Short Communication

Increasing taxon sampling using both unidentified environmental sequences and identified cultures improves phylogenetic inference in the Prorodontida (Ciliophora, Prostomatea)

Zhenzhen Yi a,b,c,1, Micah Dunthorn a,* ,1, Weibo Song b , Thorsten Stoeck a

a Department of Ecology, University of Kaiserslautern, 67663 Kaiserslautern, Germany
b Laboratory of Protozoology, Ocean University of China, Qingdao 266603, China
c Laboratory of Protozoology, South China Normal University, Guangzhou 510631, China

A R T I C L E  I N F O

Article history:
Received 5 May 2010
Revised 15 July 2010
Accepted 4 August 2010
Available online 10 August 2010

Keywords:
Taxon sampling
Prorodontida
Environmental sequences
Phylogeny
Ciliophora

A B S T R A C T

Taxon sampling for molecular phylogenetic inferences of microbial eukaryotes is limited because of the difficulties in finding specific taxa and culturing them. By contrast, unidentified sequences are easily collected during environmental diversity surveys. Here taxon sampling within prorodontid ciliates is increased using identified cultured isolates, and complemented with unidentified environmental sequences, with the nuclear small subunit rDNA locus. With identified cultured isolates there is support for the morphologically-circumscribed Colepidae. Increasing taxon sampling with unidentified environmental sequences is shown to change both topology and node support in clades that have low sampling for identified cultured isolates. This approach to increasing taxon sampling using unidentified environmental sequences can be used in other ciliate clades in which there is also low taxon sampling.

© 2010 Elsevier Inc. All rights reserved.

1. Introduction

Taxon sampling is known to affect the accuracy of phylogenetic inference (Graybeal, 1998; Heath et al., 2008; Hedtke et al., 2006; Hillis et al., 2003; Rannala et al., 1998). And theoretical approaches have been developed to specify which species should be targeted during diversity surveys. Here taxon sampling within prorodontid ciliates is increased using identified cultured isolates, and complemented with unidentified environmental sequences, with the nuclear small subunit rDNA locus. With identified cultured isolates there is support for the morphologically-circumscribed Colepidae. Increasing taxon sampling with unidentified environmental sequences is shown to change both topology and node support in clades that have low sampling for identified cultured isolates. This approach to increasing taxon sampling using unidentified environmental sequences can be used in other ciliate clades in which there is also low taxon sampling.

An exemplar clade of ciliates in which there are few sequences from identified isolates, but increasing numbers of environmental sequences, is the Prorodontida Corliss, 1974 (Ciliophora, Prostomatea). Prorodontids are diagnosed by their oral extrusomes and a bross (brush) of clavate cilia (Huttenlauch and Bardele, 1987; Lynn, 2008). Because of their simplified and apical oral structures, prorodontids where once thought to be an early branching ciliate lineage (Corliss, 1979). Later morphological (Bardele, 1989; Hiller, 1993; Huttenlauch and Bardele, 1987) and molecular (Baroin-Toumacheau et al., 1992; Lynn et al., 1999) analyses concluded that their simplified morphology is secondarily derived. Previous molecular phylogenetic inferences of the prorodontids using identified cultures have sampled 14 species from Balanion Wulf, 1919, Colepa Nitzsch, 1827, Cryptocaryon Brown, 1951, Leviceleps Foissner et al., 2008, Holophrya Ehrenberg, 1831, and Nolanda Small and Lynn, 1985 (Foissner et al., 2008b; Kim et al., 2007; Stechmann et al., 1998; Wright and Colomi, 2002). To the date of initiating this study (May 2009), on the other hand, there were 39 publicly available unidentified environmental SSU-rDNA sequences tentatively assigned to the prorodontids.
With our increased taxon sampling here of identified cultured isolates of prorodontids, we aim to test monophyly of the Colepidae Ehrenberg, 1838 as circumscribed by Foissner et al. (2008b). With the inclusion of unidentified environmental sequences, we aim to test the hypothesis that increasing taxon sampling using environmental sequences will change the SSU-rDNA topology and increase node support within the prorodontids.

