



Morphological and molecular phylogenetic characterization of *Sarcocystis kani* sp. nov. and other novel, closely related *Sarcocystis* spp. infecting small mammals and colubrid snakes in Asia

Thomas Jäkel^{a,c,1,*}, Lisa Raisch^a, Sarah Richter^a, Mareike Wirth^a, Damaris Birenbaum^a, Sulaiman Ginting^b, Yuvaluk Khoprasert^c, Ute Mackenstedt^a, Marion Wassermann^{a,d}

^a University of Hohenheim, Institute of Biology, Department of Parasitology, Stuttgart, Germany

^b Islamic University of North Sumatra, Medan, Indonesia

^c Department of Agriculture, Plant Protection Research and Development Office, Bangkok, Thailand

^d University of Hohenheim, Center of Biodiversity and Integrative Taxonomy, Stuttgart, Germany

ARTICLE INFO

Keywords:
Sarcocystis
 Colubrid snakes
 Phylogeny
 18S rRNA
 cox1
 Protein structure

ABSTRACT

We investigated the morphology and phylogenetic relationships of novel and previously recognized *Sarcocystis* spp. infecting small mammals and colubrid snakes in Asia. The nuclear 18S rRNA and mitochondrial *cox1* of *Sarcocystis* sp.1 from mangrove snakes (*Boiga dendrophila*) in Thailand and *Sarcocystis* sp.2 from a ricefield rat (*Rattus argentiventer*) in Sumatra were partially sequenced. Sporocysts of *Sarcocystis* sp.1 induced development of sarcocysts in experimentally infected rats, which showed a unique ultrastructure that was observed previously by S.P. Kan in rats from Malaysia; therefore, we describe this species as *Sarcocystis kani* sp. nov. Its integration into the 18S rRNA phylogeny of *Sarcocystis* spp. cycling between small mammals and colubrid snakes helped clarify relationships among the so-called *S. zuoi*-complex of molecularly cryptic species: *Sarcocystis kani* sp. nov., *S. sp.2*, *S. attenuati*, *S. scandentiborneensis*, and *S. zuoi* were all included in this clade. Tree topology was resolved into dichotomies congruent with the morphological disparities between the taxa. However, *cox1* gene sequencing (including newly sequenced *S. singaporensis* and *S. zamani*) revealed that *Sarcocystis kani*, *S. attenuati*, and *S. scandentiborneensis* were identical suggesting a recent, common ancestry. To identify other distinctive features, lineage-specific molecular patterns within both genes were examined revealing that all 18S rRNA sequences of the *S. zuoi* – complex possess a unique, 7-nt long motif in helix 38 of domain V7 that was different in *S. clethrionomyelaphis* which branched off basally from the complex. Three-dimensional homology modelling of COX1 protein structure identified amino acid substitutions within the barcode area specific for the *S. zuoi*-complex and substantial divergence in structurally important amino acids between *Sarcocystis* species of snakes as definitive hosts and other lineages of the Sarcocystidae. We discuss the utility of selected genes for species delimitation of the *Sarcocystis* spp. under investigation, which probably evolved during recent radiations of their intermediate and definitive hosts.

1. Introduction

The genus *Sarcocystis* comprises a group of more than 200 species of protozoan parasites that infect mammals, reptiles, and birds. *Sarcocystis* spp. usually require two hosts to complete their life cycle (Odening, 1998; Dubey et al., 2016). Although mammals are often regarded the predominant hosts, this may be attributable to a higher research interest in *Sarcocystis* spp. of medical and veterinary importance rather than the

actual situation in nature. In fact, while 74% of the intermediate hosts and 27% of the definitive hosts of *Sarcocystis* species known by 2012 were mammals, definitive hosts of 56% of recognized *Sarcocystis* spp. were unknown (Prakas and Butkauskas, 2012). Larger-scale wildlife surveys have revealed that reptiles and birds may carry the bigger share of extant *Sarcocystis* spp. (Munday et al., 1978, 1979; Upton et al., 1992; McAllister et al., 1995). Furthermore, if one takes into consideration the eimeriid coccidia with about 2000 known species (Upton, 2000),

* Corresponding author. University of Hohenheim, Institute of Biology, Department of Parasitology, Stuttgart, Germany.

E-mail address: thom.jaekel@t-online.de (T. Jäkel).

¹ Current address: Rice Department 50, Phaholyothin Road, Ladyao, Khet Chatuchak, 10900 Bangkok, Thailand.

<https://doi.org/10.1016/j.ijppaw.2023.10.005>

Received 7 June 2023; Received in revised form 24 August 2023; Accepted 9 October 2023

Available online 13 October 2023

2213-2244/© 2023 The Authors. Published by Elsevier Ltd on behalf of Australian Society for Parasitology. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

reptiles are the vertebrates with the most described coccidian species (Megia-Palma et al., 2015).

Sarcocystis spp. with a snake-rodent life cycle are among the most prevalent Apicomplexan parasites found in rodents in the Indo-Australian Archipelago (O'Donoghue et al., 1987; Jäkel et al., 1997; Paperna et al., 2004), also showing prominent presence in environmental samples (Lee, 2019). However, despite increased molecular analysis of new taxa proper identification of *Sarcocystis* spp. remains difficult when availability of morphological or biological data is limited or non-existent. A case in point is *S. zuoi*, which was originally described from China where it develops in the colubrid snake *Elaphe carinata* as definitive host and rats (order Rodentia) as intermediate hosts (Hu et al., 2012). Subsequently, other researchers collected partial 18S rRNA gene sequences with high similarity to *S. zuoi* but without corresponding morphological data, assuming a wide distribution range for this species including Thailand, Malaysia, China, and Japan (Lau et al., 2013; Abe et al., 2015; Watthanakaiwan et al., 2017). However, we previously discovered a new species (*S. scandentiborneensis*) in treeshrews (Scandentia) that is very similar to *S. zuoi* in its nuclear 18S rRNA gene but morphologically distinct from this species (Ortega Pérez et al., 2020); the same applies to *S. attenuati*, another recently discovered species in shrews (Eulipotyphla) (Hu et al., 2022). This suggested that species diversity might be higher and intermediate host range wider than previously assumed among what we call the *S. zuoi* - complex (so named due to high sequence but not morphological similarity). As known so far, this group of parasites commonly infects rat snakes of the genus *Elaphe* and related genera which are widely distributed in Asia (Uetz et al., 2023). We have shown previously that there are at least two distinct lineages of *Sarcocystis* spp. developing in snakes as definitive hosts in the phylogenetic tree of the 18S rRNA gene, which were termed S1 and S2 (Wassermann et al., 2017). Various *Sarcocystis* spp. from pythonid and colubrid snakes in Asia cluster within lineage S1 including all known members of the *S. zuoi* - complex (Ortega Pérez et al., 2020).

Here, we examined two new *Sarcocystis*-isolates infecting colubrid snakes and rodents in Southeast Asia, also applying new methodological approaches. We describe a new *Sarcocystis* species from mangrove snakes (*Boiga dendrophila*) in Thailand and named it after the investigator who first observed the parasite in rodents in Malaysia (Kan, 1979), embedding this into a broader examination of the morphological and molecular (nuclear 18S rRNA and mitochondrial *cox1*) characteristics of established *Sarcocystis* spp. and unidentified isolates that cluster in the *S. zuoi* - complex. Because both selected genes play important, functional roles in the tree of life and are of major phylogenetic relevance (e. g., Cavalier-Smith, 2014; Pentinsaari et al., 2016), we additionally studied lineage-specific patterns of 18S rRNA secondary structure and COX1 protein sequence variability among the *Sarcocystis* spp. of interest, aiming to include a functional perspective in phylogenetic reconstruction and extending analysis beyond a purely taxonomic assessment of 'markers' for species delimitation.

2. Materials and methods

Ethical statement

Animal experimentation was performed at the Agricultural Zoology Research Group, Department of Agriculture, Ministry of Agriculture and Cooperatives, Thailand, according to the 'Ethical Principles and Guidelines for the Use of Animals' as published by the National Research Council of Thailand (Anonymous, 1999) and the 'Guide for the Care and Use of Laboratory Animals' of the National Academy of Sciences USA (Clark et al., 1997). Permission was granted under the bilateral collaboration agreement between the Department of Agriculture and German International Cooperation (GIZ) (Project No. 2002.2156.4). Animals were obtained from the National Laboratory Animal Center of Mahidol University in Salaya, Nakhonpathom Province, Thailand, which is accredited internationally and maintains a quality-controlled rodent

breeding facility.

2.1. Origin of parasite isolates, tissue samples, and purification of sporocysts

Sarcocystis oocysts/sporocysts were recovered from the intestines of two road-killed adult mangrove snakes, *Boiga dendrophila*, which had been collected near Khao Sok National Park in southern Thailand. The snake carcasses were transported to the laboratory on ice, where their intestines were examined microscopically by wet smears. Intestinal scrapings containing sporocysts were further purified on a discontinuous sucrose density gradient to reduce bacterial contamination. For this, Sheather's basic sucrose solution was prepared by dissolving 500 g sucrose in 320 ml tap water and 6 ml formalin (37%). In a 50 ml Falcon centrifuge tube, 15 ml of Sheather's solution A (basic sucrose solution diluted 1:4 with 0.1 M PBS containing 1% Tween 80 source) was layered gently over 15 ml of Sheather's solution B (basic sucrose solution diluted 1:2 with 0.1 M PBS and 1% Tween 80). About 10 ml of a sporocyst suspension was placed on top of the two sucrose layers and the Falcon tube centrifuged at 1200 g for 30 min using a swing-out rotor without brake. Highly purified sporocysts were recovered from the upper layers of the gradient and washed twice in distilled water by centrifugation at 1200g for 10 min. The pellet was re-suspended in 1 ml distilled water and stored at 4 °C until further use. The same procedure was used to isolate sporocysts of *S. singaporensis* for DNA extraction and amplification of the *cox1* gene. These sporocysts originated from a previous study and were collected from a wild-caught reticulated python in Bangkok (Wassermann et al., 2017). The DNA used for PCR amplifications of the *cox1* sequences was the same as that used for amplification of the 18S rRNA gene reported previously (KY513624).

A wild-caught, adult ricefield rat (*Rattus argentiventer*), which had been kill-trapped by a farmer in a ricefield near the city of Medan (not for the purpose of this study), northern Sumatra, was found to contain sarcocysts of *Sarcocystis zamani* and another unknown *Sarcocystis* sp., here designated as *Sarcocystis* sp.2. Single sarcocysts of macroscopic *S. zamani* and fresh tissue pieces containing sarcocysts of *Sarcocystis* sp.2 were stored in 70% ethanol in phosphate-buffered saline (PBS) for preservation of DNA.

2.2. Experimental animals and husbandry

All experiments involving transmission of *Sarcocystis* spp. to a rodent host were performed using out-bred Sprague-Dawley rats, which were obtained from the National Laboratory Animal Center of Mahidol University in Thailand. The experimental rats were kept under coccidia-free conditions in isolated and climate-controlled rooms of the Agricultural Zoology Research Group, at the Department of Agriculture in Bangkok, Thailand. The animals were housed two to a cage (Makrolon polycarbonate: 60 cm × 38 cm × 20 cm) under conditions of controlled temperature and humidity and exposed to a 12:12-h light/dark cycle. They received a standard diet and water *ad libitum*.

2.3. Infection experiments and processing of tissue samples for light and electron microscopy

Sporocysts isolated from two *Boiga dendrophila* were used to inoculate 12 male rats, whereby six rats were used for each isolate. Another six rats were inoculated with PBS and kept in the same room serving as negative controls. The animals were inoculated with about 1000 purified sporocysts each using a stomach tube. Afterwards, the animals were monitored daily. Two rats each were euthanized after 3.5, 4, and 14 months with CO₂ and necropsied. At earlier examination intervals the animals had attained a body weight of 330–350 g; after 14 months, they weighed between 480 and 510 g.

We examined fresh muscle tissue from esophagus, tongue, diaphragm, muscles of the abdominal wall, femoral muscles, and masseter

for presence of sarcocysts under a light microscope. The size of freshly released cystozoites was determined after rupturing individual sarcocysts in phosphate-buffered saline at room temperature. Samples of muscle tissue were fixed and processed for electron microscopy as described previously (Jäkel, 1995).

2.4. Light microscopic measurement and statistical analysis of mean sizes of parasite stages (sporocysts, cystozoites)

A stage micrometer was used to calibrate the divisions on an eyepiece graticule of a light microscope. Live sporocysts and cystozoites of the *Sarcocystis* spp. isolated in this study were measured at 400x and 1000x magnification. The corresponding data on other species were obtained from published sources as indicated below. Statistical testing and comparison of mean sizes (length and diameter) of the transmission stages of different *Sarcocystis* spp. was conducted by One-Way Analysis of Variance (ANOVA) using Sigma Plot (Systat Software, Inc. USA, version 12.0), whereby the calculations compared mean values including the standard deviation (SD) and the number of observations. The test procedure assumed that the data underlying the means were normally distributed; this was confirmed for all raw measurements of this study at the indicated sample sizes. Because some of the referenced sources did not report SD, we approximated SD through a given range with the formula: $(\max - \min)/4$ (Wan et al., 2014).

2.5. DNA preparation from sarcocysts in muscle tissue and oocysts/ sporocysts

DNA from a single, macroscopic sarcocyst of *S. zamani* and from tissue pieces containing several sarcocysts of *Sarcocystis* sp.2 was extracted via Phenol-Chloroform extraction and ethanol precipitation: About $3 \times 3 \times 3$ mm tissue pieces were further cut into smaller pieces and placed into a 1.5 ml tube. To 500 µl lysis buffer (10 mM HCl pH 7.5; 10 mM EDTA pH 8.0; 50 mM NaCl; 2% SDS pH 7.5), 60 µl proteinase K (~20 mg/ml) and 10 µl DTT were added, and incubated for 3–4 h at 56 °C until the tissue was dissolved. Then 570 µl Phenol-Chloroform-Isoamyl alcohol (25:24:1) was added, mixed, and centrifuged for 10 min at 16,000 g. Fresh 2 ml tubes were prepared, and 1000 µl of 100% ethanol (–20 °C cold) and 50 µl of 3 M sodium acetate was added. The upper DNA containing aqueous phase from the previous centrifugation step was carefully transferred into the prepared tubes, mixed carefully, and kept at –20 °C over night (ethanol precipitation). Then, the tubes were centrifuged 30 min at 4 °C at 16,000 g, ethanol was carefully discarded, and 1 ml 70% ethanol added and centrifuged for 5 min at 16,000 g. The ethanol was discarded and the pellet dried for 30 min at 45 °C. The dried DNA pellet was dissolved in 50 µl nuclease free water and stored at –20 °C until use.

