Ticks and Tick-borne Diseases xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Ticks and Tick-borne Diseases

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

Short communication

Molecular characterization of *Theileria equi* in horses from the state of Rio de Janeiro, Brazil

Maristela Peckle[a,*], Marcus Sandes Pires[a], Claudia Bezerra da Silva[b], Renata Lins da Costa[a], Gabriela Lopes Vivas Vitari[a], Marcus Vinicius Xavier Senra[c], Roberto Júnio Pedroso Dias[c], Huarrisson Azevedo Santos[b], Carlos Luiz Massard[a]

[a] Department of Animal Parasitology, Veterinary Institute, Federal Rural University of Rio de Janeiro – UFRRJ, Seropedica, Rio de Janeiro, Brazil
[b] Department of Epidemiology and Public Health, Veterinary Institute, Federal Rural University of Rio de Janeiro – UFRRJ, Seropedica, Rio de Janeiro, Brazil
[c] Department of Zoology, Biologic Sciences Institute, Federal University of Juiz de Fora – UFJF, Juiz de Fora, Minas Gerais, Brazil

ARTICLE INFO

Keywords:
Equine piroplasmosis
Theileriosis
Phylogeny
18S rRNA gene
Genetic diversity

ABSTRACT

Theileria equi is one of the etiologic agents of the equine piroplasmosis. This infectious disease is transmitted by ticks and is a worldwide problem in the international horse movement. The *18S rRNA* gene of *T. equi* is often used for genotyping and phylogenetic purpose. This study aimed to analyze the degree of the heterogeneity of the *18S rRNA* gene of *T. equi* in horses from the state of Rio de Janeiro, Brazil. The complete *T. equi 18S rRNA* sequences were obtained from twenty naturally infected horses. The PCR amplicons were cloned and sequenced. The phylogenetic analyses were performed using a set of *T. equi 18S rRNA* sequences and other related organisms available in ARB-Silva database. There were twelve distinct *T. equi 18S rRNA* gene sequences circulating in horses in the state of Rio de Janeiro, Brazil. Monophyletic clades with 2% evolutionary divergence between clades and high bootstrap value were the support to divide *T. equi* sequences in three distinct clades. The sequences from this study grouped into clades I (70%, n = 14/20) and II (30%, n = 6/20). All of the *T. equi* sequences grouped within a node other than the theileriids. This study reported a clear division of two distinct genotypes of *T. equi 18S rRNA* sequences in state of Rio de Janeiro, Brazil, and it demonstrates that distinct isolates of *T. equi* can coexist in the same geographic region.

Constant economic losses occur in the equine industry due to infectious diseases that include equine piroplasmosis (EP). This disease causes fever, anemia, jaundice, hepatomegaly, splenomegaly, and hemoglobinuria in horses, donkeys, mules, zebras, and, as recently reported, in camels (Sloboda et al., 2011). Equine piroplasmosis is considered to be the major impediment to the international movement of horses because EP-positive animals are banned from entering countries that are considered free of the disease (Friedhoff et al., 1990; Knowles, 1996). *Theileria equi* (Laveran, 1901) (Mehlhorn and Schein, 1998) and *Babesia caballi* (Nuttall and Strickland, 1912), the etiological agents of EP, are obligate intracellular protozoans (Apicomplexa: Piroplasmorida) that infect the erythrocytes of equines (Friedhoff and Soule, 1996; Schein, 1988).

The *18S rRNA* gene is a target that is widely since nucleotide substitution rates are low, and there is no evidence of lateral gene transfer between lineages (Allsopp and Allsopp, 2006). Despite these facts, the variable regions of this gene are often used for phylogenetic studies, particularly the *18S rRNA* gene of *T. equi*, which has exhibited a high degree of heterogeneity in sequences from different regions of the world (Bhoora et al., 2009; Criado-Fornelio et al., 2003; Nagore et al., 2004).

In Brazil, although some molecular studies have been conducted regarding the identification of *T. equi* in horses (Baldani et al., 2010; Ferreira et al., 2016; Heuchert et al., 1999; Peckle et al., 2013), the genetic diversity of the *18S rRNA* gene remains unknown. This fact raises a concern regarding molecular diagnosis based on *18S rRNA* sequences of *T. equi*. The existence of variation in the sequences of the *18S rRNA* gene of *T. equi* could complicate the standardization of molecular diagnostic techniques that use this gene as a target and raise the chances of false-negative diagnoses. Studies reported that PCR probes designed based on the *18S rRNA* gene failed to hybridize *T. equi*-positive samples due to some sequence differences in this region (Bhoora et al., 2009; Liu et al., 2016).

