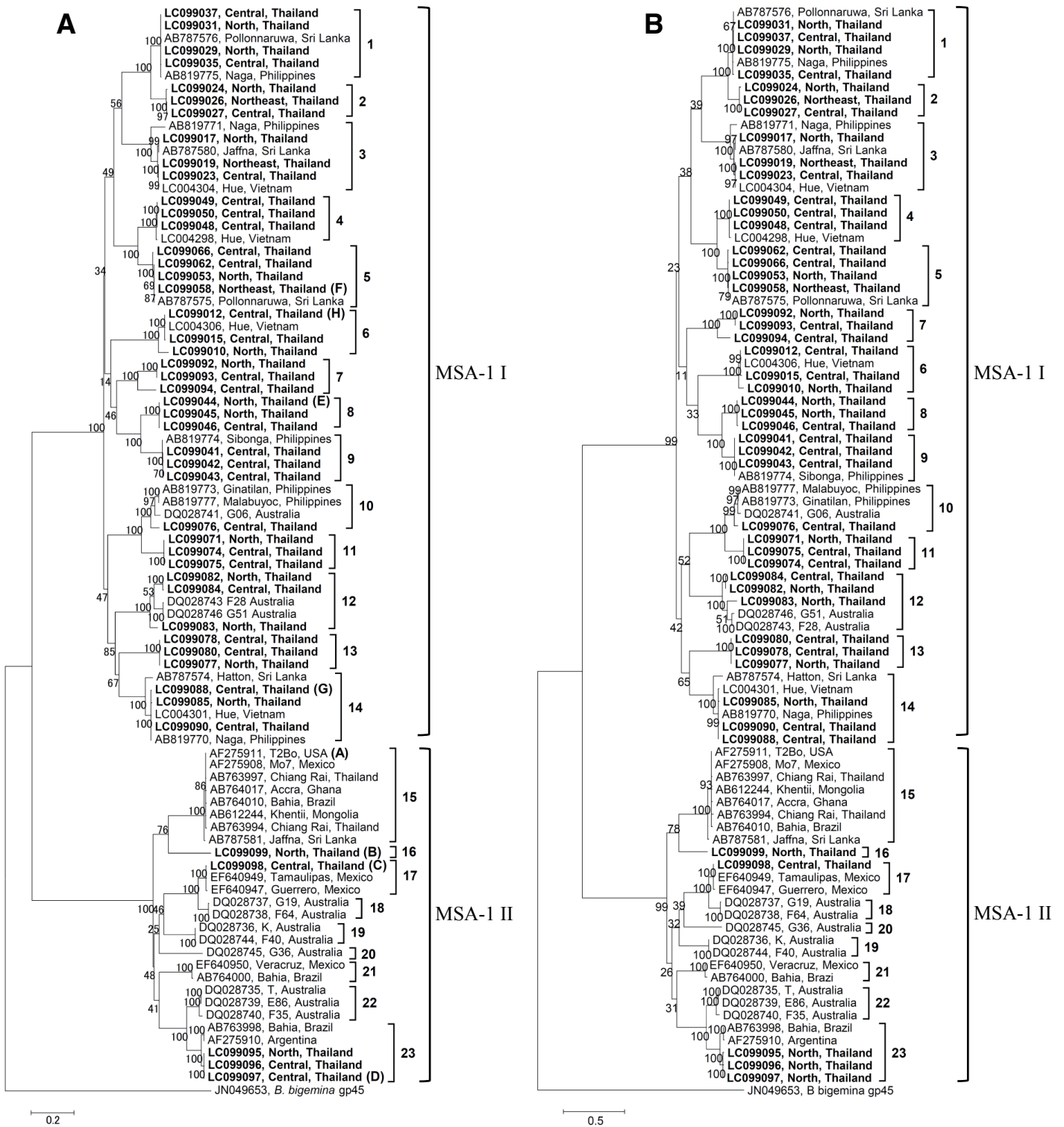was observed with rMSA-1 proteins E, G, and H.