A small aliquot of the solution containing single oocysts/sporocysts were transferred to a Petri dish and diluted with water. The solution was screened for the presence of oocysts/sporocysts under an inverted microscope. If multiple oocysts/sporocysts were present in the field of view, the solution was further diluted so that oocysts/sporocysts were clearly separated and only individual oocysts/sporocysts were present in the aspiration radius of the pipette. Detected single oocysts/sporocysts were then aspirated with a pipette in a volume of 1 µl under an inverted microscope, transferred to 0.2 ml reaction tubes containing 5 µl of 0.02 M NaOH, and subsequently lysed by 95 °C for 15 min. The lysate was used directly as template for the following PCRs. For each sample, 30 sporocysts were randomly chosen and subjected to PCR. For those samples showing no or only few positive reactions, the number of single sporocysts for PCR analysis was increased up to 100.

2.6. PCR and sequencing of the 18S rRNA and *cox1* genes

2.6.1. Nuclear 18S rRNA gene

Because single sporocysts only contain minute amounts of DNA, we

chose a nested PCR approach. Since amplification of a long gene fragment in one piece was rarely successful, the 18S rRNA gene was amplified in two overlapping fragments; the first covering the front part (5' end) of the gene with an approximate length of 1200 bp, and the second the back part (3' end) with 900 bp in length. For each fragment, we used two pairs of primers, one for the first and the other for the following nested PCR. The primer sequences have been published previously (Wassermann et al., 2017). Accordingly, the 18S rRNA gene was amplified in two overlapping fragments. The first PCR step was performed in a volume of 25 µl, containing 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 2 mM MgCl₂, 200 µM of each dNTP, 6.25 pmol of each external primer, 0.625 U Ampli-Taq Polymerase (Applied Biosystems), and 1.5 µl of the sporocyst lysate. For the nested PCRs (the second step) volumes were increased to 50 µl, containing 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 2 mM MgCl₂, 200 µM of each dNTP, 12.5 pmol of each internal primer and 1.25 U Ampli-Taq Polymerase and 3 µl of the first PCR product. For the amplification of the DNA extracted from sarcocysts, only the second PCR was performed. The volume of reaction and concentration of reagents were identical as described above and 3 µl of DNA was added as template.

All amplification reactions started with an initial denaturation at 95 °C for 5 min, followed by 35 cycles of denaturation for 30 s at 95 °C, annealing for 30 s at 50 °C, elongation for 60 s at 72 °C, and a final elongation step for 5 min at 72 °C. PCR products were purified with the 'High Pure PCR Product Purification Kit' (Roche, Germany) according to the manufacturer's instructions, and sequenced according to the Sanger method by Eurofins GATC Biotech AG (Konstanz, Germany). To join the two fragments, the sequences were aligned, the overlapping region identified and concatenated to a single sequence. In case of sporocysts, the two fragments had to be obtained from a single oocyst/sporocyst. All concatenated sequences identified in this study were regarded as valid when at least two concatenated sequences from different sporocysts showed 100% identity.

2.6.2. Mitochondrial cytochrome c oxidase subunit 1 gene (*cox1*)

For PCR amplification of the *cox1* gene, the following primers were used (Gjerde, 2013b): First PCR, forward (SF1: ATGGCGTACAA-CAATCATAAAGAA), reverse (CO1Rm: CCCAGAGATAATCAAAAATGGAA); second, semi-nested PCR, forward (SF1: ATGGCGTACAACAATCATAAAGAA), reverse (SR5: TAGGTATCATG-TAACGCAATATCCAT). There were two sequential PCR reactions: an initial amplification using the primers as indicated above, followed by a semi-nested PCR reaction. For all reactions, the concentrations of the used ingredients for the PCR and the conditions during amplification were as described for the 18S rRNA gene amplification. For the amplification of the DNA obtained from the sarcocysts, again only the second/semi-nested PCR was performed. After purification of the PCR products, they were sequenced by Eurofins GATC Biotech GmbH (Konstanz, Germany).

2.7. Sequence alignment and molecular phylogenetic analyses

The 18S rRNA gene sequences of the *Sarcocystis* isolates under investigation were aligned with published Apicomplexan species available at GenBank (Clark et al., 2016) using the multiple sequence alignment algorithm of the 'R-Coffee' web server, which considers predicted secondary structures of RNA (Tommaso et al., 2011). We repeated the alignment procedure several times, selecting three alignments with the smallest proportion of ambiguous positions. Conserved areas were inspected visually for potential misalignments. The alignments comprised 43 sequences and 1949 to 2120 characters, including the new *Sarcocystis* isolates from a mangrove snake in Thailand and a ricefield rat in Sumatra and relevant sequences of closely related taxa, which were selected because they showed high similarity scores with the new sequences after initial screening by BLAST (Johnson et al., 2008): *S. cf. Zuoi* from a rat snake in Malaysia (KC878488, KC878488; Lau et al.,

2013); *S. scandentiborneensis* from treeshrews in Borneo (MN733816, MN733817); *S. attenuati* from a shrew and colubrid snake (*Elaphe taeniura*) in China, respectively (MZ826982, MZ826985); four sequences of *S. cf. zuoi* isolated from *Rattus tiomanicus* in Thailand (KU341118-21); *S. cf. zuoi* isolated from *Elaphe taeniura* in Japan (LC054267); the sequences attached to the original description of *S. zuoi* from China (JQ029112, JQ029113), and an isolate of *S. clethrionomyelaphis* from *Elaphe taeniura* in China (KP057504). Representative sequences of two species that parasitize pythonid snakes and are basally situated in this *Sarcocystis*-lineage S1 formed by the above species, *S. singaporensis* (KY513624, AF434057) and *S. zamani* (KU244524), were also included. To account for intraspecific variability of *S. zamani* and complement a corresponding *cox1* sequence isolated from the same ricefield rat from Sumatra mentioned above, we amplified an 18S rRNA sequence from the same sarcocyst and host. To account for members of the second *Sarcocystis*-lineage S2 from snakes (Wassermann et al., 2017), we selected *S. pantherophisi* (KU891600, isolated from an Eastern rat snake in North America by Verma et al., 2017), two *Sarcocystis* sequences from viperid definitive hosts in Africa (*S. atheridis*, AF120114; *Sarcocystis ex-Pseudocerastes*, AF513491), and the Indo-Australian *S. nesbitti* from a pythonid snake and human sample, respectively (KY513626, HF544324). Carnivore-ruminant *Sarcocystis* spp. were represented by *S. cruzi* (AF017120), *S. tenella*, (L24383), *S. gigantea* (L24384), and *S. hirsuta* (KT901165), the bird-ruminant group by *S. ovalis* (LC184602). *Sarcocystis* spp. that form a monophyletic clade with *S. neurona* included *S. speeri* (KT207459; sister species of *S. neurona*), *S. cymruensis* (MT372785), *S. halioti* (MH130211), *S. rileyi* (KJ396583), and *S. strixi* (MF162315). Sequences of the Toxoplasmatinae included *Cystoisospora canis* (KT184368), *Hammondia heydorni* (KT184370), *Neospora caninum* (A24380), and *Toxoplasma gondii* (M97703). Eimeriid species of fish (*Goussia janae*, AY043206), amphibian (*Eimeria ranae*, EU717219), reptilian (*Caryospora bigenetica*, AF060976), and avian (*E. tenella*, U67121) hosts served as outgroup. If possible, we selected sequences of the same taxa for the analysis of the *cox1* gene tree (below).

The 18S rRNA phylogenetic tree was reconstructed using Bayesian Inference (BI), executed from within MrBayes Version 3.2 (Ronquist et al., 2012). While alignments were around 2000 positions long, sites with gaps were excluded from analysis. Because some of the *Sarcocystis* sequences were relatively short, ‘unique site patterns’ identified by MrBayes ranged between 686 and 740 positions. We selected the general time-reversible substitution model with discrete gamma distribution and invariable sites (GTR + G + I; nst = 6 gamma categories) in addition to an assumed among-site rate variation between different lineages (covarion = yes; the so-called ‘covarion’ model). This substitution model has been found appropriate for the coccidia (Morrison et al., 2004). Bayesian Inference was computed in three replicate runs: all sequences were subjected to 1,000,000 or more Markov Chain Monte Carlo generations (heated chains = 4, chain temperature = 0.2, unconstrained branch length with exponential = 10) with a sampling frequency of 1000; burn-in length was set at 10%. Completion of an analysis was regarded successful, if the average standard deviation of split frequencies remained below $P = 0.01$ and the convergence diagnostic (potential scale reduction factor, PSFR) indicated a good sample of the posterior probability distribution (all values close to 1.0). Resulting trees were visualized using the program FigTree (Version 1.4.2, by Andrew Rambaut, Institute of Evolutionary Biology, University of Edinburgh; <http://tree.bio.ed.ac.uk/>). Using the same sequences and alignment, we also performed a Maximum Likelihood (ML) analysis, which was restricted to 1271 characters to accommodate for the shorter sequences in the alignment. The appropriate model for nucleotide substitution rates was determined by ML analysis as implemented in MEGA-7 (Kumar et al., 2016) using the model with the lowest BIC (Bayesian Information Criterion) and AICc (Akaike Information Criterion, corrected) values: the Hasegawa-Kishino-Yano model with a discrete gamma distribution rate variation among sites and accounting for a proportion of invariant sites (HKY + G + I) over five Gamma

categories. Site coverage cut-off was 90%. The number of bootstrap replications was set to 1000 and tree reconstructions repeated three times. Pairwise sequence analysis of transition/transversion ratios (Ti/Tv) was executed using the statistics function of the MEGA-7 software package and interpretation of results followed a rationale as described previously (Ortega Pérez et al., 2020). Tajima’s relative rate test (Tajima, 1993) was carried out with selected sequences to test the hypothesis that a pair of sequences has undergone different rates of evolution.

Sequences of mitochondrial *cox1* of *Sarcocystis* spp. from snakes as definitive hosts are still limited. Therefore, we sequenced the *cox1* genes of *S. singaporensis* and *S. zamani* in addition to the two new species associated with colubrid snakes. All sequence alignments were executed within MEGA-7 using MUSCLE. We checked translation of nucleotides into amino acids (aa) under different open reading frames with the EXPASY online tool of the Swiss Institute of Bioinformatics (<http://web.expasy.org/translate/>; last accessed December 22, 2022). Phylogenetic analysis of *cox1* was based on nucleotides of 25 sequences, whereby analysis was restricted to 603 characters (site coverage cut-off: 98–99%), which is slightly shorter than the barcode area of *cox1* comprising 219 aa or 657 bp (Pentinsaari et al., 2016). We selected *S. speeri*, KT207461, as substitution for *S. neurona* due to lack of sequences of appropriate length for the latter species. Sequences representing snakehost lineage S1 included *S. attenuati* (MZ889672), *S. scandentiborneensis* (MN732561), and the four new *cox1* sequences of this study (*Sarcocystis* sp.1 and sp.2, *S. singaporensis*, *S. zamani*); the only sequence available for snakehost lineage S2 was *S. pantherophisi* (KU891603). For ML analysis, the Hasegawa-Kishino-Yano substitution model with a discrete gamma distribution rate variation among sites and accounting for a proportion of invariant sites (HKY + G + I) over four Gamma categories was used; the HKY + G + I model showed the lowest BIC and AICc values in model selection tests. Tree inference options were set to Nearest-Neighbor-Interchange (NNI) as heuristic method and an initial tree was constructed by Neighbor-Joining (NJ). Bootstrapping was set to 1000 replicates.

2.8. Secondary structure modelling of the 18S rRNA

The small subunit (SSU) 18S rRNA gene sequences were aligned and their secondary structure modeled using SSU-ALIGN (version 0.1.1, February 2016, eddylab.org/software/ssu-align, Eric P. Nawrocki; Howard Hughes Medical Institute), which uses models of the primary sequence and secondary structure conservation of 18S rRNA as inferred by comparative sequence analyses and confirmed by crystal structure determination from the Comparative RNA Website (<https://crw-site.chemistry.gatech.edu/>; last accessed December 12, 2022) (Cannone et al., 2002). As input dataset of unaligned target SSU sequences we included *Sarcocystis* sp.1, *S. sp.2*, *S. scandentiborneensis* (MN733816), *S. attenuati* (MZ826985), *S. zuoi* (JQ029112), *S. clethrionomyelaphis* (KP057504), *S. pantherophisi* (KU891600), *S. nesbitti* (KY513626), *S. atheridis* (AF120114), and *Toxoplasma gondii* (M97703). Then, structure-based, multiple sequence alignments were created using so-called covariance models (Nawrocki et al., 2009). Both primary sequence conservation and secondary structure conservation between the target and the model (here: 1881 nt-long, structural template of the Eukarya) contributed to an alignment score. Finally, a graphical representation of each individual target sequence and its predicted secondary structure as it aligned to the default Eukarya model was created including a scoring statistic of the alignment. The numbering of nucleotides and nomenclature of helices followed the respective template. The program was executed on a Linux-based operating system using a command line interface. Target sequence files were in the recommended FASTA-format.

For 18S rRNA motif search and comparison, we checked the taxonomic framework of the SILVA RNA Database (version 138 released December 2019 with over 9,400,000 18S rRNA sequences; Yilmaz et al.,

2014) for species-specificity of 12–13 nt-long sequence fragments that included the 7-nt motif that had been suggested previously as specific for the *S. zuoi* - group (Ortega Pérez et al., 2020). Additionally, we conducted sequence comparisons with the reference sequence and predicted secondary structure of coccidia (*Toxoplasma gondii*, X75453) as published on the Comparative RNA Website.

