Infection, Genetics and Evolution 79 (2020) 104161

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Research paper

Use of molecular markers can help to understand the genetic diversity of *Babesia bovis*

Paulo Vitor Marques Simas[a,b,*,1,2], César Cristiano Bassetto[a,b,1], Rodrigo Giglioti[c], Cintia Hiromi Okino[b], Henrique Nunes de Oliveira[a], Márcia Cristina de Sena Oliveira[b]

[a] Faculdade de Ciências Agrárias e Veterinárias, Universidade Estadual Paulista Júlio de Mesquita Filho, Jaboticabal, São Paulo, Brazil
[b] Embrapa Pecuária Sudeste, São Carlos, São Paulo, Brazil
[c] Centro de Pesquisa de Genética e Reprodução Animal, Instituto de Zootecnia, Nova Odessa, São Paulo, Brazil

ARTICLE INFO

Keywords:
Babesia bovis
Molecular markers
Diversity
Virulence
Attenuation

ABSTRACT

Cattle babesiosis is a tick-borne disease responsible for significant losses for the livestock industries in tropical areas of the world. These piroplasms are under constant control of the host immune system, which lead to a strong selective pressure for arising more virulent or attenuated phenotypes. Aiming to better understand the most critical genetic modifications in *Babesia bovis* genome, related to virulence, an *in silico* analysis was performed using DNA sequences from GenBank. Fourteen genes (*sbp-2, sbp-4, trap, msa-1, msa-2b, msa-2c, Bv80* (or *Bb-1*), *18S rRNA, acs-1, ama-1, β-tub, cp-2, p0, rap-1a*) related to parasite infection and immunogenicity and ITS region were selected for alignment and comparison of several isolates of *Babesia bovis* from different geographic regions around the world. Among the 15 genes selected for the study of diversity, only 7 genes (*sbp-2, sbp-4, trap, msa-1, msa-2b, msa-2c, Bv80*) and the *ITS region* presented sufficient genetic variation for the studies of phylogeny. Despite this genetic diversity observed into groups, there was not sufficient information available to associate molecular markers with virulence of isolates. However, some genetic groups no were correlated with geographic region what could indicate some typical evolutionary characteristics in the relation between parasite-host. Further studies using these genes in herds presenting diverse clinical conditions are required. The better understanding of evolutionary mechanisms of the parasite may contribute to improve prophylactic and therapeutic measures. In this way, we suggest that genes used in our study are potential markers of virulence and attenuation and have to be analyzed with the use of sequences from animals that present clinical signs of babesiosis and asymptomatic carriers.

1. Introduction

Cattle fever caused by intraerythrocytic tick-transmitted protozoan *Babesia bovis* and *B. bigemina* are distributed worldwide and causes significant economic impact on the livestock industry (Bock et al., 2004; Gohil et al., 2013).The clinical signs include anemia, due to intravascular hemolysis caused by cyclical replication of the parasite, consequently, infected animals become severely depressed and may die. *B. bovis* infections are accompanied by the accumulation of parasitized RBCs in the microvasculature and the development of severe clinical complications such as cerebral babesiosis, respiratory distress syndrome and multi-organ failure (Bock et al., 2004; Gohil et al., 2010).

These parasites undergo a constant pressure of the immune system of their hosts, which leads to the emergence of genetic variations.(Lau et al., 2011; Pedroni et al., 2013). According to Lau et al. (2011), pathogen adaptation is a dynamic and continuous process; the acquisition and/or loss of virulence are dynamic within a population of pathogens, varying with the host genetics, the host's immune status and transmission efficiency. Modifications associated to emergence of different phenotypes can be originated during the sexual or asexual stages of the parasite life cycle due to various mechanisms. During mitosis, mutation, chromosome breakage, sister chromatid exchange, gene conversion or interchromosomal exchange may occur, whereas, in meiosis, chromosome rearrangement occurs, recombination between homologous or heterologous chromosomes, exchange of sister chromatids or gene conversion (Taylor et al., 2000).

* Corresponding author.
E-mail address: simaspvm@gmail.com (P.V.M. Simas).
[1] The authors have the same contribution on this work.
[2] Present address: Washington Luiz Road, Km 234 no number, Canchim Farm, Postal Code: 13560-970, São Carlos City, São Paulo State, Brazil.

https://doi.org/10.1016/j.meegid.2019.104161
Received 2 August 2019; Received in revised form 13 December 2019; Accepted 23 December 2019
Available online 24 December 2019
1567-1348/

Table 1
GenBank sequences reference number of Gene/Region from *Babesia bovis*used to conduct this study.

Gene/Region	Reference number on GenBank
sbp-2 (Spherical Body Protein-2)	AB742543.1 to AB742551.1, AB772317.1 to AB772322.1, KT460089.1 to KT460093.1 and MG725962.1
sbp-4 (Spherical Body Protein-4)	KF626629.1 to KF626638.1, KY484522.1 to KY484534.1, KY562844.1 to KY562847.1
trap (Thrombospondin-Related Anonymous Protein)	FJ423465.1 to FJ423472.1, FJ42346574.1, FJ588019.1 to FJ588023.1
msa-1 (Merozoite Surface Antigen-1)	AB612244.1 to AB612251.1, AB763993.1 to AB763997.1, AB764000.1, AF275911.1, EF640942.1 to EF640954.1, LC099095.1 to LC099097.1, LC099099.1
msa-2b (Merozoite Surface Antigen-2b)	KX384021.1, KX384023.1 to KX384026.1, NW 001820854.1
msa-2c (Merozoite Surface Antigen-2c)	HM352731.1 to HM352735.1, KX384028.1 to KX384033.1, KX384035.1 to KX384039.1
Bv-80 (Rhoptry Protein Of Merozoite)	AY727909.1, JX088041.1, JX088042.1, JX088044.1 to JX088046.1, JX088049.1, KF111549.1 to KF111555.1
ITS region (Internal Transcribed Spacer)	EF458270.1, EF458275.1 to EF458278.1, EF458280.1, EF458285.1 to EF458292.1, EF458297.1 to EF458302.1, HQ264121.1, HQ264122.1, KF928960.1

Apicomplexan hemoparasites exist as subpopulations expressing distinct virulence phenotypes, which may be resulted from gene mutation or differential regulation of gene expression (Carson et al., 1990). Changes in the population structure of *B. bovis* throughout the genome were demonstrated for the first time, and no specific gene responsible for virulence was identified (Lau et al., 2011). However, several genes related to replication and interaction between host-parasite should be investigated since may be useful for labeling and differentiating these subpopulations of parasites with phenotype of stable virulence in asymptomatic carriers host.

