

Molecular differentiation of *Sarcocystis buffalonis* and *Sarcocystis levinei* in water buffaloes (*Bubalus bubalis*) from *Sarcocystis hirsuta* and *Sarcocystis cruzi* in cattle (*Bos taurus*)

Bjørn Gjerde¹ · Mosaad Hilali² · Ibrahim E. Abbas³

Received: 30 January 2016 / Accepted: 9 March 2016 / Published online: 16 March 2016
© Springer-Verlag Berlin Heidelberg 2016

Abstract The purpose of the present study was to obtain sarcocysts of *Sarcocystis buffalonis* and *Sarcocystis levinei* from water buffaloes and characterize the isolates by molecular methods in order to determine whether the two species were genetically different from *Sarcocystis hirsuta* and *Sarcocystis cruzi*, respectively, from cattle, which had been characterized before. About 35 macroscopically visible (3–4 × 1–2 mm) and 20 barely visible (1–3 × 0.2 mm) sarcocysts were excised from the esophagus of 18 naturally infected and freshly slaughtered adult water buffaloes at three slaughterhouses in Egypt. Genomic DNA was extracted from the sarcocysts, and all isolates were first characterized at the mitochondrial cytochrome c oxidase subunit I gene (*cox1*) gene through PCR amplification and direct sequencing. Selected isolates were subsequently further characterized at the 18S and 28S ribosomal (r) RNA genes and the internal transcribed spacer 1 (ITS1) region of the nuclear rDNA unit by direct sequencing or cloning. Only six of the isolated macroscopic sarcocysts belonged to *S. buffalonis*, whereas the others belonged to *Sarcocystis fusiformis*. Twelve of the smaller

cysts belonged to *S. levinei* and seven to *Sarcocystis sinensis*. The characterization of the sarcocysts of *S. sinensis* and some of the sarcocysts of *S. fusiformis* have been reported before. Fifteen additional sarcocyst isolates of *S. fusiformis* were characterized at *cox1* in the present study and found to be identical or closely similar to previous isolates. At *cox1*, the sequence identity between the six isolates of *S. buffalonis* was 99.8–100 % (two haplotypes), whereas the identity between the 12 isolates of *S. levinei* was 99.0–100 % (10 haplotypes). The identity between *cox1* sequences of *S. buffalonis* and *S. hirsuta* ($n=56$) was 92.9–93.6 % (on average 93.4 %), and the identity between *cox1* sequences of *S. levinei* and *S. cruzi* ($n=22$) was 92.9–94.0 % (on average 93.5 %). The phylogenetic analyses placed with high support the *cox1* sequences of *S. buffalonis* and *S. hirsuta* into two monophyletic sister groups, and the same was true for the *cox1* sequences of *S. levinei* and *S. cruzi*. Hence, the study established that *S. buffalonis* and *S. levinei* are distinct species different from *S. hirsuta* and *S. cruzi*, respectively. Nucleotide sequences of *S. buffalonis* could be distinguished from those of *S. hirsuta* also at the 28S rRNA gene (clearly different) and the ITS1 region (small and uncertain difference) but not at the 18S rRNA gene. Sequences of *S. levinei* could be distinguished from those of *S. cruzi* both at the 18S and 28S rRNA genes (ITS1 region not examined). However, the *cox1* gene was superior to the 18S and 28S rRNA genes as regards the ability to unambiguously delimit the species within each species pair, since at the latter markers, the number of consistent nucleotide differences between the species was low and there was a slight overlap between the intraspecific and interspecific sequence divergence. Comparison of the newly generated 18S rRNA gene sequences of *S. levinei* from water buffaloes with similar sequences deposited in GenBank suggested that *S. levinei* and *S. cruzi* are not strictly intermediate host specific but might occasionally infect cattle and water buffaloes, respectively.

Electronic supplementary material The online version of this article (doi:10.1007/s00436-016-4998-1) contains supplementary material, which is available to authorized users.

✉ Bjørn Gjerde
bjorn.gjerde@nmbu.no; bkgjerde@hotmail.no

¹ Department of Food Safety and Infection Biology, Norwegian University of Life Sciences, P.O. Box 8146 Dep., 0033 Oslo, Norway

² Department of Parasitology, Faculty of Veterinary Medicine, Cairo University, 12211 Giza, Egypt

³ Faculty of Veterinary Medicine, Mansoura University, Mansoura, Egypt

Keywords *Sarcocystis buffalonis* · *Sarcocystis levinei* · *Sarcocystis hirsuta* · *Sarcocystis cruzi* · *Sarcocystis fusiformis* · Water buffalo · *Bubalus bubalis* · Cattle · *Bos taurus* · Egypt · Phylogeny · 18S rRNA · 28S rRNA · ITS1 · *cox1*

Introduction

Protozoan parasites of the genus *Sarcocystis* have an obligatory two-host life cycle comprising sexual development and oocyst formation in the intestinal mucosa of their definitive hosts and asexual multiplication in vascular endothelial cells (schizont stage) and striated muscle cells (sarcocyst stage) of their intermediate hosts. Definitive hosts become infected through ingestion of sarcocysts in muscle tissues, whereas intermediate hosts become infected through ingestion of oocysts/sporocysts in fecally contaminated feed or water.

The Indian water buffalo (*Bubalus bubalis*) is the natural intermediate host of the four named species *Sarcocystis fusiformis*, *Sarcocystis buffalonis*, *Sarcocystis levinei*, and *Sarcocystis sinensis* (Gjerde 2015, 2016). In addition, a *Sarcocystis hominis*-like species has been found by transmission electron microscopy (TEM) of sarcocysts from water buffaloes in the Philippines (Parairo et al. 1988; see Gjerde 2016) and by analysis of the partial 18S ribosomal (r) RNA gene of sarcocyst isolates from water buffaloes in China (Yang et al. 2001a, b, 2002). This species might be identical with *S. hominis* in cattle, since the latter species has been shown to infect water buffaloes via sporocysts excreted by humans (Chen et al. 2003). Likewise, the three named species *S. buffalonis*, *S. levinei*, and *S. sinensis* are closely similar to the species *Sarcocystis hirsuta*, *Sarcocystis cruzi*, and *Sarcocystis bovifelis/Sarcocystis bovini*, respectively, in cattle as regards sarcocyst morphology and 18S rRNA gene sequences (Holmdahl et al. 1999; Yang et al. 2001a, b, 2002; Li et al. 2002; Jehle et al. 2009; Xiang et al. 2011; Moré et al. 2014; Gjerde 2015, 2016). Hence, some of these papers have suggested or concluded that these species are shared between the two hosts (Yang et al. 2001a, b, 2002; Li et al. 2002; Jehle et al. 2009; Xiang et al. 2011; Moré et al. 2014).

However, in a recent comprehensive molecular comparison of *S. sinensis* from water buffaloes with two similar species in cattle, *S. bovifelis* and *S. bovini*, it was shown that *S. sinensis* was indeed a distinct species (Gjerde 2015). Using the partial mitochondrial cytochrome c oxidase subunit I gene (*cox1*) as a genetic marker, sequences of *S. sinensis* differed by about 11 % from sequences of *S. bovifelis* and *S. bovini*, and the phylogenetic analysis placed the *cox1* sequences of the three species into three separate monophyletic clusters. In contrast, there were only minor differences between these species at the nuclear 18S and 28S rRNA genes, and no obvious differences at all between the three species at the internal transcribed

spacer 1 (ITS1) region of the nuclear ribosomal DNA unit. Likewise, *S. fusiformis* was found to be closely similar to *Sarcocystis cafferi* from the African buffalo (*Syncerus caffer*) with respect to sarcocyst morphology and 18S rRNA gene sequences (Dubey et al. 2014) but proved to be clearly different from the latter species based on partial *cox1* sequences (Gjerde et al. 2015). These studies (Gjerde et al. 2015; Gjerde 2015), as well as previous comparisons of two pairs of morphologically similar *Sarcocystis* spp. in reindeer and red deer (Gjerde 2013; Gjerde 2014), have shown that the *cox1* marker is better able than the 18S rRNA gene to separate recently diverged and thus closely related *Sarcocystis* spp. in closely related intermediate hosts.

The major aim of the present study was therefore to obtain sarcocyst isolates of *S. buffalonis* and *S. levinei* from water buffaloes and characterize these isolates at the partial *cox1* gene in order to determine whether these species were distinct from *S. hirsuta* and *S. cruzi*, respectively, in cattle, which already had been thoroughly characterized at this marker in previous studies (Gjerde 2013, 2015). An additional aim was to characterize selected isolates identified through the *cox1* marker also at the 18S and 28S rRNA genes and ITS1 region, in order to determine whether isolates from the two hosts could be separated on the basis of these markers. A characterization of the 18S rRNA gene would also allow the new isolates of *S. buffalonis* and *S. levinei* to be compared with GenBank sequences obtained from similar species in water buffaloes in previous studies.

