Ultrastructure and 28S rDNA Phylogeny of Two Gregarines: Cephaloidophora cf. communis and Heliospora cf. longissima with Remarks on Gregarine Morphology and Phylogenetic Analysis

Timur G. Simdyanov¹, Andrei Y. Diakin², Vladimir V. Aleoshin³

¹ Department of Invertebrate Zoology, Faculty of Biology, Lomonosov Moscow State University, Moscow, Russian Federation; ² Department of Botany and Zoology, Faculty of Science, Masaryk University, Brno, Czech Republic; ³ Belozersky Institute of Physico-Chemical Biology, Lomonosov Moscow State University, Moscow, Russian Federation

Abstract. 18S rRNA gene sequences (SSU rDNA) in gregarines are problematic for phylogenetic analysis, mainly due to artifacts related to long branch attraction (LBA). In this study, we sequenced 18S rRNA (SSU rRNA), 5.8S rRNA, and 28S rRNA (LSU rRNA) genes of two gregarine species from crustacean hosts (gregarine superfamily Cephaloidophoroidea): Cephaloidophora cf. communis from a marine cirripedian Balanus balanus (White Sea), and Heliospora cf. longissima from the freshwater amphipods, Eulimnogammarus verrucosus and E. vittatus (Lake Baikal). Phylogenetic analyses of SSU rDNA sequences failed to produce a robust tree topology, for a limited taxon sample (31 operational taxonomic units (OTU), based on 1,604 sites), while LSU (2,869 sites), and concatenated dataset based on SSU, 5.8S, and LSU (4,627 sites) produced more consistent tree topologies for the same taxon sample. Analyses testing for LBA-influence were negative, therefore we suggested that the main reason of the failed topologies in SSU rDNA analyses is insufficient data (insufficient taxon sampling and limited molecular data), rather than LBA. Possible advantages of Bayesian analyses, compared to Maximum Likelihood, and usage of LSU rDNA within the context of apicomplexan phylogenetics were discussed. One of the advantages of LSU is likely its lower rate of evolution in long-branching apicomplexans (e.g., gregarines), relative to other (non-long-branching) apicomplexans, compared to SSU rDNA. Ultrastructure of the epicytic folds was studied. There are 3 to 5 apical arcs (also known as rippled dense structures) and 2 to 5 apical filaments in the tops of the folds. This small number of the apical structures fits into morphological diversity of the epicyste in other Cephaloidophoroidea, but this is not a synapomorphy of the group because this was also detected in several unrelated gregarines. C. cf. communis was found to contain a septum between the epimerite and the protomerite, which has not been reported in other gregarines. More exact terminology, which takes into account number of body sections and septa, is proposed for morphological descriptions of trophozoites and free mature gamonts of gregarines. In accordance with this, C. cf. communis gamonts are tricystid and biseptate, whereas H. cf. longissima gamonts are tricystid and uniseptate, similar to other eugregarines.

Key words: Apicomplexa, gregarines, crustacean host, morphology, ultrastructure, LSU rDNA, molecular phylogeny.
INTRODUCTION

Gregarines are obligate, unicellular apicomplexan parasites that predominantly infect the digestive system and coelomic cavities of a broad range of marine, freshwater, and terrestrial invertebrates. The feeding and growing stages of gregarines (trophozoites) can reach a large size, compared to the majority of unicellular organisms (up to several millimeters; however, about sizes ranging from 200–500 μm are common). Many of these trophozoite stages are also complex at the ultrastructural and superficial level. In general, gregarines are monoxenous (except members of the family Porosporidae), epicellular parasites; their trophozoites have ovoid or elongated shape and are heteropolar: the cell forebody is used to attach to the host cell.

Traditionally, trophozoites have been divided into either non-segmented (monocystid) forms, or into segmented (polycystid) forms that are compartmentalized by constrictions and/or septa; the latter are further divided into dicystid, tricystid, and multisegmented types (Grassé 1953). Since 1971, all these terms have been replaced with “aseptate” and “septate” (Levine 1971), and are widely used among workers in the field (Perkins et al. 2000). However, the traditional terminology has been partially restored in the new edition of “Treatise on Zoology” (Desportes and Schrével 2013).

Septate forms (e.g. Gregarina, Cephaloidophora, Stylocephalus) belong to the largest gregarine order, the Eugregarinida (or Eugregarinorida) Léger 1900, which also comprises many aspartate representatives (e.g., Monocystis, Lecudina, and Urospora). The anterior end of the trophozoite in eugregarines, responsible for the attachment to the host cell has a different cytoplasmic structure: its cytoplasm is more translucent (hyaline), while the rest of the cell is more opaque, due to the storage carbohydrate (amylopectin) granules (Mercier et al. 1973). This part of the cell is called the “mucron” in aspartate forms, or “epimerite” in septate ones (Levine 1971). Epimerites of gregarines vary in size and shape, being large or small, prolonged or lens-like. They can bear hooks or projections, and often appear to be separated from the rest of the body by a septum.

Tricystid septate forms have a body part called the protomerite, behind the epimerite. The protomerite is followed by the deutomerite, which contains the nucleus. The protomerite and deutomerite are separated by a septum. There are several electron microscopy studies that observed the ultrastructure at the anterior end of septate gregarines. These studies demonstrated that the septum is composed primarily of tightly packed, thin fibrils. However, no such septum has been observed between the epimerite and protomerite (Grassé and Théodoridès 1959, Devauchelle 1968a, Desportes 1969, Baudoin and Ormières 1973, Hoshide 1975, Hildebrand 1976, Ormières 1977, Tronchin and Schrével 1977, Valigurová and Koudela 2005, Valigurová 2012). In Epicyclus araeoceri, separate radial microtubules, which arise from a putative remnant of the conoid situated at the base of the epimerite, form a structure referred to as the “pseudoseptum” (Ormières and Daumal 1970a, b).

In multisegmented (or “metamerized”) gregarines (e.g., Metamera and Taeniocystis), the deutomerite is subdivided by several additional septa. However, no work has been done on the ultrastructure of these septa. The presence of a septum in the protomerite/deutomerite (using only light microscopy) is considered a diagnostic character of suborders Aseptata (Aspotorina) and Septata (Septatorina) (Levine 1985, Clpton 2000). When the growth period is over, the epimerite of septate gregarines usually disappears (Grassé 1953), or is condensed (Devauchelle 1968a, Valigurová et al. 2009, Valigurová 2012), and the gregarine cells acquire a dicystid-like morphology and (commonly) start to move in the intestine lumen. Under Levine’s nomenclature (Levine 1971) this stage is called a “gamont”.

The overwhelming majority of eugregarines have gliding motility, and a specific surface structure that is composed of numerous (sometimes hundreds) longitudinal pellicular folds known as apicytic folds. Gregarines have a characteristic tegument represented by a trimembrane pellicle, which consists of the plasmalemma with a cell coat, and the inner membrane complex (IMC) formed by two closely adjacent cytomembranes (Vivier 1968, Vivier et al. 1970, Schrével et al. 1983). There are micropores in the pellicle located between the apicytic folds, or (rarely) on their sides (Desportes 1969, Simdyanov 1995). Micropores are characteristic structures of apicomplexans, visible only by electron microscope. Specifically, these are small pores on the surface of a cell leading into little invaginations of plasma membrane inside the cytoplasm. The invaginations penetrate the IMC through openings collared by electron-dense material. There are electron-dense longitudinal structures within apical parts of the apicytic folds: “apical arcs” (also called “rippled dense structures”) and “apical filaments” (also called “12 nm filaments”) (Vivier 1968, Schrével et al. 1983). The apical filaments adjoin the IMC from the inner part of the cell

Motility is necessary to find a sexual partner and to form syzygy, a sexual association of gamonts (Grassé 1953, Levine 1971, Desportes and Schrével 2013). This stage is often considered characteristic for gregarines, although, among apicomplexans, it also exists in adelid coccidians, for instance, in haemogregarines (Perkins et al. 2000). Usually, the syzygy is just a mechanical association of two mature gamonts without fusion of their cytoplasms; however, the contact can be more complex in some cases (Devauchelle 1968b, Desportes 1974, Dallas and Vegni Talluri 1988, Frolov 1991). Septate eugregarines have a “caudo-frontal” (head-to-tail) syzygy where the anterior individual is called “primit” and the posterior one “satellite” (Grassé 1953, Levine 1971). There can be more than one satellite: two, three, or more (Grassé 1953). In some gregarine species, syzygy is formed in feeding (trophozoite) and growth stages (early syzygy). In others, they emerge immediately before formation of the gametocyst (late syzygy). The general ultrastructure of the syzygy partners in septate gregarines is largely similar to that of trophozoites in the same species (Devauchelle 1968b, Desportes 1974, Philippe 1983, Dallas and Vegni Talluri 1988).