2.9. Analysis of amino acid substitutions in the barcode area and homology modelling of COX1 protein structure

Amino acid substitutions in the barcode area (six helices) between chemically/structurally different aa groups of different lineages of tissue cyst-forming coccidia including the *Sarcocystis* spp. under investigation were identified using the complete (Namasivayam et al., 2021) and partial (KM657810) COX1 protein sequences of *Toxoplasma gondii*. To this end, aa were categorized into the following chemical standard groups: nonpolar aliphatic (G, A, V, L, M, I), polar uncharged (S, T, C, P, N, Q), aromatic (F, Y, W), positively charged (K, R, H), and negatively charged (D, E). Sites were considered non-variable, if a given position showed variation only among aa within these groups. Initially, we checked by assembly and comparison of the published sequence blocks that the partial *T. gondii* sequence KM657810 (Gjerde and Josefsen, 2015) was identical with the corresponding areas of the full protein sequence (Namasivayam et al., 2021). In contrast to the barcode area, which consists of the first six helices and interspaced loops (Pentinsaari et al., 2016) and has been described for *S. scandentiborneensis* (Ortega Pérez et al., 2020), the complete COX1 molecule comprises 12 transmembrane helices (Timón-Gómez et al., 2018); the full coding sequence (CDS) of *T. gondii* is 496 aa long. To demarcate the boundaries between the barcode helices and interspaced loops and identify putative ligand binding sites, we subjected the *T. gondii* protein sequences to three-dimensional homology modelling of protein structure as described in the following. Using SWISS-MODEL, a fully automated protein structure homology-modelling server (Waterhouse et al., 2018) accessible via the Swiss Bioinformatics Resource Portal EXPASY (<http://www.expasy.org>; last accessed on February 11, 2023), we determined the secondary and tertiary protein structure of the COX1 protein sequence of *T. gondii*. Modelling started with the full protein sequence of *Toxoplasma*, whereby SWISS-MODEL automatically selected a suitable template for sequence comparison from a pool of 20–50 COX1 templates: the *Saccharomyces cerevisiae* III2-IV2 mitochondrial respiratory supercomplex (Hartley et al., 2019; accession number 6hu9.1). For the partial *T. gondii* sequence, bovine heart Cytochrome c oxidase at the fully oxidized state was used as template (200-s X-ray exposure dataset by Aoyama et al., 2009; accession number 3 abm.1). We also included the partial sequence of *S. scandentiborneensis* (MN732561) in the modelling process, for which the same *Saccharomyces cerevisiae* template as above was selected. Models were built based on the target-template alignment using ProMod3 (Studer et al., 2021). Insertions and deletions were re-modeled using a fragment library; side chains were then rebuilt. The geometry of the resulting model was regularized by using a force field. The global and per-residue model quality was assessed using the QMEAN scoring function (Studer et al., 2021). Ligands present in the template structure were transferred by homology to the model only if certain criteria could be met, including that those residues in contact with the ligand were conserved between the target and the template. The quaternary structure annotation of the template was used to model the target sequence in its oligomeric form; the method (Bertoni et al., 2017) is based on a supervised machine learning algorithm, Support Vector Machines. This complements the GMQE score, which estimates the accuracy of the tertiary structure of the resulting model (range 0–1, higher number indicating higher reliability). In summary, the analyses of the *T. gondii* COX1 protein helped to determine the putative ligand binding sites of the *Sarcocystis* species under investigation.

2.10. Sequences deposited

GenBank: partial nuclear 18S rRNA gene sequences of *Sarcocystis* sp.1 (ON979685), *Sarcocystis* sp.2 (ON979684), and *S. zamani* (ON972386); partial mitochondrial *cox1* gene sequences of *Sarcocystis* sp.1 (ON989199), *Sarcocystis* sp.2 (ON989200), *S. singaporensis* (ON989197), and *S. zamani* (ON989198).

3. Results

3.1. Development of sarcocysts in laboratory rats after experimental infection with sporocysts of *Sarcocystis* sp.1 from mangrove snakes

Sporocysts of *Sarcocystis* isolated from two mangrove snakes in southern Thailand were morphologically identical and had the same size (initially designated as *Sarcocystis* sp.1), each isolate measuring on average 11.8 (± 0.5 , SD) \times 7.5 (± 0.2) μm ($n = 14$) and typical in appearance for the genus (Fig. 1a). Direct comparisons with four related *Sarcocystis* spp. (sporocyst length \times diameter in μm : *S. clethrionomyelaphis*: 11.0 [± 0.8] \times 8.0 [± 0.5], $n = 25$; *S. zuoi*: 10.8 [± 0.7] \times 8.0 (± 0.7), $n = 25$; *S. attenuati*: 9.9 [± 0.4] \times 6.6 [± 0.2], $n = 30$; and *S. murinotechis*: 11.0 [± 0.5] \times 7.3 [± 0.7], $n = 10$; Hu et al., (2015), Hu et al. (2012), Hu et al. (2022); Munday and Mason (1980), respectively) showed that the sporocysts from the mangrove snakes were significantly different in size from the other species. One-Way ANOVA of the five taxa showed highly significant differences in sporocyst length ($F = 27.347$, $P < 0.001$, d. f. = 103, Power of performed test with alpha = 0.05: 1.0), whereby pairwise comparisons against *Sarcocystis* sp.1 were all significant at $P < 0.05$. Results on sporocyst diameter were similar ($F = 38.828$, $P < 0.001$, d. f. = 103, Power of performed test with alpha = 0.05: 1.0), except that sporocysts of *Sarcocystis* sp.1 and *S. murinotechis* were not significantly different in pairwise comparison.

After inoculation of rats with both sporocyst isolates, we observed sarcocysts in the striated musculature of two rats (one isolate each) that were examined 3.5 months post infection (m.p.i.). In fresh tissue

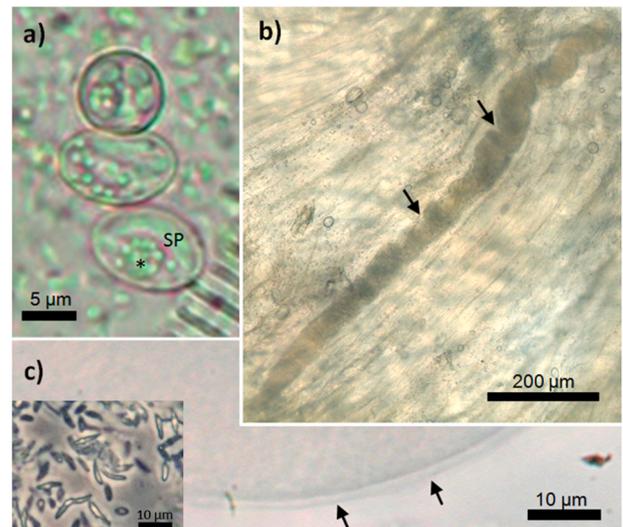


Fig. 1. Light microscopic observations on the development of *Sarcocystis* sp.1 in striated musculature of Sprague-Dawley rats; a) typical sporocysts, here in fecal smear from *Boiga dendrophila*, that were used for infection of rats; sporocysts contained a granular residual body (asterisk) and four sporozoites (SP), which are all visible in the upper sporocyst; b) full-length micrograph of a typical (native) sarcocyst of *Sarcocystis* sp.1 in striated belly musculature, the arrows indicating folds of the cysts' body; c) high magnification of the cyst wall of a native sarcocyst, the arrows pointing to the apparently smooth wall; this is the same cyst as depicted in Fig. 2a; the inset shows freshly released, live cystozoites under phase-contrast light microscopy.

preparations, sarcocysts of both isolates were morphologically identical and only located in the esophagus at low density. They were long, slender with spindle-shaped ends, and irregular in shape due to invaginations of the sarcocyst wall (Fig. 1b); they measured $830\text{--}1330 \times 89\text{--}133 \mu\text{m}$ ($n = 5$). At 4 m. p.i., the same type of sarcocyst was slightly larger in fresh preparations of the esophagus and abdominal musculature of two rats, measuring $1480\text{--}1510 \times 70\text{--}170 \mu\text{m}$ ($n = 7$). At high magnification under the light microscope, the cyst wall appeared smooth, measuring between 1.1 and $2.2 \mu\text{m}$ in thickness (Fig. 1c). At 14 m.p.i., we observed the same sarcocysts in the abdominal musculature of another two rats, whereby the size of sarcocysts had not increased further compared to earlier examination points, measuring $700\text{--}1250 \times 70\text{--}100 \mu\text{m}$ ($n = 10$). All six, negative control rats examined pairwise at the respective time intervals were free of *Sarcocystis* spp. and remained healthy throughout the experiment. Infected rats also did not show signs of disease at any time after inoculation.

Although the sarcocyst wall of *Sarcocystis* sp.1 appeared smooth under the light microscope (Fig. 1c), protrusions could be seen in semithin sections, especially, when host tissue was removed (Fig. 2a). Here, the cyst wall was between 1.2 and $2.6 \mu\text{m}$ thick (three cysts, $n = 28$), which was similar to native sarcocysts. The GS measured between 170 nm and 770 nm (three cysts, measurements $n = 28$). The ultrastructure of sarcocysts was different: the cyst wall (= primary cyst wall [PCW] plus ground substance [GS]) showed characteristic protrusions, which often possessed a broad (at times reticulated) base and were irregular-shaped in longitudinal sections (Fig. 2b). In cross-sections, protrusions appeared slenderer and finger-like (Fig. 2c). Regardless of the

orientation of sections, the PCW consisted of a chain of electron-dense knobs that were arranged in an undulating fashion. The potential reason for invisibility of the protrusions in their native state was revealed in ultrastructural preparations: apparently, the space between protrusions was filled with fluid matter as suggested by the presence of small granules and vesicles (Fig. 2c). This material was not visible once host tissue became detached from the cyst (Fig. 2a). The thickness of the electron-dense knobs of the PCW ranged between 37 and 60 nm (3.5 m. p.i.) and 27–53 nm (14 m.p.i.), which was not significantly different between the two examination points. Generally, metrocytes occurred rarely but more frequently at earlier examination points, suggesting that cysts had been largely mature by 3.5–4 months p. i. Already. Only cystozoites replicating by endodyogeny were observed: metrocytes undergoing cell division showed the formation of two zoites in all cases (Fig. 2d).

Live cystozoites were on average $7.7 (\pm 0.4; \text{SD}) \times 2.5 (\pm 0.2) \mu\text{m}$ ($n = 18$) large, which was recorded on two occasions (3.5 m.p.i. And 14 m.p. i), at which these stages did not differ in size. Cystozoites were slightly smaller in semithin sections, $6.2 (\pm 0.6) \times 2.4 (\pm 0.3) \mu\text{m}$ ($n = 15$). In pairwise comparisons with related *Sarcocystis* spp. (live cystozoites: *S. attenuati*: $9.2 [\pm 0.7] \times 2.5 [\pm 0.3] \mu\text{m}$, $n = 40$, Hu et al., 2022; *S. zuoi*: $8.5 [\pm 0.8] \times 2.8 [\pm 0.1] \mu\text{m}$, $n = 35$, Hu et al., 2012; *S. clethrionomyelaphis*: $9.5 [\pm 1.0] \times 2.7 [\pm 0.1] \mu\text{m}$, $n = 40$, Hu et al., 2015) cystozoites of *Sarcocystis* sp.1 differed significantly in length (One-Way ANOVA, $F = 16.694$, $P < 0.001$, d. f. = 121, Power of performed test with $\alpha = 0.050$: 1.000; all pairwise comparisons $P < 0.05$), while cystozoite diameter was similar to other species, except for *S. zuoi* ($P < 0.05$). When

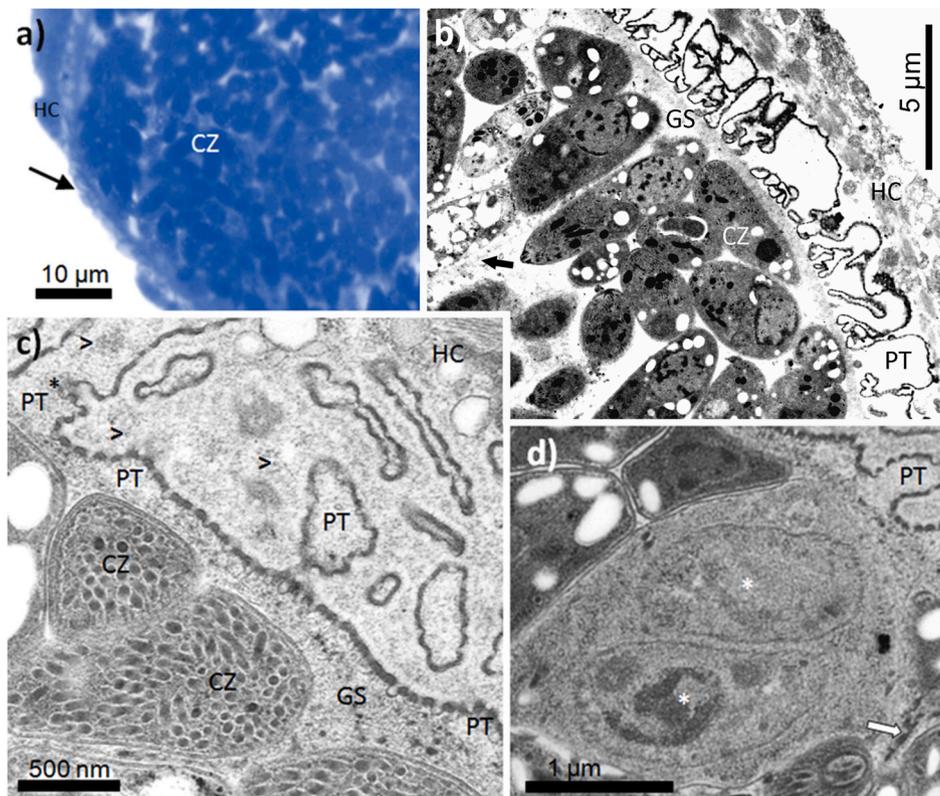


Fig. 2. Ultrastructure of *Sarcocystis* sp.1 from the mangrove snake in abdominal musculature of a Sprague-Dawley rat; **a)** One μm -thin section through a resin-embedded, toluidine-stained sarcocyst showing densely-packed cystozoites (CZ) that were contained in septate compartments; the arrow indicates the cyst wall with small protrusions, which are clearly visible because host cell (HC) tissue is removed at this position; **b)** gross view of the cyst wall and cystozoites in a longitudinal section; the black arrow indicates a thin septum that separated the compartments filled with cystozoites; protrusions (PT) were broad, short, and irregular-shaped, often with a reticulate base that rested on a thin layer of ground substance (GS); **c)** cross-section of sarcocyst, showing the primary cyst wall at higher magnification to consist of electron-dense, knob-like structures with intermittent invaginations; it appeared as if the primary wall was fenestrated (asterisk) allowing exchange of fine granular material (arrowheads) between the interior and exterior of the cyst; the exterior space between the protrusions was entirely filled with granular substance. **d)** metrocytes exclusively divided by endodyogeny as only cells with two developing zoites (asterisks) were observed; the white arrow points at deposits of highly electron-dense matter that could form larger clusters in the GS of the septae.

measuring formalin-fixed cystozoites in histological sections, *Sarcocystis* sp.1 was significantly longer and wider than *S. scandentiborneensis* (Ortega Pérez et al., 2020) (Two-tailed *t*-Test: length, $t = 4.849$, $P < 0.001$, d. f. = 98; diameter, $t = 17.081$, $P < 0.001$, d. f. = 98; Power of performed two-tailed tests with $\alpha = 0.05$: 1.0). Cystozoites of *Sarcocystis* sp. 1 were densely packed in septate chambers of the sarcocyst (Fig. 2a). At the ultrastructural level, cystozoites were on average 6019 (± 757) nm long and 2147 (± 290) nm wide ($n = 12$). We observed 3–4 rhoptries, 1–3 dense granules, and numerous micronemes (Fig. 2b and c). Cross-sections through the anterior part of cystozoites showed on average 61 micronemes (range: 48–96, $n = 11$) 3.5 m.p.i., and 72 micronemes (range: 45–90, $n = 9$) 14 m.p.i., which was not significantly different between the two examination points.

