

Phylogeny of fish-infecting *Calyptospora* species (Apicomplexa: Eimeriorina)

Christopher M. Whipps · John W. Fournie ·
David A. Morrison · Carlos Azevedo · Edilson Matos ·
Per Thebo · Michael L. Kent

Received: 28 November 2011 / Accepted: 9 May 2012 / Published online: 29 May 2012
© Springer-Verlag 2012

Abstract There are numerous species of apicomplexans that infect poikilothermic vertebrates, such as fishes, and possess unique morphological features that provide insight into the evolution of this important phylum of parasites. Here, the relationship of the fish-infecting *Calyptospora* species to other coccidians was investigated based on DNA sequence analysis. Genetic data from the small subunit ribosomal DNA region of the genome were obtained for three of the five nominal species in the genus *Calyptospora*. Phylogenetic analyses supported a monophyletic lineage sister to a group composed of mostly *Eimeria* species. The monophyly of *Calyptospora* species supports the validity of the family Calyptosporidae, but the sister relationship to *Eimeria* species might also suggest the Eimeriidae be expanded to encompass *Calyptospora*. The validity of the family Calyptosporidae has been questioned because it is delineated from the Eimeriidae largely based on life cycle characteristics and sporocyst morphology. In general, *Eimeria* species have a homoxenous life cycle, whereas the type species of *Calyptospora* is heteroxenous. In the absence of

experimental transmission studies, it may be difficult to demonstrate whether all *Calyptospora* species are heteroxenous. Other distinct morphological characteristics of *Calyptospora* such as an incomplete sporocyst suture, an apical opening for sporozoite release, a thin veil surrounding sporocysts supported by sporopodia, and a lack of Stieda and sub-Stieda bodies suggest there may be adequate features to delineate these taxa. Even without life cycle data for all species, the morphology and genetic data provide a means to reliably classify *Calyptospora* species. Placement in either the Calyptosporidae or Eimeriidae is discussed, along with issues relating to the phylogeny of the genus *Goussia*.

Introduction

Members of the phylum Apicomplexa are parasites of a vast range of invertebrate and vertebrate animal taxa (Levine 1988). Research has tended to focus on species that infect

C. M. Whipps (✉)
SUNY-ESF, State University of New York, College of Environmental
Science and Forestry, Environmental and Forest Biology,
246 Illick Hall, 1 Forestry Drive,
Syracuse, NY 13210, USA
e-mail: cwhipps@esf.edu

J. W. Fournie
Gulf Ecology Division,
National Health and Environmental Effects Research Laboratory,
U.S. Environmental Protection Agency,
1 Sabine Island Dr.,
Gulf Breeze, FL 32561, USA

D. A. Morrison · P. Thebo
Section for Parasitology (SWEPAR),
Department of Biomedical Sciences and Veterinary Public Health,
Swedish University of Agricultural Sciences,
75189 Uppsala, Sweden

C. Azevedo
Department of Cell Biology, Institute of Biomedical Sciences,
University of Porto,
Porto, Portugal

C. Azevedo
Zoology Department, College of Science, King Saud University,
11451 Riyadh, Saudi Arabia

E. Matos
Laboratório de Pesquisa Carlos Azevedo,
Universidade Federal Rural da Amazônia,
Belém, PA, Brazil

M. L. Kent
Center for Fish Disease Research, Department of Microbiology,
Oregon State University,
220 Nash Hall,
Corvallis, OR 97331-3404, USA

and cause disease in birds and mammals, where potential control is possible and incentive for action is strong. In other vertebrate hosts, such as fishes, there is much more of a limited understanding of apicomplexan diversity with only a few hundred described species, mostly *Eimeria* and *Goussia* species (Lom and Dyková 1992). The morphological diversity represented in other genera of fish-infecting Apicomplexa may provide clues to understanding the evolution of this important phylum.

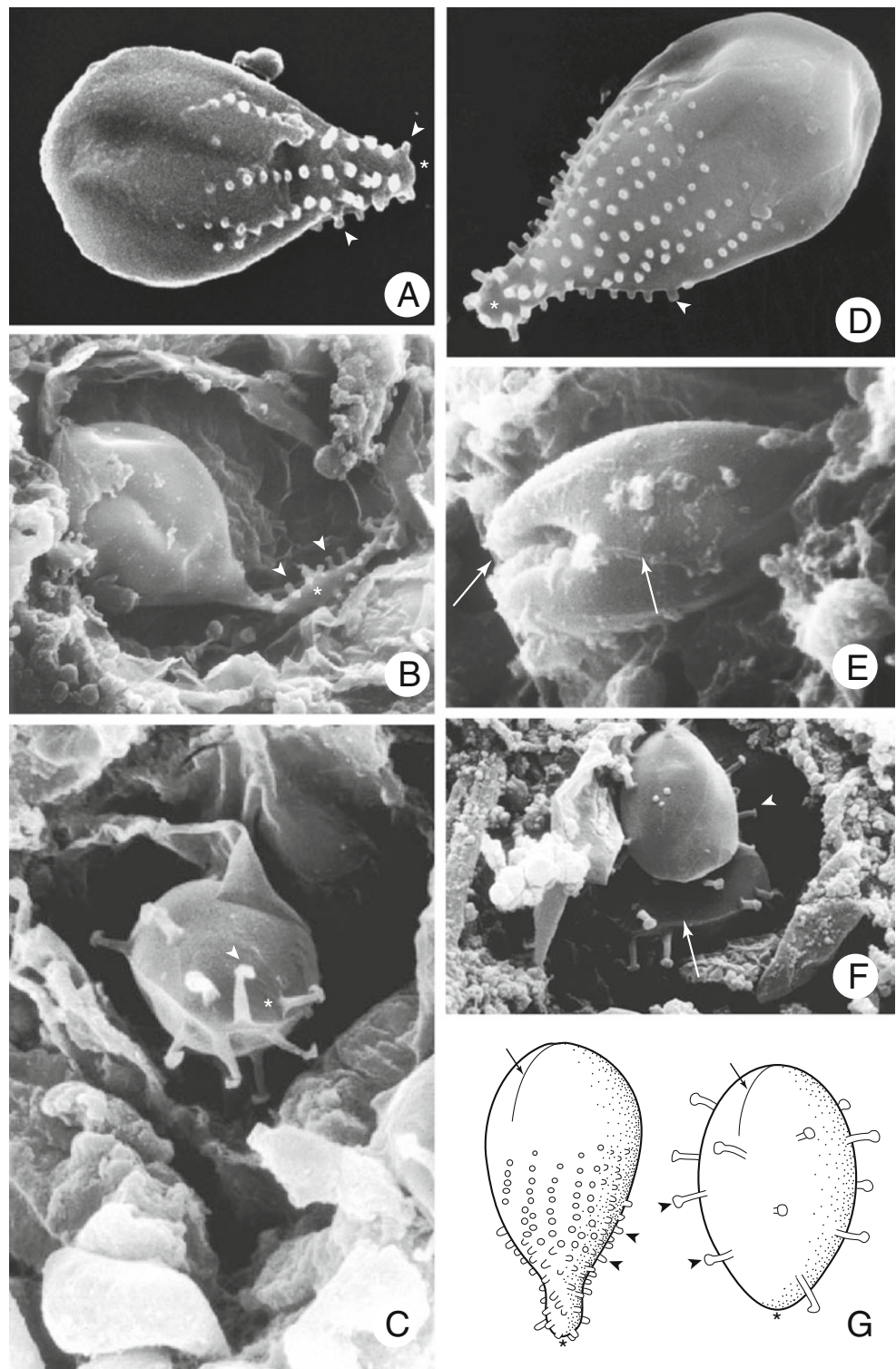
Calyptospora species are apicomplexan parasites of teleost fishes. Five species have been described to date: one occurs in estuarine fishes from coastal waters of the northern Gulf of Mexico and the southeast Atlantic, one from an inland creek in southern Mississippi, and three from freshwater systems in Brazil. An undescribed species in the arapaima is also from Brazil (de Albuquerque and de Carvalho Brasil-Sato 2010). All described species have fish definitive hosts and primarily infect the hepatic parenchyma. The impact of *Calyptospora* infections on fish species is largely unknown, but studies suggest that these can have a significant effect. Solangi and Overstreet (1980) reported that up to 85 % of the hepatic parenchyma can be replaced by oocysts in heavily infected gulf killifish. The infection provokes an inflammatory response by day 10 post-infection (p.i.) which intensifies by day 20 p.i. during gamogony. By day 30, when oocyst walls are fully formed, inflammatory infiltrates disappeared and many aggregates were encapsulated with pigmented macrophage aggregate development. Fournie and Overstreet (1993) reported that some experimentally infected fishes with abnormal coccidian development displayed more extensive host responses and pathological changes than those seen in natural definitive hosts. Gulf killifish (*Fundulus grandis*) that are infected with *Calyptospora funduli* can be severely impacted by exposure to low temperatures. Fournie (1985) showed that exposure of experimentally infected killifish to 3–4°C resulted in mortalities in infected fish, whereas uninfected fish survived. Bonar et al. (2006) associated a *Calyptospora* sp. as a contributing cause of mortality in arapaima imported to the USA from Brazil.