The genome of *B. bovis* was described in 2007 (Brayton et al., 2007) and since then some studies have been carried out to characterize genetic diversity. Some genes (*sbp-2* (spherical body protein 2), *sbp-4* (spherical body protein 4), *trap* (thrombospondin-related anonymous protein), *msa-1* (merozoite surface antigen 1), *msa-2b* (merozoite surface antigen 2b), *msa-2c* (merozoite surface antigen 2c), *Bv80* (or *Bb-1*, merozoite protein), *18S rRNA, acs-1* (acyl-CoA synthetase 1), *ama-1* (apical membrane antigen 1), *β-tub* (tubulin beta chain), *cp-2* (cysteine peptidase 2), *p0* (ribosomal phosphoprotein) and *rap-1a* (rhoptry associated protein 1a) and also *ITS region* (internal transcribed spacer (Brown et al., 1996; Hilpertshauser et al., 2007; Mazuz et al., 2012;Ramos et al., 2012; Sivakumar et al., 2013; Combrink et al., 2014; Molad et al., 2014; Molad et al., 2015; Niu et al., 2015; Yokoyama et al., 2015; Tattiyapong et al., 2016; Matos et al., 2017; Mtshali & Mtshali, 2017; Rittipornlertrak et al., 2017; Gallego-Lopez et al., 2018)) pointed for the production of highly immunogenic proteins have been targets of these studies. However, it remains not elucidated which gene or which of these genes would be a good marker for genotyping studies regarding attenuation or virulent phenotypes. In this context, this work was carried out a comparison of several *B. bovis* sequences of thirteen genes and the *ITS region* obtained from different geographic regions, deposited in GenBank. *In silico* analyzes were performed to identify genes that presentenough genetic diversity to be used in molecular marker studies for virulence or attenuation.

2. Material and methods

2.1. Ethical statement and declaration of Interest

The work was conducted by *in silico* analysis and then there was no necessary to submit this research to any ethical statement. Moreover all authors confirm none declaration of interest.

2.2. Phylogenetic analysis

We selected several genes and a genomic region related to parasite infection and immunogenicity. We constructed phylogenetic trees using the genes *sbp-2* (spherical body protein 2), *sbp-4* (spherical body protein 4), *trap* (thrombospondin-related anonymous protein), *msa-1* (merozoite surface antigen 1), *msa-2b* (merozoite surface antigen 2b), *msa-2c* (merozoite surface antigen 2c) and *Bv80* (or *Bb-1*, merozoite protein) and also ITS region (internal transcribed spacer). These genes/regions were selected based in previous diversity analysis from published data available on database (Bing-Zang et al., 2016; Brown et al., 1996; Molad et al., 2015; Flores et al., 2012; Niu et al., 2015; Ramos et al., 2012; Tattiyapong et al., 2016; Mazuz et al., 2012; Sivakumar et al., 2013; Molad et al., 2014; Yokoyama et al., 2015; Combrink et al., 2014; Hilpertshauser et al., 2007; Mtshali & Mtshali, 2017; Matos et al., 2017; Rittipornlertrak et al., 2017; Gallego-Lopez et al., 2018). The genes *18S rRNA*, *acs-1* (acyl-CoA synthetase 1), *ama-1* (apical membrane antigen 1), *β-tub* (tubulin beta chain), *cp-2* (cysteine peptidase 2), *p0* (ribosomal phosphoprotein) and *rap-1a* (rhoptry associated protein 1a) were previously selected but after some analysis they were excluded (Table 1S – Supplementary material).

The sequences were obtained from GenBank and the inclusion's criteria were the longest sequences for each gene/region and that did not disturb the alignment. The sequences used according GenBank' references were grouped into Table 1.

The phylogenetic analysis was conducted in MEGA7 (Kumar et al., 2016) using the Maximum Likelihood method with 1000 replicates of bootstrap. The best evolution model was inferred by JModelTest (Darriba et al., 2012).

3. Results and discussion

In the case of epidemiological studies, it is very common to find certain strains of parasites circulating regionally (Howes et al., 2016). However, a common feature of the selected molecular markers seemed to be that the genotypes were not restricted to geographical regions and in these cases they could be associated with the strain phenotype of virulence, as observed in previous studies (Wilkowsky et al., 2009). On the other hand, in many cases, there was not sufficient information related to the sequence's source.

Allred (2019) purposed a distinction between “variable” and “variant” proteins that may contribute to persistence of parasites into their hosts. A “variable” multigene protein family contains two ormore members, which differs in sequence and rarely recombine each other and are stable in sequence during the course of an infection within an individual host, and may beco-expressed within an individual parasite. On the other hand, a “variant” multigene protein family contains more than several members, which differs in sequence, is generally expressed monoparalogously or monoallelically and may recombine with others as an integral aspect of the family's biology, and the expressed member is routinely altered in sequence during the course of an infection within anindividual host. The genes rap1, msa1, sbp2, and 6-cys fall into the variable category and their diversity coupled with their relativestability make them good markers for use in epidemiologic studies at the isolate level. In this way, is very important to understand the genetic pattern of these genes.

The analysis performed using the genes *18S rRNA*, *acs-1*, *ama-1*, *β-tub*, *cp-2*, *p0* and *rap-1a* (Supplementary material) were excluded because the alignments showed sites highly conserved and due the low

diversity and due this low diversity level the phylogenetic tree presented no resolution in clades.

Still, the sequences deposited on GenBank for *Bv80*, *sbp-2*, *sbp-4*, *trap*, *msa-1*, *msa-2b*, *msa-2c* and *ITS region* presented low level of conservation into some sites and they were analyzed more carefully.

3.1. Spherical body protein-1 or Bv80 or Bb-1

Bv80, also known as *Bb-1*, is a single-copy gene encoding the *B. bovis* spherical body protein-1, with conserved 5′- and 3′-terminal regions separated by variable-length tandem repetitive sequences, which are responsible for the polymorphism of the corresponding genes (Brayton et al., 2007). This variable portion that includes the repetitive sequences of *Bv80* has been used to differentiate subpopulations of *B. bovis* (Wilkowsky et al., 2009). The virulent field isolates and the attenuated strains were already identified based in intra- and inter-strain diversity in the *Bv80* gene (Mazuz et al., 2012), and the selection has been associated with the reduction of allelic *Bv80* diversity (Baravalle et al., 2012).

We analyzed the *Bv80* nucleotide diversity using 14 sequences deposited on GenBank. Although there are few available provided by only two different studies, we noted that there was a separation in two specific clades between attenuated strains and isolated field strains (Fig. 1). However, just a sequence (AY727909.1) that presented the most significant distance between field strains is related to a vaccine strain used in Israel. One possible explanation for this fact is that the variable portion of the gene has tandem repeats. This sequence, because it did not have repetition insertionevidenced in the attenuated strains, was grouped with the field isolates, already indicating a distance of these sequences. New studies should be conducted to understand the evolutionary profile of this gene to be used like a molecular marker attenuation and virulence.

3.2. Internal transcribed spacer 1 and 2

The rRNA internal transcribed spacer 1 (*ITS1*) and internal transcribed spacer 2 (*ITS2*) have been used as a diversity marker in phylogenetic analyzes for the separation of closely related species, recognition of new species, determination of conspecificity between isolates, discrimination within a species, and differentiation between piroplasm species and subspecies (Zahler et al., 1998; Collins and Allsopp, 1999; Aktas et al., 2007). This sequence is useful in molecular studies since it is flanked by SSU rRNA and large subunit rRNA genes, highly conserved between closely related species. The ITS regions are separated by the 5.8S gene but are not subject to the same functional constraints as the rRNA genes, making this region subject to higher evolutionary rates and high level of nucleotide diversity (Hillis and Dixon, 1991).