Materials and methods

Collection of sarcocysts from water buffaloes

In November 2014, macroscopically visible sarcocysts (3–4 mm long and 1–2 mm wide) consistent with those of *S. buffalonis* (Huong et al. 1997a) were excised from the esophagus of eight freshly slaughtered naturally infected adult water buffaloes at the El Warak slaughterhouse, Giza, Egypt. In addition, small pieces of esophageal muscle containing barely visible microscopic sarcocysts (1–3 mm long and about 0.2 mm wide) consistent with those of *S. levinei* (Dissanaike and Kan 1978; Huong et al. 1997b) and *S. sinensis* (Chen et al. 2011) were collected. In April 2015, additional macroscopic *S. buffalonis*-like sarcocysts were excised from the esophagus of 10 freshly slaughtered adult water buffaloes at slaughterhouses in Mansoura City and Belkas City in Dakahlia Governorate, Egypt (135 km northwest of Cairo City). The excised sarcocysts and muscle tissue were immediately fixed in 70 % ethanol and pooled in a few microcentrifuge tubes (mainly according to host animal and sarcocyst size) before being shipped to the first author for molecular examination. Prior to DNA extraction, the fixed material in each tube was

poured into a Petri dish and examined under a stereo microscope in order to place individual sarcocysts into separate tubes and remove most of the muscle tissue still surrounding some of the larger sarcocysts. Moreover, the muscle tissue was carefully examined for the presence of any microscopic sarcocysts that might have passed undetected during sampling. Some of the smaller sarcocysts were so deeply embedded in the muscle tissue that they were difficult to discern in the fixed material, and hence, the adjacent muscle tissue was not removed before DNA extraction. In total, about 35 macroscopic *S. buffalonis*-like sarcocysts and about 20 fairly small sarcocysts were available for molecular characterization.

Molecular examination of sarcocysts

After removal of ethanol, genomic DNA was extracted from the sarcocysts using QIAmp DNA Mini Kit (Qiagen, Germany) according to the manufacturer's tissue protocol. The DNA samples were subsequently kept frozen at -20°C in between their use as templates for PCR amplifications. Four DNA regions were amplified as described previously, including the primer sequences (Gjerde 2015). From all available isolates, an ~1100-bp-long portion of *cox1* was amplified with primer pair SF1/SR9 in order to identify the sarcocysts to species and study the variability of this marker. From selected isolates of *S. buffalonis* and *S. levinei*, the complete (~1880 bp) 18S rRNA gene was amplified in two overlapping fragments using primer pairs ERIB1/S2r and S3f/Primer Bsarc, whereas the partial 28S rRNA gene (~1650 bp) was amplified with primer pair KL1/KL3. From selected isolates of *S. buffalonis*, the complete ITS1 region and short segments of the flanking 18S and 5.8S rRNA genes (~800 bp) were amplified with primer pair 18ShsF/5.8SR2. Initially, attempts were made to amplify this region with primer pair SU1F/5.8SR2, which targets many *Sarcocystis* spp., but only sequences of an unidentified DNA region were obtained, which was also the case with the related species *S. hirsuta* in a previous study (Gjerde 2015). That problem was then solved by designing forward primer 18ShsF, targeting a variable region 92–113 nucleotides from the 3' end of the 18S rRNA gene of *S. hirsuta*. This primer also had a complete match with sequences of *S. buffalonis* and was therefore used for this species also.

All procedures concerning PCR amplification and the evaluation, purification, and sequencing of PCR products were as previously described (Gjerde 2015). The same was true for procedures concerning cloning of selected PCR products and the subsequent purification and sequencing of plasmid DNA. Thus, due to intrainolate sequence variation, the 5' end half of the 18S rRNA gene (about 1000 bp generated with primer pair ERIB1/S2r), the ITS1 region, and the partial 28S rRNA gene of *S. buffalonis*, as well as the partial 28S rRNA gene of *S. levinei*, were cloned before sequencing in order to

obtain unambiguous sequences. However, for some isolates of one or both species, amplicons from the 18S and 28S rRNA genes were also sequenced directly (Table 1). The sequences were assembled using the Alignment Explorer of the MEGA5 software (Tamura et al. 2011) and compared with sequences in GenBank using the nucleotide BLAST program as previously described (Gjerde 2013). The new sequences from the rDNA unit of *S. buffalonis* and *S. levinei* were also visually compared with homologous GenBank sequences of *S. hirsuta* and *S. cruzi*, respectively, in multiple alignments in the Alignment Explorer of MEGA5 in an attempt to uncover differences and similarities between sequences of each species pair (Gjerde 2015). The software package DnaSP (DNA Sequence Polymorphism) version 5.10.01 (Librado and Rozas 2009) was used for the analysis of nucleotide polymorphisms (sequence variation) among the newly generated *cox1* sequences of *S. fusiformis*, *S. buffalonis*, and *S. levinei* and for comparison of these sequences with previous sequences of the same (*S. fusiformis*) or related species (*S. hirsuta* and *S. cruzi*). These analyses included enumeration of variable (polymorphic) sites and haplotypes among the sequences and estimation of different measures of sequence divergence within and between species (Gjerde 2013).

Phylogenetic analyses

Phylogenetic analyses were conducted separately on nucleotide sequences of *cox1*, the 18S rRNA gene, the 28S rRNA gene, and the ITS1 region, respectively, by means of the MEGA5 software (Tamura et al. 2011). All sequences included in the analyses of each DNA region and their GenBank accession numbers are listed in Table S1 in the Supplementary material. In all analyses, the phylogeny was tested with the bootstrap method using 1000 bootstrap replications.

For *cox1*, a total of 563 partial sequences from 37 taxa were initially included in the analysis, including 33 new *cox1* sequences generated from the three *Sarcocystis* spp. examined at this marker in this study. However, in order to reduce computation time, identical superfluous sequences were removed, so that each haplotype (based on nucleotides 1–1020; see below) was only represented by a single sequence. Hence, for the final analysis, a total of 297 sequences (haplotypes) were used. A codon-based multiple alignment of all sequences was obtained by using ClustalW within MEGA5 as described previously (Gjerde 2013). Since some sequences were only 1020-bp long, sequences longer than this were truncated at their 3' end, so that the final alignment comprised 1020 positions with no gaps. The phylogenetic tree was reconstructed using the maximum parsimony (MP) method with the subtree-pruning-regrafting (SPR) algorithm. All codon positions were used. The intestinal coccidium *Eimeria tenella* of chickens was used as outgroup species to root the tree.

Table 1 Overview of the molecular characterization at four DNA regions of different sarcocyst isolates from water buffaloes (Bb) and the resulting GenBank sequences

Species	<i>cox1</i>	18S rRNA gene	ITS1	28S rRNA gene
<i>S. buffalonis</i>	<i>n</i> = 6; Bb5.3, Bb9.2, Bb12.1, Bb14.2, Bb15.2, and Bb18.1 KU247868–KU247873	<i>n</i> = 13; Bb12.1, Bb14.2 (1 clone), Bb15.2, and Bb18.1 (10 clones) KU247901–KU247913	<i>n</i> = 10; Bb14.2 (2 clones) and Bb18.1 (8 clones) KU247923–KU247932	<i>n</i> = 4; Bb12.1 and Bb18.1 (3 clones) KU247933–KU247936
<i>S. levinei</i>	<i>n</i> = 12; Bb7.1, Bb12.2, Bb16.3, Bb19.2–Bb19.5, and Bb20.1–Bb20.5 KU247874–KU247885	<i>n</i> = 9; Bb7.1, Bb12.2, Bb16.3, Bb19.2–Bb19.5, Bb20.1, and Bb20.2 KU247914–KU247922	nd nd	<i>n</i> = 9; Bb16.3, Bb19.2, Bb19.4, Bb19.5, and Bb20.1 (5 clones) KU247937–KU247945
<i>S. fusiformis</i>	<i>n</i> = 15; Bb21.1–Bb21.6, Bb21.16, and Bb22.1–Bb22.8 KU247886–KU247900	nd nd	nd nd	nd nd

The PCR products were either sequenced directly or cloned before sequencing (from selected isolates as stated)

nd not done

As regards the 18S rRNA gene, a total of 102 near-full-length sequences from 52 taxa were used in the analysis, including 22 new sequences of the two *Sarcocystis* spp. examined in the present study. A multiple sequence alignment was generated with the ClustalW program within MEGA5, using a gap opening penalty of 10 and a gap extension penalty of 0.1 and 0.2 for the pair-wise and multiple alignments, respectively. Most sequences were truncated slightly at both ends, so that most sequences started and ended at the same homologous nucleotide positions, corresponding to positions 91 and 1792, respectively, of GenBank sequence KT901117 of *S. bovifelis*. The final alignment comprised 1985 aligned positions, including gaps. The phylogenetic tree was reconstructed using the MP method with the SPR algorithm. All sites were used. A sequence of *E. tenella* of chickens was used as outgroup species to root the tree.