A

100 ⌐ LC099037, Central, Thailand ⌐
     LC099031, North, Thailand
100 ⌐ AB787576, Pollonnaruwa, Sri Lanka   �‾ 1
     LC099029, North, Thailand
     LC099035, Central, Thailand
     AB819775, Naga, Philippines
100 ⌐ LC099024, North, Thailand   ⌐ 2
     LC099026, Northeast, Thailand
97   LC099027, Central, Thailand
56   AB819771, Naga, Philippines
99 ⌐ LC099017, North, Thailand   ⌐
100 AB787580, Jaffna, Sri Lanka   3
     LC099019, Northeast, Thailand
100 LC099023, Central, Thailand
49 99 LC004304, Hue, Vietnam
100 ⌐ LC099049, Central, Thailand   ⌐
100 LC099050, Central, Thailand   4
100 LC099048, Central, Thailand
     LC004298, Hue, Vietnam
100 ⌐ LC099066, Central, Thailand   ⌐
     LC099062, Central, Thailand
100 LC099053, North, Thailand   5
34 69 LC099058, Northeast, Thailand (F)
    87 AB787575, Pollonnaruwa, Sri Lanka
100 ⌐ LC099012, Central, Thailand (H) ⌐
     LC004306, Hue, Vietnam   6
100 LC099015, Central, Thailand
     LC099010, North, Thailand
14 100 ⌐ LC099092, Central, Thailand ⌐
         LC099093, Central, Thailand   7
         LC099094, Central, Thailand
46 100 ⌐ LC099044, North, Thailand (E) ⌐ 8
         LC099045, North, Thailand
         LC099046, Central, Thailand
100 100 ⌐ AB819774, Sibonga, Philippines ⌐
          LC099041, Central, Thailand   9
     100 LC099042, Central, Thailand
      70 LC099043, Central, Thailand
100 ⌐ AB819773, Ginatilan, Philippines ⌐
97   AB819777, Malabuyoc, Philippines   10
100 DQ028741, G06, Australia
100 LC099076, Central, Thailand
     100 ⌐ LC099071, North, Thailand   ⌐
          LC099074, Central, Thailand   11
     100 LC099075, Central, Thailand
47 100 ⌐ LC099082, North, Thailand   ⌐
53   LC099084, North, Thailand
         DQ028743 F28 Australia   12
    100 DQ028746 G51 Australia
    100 LC099083, North, Thailand
85 100 ⌐ LC099078, Central, Thailand ⌐
         LC099080, Central, Thailand   13
         LC099077, North, Thailand
67 AB787574, Hatton, Sri Lanka
    100 ⌐ LC099088, Central, Thailand (G) ⌐
         LC099085, North, Thailand
    100 LC004301, Hue, Vietnam   14
         LC099090, Central, Thailand
         AB819770, Naga, Philippines

MSA-1 I

     AF275911, T2Bo, USA (A)   ⌐
     AF275908, Mo7, Mexico
86   AB763997, Chiang Rai, Thailand
     AB764017, Accra, Ghana   15
100 AB764010, Bahia, Brazil
     AB612244, Khentii, Mongolia
76   AB763994, Chiang Rai, Thailand
     AB787581, Jaffna, Sri Lanka   ⌐ 16
     LC099099, North, Thailand (B)
     LC099098, Central, Thailand (C) ⌐ 17
100 EF640949, Tamaulipas, Mexico
100 EF640947, Guerrero, Mexico
     DQ028737, G19, Australia   ⌐ 18
100 DQ028738, F64, Australia
25   DQ028736, K, Australia   ⌐ 19
100 DQ028744, F40, Australia
     DQ028745, G36, Australia   ⌐ 20
48   EF640950, Veracruz, Mexico   ⌐ 21
100 AB764000, Bahia, Brazil
41   DQ028735, T, Australia   ⌐
100 DQ028739, E86, Australia   22
100 DQ028740, F35, Australia
100 AB763998, Bahia, Brazil   ⌐
     AF275910, Argentina
100 LC099095, North, Thailand   23
100 LC099096, Central, Thailand
100 LC099097, Central, Thailand (D)
     JN049653, B. bigemina gp45