In our study, we performed a phylogenetic analysis using 23 sequences from GenBank. All sequences were deposited by the same researchers and were related to samples collected from several locations of America (Argentina, Brazil, Mexico, Uruguay and USA), Africa (South Africa), Asia (India) and Oceania (Australia). Curiously, the strains were not strictly grouped by location indicating that the evolutionary profile of this genomic region presents something different related to the interaction parasite-host. The phylogenetic tree showed two clades (A and B) and each one presented subclades (AI, AII, BI and BII) and two sub groups for subclade AI and AII, respectively (AI-1, AI-2, AII-1 and AII-2), as evidenced in Fig. 2. Intending to confirm this classification, a greater number of sequences should be included in future analyzes, from different sites (including attenuated and pathogenic strains).

3.3. Merozoite surface antigens 1, 2b and 2c

There is a protein group distributed on the merozoite surface that has been a target to vaccine development due to their antigenic properties, *MSA-1* and *MSA-2*. The blocking induction of merozoites and sporozoites invasion in immunized animals is accomplished by the interaction of the antibodies with the MSA proteins (Mosqueda et al., 2002). These proteins present the regions amino- and carboxy-terminal highly conserved but the exposed portion (central region) is variable (Hines et al., 1992). The hypervariable part presents a high number of proline, which is evolved in the binding site into the host cell (LeRoith

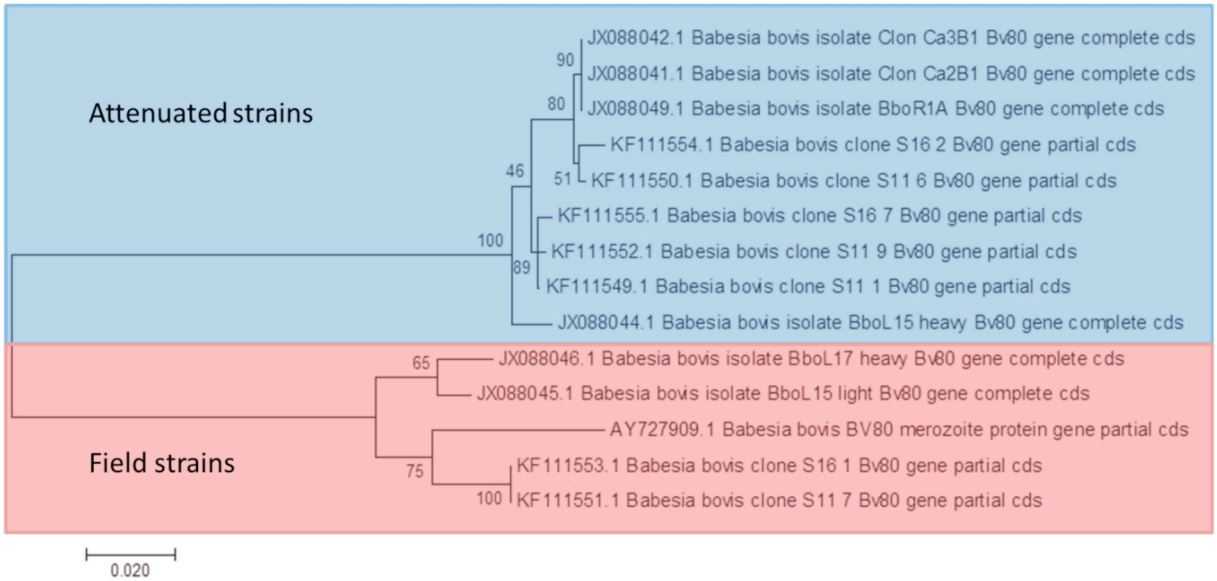

Fig. 1. Phylogenetic analysis of *Bv80* gene from *Babesia bovis* using sequences deposited on GenBank. A: Attenuated strains; B: Field strains. The evolutionary history was inferred by using the Maximum Likelihood method based on the Kimura 2-parameter model. The percentage of trees in which the associated taxa clustered together is shown next to the branches. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+G, parameter = 0.3306)) according to JModelTest appointment. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 14 nucleotide sequences. All positions containing gaps and missing data were eliminated. There were a total of 682 positions in the final dataset. Evolutionary analyses were conducted in MEGA7.

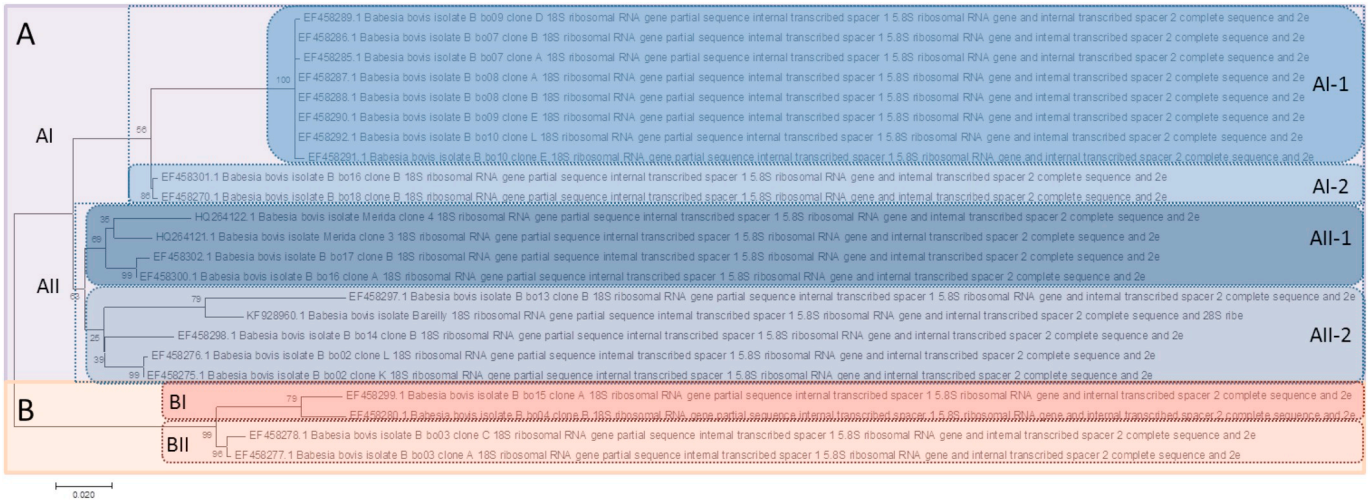

Fig. 2. Phylogenetic analysis of *ITS* regions from *Babesia bovis* using sequences deposited by GenBank. A: Genogroup A; AI: subgroup I of genogroup A; AII: subgroup II of genogroup A; B: Genogroup B; BI: subgroup I of genogroup B; BII: subgroup II of genogroup B. The evolutionary history was inferred by using the Maximum Likelihood method based on the Tamura 3-parameter model. The percentage of trees in which the associated taxa clustered together is shown next to the branches. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+G, parameter = 0.1657)). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 23 nucleotide sequences. All positions containing gaps and missing data were eliminated. There were a total of 583 positions in the final dataset. Evolutionary analyses were conducted in MEGA7.

et al., 2006). The *msa-1* and *msa-2* genes are single-copy genes located into chromosome 1 and they are members of the variable merozoite surface antigen family. The *msa-2* locus is organized into four tandemly arranged and expressed genes (two *msa-2a* genes, *msa-2b*, and *msa-2c*) (Brayton et al., 2007). The *MSA-2a* and *MSA-2b* also can be associated with erythrocytes invasion. (Florin-Christensen et al., 2002). The hypervariable gene portion was associated to the characteristics related to immunogenicity and neutralization by antibodies indicates this region to undergo strong selection pressure and could be considered as an excellent indicator of attenuation or pathogenicity.

