Veterinary Parasitology 266 (2019) 103–110

Contents lists available at ScienceDirect

Veterinary Parasitology

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

Research paper

Paralogs vs. genotypes? Variability of *Babesia canis* assessed by 18S rDNA and two mitochondrial markers

Kristýna Hrazdilová[a,b,*], Izabella Myśliwy[c], Joanna Hildebrand[c], Katarzyna Buńkowska-Gawlik[c], Bartłomiej Janaczyk[d], Agnieszka Perec-Matysiak[c], David Modrý[a,e,f]

[a] CEITEC-VFU, University of Veterinary and Pharmaceutical Sciences Brno, Palackého tř. 1946/1, 612 42, Brno, Czech Republic
[b] Department of Virology, Veterinary Research Institute, Hudcova 296/70, 621 00, Brno, Czech Republic
[c] Department of Parasitology, Institute of Genetics and Microbiology, University of Wrocław, Przybyszewskiego 63, 51-148, Wrocław, Poland
[d] Vetlab, Veterinary Diagnostic Laboratory, Parkowa 13, 51-616, Wrocław, Poland
[e] Department of Pathological Morphology and Parasitology, University of Veterinary and Pharmaceutical Sciences Brno, Palackého tř. 1946/1, 612 42, Brno, Czech Republic
[f] Biology Centre, Institute of Parasitology, Czech Academy of Sciences, Branišovská 31, 370 05, České Budějovice, Czech Republic

ARTICLE INFO

Keywords:
Babesia canis
18S rDNA paralog
Genotype
COI
Cytb

ABSTRACT

Canine babesiosis caused by *Babesia canis* sensu stricto became an emerging disease of dogs across Europe calling for attention also in countries where it was an only rare imported disease. An easy accessibility of molecular methods and the growing amount of sequencing data led to the description of intraspecific variability in 18S rDNA sequences designated as "genotypes". Using material from a homogenous cohort of dogs with microscopically confirmed canine babesiosis caused by *B. canis*, we evaluated *Babesia* intraspecific variability and amplification sensitivity of three different genes (18S rDNA, COI, Cytb) to assess their potential as diagnostic or phylogenetic markers. In raw sequencing data obtained, we observed at least 3 ambiguous positions in up to 86% of chromatograms within the ~560 bp fragment of 18S rDNA suggesting the existence of several, not identical copies of this gene. Our COI haplotype analysis resulted in a star-like pattern indicating a recent origin of most haplotypes, but not supporting the existence of two dominant haplotypes. Similarly, the Cytb sequences obtained from samples with all variants of 18S rDNA were identical. We corroborate previous observations from three other European countries and bring the evidence of the existence of 18S rDNA paralogs in *B. canis* genome replacing currently used "genotype" theory.

1. Introduction

Canine babesioses are caused by several species of tick-borne apicomplexans of the genus *Babesia*. Traditional diagnostic methods based on microscopy of blood smears allow distinguishing two groups of causative agents, large (3–5 µm; e.g. *Babesia canis* sensu lato) and small (1.5–2.5 µm; e.g. *Babesia gibsoni*) piroplasmids. Nowadays, molecular techniques enable differentiation of morphologically indistinguishable species (*B. canis*, *B. rossi*, *B. vogeli*), previously classified as subspecies of *B. canis* based on clinical presentation, geographical distribution and vector specificity (Solano-Gallego and Baneth, 2011). However, studies addressing intraspecific genetic diversity of individual taxa are few.

Babesiosis caused by *Babesia canis* sensu stricto (previously referred to as *B. canis canis*) became an emerging disease of dogs across Europe calling for attention also in countries where it was only rarely imported (Mitkova et al., 2017; Paulauskas et al., 2014). Reported prevalence ranges from 0.1 to 88% depending on geography, the abundance of the tick vector, and study design (Solano-Gallego et al., 2016). Recent climatic changes together with anthropogenic impact are suggested as a predominant driving force of geographical spread of *Dermacentor reticulatus*, the vector of *B. canis* (Dautel et al., 2006; Matijatko et al., 2012; Randolph and EDEN-TBD sub-project team, 2010). Clinical manifestation varies from mild/subclinical to severe or even fatal cases with remittent fever, progressive anaemia, hemoglobinuria, marked splenomegaly and hepatomegaly (Kubelová et al., 2013; Łyp et al., 2015; Máthé et al., 2006). Correct determination of *Babesia* species is an essential part of understanding the epidemiology of the infection. Easy accessibility of molecular methods and growing amount of sequencing data inspired a range of the phylogenetic studies on *B. canis* (Solano-Gallego et al., 2016).

* Corresponding author at: CEITEC-VFU, University of Veterinary and Pharmaceutical Sciences Brno, Palackého tř. 1946/1, 612 42, Brno, Czech Republic.
E-mail addresses: kristyna@hrazdilova.cz, hrazdilovak@vfu.cz (K. Hrazdilová).

https://doi.org/10.1016/j.vetpar.2018.12.017
Received 29 October 2018; Received in revised form 27 December 2018; Accepted 28 December 2018

Most of the available molecular data are based on the small ribosomal subunit RNA gene 18S (18S rDNA). In 2008, Adaszek et al. defined two restriction groups based on dinucleotide polymorphism within the 18S rDNA gene (Adaszek and Winiarczyk, 2008). Since then, these "genotypes" were associated with a clinical course of the disease (Adaszek et al., 2009), uneven seasonal distribution (Hornok et al., 2015), and new genotype discrimination methods were developed (Adaszek and Winiarczyk, 2010). Recently, third "genotype" was described at the same nucleotide positions (Łyp et al., 2015). In this context, the broad existence of 18S rDNA paralogs across phylum Apicomplexa (El-Sherry et al., 2013; Kibe et al., 1994; Le Blancq et al., 1997; Rooney, 2004) remains overlooked as well as the evidence for paralogs existence in *Babesia* spp. (Brayton et al., 2007; Dalrymple, 1990; Reddy et al., 1991).

Here we aim to review the available GenBank deposited sequences of 18S rDNA of *B. canis* with respect to the sequence variability. We investigate the existence of 18S rDNA paralogs and evaluate the reality of 18S rDNA "genotypes" by addition of two mitochondrial markers into our analyses. Using material from a homogenous cohort of dogs with microscopically confirmed canine babesiosis caused by *B. canis*, we evaluated *Babesia* intraspecific variability and tested the sensitivity of amplification of three different genes to assess their potential as diagnostic or phylogenetic markers.

2. Material and methods

2.1. GenBank review

All sequences of 18S rDNA assigned as *B. canis* or *B. canis canis* were retrieved from GenBank database 13th June 2018. If available, host, country of origin and publication status were recorded for each sequence.

2.2. Sample set

Blood samples of dogs originated from the area of Masovian and Lower Silesian voivodeships in Poland were subjected to routine laboratory examinations in the veterinary diagnostic laboratory Vetlab (Poland). Ninety nine blood samples obtained from microscopically *B. canis* positive dogs were frozen and delivered to the Department of Parasitology, University of Wrocław for DNA isolation procedures.