A recent system of gregarine classification has been developed by Levine throughout the 1970s and 1980s (Levine 1970, 1979, Levine et al. 1980, Levine 1984, 1985, 1988a, b) with minor modifications accepted later (Perkins et al. 2000, Adl et al. 2012). This system relies on morphological characters of trophozoites, gamonts and syzygy, as well as on certain lifecycle traits. For instance, the order Eugregarinida (=Eugregarinarorida) is divided into suborders Aseptata (=Aseptatorina) and Septata (=Septatorina). This assumes that the presence of the protomerite/deutomerite septum is a synapomorphy of the monophyletic taxon “septate eugregarines”. The suborder Septata includes four superfamilies: Porosporicae (heteroxenous, one family), Gregarinicae (early syzygies), Stenophoricae (late syzygies), and Fusionicae (gamont cytoplasms fuse in syzygy; one family).

Previous work also established other classification schemes. The most common one was the system by Grassé (1953), based on the concept of co-evolution of gregarines and their hosts. Morphological characters were of secondary importance in this system; it included no suborders or superfamilies, and the families were recognized primarily according to the taxonomic position of their hosts. The presence of the septum was not considered a major taxonomic character. For instance, the asperate gregarine family Lecudinidae (parasites of polychaetes) comprised the genus Sycia with a septum, while the family of septate gregarines, Actinocophalidae (parasites of insects), comprised Schneideria, which lacks a septum, a typical tricystid, Actinocophalus, and the multisegmented Taeniocystis. Grassé (1953) proposed that certain asperate gregarines secondarily lost the septum (“pseudomonocystid” gregarines). This was proposed for Lankesteria (now Ascogregarina) culicis. Overall, this system upheld that the septa are a labile morphologic formation that repeatedly appeared and disappeared in the evolution of gregarines. Grassé’s evolutionary hypotheses were put forth before the use of electron microscopy and molecular methods became established in protist studies.

Still, ultrastructural characters are not entirely applicable for taxonomic studies. This is due to the fact that most gregarine species have been described using light microscopy, which is generally insufficient for phylogenetic inferences (Grassé 1953, Levine 1985, Perkins et al. 2000). Moreover, electron microscopic studies of gregarines are fragmentary, and subsequently, our knowledge about different taxa contains uneven levels of detail. Therefore, the more recent use of molecular tools will start to ‘fill in the gaps’ in our knowledge of species-level diversity and evolutionary relationships between distinct and closely-related gregarine lineages.

Contemporary use of molecular phylogenetics in gregarines largely relies on the analysis of small subunit (SSU) rDNA sequences (Leander et al. 2003a, b, Leander et al. 2006, Leander 2007, Rueckert and Leander 2008, 2009, Clopton, 2009, Rueckert et al. 2010, Rueckert and Leander 2010, Rueckert et al. 2011a, b, Wakeman and Leander 2012, Rueckert et al. 2013, Wakeman and Leander 2013). However, these phylogenetic analyses show high divergence rates among of SSU rDNA sequences in many gregarine clades, mak-
ing nodal-support and the overall topology for some parts of tree susceptible to artifacts such as long branch attraction (LBA). Recent SSU rDNA analyses demonstrated a robust clade of gregarines (Cephaloidophoroidea) parasitic in crustaceans. The Cephaloidophoroidea comprises both aseptate and septate trophozoite forms (Rueckert et al. 2011b). In phylogenetic trees, this branch is extremely long and occupies an isolated position (i.e., nodal support between this clade and other gregarine clades is unresolved).

Here we present ultrastructural data on two gregarine species: Cephaloidophora cf. communis Marek-Diadi, 1908 and Heliospora cf. longissima (von Siebold in Kölliker 1848) Goodrich, 1949, as well as the results from phylogenetic analyses of large subunit (LSU) rDNA and SSU rDNA datasets.

MATERIALS AND METHODS

Collection and isolation of organisms

Both gregarine species were isolated from different crustacean hosts in marine and freshwater habitats. Cephaloidophora cf. communis was isolated from the intestine of Balanus balanus Linnaeus, 1758 (Cirripedia) collected from the White Sea (White Sea Biological Station of Moscow State University, Velikaya Salma Straight, Kandalaksha Gulf, 66°33′12″N, 33°06′17″E, Russia) in 2006. Heliospora cf. longissima was isolated from the intestines of two freshwater amphipods endemic to Lake Baikal, namely, Eulimnogammarus verrucosus Gerstfeldt, 1858, and Eulimnogammarus vittatus Dybowsky, 1874; both species were collected near the Bolshie Koty village (51°54′12″N, 105°04′30″E), Lake Baikal, Russia in 2005.

The specimens of each species were released into seawater (C. cf. communis) or saline solution (150 mM NaCl) (H. cf. longissima) by tearing apart the intestine of the hosts with fine-tipped needles under a stereomicroscope (MBS-1, LOMO, Russia). The gut material was examined under the stereomicroscope and the parasites were isolated using fine glass pipettes, rinsed three times in either filtered seawater or saline solution depending on the hosts' habitat (marine or freshwater), photographed, fixed for electron microscopy, or prepared for DNA extraction.

Light microscopy

Some micrographs of C. cf. communis and H. cf. longissima were produced using light microscopes (Karl Zeiss, Jena and LOMO, USSR) and a Nikon Coolpix 7900 camera. Differential interference contrast (DIC) light micrographs of C. cf. communis were produced with a Leica DM 2000 light microscope connected to a Leica DFC 420 camera (Leica Microsystems, Germany).

Transmission electron microscopy

Several specimens of C. cf. communis (gamonts) and H. cf. longissima (gamonts and syzygy) from E. verrucosus were prepared for transmission electron microscopy (TEM). The specimens were fixed with 2.5% (v/v) glutaraldehyde in 0.05 M cacodylate buffer (pH 7.4) containing 1.28% (w/v) NaCl in an ice bath, in the dark. Fixative was replaced with a fresh portion after 1 hour; total time for fixation was 2 hours. Fixed samples were rinsed three times with the cacodylate buffer, and post-fixed with 2% (w/v) OsO4 in cacodylate buffer (ice bath, 2 hours). After dehydration in a graded ethanol series, the fixed parasites were embedded in Epon resin using a standard procedure. Ultrathin sections obtained using an LKB-III ultramicrotome (LKB, Sweden) were contrasted with 4% (w/v) water solution of uranyl acetate (40 minutes at 37°C) and 0.04% (w/v) water solution of lead citrate (Reynolds 1963) for 20 minutes at room temperature in the dark. The sections were examined under a JEM-100B or a JEM 1011 electron microscopes (Jeol, Japan).

DNA isolation, PCR, and sequencing

Individuals of each species were isolated from the dissected host intestines, washed three times in the appropriate medium, and deposited into 1.5 ml microcentrifuge tubes: 93 individuals of C. cf. communis from B. balanus, 20 individuals of H. cf. longissima from E. verrucosus, and 25 trophozoites of H. cf. longissima from E. vittatus. All samples were subjected to alkaline lysis (Floyd et al. 2002) with modifications (Petrov et al. 2007). The lysates were used directly for PCR.