Based on these considerations, we performed phylogenetic analyses using all sequences from GenBank for *msa-1*, *msa-2b* and *msa-2c*. Regarding to *msa-1* sequences, we could identify sequences from several locals, virulent strains, field isolates, from epidemiological screening and as we observed in the *ITS* regions analyzes, there was not grouping by geographic site. The*msa-1* presented three different clades, which we suggest like I, II and III (Fig. 3). The clade I contain sequences from four different work groups, with deposited sequences from Thailand, Brazil, Ghana, Mexico, USA and Mongolia and it was divided into two well-defined groups, I-1 and I-2; into the group I-1 was included the sequence T2B, the virulent strain from Texas (USA). Even, the clade II was generated just with sequences related to field isolates from Mexico. Like the clade I, the clade III was divided into two groups, III-1 and III-2 with strains from Brazil, Mexico, and Thailand, with the III-2 group contained just sequences from Thailand. We hadn't enough information related to clinical conditions of animals, *e.g.*, if the sequences were from sick or healthy animals. However, we considered that this gene is under intense selection pressure and future studies can indicate differences between attenuated and pathogenic strains.

The analysis performed with *msa-2b* gene indicated the formation of two clusters (Fig. 4); however, just six sequences were used in this analysis and all of these are from just a single research group from Brazil related to the unpublished data about genetic diversity and phylogenetic analysis of *B. bovis* in cattle from Mozambique. The samples were obtained from healthy animals, without clinical signals, and from endemic regions. The cluster I presented the strain T2Bo, the attenuated strain from Texas (USA). Due to the origin and the small number of sequences, it is not possible to conclude if there is any relation between attenuation and pathogenicity in the formation of the clades. Strains from endemic regions are present in animals that act as reservoirs but may eventually lead to clinical cases, either by the appearance of a variant with an adaptive advantage for pathogenicity or by cattle immunological failures. By the way, the gene presented enough variation to conduct studies related to the genetic diversity of *B. bovis*.

For the gene *msa-2c*, we used 16 sequences from GenBank. These sequences are from Mozambique and the all regions from Brazil (South, Southeast, North, Northeast and Midwest). We also obtained two clades (Fig. 5). In the first one, the sequences from Brazil formed a monophyletic group, with low variation degree, but other four sequences from Mozambique grouped with a reduced resolution into the cluster I-1. The Brazilian isolates were obtained from endemic regions with animals naturally infected. However, in an isolated experiment conducted with splenectomized animals, there was developing babesiosis and some animal died. These sequences are more distant phylogenetically of the other strains into the same clade, which could indicate some variation that confers an adaptive advantage of the parasite and increases its capacity to cause disease, since the splenectomy reduces the selection pressure acting on the parasite. The clade I-2 presented just sequences from Mozambique and the sequences are very similar with no evidence of differentiation between them. The cluster II contains only sequences from Mozambique and healthy animals. It is worth noting that the search for attenuation and virulence markers *in silico* or *in vitro* studies is essential since the use of animals for this purpose has many ethical implications.

3.4. Thrombospondin-related anonymous protein (trap)

The *trap* gene is located in chromosome 2 and contains five exons, 2128 nucleotides, which encode 660 amino acids. The region between 1 and 26 is associated with signal peptide and from 27 to 660 is related to the protein chain (Brayton et al., 2007). This is a protein family member identified into the micronemal structure in all apicomplexan parasites and is involved in interaction cell-matrix process (Morahan et al., 2009). These proteins can be immunogenic potential since recombinant proteins of homologues genes from *B. bovis* and *B. gibsoni* were capable to inhibit the parasite host invasion (Zhou et al., 2006). Considering that these proteins are expressed on the surface of the primary stages of the parasite and thus have immunogenic potential, the gene is under strong selection pressure of the host's immune system

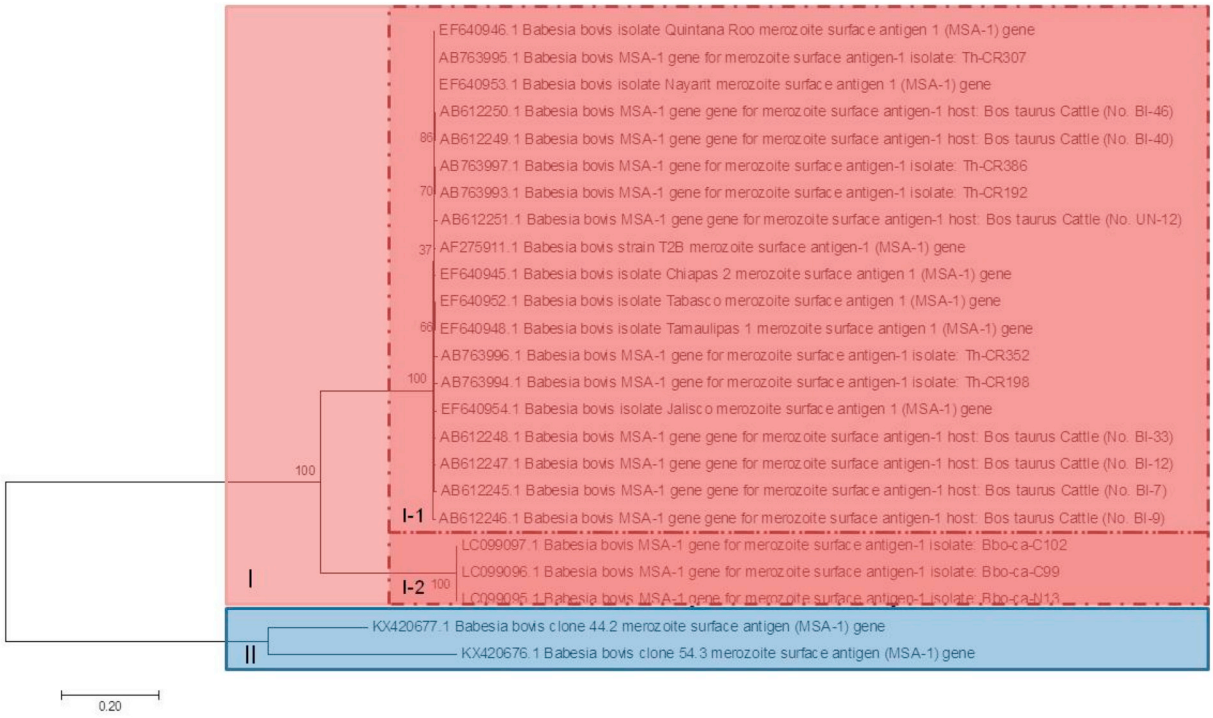

Fig. 3. Phylogenetic analysis of *msa-1* gene from *Babesia bovis* using sequences deposited on GenBank. I, II e III: genogroups I, II and II, respectively; I-1 and I-2: subgroups 1 and 2 of genogroup I, respectively; III-1 and III-2: subgroups 1 and 2 of genogroup III, respectively. The evolutionary history was inferred by using the Maximum Likelihood method based on the Tamura 3-parameter model. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+G, parameter = 1.2907)). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 32 nucleotide sequences. All positions containing gaps and missing data were eliminated. There were a total of 927 positions in the final dataset. Evolutionary analyses were conducted in MEGA7.