2.3. DNA isolation and PCR protocols

DNA was isolated using Quick Blood DNA Purification Kit (EURx, Poland) from 200 µl of the whole blood following manufacturer's instruction and stored in −20 °C. All samples were screened using universal, sensitive assay detecting a wide range of piroplasmids targeting nuclear 18S rDNA. Subsequently, all positive samples were screened by PCR assay targeting mitochondrial Cytochrome c oxidase subunit I marker (COI) and, based on the results discussed further, a representative subset of 25 samples was used to amplify a partial sequence of mitochondrial Cytochrome *b* (Cytb). Primers for Cytb amplification were designed based on nucleotide alignment of available *B. canis* sensu lato sequences in conserved regions using Geneious software (Kearse et al., 2012) with manual editing. Primer sequences, used annealing temperatures and length of amplicons are listed in Table 1. All first round reactions were performed in 15 µl reaction volume, consisting of 1 µl template DNA, 400 nM of each primer, and 7.5 µl of 2× PCRBIO Taq Mix Red (PCR Biosystems, UK). Second rounds of nested PCRs were done in 25 µl using 1 µl of PCR product from the first round as template, 400 nM of each primer, and 12.5 µl of 2× PCRBIO Taq Mix Red (PCR Biosystems, UK). PCR products were visualized on 1% agarose gel stained by MidoriGreen Advance (Nippon Genetics Europe, Germany).

2.4. Sequencing and sequence analysis

All PCR products of expected size were purified from gel using the Gel/PCR DNA Fragment Extraction Kit (Geneaid Biotech, Taiwan) and directly sequenced by Macrogen capillary sequencing services (Macrogen Europe, the Netherlands) using the amplification primers in both directions. Samples with low quality sequencing results or resulting in clearly mixed chromatogram signals were cloned using the pGEM®-T Easy Vector System (Promega, USA). Additionally, 18S rDNA PCR product of a single selected sample with high quality chromatogram containing GA/AG double peaks was cloned and processed the same way. Cloned plasmid DNA was extracted and purified from the bacterial culture by GenElute™ Plasmid Miniprep Kit (Sigma-Aldrich, USA) and sequenced using universal T7/SP6 primers. All obtained sequencing results were carefully edited (ambiguous nucleotides were assigned by degenerated base code) using Geneious 9.1.2 (Kearse et al., 2012).

Alignments of non-coding (18S rDNA) sequences were generated using the ClustalW algorithm. For COI and Cytb sequences, the nucleotide dataset was translated into amino acids (using mold protozoan mitochondrial genetic code) and aligned using the ClustalW algorithm; subsequently, aligned sequences were translated back to nucleotides (TransAlign, Geneious 9.1.2)(Kearse et al., 2012).

Sequence variation of 900 nt long fragment of COI was visualized as median-joining network inferred by PopART software (Leigh and Bryant, 2015). All graphic outputs were processed by InkScape v.092.1 (http://www.inkscape.org/).

2.5. Sensitivity assay

To compare the sensitivity of detection of used markers, PCR products from the first round of PCR of several samples were mixed (to obtain enough material), visualized on a gel, purified, cloned and sequenced as described above. Two clones for each marker were used to assess the assay sensitivity. To eliminate the potential effect of secondary structure (supercoil) of plasmid template DNA, pGEM clones containing desired inserts were linearized by restriction endonuclease *Nco*I (New England Biolabs, USA) which cuts just once within the vector sequence and does not affect any of insert sequences. The success of restriction digestion was confirmed by agarose gel electrophoresis. The concentration of purified and linearized clones was assessed in triplicates using 1×, 10× and 100× dilution by the Invitrogen Qubit dsDNA HS Assay (ThermoFisher Scientific, USA) with reliability across replicates higher than 0.99. Plasmid DNA copy number was calculated using the formula: $N = (m/M_r)/N_A$, where N is number of molecules, m is amount of dsDNA [g], M_r is relative molecular weight (computed as M_r = length of target sequence including pGEM®-T Easy vector DNA [bp] × average relative molecular weight per base [662]), and N_A is Avogadro's number [$6.022 \times 10^{23}\,mol^{-1}$]. Final template DNA was prepared by series of dilutions of the linearized pGEM®-T Easy plasmid containing insert into the background DNA isolated from *Babesia* spp. negative dog to the final concentration of 1 to 100 000 copies in 1 µl. Nested PCRs were performed as described above, 10 µl of each reaction were visualized on 1% agarose gel stained by MidoriGreen Advance (Nippon Genetics Europe, Germany).

3. Results

The GenBank search for "18S *Babesia canis*" or "18S *Babesia canis canis*" in mid-June 2018 retrieved 176 sequences of length ranging from 188 to 1714 nt. Eleven most divergent sequences were excluded from the dataset as not *B. canis*, based on nucleotide sequence alignment and BLAST search, resulting in final 165 sequences; 56% of these sequences were already published. Majority of the *B. canis* sequences originates either from the dog as expected host, or from its vector *D. reticulatus* (Fig. 1). Relatively high number of sequences (7/129) originates from

Table 1
Nucleotide sequences of PCR primers used for amplification and sequencing, length of products and annealing temperatures used.

	Primer name	Primer sequence (5´→3´)	Annealing temperature/product length	Reference
18S rDNA	BTH_F	CCTGMGARACGGCTACCACATCT	60 °C/ ~ 690 bp	Criado-Fornelio et al. (2003)
	BTH_R	TTGCGACCATACTCCCCCCA		
	GF1	GTCTTGTAATTGGAATGATGG	50 °C/ ~ 560 bp	Zintl et al. (2011)
	GR2	CCAAAGACTTTGATTTCTCTC		Bonnet et al. (2007)
COI	Bab_For1	ATWGGATTYTATATGAGTAT	45 °C/ ~ 1250 bp	Modified, Gou et al. (2012)
	Bab_Rev1	ATAATCWGGWATYCTCCTTGG		
	Bab_For2	TCTCTWCATGGWTTAATTATGATAT	49 °C/ ~ 980 bp	
	Bab_Rev2	TAGCTCCAATTGAHARWACAAAGTG		
Cytb	Bc_cytB_F1	TGGTCWTGGTATTCWGGAATG	50 °C/ ~ 700 bp	This study
	Bc_cytB_R1	AAGMYARTCTYCCTAAACATCC		
	Bc_cytB_F2	RATKAGYTAYTGGGGAGC	48 °C/ ~ 580 bp	
	Bc_cytB_R2	GCTGGWATCATWGGTATAC		

horses, but only single sequence from Spain was published (Criado-Fornelio et al., 2003), remaining samples from Italy or of unknown geographical origin keep unpublished status. Except of eight samples from Turkey and one from Israel, all samples originate from Europe. Dinucleotide region defining so-called "genotypes" was covered in 163 sequences with the dominance of GA (55.8%) and AG (39.3%) variants (Table 2 Table 2, Supplementary material).

From 99 dog samples screened in this study by nested PCR assay targeting nuclear 18S rDNA, 97 samples resulted in an amplicon of the expected size. Direct sequencing produced high quality, 484–559 nt long sequences from 94 of them; remaining three samples were sequenced after cloning.

