Published in
Protist 162, issue 1, 131-141, 2011
which should be used for any reference to this work

Comparing Potential COI and SSU rDNA Barcodes for Assessing the Diversity and Phylogenetic Relationships of Cyphoderiid Testate Amoebae (Rhizaria: Euglyphida)

Thierry J. Heger[a,b,c,d,e,1], **Jan Pawlowski**[d], **Enrique Lara**[c], **Brian S. Leander**[e], **Milcho Todorov**[f], **Vassil Golemansky**[f], and **Edward A.D. Mitchell**[a,b,c]

[a] WSL, Swiss Federal Institute for Forest, Snow and Landscape Research, Ecosystem Boundaries Research Unit, Wetlands Research Group, Station 2, CH-1015 Lausanne, Switzerland
[b] École Polytechnique Fédérale de Lausanne (EPFL), Laboratory of Ecological Systems, Station 2, CH - 1015 Lausanne, Switzerland
[c] Laboratory of Soil Biology, Institute of Biology, Rue Emile Argand 11, University of Neuchâtel, CH-2009 Neuchâtel, Switzerland
[d] Department of Zoology and Animal Biology, University of Geneva, Sciences III, CH-1211 Geneva 4, Switzerland
[e] Departments of Botany and Zoology, University of British Columbia, 6270 University Boulevard, Vancouver, BC, Canada V6T 1Z4
[f] Bulgarian Academy of Sciences, Institute of Biodiversity and Ecosystem Research, 2 Gagarin St., 1113 Sofia, Bulgaria

The mitochondrial Cytochrome Oxidase Subunit 1 gene (COI) has been promoted as an ideal "DNA barcode" for animal species and other groups of eukaryotes. However, the utility of the COI marker for species level discrimination and for phylogenetic analyses has yet to be tested within the Rhizaria. Accordingly, we analysed mitochondrial COI gene sequences and nuclear small subunit rDNA (SSU) sequences from several morphospecies of euglyphid testate amoebae (Cercozoa, Rhizaria) in order to evaluate the utility of these DNA markers for species discrimination and phylogenetic reconstructions. Sequences were obtained from eleven populations belonging to six *Cyphoderia* morphospecies that were isolated from field samples in North America and Europe. Mean inter-population COI sequence dissimilarities were on average 2.9 times greater than in the SSU, while the intra-population sequence dissimilarities were higher in the SSU (0-0.95%) than in the COI (0%); this suggests that the COI fragment is valuable for discriminating Cyphoderiidae isolates. Our study also demonstrated that COI sequences are useful for inferring phylogenetic relationships among Cyphoderiidae isolates. COI and SSU tree topologies were very similar even though the COI fragment used in these analyses (500 bp) was much shorter than the SSU sequences (1600 bp). Altogether, these results demonstrate the utility of the COI as a potential taxonomic DNA barcode for assessing cyphoderiid species diversity and for inferring phylogenetic relationships within the group.

Key words: barcoding; COI; Cyphoderiidae; Euglyphida; Rhizaria; testate amoebae

Introduction

Facilitating the identification of living species is one of the major goals of modern systematics. DNA-based identification represents an alternative taxonomic approach for identifying the huge diversity of living organisms (Tautz et al. 2003). Hebert et al. (2003a) proposed the mitochondrial Cytochrome c Oxidase Subunit 1 (COI) as an ideal "DNA barcode" for animal species. Indeed, COI proved to be efficient in many taxonomic groups of animals, such as birds (Hebert et al. 2004a; Kerr et al. 2007), fishes (Ivanova et al. 2007), bryozoans (Gómez et al. 2007) and insects (Hebert et al. 2004b; Smith et al. 2005). This marker, however, appeared to be too conserved in other groups of animals, such as scleractinian corals (Shearer and Coffroth 2008), and in land plants (Newmaster et al. 2008). This stresses the need of testing any potential barcode for species delimitation before utilizing the genetic marker in contingent research (e.g., a study on biogeography or paleoecology). Some recent studies have demonstrated the usefulness of the COI for species identification in different groups of microbial eukaryotes, such as ciliates (Barth et al. 2006; Chantangsi et al. 2007; Chantangsi and Lynn 2008), diatoms (Evans et al. 2007; Evans et al. 2008; Evans and Mann 2009), dinoflagellates (Lin et al. 2009) and amoebozoans (Nassonova et al. 2010). The utility of the COI marker for species level discrimination and phylogenetic inferences has yet to be tested within a member of the Rhizaria, the eukaryotic supergroup to which the euglyphid testate amoebae belong.

Euglyphid testate amoebae are diverse and abundant in different terrestrial and aquatic environments and are found in virtually all habitats on Earth from the tropics to the poles (Meisterfeld 2002). They are an important component of peatland microbial communities (Mitchell et al. 2003, 2008) and are functionally important in soils (Wilkinson 2008). Moreover, the response of testate amoebae to different ecological gradients and pollutants make them a useful tool for paleoecological studies and pollution monitoring (Mitchell et al. 2008; Nguyen-Viet et al. 2004, 2008). The current taxonomy for euglyphids is largely based on shell characters. Shells (also called tests) are composed of secreted plates which often differ in shape, size and arrangement from one species to another (Meisterfeld 2002). However, the delimitation of morphospecies can be very difficult due to the absence of diagnostic characters. Comparative molecular and morphological analyses have shown cryptic (species which cannot be discriminated by morphology alone) and pseudo-cryptic species (species with subtle morphological dissimilarities, possibly visible only by scanning electron microscopy) within different euglyphid morphospecies (Heger et al. 2010; Todorov et al. 2009; Wylezich et al. 2002). There is therefore a need for developing a tool for DNA-based identification, which may be useful for evaluating the diversity and assessing the biology and ecology of euglyphid testate amoebae. Until now, the molecular diversity of euglyphid testate amoebae was based only on the small and large subunit rRNA sequences (Heger et al. 2010; Lara et al. 2007; Wylezich et al. 2002; Yoon et al. 2009). However, the slow rate of evolution and the presence of diverging copies (Pecher et al. 2004; Scholin et al. 1993) in the ribosomal genes of some groups of eukaryotes limit the effectiveness of these markers for species distinction and overall estimates of diversity.