The rDNA sequences were amplified in a PCR series, following methods and primers for SSU rDNA amplification described earlier (Rueckert et al. 2011b). For the rest of the ribosomal operon, a set of overlapping fragments was obtained (Fig. 1). All fragments were amplified with an Encyclo PCR kit (Evrogen, Russia) in a total volume of 25 µl using a DNA Engine Dyad thermocycler (Bio-Rad) using the following protocol: initial denaturation at 95°C for 3 min; 40 cycles of 95°C for 30 sec, 50°C for 30 sec, and 72°C for 1.5 min; and a final extension at 72°C for 10 min. The fragments were amplified using primers d6 (5′– CGTTCTTAGTTGGTGG –3′) and 28r2 (5′– TACTTGTYBRCTATCG –3′) for fragment I, d71c (5′– gta...
caccacgcctcggttct 3′) and 28r3.2 (5′– ACTCTYRTGCTCGTCCTCACTCATATCACTCATAGTTAC
3′) for fragments II and V, and 28d5 (5′– CCGCTAAGGATACGGAATAGGTGAACAC AC–3′) and 28r11 (5′– GTCTAAGGATACGGAATAGGTGAACAC
AC–3′) for fragment III. The LSU rDNA primer sequences are based on van der Auwera et al. (1994).

PCR products of the expected sizes (~800 bp for fragment I, ~900 bp for fragment IV, and ~1,600 bp for fragments II, III, and V, also see Fig. 1) were gel isolated using a Cytokine DNA isolation kit (Cytokine, Russia) and sequenced using an ABI PRISM BigDye Terminator v. 3.1 reagent kit on an automatic sequencer Applied Biosystems 3730 DNA Analyzer. DNA sequences were preliminarily identified by BLAST analysis, including the built-in NJ-tree tool. Following this initial BLAST search, fragments were assembled with overlapping counterparts, and with matching parts of SSU rDNA sequences that were obtained earlier from the same gregarine samples (GenBank accession numbers: HQ891113.1 – HQ891115.1). The contigs of the ribosomal operons (SSU rDNA + ITS1 + 5.8S rDNA + ITS2 + LSU rDNA) were constructed for each gregarine sample (GenBank accession numbers: HQ891113.2 – HQ891115.2). Sequence accession numbers from this study are presented in the Table 1.

Table 1. List of SSU, 5.8S, and LSU rDNA sequences used in phylogenetic analyses.

<table>
<thead>
<tr>
<th>Operational taxonomic unit (OTU)</th>
<th>GenBank accession number</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SSU</td>
</tr>
<tr>
<td>Akashiwo sanguinea</td>
<td>AY831412</td>
</tr>
<tr>
<td>Alexandrium catenella</td>
<td>AY347308</td>
</tr>
<tr>
<td>Ascogregarina taiwanensis</td>
<td>EF666482</td>
</tr>
<tr>
<td>Babesia bigemina</td>
<td>The Sanger Institute Babesia bigemina genome project (<a href="http://www.sanger.ac.uk/Projects/B_bigemina/)">www.sanger.ac.uk/Projects/B_bigemina/)</a></td>
</tr>
<tr>
<td>Bigelowiella natanis</td>
<td>U02075</td>
</tr>
<tr>
<td>Chromera velia</td>
<td>DQ174731</td>
</tr>
<tr>
<td>Vitrella brassicaformis</td>
<td>HM245049</td>
</tr>
<tr>
<td>Cryptosporidium parvum</td>
<td>AF040725</td>
</tr>
<tr>
<td>Cylindrotheca closterium</td>
<td>DQ082742</td>
</tr>
<tr>
<td>Eimeria tetella</td>
<td>AF026388</td>
</tr>
<tr>
<td>Gonyaulax polyedra</td>
<td>AF377944</td>
</tr>
<tr>
<td>Gregarina niphandrodes</td>
<td>AF129882</td>
</tr>
<tr>
<td>Hyphochytrium catenoides</td>
<td>X80344</td>
</tr>
<tr>
<td>*Isospora belli</td>
<td>–</td>
</tr>
<tr>
<td>*Isospora felis</td>
<td>L76471</td>
</tr>
<tr>
<td>*Mallomonas asmundae</td>
<td>–</td>
</tr>
<tr>
<td>*Mallomonas striata</td>
<td>M87333</td>
</tr>
<tr>
<td>Paramaecum tetraurelia</td>
<td>AF149979</td>
</tr>
<tr>
<td>Perkinsus atlanticus</td>
<td>AF509333</td>
</tr>
<tr>
<td>Pfiesteria piscicida</td>
<td>AY112746</td>
</tr>
<tr>
<td>Phytophthora megasperma</td>
<td>M54938</td>
</tr>
<tr>
<td>Plasmodium berghei</td>
<td>M19712</td>
</tr>
<tr>
<td>Plasmodium falciparum</td>
<td>M19172</td>
</tr>
<tr>
<td>Procercentrum micans</td>
<td>AY803739</td>
</tr>
<tr>
<td>*Sarcocystis canis</td>
<td>–</td>
</tr>
<tr>
<td>*Sarcocystis muris</td>
<td>M64244</td>
</tr>
<tr>
<td>*Spathidium amphoriforme</td>
<td>–</td>
</tr>
<tr>
<td>*Spathidium sp.</td>
<td>Z22931</td>
</tr>
<tr>
<td>Sterkiella histrionomuscorni</td>
<td>FJ545743</td>
</tr>
<tr>
<td>Theileria annulata</td>
<td>KF429795</td>
</tr>
<tr>
<td>Theileria parva</td>
<td>L02366</td>
</tr>
<tr>
<td>Toxoplasma gondii</td>
<td>M97703</td>
</tr>
</tbody>
</table>

*The alignment included sequences composed of genes from two closely related species.
Molecular phylogenetic analyses

Three sets of alignments were used for phylogenetic analyses: SSU, LSU, and concatenated SSU, 5.8S, and LSU rDNA sequences: 1,604; 2,869, and 4,627 sites, respectively (after removing hypervariable regions and gaps). The alignments were constructed using MUSCLE 3.6 (Edgar 2004) and manually tuned with BioEdit 7.0.9.0 (Hall 1999). Ribosomal DNA sequences of heterokonts and rhizarians were used as outgroups. The taxonomic range was substantially limited by the available alveolate sequences of 5.8S and LSU rDNA. We managed to find just two gregarine LSU rDNA sequences: Gregarinina niphandrodes and Ascogregarina taiwansen-sis (Toso and Omoto 2007, Templeton et al. 2010). To enhance the whole dataset, missing gene sequences for three outgroups (heterokonts, ciliates and coccidians) were taken from closely related species (i.e., representatives of the same genus or closely related genera, which are presented in GenBank). The DNA sequences were tested to see if they localized in the same clusters (i.e. heterokonts, ciliates and coccidians), both in SSU and LSU phylogenetic analyses. Following this, the sequences were concatenated and added to ribosomal operon multigene alignment (Table 1). The use and validity of composite taxa in phylogenetic analysis has been previously discussed and applied earlier (Philippe et al. 2004, Leander and Keeling 2004, Delsuc et al. 2006, Bourlat et al. 2008, Campbell and Lapointe 2009).

For the SSU rDNA alignment, we preliminarily examined a large taxon sampling of alveolates (>300 OTUs, including majority of sequenced gregarine species), using certain features of predicted secondary structure of SSU rDNA during manual aligning, especially of problematic regions in gregarines: aligning using automatic software failed there. The resulting tree topology was compatible with previously published work (Rueckert et al. 2011b). We then removed a majority of the OTUs from the alignment, retaining the 31 OTUs for the phylogenetic analyses described below.

Maximum-likelihood (ML) analyses were performed with RAxML 7.2.8 (Stamatakis 2006) under GTR+I+Γ and GTR+Γ models, with 12 categories of discrete gamma distribution. The ML analyses included bootstrapping with 1,000 replicates and 100 independent runs of RL each. Bayesian inference (BI) analysis was conducted using MrBayes 3.2.1 program (Ronquist and Huelsenbeck 2003), under a GTR+I+Γ model. The program was set to operate using the following parameters: nst = 6, ngammacat = 12, rates = invgamma, covarion = yes; parameters of Metroplis coupling Markov chains Monte Carlo (mcmc): nchains = 4, nruns = 2, temp = 0.2, ngen = 3,000,000, samplefreq = 1,000, burninfrac = 0.5 (first 50% of 3,000 sampled trees, i.e. first 1,500 ones, were discarded in each run). Standard deviation of split frequencies averaged 0.001 in the LSU rDNA analysis and of 0.002 in the concatenated ribosomal DNA analysis.