Thus, the objectives of this study were to analyze the degree of the heterogeneity of the *18S rRNA* gene of *T. equi* in samples from two geographical regions from the state of Rio de Janeiro, Brazil, and to compare them with other isolates from around the world, and highlight

* Corresponding author.
E-mail address: maristelapeckle@yahoo.com.br (M. Peckle).

https://doi.org/10.1016/j.ttbdis.2017.11.011
Received 25 April 2017; Received in revised form 9 September 2017; Accepted 20 November 2017

the genotypes by phylogenetic analysis.The twenty DNA samples from horses infected with *T. equi* used in this study were obtained from a former study conducted by Peckle et al. (2013). Ten samples were selected from the municipality of Petropolis, and the other ten were from the municipality of Seropedica; both sets of sample were from the state of Rio de Janeiro, Brazil. Seropedica is characterized as tropical (Aw), and Petropolis is considered (Cwb) (Peel et al., 2007). Both places were selected to obtain isolates of *T. equi* with greater genetic variability.

The complete *T. equi 18S rRNA* gene was amplified following the procedures of Bhoora et al. (2009). The positive control used in the reaction was obtained from the blood DNA of a chronically *T. equi* infected horse. The number of copies of the *18S rRNA* gene in the positive control was approximately 1000 copies of the DNA target (Peckle et al., 2013). Ultrapure water was used as a negative control.

The PCR amplicon were cloned into the plasmid pGEM-T® Easy Vector System (Promega, Madison, WI, USA). The plasmid DNA extraction from the transformed bacteria was performed using the QIAprep Spin Plasmid Miniprep kit (Qiagen) according to the manufacturer's recommendations. The plasmid DNA of each sample was sequenced with six different primers: NBabesia1F, BT18S3R, BT18S3F, 18SRev-TB, BT18S2F and BT18S2R (Bhoora et al., 2009) via the Sanger method (Sanger et al., 1977) Sequencing with different primers was performed to ensure a greater coverage and reliability of the data for contig assembly.

The electropherograms were analyzed, edited and assembled into contigs using DNA Baser Sequence Assembly 3.0 (Heracle Biosoft SRL, Romania). BLASTN was used to identify similar sequences in public databases (Altschul et al., 1990).

The phylogenetic analyses were performed using a set of Piroplasmorida *18S rRNA* gene sequences downloaded from the ARB-Silva database (Quast et al., 2013), which included the 48 *Theileria equi* sequences (ingroup), 8 other Theileriidae sequences, 8 Babesiidae sequences and 2 other apicoplexans (outgroup), plus the sequences generated in this work. The alignment of the sequences and all subsequent dataset manipulations were performed using the ARB software package (Ludwig et al., 2004), and these manipulations included manual inspections for the removal of unaligned sites, columns with gaps, primer-related sequences, and ambiguous sites. The dataset was then trimmed to produce an equi-length dataset (100 taxa and 1202 characters). The best evolutionary model was then selected using the jModelTest_2.1.4 software (Darriba et al., 2015). Maximum likelihood (ML) analysis was then performed using PhyML v3.0 (Guindon et al., 2010) with the GTR + G + I nucleotide substitution model. The reliability of the internal branches was assessed using the nonparametric bootstrap method with 1000 pseudoreplicates. The genetic distance table was calculated using the ARB software package (Ludwig et al., 2004). Similarly, the intergroup and intragroup distances and overall distances were calculated.

The twenty new *18S rRNA* gene sequences of *T. equi* obtained in this study have been deposited in GenBank under the following accession numbers: KX722511 to KX722525, KJ573370, KJ573371, KJ573372.2, KJ573373.2, and KJ573374.2 (Table 1).

Among all analyzed samples, thirteen were deposited as complete sequences of the *T. equi 18S rRNA* gene, and seven other were partial gene sequences (Table 1). Among our samples, there were twelve distinct *T. equi 18S rRNA* gene sequences circulating in horses in the state of Rio de Janeiro, Brazil. The lowest identity value shared by these sequences was 95.8% and that was between RJ12 and RJ17. In contrast, RJ2 and RJ6 exhibited 100% identity to each other, and a group composed of RJ3 and RJ7, RJ8, RJ9, RJ11, RJ15, RJ16, and RJ20 exhibited the same.