In this study, we tested the utility of the COI marker for species level discrimination for eleven populations (one population corresponds to several individuals of a given morphospecies collected from one site) belonging to six morphotaxa. Furthermore, we compared the effectiveness of the mitochondrial COI and the nuclear SSU rDNA for assessing phylogenetic relationships among the Cyphoderiidae. We also evaluated the validity of the morphology-based taxonomy by comparing the SSU rDNA and COI gene phylogenies to the shell morphology documented by scanning electron microscopy.

Results

A total of thirty new Cyphoderiidae COI sequences (502-760 bp) were obtained from ten *Cyphoderia* populations and one *Pseudocorythion acutum* outgroup population (see Methods). The corresponding SSU rDNA sequences (almost full length) were recently obtained by Heger et al. (2010). With the exception of *Cyphoderia* cf. *compressa*, some representatives of these populations were characterized by scanning electron microscopy (Fig. 1) (Heger et al. 2010; Todorov et al. 2009).

Sequence Dissimilarities

The rDNA and COI sequence data demonstrated that the eleven populations investigated were relatively well differentiated (Fig. 2). The COI and SSU sequence revealed a contrasting pattern of intra- and inter-population dissimilarities: Intra-population SSU rDNA sequence dissimilarities ranged between 0 to 0.95%, whereas COI intra-population sequences were always identical. SSU rDNA inter-population sequence dissimilarities ranged between 0.4 and 9.9% and COI inter-population sequence dissimilarities ranged between 3% and 18.8%. Mean inter-population COI sequence dissimilarities among populations were on average more than 2.9 times greater than the SSU rDNA inter-population sequence dissimilarities (Fig. 3).

Molecular Phylogenies Based on SSU rRNA and COI Sequences

Phylogenetic trees based on a 1600 bp fragment of the SSU rDNA and on a 500 bp fragment of the COI showed very similar topologies (Fig. 2A and 2B). The freshwater *Cyphoderia* isolates were separated into two strongly supported subclades in the COI tree (subclade 1: 93% Bootstrap support (BS) and 0.99 posterior probabilities (PP); subclade 2: 100% BS and 1.00 PP) as well as in the SSU rDNA tree (subclade 1: 100% BS and 1.00 PP; subclade 2: 89% BS and 0.76 PP). There is only one main topological difference between the results obtained from the two genes. In the SSU rDNA tree, the marine supralittoral species *C. compressa* and *C.* cf. *compressa* branched as a sister group to the freshwater *Cyphoderia* subclade 1. By contrast, *C. compressa* had a basal position within the *Cyphoderia* clade and *C.* cf. *compressa* branched as a sister taxon to the two freshwater *Cyphoderia* subclades in the COI tree. The first freshwater subclade comprised *Cyphoderia amphoralis*, *Cyphoderia trochus* ssp. *palustris* and *Cyphoderia ampulla* from Nova Scotia, Rhodopes and Vitosha and was composed of isolates having a shell built of overlapping or slightly overlapping scales (Figs 1, 2A and 2B). Within this subclade, the phylogenetic relationships among populations were generally weakly or not supported by either the Bayesian or the maximum likelihood analyses. However, *Cyphoderia amphoralis* from Rila and *Cyphoderia ampulla* from Vitosha clustered together with high posterior probabilities in both COI and SSU rDNA trees. The second freshwater subclade comprised the freshwater species *C. major* and *C. ampulla* from Lake Geneva and Moiry. The shell ultrastructure of these specimens was characterized by non-overlapping scales (Figs 1, 2A and B). In the COI and SSU rDNA trees, *Cyphoderia major* branched as a sister group to the strongly supported clade composed of *Cyphoderia ampulla* from Lake Geneva and Moiry.

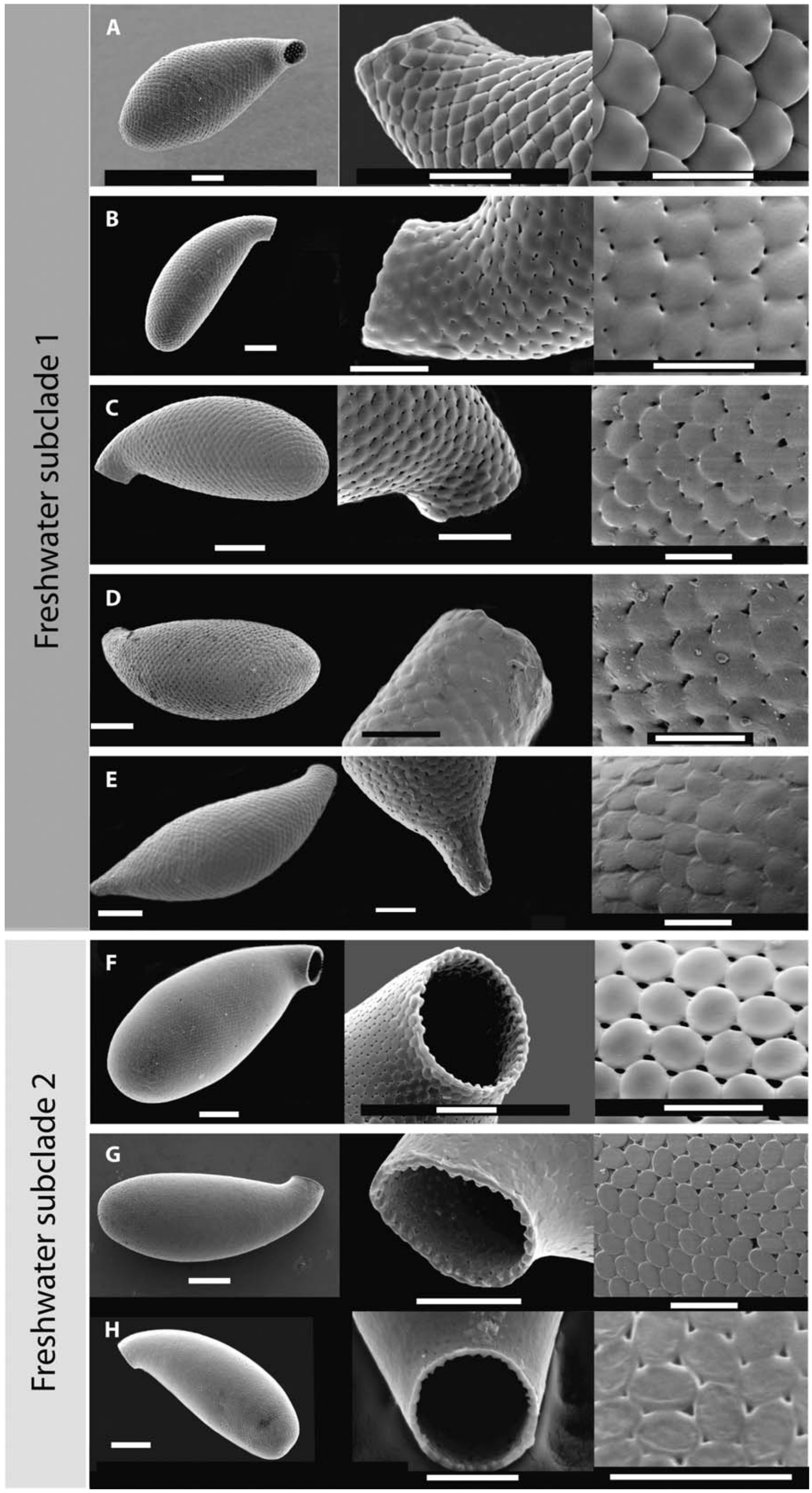
Influence of the Length of the COI and SSU Alignments on the Robustness of the Subclade 1 and 2