The genus *Calyptospora* Overstreet, Hawkins and Fournie 1984 was proposed to accommodate *Eimeria* species of fishes possessing atypical characteristics. Specifically, sporocysts (Fig. 1) with a posterior thickening or extension and sporopodia surrounded by a membranous veil lack Stieda and sub-Stieda bodies and have an apical opening for sporozoite release continuing to the opposite pole first as a tapering suture, then as a protruding ridge (Overstreet et al. 1984). The family Calyptosporidae was also proposed to accommodate *Calyptospora* and *Goussia*, distinguishing these genera from the Eimeriidae in having species with a heteroxenous life cycle with an invertebrate intermediate host. Indeed, the type species *C. funduli* has a heteroxenous life cycle involving the

grass shrimp, *Palaemonetes pugio* (Fournie and Overstreet 1983). Sporozoites develop in the invertebrate digestive tract and do occur intracellularly in basal cells (Fournie et al. 2000). A two-host life cycle, however, has yet to be demonstrated for other species of *Calyptospora*, weakening the status of the family Calyptosporidae based primarily on that characteristic. Fournie et al. (1985) did observe numerous motile sporozoites free in the intestinal contents of experimentally infected *Palaemonetes kadiakensis* and suggested that benthic crustaceans are likely candidates for intermediate hosts in *Calyptospora empristica* because the fish-host must apparently die before oocysts can be dispersed. Still, the validity of the Calyptosporidae based on a heteroxenous life cycle has been questioned primarily because of the lack of asexual multiplication in the intermediate host (Lom and Dyková 1992). However, for *C. funduli*, an intermediate host is required for transmission (Fournie and Overstreet 1983) and a developmental period of about 5 days in grass shrimp is necessary to become infective to killifishes. Fournie et al. (2000) subsequently demonstrated that sporozoites do occur within gut basal cells and suggested that development into an infective sporozoite takes place in these basal cells.

To address some of these challenges of coccidian classification, DNA sequence analysis has become commonplace for investigations into the systematics of these species (Morrison 2009). Molecular data supplement morphological analyses and help identify taxonomically significant characteristics. For example, Morrison et al. (2004) noted some phylogenetic patterns with regard to host specificity and life cycle. Species within a particular family or genus of coccidia tend to cluster by broad taxonomic groups of hosts (i.e., mammals, birds, reptiles, etc.), albeit an imperfect pattern. With regard to life cycle, heteroxenous sarcocystids form a distinct lineage from homoxenous eimeriids, again with some exceptions; notably, the genus *Iso-spora* within the Sarcocystidae is homoxenous. From the perspective of sporocyst morphology, Jirků et al. (2002) emphasized the importance of the excystation structures in coccidian taxonomy. Excystation structures may be useful for *Calyptospora* species which possess a unique membrane-covered, oblong apical opening for sporozoite excystation (Fournie et al. 1985). An incomplete suture extends posteriorly from the opening to about midway on the sporocyst wall where it becomes fused, and a prominent ridge occurs on the posterior one third of the sporocyst wall (Hawkins et al. 1983). This is reminiscent of a bivalved *Goussia* species, but *Calyptospora* is ultimately univalved like *Eimeria* because the suture is partially fused. If these features represent a stage of coccidian evolution between valved and non-valved species, then we expect this to be reflected in the phylogeny of the coccidia. Likewise, these features alone may warrant distinct familial status. It was our goal to estimate the phylogeny of *Calyptospora* using DNA sequence analysis and evaluate the taxonomic relationships of fish species to other coccidians.

Fig. 1 Scanning electron micrographs of *Calyptospora* species sporocysts indicating posterior extension or thickening (asterisks), partial suture (arrows), and sporopodia (arrowheads). **a** Lateral view of *C. spinosa* with posterior projection and sporopodia from Azevedo et al. (1993). **b** *C. empristica* showing elongate, flattened posterior extension with sporopodia from Fournie et al. (1985). **c** Posterior view of *C. funduli* showing lack of extension and characteristic sporopodia. **d** Lateral view of *C. serrasalmi* showing more developed posterior extension and numerous sporopodia from Casal et al. (2007) **e** *C. empristica* showing oblong apical opening and incomplete suture. **f** *C. funduli* showing sutural ridge on sporocyst. **g** Generalized line drawings of contrasting *Calyptospora* sporocyst types showing *C. funduli*-type morphology (right) presenting a posterior thickening as opposed to an obvious extension, and capitate sporopodia. Sporocyst with *C. serrasalmi*-like morphology (left) with pronounced posterior extension and acapitate sporopodia



Materials and methods

Collection

F. grandis (Gulf killifish) were collected in October 2005 using baited minnow traps from Big Sabine Point, Escambia County, Florida. Livers were removed from 20 specimens,

and wet mounts were examined to identify sporulated oocysts of *C. funduli*. Livers with heavy infections were cut into small pieces, placed in reagent grade 95 % ethanol, and shipped to Oregon State University for sequencing.