3.2. 18S rRNA phylogeny and comparison of evolutionary rates of *Sarcocystis* sp.1, *Sarcocystis* sp.2, and related species

We amplified a partial 18S rRNA gene sequence of *Sarcocystis* sp.1 (1209 nt, ON979685). Although the sequence is incomplete (covering the range from bp 485–1693 of the complete *S. neurona* sequence, accession number U07812), it includes parts of domain V3, the variable sites of domain V4, and continues until domain V9 of the predicted secondary structure of the 18 S rRNA molecule of *T. gondii* (Gagnon et al., 1996). Under Bayesian inference (BI), *Sarcocystis* sp.1 clustered with two partial sequences isolated from a colubrid snake in Malaysia (KC878487, KC878488) (Fig. 3), which were labeled as *S. zuoi* (Lau et al., 2013). However, we use the designation *Sarcocystis* cf. *zuoi* for this and all other isolates, for which no morphological evidence is available. Sequence KC878488 was highly similar to *Sarcocystis* sp.1 in pairwise BLAST comparison (957 identities out of 958 bp: 99.90%) with good branch support (posterior probability: 0.98–0.99); sequence KC878487

showed a slightly lower identity score (953 identities out of 958 bp: 99.48%). Furthermore, *Sarcocystis* sp.1 was closely related to *S. attenuati* and *S. scandentiborneensis*, because these sequences formed a monophyletic subclade (posterior probability: 0.97–0.99; Fig. 3), within which *S. attenuati* and *S. scandentiborneensis* clustered together. Sarcocysts were also observed in the muscles of a ricefield rat (*Rattus argentiventer*) from northern Sumatra. Unfortunately, we could not determine cyst morphology beyond the observation that live sarcocysts were microscopic, with a cyst wall that appeared striated. This species was labeled *Sarcocystis* sp.2. (1724 nt, ON979684). *Sarcocystis* sp.2 clustered with sequences KU341118/19/21 (*Sarcocystis* cf. *zuoi* from Thailand), forming a subclade with high branch support (Fig. 3). Sequence similarity of *Sarcocystis* sp.2 versus KU341118 was 98.32% (1639 identities out of 1667 bp), 98.51% if compared with KU341119 (1657 identities out of 1682 bp), and 98.43% versus KU341121 (1625 identities out of 1651 bp). However, the sarcocysts associated with these three sequences were reported to be smooth-walled (Wathanakaiwan et al., 2017). Another sequence, KU341120, which belongs to the previously mentioned sequence triplet from Thailand and all of which were considered conspecific, was however separated from them in the resulting phylogram. Hence, tree topology indicated that the *S. zuoi* - complex consists of at least five molecularly and morphologically distinct *Sarcocystis* species: *Sarcocystis* sp. 1, *S. sp.2*, *S. attenuati*, *S. scandentiborneensis*, and *S. zuoi*. The isolate LC054267 from Japan clustered with the original sequences of *S. zuoi* suggesting that it could be conspecific with the latter; however, this sequence was truncated and lacked the nt-motif characteristic for the *S. zuoi* - complex (see below). Since sequence KU341120 branched off basally from a clade containing four of the five mentioned species, it may potentially represent a sixth species (posterior probability: 0.97–0.99). In view of cyst wall structure, sequences KU341118/19/21 apparently belonged to a species different

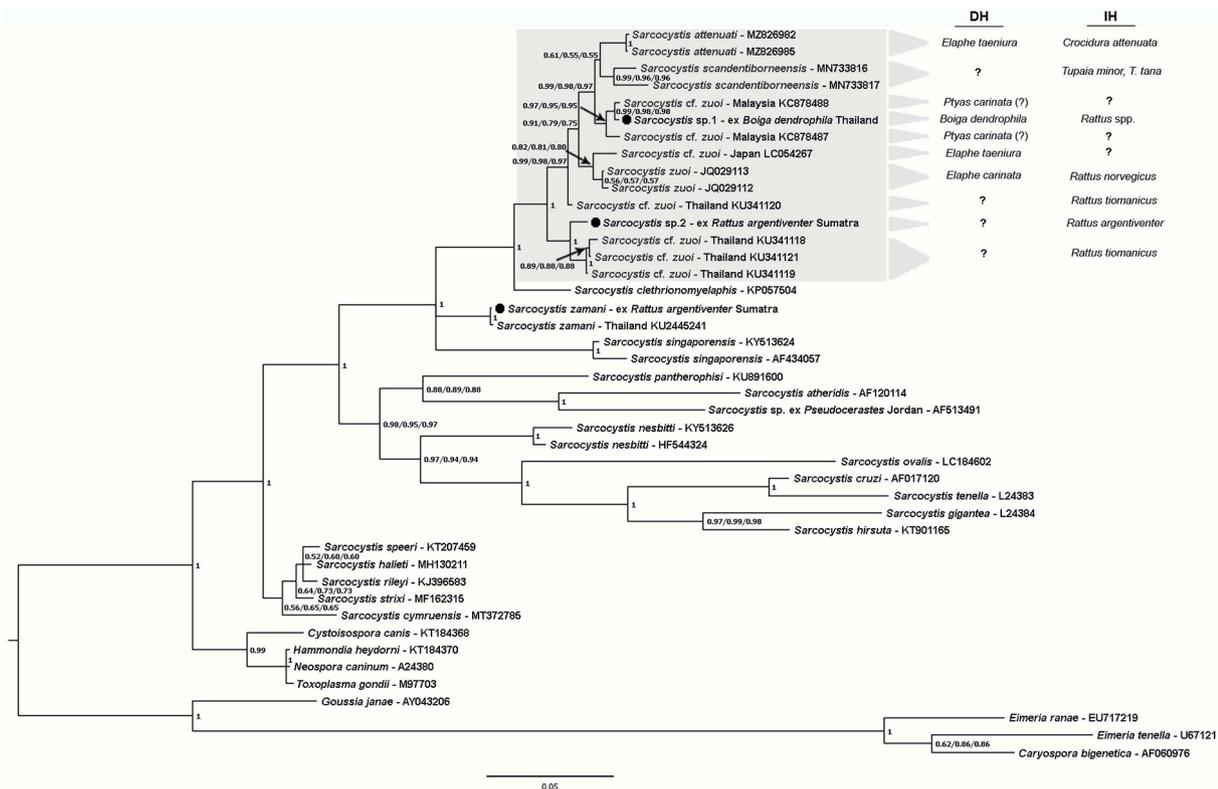


Fig. 3. Bayesian Inference (BI) of the 18S rRNA phylogeny of *Sarcocystis* spp. infecting colubrid snakes and small mammals in Asia; the *S. zuoi* - complex of species is highlighted by the shaded box. This complex excludes *S. clethrionomyelaphis*, which branches off basally. All new sequences, including a new isolate of *S. zamani* from Sumatra, are highlighted by black symbols. The known definitive and/or intermediate hosts associated with the selected sequences of the *S. zuoi*-group are indicated. Eimeriid species from phylogenetically diverse hosts served as outgroups. Bayesian posterior probabilities of three independent analyses (and alignments) are indicated behind each node, showing only one value if results of replicates were identical.

from *Sarcocystis* sp.2 and may represent a seventh taxon. All taxa were sister to *S. clethrionomyelaphis*, which branched off basally from the shared, monophyletic clade of *Sarcocystis* spp. presumably cycling between colubrid snakes and small mammals in Asia. We have added the known definitive and intermediate hosts of the *S. zuoi* - complex to the phylogenetic tree to highlight associations between sequences and hosts (Fig. 3). Notably, *Sarcocystis* spp. of snakehost lineage S2 and sequences of ruminants formed a monophyletic clade with the longest branch lengths of the tree. Tree topology under ML analysis was like the BI tree, except that the branching of *Sarcocystis* sp.1, *S. scandentiborneensis*, and *S. attenuati* was not resolved resulting in polytomy.

To scrutinize more closely the relationships among selected 18S rRNA sequences, we compared evolutionary rates by using Tajima's Relative Rate Test and inspection of transition/transversion ratios (Ti/Tv) of sequence pairs. *Sarcocystis* sp.1 and sequence KC878488 from Malaysia showed similar evolutionary rates (Supplementary Table S1) and only differed by a single transversion, additional evidence that the two sequences could belong to the same species (under the restriction that only partial sequences were compared). A similar result was obtained with sequence KC878487 from Malaysia, although differences were greater than in the previous comparison. *Sarcocystis* sp.1, *S. scandentiborneensis* (MN733817), and *S. attenuati* (MZ826982) showed significantly different evolutionary rates in pairwise comparisons when using *S. clethrionomyelaphis* or *S. singaporensis* as outgroup (Table S1). However, evolutionary rates were similar if *Sarcocystis* sp.1 was compared with other haplotypes of *S. scandentiborneensis* (MN733816) and *S. attenuati* (MZ826985), potentially indicating incomplete sorting of alleles among these taxa. If a closely related sequence was used as outgroup to test for subtle differences in substitution rates (e.g., KU341120 which was basal to the original *S. zuoi* according to the tree), evolutionary rates of *Sarcocystis* sp. 1 and *S. zuoi* (JQ029113) were significantly different. Similarly, *Sarcocystis* sp.2 was significantly different from sequences KU341118 and KU341119 when the closely related sequence KU341121 was used as outgroup, while Ti/Tv values were relatively low (*Sarcocystis* sp.2 vs. KU341118: 0.32, Ti = 7, Tv = 22; *S. sp.2* vs. KU341119: 0.37, Ti = 7, Tv = 19; *S. sp.2* vs. KU341121: 0.35, Ti = 7, Tv = 20). When comparing the largely complete 18S rRNA gene sequence of *Sarcocystis* sp.2 as representative of snakehost lineage S1 with those of lineage S2 (e.g., *S. pantherophisi*, *S. atheridis*), the latter showed significantly higher evolutionary rates ($P < 0.000005$). The same was true for sequences of ruminant-carnivore *Sarcocystis* spp., where sequence divergence was even more pronounced (Table S1).

Moreover, secondary structure alignment of the 18S rRNA (using SSU-ALIGN) of *Sarcocystis* sp.2 and (the shorter) *Sarcocystis* sp.1 indicated that these sequences contained relatively few nt positions that deviated from the consensus sequence of the structural model. Since paralogous sequences (pseudogenes) can exhibit high substitution rates at random (not functionally constrained) positions (Buckler et al., 1997), frequent non-canonical base pairings in the stem regions and single, deviant nt positions could potentially indicate presence of a paralogous sequence: such positions were rare accounting for 0.94% of 1700 nt and 1.17% of 1209 nt, respectively. The corresponding percentages for other species were 1.43% for *S. scandentiborneensis* (1820 nt; MN733816), 1.82% for *S. attenuati* (1868 nt; MZ826985), and 1.75% for *S. clethrionomyelaphis* (1773 nt; KP057504) while this value reached 4.29% in *Toxoplasma gondii* (1793 nt; M97703). All in all, these data suggested that the sequences of *Sarcocystis* sp.1 and *Sarcocystis* sp.2 were not paralogs and that *Sarcocystis* sp.2 was not conspecific with KU341118/19/21, with which it clustered (see above).

3.3. Shared 18S rRNA motif in helix 38 of *Sarcocystis zuoi*-species complex

Members of the *Sarcocystis zuoi* - species complex can be identified by the sequence motif 5'-AAUUCGU-3' in their 18 S rRNA, apparently a synapomorphy of this clade (Ortega Pérez et al., 2020). Here, we

observed this motif again in *Sarcocystis* sp.1, *Sarcocystis* sp.2, and all other taxa of the complex (except sequence LC054267 which was truncated). We were able to map this motif to a hairpin loop position on helix 38 in the V7 domain of the 18S rRNA by secondary structure-modelling in SSU-ALIGN. Unlike *S. clethrionomyelaphis*, which is not a member of the *S. zuoi* - complex, all member taxa exhibited an additional Cytosine at the tip of the loop (Fig. 4). According to the predicted secondary structure of the sequence of *Toxoplasma gondii* (X75453), this hairpin loop has a variable arc area that is less than 80% conserved (Cannone et al., 2002). To check for species-specificity of the motif beyond our own collection of *Sarcocystis* sequences, we submitted a BLAST search for a slightly longer motif (13 nt, including neighboring nucleotides) to the SILVA RNA Database of quality checked and aligned ribosomal RNA sequences (Yilmaz et al., 2014). A search strategy that did not allow mismatch between base pairs returned nine sequences out of 9.4 million 18S rRNA gene sequences. Out of the nine sequences, four were of Apicomplexan origin, all *Sarcocystis*: *S. zuoi* (JQ029112), *S. cf. zuoi* (KU341119), and two unidentified *Sarcocystis*-isolates collected from raccoons (*Procyon lotor*) in Japan (AB251613, AB251614). The other five sequences related to fungal species. For comparison, a search allowing one base pair mismatch resulted in 1538 sequences (bacteria, fungi, and other eukaryotes) confirming the high specificity of the 7-nt motif for members of the *S. zuoi*-complex.

3.4. Phylogeny of the mitochondrial *cox1* gene

We amplified four new sequences of the mt *cox1* gene: *Sarcocystis* sp.1 from *Boiga dendrophila* (927 bp), *Sarcocystis* sp.2 (986 bp) from the same sample (ricefield rat) collected in Sumatra as the 18S rRNA gene sequence, *S. singaporensis* (994 bp), and *S. zamani* (955 bp) (see section 2.10 for accession numbers). Intriguingly, the partial sequence of *Sarcocystis* sp.1 was identical with *cox1* of *S. scandentiborneensis* (MN732561) and *S. attenuati* (MZ889673) suggesting a shared, recent ancestry of the three species. The *cox1* genes of *Sarcocystis* sp.1 and *Sarcocystis* sp.2 only differed in 16 bp (1.73%) over a length of 927 aligned positions, while differences to species from pythonid hosts were greater: *Sarcocystis* sp.1 versus *S. singaporensis*, 56 bp exchanges in 927 (6.04%) positions; *Sarcocystis* sp.1 versus *S. zamani*, 40 bp exchanges in 927 (4.31%) positions.