For the 28S rRNA gene, a total of 99 partial sequences of 39 taxa were included in the analysis, including 13 new sequences of the two *Sarcocystis* spp. examined in the present investigation. A multiple sequence alignment was generated with the ClustalW program within MEGA5, using a gap opening penalty of 10 and a gap extension penalty of 0.8 for both the pair-wise and the multiple alignment. Most sequences were truncated at both ends, so that nearly all sequences started and ended at the same nucleotide positions, corresponding to positions 50 and 1561, respectively, of GenBank sequence KT901245 of *S. bovifelis*. The final alignment comprised 1810 aligned positions, including gaps. The phylogenetic tree was reconstructed using the MP method with the SPR algorithm. All sites were used. A sequence of *E. tenella* of chickens was used as outgroup to root the tree.

As regards the ITS1 region, a total of 98 sequences from seven taxa were included in the analysis, including 10 new sequences of *S. buffalonis* from the present investigation. A multiple sequence alignment was generated with the MUSCLE program within MEGA5, using the default settings,

but slightly corrected manually, so that closely similar sequences were treated in the same manner. The GenBank sequences were truncated at both ends so that only the ITS1 region was included in the analysis. There were a total of 761 positions in the final dataset. The phylogenetic tree was reconstructed using the MP method with the SPR algorithm. All sites were used. A sequence of *S. cruzi* was used as outgroup to root the tree.

Results

Partial *cox1* gene

DNAs from a total of 30 *S. buffalonis*-like sarcocysts were successfully amplified with primer pair SF1/SR9 and sequenced with good results, whereas a few other isolates of such cysts resulted in no visible amplification products or poor sequences due to a mixture of DNA from two or more species. However, only 6 of the 30 isolates were identified as sarcocysts of *S. buffalonis* (inferred from both *cox1* and 18S rRNA gene sequences), whereas the other 24 macroscopic sarcocysts belonged to *S. fusiformis*. The latter species also predominated in the mixed sequences. Nine *cox1* sequences of *S. fusiformis* were obtained from sarcocysts collected on the first sampling occasion in November 2014 and were included in the previous molecular study of this species (Gjerde et al. 2015). The 15 new *cox1* sequences of *S. fusiformis* (KU247886–KU247900) from this study comprised seven haplotypes, of which five were new and two were known from the previous 33 *cox1* sequences (KR186081–KR186113), making it a total of 18 haplotypes among 48 *cox1* sequences of this species. The 18 haplotypes differed from each other at 1–7 of 1038 nucleotide positions (99.3–99.9 % identity), and there were a total of 19 variable (polymorphic) sites.

The six new *cox1* sequences of *S. buffalonis* (KU247868–KU247973) comprised only two haplotypes. Five sequences were identical and differed from the remaining sequence (KU247870) at only two of 1038 nucleotide positions (99.8 % identity). They were most similar to sequences of *S. hirsuta* (KC209634 and KT901023–KT901077), sharing an identity of 92.9–93.6 % (on average 93.4 %), followed by those of *S. fusiformis* (89.2–89.6 % identity). Compared with the 56 sequences of *S. hirsuta*, they differed at 66–76 of 1038 nucleotide positions (on average 68.5 positions), and there were 60 fixed nucleotide differences between the two populations, that is, positions where all sequences of *S. buffalonis* differed from all sequences of *S. hirsuta*. By comparison, the intraspecific identity of all sequences of *S. hirsuta* was 99.1–100 % (Gjerde 2015).

Twelve microscopic sarcocysts were found to belong to *S. levinei* based on *cox1* and 18S rRNA gene sequences, whereas seven such cysts belonged to *S. sinensis* as reported elsewhere (Gjerde 2015). The 12 *cox1* sequences of *S. levinei* (KU247874–KU247885) comprised 10 haplotypes, differing from each other at 1–10 positions (99.0–100 % identity between all sequences). They were most similar to sequences of *S. cruzi* (KC209597–KC209600 and KT901078–KT901095), sharing an identity of 92.9–94.0 % (on average 93.5 %), followed by those of *Sarcocystis rangi* and *Sarcocystis hjorti* (89–90 % identity). The sequences of *S. levinei* differed from the sequences of *S. cruzi* at 62–74 of 1038 nucleotide positions (on average 67.5 positions), of which 56 were fixed nucleotide differences. By comparison, the 22 *cox1* sequences of *S. cruzi* comprised 15 haplotypes differing at 1–15 of 1038 nucleotide positions (98.6–100 % intraspecific sequence identity; Gjerde 2015).

Complete 18S rRNA gene

Initially, the 18S rRNA genes of five isolates of *S. buffalonis* were PCR amplified and sequenced directly, but only the sequences of the 3' end half of the gene were of good quality throughout their length, whereas those from the 5' end half became poor to unintelligible about 200 bp downstream from both primers due to indels, which caused different sequence variants to become juxtaposed and superimposed on each other in the sequence reads. Hence, the sequences of only two isolates (Bb12.1 and Bb15.2) could be fairly accurately estimated after direct sequencing (KU247901 and KU247903). Amplicons from two other isolates (Bb14.2 and Bb18.1) were cloned, and 1 and 11 clones, respectively, were obtained and sequenced from these isolates. Two of the clones from isolate Bb18.1 were identical, and hence, sequences from a total of 11 clones were submitted to GenBank (Table 1). The full-length sequences were 1880–1886 bp long and differed by three 3-bp-long indels, as well as a moderate number of nucleotide substitutions. The three indels were located at nucleotide

positions 214–216 and 251–253 and between positions 855 and 856, respectively, of sequence KU247911. Overall, the 13 new sequences of *S. buffalonis* shared an identity of 98.7–99.9 % with each other and 99.1–99.9 % identity with the only full-length sequence of this species in GenBank (AF017121). Likewise, they shared 98.8–99.7 % identity with two ~1550-bp-long sequences from water buffaloes in China, which had been submitted to GenBank (Yang et al. 2001b) as sequences of *S. hirsuta* (AF170940 and AF176941). Their identity with 14 full-length sequences of *S. hirsuta* from cattle (AF017122, JX855283, KC209741, and KT901156–KT901166) was 98.3–99.6 % (Table S2 in the Supplementary material). The indels were located at the same positions in sequences of both species, but whereas 16 of 17 sequences derived from water buffaloes had a deletion at the first two indels, only 1 of 16 sequences of *S. hirsuta* derived from cattle had deletions in these positions. The single sequence of each species that deviated from the majority (KU247911 of *S. buffalonis* and KT901160 of *S. hirsuta*) thus shared the highest identity (up to 99.6 %) with sequences of the opposite species. Similarly, there were a few nucleotide positions in which the majority of sequences of one species differed from those of the other species, but there were no positions in which all sequences of the two species differed from each other. Hence, an unambiguous differentiation of *S. buffalonis* from *S. hirsuta* on the basis of the 18S rRNA gene sequences did not seem to be possible.

The complete 18S rRNA gene was amplified and sequenced from nine isolates of *S. levinei*, yielding nine 1867-bp-long sequences (KU247914–KU247922; Table 1). Eight of them were identical, whereas one sequence (KU247918) differed by a single substitution (at nucleotide position 140) from the others (99.9–100 % identity). This seemed to be a polymorphic site as suggested by a double peak (A/G) in the chromatograms of several isolates. The new sequences of *S. levinei* were first compared in a multiple alignment with 10 sequences of *S. cruzi* from cattle (KC209738–KC209740 and KT901167–KT901173), which had been obtained from isolates that were also examined at *cox1* (Gjerde 2013, 2015), and thus known to be different from *S. levinei* at this marker. This comparison revealed consistent differences between the two sets of sequences in three regions, consisting of three indels and five nucleotide substitutions (Fig. S1 in the Supplementary material). Downstream of position 253, two nucleotides were deleted in sequences of *S. levinei* compared with those of *S. cruzi* (–TAT versus AACAG), but at the two other indels, one and three nucleotides, respectively, were inserted in sequences of *S. levinei* compared with those of *S. cruzi* (TT/–C and TTA/– –). Thus, the full-length gene of *S. levinei* was two nucleotides longer than that of *S. cruzi*, and there was about 99.3 % identity between the sequences of the two species (Table S2 in the Supplementary material). At the aforementioned polymorphic site at position 140, all

sequences of *S. cruzi* had a G, whereas all but one sequence of *S. levinei* had an A.