and may be a good virulence molecular marker. In our analyzes (Fig. 6), we identified variation between field isolates (clade II), attenuated strains and virulent strains (without specific separation within clade 1, although subgroups were observed: I-1 contains two vaccine strains and one virulent -FJ423468.1-; I-2 only virulent strains; I-3 attenuated and virulent from multiple locals). One possible explanation can be few sequences and small fragments with only 445 nucleotides were used (just a quarter of the gene). Thus, we suggest that analyzes with larger fragments and with a greater sequences number from diseased animals and asymptomatic carriers could elucidate this question and indicate if

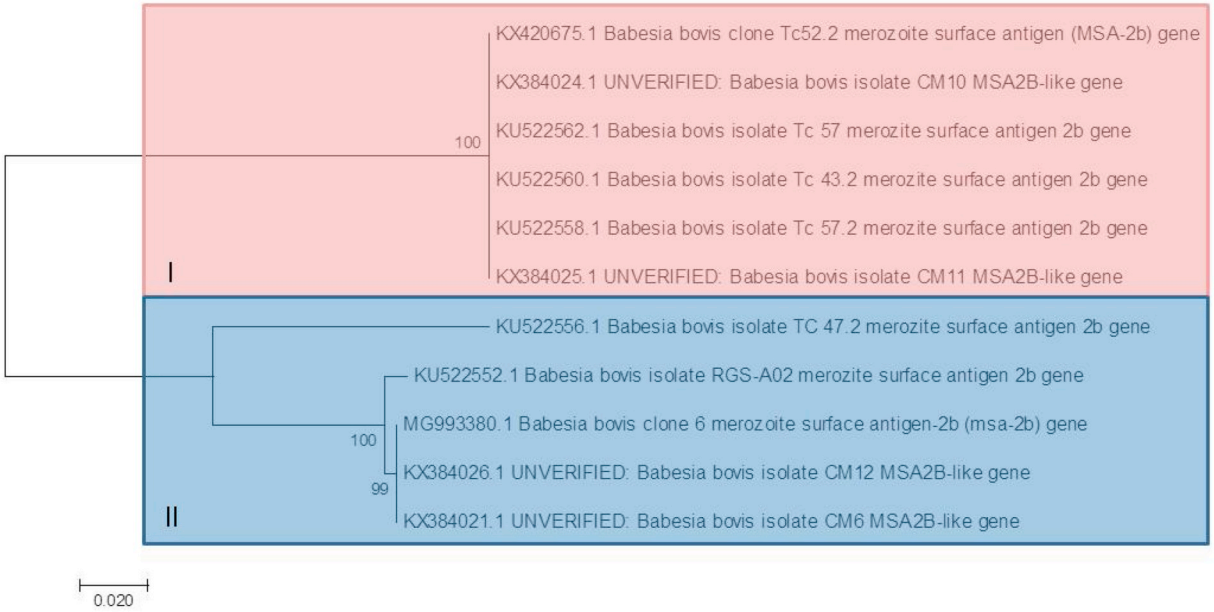

Fig. 4. Phylogenetic analysis of *msa-2c* gene from *Babesia bovis* using sequences deposited on GenBank. I and II: genogroups I and II, respectively. The evolutionary history was inferred by using the Maximum Likelihood method based on the Tamura 3-parameter model. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved six nucleotide sequences. All positions containing gaps and missing data were eliminated. There were a total of 737 positions in the final dataset. Evolutionary analyses were conducted in MEGA7.

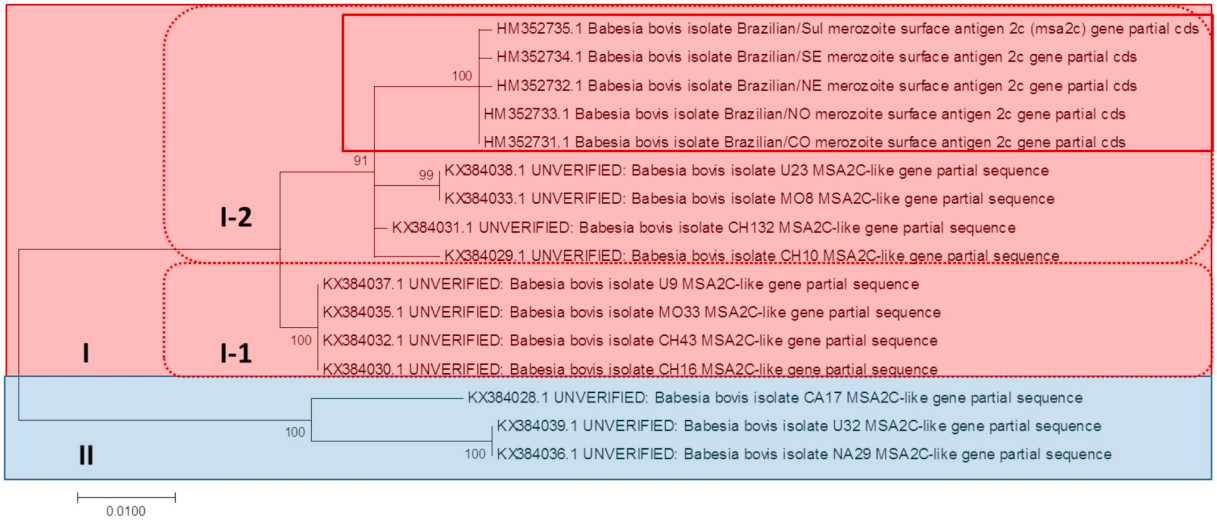

Fig. 5. Phylogenetic analysis of *msa-2c* gene from *Babesia bovis* using sequences deposited on GenBank. I and II: genogroups I and II, respectively; I-1 and I-2: subgroups 1 and 2 of genogroup I, respectively. The evolutionary history was inferred by using the Maximum Likelihood method based on the Tamura 3-parameter model. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 16 nucleotide sequences. All positions containing gaps and missing data were eliminated. There were a total of 750 positions in the final dataset. Evolutionary analyses were conducted in MEGA7.