For SSU rDNA, two Bayesian analyses were performed on the 31 OTUs alignment: one with the same parameters as above (average standard deviation was 0.03) and another one with increased numbers of generations and independent runs (nruns = 8, ngen = 7,000,000). Standard deviation of split frequencies averaged 0.007. ML analysis was performed with the parameters specified above. In addition, we performed ML analyses of reduced alignments of SSU rDNA, where non-alveolate outgroups, and long-branch forming representatives (the crustacean gregarines, Gregarinina niphandrodes, and Plasmodium spp.) were excluded; the parameters of the computations were the same as in the LSU rDNA and operon analyses. The topology, which is identical to that LSU rDNA and ribosome operon trees, was used as a constraint for branch lengths and node supports calculation in SSU rDNA trees under the same parameters. Topology tests for the 31 OTU Bayesian, ML, and user trees were performed using TREEFINDER (Jobb et al. 2004, Jobb 2011). For comparison of branch lengths in different trees and evaluation of molecular evolution rates, the measurements of the branch lengths were performed manually using scale bars.

RESULTS

Light microscopy (LM)

Solitary cells of C. cf. communis (Fig. 2A, B) demonstrated typical tricystid organization, since they are subdivided in three distinct compartments: small lenticular epimerite (according to the generic diagnosis in: Clopton 2000), protomerite, and deutomerite, which are separated from one another by two septa. The septa are more distinctly visible under DIC. Syzygies were not observed by us and therefore were not studied.

Solitary individuals and syzygy primites of H. cf. longissima (Fig. 2C, D) had similar tricystid organization with two septa, but syzygy satellites did not have an epimerite; therefore we classified them as dicystid. Unfortunately, DIC studies were impossible during the material collecting on Lake Baikal.

Transmission electron microscopy (TEM)

Cephaloidophora cf. communis (Fig. 3)

The cross-sections through the middle of the cell demonstrated a typical eugregarine tegument structure: a trimembrane pellicle (~30 nm thick) covered by a cell coat (~10 nm thick), forming numerous epicytic folds (Fig. 3A–C). Cross-sections of the folds show a finger-like shape: they have parallel lateral sides and slightly swollen rounded tops. The height of the folds significantly varies from ~400 nm in the middle of the body to ~260 nm (and probably less) closer to the anterior end of the cell, while the thickness of the folds remains the same and averages ~130 nm (tops) and ~100 nm (stems). There are three distinct apical arcs in the fold tops and two apical filaments, which are very difficult to detect; however, it is possible to discern them in some photos (Fig. 3B, C). The two apical filaments face spaces between the three apical arcs (they are not in front of apical arcs themselves). Beneath the pellicle is a layer (35 nm thick) forming crossbars at the bases of the folds (Fig. 3B). Micropores are located at
Fig. 2. Light microscopy of the gregarine studied: free individuals (gamonts) of Cephaloidophora cf. communis (A, common light microscopy; B, DIC microscopy); a free gamont (C) and a syzygy (D) of Heliospora cf. longissima. Epimerite (ep), promerite (pr), deutomerite (de), septum between poto- and deutomerite (s1), and septum between proto- and epimerite (s2) are visible.

the bottom of the grooves between the folds (Fig. 3B), without any regular pattern.

The cytoplasm is subdivided into ecto- and endoplasm. This stratification in gregarines is caused by the absence or presence of storage carbohydrate (amylopectin) grains, respectively (Vivier 1968). The thickness of the ectoplasm, free of amylopectin grains, was about 1 μm. The endoplasm contains numerous amylopectin grains, rounded shape up to 0.7 μm in diameter (Fig. 3A). The cytoplasm also contained numerous vacuoles with dense inclusions.

Longitudinal sections of the cell demonstrated two distinct septa: one between the epi- and protomerite and another between the proto- and deutomerite (Fig. 3D, E, G–I). Both septa appear to be fibrillar in nature (Fig. 3G, H, I). No contacts were observed between the septa and the pellicle (Fig. 3H). The anterior and middle parts of epimerite are filled by unusual objects, which look similar to micronemes (typically absent in gregarine trophozoites and gamonts) and tightly adjoin one another (Fig. 3F). There are pores in the epimerite pellicle previously observed by SEM (Rueckert et al. 2011b). Probably, the homogenous content of the microneme-like objects can be released through these pores (Fig. 3F). The posterior part of the epimerite cytoplasm comprises electron-translucent vacuoles (Fig. 3H). The protomerite cytoplasm has a foam-like structure (Fig. 3G, H, I), possibly formed by highly developed endoplasmic reticulum. Additionally, there are lipid drops and amylopectin grains in the middle and in the rear part of the protomerite (Fig. 3E). The number of these inclusions increased towards the posterior, and were in particularly high abundance in the deutomerite (Fig. 3D, E).

**Heliospora cf. longissima from Eulimnogammarus verrucosus** (Fig. 4).

The transversal and longitudinal sections were studied. Unfortunately, the fixation was not very successful, and these samples could not be recollected. The cross-sections demonstrated pellicular epicytic club-shaped folds with swollen tops containing about 5 apical filaments and 4–6 hardly discernible apical arcs (Fig. 4A, B). The apical filaments face spaces between the apical arcs rather than the apical arcs themselves. The height of the folds varies from ~260 to ~500 nm, depending on location of the section: closer to the posterior end of the cell the folds are significantly lower than in the middle of the body, while the thickness remains the same (~140 nm tops and ~90 nm stems). Micropores were not observed. The subjacent layer of the pellicle was well developed, however, unlike C. cf. communis it is less electron-dense; its thickness was about 45 nm. The subdivision of the cytoplasm into ecto- and endoplasm
was not observed. Large (up to 0.75 μm) carbohydrate granules come close to the subjacent layer under the pellicle.

The longitudinal sections of the two separated syzygy partners, primite and satellite, were also studied (the syzygy was fragmented during the embedding procedure). The primite demonstrates differences of cytoplasm structure between its epimerite, protomerite, and deutomerite (Fig. 4C), despite the section missing the very apex of the cell. Just basal part of epimerite is present on the section; its cytoplasm looks homogeneous and does not contain any organelles or inclusions. A distinct septum between the epi- and protomerite is apparently absent (Fig. 4C). The protomerite cytoplasm is subdivided into two zones: anterior and posterior. The anterior zone looks darker and more homogeneous; the posterior zone is lighter and apparently contains many membranous structures (Fig. 4D, E). The protomerite contains large electron-dense globules (Fig. 4C, D), which are arranged mainly near its approximate anterior and posterior borders. However, several ones may be observed in the middle part. Such globules in the deutomerite as well, and they also occur near the deuto-/protomerite interface, which may be observed as a (barely) visible, loose septum presumably consisting of thin fibrils (Fig. 4E). There are many large amylopectin granules within the deutomerite, unlike the epimerite and protomerite.

The fine structure of the satellite forebody sharply differs from that of primite. There is no epimerite. The protomerite is separated from the deutomerite by a thick granulated (or fibrillar) zone of the cytoplasm, which looks like a septum. The protomerite of the satellite contains no amylopectin grains, while the deutomerite is rich with them. The protomerite of the satellite is highly vacuolated. The vacuoles are large, electron-light, and contain large electron-dense globules (Fig. 4F, G).

**Molecular phylogeny**

ML analyses for all genes was conducted under GTR+Γ and GTR+Γ+I models. No significant changes were observed between these two models. The tree topologies were the same and the bootstrap supports varied only slightly (data not shown). Accordingly, one model, GTR+Γ+I, was used in the Bayesian analyses. The results presented below were obtained using this model only.

**Analyses of LSU rDNA and ribosomal operon (Fig. 5)**

Internal transcribed spacers (ITS1 and ITS2), 5.8S, and LSU rDNA sequences of both *Heliospora cf. longissima* samples (from *Eulimnogammarus verrucosus* and *E. vittatus*) were almost identical throughout their length (two substitutions out of the 2,833 included nucleotide sites for each sample). This confirms that the same gregarine species was present in two crustacean hosts. The phylogenetic trees were constructed using Bayesian Inference (BI) and Maximum Likelihood (ML) with bootstrapping for two datasets containing the same 31 OTUs. All the constructed trees had identical topologies, independent of the inference method. The apicomplexan clade comprised chromerids and sporozoans, the latter included the clades of cryptosporidia+gregarines and coccidians+hematozoans. The trees contain the long branches of *Plasmodium* spp., *Gregarina niphandrodes*, ...
and the two gregarines in this study parasitizing crustacean hosts, *Cephaloidophora* cf. *communis* and *Heliospora* cf. *longissima*.