The new sequences of *S. levinei* were next compared in a multiple alignment with about 100 other *S. cruzi*-like sequences retrieved from GenBank, of which 90 originated from sarcocyst isolates from cattle and 10 from isolates from water buffaloes. All but one of the sequences from cattle had been assigned to *S. cruzi*, the exception being sequence KM434885, which was designated *Sarcocystis* cf. *cruzi*. The latter designation had also been used for seven sequences from water buffaloes in China (HM447190–HM447196), whereas three sequences from this host had been attributed to *S. cruzi* (AF176932, AF176935, and KT306827). Eight previous sequences of *S. cruzi* (AF017120, AB682779–AB682782, JX679467, JX679468, and KP640133) comprised the (near) full-length 18S rRNA gene and thus included all regions in which the abovementioned sequences of *S. cruzi* had been found to differ from those of *S. levinei*. About 60 of the sequences of *S. cruzi* were from Malaysian cattle (*Bos indicus*) and covered only about 1000 bp in the middle region of the gene (Ng et al. 2015) and consequently did not include the two distinguishing indels close to either end of the gene. However, 31 other sequences of *S. cruzi* from cattle (EF622146–EF622176), which had been generated in order to compare the ITS1 region of different isolates (Rosenthal et al. 2008), comprised the distinguishing indel in the terminal portion of the gene. None of the 10 previous sequences from water buffaloes, on the other hand, extended to this indel located about 100 bp from the 3' end of the gene.

The comparisons showed that eight of 10 sequences from water buffaloes (AF176935, HM447190–HM447194, HM447196, and KT306827) were closely similar to the present sequences of *S. levinei*, whereas one sequence (AF176932) was consistent with *S. levinei* at the first indel but with *S. cruzi* in the variable region between positions 750 and 776. Moreover, one sequence (HM447195) was fully consistent with sequences of *S. cruzi* in both variable regions. Conversely, one sequence from cattle in India (KM434885) was closely similar to the new sequences of *S. levinei*, whereas another sequence from cattle in India (KT306829) was consistent with *S. levinei* in the first variable region and with *S. cruzi* in the second variable region and thus similar to the abovementioned hybrid-like sequence from a water buffalo. All remaining sequences from cattle covering one or both variable regions near both ends of the gene were consistent with the abovementioned sequences of *S. cruzi* from isolates that had also been examined at *cox1*. The same was true for the majority of sequences from cattle concerning the variable region between positions 750 and 780. However, some sequences (e.g., JX679468 and KJ917907–KJ917916) differed from this common sequence type by various inserts/deletions but were nevertheless different from sequences of *S. levinei*.

Complete ITS1 region

The ITS1 region of two isolates (Bb14.2 and Bb18.1) of *S. buffalonis* was PCR amplified with primer pair 18ShsF/5.8SR2 and cloned, and a total of 11 clones were sequenced. Two of the clones from isolate Bb18.1 were identical (represented by KU247927), and hence, 10 unique sequences were submitted to GenBank (KU247923–KU247932; Table 1). These full-length sequences were 780–815 bp long, including primers, of which the variable ITS1 region comprised 549–584 bp. The difference in length was due to 12 indels, each 1–22 bp long. In addition, the sequences differed by several substitutions, resulting in a wide range of sequence identities (91–99 %) between individual clones. By comparison, the ITS1 region of 24 sequences of *S. hirsuta* (KT901209–KT209232) from a previous study (Gjerde 2015) was 541–572 bp long and shared 10 of the indels with sequences of *S. buffalonis*. The identities between the clones of the two species were 89–95 %. Certain sequence features were common in one species and rare in the other, but all except one were not unique for one species. The only consistent difference was that all sequences of *S. buffalonis* had a string of 10–13 consecutive Ts, whereas this string was interrupted by CG (21 sequences), C (2 sequences), or CCG (1 sequence) in sequences of *S. hirsuta* (e.g., TTT-TTTTTTTT versus TTCGTTTTTTTT). Since there were some variations among the sequences of *S. hirsuta*, this difference might not be present in other variants of this species.

Partial 28S rRNA gene

Five sequences of the partial 28S rRNA gene of *S. buffalonis* were obtained, one by direct sequencing of PCR amplicons from one isolate and four after cloning of amplicons from another isolate (Table 1). Two of the latter clones were identical (represented by KU247934), and hence, a total of four different sequences were submitted to GenBank (KU247933–KU247936). The sequences were 1604 or 1610 bp long with primers included, the difference in length being due to two 3-bp-long indels, appearing as deletions in the shortest sequence (KU247936). The four sequences of *S. buffalonis* shared an identity of 99.3–99.9 % with each other. Their identity with six homologous sequences of *S. hirsuta* (KT901264–KT901269) from a previous study (Gjerde 2015) was 97.7–98.3 %. The latter sequences were 1608–1617 bp long and all of them differed from those of *S. buffalonis* by seven substitutions and two inserts, which were 3 and 4 bp long, respectively. Some sequences of *S. hirsuta* also differed from all sequences of *S. buffalonis* by an additional 6-bp-long insert and a 3-bp-long deletion, whereas a 3-bp-long insert in four of five sequences of *S. buffalonis* was absent from all sequences of *S. hirsuta*. Based on these consistent differences, the two species could be separated from each other.

Ten sequences of the partial 28S rRNA gene of *S. levinei* were obtained, four by direct sequencing of PCR amplicons from four isolates and six after cloning of amplicons from a fifth isolate (Table 1). Two of the latter clones were identical and are represented by a single sequence (KU247941). Two identical sequences were also obtained from two separate sarcocyst isolates by direct sequencing, but both were submitted to GenBank. Thus, a total of nine sequences were submitted (KU247937–KU247945). These sequences shared an identity of 99.1–100 % with each other and an identity of 97.7–98.4 % with the predominant sequence type found in 11 of 16 homologous sequences of *S. cruzi* (KT901270–KT901285) in a previous study (Gjerde 2015). The identity with the five aberrant sequences of *S. cruzi* (KT901272, KT901275, KT901280, KT901282, and KT901285) was only 95.1–95.9 %. The latter sequences have therefore not been included in the following comparison or in the phylogenetic analysis using 28S rRNA gene sequences (next section). The sequences of *S. levinei* were 1661–1666 bp long with primers included, whereas those of the predominant type of *S. cruzi* were 1652–1657 bp long. The differences between sequences of *S. levinei* were due to two indels (2 and 3 bp) and several substitutions, whereas the differences between sequences of *S. levinei* and *S. cruzi*, respectively, were due to the same and additional indels and substitutions. There were five substitutions and one 6-bp-long indel that separated all sequences of *S. levinei* from the abovementioned 11 typical sequences *S. cruzi*, as well as GenBank sequence AF076903 of *S. cruzi*. As regards the latter indel, sequences of *S. levinei* (positions 1226–1232 in KU247937) displayed AGATATT, whereas those of *S. cruzi* had G followed by a 6-bp-long deletion in the homologous region. From these consistent differences, it was possible to separate *S. levinei* from *S. cruzi* on the basis of the 28S rRNA gene sequences.

Phylogenetic relationships

The phylogenetic analysis based on partial *cox1* sequences (Fig. 1), placed with near-maximum support, all sequences of *S. buffalonis* as a monophyletic sister group to all sequences of *S. hirsuta*. Both taxa were sister to sequences of *S. fusiformis*, and together, these three species formed a sister group to *S. gigantea* within a clade comprising species with feline or unknown definitive hosts. This clade was separated with maximum support from the clade comprising species with canids as known or presumed definitive hosts. In the latter clade, all sequences of *S. levinei* were placed with near-maximum support as a monophyletic sister group to all sequences of *S. cruzi*. Both species formed a sister group to *S. hjorti* and *S. rangi* from cervid hosts, but the relationships between these species were not well resolved.

In the phylogenetic analyses based on near full-length 18S rRNA gene sequences (Fig. S2 in the Supplementary

material), sequences of *S. buffalonis* and *S. hirsuta* tended to cluster according to species but were not completely separated. Thus, one sequence of *S. buffalonis* (KU247911) was placed within a cluster comprising 13 of 14 sequences of *S. hirsuta*, whereas two other sequences of *S. buffalonis* were placed as a sister to this group. Similarly, one sequence of *S. hirsuta* (KT901160) was placed within a cluster comprising 11 of 14 sequences of *S. buffalonis*. Both species formed a sister group to *S. fusiformis* and *S. cafferi*. All sequences of *S. levinei*, on the other hand, formed with fairly high bootstrap support a monophyletic sister group to all sequences of *S. cruzi*. Both species were sister to a group of four species (*Sarcocystis alceslatrans*, *Sarcocystis capreolicanis*, *S. hjorti*, and *S. rangi*) from cervid hosts with a closely similar sarcocyst morphology (Gjerde 2012).