2. Materials and methods

2.1. Taxon sampling and terminology

Five new prorodontid isolates were collected for this study (Table 1). Apocoleps sp. 1, Apocoleps sp. 2, Apocoleps magnus, and Nolandia sp. were isolated from sandy littoral sediments at Jiaozhou Bay near Qingdao, China (36°08′N; 120°43′E), in 2006 and 2007. Tiarina fusa was isolated from coastal waters off Guangzhou, China (22°42′N; 114°32′E) in 2008. Identifications were based on live protargol impregnated specimens following standard protocols (Foissner, 1991). Descriptions of the new species will be presented elsewhere. In addition, 14 identified prorodontid GenBank accesses were used (Table 1).

Table 1

<table>
<thead>
<tr>
<th>Taxon name</th>
<th>GenBank #</th>
</tr>
</thead>
<tbody>
<tr>
<td>Apocoleps magnus</td>
<td>AM292311</td>
</tr>
<tr>
<td>Apocoleps sp. 1</td>
<td>AM292312</td>
</tr>
<tr>
<td>Apocoleps sp. 2</td>
<td>AM412525</td>
</tr>
<tr>
<td>Balanion maenasii</td>
<td>AM412525</td>
</tr>
<tr>
<td>Coleps hirtus</td>
<td>AM293211</td>
</tr>
<tr>
<td>Coleps hirtus viridis</td>
<td>U97109</td>
</tr>
<tr>
<td>Coleps spetax</td>
<td>AM293212</td>
</tr>
<tr>
<td>Coleps sp. 1</td>
<td>EU004074</td>
</tr>
<tr>
<td>Coleps sp. 2</td>
<td>X7664</td>
</tr>
<tr>
<td>Coleps sp. 3</td>
<td>DQ487194</td>
</tr>
<tr>
<td>Cryptocaryon iritans</td>
<td>AF351579</td>
</tr>
<tr>
<td>Furgasonia blochmanni</td>
<td>X65150</td>
</tr>
<tr>
<td>Holophrya ovum</td>
<td>U97111</td>
</tr>
<tr>
<td>Holophrya teres</td>
<td>X77140</td>
</tr>
<tr>
<td>Leucoleps bwar</td>
<td>AB354737</td>
</tr>
<tr>
<td>Nolandia noldia</td>
<td>AM292313</td>
</tr>
<tr>
<td>Nolandia-like sp.</td>
<td>AM292315</td>
</tr>
<tr>
<td>Pelagiothrix alveolata</td>
<td>A8486009</td>
</tr>
<tr>
<td>Platyophrya vorax</td>
<td>AF060454</td>
</tr>
<tr>
<td>Tiarina fusa</td>
<td>FJ858217</td>
</tr>
<tr>
<td>Urotricha sp.</td>
<td>EU004081</td>
</tr>
</tbody>
</table>

With our increased taxon sampling here of identified cultured isolates of prorodontids, we aim to test monophyly of the Colepidae Ehrenberg, 1838 as circumscribed by Foissner et al. (2008b). With the inclusion of unidentified environmental sequences, we aim to test the hypothesis that increasing taxon sampling using environmental sequences will change the SSU-rDNA topology and increase node support within the prorodontids.

2.2. Amplification and alignments

One or more cells of each culture were repeatedly washed in sterile seawater. Genomic DNA was extracted using REDExtract-N-Amp Tissue PCR Kit (Sigma, St. Louis, USA). The SSU-rDNA locus was amplified using EUKA and EUKB primers from Medlin et al. (1988), with high-fidelity TaKaRa ExTaq polymerase (TaKaRa, Otsu, Japan). As the EUA forward primer yielded a product only for Tiarina fusa, the amplifications for the other four identified morphospecies were performed with the EUF primer. The amplification profile consisted of an initial step at 94°C for 5 min, 30 cycles of 1 min at 94°C, 2 min at 58°C and 2.5 min at 72°C, then a 15 min extension at 72°C. Purified products were inserted into the pUCm-T vector (Shanghai Sangon Biological Engineering and Technical Service Company, China) and sequenced bidirectionally (Invitrogen sequencing facility, Shanghai, China) using vector and SSU-rDNA primers.