Crenicichla lepidota (Pike cichlid) were collected in January and February 1993 in the Amazon River about 100 km from the Atlantic coast. Small fragments of infected liver and

different isolated developmental stages of the parasite, identified as *Calyptospora spinosa*, were prepared for TEM and SEM. The white piranha, *Serrasalmus striolatus*, were collected in February 2007 in the lagoonal region near the city of Recife, and in January of the same year, the fish were collected in the Amazon River near the city of Belém, both in Brazil. The parasites found in these two specimen fish were prepared for TEM and SEM and identified as *Calyptospora serrasalmi*. In each case, a subsample of infected tissue was preserved in 95 % ethanol.

DNA extraction and sequencing

Nucleic acid was extracted from *Calyptospora* infected tissues with the Qiagen DNeasy Blood & Tissue Kit (Qiagen Inc., Valencia, CA, USA) following the manufacturer's instructions. PCR was performed using primers specific to the small subunit ribosomal DNA (ssrDNA) of apicomplexan parasites. All samples were run in 50 μ L reaction volumes containing 1.5 mM MgCl₂, 0.2 mM dNTP, 1.25 U *Taq* polymerase, 0.5 μ M of each primer, and 3 μ L of template DNA. Amplifications were performed on an MJ Research Peltier 200 Thermocycler with initial denaturation at 95°C for 3 min, followed by 35 cycles of 94°C for 30 s, 55°C for 45 s, 72°C for 1 min and 15 s, and a final extension at 72°C for 7 min. PCR products were excised from an agarose gel and purified using the QIAquick Gel Extraction Kit (Qiagen Inc., Valencia, CA, USA).

Initially, PCR primers 18E (5'-CTG GTT GAT CCT GCC AGT) and Coc2r (5'-CTT TCG CAG TAG TTC GTC) were used to generate a DNA sequence from the five prime regions of the ssrDNA. Direct sequencing of this product using the 18E primer yielded inconsistent results, thus a *Calyptospora*-specific primer, Cal1F (5'-TAC ATG CGT AAA TGG ATT TGC), was designed based on the preliminary sequence data and used with the reverse primer 18R (Whipps et al. 2003). These products were cloned using the QIAGEN PCR Cloning Plus Kit (Qiagen Inc., Valencia, CA, USA) and at least three clones from each transformation were sequenced with plasmid primers T7 and SP6, and internal primers 18K and 18I (Jirků et al. 2006). Sequencing used the AP Biotech® DYEnamic ET Terminator cycle sequencing chemistry with Thermo Sequenase II (Amersham Biosciences, Piscataway, NJ, USA) on an ABI PRISM® 377 DNA Sequencer (Applied Biosystems, Foster City, CA, USA). Sequences were assembled by visual inspection using BioEdit (Hall 1999) and verified as apicomplexan using the GenBank basic local alignment search tool.

To further address patterns of phylogenetic grouping by host taxon, we also included an avian-infecting *Eimeria* sp. sequence from woodpecker. Oocysts were obtained from two fecal samples from a nestbox at Nordens Ark wildlife park, Hunnebostrand, Sweden in January and February

2004. The nestbox contained a pair of *Dendrocopos leucotos* (white-backed woodpecker), originally from southern and central Norway in 2003. The first sample (1.5 g feces) contained 1×10^6 oocysts/g feces, and the second sample (1.1 g feces) contained 5×10^6 oocysts/g. The oocysts were recovered, sporulated (50 %), and cleaned. DNA was extracted and the ssrDNA gene sequenced as described by Morrison et al. (2004).

Sequence alignment

The alignment of the Coccidiasina initially followed the secondary-structure model of Gutell et al. (1994), obtained from the Comparative RNA Website (Cannone et al. 2002), where a model and an associated alignment of almost-complete rRNA sequences is available for the coccidia. The alignment was refined using the model for the variable area V4 of Wuyts et al. (2000) (which did not have any structure indicated in most of the initial models). The alignment was then manually optimized for best fit to the consensus structure, taking particular note of compensatory base changes where there was length variation in stems. For regions with variable lengths, hypotheses of base pairing were evaluated using the Mfold program (Zuker 2003), which folds RNA based on energy minimizations. All structure nomenclature (for stems and regions) follows Wuyts et al. (2000).

The remaining almost-complete database sequences (defined as >1,500 nt long) were then added to this initial alignment using the pairwise alignment feature of the MacClade v. 4.08 program (Maddison and Maddison 2000) and then manually optimized for fit to the consensus structure. Most of the new sequences were obtained from the SILVA database (Pruesse et al. 2007), although direct searches of the main sequence databases were also undertaken. The alignment and consensus structure were modified, if necessary, to accommodate sequences that deviated strongly from those already in the alignment.

Sequences unidentified at least to genus were excluded from the alignments, as were some sequences that proved to be difficult to align, possibly due to sequencing errors (e.g., X65508). Sequence AY618554 was also excluded, as BLAST searches showed it to be a basidiomycete fungus rather than a species of *Cystoisospora* (as claimed). For some sequences, up to ten of the terminal nucleotides were excluded when they deviated strongly from the conserved motifs in the other sequences, possibly due to sequencing errors. For FJ009244, the sequence has ~104 nucleotides missing from stems 23–29. The final alignment consisted of 454 sequences and 2,350 aligned positions. Thirteen sequences of *Cryptosporidium* were used as the outgroup (Barta and Thompson 2006), sampled from the list suggested by Slapeta (2008).

Phylogenetic analysis

Sequence locations in which positional homology was ambiguous across all taxa were identified as regions of ambiguous alignment (RAA) or regions of expansion and contraction (REC) (Gillespie 2004). The REC notably included stems 10, E10-1, 12, E23-2, E23-7, 43, 46, and 49, all of which are widely recognized as variable regions in SSU rRNA. Indeed, it proved to be impossible to derive a consensus secondary-structure model for stems E23-1 to E23-7 (in the V4 region) within the Sarcocystidae. These RAA and REC regions were excluded from some of the phylogenetic analyses. Excluded from all analyses were autapomorphies, as many of these appeared to be sequencing errors; this included long tandem repeats that occurred in some sequences (e.g., DQ111754). Two data sets were prepared for analysis. The full data set consisted of all 454 sequences and 2,350 aligned positions, excluding the RAA and REC regions. The edited data set consisted of 31 representative sequences from the clades identified in the analysis of the full data set. The RAA regions were excluded from the analysis of this data set.