Our ambition was to compare visual inspection of obtained chromatograms (ab1 files) with nucleotide sequences delivered as accompanying text files by sequencing company. Three apparent ambiguous sites were detected in chromatograms across the dataset in position 109 and in dinucleotide in position 129–130 of our alignment, corresponding to the nucleotides 589, 609 and 610 of the longest available 18S rDNA sequence GenBank acc. no. AY072926 (dog, Croatia). Clear double peaks AG/GA (assigned as RR) were detected in 87% of chromatograms in positions 129–130 (Fig. 2A, Table 2), double peak C/T (Y) in position 109 was less frequent (32%, Fig. 2A). In automatically read sequence txt files, positions 129–130 were read as GA in 86% of sequences, AG in 3%, RR in 1% and in remaining 10% as other combination of nucleotides (Table 2). Position 109 was automatically read as ambiguous Y in 7%. To clarify the combinations in these three ambiguous positions, we sequenced 38 clones originating from single sample with high quality sequencing chromatogram containing clear double peaks. In position 109, where double peak C/T was originally observed, exclusively C was found. Positions 129–130 (double peaks RR) displayed variant GA in 24 clones (63%) and variant AG in 13 clones (34%), single clone displayed variant GG. Surprisingly, in a position 146 (Fig. 2B), another polymorphic site was revealed. Nucleotide G was detected in all clones carrying GA variant in positions 129–130 and nucleotide A was in all clones carrying AG; the single clone with GG (positions 129–130) carry G in a position 146 (Fig. 2B).

Two other SNP sites supported by high-quality single peak chromatograms were detected in position 483 (position 963 in AY072926; A in one sequence, remaining 94 sequences were in accordance with the GenBank deposited sequences with T in this position) and dinucleotide in positions 501–502 (positions 981–982 in AY072926; GG in one sequence, AA in one sequence, remaining 90 sequences were in accordance with the GenBank deposited sequences with GA dinucleotide in this positions).

Amplification of mitochondrial COI marker was successful in 93 out of 99 samples. Direct sequencing using amplification primers was successful in 91 cases; two PCR products were cloned before sequencing. Resulting 900 nucleotide long alignment (guided by amino acid translation) corresponds to positions 241–1140 of full length COI sequence (GenBank acc. no. KC207822). In our dataset, neither ambiguous sites nor mixed chromatograms were observed for COI marker. Pairwise nucleotide sequence distances were minimal, varying in range 0–0.68%; all thirteen detected SNPs are listed in Table 3 including their amino acid mutation effect. Within the same fragment, the interspecific variability among GenBank sequences of COI of *B. canis*, *B. vogeli*, *B. rossi* and *B. gibsoni* (acc. no. KC207822, KC207825, KC207823, KP666169, respectively) ranges from 8 to 16% at the nucleotide level. Relationships between individual genotypes were displayed as haplotype network computed by the median-joining method (Fig. 3). Fifty-six samples are identical; eight less frequent genotypes were detected forming a star-shape map. In case of clear, no double peak chromatogram of 18S rDNA was observed; variant of dinucleotide 129–130 was assigned to corresponding COI haplotype with no clear correlation of 18S and COI variants (Fig. 3).

Second mitochondrial marker, cytochrome *b*, was amplified only in 25 pre-selected samples covering all variants of chromatogram results in 18S rDNA and all COI haplotypes. Fragment of 555 nucleotides corresponds to the positions 145–699 in ORF direction of a full-length

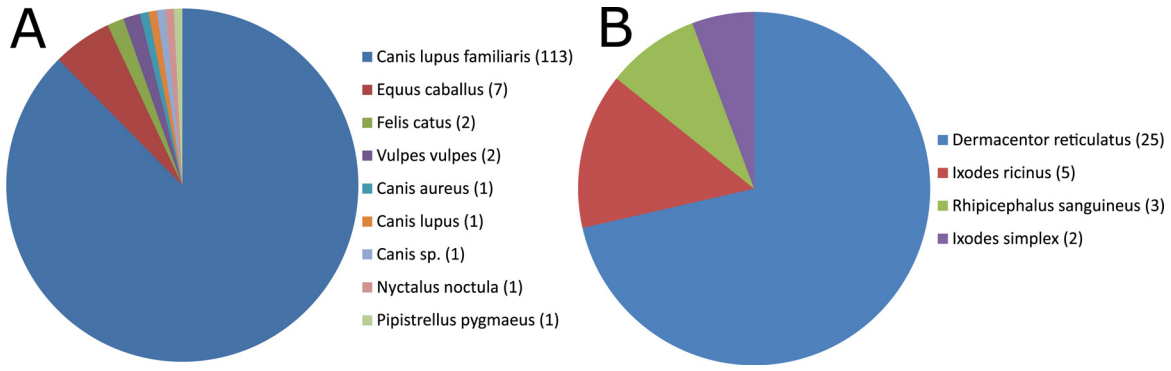

Fig. 1. Host (A) and vector (B) distribution of GenBank available *B. canis* sequences (as of 13th June 2018). The number of sequences is in brackets within the legend.

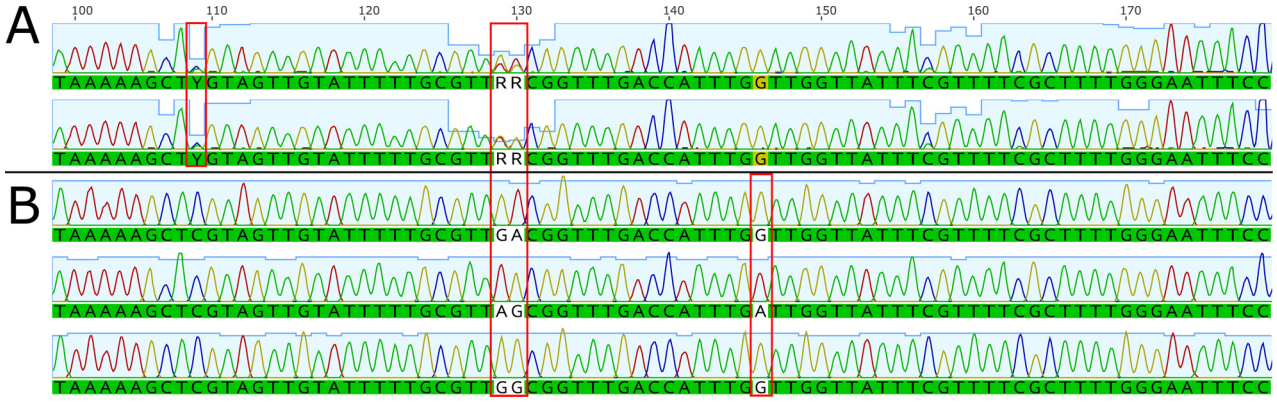

Fig. 2. Chromatograms of *B. canis* 18S rDNA sequences from dogs. A. Results of direct sequencing. SNPs (double peaks G/A in positions 129–130) correspond to the positions used to define *B. canis* genotypes. The nucleotide at position 109 (589 in AY072926) also displays clear double peak C/T. B. Results of cloned sequences. Newly uncovered SNP in position 146 is highlighted.

Table 2
Reading alternatives at positions 609–610 of 18S rDNA gene (numbering corresponds to the sequence AY072926) and their distribution (%) as retrieved from GenBank and from our dataset, presented as (i) visually read original ab1 file (column Chromatogram) and (ii) text file delivered by sequencing company (column Txt file).

	GenBank (n = 163)	Our dataset (n = 97)	
		Chromatogram	Txt file
RR	2.45	87.6	1.0
GA	55.80	10.3	85.6
AG	39.30	2.1	3.1
Other	2.45	0.0	10.3

Table 3
Single nucleotide polymorphisms in 900 nt fragment of 93 COI sequences; positions refer to complete COI sequence (acc. no. KC207822), variant in this reference sequence is highlighted in bold. Amino acids are described in standard single letter code.