In the LSU tree (Fig. 5A), most alveolate clades were highly supported in both Bayesian (PP = 0.95–1.0) and ML analysis (BP = 95–100%). The only exception was the clade containing all gregarines with PP = 0.92 and BP = 56%. The clades of chromerids, and of cryptosporidians + gregarines, were highly supported by Bayesian analyses, but moderately supported by ML analyses (PP = 0.95 and 1.0, and BP = 76% and 88%, respectively). Three sequences from the Cephaloidophoroidea formed a robust clade with full support of both Bayesian and ML analyses. This clade grouped with another long branch, *Gregarina niphandrodes*, with high support (PP = 1, BP = 95%).

In the ribosomal operon tree (Fig. 5B), the support values were similar to the LSU tree, but the gregarine clade had a higher PP (1.0 instead of 0.92) and BP (71% instead of 56%). However, support for the long branches of *G. niphandrodes* and Cephaloidophoroidea, the BP value decreased from 95% in LSU rDNA tree to 85%.

**Analyses of SSU rDNA (Fig. 6)**

In order to obtain comparable results, we analyzed SSU rDNA for the same sample of 31 OTUs. Bayesian and ML analyses yielded different results. The topology of the Bayesian consensus tree (Fig. 6A) was similar to those of the LSU rDNA and ribosomal operon trees; the only exception being chromerids, which appeared at the base of dinoflagellates, and the branch containing the Cephaloidophoroidea. The Cephaloidophoroidea grouped with coccidians, rather than with other gregarines. Nodal support at deeper positions within the tree had low resolution, although the support values for the Myzozoa clade, ciliates, as well as for alveolates, as a whole, were high. The increased number of generations or independent runs had no effect on the support values (data not shown).

The ML tree of 31 OTUs does not correspond to the current views on alveolate phylogeny (i.e., the relationship between dinoflagellate and apicomplexan clades was poorly supported). Support for many of the subgroups within these larger clades was also uncertain in our analyses. Chromerids appeared within the dinoflagellate clade, as well as a clade containing gregarines and cryptosporidians. The Cephaloidophoroidea had low support at the tip of dinoflagellates.

In order to evaluate the influence of long branch attraction (LBA), we sequentially removed non-alveolate outgroups and each long branch from the alignment (*Plasmodium* spp, *Gregarina niphandrodes*, and crustacean gregarines), as recommended by Bergsten (2005), and recalculated the ML trees each time, however this did not yield any improvement or result in a considerable change to the tree topology (data not shown).

Because of inconsistency is present in the Bayesian and ML SSU rDNA tree topologies, as well as discrepancy between those and LSU rDNA and ribosomal operon trees, we tested all three topologies (Figs 5 and 6), including a constraint SSU rDNA tree. The constraint SSU rDNA tree had the same branching order as both the LSU rDNA and ribosomal operon trees. None of the topology tests rejected any of these topologies (Table 2). This means that the LSU rDNA and ribosomal operon tree topology is also valid for SSU rDNA phylogeny of the 31 OTUs; although, it was not elected by the programs as a topology of “the best tree”, in this case.

Additional to the topology testing, we compared PP and BP supports of main clades in the three trees that shared the same topology (i.e., the constraint SSU rDNA tree and the trees resulting from the LSU rDNA and ri-
A

Theileria annulata
Theileria parva
Babesia bigemina
Plasmodium falciparum
Plasmodium berghei
Toxoplasma gondii
Isospora felis
Sarcocystis muris
Eimeria tenella
Ascogregarina taiwanensis
Cryptosporidium parvum
Chromera velia
Alexandrium catenella
Gonyaulax polyedra
Pfiesteria piscicida
Akashiwo sanguinea
Prorocerium micans
Sterkiella histriomuscorum
Spathidium amphoriforme
Paramecium tetraurelia
Hyphochytrium catenoides
Phytophthora megasperma
Mallomonas assemblae
Cylindrotheca closterium
Bigelowiella natans
Heterokonts
Rhizaria

B

Theileria annulata
Theileria parva
Babesia bigemina
Plasmodium falciparum
Plasmodium berghei
Toxoplasma gondii
Isospora felis + I. bellii
Sarcocystis muris + S. canis
Eimeria tenella
Ascogregarina taiwanensis
Cryptosporidium parvum
Chromera velia
Alexandrium catenella
Gonyaulax polyedra
Pfiesteria piscicida
Akashiwo sanguinea
Prorocerium micans
Sterkiella histriomuscorum
Spathidium amphoriforme + Spathidium sp.
Paramecium tetraurelia
Hyphochytrium catenoides
Phytophthora megasperma
Mallomonas assemblae
Cylindrotheca closterium
Bigelowiella natans
Heterokonts
Rhizaria

LSU rDNA
Ribosomal operon
(SSU+5.8S+LSU rDNA)
bosome operon analyses) (Table 3). The higher support values tended to coincide with increasing the amount of sites included in the analyses. This trend was especially apparent in the apicomplexan clades, and clades containing their close relatives, chromerids. Support values for clades based on LSU rDNA were considerably higher than those based on SSU rDNA. The clade containing the Cephaloidophoroidea clade was robust throughout all three trees. It was also noticed that PP values were higher than BP values in each case.

Finally, we compared rates of molecular evolution in LSU and SSU rDNA among apicomplexans using branch lengths. For this, branch lengths of apicomplexan OTUs were measured using scale bars from the root (common ancestor) of sporozoans (parasitic apicomplexans) to the leaves (tips) in the Bayesian trees of identical topologies: SSU rDNA constraint, LSU rDNA, and ribosomal operon trees. The branch length values (absolute evolutionary rates) were expressed in substitutions/site (Table 4).

### Table 2. Results of SSU rDNA tree topology tests.

<table>
<thead>
<tr>
<th>Tree topology</th>
<th>– ln L</th>
<th>bp^a</th>
<th>ELW^b</th>
<th>kh^c</th>
<th>sh^d</th>
<th>WSH^e</th>
<th>AU^f</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fig. 6A (Bayesian)</td>
<td>17158.31</td>
<td>0.261015</td>
<td>0.2628862</td>
<td>0.401667</td>
<td>0.607025</td>
<td>0.617512</td>
<td>0.5124509</td>
</tr>
<tr>
<td>Fig. 6B (ML)</td>
<td>17155.49</td>
<td>0.51943</td>
<td>0.5164639</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>0.6466904</td>
</tr>
<tr>
<td>as on Fig. 5 (constraintg)</td>
<td>17160.78</td>
<td>0.219555</td>
<td>0.2204883</td>
<td>0.333012</td>
<td>0.444848</td>
<td>0.475636</td>
<td>0.3695847</td>
</tr>
</tbody>
</table>

^a Bootstrap Probability (Felsenstein 1985); ^b Expected-Likelihood Weights (Strimmer and Rambaut 2002); ^c p-value of the Kishino-Hasegawa test (Kishino and Hasegawa 1989); ^d p-value of the Shimodaira-Hasegawa test (Shimodaira and Hasegawa 1999); ^e p-value of the Weighted Shimodaira-Hasegawa Test (Shimodaira and Hasegawa 1999); ^f p-value of the Approximately Unbiased test (Shimodaira 2002); ^g constraint according to the LSU rDNA / ribosome operon tree topology. P < 0.05 discards current topology.

### Table 3. Posterior probability (PP) and bootstrap percentage (BP) support of main clades in the SSU rDNA, LSU rDNA, and ribosomal operon trees*.