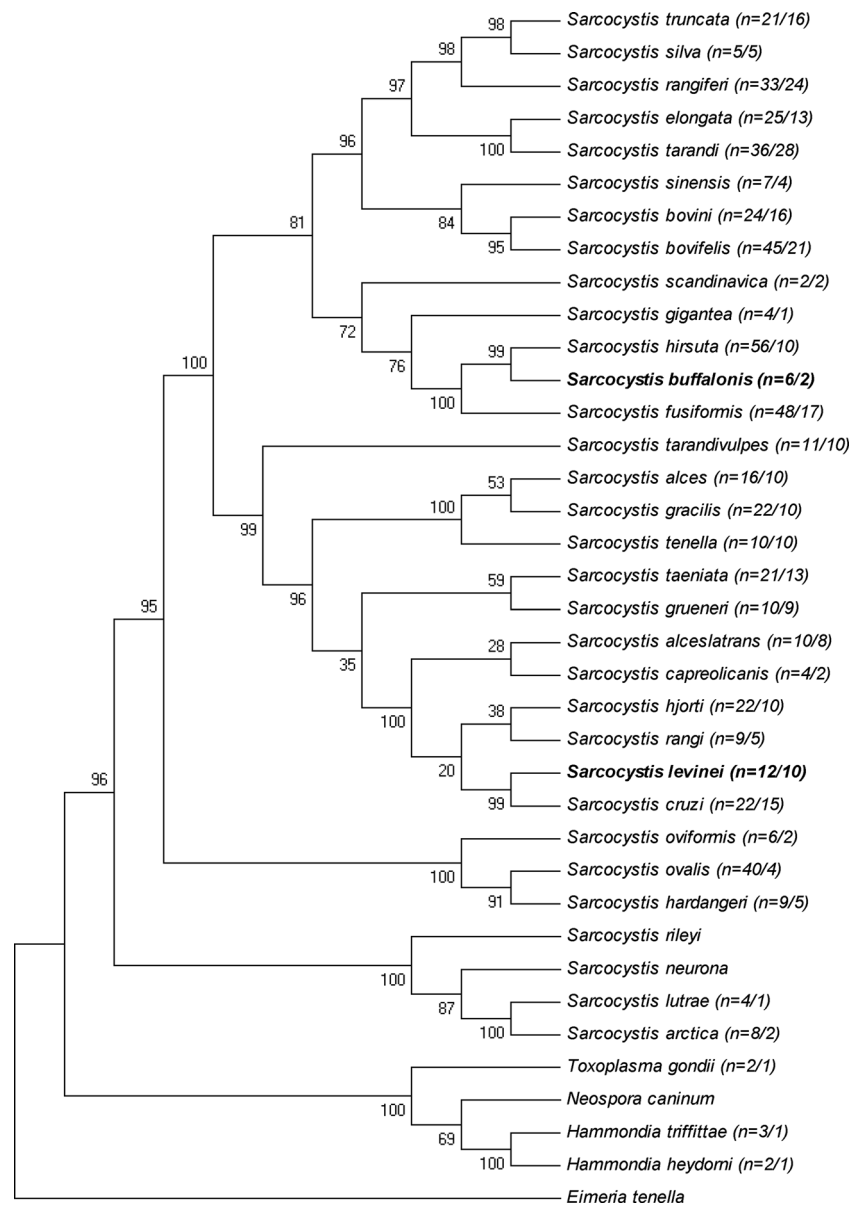
In the phylogenetic tree inferred from partial 28S rRNA gene sequences (Fig. 2), the four sequences of *S. buffalonis* were placed with high support in a monophyletic cluster, which was sister to a cluster comprising all six sequences of *S. hirsuta*. Both species were sister to sequences of *S. fusiformis*. Likewise, sequences of *S. levinei* and *S. cruzi*, respectively, were placed separately with fairly high support into two monophyletic clusters (sister groups). The two species were sister to a single sequence of *S. rangi*. Sequences of *S. bovifelis* and *S. bovini* were interleaved with each other but were placed with maximum support separately from those of *S. sinensis*.

In the phylogeny inferred from the ITS1 sequences (Fig. S3 in the Supplementary material), all sequences of *S. buffalonis* were placed with near-maximum support in a monophyletic cluster, which was sister to a similar cluster of all sequences of *S. hirsuta*. Both species formed a sister group to sequences of *S. fusiformis*. Sequences of *S. bovifelis*, *S. bovini*, and *S. sinensis*, on the other hand, were not placed in monophyletic clusters. Thus, one sequence of *S. bovini* was interleaved with sequences of *S. bovifelis*, whereas three other sequences of that species were interleaved with sequences of *S. sinensis* (data not shown). Sequences of *S. bovifelis* and *S. sinensis*, on the other hand, were not interleaved with each other.

Discussion

Using the mitochondrial *cox1* gene, this study has shown unequivocally that *S. buffalonis* and *S. levinei* in water buffaloes are distinct from *S. hirsuta* and *S. cruzi*, respectively, in cattle. The findings at this marker were corroborated by additional consistent differences between *S. buffalonis* and *S. hirsuta* at the partial 28S rRNA gene and between *S. levinei* and *S. cruzi* at both the 18S and 28S rRNA genes. In previous studies of these species pairs, only the 18S rRNA gene has been examined (Holmdahl et al. 1999; Yang et al.

Fig. 1 Phylogenetic tree for selected members of the *Sarcocystidae* based on partial sequences of *cox1* and inferred using the maximum parsimony method. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) is shown next to the branches. GenBank accession numbers for all sequences are given in Table S1 in the Supplementary material. Subtrees formed by two or more sequences/haplotypes of the same species have been collapsed, but the numbers of sequences/haplotypes included are given behind the taxon names. Sequences of *S. buffalonis* and *S. levinei* from the present study are in **boldface**

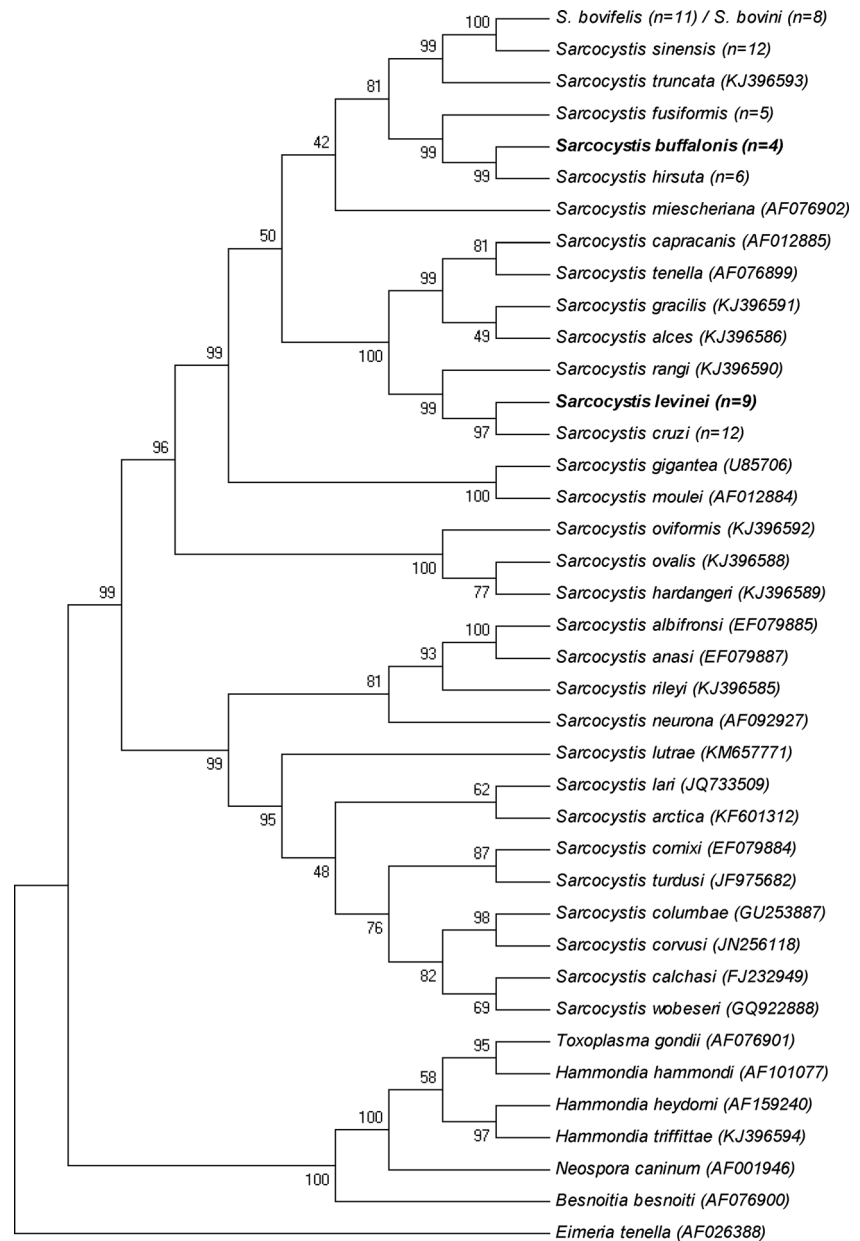


2001b, 2002; Li et al. 2002; Jehle et al. 2009; Xiang et al. 2011).