Position	Variant frequency	Amino acid mutation
259	C (2) / **T** (91)	Silent (L)
271	C (1) / **T** (92)	Silent (L)
300	A (4) / **G** (88) / T (1)	Silent (P)
417	C (1) / **T** (92)	Silent (T)
486	A (1) / **G** (92)	**M** - > I
517	T (2) / **G** (91)	**A** - > S
705	A (14) / **G** (79)	Silent (Q)
711	A (5) / **G** (87)	Silent (L)
781	C (2) / **T** (91)	Silent (L)
948	G (9) / **A** (84)	Silent (L)
978	G (2) / **A** (91)	Silent (V)
1030	T (14) / **G**(79)	**A** - > S
1104	A (3) / **T** (90)	Silent (I)

coding sequence of Cytb (acc.no. KC207822). Four silent mutations (not changing coded amino acid) in two positions were observed (Fig. 4), resulting in pairwise nucleotide distance 0–0.36%. These Cytb variants doesn't correlate neither with the 18S rDNA nor with the COI haplotypes observed. Within the same fragment, the interspecific variability among GenBank sequences of Cytb of *B. canis*, *B. vogeli*, *B. rossi* and *B. gibsoni* (acc. no. KC207822, KC207825, KC207823, AB499087, respectively) ranges from 9 to 20% at the nucleotide level (Fig. 4).

All unique sequences were deposited to GenBank (accession numbers MK024714-6 for 18S rDNA, MK024717-24 for COI, and MK024725-7 for Cytb).

The sensitivity of nested PCR assays targeting nuclear 18S rDNA and two mitochondrial markers, COI and Cytb, was compared using

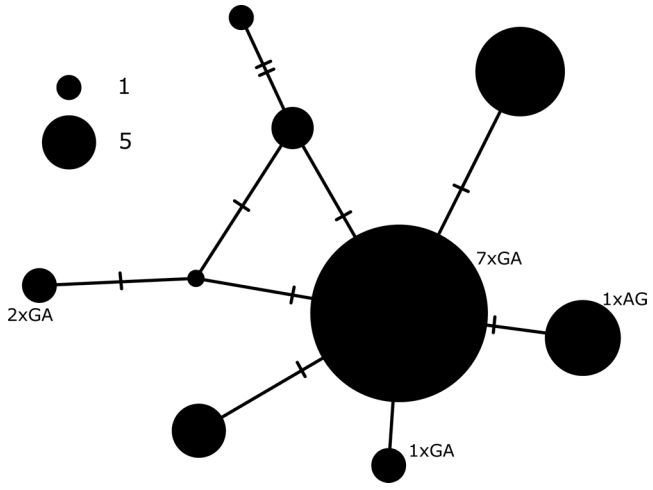

Fig. 3. Median-joining haplotype network of *B. canis* mitochondrial partial COI sequences. Size of circles corresponds to the haplotype frequency; hatch marks represent mutational steps separating haplotypes. Clear (non-double peak) 18S rDNA "genotypes" obtained from corresponding samples are indicated as GA/AG.

linearized clones of first-round PCR products cloned into pGEM®-T Easy vector. Two independent clones of each marker were diluted to the background DNA isolated from *Babesia* spp. negative dog to a final concentration of 1, 10, 10^2, 10^3, 10^4 and 10^5 molecules in one microliter used as a template for PCR. The 18S rDNA assay showed highest sensitivity varying from 1 to 10 molecules per reaction detected after the second round of PCR. After the first round of PCR, a band of similar intensity was observed in all template dilutions including background DNA itself (Fig. 5). Sequencing of the product of the first PCR confirmed unspecific amplification of host DNA, co-amplified with *B. canis* 18S rDNA in spiked samples. First round PCR product of mitochondrial markers was detectable in the reaction containing 10^3 copies for Cytb and 10^5 copies for COI. After the second round of PCR, Cytb was detectable from 10^2 to 10^3 copies, COI assay detection limit was 10^3 copies. The representative result of sensitivity assay for each marker is presented in Fig. 5.

4. Discussion

Emergence of canine babesiosis caused by *Babesia canis* sensu stricto across Europe follows recent range expansion of its dominant (if not exclusive) vector, *Dermacentor reticulatus* (Rubel et al., 2016). Domestic dogs are dominant hosts under recent conditions throughout Europe,

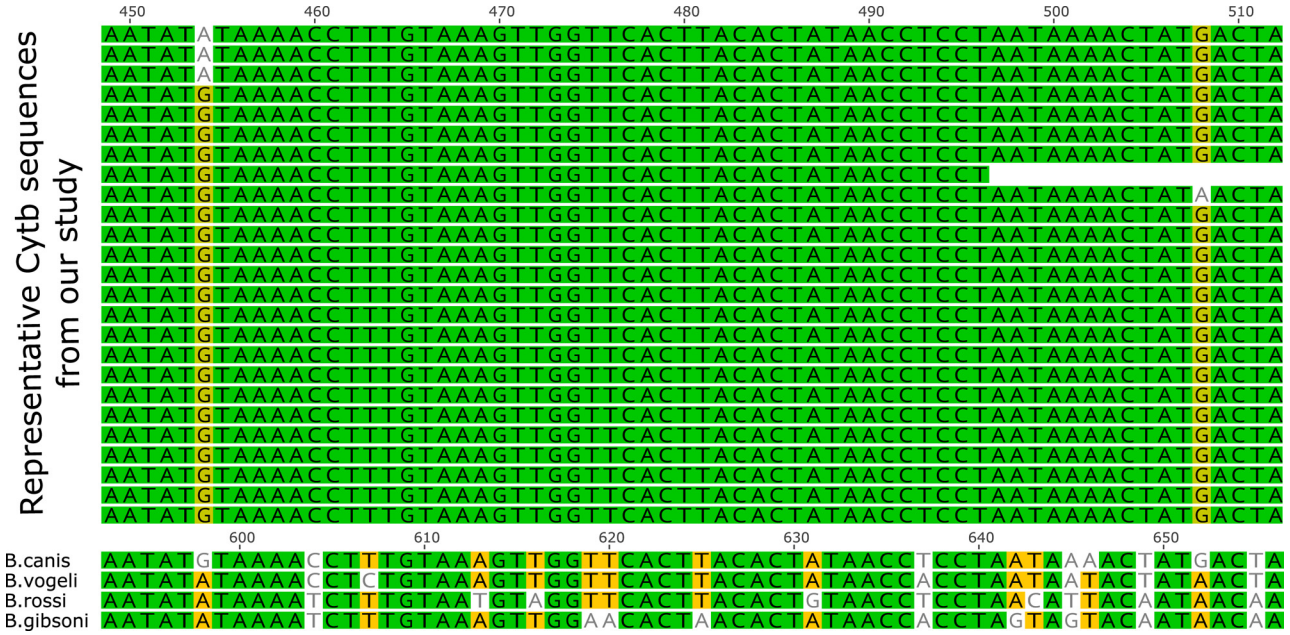

Fig. 4. Intra- and interspecific nucleotide sequence variability of Cytb of sequences of canine *Babesia* spp. Identical residues are highlighted in green. Numbering corresponds to a partial sequence of B. canis Cytb alignment from this study (upper part, 555 nt) and whole gene sequence alignment of representative Cytb sequences (*B. canis* KC207822, *B. vogeli* KC207825, *B. rossi* KC207823, *B. gibsoni* KP666169; the lower part, 1092 nt) (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

however, presence of *B. canis* in wolves and jackals (e.g. Mitková et al., 2017) together with distribution of the vector suggest the Palearctic canines of the genus *Canis* as original hosts.