Phylogenetic analyses were performed using Bayesian analyses via the MrBayes v3.2 program (Ronquist and Huelsenbeck 2003). Following Beiko et al. (2006), analyses consisted of multiple, relatively short Markov Chain Monte Carlo (MCMC) runs each with a small number of chains. The Tracer v. 1.4 program (Rambaut and Drummond 2007) was used to check for convergence of the pooled data from the multiple runs. The number of runs for each analysis was calculated so that the expected sample size values for all of the model parameters were $\gg 100$. The default priors were used in all analyses. Two analyses were performed for the full data set. The first analysis used a single partition with the GTR+G+I+covarion substitution model. This analysis consisted of 19 runs of two chains each, with 1,000,000 generations per run and a burn-in of 400,000 generations. This yielded 114,000 trees for the consensus tree. The convergence of the MCMC runs for the one-partition Bayesian analysis was fairly slow, necessitating a relatively long burn-in. The chains sampled from two likelihood islands of trees: six runs with 36,000 trees from one island and 13 runs with 78,000 trees from the other. Each MCMC run sampled one island or the other. The composition of the clades with large posterior probabilities was the same for both islands, but with some differences in branch order both between and within those clades.

The second analysis used two partitions, one for the unpaired rRNA positions with the GTR+G+I model and one for the paired (stem) positions with the doublet+G+I model. All model parameters were unlinked across the partitions, except for the tree topology. This analysis consisted of 15 runs of two chains each, with 1,000,000 generations

per run and a burn-in of 250,000 generations. The runs of the two-partition analysis converged more quickly than the one-partition analysis and were sampled from a single tree island which was rather variable, with a final standard deviation of the split frequencies=0.063. This yielded 112,500 trees for the consensus tree. Phylogenetic trees were drawn with FigTree v. 1.2.1 (Rambaut 2009).

To evaluate the relationships of the *Calyptospora* sequences to their sisters in more detail, an edited alignment containing fewer taxa was analyzed. The edited data set allowed a greater proportion of the alignment to be used because the lengths of the REC regions were greatly reduced for the subset of closely related taxa. Only a single analysis was performed for the edited data set. The setup was the same as for the second analysis described above. It consisted of five runs of two chains each, with 1,000,000 generations per run and a burn-in of 50,000 generations. This yielded 66,500 trees for the consensus tree. Phylogenetic trees were drawn with FigTree v. 1.2.1 (Rambaut 2009).

Previous analyses of the coccidia (Morrison et al. 2004; Morrison 2006, 2009) have shown that many of the potential analytical problems are associated with differences between the outgroup sequences and those of the ingroup. To assess any possible artifacts, the two-partition analysis was repeated for six runs without the outgroup sequences (yielding 45,000 trees). The topology remained mostly unchanged, except for sampling error associated with poorly supported terminal nodes, indicating that any potential differences between the outgroup and ingroup have not affected the resulting estimate of phylogeny.

Within-species genetic diversity was quantified for those species with at least three sequences, via three possible diversity measures: (1) average pairwise *p*-distance (calculated using PAUP* v. 4.0b10; Swofford 2002); (2) average nucleotide diversity (using DnaSP v. 4.50.3; Rozas et al. 2003); and (3) average pairwise patristic distance on the phylogenetic tree (using TreeEdit v. 1.0a10; Rambaut and Charleston 2001). The difference between average pairwise *p*-distance and average nucleotide diversity is that the latter incorporates a Jukes–Cantor type of correction. Patristic distance excludes the alignment sites coded as REC, ambiguous or autapomorphic.

Results

Phylogeny of the coccidia

Phylogenetic analyses of *ssrDNA* sequence from 454 apicomplexan species using a two-partition analysis (Fig. 2) place *Calyptospora* species in a distinct lineage, sister to the Eimeriidae and Lankesterellidae. This relationship was well

supported (posterior probability=0.82) in our two-partition analysis, but in the one-partition analysis (not shown), the *Calyptospora* sequences formed a clade with *Goussia neglecta*, *Goussia noelleri*, and *Goussia* sp., although this was poorly supported (posterior probabilities of 0.46 and 0.51 for the two islands).

Elsewhere in the tree, there were other minor inconsistencies between our one- and two-partition analyses involving the placement of some individual terminals. The only major branch-order differences from the one-partition analysis were the placement of *Rhytidocystis* (one sequence), which was the sister to the rest of the coccidia in the one-partition analysis (rather than the sister to *Aggregata*), and *Aggregata* (two sequences) was the sister to *Hepatozoon* (rather than the sister to *Rhytidocystis*). The *Rhytidocystis*+*Aggregata* clade was poorly supported (posterior probability=0.55) and may represent an example of long-branch attraction in the two-partition analysis (the branch lengths are not shown in Fig. 2). The alternative placement of the three *Rhytidocystis* and *Aggregata* sequences was within the Adelorina. The two-partition analysis was repeated for six runs without the possibly long-branch sequences from *Rhytidocystis* and *Aggregata* (yielding 51,000 trees), yielding the same overall tree topology (excluding sampling error), but the support for the Adelorina increased from 0.55 to 1.00. Removal of the outgroup Cryptosporiidae yielded results similar to when the outgroup was included, with the only exception of the placement of the three long-branch taxa *Rhytidocystis* and *Aggregata*, which appeared as the sister to *Hepatozoon* with posterior probability of 1.00.

In our reduced data set, the analysis focused on *Calyptospora*; the analysis runs converged quickly (standard deviation of the split frequencies=0.007) and sampled from only one tree island. The resulting consensus tree is shown in Fig. 3, along with the branch lengths. The relationship of *Calyptospora* to the three amphibian-host *Goussia* sequences was unresolved (Fig. 3), with a slight preference for *Calyptospora* as the sister to *Goussia*+Eimeriidae sequences (support, 0.56) rather than as the sister to *Goussia* alone (support, 0.43).