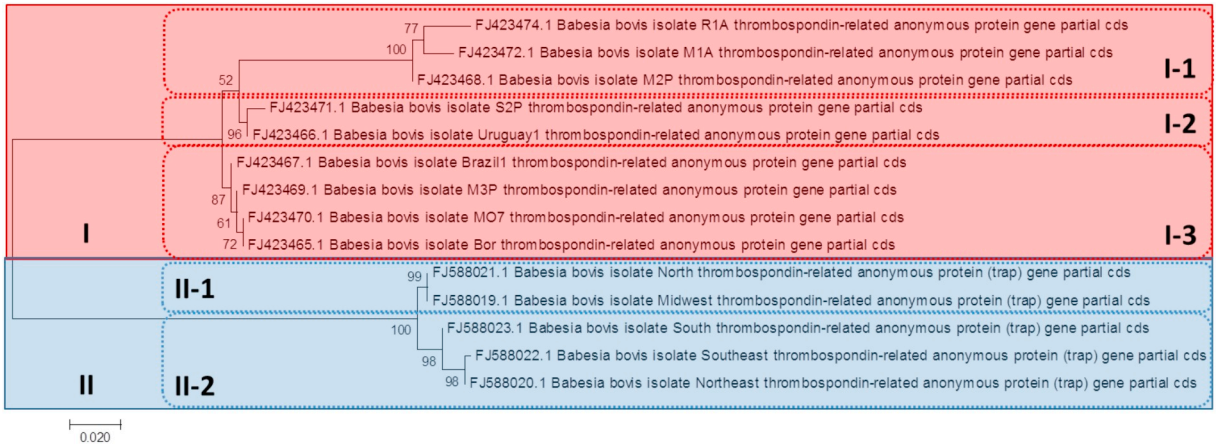

Fig. 6. Phylogenetic analysis of *trap* gene from *Babesia bovis* using sequences deposited on GenBank. I and II: genogroups I and II, respectively; I-1, I-2 and I-3: subgroups 1, 2 and 3 of genogroup I, respectively; II-1 and II-2: subgroups 1 and 2 of genogroup II, respectively. The evolutionary history was inferred by using the Maximum Likelihood method based on the Tamura 3-parameter model. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 14 nucleotide sequences. All positions containing gaps and missing data were eliminated. There were a total of 445 positions in the final dataset. Evolutionary analyses were conducted in MEGA7.

this gene can be used as a good marker of pathogenicity in genetic diversity studies of *B. bovis*.

3.5. Spherical body protein 2 and 4

The apicomplexans present a group of dense granules that are homologues to *SBPs* proteins identified within spherical body organelles in *B. bovis* (Liu et al., 1997). *SBP2* was identified in the erythrocyte membrane, which may play the same vital roles in adhesion. Furthermore, *SBP* proteins contribute to growth, development, survival parasites within erythrocytes and cause erythrocytic membrane permeability alterations (Lobo et al., 2012).

This gene family expresses 13-member located on chromosome 2 and chromosome 3. In chromosome 3 occur in three clusters of two, four and five truncated genes. Studies conducted by Pedroni et al. (2013) and Gallego-Lopez et al. (2018) showed the truncated gene *sbp-2 t* (7, 9 and 11 – BBOV_III006460, BBOV_III006500, and BBOV_III006540, respectively) were transcribed in a manner upregulated in attenuated strains when compared to the virulent parental strains (Pedroni et al., 2013; Gallego-Lopez et al., 2018). Little is known of why these genes were upregulated in the Babesia bovis attenuation process, indicating the need for diversity studies to understand the genetic mechanisms involved in the virulence attenuation process.

We selected sequences from four different works and constructed a phylogeny. We could observe a significant variability between the sequences in three clades and several subclades (Fig. 7). However, all sequences were obtained of healthy animals from epidemiological studies, endemic regions (Nagano et al., 2013); related this last one, the study was performed in farms without any ticks control system, indicating that the strains have mainly undergone selection pressure from the host immune system. The sequences from this study were clustered into two of three clades; the clade II was formed with just a sequence from Bahia, Northeast region from Brazil (AB772321.1). All subclades were composed by strains from different countries, as in the case of

Fig. 7. Phylogenetic analysis of *sbp-2* gene from *Babesia bovis* using sequences deposited on GenBank. I and II: genogroups I and II, respectively; I-1 and I-2: subgroups 1 and 2 of genogroup I, respectively; II-1 and II-2: subgroups 1 and 2 of genogroup II, respectively. The evolutionary history was inferred by using the Maximum Likelihood method based on the Tamura 3-parameter model. A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories (+G, parameter = 0.2559)). The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 21 nucleotide sequences. All positions containing gaps and missing data were eliminated. There were a total of 581 positions in the final dataset. Evolutionary analyses were conducted in MEGA7.

subclade III was observed strains from Vietnam and Brazil, indicating something similar related to the adaptive profile of the gene *sbp-2* in *B. bovis* evolution. Even using a small fragment of the gene (581 nucleotides), there was a great diversity among the strains coming from different geographic regions, and that were not grouped by region. All these data suggest that *sbp-2* should be better studied to indicate its relationship with attenuation profile of this parasite.

The *sbp-4* gene is located into chromosome 4 that presents just an exon with 1508 nucleotides. The CDS is located between positions 147..1265, coding a peptide with 372 amino acids (Brayton et al., 2007). *SBP4* was identified to be secreted throughout the cytoplasm. Although many studies have been shown that SBP proteins are good vaccine candidates, but the genetic and functional roles must be better evaluated.

Our analyses were conducted using 27 nucleotides sequences with a fragment 493 nucleotides of *sbp-4* gene from Indonesia and South Africa. We identified two different clades with two subclades each one (Fig. 8). Due to the small fragment and low geographic variation between the isolates (strains from only two different sites); we observed a lower degree of diversity within each clade and subclade when we compared with phylogenetic analysis using *sbp-2*. However, the two clades were very divergent. The subclades I-1, II-1 and II-2 were formed with just sequences from Indonesia and the subclades I-2 and II-3, with sequences from South Africa. So, the real diversity of *B. bovis* using *sbp-2* only will be understood using a more significant fragment and a greater geographic variation of isolates.

The persistence and genomic evolution of Babesia genus are complex and do not fully understand. The interaction between host and pathogen is complex and exerts a strong impact on the evolution of the parasite genomes (Allred, 2019).

4. Conclusions

We selected genes according to their functional and immunogenic abilities to conduct phylogenetic analysis and to understand the genetic diversity of *Babesia bovis* isolates from different geographic regions of the world. Among the 15 selected genes, only eight presented sufficient genetic variation for phylogeny studies. It was possible to observe that GenBank does not have enough data to conclude on molecular markers for attenuation or virulence. Some genetic groups presented no correlation with geographic region, which could indicate some common evolutionary characteristic related to the interaction parasite-host independent of location. So important as understanding host defense mechanisms is to understand the evolutionary mechanisms of the parasite and also provide effective prophylactic and therapeutic measures to livestock producers. The genes evaluated presented enough diversity to carry out diversity studies, however, only analyzes performed with the BV80 genes used sequences from animals or mechanisms of attenuation of the parasite in different clinical conditions of the host. Thus, we suggest that new studies using these genes in different parasite populations, clinical conditions, and level of pathogenicity can be developed.

Acknowledgements and funding

This work was supported by the São Paulo State Research Support Foundation (FAPESP) (Grant # 2016/07216-7, 2018/16582-2), the National Council for Scientific and Technological Development (CNPq) (Grant # 153231/2018-1) and the Brazilian Agricultural Research Corporation (Embrapa SEG 02.12.02.008.00.00).