In this study, we focused on the intraspecific variability of *B. canis* based on sequences of nuclear 18S rDNA and two mitochondrial markers (COI and Cytb). In the GenBank database, the 18S rDNA is the most abundant *B. canis* sequence available, others are represented by sporadic sequences of merozoite surface protein Bc28.1 (n = 3), adenosine kinase (n = 1), secreted antigen 1 (n = 1), heat shock protein 70 (n = 2), inosine 5′-monophosphate dehydrogenase (n = 1), complete mitochondrion genome (n = 1) and a partial cytochrome oxidase subunit I (n = 2).

The 18S rDNA gene has been used extensively for classification of apicomplexans parasites, being well suited for resolving phylogenies at higher taxonomic levels (Morrison et al., 2004; Schnittger et al., 2012). However, insufficient sequence variation among closely related species (Gou et al., 2012), together with the presence of paralogs disqualify, in part, 18S rDNA gene for species discrimination or studies on intraspecific variability. Multiple, divergent gene copies (paralogs) of 18S rDNA were described within the genome in all major Apicomplexa groups: cryptosporidia (Le Blancq et al., 1997; Stenger et al., 2015; Xiao et al., 1999), coccidia (El-Sherry et al., 2013; Ogedengbe et al., 2018), plasmodia (McCutchan et al., 1995; Rooney, 2004), and also in piroplasmids (Dalrymple, 1990; Kibe et al., 1994; Reddy et al., 1991). In malarial parasites, these paralogs encode structurally distinct rRNA molecules that are expressed at different life cycle stages and can have even greater sequence divergence than homologous rDNA copies from closely related species (McCutchan et al., 1995).

In *B. canis*, the diversity in 18S rDNA was interpreted as “genotypes” (Adaszek et al., 2009; Adaszek and Winiarczyk, 2008; Hornok et al., 2015; Łyp et al., 2015). These hypothetic genotypes rely on dinucleotide polymorphism in positions 609–610 (corresponding to sequence AY072926), when GA variant is called “A” and AG is called “B”. Third variant TT in the same positions remained unnamed (Łyp et al., 2015).

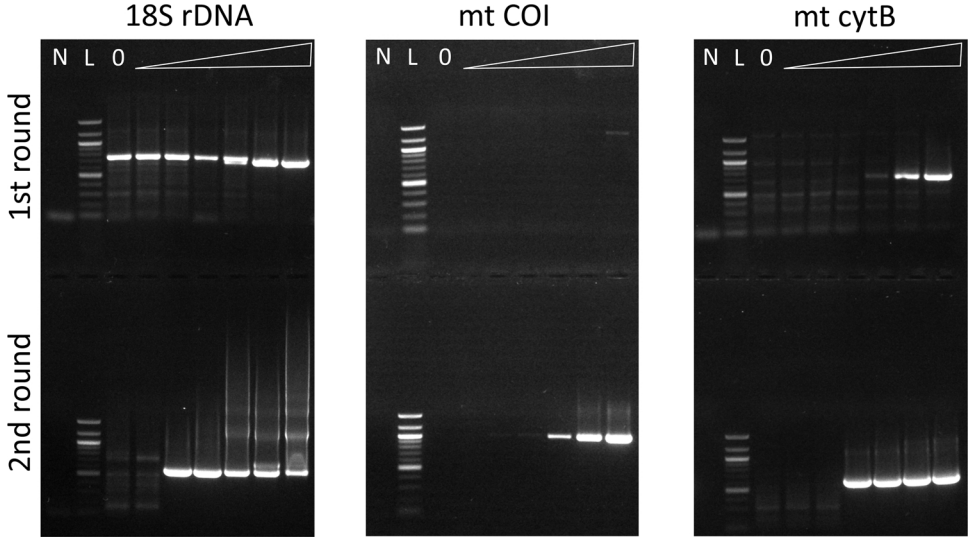

Fig. 5. Assay sensitivity comparison for nuclear (18S rDNA) and mitochondrial markers (COI, Cytb). From both PCR rounds, 10 µl were analyzed on 1% agarose gel stained by Midori Green Advance DNA stain. N - negative control with no template DNA; L - 100 bp DNA ladder (thick bands represents 500 bp and 1000 bp); 0 - background DNA of *Babesia* spp. negative dog; template dilution represents 1, 10, 10^2, 10^3, 10^4 and 10^5 copies of template DNA in the reaction (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

In raw sequencing data obtained, we observed at least three ambiguous positions in up to 86% of chromatograms within the ~560 bp fragment of 18S rDNA suggesting the existence of several, not identical copies of this gene (Fig. 2). AG/GA double peaks in positions defining "genotypes" were already reported in *B. canis* sequences from naturally infected dogs in Croatia (Beck et al., 2009), Switzerland (Schaarschmidt et al., 2013) and Lithuania (Paulauskas et al., 2014). Sequencing of clones originating from a single sample confirmed existence of two major and one rare 18S rDNA variants (Fig. 2B). Such findings are rather suggestive for a presence of paralogs, as described in other piroplasmids, where three copies of ribosomal RNA coding units in *B. bovis* (Brayton et al., 2007; Dalrymple, 1990), *B. bigemina* (Reddy et al., 1991), and two copies in *T. parva* (Kibe et al., 1994) were observed, similar to other apicomplexans. The "universal" existence of observed 18S rDNA variants throughout the distribution of *Babesia canis* deserves further attention.

Analysis of sequencing results by terminal users is based either on visual inspection of an ab1 file or simply on adoption of obtained, automatically read sequence provided in the txt file. The visual reading of chromatograms (ab1 files) we obtained collide with automatically read sequences (txt files). As obvious from Table 2, the double peak was apparent in 85 samples, however, in 84 of this samples (99%), the double peak presence was not indicated in accompanying txt file.

Sequences in the text file provided by the sequencing companies result from an algorithm set up to single base-call method reading each nucleotide position as one definite base based on Phred score as a basis for a quality value assessment (Rockmann et al., 2018). Then, the sequences from different studies deposited in GenBank differ de facto stochastically based on the approach used for sequencing data analysis and resulting discrepancies are obvious from our data as well. To prevent this bias in future studies, we can only recommend visual inspection of chromatograms prior the sequence analysis. Alternatively, modified base calling can be provided (upon request) by sequencing company or by additional analysis (e.g. plugin Heterozygotes in Geneious software). Then, a degenerated base is indicated in situations when the second highest peak in the same position is more than a certain proportion of the highest peak.