<table>
<thead>
<tr>
<th>Clade</th>
<th>SSU rDNA constraint tree (PP/BP)</th>
<th>LSU rDNA tree (PP/BP)</th>
<th>Ribosomal operon tree (PP/BP)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alveolates</td>
<td>1/98</td>
<td>1/100</td>
<td>1/100</td>
</tr>
<tr>
<td>Ciliates</td>
<td>1/99</td>
<td>1/100</td>
<td>1/100</td>
</tr>
<tr>
<td>Myozoa</td>
<td>1/69</td>
<td>1/100</td>
<td>1/99</td>
</tr>
<tr>
<td>Dinoflagellate (including perkinsids)**</td>
<td>0.93/48</td>
<td>1/87</td>
<td>1/99</td>
</tr>
<tr>
<td>Apicomplexans (including chromerids)**</td>
<td>0.02/11</td>
<td>1/56</td>
<td>98/70</td>
</tr>
<tr>
<td>Chromerids**</td>
<td>0.08/14</td>
<td>0.95/76</td>
<td>98/78</td>
</tr>
<tr>
<td>Sporozoans**</td>
<td>0.56/17</td>
<td>1/76</td>
<td>1/90</td>
</tr>
<tr>
<td>Coccidian clade (including hematozoans)</td>
<td>0.50/57</td>
<td>1/98</td>
<td>1/93</td>
</tr>
<tr>
<td>Gregarine clade**</td>
<td>0.41/23</td>
<td>0.92/56</td>
<td>1/71</td>
</tr>
<tr>
<td>Crustacean gregarines</td>
<td>1/100</td>
<td>1/100</td>
<td>1/100</td>
</tr>
</tbody>
</table>

* All the three trees have the same topology of resulting tree from LSU and/or ribosomal operon analysis (Fig. 5).
** The clades are absent on the SSU rRNA RAxML best tree.

Fig. 5. Bayesian inference (BI) trees of alveolates constructed using the GTR+Г+I model for 31 OTUs: A, LSU rDNA; B, ribosome operon. Numbers at the nodes indicate Bayesian posterior probabilities (numerator) and ML bootstrap percentage (denominator). Black dots on the branches indicate Bayesian posterior probabilities and bootstrap percentages of 1 and 100%, respectively. The gregarines from crustacean hosts are framed.
Table 4. Rates of molecular evolution (substitutions/site) of apicomplexans as calculated in SSU rDNA, LSU rDNA, and ribosomal operon phylogenetic analyses.

<table>
<thead>
<tr>
<th>Branch</th>
<th>SSU rDNA constraint tree*</th>
<th>LSU rDNA tree</th>
<th>Ribosomal operon tree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cryptosporidium parvum</td>
<td>0.11</td>
<td>0.23</td>
<td>0.16</td>
</tr>
<tr>
<td>Ascogregarina taiwanensis</td>
<td>0.21</td>
<td>0.33</td>
<td>0.29</td>
</tr>
<tr>
<td>Gregarina niphandrodies</td>
<td>0.90</td>
<td>0.98</td>
<td>1.02</td>
</tr>
<tr>
<td>Heliospora cf. longissima ex Eulimnogammarus vittatus</td>
<td>1.51</td>
<td>1.73</td>
<td>1.72</td>
</tr>
<tr>
<td>Heliospora cf. longissima ex E. verrucosus</td>
<td>1.51</td>
<td>1.73</td>
<td>1.72</td>
</tr>
<tr>
<td>Cephaloidophora cf. communis</td>
<td>1.70</td>
<td>1.79</td>
<td>1.86</td>
</tr>
<tr>
<td>Isospora**</td>
<td>0.09</td>
<td>0.10</td>
<td>0.10</td>
</tr>
<tr>
<td>Toxoplasma gondii</td>
<td>0.09</td>
<td>0.10</td>
<td>0.10</td>
</tr>
<tr>
<td>Sarcocystis***</td>
<td>0.09</td>
<td>0.10</td>
<td>0.10</td>
</tr>
<tr>
<td>Eimeria tenella</td>
<td>0.16</td>
<td>0.12</td>
<td>0.14</td>
</tr>
<tr>
<td>Babesia bigemina</td>
<td>0.18</td>
<td>0.35</td>
<td>0.30</td>
</tr>
<tr>
<td>Theileria parva</td>
<td>0.12</td>
<td>0.27</td>
<td>0.22</td>
</tr>
<tr>
<td>Theileria annulata</td>
<td>0.12</td>
<td>0.26</td>
<td>0.21</td>
</tr>
<tr>
<td>Plasmodium berghei</td>
<td>0.54</td>
<td>0.63</td>
<td>0.63</td>
</tr>
<tr>
<td>Plasmodium falciparum</td>
<td>0.49</td>
<td>0.63</td>
<td>0.60</td>
</tr>
</tbody>
</table>

* The SSU rDNA constraint tree has the same topology as the LSU rDNA and ribosomal operon trees (Fig. 5); the branch lengths (from the common ancestor of sporozoans to the leaves) were calculated with the MrBayes program.
** SSU and LSU rDNAs of Isospora felis, 5.8 rDNA of I. belli.
*** SSU and LSU rDNAs of Sarcocystis muris, 5.8 rDNA of S. canis.

Comparing individual branch lengths (Table 4), the long-branching species have just slightly longer branches in the LSU rDNA than in the SSU rDNA tree (~1.12 times). For the non-long-branching apicomplexans, the situation is more complicated. Species with moderate branch lengths (Cryptosporidium, Ascogregarina, Babesia, and Theileria spp.) have branches of about 1.5–2.5 (~2 on average) times longer in the LSU than the SSU rDNA tree. The shortest branches of coccidians keep the same lengths or are even a bit shorter. The average branch lengths of the nine non-long-branching or “slow-evolving” apicomplexans ($S_a = \text{total branch length of the four coccidians, three hemosporidians (without Plasmodium spp., Cryptosporidium, and Ascogregarina, divided by 9)}$) were equal to ~0.13 and ~0.21 substitutions/site for SSU and LSU rDNA, respectively.

Branch length pattern of the LSU rDNA tree considerably differs from that of the SSU rDNA trees, including the constraint SSU rDNA tree of the same topology. Despite the absolute evolutionary rates of apicomplexan LSU rDNA being higher than SSU rDNA, the long branches in the former look relatively shorter, if to compare with the non-long-branching apicomplexans (Figs 5 and 6). Therefore, we calculated relative evolutionary rates for each long branch (G. niphandrodies, C. cf. communis, H. cf. longissima samples, and two Plasmodium species) as a ratio $L / S_a$, where $L =$ length

Fig. 6. Two different trees of alveolates inferred by different methods using the GTR+Γ+I model for the same alignment of SSU rDNA sequences for 31 OTUs: A, Bayesian inference (BI), MrBayes program; B, maximum likelihood (ML), RAxML program. Numbers at the nodes indicate Bayesian posterior probabilities (in A) or ML bootstrap percentage (in B). Black dots on branches indicate Bayesian posterior probabilities or bootstrap percentages of 1 and 100%, respectively. The gregarines from crustacean hosts marked by rectangles. Note that both methods failed to establish the admissible positions of the crustacean gregarines and some other alveolate clades here, but the BI tree shows better results.
of a certain long branch (e.g., *C. cf. communis*) and $S_a$ = averaged branch length of the nine non-long-branching apicomplexans (see above). Our results showed that the relative evolutionary rates of the long-branching apicomplexans LSU rDNA are approximately 1.5 times less than the relative evolutionary rates of their SSU rDNA (Fig. 7).

Apicomplexan branch lengths in the ribosome operon tree (Table 4) appeared to be between SSU and LSU rDNAs branch lengths, with the exception of the long-branching species (i.e., *Plasmodium* spp. and the gregarines *C. cf. communis*, *H. cf. longissima*, and *G. niphandrodes*) which have branches of the same length, or even somewhat longer (*C. cf. communis*) than those in the LSU rDNA tree. The 5.8S rDNA region in these species is likely a contributing factor. This region is extremely variable in these species, compared to other OTUs.

**DISCUSSION**

**Remarks on phylogenetic analyses**

SSU rDNA sequences from different sporozoans evolve at highly heterogeneous rates. Notably, gregarines form long branches on phylogenetic trees, especially gregarines from crustaceans (Cephaloido-phorida). For this reason, the trees comprising such clades may be affected by long branch attraction (LBA) (see review of J. Bergsten 2005). Apart from LBA, high rates of evolution in SSU rDNA (e.g., high amounts of insertions and deletions (indels) and substitutions) are characteristic, even for those regions of gregarine SSU rDNA, which are only moderately variable in other eukaryotes. Such regions are the most important for phylogenetic analyses, since they obviously contain pronounced phylogenetic signal. The large amount of indels and substitutions also makes alignment of these regions difficult or impossible. For these reasons, long-branching lineages, such as cephaloidophoroid gregarines, as well as other groups of gregarines, are "rogue" taxa.