Bootstrap values inferred by maximum likelihood analyses indicated that the robustness of the nodes supporting the monophyly of the subclades 1 and 2 were comparatively higher in the COI tree than in a corresponding SSU rDNA tree (Fig. 4A and B). With only the first 100 bp of the COI alignment, the support of the subclade 2 rises over 70% BS. By contrast, more than 500 bp of the SSU rDNA aligment is required to get a comparable bootstrap support (Fig. 4B). For subclade 2, a COI and SSU rDNA alignments of 200 bp and 800 bp, respectively, were needed to obtain a resolution higher than 70% BP (Fig. 4A).

Discussion

Recently, several studies have revealed that the DNA barcoding might represent an interesting approach for identifying species when their morphology is of limited value (Hebert et al. 2003a). Although mainly based on animals, the potential use of COI marker as a barcoding tool was also demonstrated for several groups of protist (Barth et al. 2006; Chantangsi et al. 2007; Lin et al. 2009; Nassonova et al. 2010) and red algae (Saunders 2005). However, COI sequences are available for relatively few groups of eukaryotes.

Our study is the first attempt to examine the potential usefulness of COI as a barcoding gene for the Rhizaria, one of the largest groups of eukaryotes. Our results showed that the COI marker is sufficiently variable and most likely more appropriate than the corresponding SSU rDNA gene to differentiate closely related isolates of euglyphid testate amoebae within the Cyphoderiidae. Mean inter-population COI sequence dissimilarities were indeed always higher than mean inter-population SSU rDNA sequence dissimilarities while COI mean intrapopulation sequence dissimilarities were lower than the SSU rDNA sequence dissimilarities. The apparently fast evolutionary rate of COI gene in Cyphoderiidae testate amoeba and lack of COI intra-population heterogeneity concur with previously reported data in most animals (Hebert et al. 2003b; Remigio and Hebert 2003), and the other studied groups of protists (Chantangsi et al. 2007; Evans et al. 2007; Nassonova et al. 2010). Our data suggest that the COI intra-individual polymorphism is non-existent while SSU rDNA intra-individual polymorphism remains likely. However, given the