Fig. 8. Phylogenetic analysis of *sbp-4* gene from *Babesia bovis* using sequences deposited on GenBank. I and II: genogroups I and II, respectively; I-1and I-2: subgroups 1 and 2 of genogroup I, respectively; II-1, II-2 and II-3: subgroups 1, 2 and 3 of genogroup II, respectively. The evolutionary history was inferred by using the Maximum Likelihood method based on the Kimura 2-parameter model. The tree is drawn to scale, with branch lengths measured in the number of substitutions per site. The analysis involved 27 nucleotide sequences. All positions containing gaps and missing data were eliminated. There were a total of 493 positions in the final dataset. Evolutionary analyses were conducted in MEGA7.

Declaration of Competing Interest

None.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.meegid.2019.104161.

References

Aktas, M., Bendele, K.G., Altay, K., Dumanli, N., Tsuji, M., Holman, P.J., 2007. Sequence polymorphism in the ribosomal DNA internal transcribed spacers differs among *Theileria* species. Vet. Parasitol. 147 (Issues 3–4), 221–230. https://doi.org/10.1016/j.vetpar.2007.04.007.

Allred, D.R., 2019. Variable and variant protein multigene families in *Babesia bovis* persistence. Pathogens 8, 76. https://doi.org/10.3390/pathogens8020076.

Baravalle, M.E., Thompson, C., Valentini, B., Ferreira, M., Torioni de Echaide, S., Christensen, M.F., Echaide, I., 2012. *Babesia bovis* biological clones and the inter-strain allelic diversity of the Bv80 gene support subpopulation selection as a mechanism involved in the attenuation of two virulent isolates. Vet. Parasitol. 190 (Issues 3–4), 391–400. https://doi.org/10.1016/j.vetpar.2012.06.037.

Bock, R., Jackson, L., De Vos, A., Jorgensen, W., 2004. Babesiosis of cattle. Parasitology 129, S247–S269.

Brayton, K.A., Lau, A.O.T., Herndon, D.R., Hannick, L., Kappmeyer, L.S., Berens, S.J., Bidwell, S.L., Brown, W.C., Crabtree, J., Fadrosh, D., Feldblum, T., Forberger, H.A., Haas, B.J., Howell, J.M., Khouri, H., Koo, H., Mann, D.J., Norimine, J., Paulsen, I.T., Radune, D., Ren, Q., Smith Jr., R.K., Suarez, C.E., White, O., Wortman, J.R., Knowles Jr., D.P., McElwain, T.F., Nene, V.M., 2007. Genome sequence of *Babesia bovis* and comparative analysis of apicomplexan hemoprotozoa. PLoS Pathog. https://doi.org/10.1371/journal.ppat.0030148. Oct 19.

Brown, W.C., Mcelwain, T.F., Ruef, B.J., Suarez, C.E., Shkap, V., Chitko-Mckown, C.G., Tuo, W., Rice-Ficht, A.C., Palmer, G.H., 1996. *Babesia bovis* rhoptry-associated protein 1 is immunodominant for T helper cells of immune cattle and contains t-cell epitopes conserved among geographically distant *B. bovis* strains. Infect. Immun. 3341–3350 Aug.

Carson, C.A., Timms, P., Cowman, A.F., Stewart, N.P., 1990 May. Babesia bovis: evidence for selection of subpopulations during attenuation. Exp. Parasitol. 70 (4), 404–410. https://doi.org/10.1016/0014-4894(90)90124-u.

Collins, N.E., Allsopp, B.A., 1999. Theileria parva ribosomal internal transcribed spacer sequences exhibit extensive polymorphism and mosaic evolution: application to the characterization of parasites from cattle and buffalo. Parasitology 118 (Pt 6), 541–551 Jun.

Combrink, M.P., Troskie, P.C., Pienaar, R., Latif, A.A., Mans, B.J., 2014 Jan 31. Genotypic diversity in Babesia bovis field isolates and vaccine strains from South Africa. Vet. Parasitol. 199 (3-4), 144–152. https://doi.org/10.1016/j.vetpar.2013.10.008. Epub 2013 Oct 23.

Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2012. jModelTest 2: more models, new heuristics and parallel computing. Nat. Methods 9 (8), 772. https://doi.org/10.1038/nmeth.2109. Jul 30.

Florin-Christensen, M., Suarez, C.E., 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. 70, 3566–3575. https://doi.org/10.1128/iai.70.7.3566-3575.2002.

Gallego-Lopez, G.M., Lau, A.O.T., Brown, W.C., Johnson, W.C., Ueti, M.W., Suarez, C.E., 2018. Spherical Body Protein 2 truncated copy 11 as a specific Babesia bovis attenuation marker. Parasit. Vectors 11 (1), 169. https://doi.org/10.1186/s13071-018-2782-z.

Gohil, S., Kats, L.M., Sturm, A., Cooke, B.M., 2010. Recent insights into alteration of red blood cells by *Babesia bovis*: moovin' forward. Trends Parasitol. 26 (12), 591–599. https://doi.org/10.1016/j.pt.2010.06.012.

Gohil, S., Herrmann, S., Günther, S., Cooke, B.M., 2013. Bovine babesiosis in the 21st century: advances in biology and functional genomics. Int. J. Parasitol. 43, 125–132. https://doi.org/10.1016/j.ijpara.2012.09.008.

Hillis, D.M., Dixon, M.T., 1991 Dec. Ribosomal DNA: molecular evolution and phylogenetic inference. Q. Rev. Biol. 66 (4), 411–453.

Hilpertshauser, H., Deplazes, P., Meli, M.L., Hofmann-Lehmann, R., Lutz, H., Mathis, A., 2007. Genotyping of Babesia bigemina from cattle from a non-endemic area (Switzerland). Vet. Parasitol. 145 (1-2), 59–64.

Hines, S.A., Palmer, G.H., Jasmer, D.P., McGuire, T.C., McElwain, T.F., 1992. Neutralization-sensitive merozoite surface antigens of Babesia bovis encoded by members of a polymorphic gene family. Mol. Biochem. Parasitol. 55, 85–94. https://doi.org/10.1016/0166-6851(92)90129-8.

Howes, R.E., Battle, K.E., Mendis, K.N., Smith, D.L., Cibulskis, R.E., Baird, J.K., Hay, S.I., 2016. Global epidemiology of *Plasmodium vivax*. Am. J. Trop. Med. Hyg. 95 (6 Suppl), 15–34. https://doi.org/10.4269/ajtmh.16-0141.

Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol. Biol. Evol. 33, 1870–1874. https://doi.org/10.1093/molbev/msw054.

Lau, A.O.T., Kalyanaraman, A., Echaide, I., Palmer, G.H., Bock, R., Pedroni, M.J., Rameshkumar, M., Ferreira, M.B., Fletcher, T.I., McElwain, T.F., 2011. Attenuation of virulence in an apicomplexan hemoparasite results in reduced genome diversity at the population level. BMC Genomics 12, 410. http://www.biomedcentral.com/1471-2164/12/410.