The 31 OTU trees based on SSU rDNA alone were inconsistent with (i) each other, (ii) a previously published study of 82 OTU alignment, comprising the crustacean gregarines (Rueckert et al. 2011b), and (iii) phylogenetic inferences from our preliminary analysis of the large taxon sampling. Moreover, these results did not fit expectations from morphological data. Both the trees demonstrated distorted topologies and many nodes had low support. From this, we can conclude that the reduction in sample size of SSU rDNA to 31 OTUs exerted a distinctly negative influence on the results of SSU rDNA analyses. We can also conclude that ML analyses yielded a less accurate inference, compared to the Bayesian approach, which proved to be less sensitive to size changes of taxon sampling. For example, in the ML tree topology, there is no clade uniting all sporozoans (parasitic apicomplexans) and the crustacean gregarines affiliated with dinoflagellates. In this case other gregarines, chromerids, and *Cryptosporidium*, became a sister group to all other Myzozoa (dinoflagellates + coccidians), whereas, in the Bayesian tree, the sporozoan clade exists, and the crustacean gregarines are situated within this clade. These findings fit the empirical results of Alfaro et al. (2003), which concluded (i) that the Bayesian approach is more susceptible to phylogenetic signal, (ii) it is a less biased predictor of phylogenetic accuracy than ML-bootstrapping method.
and (iii) it provides high support values for correct topological bipartitions with fewer characters than was needed for nonparametric bootstrap.

PP nodal support tends to be considerably higher than BP in the majority of the clades in our trees. This phenomenon was previously mentioned, along with the limited correlation between PP and BP values. (Alfaro et al. 2003, Cummings et al. 2003, Douady et al. 2003, Erixon et al. 2003). These values supply different aspects to a phylogenetic analysis. Alfaro et al. (2003) concluded that high Bayesian PP can indicate that the node is indeed correct; however, this was a condition based on the data set and the model of evolution. Moderate BP support of the nodes could indicate a high dependence on the underlying composition of the data matrix (i.e., taxon and gene sampling). Thus, increasing the amount of data used in the analyses, could result in either the disappearance of an artifactual node, or, conversely, the consistency of the node will increase (BP support will become higher), if the node is correct. In our study, we have the observed the latter case for majority of apicomplexan clades in Bayesian SSU, LSU, and concatenated rDNA trees. Therefore, we conclude that phylogenetic analysis employing Bayesian methodologies are the preferred method of inferring phylogenetic relationships among apicomplexans.

The 31 OTU SSU rDNA trees contain three long branches: Plasmodium spp., the crustacean gregarines (Cephaloidophoroidea), and Gregarina niphandros. Exclusion of non-alveolate outgroups and each long branch from the alignment, as it was recommended by Bergsten (2005), for detection of long branch attraction artifact (LBA) influence (the tree topology should change if it is present), did not significantly change the tree topology in our ML analyses (data not shown). Bergsten (2005) also concluded that increasing of alignment length does not improve upon, but instead, strengthens LBA effect. However, our analyses of concatenated rDNAs showed an improvement in the “best tree” topology (it was getting closer to that which would be expected, based on previous work), and support for nodes also increased (Table 3). Therefore, it is likely that one of the main reasons for “bad” topologies are datasets that are limited, with low phylogenetic signal, rather than LBA alone.

Including a greater number of characters in phylogenetic analyses can partially compensate for limited taxon sampling, despite the presence of long branches. It is also likely that improving in tree topology and support was the result of the gregarine LSU rDNA having a lower relative rate of evolution (Table 4, Fig. 7). Branch lengths among gregarine apicomplexans tended to remain the same length, while branch lengths in other apicomplexan lineages tended increase 1.5 to 2 times in length. Since variability in DNA and evolutionary rates should correspond to branch length, we can suppose that the regions of the LSU rDNA used in the phylogenetic analyses are more variable (on average about 1.5 times) than those of SSU rDNA in the majority of apicomplexans. However, the relative variability (or relative evolutionary rate) of the aligned regions of gregarine LSU rDNA is lower than that of SSU rDNA, and this resulted in a more balanced pattern of branch lengths in our LSU trees. Similar situations seem to exist among other fast-evolving apicomplexan lineages such as Plasmodium spp., as well as other gregarine lineages. We think that this is the main reason for a considerably better result in our phylogenetic analyses using LSU rDNA, in comparison with those datasets containing only SSU rDNA. Thus, our analyses containing relatively few, highly-diverged sequences have demonstrated the promising application of LSU rDNA for phylogenetic inferences among apicomplexans.

The exploration of LSU rDNA as a phylogenetic marker for apicomplexans will require further studies and broader taxon sampling. The problem lies in the limited number of LSU rDNA sequences available to date. Protein-coding genes might also contribute to the resolution of the apicomplexans backbone. However, available protein-coding sequences for gregarines are also limited and, additionally, subject to high evolutionary rates (Leander et al. 2003a).

Remarks on gregarine morphology

The monophyllic grouping of gregarines from crustaceans (Cephaloidophoroidea) prompts us to identify common morphologic traits between trophozoite and gamonts. With regard to gregarine morphology, it is necessary to distinguish between trophozoites (attached feeding stages according to Levine’s terminology (Levine 1971) and mature gamonts. The latter are free stages dedicated to the formation of syzygy, which have been detached from the host cell in a natural way. We consider all the solitary gregarine individuals studied here as gamonts because we dealt with free individuals, which were obviously full-grown, looked similar to syzygy partners, and demonstrated no obvious damages or drops of cytoplasm on their forebodies.

Previous studies on Cephaloidophora communis were conducted under magnifications too low for re-
revealing the fine cell structure of this gregarine (histological study by Lacombe et al. 2002). Our investigation has demonstrated that structure of the cells, as viewed under light and electron microscopy, correspond to each other. For instance, there are two septa in the gamonts of *C. cf. communis*, namely the protomerite/deutomerite septum, which has been documented in a diverse set of families including the Actinocyclusidae, Didymophyidae, Gregarinidae, Leidyanae, and Stylolophidae (Devauchelle 1968a, Desportes 1969, Baudoin and Ormières 1973, Ormières and Daumal 1970a, b, Hildebrand 1976, Ormières 1977, Tronechin and Schrével 1977, Valigurová and Koudela 2005, Valigurová et al. 2007, Valigurová 2012), and the distinct septum between the epimerite and the protomerite, observed in gregarines for the first time here. The small lenticular epimerite (cephaloid) is characteristic of this genus, which is reflected in the generic diagnosis of *Cephaloidophora* (Clorton 2000). This epimerite differs from the previously studied gregarine species (see above) in that it is not discarded after trophozoite detachment. It contains no large vacuoles, mitochondria, or fibril bundles, and it is filled with electron dense vesicles. These vesicles, resembling micronemes, seem to release some glutinous secretion through the pores. Previously, numerous pores have been described in *C. cf. communis* epimerite using SEM (Rueckert et al. 2011b). We observed the attachment (gluing) of living gregarines to a glass slide by the anterior end after a long-term incubation in seawater. We suppose this is likely the natural and only way of attachment for this gregarine. The previous light microscopic study of *C. communis* in the host intestine revealed no protruded epimerite (Lacombe et al. 2002). The possible reason is that trophozoites undergo intracellular development (Poisson 1924), which evidently does not require a complicated attachment apparatus. In this context, we propose to use the term “cephaloid” for the anterior end of *C. cf. communis* to emphasize its distinct difference from those of other gregarines with the discardable or condensable epimerite.

Concerning *Heliospora cf. longissima*, the situation is more difficult, because the primite and the satellite of the syzygy are organized differently and, in addition, the structure of the primite using electron microscopes (presence of just one septum) doesn’t correspond to those found in light micrographs (two septa are visible). Using light microscopy, there is a distinct light-refracting border marking some structure(s) in the cytoplasm, be it a fibrillar phragma or (probably) something else. In *H. cf. longissima*, this apparent epi-/protomerite septum could be a layer of electron-dense granules (Fig. 4C). The same granules probably reinforce a border, which is seen in light micrographs as the proto-/deutomerite septum, since the real septum seems too loose and is difficult to discern in some micrographs. Unfortunately, comparison with the data on other *Heliospora* species, which could solve this discrepancy, is impossible since these gregarines have not been studied by TEM. Regarding direct comparison with TEM data from other putatively related gregarines (e.g. *Uradiophora*), there are some doubts in this kinship, as explained below.