In both analyses, the fish-infecting *Calyptospora* species were not paired with the fish-infecting *Goussia* species (Figs. 2 and 3). Instead, two of the fish-infecting *Goussia* species were sister to all Eimeriina. The amphibian-infecting *Goussia* spp. and *Goussia metchnikovi* from fish are sister to the eimeriids. Other amphibian-infecting species are found elsewhere in the tree but tend to be clustered with basal lineages in their respective clades, i.e., *Hyaloklossia lieberkuehni*, *Eimeria ranae*. Within *Eimeria*, the bird-host species form two lineages (Fig. 2). Our *Eimeria* species from woodpecker fell within a clade of other bird-host species (the turkey-host *Eimeria meleagriditis* and *Eimeria adeneodei* and the chicken-host *Eimeria tenella*, *Eimeria necatrix*, *Eimeria mitis*,

Fig. 2 Consensus tree from the 112,500 trees sampled by the two-partition Bayesian analysis of 454 ssrDNA sequences of coccidia (Coccidiasina), with *Cryptosporidium* as the outgroup. Numbers on the branches represent the proportion of trees containing that node, and the names represent the current taxonomic classification (cf. Perkins et al. 2000). Multiple sequences from the same species have been collapsed if they form a clade, with the numbers in brackets indicating the number of sequences involved. The node labeled A or B connects the respective panels A and B to one another. Insets in each panel indicate a portion of the overall tree shown

Eimeria brunetti, *Eimeria mivati*, *Eimeria praecox*, *Eimeria maxima*, and *Eimeria acervulina*) either sister to the whole clade (support, 0.28) or nested to the clade with the exclusion of *E. meleagriditis* (support, 0.39) (as shown in Fig. 2). Bird-infecting *Eimeria reichenowi* and *Eimeria gruis* form a basal clade to all other eimeriids with the exception of the reptile-infecting *Eimeria tropidura*. Thus, within and between genera, there are some host taxon associations, but these exhibit some inconsistencies.

Intraspecific sequence variation

Calyptospora sequences formed a clade that is sister to all of the Eimeriidae species (Fig. 2), but the relationships among these sequences were not well supported by this analysis. Notably, *C. serrasalmi* did not form a monophyletic group in our two-partition analysis (Fig. 2) and monophyly was weakly supported (0.5) for *C. serrasalmi* in our reduced analysis (Fig. 3). Several of the other species shown in Fig. 2 were also intermingled with other species as paraphyletic grades, notably *Adelina bambarooniae*, *Besnoitia besnoiti*, *Sarcocystis hirsuta*, *Sarcocystis rangi*, *Toxoplasma gondii*, *Neospora caninum*, and the chicken-host *Eimeria* species (particularly *E. tenella*). The chicken-host *Eimeria* species generally showed poor resolution of their relationships.

Sequencing of *Calyptospora* species ssrDNA identified multiple alleles for all samples, at least three of which were sequenced following amplification and cloning. No two clones were identical, with sequence dissimilarities of: *C. funduli*, 0.4–0.6 %; *C. serrasalmi*, 0.3–0.6 %; and *C. spinosa*, 1.7–4.2 %. Substitutions account for most of these sequence differences, but clones from *C. spinosa* also contained ten insertion–deletions (indels) based on our alignment parameters. The clone sequences of *C. spinosa* and *C. funduli* were well supported as species (posterior probability=1.00) but those of *C. serrasalmi* were not (posterior probability=0.50). This apparent paraphyly was attributed to a single alignment position in which the “outlier” clone possessed a thymine (T) residue at position 1281 of the alignment. The other *C. serrasalmi* clones possessed an adenine (A), along with *C. funduli*, while *C. spinosa* had a guanine (G) and the remainder of the 400+ sequences from



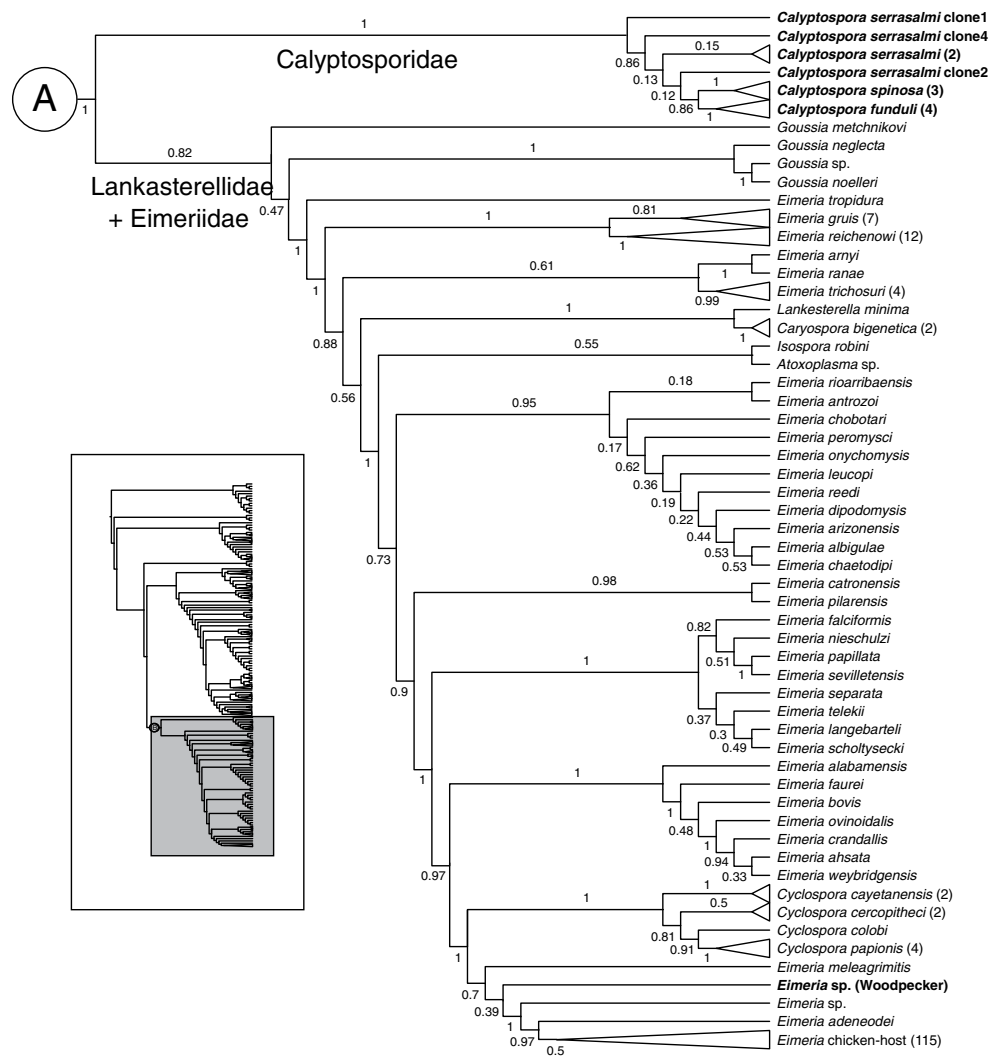


Fig. 2 (continued)

other apicomplexans had a T or rarely a cytosine (C). The relationship among the three species was unambiguous, with support of 0.80 for a sister relationship of *C. spinosa*+*C. funduli* (Fig. 3), but only 0.12 for *C. spinosa*+*C. serrasalmi* and 0.08 for *C. serrasalmi*+*C. funduli*. A comparison of 26 species (268 sequences) suitable for an analysis of genetic diversity indicated little difference between the values of the pairwise *p*-distance and the average nucleotide diversity and indicated a positive relationship between nucleotide diversity and patristic distance (Fig. 4). *C. spinosa* had appreciably more diversity than most of the analyzed species, along with *E. maxima* and *Sarcocystis singaporensis*, whereas *C. serrasalmi* and *C. funduli* were quite typical. The *E. maxima* data show distinct evidence of two groups on the tree (Fig. 2), so that the distances express both between- and within-group differences. *S. singaporensis* had extreme values, as there were three diverse sequences plus one very different one.