The species *S. buffalonis* was first described and named by Huong et al. (1997a) from water buffaloes in Vietnam. The authors recognized the morphological similarity (macroscopic sarcocysts and cyst wall ultrastructure) of this species to *S. hirsuta* in cattle as described by Böttner et al. (1987) but described *S. buffalonis* as a new species on the assumption that different *Sarcocystis* spp. were intermediate host specific. Huong et al. (1997a) did not mention the similarity in sarcocyst morphology as seen by TEM between *S. buffalonis* and the species *S. levinei* as described by Dissanaik and Kan (1978), even though they were aware of this resemblance as revealed in a subsequent paper (Huong et al. 1997a). Holmdahl et al. (1999) were the first to provide a

molecular comparison between *S. buffalonis* and *S. hirsuta*. They PCR amplified (in triplicate) and sequenced (from pooled amplicons) the complete 18S rRNA gene from a single sarcocyst isolate of both species. They found that the full-length gene of *S. buffalonis* (AF017121) differed from that of *S. hirsuta* (AF017122) “unambiguously in 13 nucleotide positions” (actually in 12 positions; 99.4 % identity) and stated that this difference “supports the characterization of *S. buffalonis* as a newly described species that infects water buffaloes.” Yang et al. (2001b), on the other hand, believed that the *S. hirsuta*-like species in water buffaloes should be considered the same as *S. hirsuta* in cattle after having sequenced and compared the partial 18S rRNA gene (~1550 bp) of two isolates from both hosts in China (AF176938 and AF176939 from cattle and AF176940 and

Fig. 2 Phylogenetic tree for selected members of the *Sarcocystidae* based on partial 28S rRNA gene sequences and inferred using the maximum parsimony method. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) is shown next to the branches. GenBank accession numbers for all sequences are given in Table S1 in the Supplementary material. Subtrees formed by multiple sequences of some species have been collapsed, but the numbers of sequences included are given behind the taxon names. Sequences of *S. buffalonis* and *S. levinei* from the present study are in **boldface**. Note that sequences of *S. bovis* and *S. bovis* are interleaved



AF176941 from water buffaloes). Yang et al. (2002) reached the same conclusion after having examined the partial 18S rRNA gene (~900 bp) by restriction fragment length polymorphism (RFLP). They compared nine *S. hirsuta*-like isolates from water buffaloes with three *S. hirsuta* isolates from cattle and found the same restriction digestion pattern in isolates from both hosts. Similarly, Jehle et al. (2009) compared the partial 18S rRNA gene (~900 bp) of these species both by RFLP and direct sequencing. They found that the two sequenced isolates from cattle were most similar (98 and 99 % identity) to previous GenBank sequences of *S. hirsuta* from cattle and that the two *S. hirsuta*-like isolates from water buffaloes were most similar (99 % identity) to three previous sequences of *S. buffalonis*/*S. hirsuta* from this host

(AF017121, AF176940, and AF176941). In addition, the isolates from both hosts shared the same restriction digestion pattern. Jehle et al. (2009) therefore presumed that the *S. hirsuta*-like species in water buffaloes was identical with *S. hirsuta* in cattle.

The present study has shown, however, that there was considerable sequence variation in the 5' end half of the 18S rRNA gene of *S. buffalonis*, which was also the case with the 18S rRNA gene of *S. hirsuta* (Gjerde 2015). Moreover, there was an overlap between the intraspecific and interspecific sequence divergence and no obvious consistent differences between the two species in any regions of the gene (Table S2 in the Supplementary material). Hence, *S. buffalonis* could not be reliably distinguished from

S. hirsuta on the basis of the 18S rRNA gene sequences. The same was largely true for the ITS1 region, although there was one possible consistent difference between sequences of the two species. Moreover, the phylogenetic analyses placed the ITS1 sequences of the two species into two separate clusters, whereas the 18S rRNA gene sequences were interleaved. At the 28S rRNA gene, on the other hand, there were several consistent differences between the two species in the available sequences, and hence, they could be separated from each other in spite of an overlap between the intraspecific and interspecific sequence divergence. Likewise, the phylogenetic analysis was placed with high-support sequences of the two species into two separate clusters (Fig. 2). Sequencing of additional isolates/clones of both species may, however, disclose more variability in the 28S rRNA gene of one or both species, and if so, this marker may be less suitable for delimitation of these species than suggested from this study. Hence, the *cox1* gene is the marker of choice for identifying the sarcocysts of *S. buffalonis* and *S. hirsuta* in water buffaloes and cattle and the oocysts/sporocysts of these species in their feline definitive hosts (Böttner et al. 1987; Huong et al. 1997a; Gjerde 2016).

The species *S. levinei* was first described and named by Dissanaïke and Kan (Dissanaïke and Kan 1978; Kan and Dissanaïke 1978) from water buffaloes in Malaysia. According to their description, the species had thin-walled microscopic sarcocysts and was transmitted by dogs. By TEM, however, the sarcocysts were reported to have broad sloping protrusions with irregular wavy outlines. Dissanaïke and Kan (1978) thought that the sarcocysts of *S. levinei* resembled the microscopic sarcocysts of *S. bovifelis* in cattle, which was then known as *S. hirsuta*, and were apparently not familiar with the species currently known as *S. hirsuta*, which has macroscopic sarcocysts (Gjerde 2015, 2016). As mentioned above, Huong et al. (1997a) later also found a *S. hirsuta*-like species in water buffaloes, which they named *S. buffalonis*. During this work, they discovered that the sarcocysts examined by TEM by Dissanaïke and Kan (1978) belonged to a different species than that described by light microscopy and found to be transmitted by dogs. Hence, in a subsequent paper, Huong et al. (1997b) redescribed *S. levinei* as a species that was ultrastructurally similar to *S. cruzi* in cattle, and they confirmed through an infection experiment that this species was transmitted by dogs. The ultrastructure of *S. levinei* sarcocysts was later further described by Claveria and Cruz (2000) from water buffaloes in the Philippines. In retrospect, Dissanaïke and Kan (1978) might have avoided the mistake of including two species in their original description of *S. levinei* if their feeding experiments had been performed in a different manner. Thus, they fed their experimental dogs either only isolated macroscopic sarcocysts of *S. fusiformis* (no sporocyst shedding), such sarcocysts together with esophageal muscle containing

microscopic sarcocysts (sporocyst shedding), or esophageal muscles containing only microscopic sarcocysts (sporocyst shedding). In contrast, the cats were fed only isolated macroscopic sarcocysts of *S. fusiformis* (sporocyst shedding). If some cats had been fed esophageal muscles without macroscopic sarcocysts, and this feeding had resulted in sporocyst shedding (due to *S. buffalonis*), these researchers might have realized that there were two species with much smaller sarcocysts than *S. fusiformis* in the material under study, although fully developed sarcocysts of *S. buffalonis* are macroscopic (Huong et al. 1997a).

The first molecular characterization of *S. levinei*-like sarcocysts from water buffaloes was by Yang et al. (2001b). They sequenced the partial 18S rRNA gene (~1550 bp) of two isolates of this species from water buffaloes (AF176932 and AF176935) and compared the sequences with those obtained from two isolates of *S. cruzi* from cattle (AF176933 and AF176934). They found a divergence of 0.2–0.6 % between sequences from the two hosts and concluded that those from water buffaloes also belonged to *S. cruzi*. However, a comparison of these sequences with those of *S. levinei* (from this study) and *S. cruzi* (Gjerde 2013, 2015), which were obtained from isolates that were also identified through the *cox1* marker, shows that both of their sequences from cattle are consistent with *S. cruzi* and that sequence AF176935 from water buffaloes is fully consistent with *S. levinei*. Sequence AF176932, on the other hand, is consistent with *S. levinei* at the first indel but with *S. cruzi* at the second variable region. Members of the same research group also compared 15 *S. cruzi*-like sarcocyst isolates from water buffaloes with 10 sarcocyst isolates of *S. cruzi* from cattle by the RFLP method in a subsequent study (Li et al. 2002). They found the same restriction digestion pattern in isolates from both hosts and again concluded that the water buffalo was a natural intermediate host of *S. cruzi*.