Freshwater subclade 1
Freshwater subclade 2
A
B
C
D
E
F
G
H

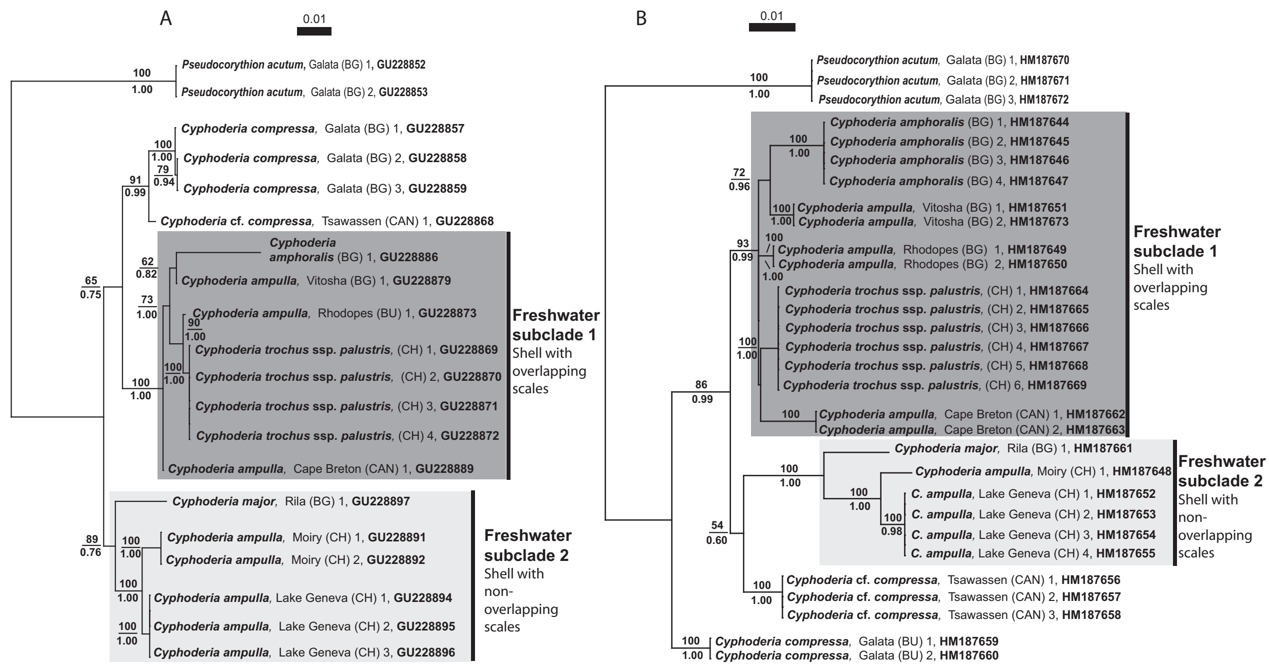


Figure 2. Cyphoderiidae phylogenies obtained from Maximum-likelihood (ML) analysis of SSU rDNA (**A**) and COI (**B**) sequences based respectively on 1600 and 500 nucleotide positions. Some ambiguously aligned regions and gaps were detected and excluded in the SSU rDNA alignment. Such gaps or unaligned regions were absent within the COI alignment. Dark grey boxes comprise freshwater individuals having overlapping or slightly overlapping scales. Light grey boxes comprise freshwater individuals having non-overlapping scales. Numbers represent values of bootstraps obtained by the maximum likelihood method where greater than 50% and posterior probabilities as calculated with Bayesian analyses. The tree was rooted with *Pseudocorythion acutum*.

fact that we used more than one cell per extraction, we can not explicitly confirm SSU rDNA intra-individual polymorphisms.

Cryptic and Pseudo-Cryptic Species

COI DNA barcode appears also to be a useful tool for revealing cryptic and pseudo-cryptic diversity. Although several *C. ampulla* isolates are clearly undistinguishable under light and scanning electron microscopy (Fig. 1), the COI marker allows distinction of all examined isolated populations. The high COI sequence dissimilarities among *Cyphoderia* populations (>3%), in agreement with SSU DNA and to some extent with morphological data demonstrate that the five *C. ampulla* populations examined in this study represent distinct species. Such COI sequence dissimilarities values are indeed similar to most threshold values usually used for discriminating protist species (Chantangsi et al. 2007; Evans et al. 2007; Nassonova et al. 2010). In agreement with the recent morphological and SSU rDNA molecular phylogenetic studies reported by Todorov et al. (2009) and Heger et al. (2010), our COI sequence data confirm the presence of a hidden diversity (i.e. indistinguishable morphospecies under the light microscope) within these *Cyphoderia* morphospecies. Light microscopy allowed us to discriminate several *Cyphoderia* morphospecies such as *C. major* (characterized by a large shell), *C. trochus* ssp. *palustris* (characterized by a terminal protuberance) or *C. amphoralis* (charac-

Figure 1. Scanning electron micrographs of eight Cyphoderiidae populations of subclade 1 **A-E.** characterized by more or less overlapping scales and population of subclade 2 **F-H.** characterized by non-overlapping scales. **A.** *Cyphoderia amphoralis* from Rila (BG). **B.** *Cyphoderia ampulla* from Vitosha (BG). C. *Cyphoderia ampulla* from Rhodopes (BG). **D.** *Cyphoderia ampulla* from Cape Breton (CAN). **E.** *Cyphoderia trochus* ssp. *palustris* from Marchairuz (CH). **F.** *Cyphoderia major* from Rila (BG). **G.** *Cyphoderia ampulla* from Moiry (CH). **H.** *Cyphoderia ampulla* from Lake Geneva (CH). Scale bars on the left, at the centre and on the right correspond respectively to 20 μm, 10 μm and 5 μm (pictures from Todorov et al. (2009) and Heger et al. (2010)).

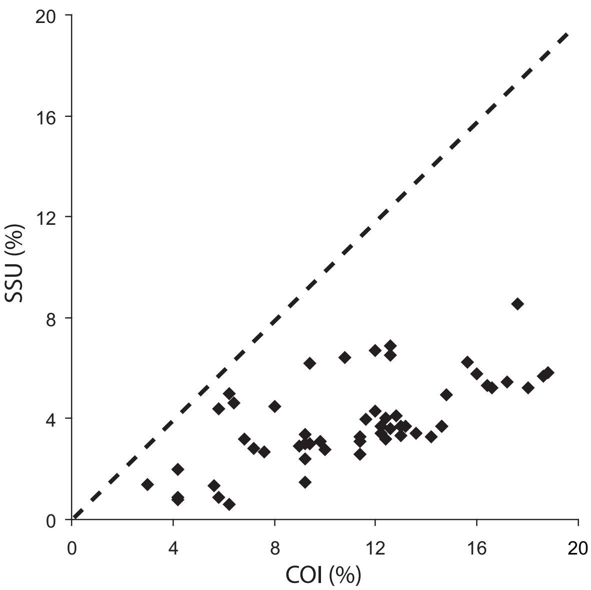


Figure 3. Relationships between mean interpopulation COI (500 bp) sequence dissimilarities and mean SSU rDNA (1630 bp) sequence dissimilarities of 11 Cyphoderiidae populations. The dash line indicates identical levels of dissimilarities. COI sequence dissimilarities are greater than SSU rDNA sequence dissimilarities.

terized by clearly imbricated scales). Scanning electron microscopy was necessary to distinguish *C. ampulla* populations of subclade 1 (species with overlapping scales) and *C. ampulla* populations of subclade 2 (species with non-overlapping scales). Within these two subclades, the discrimination of *C. ampulla* populations (corresponding to different "species") was very difficult. Nevertheless, extensive biometrical analyses revealed significant morphological differences between *C. ampulla* specimens from Rhodopes and Vitosha (Todorov et al. 2009). Altogether this example shows that the morphological-based identification of some testate amoebae remains very difficult and/or time consuming because (1) morphological criteria visible by light microscopy do not discriminate between different clades of *C. ampulla*, and (2) the taxonomic significance of ultrastructural characters, such as scale morphology and arrangement, has not been fully explored. Our combined SEM and molecular phylogenetic results suggest that detailed SEM-based morphological studies combined with COI sequence data should clarify the interrelationships of species within *Cyphoderia*, and this is expected to hold true for other euglyphid testate amoebae as well.