Discussion

Phylogenetic analysis of *Calyptospora* species *ssrDNA* reveals a monophyletic taxon distinct from other coccidian species. The *Calyptospora* clade did not form a monophyletic grouping with *Goussia* species nor *Eimeria* species (Fig. 2). There was strong support for a sister relationship between *Calyptospora* and amphibian *Goussia*+*Eimeria*, suggesting a divergence of *Calyptospora* species prior to the *Goussia*+*Eimeria* split. This warrants some careful consideration of the evolution of these apicomplexan genera and their classification.

In the context of excystation structures, *Calyptospora* appears to represent a unique loss of the valved sporocyst form. Eimeriids which possess a Stieda or sub-Stieda body appear to have evolved independently as there are two-valved *Goussia* species at the base of the clade composed mostly of *Eimeria* species. Thus, one could propose the following scenario for the evolution of these lineages. A

amphibian-infecting *Goussia* goes to support a general trend of close host–parasite co-phylogeny, at least as regards the major taxonomic groupings of both the definitive and intermediate hosts, within certain coccidian lineages. This host clustering was observed for bird-host *Eimeria* when we included a previously unpublished *Eimeria* sequence from woodpecker in our analysis (Fig. 2) and has been suggested for other *Eimeria* and *Sarcocystis* (Holmdahl et al. 1999; Slapeta et al. 2003; Morrison et al. 2004; Dahlgren et al. 2008; Dahlgren and Gjerde 2009).

The basics of the current classification of the Coccidia are supported by our analyses. In the two-partition analysis (Fig. 2), the phylogenetic tree roughly follows the current classification of the species (see Perkins et al. 2000). The main exceptions are the placement of *Rhytidocystis* and *Aggregata*, likely due to long-branch attraction as previously noted by Morrison (2009). It is not immediately obvious how to deal with this situation, which might be exacerbated by GC biases, and therefore, most previous analyses have simply deleted the problematic taxa. Our analyses show that deleting these taxa did not change the rest of the phylogenetic tree, and so it is only the placement of the long-branch taxa that remains ambiguous. *Aggregata* is currently placed in a separate family of the Eimeriorina rather than in the Adeleorina. *Rhytidocystis* is usually placed in a separate order, Agamococcidiorida, to the rest of the coccidia sampled, which are all from the Eucoccidiorida (Perkins et al. 2000). The other notable departures from traditional classification are the position of *Lankesterella minima* (family Lankesterellidae) within the Eimeriidae and the relationship of the fish-host *Goussia* species as the sister to the rest of the Eimeriorina, rather than within the Eimeriidae.

Within the *Calyptospora* clade (Fig. 3), the ssrDNA sequences obtained for each species exhibited some variation. However, each species was monophyletic (Fig. 3). Within-species sequence variation has also been reported for *Sarcocystis* species (Slapeta et al. 2002) and may be explained by a mixed infection of two closely related species or the presence of paralogous copies of the ssrDNA. Our analyses also detected potential problems with the monophyly of some of those *Sarcocystis* species with multiple isolates from cervids. In particular, the *Sarcocystis hjorti* and *S. rangi* sequences were intermixed, along with the *Sarcocystis cruzi* sequence. Also, the *Sarcocystis rangiferi* sequences were mixed with *S. sinensis* and the *S. tarandi* sequences were mixed with an unnamed species isolated from Sika deer. The latter situation was also noted by Dahlgren and Gjerde (2010). These species have all been described from both morphological and molecular data (Dahlgren and Gjerde 2007, 2010), but the ssrDNA sequences differed by much less than 5 % for these species pairs. Either the ssrDNA sequences are not useful for molecular characterization of these particular species or their validity as separate species is questionable. The lack of monophyly

for species of *Toxoplasma*, *Neospora*, and *Hammondia* was ascribed by Morrison (2005) to paralogy, and our analysis here matches the results of his network analysis. The ribosomal DNA array is also known to occur in multiple copies and in different locations within the genome. Future molecular studies on *Calyptospora* species might include the sequencing of many more clones or hybridization analysis to identify chromosomal locations of the rDNA to determine the mechanism that would explain our observations. We do not believe that our results are an artifact of mixed infections because all samples tested exhibited this same characteristic. This ssrDNA sequence variation may represent yet another unifying feature of the Calyptosporidae but does present challenges for barcoding, i.e., cloning and sequencing several clones.

As a general pattern, life cycles of the sarcocystids are heteroxenous whereas eimeriids are homoxenous. As indicated in the estimated phylogeny here (Fig. 2) and in earlier studies (Barta 2001; Morrison et al. 2004), there are exceptions to this homoxenous–heteroxenous division. This suggests gains and losses of a complex life cycle throughout the coccidian tree of life, making the placement of the Calyptosporidae based on life cycle alone difficult. In proposing the family, Overstreet et al. (1984) suggested that the Calyptosporidae are ancestral to both the Eimeriidae and Sarcocystidae because of a heteroxenous life cycle. A two-host life cycle has been demonstrated for the type species *C. funduli* (Fournie and Overstreet 1983; Fournie et al. 2000) but not for the other four species: *C. empristica* (Fournie et al. 1985), *C. serrasalmi* (Cheung et al. 1986), *C. spinosa* (Azevedo et al. 1993), and *Calyptospora tucunarensis* (Békési and Molnár 1991). Another fish parasite *Goussia carpelli*, which would have at one time been included in the Calyptosporidae, can be transmitted to naïve fish directly fecal orally as well as through an intermediate host (*Tubifex tubifex* worms), but reinfection can only be achieved via an intermediate host (Steinhagen and Körting 1990). Thus, a purely homoxenous or heteroxenous life cycle may not always be applicable, creating further challenges for classification. As an aside, since Bonar et al. (2006) evaluated new histologic sections of liver infected with *C. tucunarensis* and reported the presence of numerous sporopodia on the posterior end of the sporocyst, the validity of this species is questionable. Because transmission studies would be required for familial classification, but descriptions are typically based on morphology, the status of these species could potentially remain uncertain in perpetuity.

