International Journal of Systematic and Evolutionary Microbiology (2010), **60**, 1962–1977 **DOI** 10.1099/ijs.0.013888-0

An SSU rDNA barcoding approach to the diversity of marine interstitial cercozoans, including descriptions of four novel genera and nine novel species

Chitchai Chantangsi† and Brian S. Leander

Departments of Zoology and Botany, University of British Columbia, Biological Sciences Bldg, 6270 University Blvd, Vancouver, BC V6T 1Z4, Canada

Correspondence
Chitchai Chantangsi
chitchai.c@chula.ac.th *or*
chantangsi01@hotmail.com

Environmental DNA surveys have revealed a great deal of hidden diversity within the Cercozoa. An investigation into the biodiversity of heterotrophic flagellates in marine benthic habitats of British Columbia, Canada, demonstrated the presence of several undescribed taxa with morphological features that resemble the cercozoan genera *Cryothecomonas* and *Protaspis*. Nine novel species of marine interstitial cercozoans are described that are distributed into five genera, four of which are new. Phylogenetic analyses of small subunit rDNA sequences derived from two uncultured isolates of *Protaspis obliqua* and nine novel cercozoan species (within four novel genera) provided organismal anchors that helped establish the cellular identities of several different environmental sequence clades. These data, however, also showed that the rarity of distinctive morphological features in cryomonads, and other groups of cercozoans, makes the identification and systematics of the group very difficult. Therefore, a DNA barcoding approach was applied as a diagnostic tool for species delimitation that used a 618 bp region at the 5′ end of the SSU rDNA sequence. Nucleotide sequence analysis of this region showed high intergeneric sequence divergences of about 7 % and very low intraspecific sequence divergences of 0–0.5 %; phylogenetic analyses inferred from this barcoding region showed very similar tree topologies to those inferred from the full-length sequence of the gene. Overall, this study indicated that the 618 bp barcoding region of SSU rDNA sequences is a useful molecular signature for understanding the biodiversity and interrelationships of marine benthic cercozoans.

INTRODUCTION

Marine benthic environments are complex and poorly understood ecosystems that contain a diverse array of micro-organisms (Fenchel, 1987; Hondeveld *et al.*, 1992). Environmental DNA surveys of these systems have demonstrated a significant degree of hidden diversity, especially of so-called cercozoan amoeboflagellates (Bass & Cavalier-Smith, 2004; Berney *et al.*, 2004; Šlapeta *et al.*, 2005). The group Cercozoa was initially recognized as a monophyletic assemblage based on molecular phylogenetic studies inferred from rDNA genes (Cavalier-Smith, 1998a, b). Subsequent studies have shown that several protists *incertae sedis* are also members of this group (e.g. organisms of the genus *Ebria*) and that there are many novel environmental sequence clades that have yet to be described at the microscopic level (Bass & Cavalier-Smith, 2004; Hoppenrath & Leander, 2006a, b). Cryomonads (e.g. members of the genus *Protaspis*) are a group of cercozoan biflagellates that glide along substrates with heterodynamic flagella and are common predators in marine benthic habitats (Hoppenrath & Leander, 2006a). These particular cercozoans possess several morphological characteristics (e.g. flattened cell shape, mode of locomotion and a conspicuous nucleus with condensed chromatin) that resemble distantly related eukaryotes living in the same environments, namely some phagotrophic euglenids, dinoflagellates and katablepharids. These convergent morphological and behavioural features have led to difficulties in the classification and identification of benthic biflagellates in general, but especially in cercozoan lineages, which are very poorly understood.

†Present address: Department of Biology, Faculty of Science, Chulalongkorn University, Phayathai Road, Pathumwan, Bangkok, 10330 Thailand.

Abbreviations: *cox1*, cytochrome *c* oxidase subunit 1; K2P, Kimura two-parameter; SSU rDNA, small subunit rDNA; LM, light microscopy; ML, maximum-likelihood.

The GenBank/EMBL/DDBJ accession numbers for the SSU rDNA sequences of the marine benthic interstitial cercozoans examined in this study are FJ824121–FJ824131.

A supplementary table giving details of the SSU rDNA nucleotide sequences included in the analyses performed for this study is available with the online version of this paper.