MSA-1 II

0.2

B

AB787576, Pollonnaruwa, Sri Lanka ⌐
67 LC099031, North, Thailand
   LC099037, Central, Thailand   1
100 LC099029, North, Thailand
   AB819775, Naga, Philippines
100 LC099035, Central, Thailand
   LC099024, North, Thailand   ⌐ 2
   LC099026, Northeast, Thailand
100 LC099027, Central, Thailand
39 AB819771, Naga, Philippines
97 LC099017, North, Thailand   ⌐
100 AB787580, Jaffna, Sri Lanka   3
   LC099019, Northeast, Thailand
100 LC099023, Central, Thailand
38 97 LC004304, Hue, Vietnam
   LC099049, Central, Thailand   ⌐
   LC099050, Central, Thailand   4
   LC099048, Central, Thailand
   LC004298, Hue, Vietnam
100 LC099062, Central, Thailand   ⌐
   LC099066, Central, Thailand
   LC099053, North, Thailand   5
23 LC099058, Northeast, Thailand
79 AB787575, Pollonnaruwa, Sri Lanka
100 LC099092, North, Thailand   ⌐
100 LC099093, Central, Thailand   7
   LC099094, Central, Thailand
99 LC099012, Central, Thailand   ⌐
11 LC004306, Hue, Vietnam
100 LC099015, Central, Thailand   6
   LC099010, North, Thailand
33 100 LC099044, North, Thailand   ⌐ 8
   LC099045, North, Thailand
   LC099046, Central, Thailand
100 LC099041, Central, Thailand   ⌐
   LC099042, Central, Thailand   9
100 LC099043, Central, Thailand
   AB819774, Sibonga, Philippines
99 AB819777, Malabuyoc, Philippines ⌐
97 AB819773, Ginatilan, Philippines   10
99 DQ028741, G06, Australia
100 LC099076, Central, Thailand
   LC099071, North, Thailand   ⌐
   LC099075, Central, Thailand   11
100 LC099074, Central, Thailand
52 100 LC099084, Central, Thailand ⌐
   LC099082, North, Thailand
100 LC099083, North, Thailand   12
   DQ028746, G51, Australia
51 DQ028743, F28, Australia
100 LC099080, Central, Thailand   ⌐
42 LC099078, Central, Thailand   13
   LC099077, North, Thailand
65 AB787574, Hatton, Sri Lanka
   LC004301, Hue, Vietnam
100 LC099085, North, Thailand
   AB819770, Naga, Philippines   14
99 LC099090, Central, Thailand
   LC099088, Central, Thailand

MSA-1 I

AF275911, T2Bo, USA   ⌐
AF275908, Mo7, Mexico
93 AB763997, Chiang Rai, Thailand
   AB612244, Khentii, Mongolia   15
100 AB764017, Accra, Ghana
   AB763994, Chiang Rai, Thailand
78 AB764010, Bahia, Brazil
   AB787581, Jaffna, Sri Lanka
   LC099099, North, Thailand   ⌐ 16
   LC099098, Central, Thailand   ⌐ 17
100 EF640949, Tamaulipas, Mexico
100 EF640947, Guerrero, Mexico
99 39 DQ028737, G19, Australia   ⌐ 18
100 DQ028738, F64, Australia
32 DQ028745, G36, Australia   ⌐ 20
   DQ028736, K, Australia   ⌐ 19
100 DQ028744, F40, Australia
26 EF640950, Veracruz, Mexico   ⌐ 21
100 AB764000, Bahia, Brazi
31 DQ028735, T, Australia   ⌐
100 DQ028739, E86, Australia   22
   DQ028740, F35, Australia
100 AB763998, Bahia, Brazil   ⌐
   AF275910, Argentina
100 LC099095, North, Thailand   23
100 LC099096, North, Thailand
100 LC099097, North, Thailand
   JN049653, B bigemina gp45