LeRoith, T., Berens, S.J., Brayton, K.A., Hines, S.A., Brown, W.C., Norimine, J., McElwain, T.F., 2006. The *Babesia bovis* Merozoite surface antigen 1 hypervariable region induces surface-reactive antibodies that block Merozoite invasion. Infect. Immun. 74 (6), 3663–3667.

Liu, Z., Zhao, J., Ma, L., Yao, B., 1997. Studies on buffalo babesiosis in Hubei Province, China. Trop. Anim. Health Prod. 29, 33S–36S. https://doi.org/10.1007/BF02632914.

Lobo, C.A., Rodriguez, M., Cursino-Santos, J.R., 2012. Babesia and red cell invasion. Curr. Opin. Hematol. 19, 170–175. https://doi.org/10.1097/MOH.0b013e328352245a.

Matos, C.A., Gonçalves, L.R., Alvarez, D.O., Freschi, C.R., Silva, J.B.D., Val-Moraes, S.P., Mendes, N.S., André, M.R., Machado, R.Z., 2017 Oct-Dec. Longitudinal evaluation of humoral immune response and merozoite surface antigen diversity in calves naturally infected with Babesia bovis, in São Paulo, Brazil. Rev. Bras. Parasitol. Vet. 26 (4), 479–490. https://doi.org/10.1590/S1984-29612017069.

Mazuz, M.L., Molad, T., Fish, L., Leibovitz, B., Wolkomirsky, R., Fleiderovitz, L., Shkap, V., 2012 Mar. Genetic diversity of Babesia bovis in virulent and attenuated strains. Parasitology. 139 (3), 317–323. https://doi.org/10.1017/S0031182011001995. Epub 2011 Nov 11.

Molad, T., Fleiderovitz, L., Leibovich, B., Wolkomirsky, R., Erster, O., Roth, A., Mazuz, M.L., Markovics, A., Shkap, V., 2014 Sep 15. Genetic polymorphism of Babesia bovis merozoite surface antigens-2 (MSA-2) isolates from bovine blood and Rhipicephalus annulatus ticks in Israel. Vet. Parasitol. 205 (1-2), 20–27. https://doi.org/10.1016/j.vetpar.2014.07.016. Epub 2014 Jul 24.

Morahan, B.J., Wang, L., Coppel, R.L., 2009. No TRAP, no invasion. Trends Parasitol. 25, 77–84. https://doi.org/10.1016/j.pt.2008.11.004.

Mosqueda, J., McElwain, T.F., Stiller, D., Palmer, G.H., 2002. *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, 1599–1609. https://doi.org/10.1128/IAI.70.3.1599-1603.2002.

Mtshali, P.S., Mtshali, M.S., 2017 Dec 8. In silico and phylogenetic analyses of partial BbRAP-1, BbCP2, BbSBP-4 and BbβTUB gene sequences of Babesia bovis isolates from cattle in South Africa. BMC Vet. Res. 13 (1), 383. https://doi.org/10.1186/s12917-017-1261-7.

Nagano, D., Sivakumar, T., De 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 (11), 1463–1470. https://doi.org/10.1292/jvms.13-0251.

Niu, Q., Liu, Z., Yu, P., et al., 2015. Genetic characterization and molecular survey of Babesia bovis, Babesia bigemina and Babesia ovata in cattle, dairy cattle and yaks in China. Parasit. Vectors 8, 518. https://doi.org/10.1186/s13071-015-1110-0. Published 2015 Oct 9.

Pedroni, M.J., Sondgeroth, K.S., Gallego-Lopez, G.M., Echaide, I., Lau, A.O., 2013. Comparative transcriptome analysis of geographically distinct virulent and attenuated *Babesia bovis* strains reveals similar gene expression changes through attenuation. BMC Genomics 14 (1), 763. https://doi.org/10.1186/1471-2164-14-763.

Ramos, C.A., Araújo, F.R., Alves, L.C., de Souza, I.I., Guedes Jr., D.S., Soares, C.O., 2012. Genetic conservation of potentially immunogenic proteins among Brazilian isolates of *Babesia bovis*. Vet. Parasitol. 187 (3–4), 548–552. https://doi.org/10.1016/j.vetpar.2012.01.020. Jul 6.

Rittipornlertrak, A., Nambooppha, B., Simking, P., Punyapornwithaya, V., Tiwananthagorn, S., Jittapalapong, S., Chung, Y.T., Sthitmatee, N., 2017. Low levels of genetic diversity associated with evidence of negative selection on the *Babesia bovis* apical membrane antigen 1 from parasite populations in Thailand. Infect. Genet. Evol. 54, 447–454. https://doi.org/10.1016/j.meegid.2017.08.009.

Sivakumar, T., Lan, D.T., Long, P.T., et al., 2013. PCR detection and genetic diversity of bovine hemoprotozoan parasites in Vietnam. J. Vet. Med. Sci. 75 (11), 1455–1462. https://doi.org/10.1292/jvms.13-0221.

Tattiyapong, M., Sivakumar, T., Takemae, H., Simking, P., Jittapalapong, S., Igarashi, I., Yokoyama, N., 2016 Jul. Genetic diversity and antigenicity variation of Babesia bovis merozoite surface antigen-1 (MSA-1) in Thailand. Infect. Genet. Evol. 41, 255–261. https://doi.org/10.1016/j.meegid.2016.04.021. Epub 2016 Apr 19.

Taylor, H.M., Kyes, S.A., Newbold, C.I., 2000. Var gene diversity in Plasmodium falciparum is generated by frequent recombination events. Mol. Biochem. Parasitol. 110 (2), 391–397002E. https://doi.org/10.1016/S0166-6851(00)00286-3.

Wilkowsky, S.E., Moretta, R., Mosqueda, J., Gil, G., Echaide, I., Lía, V., Falcón, A., Florin Christensen, M., Farber, M., 2009. A new set of molecular markers for the genotyping of *Babesia bovis* isolates. Vet. Parasitol. 161 (1–2), 9–18. https://doi.org/10.1016/j.vetpar.2008.12.024.

Yokoyama, N., Sivakumar, T., Tuvshintulga, B., Hayashida, K., Igarashi, I., Inoue, N., Long, P.T., Lan, D.T.B., 2015 Mar. Genetic variations in merozoite surface antigen genes of Babesia bovis detected in Vietnamese cattle and water buffaloes. Infect. Genet. Evol. 30, 288–295. https://doi.org/10.1016/j.meegid.2014.12.035. Epub 2015 Jan 6.

Zahler, M., Schein, E., Rinder, H., Gothe, R., Zhou, J., Fukumoto, S., Jia, H., Yokoyama, N., 1998. Characteristic genotypes discriminate between Babesia canis isolates of differing vector specificity and pathogenicity to dogs. Parasitol. Res. 84 (7), 544–548.

Zhou, J., Fukumoto, S., Jia, H., Yokoyama, N., Zhang, G., Fujisaki, K., Lin, J., Xuan, X., 2006. Characterization of the *Babesia gibsoni* P18 as a homologue of thrombospondin related adhesive protein. Mol. Biochem. Parasitol. 148 (2), 190–198. https://doi.org/10.1016/j.molbiopara.2006.03.015.