Phylogenetic Relationships

Our study revealed that COI sequences not only allow distinguishing different isolates but that the information is also useful for inferring phylogenetic relationships among isolates of the Cyphoderiidae. COI and SSU rDNA tree topologies were indeed very similar although COI fragments used in these analyses (500 bp) were much shorter than the SSU rDNA fragments (1600 bp). The main difference between the SSU rDNA tree and the COI tree is related to the phylogenetic position of *C. compressa* (sensu lato). In the COI tree, isolates of *Cyphoderia compressa* (sensu lato) were not closely related while in the SSU rDNA tree, the isolates of *Cyphoderia compressa* (sensu lato) branched together. The phylogenetic topology shown in the SSU rDNA tree is concordant with morphology. Although *C.* cf. *compressa* was not characterized with scanning electron microscopy, our light microscopic observations performed at the time of the isolation indicated that *C* cf. *compressa* is very similar to *C. compressa* (laterally compressed shell with kidney shaped scales and a relative long size). Our molecular phylogenetic analyses also demonstrated that a much shorter COI than SSU rDNA alignment is needed to obtain good resolutions of the subgroup 1 and 2. In order to get a phylogenetic support of the freshwater subgroup 1 higher than 70%, a COI alignment six time shorter than the SSU rDNA alignment is needed (100 COI bp and 600 SSU rDNA bp) while for the freshwater subclade 2, a COI alignment four times shorter than the SSU rDNA alignment is needed (200 COI bp, 800 SSU rDNA bp).

Limitations of the Primers Used

An ideal COI barcoding fragment contains both variable regions allowing species discrimination as well as highly conserved regions used as the targets for primers. For example, LCO and HCO standardized COI barcoding primers are usable for various Metazoan taxa as well as for closely related Amoebozoa lineages (Folmer et al. 1994; Nassonova et al. 2010). However, these primers did not work for Cyphoderiidae testate amoebae and different primers relatively specific to Cyphoderiidae and Euglyphida species were developed in order to amplify the COI fragment of different isolates. The current lack of universal euglyphid primers clearly represents the main limitation for the use of the COI DNA marker as a barcoding tool for Euglyphida testate amoebae. Such problems were also reported in other groups of protists. In ciliates and diatoms,

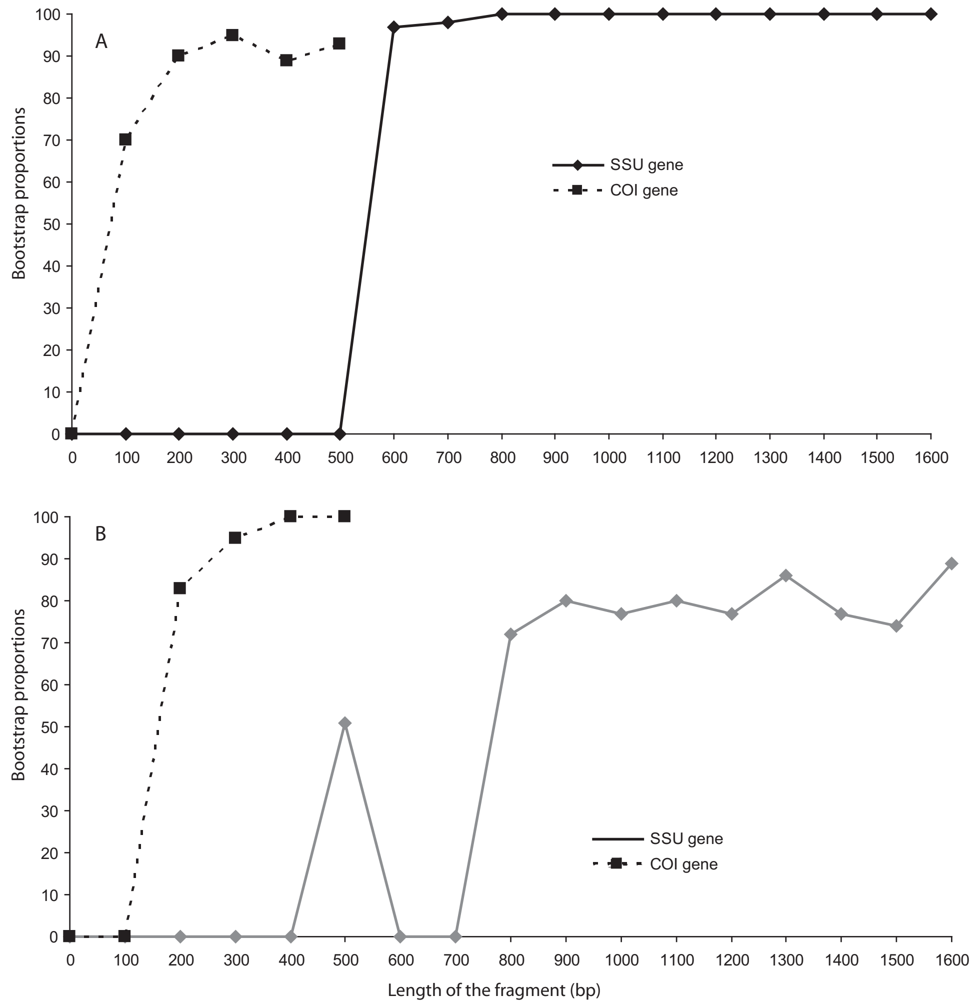


Figure 4. Influence of the size of the COI and SSU fragments on the phylogenetic resolution of the freshwater subclade 1 (graph **A**) and 2 (graph **B**). Bootstrap supports were inferred by maximum-likelihood analyses.

the developed COI primers work also for a limited number of genera (Barth et al. 2006; Chantangsi et al. 2007; Chantangsi and Lynn 2008; Guggiari and Peck 2008; Lynn and Strüder-Kypke 2006). More universal primers need to be developed in order to use this barcoding approach for all Euglyphida, and perhaps for the Rhizaria in general.