Due to length variability of sequences, we restricted ML analysis for phylogeny reconstruction of the *cox1* gene tree to a 603 nt-long alignment of the barcode area of 25 *Sarcocystis* spp., where possible using the same taxa as for the 18S rRNA tree. Like the 18S rRNA phylogeny, *Sarcocystis* sp.1, *S. attenuati*, and *S. scandentiborneensis* formed a monophyletic subclade with *Sarcocystis* sp.2 as direct sister species (Fig. 5). *Sarcocystis singaporensis* and *S. zamani* branched off before, and together with *S. pantherophisi* (representative of lineage S2) the snakehost *Sarcocystis* spp. formed a monophyletic clade that shared a direct common ancestor with *Sarcocystis* spp. clustering with *S. speeri*, a close relative of *S. neuron*. In contrast to the 18S rRNA phylogeny, *Cystoisospora canis* clustered with the Sarcocystinae (Fig. 5), a topology that remained stable if the number of sites was increased. Notably, expansion of the tree with key taxa (e.g., *S. singaporensis*, *S. zamani*, *S. ovalis*) and inclusion of the fish parasite *Goussia bayae* (MH792860) and other selected eimeriid species as outgroup stabilized tree topology in that most members of the Toxoplasmatinae appeared as basal group of the Sarcocystidae (Fig. 5). This was different to a previous investigation (Ortega Pérez et al., 2020; see discussion section 4.3).

3.5. Lineage-specific evolution of the COX1 protein among Sarcocystidae

Since the *cox1* phylogenetic tree (and to a lesser degree also the 18S rRNA phylogeny) indicated that ruminant *Sarcocystis* spp. exhibited significantly higher evolutionary rates than other species of Sarcocystidae, we wanted to know what impact this would have on the aa composition of the COX1 protein and whether there existed lineage-

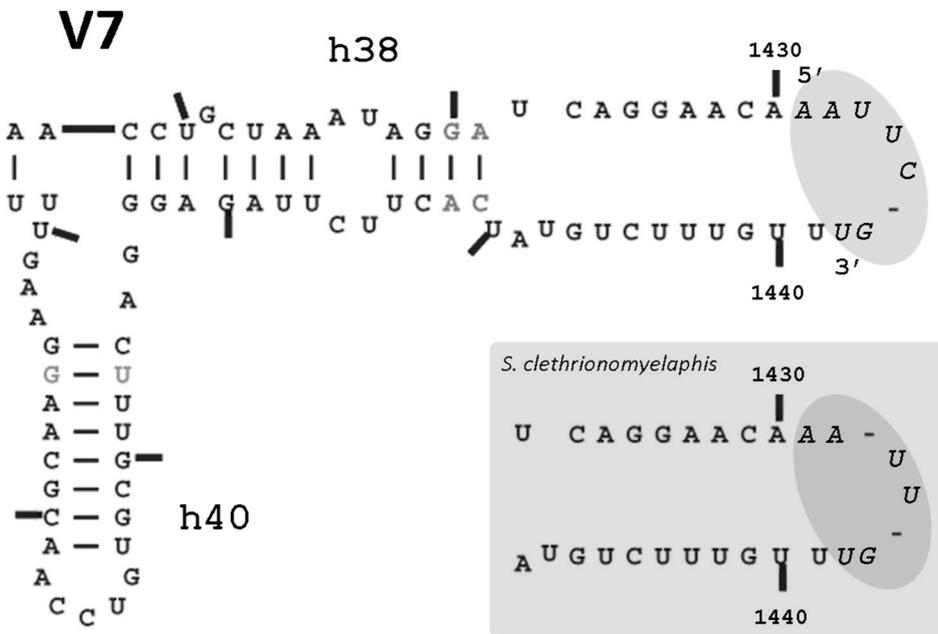


Fig. 4. Predicted secondary structure of helix 38 in domain V7 of the 18S rRNA of *Sarcocystis* sp.1, *Sarcocystis* sp.2, *S. attenuati*, *S. scandentiborneensis*, and *S. zuoi* in comparison to *S. clethrionomyelaphis* (shaded inset). The 7-nt long motif 5'-AAUUCGU-3' (relative to all Apicomplexan taxa examined; nt in italic letters in shaded oval) mapped to a hairpin loop position of helix 38 and was characteristic for all species of the *S. zuoi* – complex; the motif was one nt (Cytosine) shorter in helix 38 of *S. clethrionomyelaphis*. Sequences were aligned to a secondary structure model of the Eukarya using SSU-ALIGN. The numbering of nucleotides (bars) and helices is based on the 1881 nt-long structural template. Position 1430 corresponds to position 1357 of the predicted secondary structure of *Toxoplasma gondii* (RH strain) as published by Gagnon et al. (1996).



Fig. 5. Maximum Likelihood (ML) phylogenetic reconstruction of the mitochondrial Cytochrome C oxidase Subunit I (*cox1*) gene tree of the *Sarcocystis* spp. under investigation. Where possible, sequences of the same species were used as shown in the 18S rRNA gene tree, including the outgroup. The newly sequenced *Sarcocystis* spp. are marked with black symbols. A total of 25 nucleotide sequences and 603 sites of the barcode area (all codon positions included) was analyzed with 1000 bootstrap replicates. Branch support from three replicate analyses is shown. The evolutionary history was inferred by using ML based on the HKY model. The tree with the highest log likelihood (−6438.6551) is shown. A discrete Gamma distribution was used to model evolutionary rate differences among sites (four categories [+G, parameter = 0.9037]). The rate variation model allowed for some sites to be evolutionarily invariable ([+I], 17.4959% sites). All positions with less than 98% site coverage were eliminated. Branch lengths are measured in the number of substitutions per site. Note the long branch lengths of ruminant *Sarcocystis* spp. in comparison to other members of the Sarcocystidae as well as eimeriid coccidia from phylogenetically diverse hosts.

specific patterns in protein structure. As initial step, we used SWISS-MODEL, a fully automated protein structure homology-modelling server (Waterhouse et al., 2018), to determine the three-dimensional

conformation of the structural elements of the barcode area (six helices and interconnected loops) of the Sarcocystidae and assess whether model quality would be sufficiently high to identify putative ligand

binding sites. When the complete protein sequence of *T. gondii* was aligned with a template from *Saccharomyces cerevisiae* (SWISS-MODEL accession number 6hu9.1), the resulting aa alignment (474 aa, 94% coverage, 42% sequence similarity; GMQE score = 0.70) identified one (Cu_B) of the two copper binding sites (Maghool et al., 2020), which involved *Toxoplasma* - Histidine residues 253, 302 and 303 located on helices 6 and 7, respectively (Supplementary Fig. S1). No other ligand binding sites could be identified by homology modelling due to high variability of the Apicomplexan protein sequence relative to the template. Therefore, in a second step, we mapped putative ligand binding sites of the selected SWISS-MODEL templates onto the alignment of 18 COX1 protein sequences of the Sarcocystidae (Fig. 6A). This resulted in 21 putative heme binding sites, of which 19 are shown (in helices 1, 2, 6 and loop 3–4); two additional sites were close to the N-terminus of helix 1, which was truncated. Apparently, the putative heme binding sites were not only highly conserved among the Sarcocystidae but also in the template species: only one of the binding sites of the yeast template showed a substitution towards a chemically different aa group, while bovine COX1 exhibited four such changes. Only ruminant *Sarcocystis* spp. showed aa changes at putative heme binding sites that involved a switch between chemically different aa groups, e.g., in helix 2 at position 45, and in helix 6 at position 234 of the barcode area of the sequence of *S. hirsuta* (Fig. 6A).

In combination with the partial sequences of *T. gondii* (KM657810)

and *S. scandentiborneensis* (MN732561) and alignment with appropriate templates (*Bos taurus* - 3 abm.1 and *Saccharomyces cerevisiae* - 6hu9.1, respectively) we could delimit the parasite-specific component boundaries of the barcode area. A protein sequence alignment of the complete barcode area is shown for the selected 18 species of Sarcocystidae (Supplementary Fig. S2). Subsequently, we produced a ‘heat map’ for visualization of variable amino acid sites by counting all aa substitutions that involved a change from one chemical aa group to another relative to *Toxoplasma gondii* as reference sequence (Fig. 6B). As could be expected from the long branches of the *cox1* gene tree, ruminant *Sarcocystis* spp. showed the highest aa variability, whereby canine-ruminant species (*S. cruzi*, *S. tenella*) reached maximum values, especially affecting helices 2 to 5 and loops 1–2, 4–5, and 5–6. All snakehost *Sarcocystis* spp. showed similar patterns, except *S. zamani* which had an additional aa substitution in helix 5 (Fig. 6B). Species of the *S. neurona*-clade were like the previous group, except that variability was lower in helix 4 and loop 4–5. Interestingly, *Cystoisospora canis* exhibited a variability pattern that was closer to the Sarcocystinae than the Toxoplasmatinae; however, the sequence was incomplete in the anterior barcode area. We observed lineage-specific aa substitution sites of the *S. zuoi*-complex (*Sarcocystis* sp.1, *S. sp.2*, *S. attenuati*, *S. scandentiborneensis*) at positions 139 (Glycine) in helix 4 and 176 (Methionine) in helix 5 while these changes occurred within the same chemical aa group (Fig. S2). Similar lineage-specific substitutions were also present in other groups, e.g., the

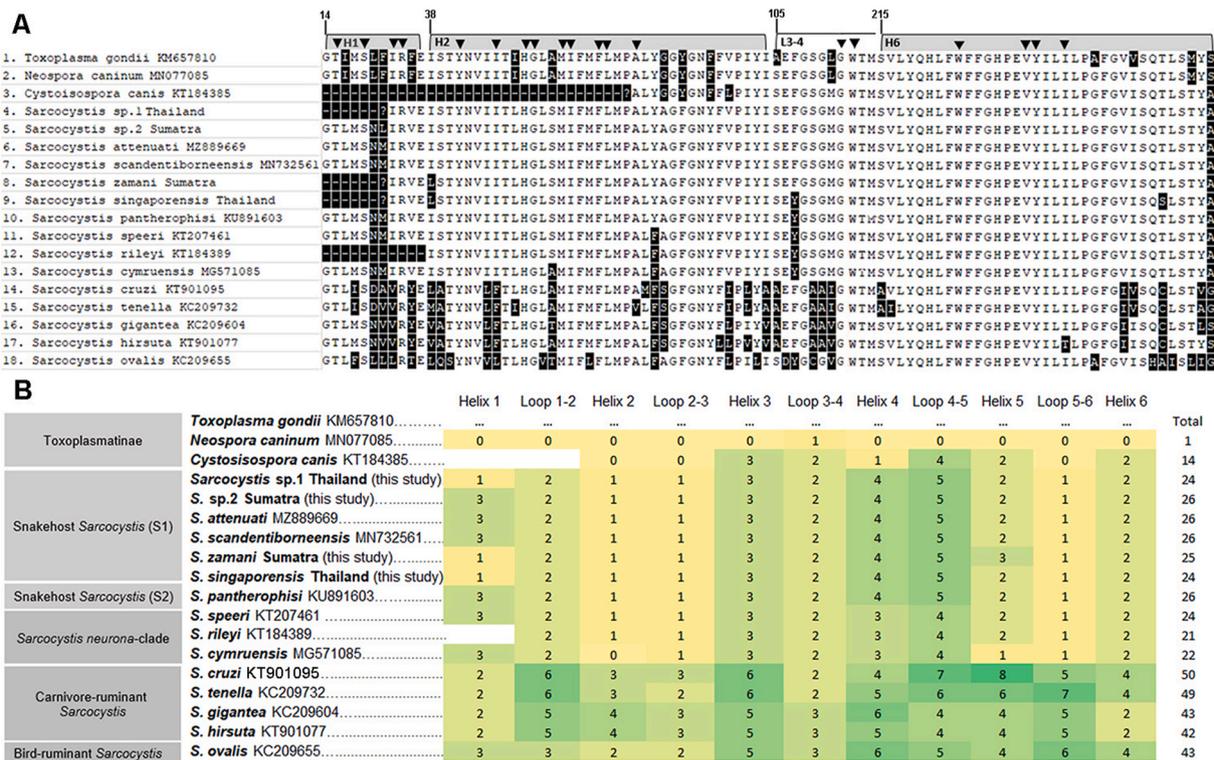


Fig. 6. Mapping of (A) potential heme ligand binding sites and (B) amino acid variability among different lineages of tissue cyst-forming coccidia in the barcode area of the mitochondrial COX1 protein. A) Map of putative heme ligand binding sites (arrowheads) in a protein sequence alignment of selected taxa used in the phylogenetic tree of *cox1*. Identical/conservative aa positions are highlighted by light background, variable positions and gaps are shown against black background. Helix 1 (H1) is shown partial, starting at position 14 of the global barcode alignment (Pentinsaari et al., 2016); aa sequences of helices 2 (H2) and 6 (H6) are shown in full length, while putative heme binding sites of loop 3–4 (L3-4) were in its anterior part only. Domain boundaries and putative ligand binding sites were derived from COX1 of template organisms *Saccharomyces cerevisiae* and *Bos taurus* by sequence alignment against *Toxoplasma gondii* applying three-dimensional homology modelling of protein structure. The complete alignment of the six helices of the barcode area is shown in Supplementary Fig. S2. Note that aa numbering of the barcode area of the Apicomplexan sequence is different to the global alignment, because the former showed one additional aa and domain boundaries were slightly altered. The nucleotide sequence KC209732 of *S. tenella* (GenBank) is also registered in the barcode reference database BOLD (accession number JRPAA5858-15; <http://boldsystems.org>); the aa barcode position 14 (Glycine) shown here corresponds to position 16 of the aa translation of GenBank record KC209732. B) ‘Heat map’ of aa changes (darker shades of green = more changes, number of changes indicated) in the barcode area among chemically/structurally different aa groups in different lineages of tissue cyst-forming coccidia relative to the COX1 protein sequence of *T. gondii*. Because helix 1 was truncated, records of aa changes in this area are incomplete.

canine-ruminant and feline-ruminant *Sarcocystis* spp.

3.6. Taxonomic description: *Sarcocystis kani* sp. nov

***Sarcocystis kani* sp. nov.** (*Sarcocystis* sp.1: Figs. 1–6, Fig. S2, Table S1).

Diagnosis: Oocyst containing two sporocysts with four sporozoites each; sporocysts measuring on average $11.8 \times 7.5 \mu\text{m}$; sarcocysts macroscopic, filiform, up to 1.5 mm long and 70–170 μm wide in the esophagus and abdominal musculature; cyst wall smooth and 2.2 μm wide in fresh preparations, but showing protrusions in histological sections; banana-shaped, live cystozoites on average $7.7 \times 2.5 \mu\text{m}$ in size; ultrastructurally, cyst wall with characteristic, broadly-branched, rectangular to irregular-shaped protrusions with a maximum length of 2.6 μm ; sarcocyst septated whereby septae may contain areas of accumulation of highly electron-dense matter.