DNA barcoding has been proposed as an alternative and more precise approach for the delimitation and identification of species. This strategy is expected to be particularly advantageous for understanding the diversity of uncultured microeukaryotic lineages that lack sufficient morphological details for species delimitation (Godfray, 2002; Hebert *et al.*, 2003a, b; Tautz *et al.*, 2002, 2003). Hebert *et al.* (2003b) proposed the mitochondrial cytochrome *c* oxidase subunit 1 (*cox1*) gene as a pragmatic and widely applicable 'DNA barcode' for animal species. Although this gene has proved useful for the identification of animal species and even some protists (Barth *et al.*, 2006; Chantangsi *et al.*, 2007; Lynn & Strüder-Kypke, 2006; Saunders, 2005), the gene is not applicable for all groups of eukaryotes due to several limitations (Scicluna *et al.*, 2006). For example, some groups of eukaryotes are amitochondriate and therefore lack the *cox1* gene, and most other groups of eukaryotes have yet to be studied at the level of the mitochondrial genome. Moreover, a few groups of eukaryotes, such as ciliates, have different-sized insertions within the *cox1* gene that create difficulties in the PCR amplification of the barcode sequences (Cummings, 1992; Norman & Gray, 1997). In addition, the mitochondrial genome of a free-living diplonemid *Diplonema papillatum* has been found to show fragmented coding regions for the *cox1* gene (Marande & Burger, 2007).

Nuclear small subunit rDNA (SSU rDNA) sequences are the molecular markers that are most widely used to study the phylogenetic relationships of eukaryotes and prokaryotes. The SSU rRNA gene is present in all organisms and plays a highly conserved role in protein translation that is critical for the survival of cells; therefore, this sequence can be compared across the entire tree of life. This gene is also present in numerous copies within the genome and contains highly conserved regions that facilitate the design of universal primers for PCR amplification (Long & David, 1980; Minelli, 1993; Sogin *et al.*, 1986). SSU rDNA sequences also contain regions of sequence variation that are sufficient and advantageous for DNA barcoding; for instance, Scicluna *et al.* (2006) demonstrated the utility of a 600 bp segment of SSU rDNA sequences for identifying subtypes of the *Blastocystis hominis* species complex.

In order to better understand the microeukaryotic components of marine benthic ecosystems, we investigated the phenotypic and genetic diversity of uncultured marine interstitial cercozoans that showed morphological and behavioural similarities to previously described species of the genera *Protaspis* and *Cryothecomonas* (Hoppenrath & Leander, 2006a). In particular, we examined the potential for coupling high resolution light microscopy (LM) with SSU rDNA barcodes (e.g. a 5′-618 bp fragment) for facilitating the systematics of these uncultured lineages. We also used light micrographs and SSU rRNA gene sequences to help establish the cellular identities of several environmental DNA sequence clades. Overall, this combined approach enabled us to (i) describe nine novel species within four novel genera and one described genus of uncultured marine interstitial cercozoans and (ii) establish an efficient and effective protocol for advancing the systematics of this group.

METHODS

Sampling and LM. Sand samples were collected from several habitats around British Columbia, Canada, during 2006–2007. Organisms were extracted from the sand samples through a 48 μm mesh using a melted seawater-ice method as described by Uhlig (1964). Briefly, 2–3 spoons of sand samples were placed into an extraction column wrapped with the mesh. Two to three seawater ice cubes were then put on top of the sand samples and left to melt over several hours. The organisms of interest were separated through the mesh and concentrated in a Petri dish that was filled with seawater and placed underneath the extraction column. The Petri dish containing the organisms was then observed using a Leica DMIL inverted microscope. Cells were isolated individually and placed on a slide for LM using phase-contrast and differential interference contrast (DIC) microscopy with a Zeiss Axioplan 2 imaging microscope connected to a Leica DC500 colour digital camera.