An alternative explanation of double peak presence in sequencing results is a co-infection by two *B. canis* lineages in the respective dog. It seems fairly unlikely that majority of dogs (86% in our study, 66% in Lithuania) and all ticks in Switzerland study (Paulauskas et al., 2014; Schaarschmidt et al., 2013) are simultaneously infected by two different strains. Our results together with the above-stated arguments bring a substantial doubt about the existence *B. canis* 18S rDNA "genotypes" as such, and, consequently, about the differences observed in their occurrence and seasonality. Rather, the uneven detection of AG/GA variants attributes to base calling setting and to amplification bias. This inherent stochasticity of the PCR process (Best et al., 2015; Peccoud and Jacob, 1996) affects obtained results on two independent levels: (i) in the heterogeneity of PCR itself, and, (ii) in a sampling of the amplified product for sequencing. When one particular variant of the gene is randomly or preferentially amplified in the early cycles of the PCR, this particular one prevails in the final sequencing results.

Compared to nuclear 18S rDNA, mitochondrial genes are evolving faster, resulting in higher sequence divergence (Gray, 1999). To scrutinize the alternative scenario of co-occurrence of several *B. canis* variants, we sequenced two mitochondrial markers and explored their sequence diversity. The COI gene is the most often used barcoding marker in the animal kingdom (Hebert et al., 2003b, 2003a) and it was used for closely related species differentiation of apicomplexans parasites (Calzada et al., 2014; Ogedengbe et al., 2018; Schreeg et al., 2016). Mitochondrial markers (especially COI) were recently used to elucidate also the phylogenetic relationships within the order Piroplasmida (Annoscia et al., 2017; Gou et al., 2012; Hikosaka et al., 2010; Schreeg et al., 2016). Our COI haplotype analysis resulted in a star-like pattern suggesting a recent origin of most haplotypes and indicating a population expansion in recent history (Miro et al., 2011), but not supporting the existence of two dominant haplotypes. Moreover, the samples with 18S rDNA sequences that were read without ambiguities either as GA or as AG were scattered among the several haplotypes within COI network. Similarly, the Cytb sequences obtained from samples with all variants of 18S rDNA (AG, GA and RR) were identical.

The 18S rDNA is most common (if not only) marker used for diagnostics of piroplasmid infections in clinical praxis (Lempereur et al., 2017). This gene has both conserved and variable regions, the former allowing universal primer design across genera and the latter providing phylogenetic information. However, the usefulness of the 18S rDNA for discrimination of piroplasmids at species (or lower) level was repeatedly questioned (Schnittger et al., 2012; Schreeg et al., 2016; Uilenberg et al., 2018). There is relatively little information on the use of other DNA markers for studies on piroplasmid diversity. Pioneering work of Gou et al. (2012), comparing species identification efficiency of the most common nuclear markers (18S rDNA, 28S rDNA, ITS) and mitochondrial COI of piroplasmids, is mostly ignored. Few other studies used complete mitochondrial genomes to solve the higher phylogeny of piroplasmids (Hikosaka et al., 2010; Schreeg et al., 2016), but numbers of available sequences are very limited.

Amplification of three different markers (18S rDNA, COI, Cytb) from a rather homogenous set of *B. canis* infected dogs allows also evaluation of their applicability as diagnostic and/or population variability markers. The high sensitivity (as low as one molecule per reaction) and a wide range of target species of 18S rDNA assay (data not shown) confirm its strength as a first choice diagnostic marker. On contrary, the presence of several different gene copies (paralogs) disqualifies its use for phylogenetic and infraspecific variability studies. In the same manner, high interspecific variability makes COI a promising candidate. However, in contrast to 18S rDNA and in agreement with other studies (Gou et al., 2012), our COI amplification assay suffered from considerably lower sensitivity (10^3 copies in reaction). The lower efficiency of amplification using degenerated primers and high AT content of this gene (Pan et al., 2014) stays probably behind observed sensitivity difference. Finally, the Cytb displayed middle sensitivity compared to 18S rDNA and COI and lowest intraspecific variability and seems to be a good candidate as an alternative diagnostic target for *B. canis*.

Apparently, neither the 18S rDNA nor any other single marker can answer a broad range of questions we ask. We bring the evidence of the existence of 18S rDNA paralogs in *B. canis* genome, based on detailed sequence analysis of three markers, replacing currently used "genotype" theory. Three observed ambiguous sites in a partial 18S rDNA sequence of *B. canis* are probably only a part of variable sites across the entire gene. The final number of different 18S rDNA copies, their variability, and general understanding of this phenomenon across apicomplexan parasites requires further studies. Only a growing number of whole genome sequences of a wide range of piroplasmid species can provide ultimate answers.

Funding

All grants funding this study are stated under Acknowledgment section.

Conflict of interest

All authors declare that he/she has no conflict of interest.

Ethical approval

This article does not contain any studies with human participants or animals performed by any of the authors.

Declaration of interests

None.

Acknowledgements

This research was carried out under the project CEITEC 2020 (LQ1601) with financial support from the Ministry of Education, Youth and Sports of the Czech Republic under the National Program of Sustainability II, under the project LO1218 with financial support from the Ministry of Education, Youth and Sports of the Czech Republic under the NPU I program, and further co-financed from the European Social Fund and the state budget of the Czech Republic (project OPVK CZ.1.07/2.3.00/20.0300). We acknowledge a grant for the development of research organization (RVO: RO0516). This study was partly supported by the Polish Ministry of Science and Higher Education (1016/S/IGM/2016).

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.vetpar.2018.12.017.

References

Adaszek, L., Winiarczyk, S., 2008. Molecular characterization of *Babesia canis canis* isolates from naturally infected dogs in Poland. Vet. Parasitol. 152, 235–241. https://doi.org/10.1016/j.vetpar.2007.12.024.

Adaszek, Ł., Winiarczyk, S., 2010. Application of the SYBR Green real-time HRM PCR technique in the differentiation of the *Babesia canis canis* protozoa isolated in the areas of eastern Poland. Parasitol. Res. 106, 1253–1256. https://doi.org/10.1007/s00436-010-1784-3.

Adaszek, Ł., Winiarczyk, S., Skrzypczak, M., 2009. The clinical course of babesiosis in 76 dogs infected with protozoan parasites Babesia canis canis. Pol. J. Vet. Sci. 12, 81–87.

Annoscia, G., Stefania, M., Cantacessi, C., Olivieri, E., Teresa, M., Dantas-torres, F., Otranto, D., 2017. A new PCR assay for the detection and differentiation of *Babesia canis* and *Babesia vogeli*. Ticks Tick. Borne Dis. 8, 862–865. https://doi.org/10.1016/j.ttbdis.2017.07.002.

Beck, R., Vojta, L., Mrljak, V., Marinculić, A., Beck, A., Živičnjak, T., Cacciò, S.M., 2009. Diversity of *Babesia* and *Theileria* species in symptomatic and asymptomatic dogs in Croatia. Int. J. Parasitol. 39, 843–848. https://doi.org/10.1016/j.ijpara.2008.12.005.

Best, K., Oakes, T., Heather, J.M., Shawe-Taylor, J., Chain, B., 2015. Computational analysis of stochastic heterogeneity in PCR amplification efficiency revealed by single molecule barcoding. Sci. Rep. 5, 1–13. https://doi.org/10.1038/srep14629.

Bonnet, S., Jouglin, M., Malandrin, L., Becker, C.A.M., Agoulon, A., L'Hostis, M., Chauvin, A., 2007. Transstadial and transovarial persistence of *Babesia divergens* DNA in *Ixodes ricinus* ticks fed on infected blood in a new skin-feeding technique. Parasitology 134, 197–207. https://doi.org/10.1017/S0031182006001545.