Xiang et al. (2009) used both RFLP and sequencing of the partial 18S rRNA gene (~850 bp) to identify sporocysts excreted by dogs fed *S. cruzi*-like sarcocysts from water buffaloes, as well as the sarcocysts themselves. They found the restriction digestion pattern and sequences to be consistent with those of *S. cruzi* from previous studies but described the species in water buffaloes only as *S. cruzi*-like. Jehle et al. (2009) also used both RFLP and sequencing of the partial 18S rRNA gene (~900 bp) in order to identify *S. cruzi*-like sarcocyst isolates from water buffaloes in Vietnam and compare them with isolates of *S. cruzi* in cattle. No sequences from this study were deposited in GenBank, but they found that five *S. cruzi*-like sequences were “nearly identical with the corresponding sequences of *S. cruzi* (AF176932: 99 % identity, AF176935: 100 %).” However, as stated above, these sequences were derived from water buffaloes (Yang et al. 2001b), and sequence AF176935 is fully consistent with sequences of *S. levinei* from the present study. Hence, Jehle et al.

(2009) actually showed that the *S. cruzi*-like sarcocysts from water buffaloes belonged to *S. levinei* rather than to *S. cruzi* as they claimed.

Xiang et al. (2011) compared *S. cruzi*-like sarcocysts from water buffaloes in China with *S. cruzi* from cattle, both by light and transmission electron microscopy and by sequencing of the partial 18S rRNA gene (750–800 bp). In addition, they fed such sarcocysts from each host to two dogs and compared the morphology of oocysts and sporocysts derived from those infections. Seven sequences were generated from *S. cruzi*-like sarcocysts (HM447190–HM447196) and compared with 12 sequences of *S. cruzi* (HM447197 and HM447199–HM447209). However, this comparison seems to have been based mainly on the relative phylogenetic position of the different sequences rather than on a careful comparison of individual sequences by BLAST or by eye in a multiple alignment. Moreover, for the phylogenetic analysis, which also included many other taxa, only ungapped positions in the multiple alignment were used, and hence, some of the differences (indels) between the sequences from water buffaloes and cattle were missed. In the resulting phylogenetic tree (Fig. 6 in their paper), sequences derived from water buffaloes were interleaved with sequences from cattle. The apparent lack of consistent differences between sequences from the two hosts, in combination with the closely similar sarcocyst and sporocyst morphology irrespective of host origin, led Xiang et al. (2011) to conclude that all the examined *S. cruzi*-like sarcocysts from water buffaloes belonged to *S. cruzi*. Still, the resulting sequences deposited in GenBank were designated *Sarcocystis* cf. *cruzi*.

A careful comparison of the sequences from the study by Xiang et al. (2011) with the new sequences of *S. levinei* from the present study shows, however, that all but one sequence from water buffaloes are consistent with *S. levinei*. The exception is sequence HM447195, which is fully consistent with *S. cruzi*, as are all sequences derived from cattle. Provided that sequence HM447195 indeed derives from a water buffalo, the study by Xiang et al. (2011) shows that this host may become infected by *S. cruzi* but is mainly infected by *S. levinei* rather than by *S. cruzi* as claimed by the authors. A similar conclusion may be drawn from the present study, as well as from the studies by Yang et al. (2001b) and Jehle et al. (2009). Likewise, a single sequence (KT306827; unpublished study) from a water buffalo in India attributed to *S. cruzi* is consistent with *S. levinei*. However, another sequence from cattle in India (KM434885; unpublished study) is fully consistent with the new sequences of *S. levinei*. This sequence has been designated *Sarcocystis* cf. *cruzi*, possibly because it is 100 % identical in the overlapping region with three of the six aforementioned putative sequences of *S. levinei* from the study by Xiang et al. (2011), which also bear this designation.

Thus, from these sequence comparisons, it seems that cattle may occasionally act as intermediate hosts for *S. levinei*

(KM434885), whereas water buffaloes may occasionally act as intermediate hosts for *S. cruzi* (HM447195). In addition, one sequence from a water buffalo (AF176932) and one sequence from cattle (KT306829) were identical to the new sequences of *S. levinei* at the first indel, but similar to *S. cruzi* in the second variable region, although different from each other. The identity of these sequences is uncertain, but they may also represent sequences of *S. levinei*. Anyway, the available data suggest that, although *S. cruzi* and *S. levinei* are not strictly intermediate host specific, they are better adapted to cattle and water buffaloes, respectively, than to the opposite host. Hence, each of these species may only cause low-level infections in its aberrant host and thus be difficult to detect by examining just a few sarcocysts. Xiang et al. (2011), on the other hand, interpreted the apparent successful cross-transmission experiments of cattle and water buffaloes with *S. cruzi*/*S. levinei* reported by Xiao et al. (1991, 1993) as being due to the presence of a single species (*S. cruzi*) infecting both hosts. By using the *cox1* marker for identification, it will now be possible to unambiguously determine if *S. cruzi* and *S. levinei* are able to infect water buffaloes and cattle, respectively, either through a direct identification of tissue stages in the intermediate hosts or through identification of intestinal oocysts/sporocysts in dogs fed tissues from infected intermediate hosts.

In summary, the present study has shown unequivocally through the use of the *cox1* marker that *S. buffalonis* and *S. levinei* in water buffaloes are distinct from *S. hirsuta* and *S. cruzi*, respectively, in cattle. These results corroborate the findings in previous studies (Gjerde 2013, 2014, 2015; Gjerde et al. 2015), showing that *cox1* is superior to the 18S rRNA gene as regards the ability to separate closely related (recently diverged) and morphologically indistinguishable *Sarcocystis* spp. in ruminants. Previous studies aimed at proving or disproving that the morphologically similar species in these hosts were conspecific (Holmdahl et al. 1999; Yang et al. 2001a, b, 2002; Li et al. 2002; Jehle et al. 2009; Xiang et al. 2009, 2011), erroneously assumed that the 18S rRNA gene would be sufficient for delimitation of various *Sarcocystis* spp., and hence made some wrong inferences from the comparisons of sequences derived from the two hosts. Although Holmdahl et al. (1999) probably were correct in assigning the two closely similar sequences from water buffalo and cattle to *S. buffalonis* and *S. hirsuta*, respectively, this was possible mainly because they happened to sequence two variants (one from each species) that differed slightly from each other. If they had sequenced and compared more variants of each species, they may have reached a different conclusion as did Yang et al. (2001b, 2002) and Jehle et al. (2009). Thus, this study has shown that *S. buffalonis* and *S. hirsuta* cannot be reliably distinguished from each other on the basis of 18S rRNA gene sequences. Likewise, *S. sinensis* could not be distinguished from *S. bovifelis* and *S. bovini* using this marker (Gjerde 2015).

As regards *S. levinei* and *S. cruzi*, on the other hand, a differentiation of the two species seems to be possible based on 18S rRNA gene sequences. Previous studies have failed to detect these differences (Yang et al. 2001b; Li et al. 2002; Jehle et al. 2009; Xiang et al. 2009, 2011), partly because they used an insensitive RFLP method; they did not sequence and compare the 3' end of the gene, comprising an important indel; sequence comparisons were based on phylogenetic placement rather than a direct comparison of the sequences themselves (Xiang et al. 2011); and the authors may have been confused by the presence of an isolate of *S. cruzi* in water buffaloes (Xiang et al. 2011). Thus, the latter study erroneously assumed that if *S. levinei* was different from *S. cruzi*, then these species would have to be host specific and be represented by different sequence types that would segregate according to host. Since no obvious segregation was observed due to a mixed infection of water buffaloes with both *S. levinei* and *S. cruzi*, they erroneously concluded that one and the same species occurred in both hosts. In the present study, on the other hand, the isolates were first identified through the *cox1* marker and then compared at the 18S rRNA gene, which allowed consistent differences between the two species to be detected. The fact that *S. levinei* and *S. cruzi* seem to be able to infect cattle and water buffaloes, respectively, means that *S. cruzi*-like sarcocysts in cattle cannot automatically be identified as sarcocysts of *S. cruzi* in areas where both hosts occur and natural cross transmissions via dogs are possible. The same is true for *S. levinei*-like sarcocysts in water buffaloes in such areas.