DNA extraction and PCR amplification. Cells were isolated individually and washed three times in autoclaved filtered seawater. The numbers of cells from which different SSU rDNA sequences were obtained were as follows: *Protaspis obliqua* (isolate 1) one cell; *P. obliqua* (isolate 2) 10 cells; the number of isolated cells was not recorded for *Protaspis rotunda* sp. nov.; *Protaspis obaniformis* sp. nov. four cells; *Protaspis oviformis* sp. nov. 25 cells; *Botuliforma benthica* gen. et sp. nov. 14 cells; *Ventrifissura artocarpoidea* gen. et sp. nov. one cell; *Ventrifissura foliiformis* gen. et sp. nov. 12 cells; *Verrucomonas bifida* gen. et sp. nov. 15 cells; *Verrucomonas longifila* gen. et sp. nov. one cell; and *Discomonas retusa* gen. et sp. nov. one cell. DNA was extracted using the protocol provided in the Total Nucleic Acid Purification kit by EPICENTRE. PCR with the final reaction volume of 25 μl was performed for two rounds in a thermal cycler using puReTaq Ready-To-Go PCR beads (GE Healthcare Bio-Sciences, Inc.). The first round PCR was conducted using forward (i.e. NPF1 or PF1) and reverse primers (i.e. R4 or FAD) as listed in Table 1. Then, either direct PCR or a gel-purified band of an 1850 bp region from the first round PCR was used as a template for the second round PCR with the appropriate primers provided in Table 1. The thermal cycler was programmed as follows: hold at 94 °C for 4 min; five cycles of denaturation at 94 °C for 30 s, annealing at 45 °C for 1 min, and extension at 72 °C for 105 s; 35 cycles of denaturation at 94 °C for 30 s, annealing at 55 °C for 1 min, and extension at 72 °C for 105 s; and hold at 72 °C for 10 min. PCR products corresponding to the expected size were separated by agarose gel electrophoresis and cleaned using the UltraClean 15 DNA Purification kit (MO BIO Laboratories, Inc.). The cleaned DNA was cloned into a pCR2.1 vector using TOPO TA Cloning kits (Invitrogen Corporation). Plasmids with the correct insert size were sequenced using BigDye 3.1 and the vector forward and reverse primers, and internal primers with an Applied Biosystems 3730S 48-capillary sequencer (Table 1).

Sequence alignment. Sequences were assembled and edited using Sequencher (version 4.5, Gene Codes Corporation). Acquired sequences were initially identified by BLAST analysis. New SSU rDNA sequences derived from the newly found cercozoans were aligned using CLUSTAL W (Thompson *et al.*, 1994) implemented in the MEGA version 4 program (Tamura *et al.*, 2007) and further refined by eye. Four multiple sequence alignments were created for phylogenetic analyses: (i) a 69-taxon global alignment comprising sequences of representatives from all major eukaryotic groups (1134 unambiguous sites) was constructed to determine the phylogenetic affinities of the

Table 1. Oligonucleotide primers used for amplification and sequencing of SSU rDNA in this study

Annealing region was provided with reference to the SSU rDNA sequence of *Protaspis rotunda* sp. nov. (GenBank accession no. FJ824123).

Primer	Direction	Sequence 5′–3′	Annealing region
NPF1*	Forward	5′-TGCGCTACCTGGTTGATCC-3′	1–19
PF1	Forward	5′-GCGCTACCTGGTTGATCCTGCC-3′	2–23
525F	Forward	5′-AAGTCTGGTGCCAGCAGCC-3′	567–585
917FD*	Forward	5′-GCCAGAGGTGAAATTCTNGG-3′	917–936
1050F	Forward	5′-GGGGGAGTATGGTCGCAAG-3′	1130–1148
1050FD*	Forward	5′-GGGGGAGTATGGTCGCRAG-3′	1130–1148
1050MRD*	Reverse	5′-GCCTYGCGACCATACTCC-3′	1150–1133
1134FD*	Forward	5′-CGCAAGGCTGAAACHTRAAGG-3′	1143–1163
nomet1134R	Reverse	5′-TTTAAGTTTCAGCCTTGCG-3′	1161–1143
1242RD*	Reverse	5′-GTCYGGACCTGGTAAGTTTTC-3′	1242–1222
1250R	Reverse	5′-TAACGGAATTAACCAGACA-3′	1342–1324
1367RD*	Reverse	5′-TTTAGYAGGBCGAGGTCTCG-3′	1367–1348
R4	Reverse	5′-GATCCTTCTGCAGGTTCACCTAC-3′	1823–1801
FAD	Reverse	5′-TGATCCTTCTGCAGGTTCACCTAC-3′	1824–1801