Methods

Choice of isolates: Based on SSU rDNA sequences, the phylogenetic relationships of 15 Cyphoderiidae populations were studied by Heger et al. (2010). Here, eleven Cyphoderiidae populations (Table 1) of this previous study were selected to test the use of a COI marker for assessing the diversity and phylogenetic relationships of the group.

Primer design strategy: We first tested the usefulness of the COI standardized barcoding HCO and LCO primers which are commonly used for various animal phyla (Folmer et al. 1994). These primers failed to work for the cyphoderiid taxa. We therefore designed a primer set based on the three Rhizaria full length COI sequences available in GenBank (*Euglypha rotunda* GenBank Accession No AB009417, *Chlorarachnion* sp. AB009396, *Chlorarachnion reptans* AB009416) and several non-Rhizaria full length COI sequences. Because of the apparent lack of a conserved region along the Rhizaria COI alignment, we developed degenerate primers. Eucox1F and Euglycox1R or Euglycox6R (Table 2), were initially tested on *Euglypha rotunda* Wailes (CCAP 1520/1) DNA extractions. The obtained fragments were identical to the same part of the COI sequence already published in Genbank (AB009417) by Inagaki et al. (1998). In addition, more specific primers to cyphoderiidae species were designed (Eucox5F and Eucox6F, Table 2).

COI fragment amplification and sequencing: A portion of the COI gene (500-760 bp), corresponding to the region 340 to 1088 of *E. rotunda* (AF396436), was amplified by a first polymerase chain reaction (PCR) using the degenerate primers Eucox1F and Euglycox1R or Euglycox6R. In cases of poor amplification, the nested reverse degenerate primer

Table 1. List of the *Cyphoderiidae* morphotypes analysed and sampling locations.

Taxa	Sampling location	Country	Sampling date	Co-ordinates		Altitude (m)	Number of sequences (nb of extractions)	
							SSU	COI
Cyphoderia amphoralis	*Sphagnum* mosses, Rila	Bulgaria (BG)	August 2005	42°12’N	23°22’ E	1960	1(1)	4(3)
Cyphoderia ampulla	Aquatic mosses, Moiry	Switzerland (CH)	July 2006	46°08’N	07°34’ E	2310	2(2)	1(1)
Cyphoderia ampulla	*Sphagnum* mosses, Rhodopes	Bulgaria (BG)	July 2005	41°59’N	24°10’ E	1109	1(1)	2(2)
Cyphoderia ampulla	*Sphagnum* mosses, Vitosha	Bulgaria (BG)	August 2006	42°36’N	23°17’ E	1850	1(1)	2(2)
Cyphoderia ampulla	Underground waters of freshwater sand beach, Lake Geneva, St-Sulpice	Switzerland (CH)	May 2008	46°30’N	06°32’ E	375	3(3)	4(3)
Cyphoderia ampulla	Aquatic mosses, Cape Breton, Nova Scotia	Canada (CAN)	July 2008	46°48’N	60°49’ W	236	1(1)	2(1)
Cyphoderia cf. *compressa*	Underground waters of marine supralittoral sand beach, Tsawassen, Pacific Ocean	Canada (CAN)	October 2008	49°01’N	123°06’ W	0	1(1)	3(1)
Cyphoderia compressa	Underground waters of marine supralittoral sand beach, Galata, Black Sea	Bulgaria (BG)	July 2006	43°10’N	27°56’ E	0	3(3)	2(2)
Cyphoderia major	*Sphagnum* mosses, Rila	Bulgaria (BG)	August 2005	42°12’N	23°22’ E	1960	1(1)	1(1)
Cyphoderia trochus ssp. *palustris*	Wet mosses, Marchairuz	Switzerland (CH)	February 2007 and May 2008	46°33’N	06° 14’ E	1359	4(4)	6(4)
Pseudocorythion acutum	Underground waters of marine supralittoral sand beach, Galata, Black Sea	Bulgaria (BG)	May 2008	43°10’N	27°56’ E	0	2(2)	3(3)

Table 2. Sequences of the newly designed COI primers used in this study (mixed-base sites are denoted by the IUB nomenclature).

Name	Sequence (5′-3′)	Direction	Location (on *E. rotunda* AB009417)
Eucox1F	GAYATGGCKTTNCCAAGATTAAA	forward	232
Eucox5F	ACAGGWTGRACYRTTTATC	forward	331
Eucox6F	GAAGCWGGWGTWGGDACAGG	forward	316
Euglycox1R	AGCACCCATTGAHAAAACRTAATG	reverse	1081
Euglycox6R	GTTGGWACWGCTATWATCATWGT	reverse	874

Euglycox1R or Euglycox6R was applied in combination with Eucox5F and Eucox6F in a second PCR. The size of the obtained sequences varied because of the use of different combination of primers. All primers used in this study were newly designed (Table 2). The PCR cycling profile was the same for the first and second PCRs. It consisted of a 30 s initial denaturation step (95 °C); followed by 40 cycles of 95 °C for 30 s; 40 °C for 30 s; and 72 °C for 90 s and a final extension at 72 °C for 10 min. The PCR products were purified using the High Pure PCR Purification Kit (Roche, Basel, Switzerland) or the QIAquick PCR Purification Kit (Qiagen, Hilden, Germany) and then directly sequenced. Sequencing was carried out using a BigDye197 Terminator Cycle Sequencing Ready Reaction Kit (Applied Biosystems) and analysed either with an ABI-3130xl or a 3730S 48-capillary DNA sequencer (Applied Biosystems). COI sequences are deposited in GenBank with the accession numbers HM187644-HM187673.