Type (natural) definitive host: *Boiga dendrophila* (subspecies *melanota*).

Experimental intermediate host: *Rattus norvegicus*

Type locality: Southern Thailand (Khao Sok).

Other locations and natural hosts: Described by Kan (1979) as type III sarcocysts from *Rattus diardii* in Selangor, Malacca, and Perak, and from *R. exulans* in Selangor, Malaysia.

Etymology: Species named after S. P. Kan, University of Malaya, Kuala Lumpur, who first described the characteristic ultrastructure of the sarcocyst in rats from Malaysia (Kan, 1979).

Type specimens deposited: Sarcocysts in tissue pieces from a small area in the belly musculature of one *Rattus norvegicus* infected for 3.5 months; type specimens of equal status (syntypes) fixed in phosphate-buffered Glutaraldehyde at the Senckenberg Research Institute and Natural History Museum Frankfurt/M., Division Botany and Molecular Evolution, Senckenberganlage 25, D-60325 Frankfurt/Main, Germany; Accession No.: FR-0162886.

Sequences deposited at GenBank: nuclear *18S rRNA* gene sequence (ON979685), mitochondrial *cox1* gene sequence (ON989199).

ZooBank (zoobank.org) registration number: 1DC5FC33–B73B–4F58–A683–C2E18896829C.

4. Discussion

4.1. Taxonomic remarks

The ultrastructure of the ‘type III’ sarcocyst from rats in Western Malaysia presented by Kan (1979) is identical with our observations on *S. kani* sp. nov. in almost every detail, including spots of apparent accumulations of electron-dense matter within the ground substance of the cyst’s septae. Therefore, it is not surprising that molecular samples that were collected in Western Malaysia (*18S rRNA*: KC878487, KC878488) clustered with *S. kani*, implicating that the parasite’s distribution extends from Thailand southwards into the Malay peninsula; due to their high identity scores in comparison to sequence ON979685, we regard the two previous sequences conspecific with *S. kani*. The sarcocyst of *Sarcocystis kani* sp. nov. is ultrastructurally unique in comparison to all other *Sarcocystis* spp. with snake-small mammal life cycle that possess a striated cyst wall; it resembles type 20 of the common classification system (Dubey et al., 2016). We have discussed the taxa that belong to this group in detail previously (Ortega Pérez et al., 2020). Morphological disparity is especially pronounced in comparison with the closest molecular relatives, *S. attenuati* (Hu et al., 2022) and *S. scandentiborneensis* (Ortega Pérez et al., 2020), which possess completely different cyst wall protrusions. However, the basic structural elements that form the primary cyst wall of *S. kani*, the electron-dense knobs, can be also observed in other *Sarcocystis* spp. of the S1-lineage, including *S. clethrionomyelaphis* (Matuschka, 1986) and *S. singaporensis* (Beaver and Maleckar, 1981) underpinning the close relationships between these species. *Sarcocystis kani* was clearly distinguishable from sister species regarding its size of cystozoites and sporocysts.

Molecularly, *Sarcocystis kani* shared an identical (partial) *cox1* gene sequence with *S. scandentiborneensis* and *S. attenuati*. We are confident that both sequences, *cox1* and *18S rRNA* (ON989199, ON979685), relate to the same species because they were amplified from a single sporocyst. In the *18S rRNA* gene tree, *S. kani*, *S. scandentiborneensis*, and *S. attenuati* were clearly distinct forming a clade with high branch support, in which *S. attenuati* and *S. scandentiborneensis* were nested as subgroup sharing a direct common ancestor. This branching pattern is congruent with both species developing in non-murid intermediate hosts. Since *Sarcocystis* sp.2 is also integrated in the *S. zuoi*-complex, it is likely that it also develops in colubrid snakes. The tree topology regarding the positions of *Sarcocystis* sp.2 and *S. kani* was similar in both trees: *Sarcocystis* sp.2 branched off basally from the clade that contains the two species. Furthermore, congruence of the *18S rRNA* and *cox1* phylogenetic trees also related to the positions of *S. singaporensis* and *S. zamani* as basal species of lineage S1, while lineage S2 (e.g., *S. pantherophisi* and sister species) formed a separate clade.

Biologically, *S. kani*, *S. attenuati*, and *S. scandentiborneensis* use phylogenetically and physiologically distinct mammal intermediate hosts. *Sarcocystis kani* develops relatively long sarcocysts in murid rodents, sarcocysts of *S. attenuati* are of similar length in the much smaller Asian Gray shrew (Hu et al., 2022), and *S. scandentiborneensis* develops small sarcocysts in the relatively larger treeshrews (Ortega Pérez et al., 2020).

4.2. Potential recent origin of *Sarcocystis kani* and related species in Asia

We demonstrate in this study that a clade of *Sarcocystis* spp. of colubrid snakes and small mammals in Asia exhibits a low level of genetic divergence, so low that one could speak of molecularly cryptic taxa. This was especially evident from *cox1* gene sequence comparisons, where three morphologically distinct species of the *S. zuoi*-complex, including *S. kani*, shared identical sequences. Although the sequences are incomplete, they include the barcode area which has been selected as global marker for species discrimination (Hebert et al., 2003; Pentinsaari et al., 2016). The close relatedness of these *Sarcocystis* spp. is also apparent in the short branch lengths of the *18S rRNA* gene tree (and contrasted by relatively longer branches of species of lineage S2). This is probably the reason why various *Sarcocystis*-isolates from colubrid snakes in Asia were previously considered monospecific (collectively addressed as *S. zuoi*), especially, if phylogenies were based on the highly conserved parts of the *18S rRNA* gene (Ortega Pérez et al., 2020). Thus, given the high degree of genetic similarity, what could be the potential evolutionary drivers that have led to morphological disparity and differences in host preference among these species?

Because mitochondrial DNA shows a relatively rapid rate of mutation, which makes it suitable as marker for more recent evolutionary history of natural populations (Brown et al., 1979; Norman et al., 2014), we interpret low sequence variability at the *cox1* locus as an indicator for a relatively recent origin of these *Sarcocystis* spp. Perhaps, the *Sarcocystis* spp. of rat snakes from Asia (*Elaphe* spp. and other closely related genera of the Colubridae) constitute an example of adaptive radiation (Schluter, 2000) that followed the radiations of their hosts, exploiting new ecological opportunities for speciation. Indeed, the evolutionary mechanism involved in speciation of Apicomplexan parasites and other pathogens, especially in the case of closely related taxa (Woolhouse et al., 2005), have been suggested host switching or adaptive processes rather than co-speciation (Kvicerova and Hypsa, 2013; Santiago-Alarcon et al., 2014). However, Šlapeta et al. (2003) provided evidence for co-speciation events between pythonid and viperid snakes and their associated *Sarcocystis* spp. based on the *18S rRNA* tree, proposing coevolution of *Sarcocystis* spp. with snake-rodent life cycle with their definitive hosts. The latter scenario is not necessarily in conflict with potential adaptive evolutionary processes among the *Sarcocystis* spp. of colubrid snakes, since *Sarcocystis* spp. in pythons (e.g., *S. singaporensis*), African/Asian vipers (e.g., *S. atheridis*, *Sarcocystis* isolated from

Pseudocerastes), and from other colubrid snakes (e.g., *S. pantherophisi*) are apparently phylogenetic older taxa that exhibited considerably longer branch lengths in phylogenetic trees.

Regarding intermediate hosts, the occurrence of *Sarcocystis* in omnivorous treeshrews (*S. scandentiborneensis*) and insectivorous shrews (*S. attenuati*) is uncommon if one considers that most of the *Sarcocystis* spp. of lineage S1 (such as *S. kani*) use murid rodents as intermediate host (Wassermann et al., 2017). This could be due to undersampling of non-murid small mammal species in the past. The *18S rRNA* phylogenetic tree hints to another potential host switching scenario: *S. scandentiborneensis* and *S. attenuati* share a direct common ancestor that could have initiated switch from a murid to non-murid mammalian host. Since treeshrews and shrews are believed phylogenetically older mammals (Roberts et al., 2011; three studies on shrews in Kumar et al., 2017, last accessed March 12, 2023) than murid *Rattus* spp. (Verneau et al., 1998; Song et al., 2014), such a host switch could have occurred more recently which would be congruent with the absence of sequence variability at the *cox1* locus and a relatively low sequence divergence of the *18S rRNA* of these *Sarcocystis* species. The opposite scenario, origin in phylogenetically older treeshrews or shrews and host switch to murids, is difficult to reconcile with a low level of genetic divergence. A similar discussion of whether the origin of hantaviruses is either in murid rodents or in shrews and other non-rodent mammals has been resolved by the fact that phylogenetically older, non-rodent borne hantaviruses have been detected in Asia (Henttonen et al., 2008, Bennett et al., 2014). This supports our reasoning that if *Sarcocystis* were originated from shrews or tree shrews, one would expect a higher level of genetic diversity indicating that they are phylogenetically older than those from rodents; which is not the case. Given the high diversity of murid rodents of the genus *Rattus* and related genera which diversified in Asia within the last 5–6 million years (Verneau et al., 1998; Song et al., 2014), this would have provided ample opportunities for dispersal and speciation of *Sarcocystis*.

4.3. Utility of *18S rRNA*, *cox1*, and other genes for species delimitation

For studying the molecular evolution of Apicomplexan parasites only few genes have been available, with most of the taxon sampling restricted to the nuclear *18S rRNA* gene (Morrison, 2009). Newer studies have expanded to the genomic level but this is restricted in the case of the Sarcocystidae to species of medical and veterinary importance (Mathur et al., 2020; Oborník, 2020; Salomaki et al., 2021). Single-gene phylogenetic analyses have increasingly focused on *cox1*, after it was found that *Sarcocystis* spp. parasitizing ruminants could be well distinguished with this marker (Gjerde, 2013b). This raised concerns regarding utility of the *18S rRNA* gene for molecular systematics (e.g., El-Sherry et al., 2013; Poulsen and Stensvold, 2014) although it is known that various lineages of coccidia can exhibit different rates of sequence evolution (Morrison et al., 2004; Šlapeta et al., 2003) and identification of potential paralogous sequences has been discussed (Rooney, 2004). Similarly, the presence of nuclear-encoded sequence fragments of mitochondrial origin (NUMTs) also complicated characterization of *cox1* of *T. gondii* and related coccidia (Gjerde, 2013a) which has been largely resolved with publication of the complete mitochondrial genome of *T. gondii* (Namasivayam et al., 2021). Here, we argue that a universal marker gene that could best resolve the phylogeny of all cyst-forming coccidia may not exist: nuclear and mitochondrial or organellar genes may show satisfactory utility for certain lineages of parasites, while they can show, at the same time, limited discriminatory power for others. The results of our phylogenetic analyses are a case in point: when using character-based phylogenetic reconstruction methods (BI, ML), the *18S rRNA* tree could distinguish well between *Sarcocystis* spp. of snakes as definitive hosts including the newly isolated *S. kani* and *S. sp.2*, whereas *cox1* was less informative regarding the species of interest due to little sequence divergence or complete absence thereof. New observations on rodents in South America revealed a similar

situation for the *18S rRNA* and *cox1* among novel *Sarcocystis* – isolates that clustered in the *S. neuropa* clade of species (Rossoli et al., 2023). Since fast-evolving species can show incomplete lineage sorting of alleles which may lead to polyphyly or paraphyly (Maddison and Knowles, 2006; Belfiore et al., 2008), we included in our *18S rRNA* tree, where possible, two haplotypes per species to examine potential lack of complete sorting. Although we observed considerable intraspecific variation in evolutionary rates among *18S rRNA* sequences of closely related species, tree topology did not result in paraphyly for any of the established species, i.e., *S. attenuati*, *S. scandentiborneensis*, or the original *S. zuoi* isolates. Hence, it appears that lineage sorting is largely complete for the *18S rRNA* gene, while it is not in the case of the *cox1* locus of these species.

Considering other genes or intergenic markers, there are still limited data on the *ITS1* locus. Hu et al. (2022) presented a phylogeny with branching patterns that resembled the *cox1* gene tree with respect to some of the main lineages of *Sarcocystis*; however, more data from small mammals are required, also in view of exploring mitonuclear discordance which we discuss below. Watthanakaiwan et al. (2017) presented a phylogeny of the combined *18S-ITS1-5.8S-ITS2-28S rRNA* gene complex of *S. singaporensis*, *S. zamani*, and *S. cf. zuoi* from Thailand, which showed *S. cf. zuoi* clustering with the *Sarcocystis* spp. from pythonid snakes in a monophyletic fashion similar to the *18S rRNA* tree presented here. However, a more extended selection of taxa is needed to assess the utility of this combination for *Sarcocystis* spp. cycling between snakes and small mammals. General concerns have been raised regarding the combination of *rRNA* genes with the *ITS* locus, particularly combining for analysis the relatively conserved *5.8S rRNA* gene with the highly variable intergenic sequences *ITS1* and *ITS2*, which may lead to inconsistencies in the alignment and the topology of the resulting trees (Bengtsson-Palme et al., 2013). These authors also pointed out that many *ITS* sequences are incorrectly defined in the public sequence databases. Thus, while *ITS* sequences are certainly a promising tool for species delimitation, they would need to be carefully trimmed, annotated, and aligned, possibly also considering secondary-structure information since *ITS1* and *ITS2* form secondary structures with stems, bulges, and loops (Rampersad, 2014). Hu et al. (2022) amplified a variety of genes of *S. attenuati* that might be useful for future studies on the group of *Sarcocystis* spp. in focus here including mitochondrial *cox3*, *cytb*, caseinolytic protease C (*clpC*), and *rpoB* of the apicoplast. Although there is a growing number of *cytb* and *rpoB* sequences for the Apicomplexa in public databases, these sequences were not available for most of the species in this study. An additional reason for caution in putting too much weight on a single marker for species delimitation is the circumstance that any gene tree is only contributing to the whole, the species tree (Morrison, 2009). Although single genes/markers might be good proxies for a species tree in one or the other case, species delimitation aims at integrating various genes to study the coalescent processes that form different lineages under the phylogenetic species concept (Avice and Wollenberg, 1997). We think it is the great achievement of Pentinsaari et al. (2016) to point out that genes like *cox1* are important for phylogenetic analysis because its evolution and function in the tree of life can be studied together (instead of separately). In our view, the same applies to the *18S rRNA* gene, which is still of major importance when it comes to taxonomic decisions regarding the Apicomplexa (e.g., Cavalier-Smith, 2014). For these reasons, we included functional aspects of both genes, e.g., heme binding sites and the role of helix 38, respectively. Enhanced sampling and future multigene phylogenetic analyses of the snakehost *Sarcocystis* - lineages S1 and S2 will be useful for reconstructing the corresponding species trees. Ideally, that would include various haplotypes per gene and species and coalescent reconstruction to accommodate the possibility of incomplete gene segregation at species boundaries (Belfiore et al., 2008).