Sequences were aligned using Hmmer v2.3.2 (Eddy, 1998). The alignment was manually modified and ambiguously aligned positions were conservatively masked in MacClade v4.08 (Maddison and Maddison, 2003). The effect of partial sequences (<80% complete) on resulting topologies was analyzed by removing the following sequences: Coleps sp. (EU024074), Urotricha sp., and the environmental sequences EF527059, EF526810, DQ310256, DQ310245, DQ455742, EF527045, and EF527036. Since no effect was found (data not shown), these sequences were included in the final analyses. Then we compiled two alignments for analysis: Alignment 1, containing only identified cultured isolates, totaling 21 sequences including outgroups; and Alignment 2, containing both identified cultured isolates and the unidentified environmental sequences, totaling 59 sequences including outgroups.

2.3. Genealogical analyses and hypothesis testing

For both alignments, GTR + I + F was the best fitted model selected by AIC as implemented in MrModeltest v2 (Nylander, 2004). Maximum likelihood analyses, and 1000 bootstrap replicates, were conducted using RaxML-HPC v7.2.5 (Stamatakis et al., 2008). Bayesian inferences (BI) were run in MrBayes v3.1.2 (Ronquist and Huelsenbeck, 2003), running two sets of four chains for 10,000,000 generations, sampling every 100 generations. Appropriateness of chain length was investigated using AWTY (Nylander et al., 2008). The first 2,500,000 generations were discarded as burn-in. The remaining trees were used to generate a consensus tree and to calculate the posterior probabilities (PP) using a majority-rule consensus tree. Here the most likely ML tree is shown with node support from both methods.

On the Alignment 1, three constrained ML analyses were carried out: (1) Holophryidae monophyletic, as circumscribed by Lynn (2008) with Cryptocaryon, Holophrya, and Pelagiothrix forming a monophyletic clade; (2) Holophrya monophyletic; and (3) Balanion, Cryptocaryon, and Urotricha monophyletic, as suggested by the topology resulting from Alignment 2 (Table 2). Constrained topologies were compared to the non-constrained ML topology using the AU test (Shimodaira, 2002) as implemented in CONSEL v0.1i (Shimodaira and Hasegawa, 2001).

3. Results

3.1. SSU-rDNA sequence characteristics

Small subunit rDNA sequences of five newly sequenced identified isolates were deposited in GenBank (Table 1). No intra-isolate variation was found. The SSU-rDNA sequence of Coleps sp. (EU024074), an undescribed morphotype, is the same to that of
C. spetai, suggesting that these sequences are from the same species.

3.2. Phylogenetic analyses of identified prorodontids

Overall, there is variable support for the deepest nodes in the SSU-rDNA topology from Alignment 1 that contains only identified cultured isolates (Fig. 1). Balanion is sister to the other prorodontids, with the rest forming a clade with no node support. Cryptocaryon and Urotricha form a clade with no ML and low BI node support (0.71 BI). Cryptocaryon does not form a clade with the rest of the sampled Holophryidae; an AU test also supports the non-monophyly of the Holophryidae ($P < 0.001$, Table 2). The remaining Holophryidae form a clade with the Colepidae with low ML but high BI support (55 ML, 0.99 BI). The non-monophyletic Holophrya forms a clade with Pelagothrix nested within it with full support from both methods (100 ML, 1.00 BI); an AU test likewise supports this relationship ($P < 0.001$, Table 2). Colepidae is monophyletic with full support from both methods (100 ML, 1.00 BI). Coleps, Levicolops, and Nolandia nolandi form a clade with low ML but high BI support (65 ML, 0.99 BI). Apocoleps, Nolandia-like, and Tiarina form a clade with high ML and BI node support (99 ML, 1.00 BI). All Coleps spp. and Apocoleps spp. group together, respectively. Nolandia nolandi does not group with the Nolandia-like sp. newly collected here, suggesting that the characters used to circumscribe the taxon Nolandia are the plesiomorphic condition for a more inclusive clade.