Dataset constructions: Two data sets were used for phylogenetic analyses. The first one included thirty new COI fragments (500 bp). The second one included twenty SSU rDNA Euglyphida sequences (1600 bp). To test the effects of increasing alignment length on the phylogenetic resolutions of two main phylogenetic subclades, we analysed alignments of different lengths. Four additional COI alignments (100 bp, 200 bp, 300 bp and 400 bp) and sixteen SSU rDNA alignments (100 to 1600, in 100 bp steps) were built. The COI sequences were aligned using CLUSTALW (Thompson et al. 1994) and manually edited using the BIOEDIT 7.0.9 sequence alignment editor (Hall 1999). The COI sequences were easily aligned, as no insertions or deletions were detected. The SSU rDNA sequences were aligned as described in Heger et al. (2010). All trees were rooted with *Pseudocorythion acutum*, which has been shown to diverge prior to genus *Cyphoderia* (Heger et al. 2010).

Phylogenetic analyses: The best-fit models selected in MrAIC (Nylander 2004) for the COI data sets were either the General-Time-Reversible model with gamma distribution (GTR + G) or the General-Time-Reversible model with invariable sites (GTR + I + G). For the SSU rDNA datasets, the best-fit models were either the General-Time-Reversible model with gamma distribution (GTR + G), the General-Time-Reversible model with invariable sites and gamma distribution (GTR + I + G) or the Hasegawa-Kishino-Yano with invariable sites and gamma distribution (HKY + I + G). Maximum likelihood analyses were performed using Treefinder (Jobb et al. 2004). The bootstrap analysis option was used to assess nodal support on the ML trees (100 replicates, consensus level 50). In addition, Bayesian analyses were performed using the software MrBayes v. 3.1.2 (Huelsenbeck and Ronquist 2001) for the 500 bp COI and the 1600 SSU rDNA data sets. Three simultaneous Markov Chains were run for 10,000,000 generations with trees sampled every 10 generations. The first 250,000 trees were discarded as the burn in after checking that the chains had converged. The resultant trees were used to calculate the posterior probabilities (PP) for each node. The convergence of the Markov chains were graphically estimated by plotting the sample values versus the iteration values as well as by using diagnostics criteria produced by the "sump" command in MrBayes (PSRF = 1.00). Maximum-likelihood and Bayesian analyses were run through the Bioportal web-based service platform for phylogenomic analysis at the University of Oslo (www.bioportal.uio.no). The sequences dissimilarity values were obtained using BioEdit (Hall 1999). The COI data set was tested for nucleotide substitution saturation by plotting the estimated number of transitions and transversions against genetic divergence using DAMBE (Xia and Xie 2001). No saturation was observed for the COI data set (data not shown).

Acknowledgements

The authors wish to thank Jackie Guiard and José Fahrni for technical support and helpful discussions; Ralf Meisterfeld and Fabien Burki for fruitful discussions and Christophe Poupon for isolating some *Cyphoderia* specimens. We thank Barry Warner, Taro Asada and several members of the Leander lab for their help in the field. Thanks also to the CIME (EPFL) for technical help with the SEM. This work was funded by Swiss NSF projects n° 205321-109709 / 1 and 205321-109709 / 2 to E. Mitchell, PBELP2-122999 to T. Heger, PZ00P2_122042 to E. Lara and IB73A0-111064/1 (SCOPES) to J. Pawlowski, and the National Science and Engineering Research Council of Canada (NSERC) project n° 283091-09 to B.S. Leander. Additional funding to EM by CCES projects RECORD and BigLink is kindly acknowledged.

References

Barth D, Krenek S, Fokin SI, Berendonk TU (2006) Intraspecific genetic variation in *Paramecium* revealed by mitochondrial cytochrome c oxidase I sequences. J Eukaryot Microbiol **53**:20–25

Chantangsi C, Lynn DH, Brandl MT, Cole JC, Hetrick N, Ikonomi P (2007) Barcoding ciliates: a comprehensive study of

75 isolates of the genus *Tetrahymena*. Int J Syst Evol Microbiol **57**:2412–2425

Chantangsi C, Lynn DH (2008) Phylogenetic relationships within the genus *Tetrahymena* inferred from the cytochrome c oxidase subunit 1 and the small subunit ribosomal RNA genes. Mol Phylogenet Evol **49**:979–987

Evans KM, Mann DG (2009) A proposed protocol for nomenclaturally effective DNA barcoding of microalgae. Phycologia **48**:70–74

Evans KM, Wortley AH, Mann DG (2007) An assessment of potential diatom "barcode" genes (cox1, rbcL, 18S and ITS rDNA) and their effectiveness in determining relationships in *Sellaphora* (Bacillariophyta). Protist **158**:349–364

Evans KM, Wortley AH, Simpson GE, Chepurnov VA, Mann DG (2008) A molecular systematic approach to explore diversity within the *Sellaphora pupula* species complex (Bacillariophyta). J Phycol **44**:215–231

Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Mol Mar Biol Biotechnol **3**:294–299

Gómez A, Wright PJ, Lunt DH, Cancino JM, Carvalho GR, Hughes RN (2007) Mating trials validate the use of DNA barcoding to reveal cryptic speciation of a marine bryozoan taxon. Philos Trans R Soc B **274**:199–207

Guggiari M, Peck R (2008) The bacterivorous ciliate *Cyclidium glaucoma*, isolated from a sewage treatment plant: Molecular and cytological descriptions for barcoding. Eur J Protistol **44**:168–180

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

Hebert PDN, Ratnasingham S, deWaard JR (2003b) Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. Philos Trans R Soc B **270**:96–99

Hebert PDN, Cywinska A, Ball SL, DeWaard JR (2003a) Biological identifications through DNA barcodes. Philos Trans R Soc B **270**:313–321

Hebert PDN, Stoeckle MY, Zemlak TS, Francis CM (2004a) Identification of birds through DNA barcodes. PloS Biol **2**:1657–1663

Hebert PDN, Penton EH, Burns JM, Janzen DH, Hallwachs W (2004b) Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. Proc Natl Acad Sci USA **101**:14812–14817

Heger TJ, Mitchell EAD, Golemansky V, Todorov M, Lara E, Leander BS, Pawlowski J (2010) Molecular phylogeny of euglyphid testate amoebae (Cercozoa: Euglyphida) suggests transitions between marine supralittoral and freshwater/terrestrial environments are infrequent. Mol Phylogenet Evol **55**:113–122