Our phylogenetic trees include what has been recognized as potential complication for species delimitation in other organisms: mitonuclear discordance (Després, 2019). Regarding the genus *Sarcocystis*,

mitonuclear discordance had its bearing on tree topology, in particular the branching order of major clades. The most obvious: in the *cox1* gene tree bird-ruminant and carnivore-ruminant *Sarcocystis* spp. branched off first from the Sarcocystinae as monophyletic subclade (see also Ortega Pérez et al., 2020; Pan et al., 2020). In contrast, this position is occupied by species grouping with *S. neuropa* in rRNA gene trees under different sequence alignment methods and tree-building algorithms (this study; Mugridge et al., 2000; Šlapeta et al., 2003; Wassermann et al., 2017; Ortega Pérez et al., 2020). Moreover, in a previous study, a *cox1* gene tree with fewer taxa exhibited paraphyly of the Sarcocystinae (Toxoplasmatinae nested within Sarcocystinae) if all three codons were used for tree reconstruction (Ortega Pérez et al., 2020). In view of our results here, we think that this topological instability was potentially caused by long branch attraction of the (then piroplasmid) outgroup towards the rapidly evolving ruminant *Sarcocystis* spp., because it could be remedied by including further key species of *Sarcocystis* and diversification of the outgroup.

4.4. Lineage-specific evolution of 18S rRNA and COX1 protein

We have shown previously that members of the *S. zuoi*-species complex could be recognized by a 7-nt motif of the 18 S rRNA (5'-AAUUCGU-3') that did not occur in *S. clethrionomyelaphis* and other related species of lineage S1 (Ortega Pérez et al., 2020). Here, we identified this motif again in all new sequences that were part of the *S. zuoi*-clade. Our database searches confirmed its high degree of specificity. We mapped the motif's position to a hairpin loop of helix 38 in domain V7. Although V7 shows lower interspecific variability than the highly variable domains V2, V4, and V9 in the case of *S. scandentiborneensis* (Ortega Pérez et al., 2020), the arc area of the tip of helix 38 generally exhibits relatively high variability (Cannone et al., 2002). Helix 38 is part of the head of the 18S rRNA, which is important during tertiary folding (Huang and Karbstein, 2021), adopting multiple structures during translation (Hinnebusch, 2017). Helix 38 has a prominent role in the formation of a bridge with the 40S subunit and in forming a contact with the tRNA (Spahn et al., 2001), while neighboring helix 34 includes a contact zone of mRNA nucleotides with 18S rRNA in 48S/80S initiation complexes (Pisarev et al., 2008). All in all, these observations highlight a potential functional role of helix 38 in protein translation and suggest that lineage-specific evolution of the *S. zuoi*-group at this locus could be related to functional constraints; however, to confirm this, it warrants further investigation.

Lineage-specific patterns were also apparent in the structure of the COX1 protein. This analysis was largely enabled through identification of the correct nucleotide and protein sequences of COX1 (Gjerde, 2013a; Namasivayam, 2015). Based on the alignment of partial sequences that included the barcode area of the COX1 protein (Pentinsaari et al., 2016), members of the Sarcocystinae and Toxoplasmatinae exhibited several highly conserved aa sites that could putatively bind heme, which were derived from homologous positions of well characterized yeast and bovine molecules through three-dimensional homology modelling (Waterhouse et al., 2018). Since a high degree of conservation of these sites occurred across all examined taxa, this is most likely indicative of functional constraints associated with these positions, because COX1 is located at the core of energy production within cells (Srinivasan and Avadhani, 2012). However, there is a phylogenetic signal in the protein sequences in form of lineage-specific patterns of structurally/functionally important aa changes between the Toxoplasmatinae and Sarcocystinae, and between subgroups of *Sarcocystis* spp. whereby ruminant *Sarcocystis* spp. exhibited the highest aa substitution rates. This implies significant functional differences between lineages and single species of tissue cyst-forming coccidia. It is telling that the highest aa variability occurred in helices 3, 4, and 5 and loops 1–2, 4–5, and 5–6, not located at the heme binding sites. However, carnivore-ruminant *Sarcocystis* spp. also showed increased changes in Helix 2 highlighting the unique evolutionary path of this lineage. Taken together, our analyses revealed

new insights into the molecular evolution of different lineages of *Sarcocystis* species. It is becoming increasingly clear that also non-coding DNA (Guo et al., 2017) and synonymous mutations of proteins (Bailey et al., 2014) can be of functional relevance, supporting the notion that evolution and function belong together in view of species delimitation.

Declaration of competing interest

None.

Acknowledgements

We are grateful to the team of the Agricultural Zoology Research Group at the Department of Agriculture in Bangkok, Thailand, for allowing us to use their laboratory for microscopical examination of the parasite isolates from Thailand. We thank Claudia Schweitzer, who assisted in preparing samples for electron microscopy. The study was carried out with partial funding from the German Federal Ministry of Economic Cooperation and Development (BMZ) (Project No. 2002.2156.4; <http://www.bmz.de>; T.J. for German International Cooperation [GIZ]).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.ijppaw.2023.10.005>.

References

- Abe, N., Matsubara, K., Tamukai, K., Miwa, Y., Takami, K., 2015. Molecular evidence of *Sarcocystis* species in captive snakes in Japan. *Parasitol. Res.* 114, 3175–3179.
- Anonymous, 1999. Ethical Principles and Guidelines for the Use of Animals. National Research Council of Thailand, Section of Research Animal Development, Office of Economic and Peace Creation.
- Aoyama, H., et al., 2009. A peroxide bridge between Fe and Cu ions in the O₂ reduction site of fully oxidized cytochrome c oxidase could suppress the proton pump. *Proc. Natl. Acad. Sci. USA* 106, 2165–2169.
- Avise, J.C., Wollenberg, K., 1997. Phylogenetics and the origin of species. *Proc. Natl. Acad. Sci. USA* 94, 7748–7755.
- Bailey, S.F., Hinz, A., Kassen, R., 2014. Adaptive synonymous mutations in an experimentally evolved *Pseudomonas fluorescens* population. *Nat. Commun.* 5, 4076. <https://doi.org/10.1038/ncomms5076>.
- Beaver, P.C., Maleckar, J.R., 1981. *Sarcocystis singaporensis* Zaman and Colley, (1975) 1976, *Sarcocystis villivillosi* sp. n., and *Sarcocystis zamani* sp. n.: development, morphology, and persistence in the laboratory rat, *Rattus norvegicus*. *J. Parasitol.* 67, 241–256.
- Belfiore, N.M., Liu, L., Moritz, C., 2008. Multilocus phylogenetics of a rapid radiation in the genus *Thomomys* (Rodentia: geomyiidae). *Syst. Biol.* 57, 294–310.
- Bengtsson-Palme, J., Ryberg, M., Hartmann, M., Branco, S., Wang, Z., Godhe, A., De Wit, P., Sánchez-García, M., Ebersberger, I., de Sousa, F., Amend, A., 2013. Improved software detection and extraction of ITS1 and ITS 2 from ribosomal ITS sequences of fungi and other eukaryotes for analysis of environmental sequencing data. *Methods Ecol. Evol.* 4, 914–919.
- Bennett, S.N., Gu, S.H., Kang, H.J., Arai, S., Yanagihara, R., 2014. Reconstructing the evolutionary origins and phylogeography of hantaviruses. *Trends Microbiol.* 22, 473–482.
- Bertoni, M., Kiefer, F., Biasini, M., Bordoli, L., Schwede, T., 2017. Modeling protein quaternary structure of homo- and hetero-oligomers beyond binary interactions by homology. *Sci. Rep.* 7, 10480.
- Brown, W.M., George, M., Wilson, A.C., 1979. Rapid evolution of animal mitochondrial DNA. *Proc. Natl. Acad. Sci. USA* 76, 1967–1971.
- Buckler, E.S. 4th, Ippolito, A., Holtsford, T.P., 1997. The evolution of ribosomal DNA: divergent paralogues and phylogenetic implications. *Genetics* 145, 821–832.
- Cannone, J.J., Subramanian, S., Schnare, M.N., Collett, J.R., D'Souza, L.M., Du, Y., Feng, B., Lin, N., Madabusi, L.V., Müller, K.M., Pande, N., Shang, Z., Yu, N., Gutell, R.R., 2002. The comparative RNA web (CRW) site: an online database of comparative sequence and structure information for ribosomal, intron, and other RNAs. *BMC Bioinf.* 3, 2.
- Cavaliere-Smith, T., 2014. Gregarine site-heterogeneous 18S rDNA trees, revision of gregarine higher classification, and the evolutionary diversification of Sporozoa. *Eur. J. Protistol.* 50, 472–495.
- Clark, J.D., Gebhart, G.F., Gonder, J.C., Keeling, M.E., Kohn, D.F., 1997. The 1996 Guide for the Care and Use of Laboratory Animals. *ILAR J.* 38, 41–48.
- Clark, K., Karsch-Mizrachi, I., Lipman, D.L., Ostell, J., Sayers, E.W., 2016. GenBank. *Nucleic Acids Res.* 44, D67–D72.
- Després, L., 2019. One, two or more species? Mitonuclear discordance and species delimitation. *Mol. Ecol.* 28, 3845–3847.