MSA-1 II

0.5

Fig. 1

Identity (%)

Clades	1	2	3	4	5	6	7	8	9	10	11	12	13	14	16	17	23	Clades
	97.5-100	82.8-85.1	68.7-70.3	67.1-68.6	67.0-71.0	65.3-66.8	66.4-68.7	64.9-65.5	65.5-67.8	64.4-65.6	64.6-65.9	62.3-64.6	62.9-64.3	62.7-64.1	48.0-49.6	47.2-49.3	45.7-47.5	1
1	97.7-100	97.2-100	68.5-69.7	66.1-68.0	65.8-67.6	63.9-65.8	65.9-67.7	65.0-65.5	65.3-66.6	61.2-62.3	62.6-63.7	63.9-64.5	63.0-63.3	63.1-64.2	45.6-45.8	46.9-47.7	45.0-45.9	2
2	80.6-83.7	96.1-100	98.3-100	69.1-69.3	68.0-70.1	68.5-71.6	65.8-67.9	67.3-68.0	67.2-67.9	65.7-66.6	67.1-68.3	66.1-67.5	67.3-67.8	64.8-65.8	48.2-48.6	48.0-48.7	45.3-46.5	3
3	71.7-74.2	70.6-71.6	97.6-100	100	80.7-85.4	62.6-66.4	67.6-70.0	70.6-70.8	66.4	65.7	68.3-68.4	68.3-69.0	69.4-69.5	67.5-68.0	46.2	47.6	47.3-47.4	4
4	68.3-68.7	68.3-69.3	70.9-72.0	100	88.9-100	62.4-66.8	67.4-71.1	66.6-71.5	67.7-68.7	67.7-69.0	68.7-69.6	66.2-69.1	69.1-70.3	65.8-69.0	46.7-48.4	46.7-48.7	47.9-49.0	5
5	67.0-69.6	65.2-69.3	70.6-73.0	82.6-87.1	86.6-100	85.1-100	63.8-68.7	64.6-70.0	65.3-70.1	64.1-68.5	63.5-67.3	63.6-66.3	63.6-67.8	64.7-66.9	45.8-49.7	46.4-49.9	44.1-49.8	6
6	63.3-67.0	65.9-69.4	64.9-68.9	62.7-66.6	61.9-68.6	83.9-100	85.7-100	69.1-71.2	70.2-72.5	70.7-71.2	72.0-72.6	67.7-69.1	71.3-73.2	67.6-69.5	49.4-51.7	49.4-50.2	47.3-47.7	7
7	69.2-72.0	66.6-68.8	66.6-67.9	67.8-69.6	64.4-72.2	64.0-69.6	84.9-100	99.4-100	82.6-83.1	67.3-67.8	66.1-66.5	67.2-67.8	70.3-70.6	65.2-66.4	49.4-49.7	49.5-49.8	46.4-47.2	8
8	65.0-67.0	65.8-66.8	65.5-66.9	67.5-68.1	64.9-71.2	65.6-67.1	65.1-68.6	98.6-100	100	68	66.4-66.5	65.4-66.1	68.4-68.5	64.9-65.0	49.7	50.9	48.6-48.7	9
9	66.1-67.8	65.4-67.4	65.1-66.1	66.1	68.4-71.1	65.1-67.8	67.1-71.4	83.4-84.4	100	100	84.4-84.5	67.4-68.0	70.6-70.7	66.5-66.8	48.8	49.1	47.7-47.8	10
10	68.1-70.7	65.8-66.8	65.1-66.1	67.4	68.1-68.8	65.1-69.7	69.7-71.4	64.8-65.5	69.1	100	99.9-100	68.5-68.8	70.1-70.3	66.7-67.1	47.6	47	47.2-47.3	11
11	70.0-71.7	65.6-66.0	65.2-66.6	68.4-68.7	68.6-71.8	64.6-69.2	68.8-72.2	64.1-65.4	66.1	85.5-85.9	99.7-100	92.2-99.9	71.5-71.9	76.9-77.6	47.0-47.1	45.1-46.2	45.0-45.8	12
12	62.3-65.0	64.9-70.3	60.8-63.9	66.2-69.3	62.9-70.4	63.9-68.1	66.1-67.6	64.1-65.1	61.8-62.8	67.4-68.1	66.0-66.7	90.7-99.6	99.8-100	73.0-73.4	48.6-48.9	48.2-48.3	47.2-47.5	13
13	65.7-66.7	62.7-63.7	65.2-66.6	69.2-69.5	67.1-68.2	65.4-66.9	65.8-66.4	65.1-66.1	65.4	69.4-69.7	68.5-68.8	66.1-66.8	99.3-100	99.5-100	44.9-45.1	43.8-44.0	43.5-44.5	14
14	63.0-64.6	67.0-69.7	63.2-64.5	67.2-67.6	64.9-67.9	64.2-67.6	65.2-65.8	59.7-60.7	62.1-62.5	65.1	61.9-63.6	72.4-75.6	70.5-70.9	100	100	69.8	61.8-61.9	16
16	45.6-46.2	46.8-47.7	47.4-48.0	44.6	47.7-50.2	45.3-48.0	45.6-48.9	47.1-48.0	50.5	47.4	44	45.0-45.3	45.0-45.3	42.8	100	100	62.9-63.0	17
17	50.5-51.4	43.8-44.1	47.4-48.0	48.9	45.9-48.9	46.8-49.2	45.9-47.7	46.8-47.4	49.5	50.2	48.6-49.2	43.5-44.7	48.0-48.3	45	68.6	100	99.8-99.9	23
23	46.0-47.0	47.0-47.6	48.2-49.4	49.1	49.4-50.9	47.0-48.5	48.8-49.4	45.7-46.0	48.8-49.1	49.7	46.6	43.0-46.3	46.3-46.6	45.1-45.7	63.7-64.0	62.5-62.8	99.7-100	