In the phylogenetic analysis presented in Fig. 1, 68 sequences of the *T. equi 18S rRNA* gene were used, and within them, twenty sequences were from Brazil in the state of Rio de Janeiro (this study) and 48 were from other countries. The lowest distances (d = 0.0) were observed between the sequence RJ2 and six sequences from the USA (JX177670 to JX177673, ACOU01000003 and CP001669) and one from South Africa (EU888906). The distances between those and RJ12/RJ13 were also very low (d = 0.1). In contrast, the most distant sequence from Brazil was the KM046922 from Hungary with a distance that ranged from 3.6 to 4.2 in relation to all of the sequences from Rio de Janeiro (Online Resource). After alignment of the *T. equi* ingroup sequences (n = 68), it was possible to observe extensive variation in the V4 hypervariable region.

Table 1
Sample identification from this study, region of origin, contig size, register in GenBank and status of the sequence.

Sample Id	Region	Contig size (bp)	GenBank ID	Status
RJ1	Petropolis	1586	KX722511.1	Complete
RJ2	Petropolis	1498	KJ573370.1	Partial
RJ3	Petropolis	1366	KX722524.1	Partial
RJ4	Petropolis	1471	KJ573371.1	Partial
RJ5	Petropolis	1586	KX722512.1	Complete
RJ6	Petropolis	1586	KX722513.1	Complete
RJ7	Petropolis	1586	KX722514.1	Complete
RJ8	Petropolis	1585	KX722515.1	Complete
RJ9	Petropolis	1586	KX722516.1	Complete
RJ10	Petropolis	1586	KX722517.1	Complete
RJ11	Seropedica	1586	KX722518.1	Complete
RJ12	Seropedica	1586	KX722519.1	Complete
RJ13	Seropedica	1586	KX722520.1	Complete
RJ14	Seropedica	1586	KX722521.1	Complete
RJ15	Seropedica	1475	KJ573372.2	Partial
RJ16	Seropedica	1533	KX722522.1	Partial
RJ17	Seropedica	1586	KX722523.1	Complete
RJ18	Seropedica	1531	KX722525.1	Partial
RJ19	Seropedica	1460	KJ573373.2	Partial
RJ20	Seropedica	1385	KJ573374.2	Partial

In this study, monophyletic clades with 2% evolutionary divergence between clades and high bootstrap value support were considered distinct genotypes. To avoid misunderstanding in relation to the nomenclature of the genotypes, we called them clades I, II and III. The average intragroup distances were less than 1% for clades I (d = 0.8%) and II (d = 0.3%). In clade III, the average distance between the sequences was approximately 1.1%.

The phylogenetic reconstruction resulted in a phylogenetic tree with high bootstrap values that supported the clusters (Fig. 1). The *18S rRNA* gene sequences of *T. equi* that were analyzed in this study grouped together into three well-supported clades, i.e., clades I, II and III. The sequences from this study grouped into clades I and II, and the majority of the sequences observed were in clade I (70%, n = 14/20). The frequencies of the sequences observed in group I were also higher for both municipalities. In clade I, 14 sequences from Brazil grouped together with sequences from South Africa, Sudan, Kenya and the United States. In clade II, we found the other six Brazilian sequences along with sequences from South Africa, Spain, India, South Korea and the United States. In clade III, sequences from Kenya, Sudan, Hungary, Spain, Switzerland, South Korea, China and South Africa grouped together (Fig. 1).

In general, all of the *T. equi* sequences grouped within a node other than the theileriids. The *Cytauxzoon* organisms grouped between them and the other *Theileria* organisms. *Babesia bicornis* appeared to be a sister clade of *T. equi* as supported by the 78% bootstrap.

Until now, studies of the phylogenetic aspects of *T. equi* have not been performed in Brazil. Based on this study, it was possible to observe genetic differences in the *18S rRNA* gene between samples obtained from naturally infected equines of two distinct geographic regions in the state of Rio de Janeiro and with samples from other countries.

In this study, using complete or expressive partial fragments of the *18S rRNA* gene of *T. equi*, it was possible to accurately assume that all of the sequences grouped in three main clades termed I, II and III. These findings are similar to those of Bhoora et al. (2009) who obtained three groups from a cladistic analysis of the *18S rRNA* gene of *T. equi* in a