- Dubey, J.P., Calero-Bernal, R., Rosenthal, B.M., Speer, C.A., Fayer, R., 2016. *Sarcocystosis of Animals and Humans*. CRC Press, Taylor & Francis Group, Boca Raton.
- El-Sherry, S., Ogedengbe, M.E., Hafeez, M.A., Barta, J.R., 2013. Divergent nuclear rDNA paralogs in a Turkey coccidium, *Eimeria meleagridis*, complicate molecular systematics and identification. *Int. J. Parasitol.* 43, 679–685.
- Gagnon, S., Bourbeau, D., Levesque, R.C., 1996. Secondary structures and features of the 18S, 5.8s and 26s ribosomal RNAs from the Apicomplexan parasite *Toxoplasma gondii*. *Gene* 173, 129–135.
- Gjerde, B., 2013a. Characterisation of full-length mitochondrial copies and partial nuclear copies (numts) of the cytochrome b and cytochrome c oxidase subunit I genes of *Toxoplasma gondii*, *Neospora caninum*, *Hammondia heydorni* and *Hammondia triffittae* (Apicomplexa: Sarcocystidae). *Parasitol. Res.* 112, 1493–1511.
- Gjerde, B., 2013b. 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.
- Gjerde, B., Josefson, T.D., 2015. Molecular characterisation of *Sarcocystis lutrae* n. sp. and *Toxoplasma gondii* from the musculature of two Eurasian otters (*Lutra lutra*) in Norway. *Parasitol. Res.* 114, 873–886.
- Guo, C., McDowell, I.C., Nodzinski, M., Scholtens, D.M., Allen, A.S., Lowe, W.L., Reddy, T.E., 2017. Transversions have larger regulatory effects than transitions. *BMC Genom.* 18, 394.
- Hartley, A.M., Lukyanova, N., Zhang, Y., Cabrera-Orefice, A., Arnold, S., Meunier, B., Pinotsis, N., Maréchal, A., 2019. Structure of yeast cytochrome c oxidase in a supercomplex with cytochrome bc1. *Nat. Struct. Mol. Biol.* 26, 78–83.
- Hebert, P.D.N., Cywinska, A., Ball, S.L., deWaard, J.R., 2003. Biological identifications through DNA barcodes. *Proc. Roy. Soc. Lond. B* 270, 313–321.
- Henntonen, H., Buchy, P., Suputtamongkol, Y., Jittapalpong, S., Herbreteau, V., Laakkonen, J., Chaval, Y., Galan, M., Dobigny, G., Charbonnel, N., Michaux, J., 2008. Recent discoveries of new hantaviruses widen their range and question their origins. *Ann. N. Y. Acad. Sci.* 1149, 84–89.
- Hinnebusch, A.G., 2017. Structural insights into the mechanism of scanning and start codon recognition in eukaryotic translation initiation. *Trends Biochem. Sci.* 42, 589–611.
- Hu, J.J., Meng, Y., Guo, Y.M., Liao, J.Y., Song, J.L., 2012. Completion of the life cycle of *Sarcocystis zuoi*, a parasite from the Norway rat, *Rattus norvegicus*. *J. Parasitol.* 98, 550–553.
- Hu, J.J., Liu, T.-T., Liu, Q., Esch, G.W., Chen, J.-Q., 2015. *Sarcocystis clethrionomyelaphis* Matuschka, 1986 (Apicomplexa: Sarcocystidae) infecting the large oriental vole *Eothenomys miletus* (Thomas) (Cricetidae: murinae) and its phylogenetic relationships with other species of *Sarcocystis* Lankester, 1882. *Syst. Parasitol.* 91, 273–279.
- Hu, J.J., Sun, J., Guo, Y., Zeng, H., Zhang, Y., Tao, J., 2022. Infection of the asian gray shrew *Crocodyra attenuata* (Insectivora: soricidae) with *Sarcocystis attenuati* n. sp. (Apicomplexa: Sarcocystidae) in China. *Parasites Vectors* 15, 13.
- Huang, H., Karbstein, K., 2021. Assembly factors chaperone ribosomal RNA folding by isolating helical junctions that are prone to misfolding. *Proc. Natl. Acad. Sci. USA* 118, e2101164118, 25.
- Jäkel, T., 1995. Cyclic transmission of *Sarcocystis gerbilliechis* n. sp. by the Arabian saw-scaled viper, *Echis coloratus*, to rodents of the subfamily Gerbillinae. *J. Parasitol.* 81, 626–631.
- Jäkel, T., Khoprasert, Y., Sorger, I., Kliemt, D., Seehabutr, V., Suasa-ard, K., Hongnark, S., 1997. Sarcosporidiosis in rodents from Thailand. *J. Wildl. Dis.* 33, 860–867.
- Johnson, M., Zaretskaya, I., Raytselis, Y., Merezuk, Y., McGinnis, S., Madden, T.L., 2008. NCBI BLAST: a better web interface. *Nucleic Acids Res.* 36, W5–W9.
- Kan, S.P., 1979. Ultrastructure of the cyst wall of *Sarcocystis* spp. from some rodents in Malaysia. *Int. J. Parasitol.* 9, 475–480.
- Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33, 1870–1874.
- Kumar, S., Stecher, G., Suleski, M., Heddes, S.B., 2017. TimeTree: a resource for timelines, timetrees, and divergence times. *Mol. Biol. Evol.* 34, 1812–1819.
- Kvicerova, J., Hyspa, V., 2013. Host-parasite incongruences in rodent *Eimeria* suggest specific role of adaptation rather than cophylogeny in maintenance of host specificity. *PLoS One* 8, e63601.
- Lau, Y.L., Chang, P.Y., Subramaniam, V., Ng, Y.H., Mahmud, R., Ahmad, A.F., Fong, M. Y., 2013. Genetic assemblage of *Sarcocystis* spp. in Malaysian snakes. *Parasites Vectors* 6, 257.
- Lee, F.C.H., 2019. Finding *Sarcocystis* spp. on the Tioman Island: 28S rRNA gene next-generation sequencing reveals nine new *Sarcocystis* species. *J. Water Health* 17, 416–427.
- Maddison, W.P., Knowles, L.L., 2006. Inferring phylogeny despite incomplete lineage sorting. *Syst. Biol.* 55, 21–30.
- Maghool, S., Ryan, M.T., Maher, M.J., 2020. What role does COA6 play in cytochrome C oxidase biogenesis: a metallochaperone or thiol oxidoreductase, or both? *Int. J. Mol. Sci.* 21, 6983.
- Mathur, V., Kwong, W.K., Husnik, F., Irwin, N.A.T., Kristmundsson, A., Gestal, C., Freeman, M., Keeling, P.J., 2020. Phylogenomics identifies a new major subgroup of Apicomplexans, Marosporida class nov., with extreme apicoplast genome reduction. *Genome Biol. Evol.* 13 <https://doi.org/10.1093/gbe/evaa1244>.
- Matuschka, F.-R., 1986. *Sarcocystis clethrionomyelaphis* n. sp. from snakes of the genus *Elaphe* and different voles of the family Arvicolidae. *J. Parasitol.* 72, 226–231.
- McAllister, C.T., Upton, S.J., Trauth, S.E., Dixon, J.R., 1995. Coccidian parasites (Apicomplexa) from snakes in the southern and southwestern United States: new host and geographic records. *J. Parasitol.* 81, 63–68.
- Megia-Palma, R., Martínez, J., Acevedo, I., Martín, J., García-Roa, R., Ortega, J., Pese-Fernández, M., Albaladejo, G., Cooper, R.D., Paranjpe, D.A., Sinervo, B.R., Merino, S., 2015. Phylogeny of the reptilian *Eimeria*: are *Choleoimeria* and *Acroimeria* valid generic names? *Zool. Scripta* 44, 684–692.
- Morrison, D.A., 2009. Evolution of the Apicomplexa: where are we now? *Trends Parasitol.* 25, 375–382.
- Morrison, D.A., Bornstein, S., Thebo, P., Wernery, U., Kinne, J., Mattsson, J.G., 2004. The current status of the small subunit rRNA phylogeny of the coccidia (Sporozoa). *Int. J. Parasitol.* 34, 501–514.
- Mugridge, N.B., Morrison, D.A., Jäkel, T., Heckerth, A.R., Tenter, A.M., Johnson, A.M., 2000. Effects of sequence alignment and structural domains of ribosomal DNA on phylogeny reconstruction for the protozoan family Sarcocystidae. *Mol. Biol. Evol.* 17, 1842–1853.
- Munday, B.L., Mason, R.W., 1980. *Sarcocystis* and related organisms in Australian wildlife III. *Sarcocystis murinotectis* sp.n. life cycle in rats (*Rattus*, *Pseudomys* and *Mastomys* spp.) and tiger snakes (*Notechis ater*). *J. Wildl. Dis.* 16, 83–87.
- Munday, B.L., Mason, R.W., Hartley, W.J., Presidente, P.J.A., Obendorf, D., 1978. *Sarcocystis* and related organisms in Australian Wildlife: I. Survey findings in mammals. *J. Wildl. Dis.* 14, 417–433.
- Munday, B.L., Hartley, W.J., Harrigan, K.E., Presidente, P.J.A., Obendorf, D.L., 1979. *Sarcocystis* and related organisms in Australian wildlife: II. Survey findings in reptiles, amphibians and fish. *J. Wildl. Dis.* 15, 57–73.
- Namasivayam, S., 2015. Repetitive DNA and Nuclear Integrants of Organellar DNA Shape the Evolution of Coccidian Genomes. The University of Georgia, Athens, Georgia.
- Namasivayam, S., Baptista, R.P., Xiao, W., Hall, E.M., Doggett, J.S., Troell, K., Kissinger, J.C., 2021. A novel fragmented mitochondrial genome in the protist pathogen *Toxoplasma gondii* and related tissue coccidia. *Genome Res.* 31, 852–865.
- Nawrocki, E.P., Kolbe, D.L., Eddy, S.R., 2009. Infernal 1.0: inference of RNA alignments. *Bioinformatics* 25, 1335–1337.
- Norman, J.A., Blackmore, C.J., Rourke, M., Christidis, L., 2014. Effects of mitochondrial DNA rate variation on reconstruction of Pleistocene demographic history in a social avian species, *Pomastotomus superciliosus*. *PLoS One* 9, e106267.
- Obornik, M., 2020. Photoparasitism as an intermediate state in the evolution of apicomplexan parasites. *Trends Parasitol.* 36, 727–734.
- Odening, K., 1998. The present state of species-systematics in *Sarcocystis* lankester, 1882 (protista, sporozoa, coccidia). *Syst. Parasitol.* 41, 209–233.
- Ortega Pérez, P., Wibbelt, G., Brinkmann, A., Galindo Puentes, J.A., Tuh, F.Y.Y., Lakim, M.B., Nitsche, A., Wells, K., Jäkel, T., 2020. Description of *Sarcocystis scandentiborneensis* sp. nov. from treeshrews (*Tupaia minor*, *T. tana*) in northern Borneo with annotations on the utility of COI and 18S rDNA sequences for species delineation. *Int. J. Parasitol. Parasites Wildl.* 12, 220–231.
- O'Donoghue, P.J., Watts, C.H.S., Dixon, B.R., 1987. Ultrastructure of *Sarcocystis* spp. (Protozoa: Apicomplexa) in rodents from North Sulawesi and West Java, Indonesia. *J. Wildl. Dis.* 23, 225–232.
- Pan, J., Ma, C., Huang, Z., Ye, Y., Zeng, H., Deng, S., Hu, J., Tao, J., 2020. Morphological and molecular characterization of *Sarcocystis wenzeli* in chickens (*Gallus gallus*) in China. *Parasites Vectors* 13, 512.
- Paperna, I., Peh, K.S.-H., Koh, L.P., Sodhi, N.S., 2004. Factors affecting *Sarcocystis* infection of rats on small tropical island. *Ecol. Res.* 19, 475–483.
- Pentinsaari, M., Salmela, H., Mutanen, M., Roslin, T., 2016. Molecular evolution of a widely adopted taxonomic marker (COI) across the animal tree of life. *Sci. Rep.* 6, 35275.
- Pisarev, A.V., Kolupaeva, V.G., Yusupov, M.M., Hellen, C.U.T., Pestova, T.V., 2008. Ribosomal position and contacts of mRNA in eukaryotic translation initiation complexes. *EMBO J.* 27, 1609–1621.
- Poulsen, C.S., Stensvold, C.R., 2014. Current status of epidemiology and diagnosis of human sarcocystosis. *J. Clin. Microbiol.* 52, 3524–3530.
- Prakas, P., Butkauskas, D., 2012. Protozoan parasites from genus *Sarcocystis* and their investigations in Lithuania. *Ekologija* 58, 45–58.
- Rampersad, S.N., 2014. ITS1, 5.8S and ITS2 secondary structure modelling for intra-specific differentiation among species of the *Colletotrichum gloeosporioides* sensu lato species complex. *SpringerPlus* 3, 684. <https://doi.org/10.1186/2193-1801-3-684>.
- Roberts, T.E., Lanier, H.C., Sargis, E.J., Olson, L.E., 2011. Molecular phylogeny of tree shrews (Mammalia: scandentia) and the timescale of diversification in Southeast Asia. *Mol. Phylogenet. Evol.* 60, 358–372.
- Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., Huelsenbeck, J.P., 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* 61, 539–542.
- Rooney, A.P., 2004. Mechanisms underlying the evolution and maintenance of functionally heterogeneous 18S rRNA genes in Apicomplexans. *Mol. Biol. Evol.* 21, 1704–1711.
- Rossoli, J.V.B., Moré, G., Soto-Cabrera, A., Moore, D.P., Morrell, E.L., Pedrana, J., Scioli, M.V., Campero, L.M., Basso, W., Hecker, Y.P., Scioscia, N.P., 2023. Identification of *Sarcocystis* Spp. In Synanthropic (Muridae) and Wild (Cricetidae) Rodents from Argentina. <https://doi.org/10.21203/rs.3.rs-3167171/v1>. Preprint.
- Salomaki, E.D., Terpis, K.X., Rueckert, S., Kotyk, M., Varadinová, Z.K., Čepička, I., Lane, C.E., Kolisko, M., 2021. Gregarine single-cell transcriptomics reveals differential mitochondrial remodeling and adaptation in apicomplexans. *BMC Biol.* 19, 77.
- Santiago-Alarcon, D., Rodríguez-Ferraro, A., Parker, P.G., Ricklefs, R.E., 2014. Different meal, same flavor: cospeciation and host switching of haemosporidian parasites in some non-passerine birds. *Parasites Vectors* 7, 286.
- Schluter, D., 2000. *The Ecology of Adaptive Radiation*. Oxford University Press, Oxford.
- Šlapeta, J.R., Modry, D., Votýpka, J., Jírku, M., Lukes, J., Koudela, B., 2003. Evolutionary relationships among cyst-forming coccidia *Sarcocystis* spp. (Alveolata:

- Apicomplexa: Coccidia) in endemic African tree vipers and perspective for evolution of heteroxenous life cycle. *Mol. Phylogenet. Evol.* 27, 464–475.
- Song, J., Lan, Z., Kohn, M.H., 2014. Mitochondrial DNA phylogeography of the Norway rat. *PLoS One* 9, e88425.
- Spahn, C.M.T., Beckmann, R., Eswar, N., Penczek, P.A., Sali, A., Blobel, G., Frank, J., 2001. Structure of the 80S ribosome from *Saccharomyces cerevisiae*—tRNA-ribosome and subunit-subunit interactions. *Cell* 107, 373–386.
- Srinivasan, S., Avadhani, N.G., 2012. Cytochrome c oxidase dysfunction in oxidative stress. *Free Radic. Biol. Med.* 53, 1252–1263.
- Studer, G., Tauriello, G., Bienert, S., Biasini, M., Johner, N., Schwede, T., 2021. ProMod3 - a versatile homology modelling toolbox. *PLOS Comp. Biol.* 17, e1008667.
- Tajima, F., 1993. Simple methods for testing molecular clock hypothesis. *Genetics* 135, 599–607.
- Timón-Gómez, A., Nývltová, E., Abriata, L.A., Vila, A.J., Hosler, J., Barrientos, A., 2018. Mitochondrial cytochrome c oxidase biogenesis: recent developments. *Semin. Cell Dev. Biol.* 76, 163–178.
- Tommaso, P.D., Moretti, S., Xenarios, I., Orobitg, M., Montanyola, A., Chang, J.-M., Taly, J.-F., Notredame, C., 2011. T-Coffee: a web server for the multiple sequence alignment of protein and RNA sequences using structural information and homology extension. *Nucleic Acids Res.* 39, W13–W17. Web Server issue.
- Uetz, P., Freed, P., Aguilar, R., Reyes, F., Hošek, J. (Eds.), 2023. *The Reptile Database*. <http://www.reptile-database.org>. (Accessed 24 July 2023).
- Upton, S.J., 2000. Suborder eimeriorina leger, 1911. In: Lee, L.L., Leedale, G.F., Bradbury, P. (Eds.), *The Illustrated Guide to the Protozoa*, second ed., vol. 1. Allen Press Inc., Lawrence, Kansas, pp. 318–369.
- Upton, S.J., McAllister, C.T., Trauth, S.E., Bibb, S.K., 1992. Description of two new species of coccidia (Apicomplexa: eimeriorina) from flat-headed snakes, *Tantilla gracilis* (Serpentes: Colubridae) and reclassification of misnomer species within the genera *Isospora* and *Sarcocystis* from snakes. *Trans. Am. Microsc. Soc.* 111, 50–60.
- Verma, S.K., Lindsay, D.S., Mowery, J.D., Rosenthal, B.M., Dubey, J.P., 2017. *Sarcocystis pantherophisi* n. sp., from the eastern rat snakes (*Pantherophis alleghaniensis*) as definitive hosts and interferon gamma gene knockout mice as experimental intermediate hosts. *J. Parasitol.* 102, 547–554.
- Verneau, O., Catzeffis, F., Furano, A., 1998. Determining and dating recent rodent speciation events by using L1 (LINE-1) retrotransposons. *Proc. Natl. Acad. Sci. USA* 95, 11284–11289.
- Wan, X., Wang, W., Liu, J., Tong, T., 2014. Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. *BMC Med. Res. Methodol.* 14, 135.
- Wassermann, M., Raisch, L., Lyons, J.A., Natusch, D.J.D., Richter, S., Wirth, M., Preeprem, P., Khoprasert, Y., Ginting, S., Mackenstedt, U., Jäkel, T., 2017. Examination of *Sarcocystis* spp. of giant snakes from Australia and Southeast Asia confirms presence of a known pathogen - *Sarcocystis nesbitti*. *PLoS One* 12, e0187984.
- Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F. T., de Beer, T.A., Rempfer, C., Bordoli, L., Lepore, R., 2018. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* 46 (W1), W296–W303.
- Wathanakaiwan, V., Sukmak, M., Hamarit, K., Kaolim, N., Wajjwalku, W., Munagkram, Y., 2017. Molecular characterization of the ribosomal DNA unit of *Sarcocystis singaporensis*, *Sarcocystis zamani* and *Sarcocystis zuoi* from rodents in Thailand. *J. Vet. Med. Sci.* 79, 1412–1418.
- Woolhouse, M.E.J., Haydon, D.T., Antia, R., 2005. Emerging pathogens: the epidemiology and evolution of species jumps. *Trends Ecol. Evol.* 20, 238–244.
- Yilmaz, P., Parfrey, L.W., Yarza, P., Gerken, J.E.P., Quast, C., Schweer, T., Peplies, J., Ludwig, W., Glöckner, F.O., 2014. The SILVA and "All-species living tree Project (LTP)" taxonomic frameworks. *Nucleic Acids Res.* 42, D643–D648.