References

- Böttner A, Charleston WA, Hopcroft D (1987) The structure and identity of macroscopically visible *Sarcocystis* cysts in cattle. *Vet Parasitol* 24:35–45. doi:10.1016/0304-4017(87)90128-2
- Chen XW, Zuo YX, Hu JJ (2003) Experimental *Sarcocystis hominis* infection in a water buffalo (*Bubalus bubalis*). *J Parasitol* 89:393–394. doi:10.1645/0022-3395(2003)089[0393:ESHIA]2.0.CO;2
- Chen X, Zuo Y, Rosenthal BM, He Y, Cui L, Yang Z (2011) *Sarcocystis sinensis* is an ultrastructurally distinct parasite of water buffalo that can cause foodborne illness but cannot complete its life-cycle in human beings. *Vet Parasitol* 178:35–39. doi:10.1016/j.vetpar.2010.12.026
- Claveria FG, Cruz MJ (2000) *Sarcocystis levinei* infection in Philippine water buffaloes (*Bubalus bubalis*). *Parasitol Int* 48:243–247. doi:10.1016/S1383-5769(99)00025-2
- Dissanaike AS, Kan SP (1978) Studies on *Sarcocystis* in Malaysia. I. *Sarcocystis levinei* n. sp. from the water buffalo *Bubalus bubalis*. *Z Parasitenkd. Parasitol Res* 55:127–138
- Dubey JP, Lane EP, Wilpe E, Suleman E, Reininghaus B, Verma S, Rosenthal B, Mtshali M (2014) *Sarcocystis cafferi*, n. sp. (Protozoa:Apicomplexa) from the African buffalo (*Syncerus caffer*). *J Parasitol* 100:817–827. doi:10.1645/13-467.1
- Gjerde B (2012) Morphological and molecular characterization and phylogenetic placement of *Sarcocystis capreolicanis* and *Sarcocystis silva* n. sp. from roe deer (*Capreolus capreolus*) in Norway. *Parasitol Res* 110:1225–1237. doi:10.1007/s00436-011-2619-6
- Gjerde B (2013) Phylogenetic relationships among *Sarcocystis* species in cervids, cattle and sheep inferred from the mitochondrial cytochrome c oxidase subunit I gene. *Int J Parasitol* 43:579–591. doi:10.1016/j.ijpara.2013.02.004
- Gjerde B (2014) *Sarcocystis* species in red deer revisited: with a re-description of two known species as *Sarcocystis elongata* n. sp. and *Sarcocystis truncata* n. sp. based on mitochondrial *cox1* sequences. *Parasitology* 141:441–452. doi:10.1017/S0031182013001819
- Gjerde B (2015) Molecular characterisation of *Sarcocystis bovifelis*, *Sarcocystis bovini* n. sp., *Sarcocystis hirsuta* and *Sarcocystis cruzi* from cattle (*Bos taurus*) and *Sarcocystis sinensis* from water buffaloes (*Bubalus bubalis*). *Parasitol Res*. doi:10.1007/s00436-015-4881-5
- Gjerde B (2016) The resurrection of a species: *Sarcocystis bovifelis* Heydorn et al., 1975 is distinct from the current *Sarcocystis hirsuta* in cattle and morphologically indistinguishable from *Sarcocystis sinensis* in water buffaloes. *Parasitol Res* 115:1–21. doi:10.1007/s00436-015-4785-4
- Gjerde B, Hilali M, Mawgood SA (2015) Molecular characterisation of three regions of the nuclear ribosomal DNA unit and the mitochondrial *cox1* gene of *Sarcocystis fusiformis* from water buffaloes (*Bubalus bubalis*) in Egypt. *Parasitol Res* 114:3401–3413. doi:10.1007/s00436-015-4566-0
- Holmdahl OJ, Morrison DA, Ellis JT, Huang LTT (1999) Evolution of ruminant *Sarcocystis* (Sporozoa) parasites based on small subunit rDNA sequences. *Mol Phylogenet Evol* 11:27–37. doi:10.1006/mpev.1998.0556
- Huang LTT, Dubey JP, Nikkilä T, Uggla A (1997a) *Sarcocystis buffaloni* n. sp. (Protozoa:Sarcocystidae) from the water buffalo (*Bubalus bubalis*) in Vietnam. *J Parasitol* 83:471–474. doi:10.2307/3284413
- Huang LTT, Dubey JP, Uggla A (1997b) Redescription of *Sarcocystis levinei* Dissanaike and Kan, 1978 (Protozoa:Sarcocystidae) of the water buffalo (*Bubalus bubalis*). *J Parasitol* 83:1148–1152. doi:10.2307/3284375
- Jehle C, Dinkel A, Sander A, Morent M, Romig T, Luc PV, De TV, Thai VV, Mackenstedt U (2009) Diagnosis of *Sarcocystis* spp. in cattle (*Bos taurus*) and water buffalo (*Bubalus bubalis*) in northern Vietnam. *Vet Parasitol* 166:314–320. doi:10.1016/j.vetpar.2009.08.024
- Kan SP, Dissanaike AS (1978) Studies on *Sarcocystis* in Malaysia. II. Comparative ultrastructure of the cyst wall and zoites of *Sarcocystis levinei* and *Sarcocystis fusiformis* from the water buffalo *Bubalus bubalis*. *Z Parasitenkd (Parasitol Res)* 57:107–116
- Li QQ, Yang ZQ, Zuo YX, Attwood SW, Chen XW, Zhang YP (2002) A PCR-based RFLP analysis of *Sarcocystis cruzi* (Protozoa: Sarcocystidae) in Yunnan Province, PR China, reveals the water buffalo (*Bubalus bubalis*) as a natural intermediate host. *J Parasitol* 88:1259–1261. doi:10.1645/0022-3395(2002)088[1259:APBRAO]2.0.CO;2
- Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25:1451–1452. doi:10.1093/bioinformatics/btp187
- Moré G, Pantchev A, Skuballa J, Langenmayer MC, Maksimov P, Conraths FJ, Venturini MC, Schares G (2014) *Sarcocystis sinensis* is the most prevalent thick-walled *Sarcocystis* in beef for consumers in Germany. *Parasitol Res* 113:2223–2230. doi:10.1007/s00436-014-3877-x
- Ng YH, Fong MY, Subramaniam V, Shahari S, Lau YL (2015) Short communication: genetic variants of *Sarcocystis cruzi* in infected Malaysian cattle based on 18S rDNA. *Res Vet Sci* 103:201–204. doi:10.1016/j.rvsc.2015.10.009
- Parairo JR, Manuel MF, Icatlo FC Jr (1988) Ultrastructural studies on the cyst wall of *Sarcocystis* spp. in Philippine carabaos (*Bubalus bubalis*). *Phil J Vet Anim Sci* 14:40–54

- Rosenthal BM, Dunams DB, Pritt B (2008) Restricted genetic diversity in the ubiquitous cattle parasite, *Sarcocystis cruzi*. *Infect Genet Evol* 8: 588–592. doi:10.1016/j.meegid.2008.04.004
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 28:2731–2739. doi:10.1093/molbev/msr121
- Xiang Z, Chen X, Yang L, He Y, Jiang R, Rosenthal BM, Luan P, Attwood SW, Zuo Y, Zhang YP, Yang Z (2009) Non-invasive methods for identifying oocysts of *Sarcocystis* spp. from definitive hosts. *Parasitol Int* 58:293–296. doi:10.1016/j.parint.2009.03.004
- Xiang Z, He Y, Zhao H, Rosenthal BM, Dunams DB, Li X, Zuo Y, Feng G, Cui L, Yang Z (2011) *Sarcocystis cruzi*: comparative studies confirm natural infections of buffaloes. *Exp Parasitol* 127:460–466. doi:10.1016/j.exppara.2010.10.012
- Xiao BN, Zhang CG, Gong ZF, Wang M (1991) Transmission studies of *Sarcocystis cruzi* between cattle and buffalo. *Chin J Zool* 26:1–3 (in Chinese)
- Xiao BN, Zeng DN, Zhang CG, Wang M, Li Y (1993) Development of *Sarcocystis cruzi* in buffalo (*Bubalus bubalis*) and cattle (*Bos taurus*). *Acta Vet Zootech Sin* 24:185–192 (in Chinese)
- Yang ZQ, Zuo YX, Ding B, Chen XW, Luo J, Zhang YP (2001a) Identification of *Sarcocystis hominis*-like (Protozoa:Sarcocystidae) cyst in water buffalo (*Bubalus bubalis*) based on 18S rRNA gene sequences. *J Parasitol* 87:934–937. doi:10.1645/0022-3395(2001)087[0934:IOSHLP]2.0.CO;2
- Yang ZQ, Zuo YX, Yao YG, Chen XW, Yang GC, Zhang YP (2001b) Analysis of the 18S rRNA genes of *Sarcocystis* species suggests that the morphologically similar organisms from cattle and water buffalo should be considered the same species. *Mol Biochem Parasitol* 115: 283–288. doi:10.1016/S0166-6851(01)00283-3
- Yang ZQ, Li QQ, Zuo YX, Chen XW, Chen YJ, Nie L, Wei CG, Zen JS, Attwood SW, Zhang XZ, Zhang YP (2002) Characterization of *Sarcocystis* species in domestic animals using a PCR-RFLP analysis of variation in the 18S rRNA gene: a cost-effective and simple technique for routine species identification. *Exp Parasitol* 102: 212–217. doi:10.1016/S0014-4894(03)00033-X