Brayton, Ka., 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, Ha., Haas, B.J., Howell, J.M., Khouri, H., Koo, H., Mann, D.J., Norimine, J., Paulsen, I.T., Radune, D., Ren, Q., Smith, R.K., Suarez, C.E., White, O., Wortman, J.R., Knowles, D.P., McElwain, T.F., Nene, V.M., 2007. Genome sequence of *Babesia bovis* and comparative analysis of apicomplexan hemoprotozoa. PLoS Pathog. 3, 1401–1413. https://doi.org/10.1371/journal.ppat.0030148.

Calzada, E., Xayavong, M., Parise, M., Alves, M.P., Marı, A., Rodrigues, P.T., Silva, A.J., Ferreira, M.U., Alves, J.M.P., Santamaria, A.M., Calzada, J.E., Xayavong, M., Parise, M., Da Silva, A.J., Ferreira, M.U., 2014. Using mitochondrial genome sequences to track the origin of imported *Plasmodium vivax* infections diagnosed in the United States. Am. J. Trop. Med. Hyg. 90, 1102–1108. https://doi.org/10.4269/ajtmh.13-0588.

Criado-Fornelio, a., Martinez-Marcos, a., Buling-Saraña, a., Barba-Carretero, J.C., 2003. Molecular studies on *Babesia*, *Theileria* and *Hepatozoon* in southern Europe: part II. Phylogenetic analysis and evolutionary history. Vet. Parasitol. 114, 173–194. https://doi.org/10.1016/S0304-4017(03)00141-9.

Dalrymple, B.P., 1990. Cloning and characterization of the rRNA genes and flanking regions from *Babesia bovis*: use of the genes as strain discriminating probes. Mol. Biochem. Parasitol. 43, 117–124. https://doi.org/10.1016/0166-6851(90)90136-A.

Dautel, H., Dippel, C., Oehme, R., Hartelt, K., Schettler, E., 2006. Evidence for an increased geographical distribution of *Dermacentor reticulatus* in Germany and detection of *Rickettsia* sp. RpA4. Int. J. Med. Microbiol. 296, 149–156. https://doi.org/10.1016/j.ijmm.2006.01.013.

El-Sherry, S., Ogedengbe, M.E., Hafeez, M.A., Barta, J.R., 2013. Divergent nuclear 18S rDNA paralogs in a turkey coccidium, *Eimeria meleagrimitis*, complicate molecular systematics and identification. Int. J. Parasitol. 43, 679–685. https://doi.org/10.1016/j.ijpara.2013.03.005.

Gou, H., Guan, G., Liu, A., Ma, M., Xu, Z., Liu, Z., Ren, Q., Li, Y., Yang, J., Chen, Z., Yin, H., Luo, J., 2012. A DNA barcode for piroplasmea. Acta Trop. 124, 92–97. https://doi.org/10.1016/j.actatropica.2012.07.001.

Gray, M.W., 1999. Evolution of organellar genomes. Curr. Opin. Genet. Dev. 9, 678–687. https://doi.org/10.1016/S0959-437X(99)00030-1.

Hebert, P.D.N., Cywinska, A., Ball, S.L., Jeremy, R., DeWaard, J.R., Jeremy, R., DeWaard, J.R., 2003a. Biological identifications through DNA barcodes. Proc. R. Soc. B Biol. Sci. 270, 313–321. https://doi.org/10.1098/rspb.2002.2218.

Hebert, P.D.N., Ratnasingham, S., de Waard, J.R., Jeremy, R., de Waard, J.R., 2003b. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. Proc. R. Soc. B Biol. Sci. 270, S96–S99. https://doi.org/10.1098/rsbl.2003.0025.

Hikosaka, K., Watanabe, Y.I., Tsuji, N., Kita, K., Kishine, H., Arisue, N., Palacpac, N.M.Q., Kawazu, S.I., Sawai, H., Horii, T., Igarashi, I., Tanabe, K., 2010. Divergence of the mitochondrial genome structure in the apicomplexan parasites, *Babesia* and *Theileria*. Mol. Biol. Evol. 27, 1107–1116. https://doi.org/10.1093/molbev/msp320.

Hornok, S., Kartali, K., Takács, N., Hofmann-Lehmann, R., 2015. Uneven seasonal distribution of *Babesia canis* and its two 18S rDNA genotypes in questing *Dermacentor reticulatus* ticks in urban habitats. Ticks Tick. Borne Dis. 7, 16–19. https://doi.org/10.1016/j.ttbdis.2016.02.016.

Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., Drummond, A., 2012. Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28, 1647–1649. https://doi.org/10.1093/bioinformatics/bts199.

Kibe, M.K., ole-MoiYoi, O.K., Nene, V., Khan, B., Allsopp, B.A., Collins, N.E., Morzaria, S.P., Gobright, E.I., Bishop, R.P., 1994. Evidence for two single copy units in *Theileria parva* ribosomal RNA genes. Mol. Biochem. Parasitol. 66, 249–259.

Kubelová, M., Sedlák, K., Panev, A., Široký, P., 2013. Conflicting results of serological, PCR and microscopic methods clarify the various risk levels of canine babesiosis in Slovakia: a complex approach to *Babesia canis* diagnostics. Vet. Parasitol. 191, 353–357. https://doi.org/10.1016/j.vetpar.2012.09.016.

Le Blancq, S.M., Khramtsov, N.V., Zamani, F., Upton, S.J., Wu, T.W., 1997. Ribosomal RNA gene organization in *Cryptosporidium parvum*. Mol. Biochem. Parasitol. 90, 463–478. https://doi.org/10.1016/S0166-6851(97)00181-3.

Leigh, J.W., Bryant, D., 2015. POPART: full-feature software for haplotype network construction. Methods Ecol. Evol. 6, 1110–1116. https://doi.org/10.1111/2041-210X.12410.

Lempereur, L., Beck, R., Fonseca, I., Marques, C., Duarte, A., Santos, M., Zúquete, S., Gomes, J., Walder, G., Domingos, A., Antunes, S., Baneth, G., Silaghi, C., Holman, P., Zintl, A., 2017. Guidelines for the detection of *Babesia* and *theileria* parasites. Vector-Borne Zoonotic Dis. 17, 51–65. https://doi.org/10.1089/vbz.2016.1955.

Łyp, P., Adaszek Furmaga, B., Winiarczyk, S., 2015. Identification of new 18S rRNA strains of *Babesia canis* isolated from dogs with subclinical babesiosis. Pol. J. Vet. Sci. 18, 573–577. https://doi.org/10.1515/pjvs-2015-0074.

Máthé, Á., Vörös, K., Papp, L., Reiczigel, J., 2006. Clinical manifestations of canine babesiosis in Hungary (63 cases). Acta Vet. Hung. 54, 367–385. https://doi.org/10.1556/AVet.54.2006.3.7.

Matijatko, V., Torti, M., Schetters, T.P., 2012. Canine babesiosis in Europe: How many diseases? Trends Parasitol. 28, 99–105. https://doi.org/10.1016/j.pt.2011.11.003.

McCutchan, T.F., Li, J., Mcconkey, Ga, Rogers, M.J., Waters, A.P., 1995. The cytoplasmic ribosomal of *Plasmodium* spp. Parasitol. Today 11, 134–138. https://doi.org/10.1016/0169-4758(95)80132-4.