Similarity (%)

Fig. 2

A

M A B C D E F G H M

50

37

25

B

M A B C D E F G H M

50

37

25
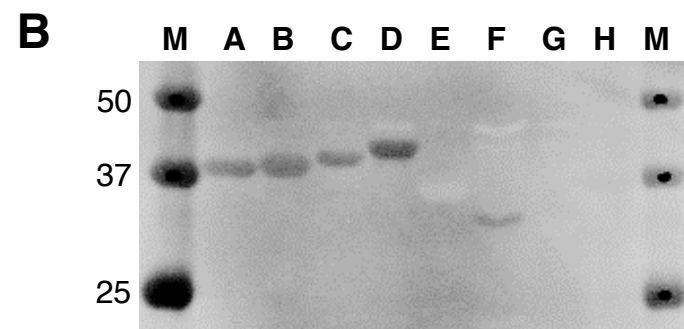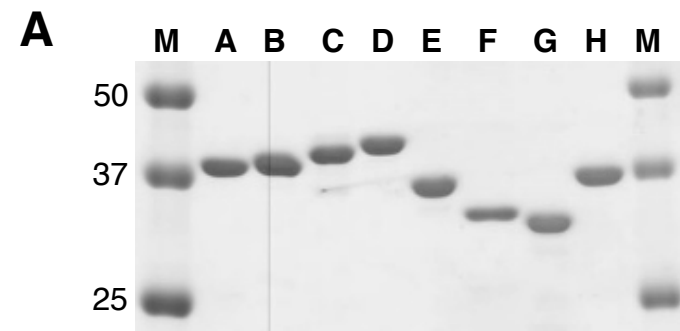
Fig. 3