The question then becomes whether the family Calyptosporidae is valid for other reasons beyond the determination of a heteroxenous life cycle. We suggest that numerous morphological features unite the members of this genus and family, as follows: sporocysts, with or without posterior extensions, covered with a thin veil supported by sporopodia, absence of a Stieda or sub-Stieda body, an apical opening for sporozoite release, and a suture that extends about one third the length of

the sporocyst wall. The sporopodia are perhaps the most distinctive feature of *Calyptospora* species. Sporopodia are projections of the sporocyst wall, vary in number among species, and support the membranous veil. We use the term in the same sense as Bonar et al. (2006), distinguishing them from the posterior extension found in all but one species (Fig. 1). *C. funduli* has a small posterior thickening and the 10–25 sporopodia all occur on the sporocyst wall; *C. empristica* has an elongate, flattened posterior extension with 25–30 lateral sporopodia; and *C. serrasalmi* and *C. spinosa* have much less developed posterior extensions and sporopodia occur on the extension and the sporocyst wall. In fact, *C. serrasalmi* has greater than 100 sporopodia on the posterior elongation and extending two thirds the length of the sporocyst. Thus, the distinctiveness of this taxon is supported by morphology and the phylogenetic analyses, but an ancestral relationship to both the Eimeriidae and Sarcocystidae was not supported.

Overstreet et al. (1984) placed both *Goussia* and *Calyptospora* within the family Calyptosporidae, but based on rDNA sequences available for six *Goussia* species, it appears that this genus represents more than one lineage (Fig. 2; Jirků et al. 2009). As such, it seems inappropriate to include *Goussia* in either the Calyptosporidae or Eimeriidae because this nominal genus clearly represents more than a single genus. If only *Calyptospora* species are included in the family, we support the validity of Calyptosporidae, with at least three lines of evidence that might distinguish it from the Eimeriidae. First, the life cycle of the type species (*C. funduli*) is heteroxenous which, at least for this species, conflicts with Levine's (1988) description of the Eimeriidae which includes homoxenous species only, accepting that there are already exceptions to this qualification as noted above. Secondly, our phylogenetic analyses support *Calyptospora* as a distinct lineage from the Eimeriidae. However, this is a sister group and expansion of the Eimeriidae could encompass these species. Lastly, Levine's (1988) description of the Eimeriidae states "sporocysts univalved, without dehiscence line." The sporocysts of *Calyptospora* are arguably univalved, but do possess a partial suture or dehiscence line. Thus, the weight of evidence suggests a distinct status for *Calyptospora*, but given existing exceptions to the description of the Eimeriidae and sister relationship of these taxa, a definitive break is not clear. As a hypothesis, one can neither reject nor accept the Calyptosporidae with absolute confidence, but we suggest maintaining this family until coccidian taxonomists can clarify the breadth of what is included in the Eimeriidae, which may include *Calyptospora* and at least some species of *Goussia*.

References

- Azevedo C, Matos P, Matos E (1993) Morphological data of *Calyptospora spinosa* n. sp. (Apicomplexa, Calyptosporidae) parasite of *Crenicichla lepidota* Heckel, 1840 (Teleostei) from Amazon River. Eur J Protistol 29:171–175
- Barta JR (2001) Molecular approaches for inferring evolutionary relationships among protistan parasites. Vet Parasitol 101:175–186
- Barta JR, Thompson RCA (2006) What is *Cryptosporidium*? Reappraising its biology and phylogenetic affinities. Trends Parasitol 22:463–468
- Beiko RG, Keith JM, Harlow TJ, Ragan MA (2006) Searching for convergence in phylogenetic Markov Chain Monte Carlo. Syst Biol 55:553–565
- Békési L, Molnár K (1991) *Calyptospora tucunarensis* n. sp. (Apicomplexa: Sporozoa) from the liver of tucunare *Cichla ocellaris* in Brazil. Syst Parasitol 18:127–132
- Bonar CJ, Poynton SL, Schulman FY, Rietcheck RL, Garner MM (2006) Hepatic *Calyptospora* sp. (Apicomplexa) infection in a wild-born, aquarium-held clutch of juvenile arapaima *Arapaima gigas* (Osteoglossidae). Dis Aquat Organ 70:81–92
- Cannone JJ, Subramanian S, Schnare MN, Collett JR, D'Souza LM, Du Y, Feng B, Lin N, Madabusi LV, Muller KM, Pande N, Shang Z, Yu N, Gutell RR (2002) The Comparative RNA Web (CRW) Site: an online database of comparative sequence and structure information for ribosomal, intron, and other RNAs. BMC Bioinformatics 3:2
- Casal G, Padovan I, Matos E, Padovan P, Matos P, Guimaraes A, Azevedo C (2007) Morphological and ultrastructural redescription of *Calyptospora serrasalmi* Cheung, Nigrelli & Ruggieri, 1986 (Apicomplexa: Calyptosporidae), a parasite found in two new host species of the genus *Serrasalmus*. Braz J Morphol Sci 24:11–16
- Cheung PJ, Nigrelli RF, Ruggieri GD (1986) *Calyptospora serrasalmi* sp. nov. (Coccidia: Calyptosporidae) from liver of the black piranha, *Serrasalmus niger* Schomburgk. J Aquat Sci 4:54–57
- Dahlgren SS, Gjerde B (2007) Genetic characterisation of six *Sarcocystis* species from reindeer (*Rangifer tarandus tarandus*) in Norway based on the small subunit rRNA gene. Vet Parasitol 146:204–213
- Dahlgren SS, Gjerde B (2009) *Sarcocystis* in Norwegian roe deer (*Capreolus capreolus*): molecular and morphological identification of *Sarcocystis oviformis* n. sp. and *Sarcocystis gracilis* and their phylogenetic relationship with other *Sarcocystis* species. Parasitol Res 104:993–1003
- Dahlgren SS, Gjerde B (2010) Molecular characterization of five *Sarcocystis* species in red deer (*Cervus elaphus*), including *Sarcocystis hjorti* n. sp., reveals that these species are not intermediate host specific. Parasitology 137:815–840
- Dahlgren SS, Gouveia-Oliveira R, Gjerde B (2008) Phylogenetic relationships between *Sarcocystis* species from reindeer and other Sarcocystidae deduced from ssu rRNA gene sequences. Vet Parasitol 151:27–35
- de Albuquerque MC, de Carvalho Brasil-Sato M (2010) First report of *Calyptospora* sp. (Apicomplexa, Calyptosporidae) in forage characid fish from the Três Marias Reservoir, São Francisco Basin, Brazil. Eur J Protistol. 46:150–152
- Fournie JW (1985) Biology of *Calyptospora funduli* (Apicomplexa) from atheriniform fishes. Doctoral dissertation, University of Mississippi, 100 pp
- Fournie JW, Overstreet RM (1993) Host specificity of *Calyptospora funduli* (Apicomplexa: Calyptosporidae) in atheriniform fishes. J Parasitol 79:720–727
- Fournie JW, Overstreet RM (1983) True intermediate hosts for *Eimeria funduli* (Apicomplexa) from estuarine fishes. J Protozool 30:672–675
- Fournie JW, Hawkins WE, Overstreet RM (1985) *Calyptospora empristica* n. sp. (Eimeriina: Calyptosporidae) from the liver of the starhead topminnow, *Fundulus notti*. J Protozool 32:542–547