Miro, F., Weijie, Q., Bo, H., Ferreri, M., Qu, W., Han, B., 2011. Phylogenetic networks: a tool to display character conflict and demographic history. Afr. J. Biotechnol. 10, 12799–12803. https://doi.org/10.5897/AJB11.010.

Mitková, B., Hrazdilová, K., D'Amico, G., Duscher, G.G., Suchentrunk, F., Forejtek, P., Gherman, C.M., Matei, I.A., Ionică, A.M., Daskalaki, A.A., Mihalca, A.D., Votýpka, J., Hulva, P., Modrý, D., 2017. Eurasian Golden Jackal as Host of Canine Vector-borne Protists. Parasit. Vectors 10. https://doi.org/10.1186/s13071-017-2110-z.

Mitkova, B., Hrazdilova, K., Novotna, M., Jurankova, J., Hofmannova, L., Forejtek, P., Modry, D., 2017. Autochthonous *Babesia canis*, *Hepatozoon canis* and imported *Babesia gibsoni* infection in dogs in the Czech Republic. Vet. Med. (Praha) 62, 138–146. https://doi.org/10.17221/152/2016-VETMED.

Morrison, D.A., Bornstein, S., Thebo, P., 2004. The current status of the small subunit rRNA phylogeny of the coccidia (Sporozoa). J. Parasitol. 34, 501–514. https://doi.org/10.1016/j.ijpara.2003.11.006.

Ogedengbe, M.E., El-Sherry, S., Ogedengbe, J.D., Chapman, H.D., Barta, J.R., 2018. Phylogenies based on combined mitochondrial and nuclear sequences conflict with morphologically defined genera in the eimeriid coccidia (Apicomplexa). Int. J. Parasitol. 48, 59–69. https://doi.org/10.1016/j.ijpara.2017.07.008.

Pan, W., Byrne-Steele, M., Wang, C., Lu, S., Clemmons, S., Zahorchak, R.J., Han, J., 2014. DNA polymerase preference determines PCR priming efficiency. BMC Biotechnol. 14, 1–16. https://doi.org/10.1186/1472-6750-14-10.

Paulauskas, A., Radzijevskaja, J., Karvelienø, B., Grigonis, A., Aleksandravičienø, A., Zamokas, G., Babickaitø, L., Sabunas, V., Petkevičius, S., 2014. Detection and molecular characterization of canine babesiosis causative agent *Babesia canis* in the naturally infected dog in Lithuania. Vet. Parasitol. 205, 702–706. https://doi.org/10.1016/j.vetpar.2014.09.001.

Peccoud, J., Jacob, C., 1996. Theoretical uncertainty of measurements using quantitative polymerase chain reaction. Biophys. J. 71, 101–108. https://doi.org/10.1016/S0006-3495(96)79205-6.

Randolph, S.E., EDEN-TBD sub-project team, 2010. Human activities predominate in determining changing incidence of tick-borne encephalitis in Europe. Euro Surveill. 15, 24–31. https://doi.org/10.2807/ese.15.27.19606-en.

Reddy, G., Chakrabarti, D., Yowell, C.A., Dame, J.B., 1991. Sequence microheterogeneity of the three small subunit ribosomal RNA genes of *Babesia bigemina*: expression in erythrocyte culture. Nucleic Acids Res. 19, 3641–3645. https://doi.org/10.1093/

nar/19.13.3641.
Rockmann, C., Endrullat, C., Frohme, M., Pospisil, H., Sciences, A., 2018. Next generation sequence analysis. Ref. Modul. Life Sci. 4, 2–5. https://doi.org/10.1016/B978-0-12-809633-8.20206-9.
Rooney, A.P., 2004. Mechanisms underlying the evolution and maintenance of functionally heterogeneous 18S rRNA genes in apicomplexans. Mol. Biol. Evol. 21, 1704–1711. https://doi.org/10.1093/molbev/msh178.
Rubel, F., Brugger, K., Pfeffer, M., Chitimia-Dobler, L., Didyk, Y.M., Leverenz, S., Dautel, H., Kahl, O., 2016. Geographical distribution of *Dermacentor marginatus* and *Dermacentor reticulatus* in Europe. Ticks Tick. Borne Dis. 7, 224–233. https://doi.org/10.1016/j.ttbdis.2015.10.015.
Schaarschmidt, D., Gilli, U., Gottstein, B., Marreros, N., Kuhnert, P., Daeppen, J.A., Rosenberg, G., Hirt, D., Frey, C.F., 2013. Questing *Dermacentor reticulatus* harbouring *Babesia canis* DNA associated with outbreaks of canine babesiosis in the Swiss Midlands. Ticks Tick. Borne Dis. 4, 334–340. https://doi.org/10.1016/j.ttbdis.2013.01.007.
Schnittger, L., Rodriguez, A.E., Florin-christensen, M., Morrison, Da, 2012. *Babesia*: a world emerging. Infect. Genet. Evol. 12, 1788–1809. https://doi.org/10.1016/j.meegid.2012.07.004.
Schreeg, M.E., Marr, H.S., Tarigo, J.L., Cohn, L.A., Bird, D.M., Scholl, E.H., Levy, M.G., Wiegmann, B.M., Birkenheuer, A.J., 2016. Mitochondrial genome sequences and structures aid in the resolution of Piroplasmida phylogeny. PLoS One 11, 1–27. https://doi.org/10.1371/journal.pone.0165702.
Solano-Gallego, L., Baneth, G., 2011. Babesiosis in dogs and cats-expanding parasitological and clinical spectra. Vet. Parasitol. 181, 48–60. https://doi.org/10.1016/j.vetpar.2011.04.023.
Solano-Gallego, L., Sainz, Á., Roura, X., Estrada-Peña, A., Miró, G., 2016. A review of canine babesiosis: the European perspective. Parasit. Vectors 9, 336. https://doi.org/10.1186/s13071-016-1596-0.
Stenger, B.L.S., Clark, M.E., Kváč, M., Khan, E., Giddings, C.W., Dyer, N.W., Schultz, J.L., McEvoy, J.M., 2015. Highly divergent 18S rRNA gene paralogs in a *Cryptosporidium* genotype from eastern chipmunks (*Tamias striatus*). Infect. Genet. Evol. 32, 113–123. https://doi.org/10.1016/j.meegid.2015.03.003.
Uilenberg, G., Gray, J., Kahl, O., 2018. Research on Piroplasmorida and other tick-borne agents: are we going the right way? Ticks Tick. Borne Dis. 9, 0–1. https://doi.org/10.1016/j.ttbdis.2018.03.005.
Xiao, L., Limor, J.R., Li, L., Morgan, U., Thompson, R.C., Lal, A.A., 1999. Presence of heterogeneous copies of the small subunit rRNA gene in *Cryptosporidium parvum* human and marsupial genotypes and Cryptosporidium felis. J. Eukaryot. Microbiol. 46, 44S–45S.
Zintl, A., Finnerty, E.J.E., Murphy, T.T.M., de Waal, T., Gray, J.S.J., Waal, Tde, Gray, J.S.J., 2011. Babesias of red deer (*Cervus elaphus*) in Ireland. Vet. Res. 42, 7. https://doi.org/10.1186/1297-9716-42-7.