The genus *Heliospora* was included in Uradiophoridae family by Grassé (1953), based on parasitism in crustaceans and its cylindrical epimerite. The drawing of *H. longissima* presented by Grassé (1953) was obviously done from individuals fixed in syzygy, which had undergone deformation (a usual situation with fixed gregarines). However, the living gregarines have rather dome-shaped or lenticular epimerite, which is rather similar to the cephaloid of *Cephaloidophora*. Judging on existing images (see Fig. 3.90 in: Desportes and Schrével 2013), the forebody morphology of representatives of the type genus *Uradiophora* strongly differs from *Heliospora*. Their forebody is elongated and bears an anterior crown of hair- or finger-like processes (the attachment apparatus), and contains a bulky axial fibrillar structure, surrounded by large vacuoles. This is visible both at the light microscopic and ultrastructural levels (in *U. maetzii*, see Desportes and Théodoridès 1985), and has never been reported for any *Heliospora* spp. This attachment apparatus has contradictory interpretations that have described this as a protomerite (Desportes and Théodoridès 1985, Desportes and Schrével 2013, page 85), or even as a mucron (Desportes and Schrével 2013, page 346). According to the TEM data, a proto-/deutomerite septum in *U. maetzii* is absent in this specific developmental stage, but it is clearly visible on the light microscopy drawings of several *Uradiophora* spp. (Desportes and Théodoridès 1985, Desportes and Schrével 2013, Fig. 3.90). Additionally, there is a difference in the epicyte structure between *U. maetzii* and *H. cf. longissima*. Epicytic folds of the former species contain axial electron-dense matter in the tops of the folds (like in *Ganymedes, Porospora, Thiriitia*) whereas epicytic folds of the latter species obviously lack this (e.g., *Cephaloidophora*) (Desportes and Théodoridès 1985, Desportes and Schrével 2013, Desportes et al. 1977, this study). Thus, the close relations between *Heliospora* and *Uradiophora* are intrigu-
ing and questionable, especially in view of the absence of molecular phylogenetic data.

Regarding terminology used in diagnoses of gregarine taxa (e.g., diagnoses of the genus Cephaloidophora in the reviews Clorton 2000, Desportes and Schrével 2013), three weak points should be noted: (i) it is not always clear what characteristic belongs to which lifecycle stage (e.g., trophozoite or gamont), (ii) the terms “trophozoites” and “gamonts” are used interchangeably, and (iii) discrepancies in light and electron micrographs can exist (appearance of a septum, which can be actually absent).

It seems logical to modify the terminology for the trophozoite and gamont morphology to reflect both the number of sections within the body and the number of septa, in order to make some of these issues less ambiguous. A possible option might be to use the terms “mono-”, “di-”, and “tricystid” gregarine morphology based on light micrographs. The terms “aseptate” and “septate”, that indicate the number of the septa, might be exclusively used for descriptions and diagnoses based on transmission electron micrographs. For instance, gamonts of Lecudina spp. are monocystid (border is absent) and aseptate; Gregarina spp. are dicystid, uniseptate; Cephaloidophora cf. communis is tricystid and bisepptate; Helispora cf. longissima is tricystid uniseptate. Future studies and species descriptions or re-descriptions are advisable, in order to provide data using transmission electron microscopy. This could improve gregarine taxonomy and our understanding of some aspects of gregarine physiology, especially the role of the septum in gregarine feeding.

In past studies, molecular phylogenetic data demonstrated that the robust clade of gregarines parasitizing crustaceans (Cephaloidophoroidea) contains four subclades (Rueckert et al. 2011b). Two of them, corresponding to the families Cephaloidophoridae and (possibly) Uradiophoridae (see above about Heliospora taxonomy), include septate forms. The third clade corresponds to the family Porosporidae. The protomerite/ deutomerite septum in this clade is implicit or even missing (e.g., in Thriotia pugettiae). All these families belong to the suborder Aseptata, which has no septa (“monicystid” or “aseptate” forms). Thus, the presence of a septum is not a trait shared by all these gregarines. This leads us to propose that the septum (septa) could have either (i) emerged in advanced Cephaloidophoroidea forms (Cephaloidophoridae and, putatively, Uradiophoridae), independent of other septate gregarines (homoplasy), or (ii) the presence of a septum is the ancestral state for this clade, but it was secondarily lost in Ganymedidae. Unfortunately, the physiological role of the septum remains unclear, which does not allow us to specifically consider the factors underlying its emergence or the advantages it provides to the parasite cell.

All gregarines from crustaceans studied by transmission electron microscopy (Porospora portunidarum, Thriotia pisae, Ganymedes vibiliae, G. eucopiae, Uradiophora maetzi, Cephaloidophora cf. communis, and Heliospora cf. longissima) have a similar structure of epicystic folds (Desportes et al. 1977, Desportes and Théodoridès 1985, Desportes and Schrével 2013, this study). On cross-sections they are narrow with bulged tops, containing a few (3–6) rippled, dense structures and 2–5 apical filaments, with ejection of Ganymedes having 9 filaments. Additionally, many of them have electron-dense axial matter in the tops of the epicystic folds, however, it is clearly absent in C. cf. communis and H. cf. longissima. On the other hand, these characteristics are present also in other, unrelated gregarines. Swollen tops are observed in Gregarina polymorpha, and Gregarina steini has epicystic folds almost identical with C. cf. communis (3 apical arcs, and 2 apical filaments). These also contain the electron dense axial matter within the tops of the folds (Valigurová et al. 2013). The aseptate gregarine Gonospora beloneides (Urosoridae) has folds with swollen tops, containing the electron-dense axial matter as well, but also contains 9 apical arcs and filaments (Corbel et al. 1979). Thus, we conclude that our data on the epicystic structure highlights the morphological diversity of gregarines belonging to Cephaloidophoroidea. However, these ultrastructural characters do not present any distinct synapomorphies of the group. Consequently, this group shares only the two following characters:

- Parasitism in the intestines of crustacean hosts.
- Close relationship based on molecular data (rDNA).

These results, while contradicting some aspects of the Levine’s taxonomical scheme of Eugregarinida, which divides this group based on the presence or absence of a septum (Levine 1985, Clorton 2000), do support views by Grassé (1953), who placed the families of gregarines from crustaceans together, rather than spread among various suborders. However, a justified decision between the systems requires further molecular phylogenetic studies, spanning the taxonomic diver-
sity of gregarines. In particular, LSU rDNA sequencing may contribute to the phylogeny of these lower apicomplexans.

Acknowledgements. This study used the Bioportal at University of Oslo (www.bioportal.uio.no) and the Chebyshev Supercomputer Center of Lomonosov Moscow State University (http://parallel.ru/cluster) to make phylogenetic computations. DNA sequencing was performed at the DNA sequencing center “Genome” (Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, www.genome-centre.ru). The electron microscopic studies were performed at User Facilities Center of Lomonosov Moscow State University under financial support of Ministry of Education and Science of Russian Federation.

The work was supported by grant NSH–1801.2014.4 from the Council of the President of the Russian Federation and grants from the Russian Foundation of Basic Researches (09-04-01682, 11-04-00778, and 15-29-02601), Ministry of Education and Science of Russian Federation (Federal Target Program “Scientific and scientific-pedagogical personnel of innovative Russia 2009–2013”), and ECIP – Centre of excellence GACR No.GBP505/12/G112 (Czech Republic). The funders had no role in study design, data collection, analyses, decision to publish, or preparation of the manuscript. Our thanks are due to Dr. Dmitry Scherbakov (Linnological Institute, Siberian branch of the Russian Academy of Sciences, Irkutsk) for his help in organizing the expedition to Lake Baikal, to Dr. Niki-ta Vassetzky (Engelhardt Institute of Molecular Biology, Russian Academy of Sciences) for translating this article into English, and to Dr. Kevin Wakeman (University of British Columbia, Canada) for reading the manuscript.

REFERENCES


Received on 9th May, 2014; revised on 16th November, 2014; accepted on 21st November, 2014