Huelsenbeck JP, Ronquist F (2001) MRBAYES: Bayesian inference of phylogenetic trees. Bioinformatics **17**:754–755

Inagaki Y, Ehara M, Watanabe KI, Hasashi-Ishimaru Y, Ohama T (1998) Directionally evolving genetic code: The UGA codon from stop to tryptophan in mitochondria. J Mol Evol **47**:378–384

Ivanova NV, Zemlak TS, Hanner RH, Hebert PDN (2007) Universal primer cocktails for fish DNA barcoding. Mol Ecol Notes **7**:544–548

Jobb G, von Haeseler A, Strimmer K (2004) TREEFINDER: a powerful graphical analysis environment for molecular phylogenetics. BMC Evol Biol **4**:18

Kerr KCR, Stoeckle MY, Dove CJ, Weigt LA, Francis CM, Hebert PDN (2007) Comprehensive DNA barcode coverage of North American birds. Mol Ecol Notes **7**:535–543

Lara E, Heger TJ, Mitchell EAD, Meisterfeld R, Ekelund F (2007) SSU rRNA reveals a sequential increase in shell complexity among the euglyphid testate amoebae (Rhizaria: Euglyphida). Protist **158**:229–237

Lin S, Zhang H, Hou YB, Zhuang YY, Miranda L (2009) High-level diversity of dinoflagellates in the natural environment, revealed by assessment of mitochondrial cox1 and cob genes for dinoflagellate DNA barcoding. Appl Environ Microbiol **75**:1279–1290

Lynn DH, Strüder-Kypke MC (2006) Species of *Tetrahymena* identical by small subunit rRNA gene sequences are discriminated by mitochondrial cytochrome c oxidase I gene sequences. J Eukaryot Microbiol **53**:385–387

Meisterfeld R (2002) Testate Amoebae with Filopodia. In Lee JJ, Leedale GF, Bradbury P (eds) The Illustrated Guide to the Protozoa. Society of Protozoologists, Lawrence, Kansas, pp 1054–1084

Mitchell EAD, Charman DJ, Warner BG (2008) Testate amoebae analysis in ecological and paleoecological studies of wetlands: past, present and future. Biodivers Conserv **17**:2115–2137

Mitchell EAD, Gilbert D, Buttler A, Amblard C, Grosvernier P, Gobat JM (2003) Structure of microbial communities in *Sphagnum* peatlands and effect of atmospheric carbon dioxide enrichment. Microbial Ecol **46**:187–199

Nassonova E, Smirnov A, Fahrni J, Pawlowski J (2010) Barcoding amoebae: comparison of SSU, ITS and COI genes as tools for molecular identification of naked lobose amoebae. Protist **161**:102–115

Newmaster SG, Fazekas AJ, Steeves RAD, Janovec J (2008) Testing candidate plant barcode regions in the Myristicaceae. Mol Ecol Resour **8**:480–490

Nguyen-Viet H, Bernard N, Mitchell EAD, Badot PM, Gilbert D (2008) Effect of lead pollution on testate amoebae communities living in *Sphagnum fallax*: An experimental study. Ecotoxicol Environ Safe **69**:130–138

Nguyen-Viet H, Gilbert D, Bernard N, Mitchell EAD, Badot PM (2004) Relationship between atmospheric pollution characterized by NO_2 concentrations and testate amoebae density and diversity. Acta Protozool **43**:233–239

Nylander JAA (2004) MrAIC.pl. Program Distributed by the Author. Evolutionary Biology Centre, Uppsala University

Pecher WT, Robledo JAF, Vasta GR (2004) Identification of a second rRNA gene unit in the *Perkinsus andrewsi* genome. J Eukaryot Microbiol **51**:234–245

Remigio EA, Hebert PDN (2003) Testing the utility of partial COI sequences for phylogenetic estimates of gastropod relationships. Mol Phylogenet Evol **29**:641–647

Saunders GW (2005) Applying DNA barcoding to red macroalgae: a preliminary appraisal holds promise for future applications. Philos Trans R Soc B **360**:1879–1888

Scholin CA, Anderson DM, Sogin ML (1993) Two distinct small-subunit ribosomal RNA genes in the North American toxic dinoflagellate *Alexandrium fundyense* (Dinophyceae). J Phycol **29**:209–216

Shearer TL, Coffroth MA (2008) Barcoding corals: limited by interspecific divergence, not intraspecific variation. Mol Ecol Resour **8**:247–255

Smith MA, Fisher BL, Hebert PDN (2005) DNA barcoding for effective biodiversity assessment of a hyperdiverse arthropod group: the ants of Madagascar. Philos Trans R Soc Ser B **360**:1825–1834

Tautz D, Arctander P, Minelli A, Thomas RH, Vogler AP (2003) A plea for DNA taxonomy. Trends Ecol Evol **18**:70–74

Thompson JD, Higgins DG, Gibson TJ (1994) Clustal W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res **22**:4673–4680

Todorov M, Golemansky V, Mitchell EAD, Heger TJ (2009) Morphology, biometry and taxonomy of freshwater and marine interstitial *Cyphoderia* (Cercozoa: Euglyphida). J Eukaryot Microbiol **56**:279–289

Wilkinson DM (2008) Testate amoebae and nutrient cycling: peering into the black box of soil ecology. Trends Ecol Evol **23**:596–599

Wylezich C, Meisterfeld R, Meisterfeld S, Schlegel M (2002) Phylogenetic analyses of small subunit ribosomal RNA coding regions reveal a monophyletic lineage of euglyphid testate amoebae (order Euglyphida). J Eukaryot Microbiol **49**:108–118

Xia X, Xie Z (2001) DAMBE: Software package for data analysis in molecular biology and evolution. J Hered **92**:371–373

Yoon HS, Nakayama T, Reyes-Prieto A, Andersen RA, Boo SM, Ishida K, Bhattacharya D (2009) A single origin of the photosynthetic organelle in different *Paulinella* lineages. BMC Evol Biol **9**:98