- Fournie JW, Vogelbein WK, Overstreet RM, Hawkins WE (2000) Life cycle of *Calyptospora funduli* (Apicomplexa: Calyptosporidae). *J Parasitol* 86:501–505
- Gillespie JJ (2004) Characterizing regions of ambiguous alignment caused by the expansion and contraction of hairpin-stem loops in ribosomal RNA molecules. *Mol Phylogenet Evol* 33:936–943
- Gutell RR, Larsen N, Woese CR (1994) Lessons from an evolving rRNA: 16 S and 23 S rRNA structures from a comparative perspective. *Microbiol Rev* 58:10–26
- Hall TA (1999) BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp Ser* 41:95–98
- Hawkins WE, Fournie JW, Overstreet RM (1983) Organization of sporulated oocysts of *Eimeria funduli* in the gulf killifish, *Fundulus grandis*. *J Parasitol* 69:496–503
- Holmdahl OJM, Morrison DA, Ellis JT, Huong LTT (1999) Evolution of ruminant *Sarcocystis* (Sporozoa) parasites based on small subunit rDNA sequences. *Mol Phylogenet Evol* 11:27–37
- Jirků M, Modrý D, Slapeta JR, Koudela B, Lukes J (2002) The phylogeny of *Goussia* and *Choleoieimeria* (Apicomplexa; Eimeriorina) and the evolution of excystation structures in coccidia. *Protist* 153:379–390
- Jirků M, Bolek MG, Whipps CM, Janovy J, Kent ML, Modry D (2006) A new species of *Myxidium* (Myxosporae: Myxidiidae), from the western chorus frog, *Pseudacris triseriata triseriata*, and Blanchard's cricket frog, *Acris crepitans blanchardi* (Hylidae) from eastern Nebraska USA: morphology, phylogeny and critical comments on amphibian *Myxidium* taxonomy. *J Parasitol* 92:611–619
- Jirků M, Jirků M, Oborník M, Lukes J, Modry D (2009) *Goussia* Labbé, 1896 (Apicomplexa, Eimeriorina) in Amphibia: diversity, biology, molecular phylogeny and comments on the status of the genus. *Protist* 160:123–136
- Levine ND (1988) The protozoan phylum Apicomplexa, vol I. CRC Press, Boca Raton
- Lom J, Dyková I (1992) Protozoan parasites of fishes. *Developments in aquaculture and fisheries science*, 26. Elsevier, Amsterdam, 315 pp
- Maddison DR, Maddison WP (2000) MacClade: analysis of phylogeny and character evolution. Sinauer Associates, Sunderland
- Morrison DA (2005) Networks in phylogenetic analysis: new tools for population biology. *Int J Parasitol* 35:567–582
- Morrison DA (2006) Phylogenetic analyses of parasites in the new millennium. *Adv Parasitol* 63:1–124
- Morrison DA (2009) Evolution of the Apicomplexa: where are we now? *Trends Parasitol* 25:375–382
- Morrison DA, Bornstein S, Thebo P, Wernery U, Kinne J, Mattsson JG (2004) The current status of the small subunit rRNA phylogeny of the coccidia (Sporozoa). *Int J Parasitol* 34:501–514
- Overstreet RM, Hawkins WE, Fournie JW (1984) The coccidian genus *Calyptospora* n.g. and family Calyptosporidae n. fam. (Apicomplexa), with members infecting primarily fishes. *J Protozool* 31:332–339
- Perkins FO, Barta JR, Clopton RE, Peirce MA, Upton SJ (2000) Phylum Apicomplexa. In: Lee JJ, Leedale GF, Bradbury P (eds) *An illustrated guide to the protozoa*, 2nd edn. Society of Protozoologists, Lawrence, pp 190–369
- Pruesse E, Quast C, Knittel K, Fuchs B, Ludwig W, Peplies J, Glöckner FO (2007) SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucl Acids Res* 35:7188–7196
- Rambaut A (2009) FigTree: tree figure drawing tool. Institute of Evolutionary Biology, University of Edinburgh, Edinburgh, Edinburgh
- Rambaut A, Charleston M (2001) Phylogenetic tree editor. Department of Zoology, University of Oxford, Oxford
- Rambaut A, Drummond AJ (2007) Tracer: MCMC trace analysis tool. Institute of Evolutionary Biology, University of Edinburgh, Edinburgh
- Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19:1572–1574
- Rozas J, Sánchez-DelBarrio JC, Messeguer X, Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* 19:2496–2497
- Slapeta J (2008) Taxonomy of the genus *Cryptosporidium* Tyzzer 1907 (Apicomplexa): revision and checklist—iCRYPTO. Available at <http://www.vetsci.usyd.edu.au/staff/JanSlapeta>. Accessed on 3 November 2008
- Slapeta JR, Kyselova I, Richardson AO, Modry D, Lukes J (2002) Phylogeny and sequence variability of the *Sarcocystis singaporensis* Zaman and Colley, 1975, 1976 ssrDNA. *Parasitol Res* 88:810–815
- Slapeta JR, Modry D, Votypka J, Jirků M, Lukes J, Koudela B (2003) Evolutionary relationships among cyst-forming coccidia *Sarcocystis* spp. (Alveolata: Apicomplexa: Coccidea) in endemic African tree vipers and perspective for evolution of heteroxenous life cycle. *Mol Phylogenet Evol* 27:464–475
- Solangi MA, Overstreet RM (1980) Biology and pathogenesis of the coccidium *Eimeria funduli* infecting killifishes. *J Parasitol* 66:513–526
- Steinhagen D, Körting W (1990) The role of tubificid oligochaetes in the transmission of *Goussia carpelli*. *J Parasitol* 76:104–107
- Swofford DL (2002) PAUP*. Phylogenetic analysis using parsimony (*and other methods). Sinauer Associates, Sunderland
- Whipps CM, Adlard RD, Bryant MS, Lester RJG, Findlay V, Kent ML (2003) First report of three *Kudoa* species from eastern Australia: *Kudoa thyrsites* from Mahi mahi (*Coryphaena hippurus*), *Kudoa amamiensis* and *Kudoa minithyrsites* n. sp. from Sweeper (*Pempheris ypsilychnus*). *J Eukaryot Microbiol* 50:215–219
- Work TM, Rameyer RA, Takata G, Kent ML (2003) Protozoal and epitheliocystis-like infections in the introduced bluestripe snapper *Lutjanus kasmira* in Hawaii. *Dis Aquat Organ* 57:59–66
- Wuyts J, De Rijk P, Van de Peer Y, Pison G, Rousseeuw P, De Wachter R (2000) Comparative analysis of more than 3000 sequences reveals the existence of two pseudoknots in area V4 of eukaryotic small subunit ribosomal RNA. *Nucl Acids Res* 28:4698–4708
- Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucl Acids Res* 31:3406–3415