Table 1. Primers used to amplify *msa-1* gene fragments for protein expression

Protein	Acc. No.	Primers sequences (5'-3')[a]	
		Forward[b]	Reverse[c]
A	AF275911	gcggatccGGTGAAGAATTGACACAATCC	gccctcgagtcaTGGCTTGGTTGTACCCTGTTGTCC
B	LC099099	gcggatccGAAGAGTTGACTCAACATCAGTCC	gccctcgagtcaGGGCTGGTTTCCCTGACC
C	LC099098	gcggatccGTTACGTCGTCAGAACCAAGG	gccctcgagtcaATAGGTGAATGAAGATTTGG
D	LC099097	ccaggaattcccTCGTCATCTGGAGATTTGG	gccctcgagtcaAAAGGTGAAGGAAGATTTGG
E	LC099044	attcccgggGCTCCCGTCCAGGTCAATTC	gccctcgagtcaAAAGGTGAATGAAGATTTGG
F	LC099058	gcggatccGAACAGGAGAATTCTGTCTCAACC	gccctcgagtcaTGGTTTGCCATGTCCATTGAGGTTACC
G	LC099088	ccaggaattcccACTTCCGTCCCGGGTAATTC	gccctcgagtcaTGGCTTGGATGGGTTAGAAC
H	LC099012	gcggatccGTCCCGGAGAACACTAAAAAGG	gccctcgagtcaGGTGGGCTTAGGGGTCTCAGCAGG

[a] The nucleotide sequences of primers corrospond to the sequences of template *msa-1* gene are indicated by upper case letters, while the restriction sites in each primer are underlined.
[b] Forward primers contained *Bam*HI (A, B, C, F, and H), *Eco*RI (D and G), and *Sma*I (E) restriction sites.
[c] All reverse primers contain an *Xho*I restriction site and a reverse complement of TGA stop codon.

Table 2. *B. bovis msa-1* gene sequences isolated from cattle populations reared in different geographical regions in Thailand

Clades[a]	No. of sequences (Accession number)		
	North	Northeast	Central
1	4 (LC099029 – LC099032)	ND	8 (LC099033 – LC099040)
2	2 (LC099024, LC099025)	1 (LC099026)	2 (LC099027, LC099028)
3	2 (LC099017, LC099018)	1 (LC099019)	4 (LC099020 – LC099023)
4	ND	ND	3 (LC099048 – LC099050)
5	4 (LC099051 – LC099054)	5 (LC099055 – LC099059)	11 (LC099060 – LC099070)
6	5 (LC099007 – LC099011)	ND	5 (LC099012 – LC099016)
7	1 (LC099092)	ND	2 (LC099093, LC099094)
8	2 (LC099044, LC099045)	ND	2 (LC099046, LC099047)
9	ND	ND	3 (LC099041 – LC099043)
10	ND	ND	1 (LC099076)
11	2 (LC099071, LC099072)	ND	3 (LC099073 – LC099075)
12	2 (LC099082, LC099083)	ND	1 (LC099084)
13	1 (LC099077)	ND	4 (LC099078 – LC099081)
14	3 (LC099085 – LC099087)	ND	4 (LC099088 – LC099091)
16	1 (LC099099)	ND	ND
17	ND	ND	1 (LC099098)
23	1 (LC099095)	ND	2 (LC099096, LC099097)

ND, not detected

[a] Clades refer to the phylogenetic clades containing *msa-1* sequences as shown in Fig. 1

Table 3. Diversity of amino acid residues of predicted B-cell epitopes of T2Bo MSA-1

B-cell epitope in A (T2Bo)		No. of conserved amino acids						
Epitope	No. of amino acids	B	C	D	E	F	G	H
TSIVLPEG	8	4	5	2	2	0	0	1
YGAVGSF	7	6	6	5	0	0	0	0
QTKLYSVLSANF	12	7	8	6	2	3	1	2
NLYKVNAL	8	6	5	4	3	1	1	2
FNATIVSGF	9	5	4	5	1	1	1	0
AVEYFKKHVYTGEHVVDVN	19	10	12	11	6	7	8	7
MAGVCKEFLSPASDFYKLVESFDAFAHAKVHAQVGNF	37	21	18	16	2	4	2	2
APPKDVTDA	9	4	1	2	2	2	1	0
TEVPAPG	7	5	2	1	1	2	2	1
ASGVQQPPA	9	4	5	2	1	0	2	0