3.3. Topology comparisons

Here we will compare the inferred topologies of Alignment 1 (containing just the identified isolate sequences), and Alignment 2 (containing both identified and unidentified environmental sequences). Balanion is sister to all other prorodontids in the topology inferred from Alignment 1 (Fig. 1). By contrast, when unknown environmental sequences are included, Balanion, Cryptocaryon, and Urotricha form a clade in the topology inferred from Alignment 2 (Fig. 2). An AU test shows that a monophyletic clade containing these three genera cannot be rejected by Alignment 1 ($P = 0.357$, Table 2). In addition, there is an increase in node support from ML bootstraps in the relationship of Balanion, Cryptocaryon, and Urotricha between Figs. 1 and 2. Another place where there is an increase in node support is the clad composed of Coleps, Levicolops, and Nolandia, albeit the increase in support is only slightly in ML bootstrap values (and taxon sampling is only increased by two Coleps environmental sequences). With only identified sequences, this clade has low ML and high BI node support (65 ML, 0.99 BI, Fig. 1). With the inclusion of unidentified environmental sequences, ML support increases (81 ML, 0.98 BI, Fig. 2). These changes demonstrate that including unidentified environmental sequences can affect both topology and node support when inferring the phylogeny of ciliates, such as with this test case using prorodontids.

4. Discussion

4.1. Evolution and taxonomy within the Prorodontida

One of the most well known subclades within the prorodontids is the Colepidae. It is in this clade where we increased taxon sampling using identified cultured isolates. Colepids are diagnosed by the presence of external armor composed of elaborate, calcium carbonate plates (Foissner et al., 2008b; Lynn, 2008). They are currently classified into about 40 valid species assigned to 10 genera.

---

**Table 2**

Approximately unbiased test results on Alignment 1, containing only identified cultured isolates.

<table>
<thead>
<tr>
<th>Topology constraints</th>
<th>Log-likelihood</th>
<th>AU value ($P$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unconstrained</td>
<td>−7901.795979</td>
<td>0.645</td>
</tr>
<tr>
<td>Holophryidae monophyletic</td>
<td>−8015.824988</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Holophrya monophyletic</td>
<td>−7979.868001</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Balanion, Cryptocaryon, and Urotricha monophyletic</td>
<td>−7906.525109</td>
<td>0.357</td>
</tr>
</tbody>
</table>

---

**Fig. 1.** Maximum likelihood phylogeny of the Prorodontida from Alignment 1, containing only identified cultured isolates. Node support values are shown as ML bootstraps/BI posterior probability; values <50 are shown as ‘–’. Newly sequenced isolates are in bold.
With the sampling of Apocoleps species, Nolandia-like sp., and Tiarina fusa, the SSU-rDNA topologies support the monophyly of the morphologically-circumscribed colepids with full node support (Figs. 1 and 2). This suggests that the combination of elaborate calcium carbonate plates and brosse cilia in species with apical to subapical oral opening arose once within ciliates.

Using some of the diagnostic characters from Foissner et al. (2008b), we can suggest hypotheses of morphological evolution within the colepids in light of the SSU-rDNA topologies here (Figs. 1 and 2). First, because Coleps, Leviceoleps, Nolandia, and Tiarina all have six tiers of armor plates, this is the plesiomorphic condition for the clade. Increasing the number of plate tiers has been derived at least once in Apocoleps. Second, the presence of armor spines is the plesiomorphic condition as all sampled taxa have them, but the spines have been secondarily lost in some species of Coleps, Leviceoleps, and Nolandia. Third, given the positions of Nolandia, Nolandia-like, and Apocoleps, the nolandia-type plate is the plesiomorphic condition within the colepids, with the hirtus-type plate evolving in Coleps and Leviceoleps, and the tiarina-type plate evolving in Tiarina.

Although we did not increase taxon sampling outside of the Colepidae, our analyses do point to non-monophyly in another prorodontid subclade: Holophryidae. Cryptocaryon does not fall into a clade with the other sequenced Holophryidae (Figs. 1 and 2). Cryptocaryonidae was erected for this taxon by Wright and Colorni (2002), while Lynn and Small (2002) suggested that Cryptocaryon should be a prostomatean incertae sedis. Later, Lynn (2008) placed Cryptocaryon into the Holophryidae. Our phylogenetic analyses (Figs. 1 and 2) and an AU test (Table 2) suggest that the inclusion of Cryptocaryon in Holophryidae renders it non-monophyletic. Cryptocaryonidae should therefore be resurrected sensu Wright and Colorni (2002).

Another non-monophyletic grouping we found is the Holophrya, as P. alveolata nests within it (Figs. 1 and 2). An AU test also supports the non-monophyly of the Holophrya (Table 2). While P. alveolata was initially described as a Urotricha, it was later transferred to Pelagophrya by Foissner et al. (1999), because of morphological similarities to other Pelagophrya species, such as conspicuous cortical algae, body size, symbiotic algae, and sparse caudal cilia. In agreement with the conclusion of Foissner et al. (1999), all our inferred trees show that this species is relatively distant from Urotricha (Figs. 1 and 2). Whether P. alveolata should be transferred to Holophrya, though, should await until further taxon sampling is accomplished in more isolates of P. alveolata (to check for accuracy of identification) and in other closely related species.

4.2. Topology changes with the addition of unidentified environmental sequences

It has long been argued if it is more important to increase the number of taxa or the number of characters to improve phylogenetic inference (Cummings and Meyer, 2005; Graybeal, 1998; Heath et al., 2008; Hedtke et al., 2006; Hillis et al., 2003; Pollock et al., 2002; Ramala et al., 1998; Rokas and Carroll, 2005). Each approach has its strengths and weaknesses for each specific taxon to which they are applied. In ciliates in particular, phylogenetic inferences since the 1980s for deep nodes have largely focused upon increasing taxon sampling (e.g., Dunthorn et al., 2008; Schmidt et al., 2007; Snelten-Struve et al., 2004; Strüder-Kypke et al., 2006; Yi et al., 2009). Still, the vast majority of ciliate species and higher taxa remain to be sampled.

With any clade of microbial eukaryotes, such as ciliates, one can attempt to design a targeted approach for taxon sampling in specific clades. Such a cultured approach is always needed, as phylo-
genetic inferences are more biologically interesting if the tips of the inferred trees are identified taxa. But finding phylogenetic inference of ciliates can change both the topology and node support. Balanionidae, Holophryidae, and the Urotrichidae do not form a clade in the topology with just identified cultured isolates (Fig. 1), while they do group together in a single clade in the topology that includes unidentified environmental sequences (Fig. 2); this monophyletic relationship cannot be rejected in the alignment of just cultured isolates (Table 2). Also, with the inclusion of unidentified environmental sequences, there is an increase in node support for the clade comprised of Balanionidae, Holophryidae, and Urotrichidae, as well as a slight increase in the clade comprised of Coleps, Leviceps, and Nolantida (Figs. 1 and 2). These differences suggest that including unidentified environmental sequences may be beneficial for phylogenetic inference in other ciliate taxa, especially in those where there is low taxon sampling using unidentified environmental sequences deposited in GenBank and should be used in ciliates and other microbial eukaryotes.

It is shown here that the inclusion of unidentified environmental sequences for phylogenetic inference of ciliates can change both topology and node support. Balanionidae, Holophryidae, and the Urotrichidae do not form a clade in the topology with just identified cultured isolates (Fig. 1), while they do group together in a single clade in the topology that includes unidentified environmental sequences (Fig. 2); this monophyletic relationship cannot be rejected in the alignment of just cultured isolates (Table 2). Also, with the inclusion of unidentified environmental sequences, there is an increase in node support for the clade comprised of Balanionidae, Holophryidae, and Urotrichidae, as well as a slight increase in the clade comprised of Coleps, Leviceps, and Nolantida (Figs. 1 and 2). These differences suggest that including unidentified environmental sequences may be beneficial for phylogenetic inference in other ciliate taxa, especially in those where there is low taxon sampling for identified cultured isolates.

Acknowledgments

We would like to thank Miao Miao, Qianqian Zhang, Feng Gao, Jiehuang, Xianrui Chen, Zicong Li, Yangang Wang, and Weiwei Liu for technical help and sample collection. Thanks also to John Archibald and two anonymous reviewers for comments and editing. This work was supported by postdoctoral fellowships from the Faculty of Biology of the University of Kaiserslautern and from the Alexander von Humboldt Foundation to M.D. the Nature Science Foundation of China (Project Nos. 30672080, 30870264) to W.S., and the Deutsche Forschungsgemeinschaft (Grant STO414/3-1) to T.S.

References