*Primers newly designed in this study.

newly isolated organisms to other eukaryotic groups; (ii) a 67-taxon cercozoan alignment consisting of cercozoan representatives and extensive short environmental sequences about 1069 bp in length (923 unambiguous sites) was constructed to determine phylogenetic affinities of the newly isolated organisms to uncharacterized taxa represented by sequences derived from environmental studies; (iii) a 35-taxon cercozoan alignment covering representatives from different cercozoan subgroups and excluding the shorter and unrelated environmental sequences (1617 unambiguous sites) was constructed to determine more robustly the phylogenetic relationships among the newly described taxa; and (iv) a 35-taxon cercozoan alignment including only the barcoding regions of 618 bp in length (583 unambiguous sites) was constructed to compare the topologies of phylogenetic relationships inferred from the 1617 bp alignment 3. Alignments 3 and 4 contained the same composition of examined taxa. All ambiguous sites were excluded from the alignments prior to phylogenetic analyses.

As for barcoding analyses, two alignment files were created for genetic distance calculations: (i) a 17-taxon cercozoan alignment of the almost complete SSU rDNA gene including 1798 bp (1745 unambiguous sites) and (ii) a 17-taxon cercozoan alignment including only the 5′-half 618 bp barcoding regions (608 unambiguous sites). Both datasets had the same composition of examined taxa. The second barcoding alignment file was constructed based on diagnostic barcoding regions of *Blastocystis hominis* (Scicluna *et al.*, 2006): its starting position was 22 bp inside the beginning of the *B. hominis* barcoding region and extended 65 bp further from the end of the *B. hominis* barcoding region. The alignment files are available upon request.

Phylogenetic analyses. MrBayes version 3.1.2 was used to perform Bayesian analyses on all four datasets (Huelsenbeck & Ronquist, 2001; Ronquist & Huelsenbeck, 2003). Four Markov Chain Monte Carlo (MCMC) chains, one cold chain and three heated chains, were run for 2 000 000 generations, sampling every 50th generation (tree). The first 4000 trees were discarded as burn-in. The remaining trees were used to compute the 50 % majority-rule consensus tree. Branch lengths of the trees were saved.

Maximum-likelihood (ML) analyses were performed on all four datasets using PhyML (Guindon & Gascuel, 2003). Input trees for each dataset were generated by BIONJ with optimization of topology, branch lengths and rate parameters selected. The General Time Reversible (GTR) model of nucleotide substitution was chosen. The proportion of invariable sites and the gamma distribution parameter were estimated from the input dataset. Eight categories of substitution rates were selected. PhyML bootstrap trees with 100 bootstrap datasets were constructed using the same parameters as the individual ML trees.

Genetic distance analyses. Sequence divergences were calculated for 1798 bp and 618 bp datasets using the Kimura two-parameter (K2P) distance model with complete deletion of positions with gaps included (Kimura, 1980).

Sequence availability. The SSU rDNA nucleotide sequences included in the analyses performed for this paper are detailed in Supplementary Table S1 (see IJSEM Online). The GenBank accession numbers for the 11 new sequences obtained from this study are: *Protaspis obliqua* isolate 1 (1826 bp; FJ824121), *Protaspis obliqua* isolate 2 (1826 bp; FJ824122), *Protaspis rotunda* sp. nov. (1824 bp; FJ824123), *Protaspis obaniformis* sp. nov. (1823 bp; FJ824124), *Protaspis oviformis* sp. nov. (1823 bp; FJ824125), *Botuliforma benthica* gen. et sp. nov. (1827 bp; FJ824126), *Ventrifissura artocarpoidea* gen. et sp. nov. (1824 bp; FJ824127), *Ventrifissura foliiformis* gen. et sp. nov. (1822 bp; FJ824128), *Verrucomonas bifida* gen. et sp. nov. (1827 bp; FJ824129), *Verrucomonas longifila* gen. et sp. nov. (1828 bp; FJ824130) and *Discomonas retusa* gen. et sp. nov. (1824 bp; FJ824131).

RESULTS

Identification of the examined flagellates

Eleven different isolates of uncultured cercozoan flagellates were isolated and characterized by LM (Figs 1–3) and SSU rDNA gene sequences (Table 2; Figs 4–5). All organisms were found gliding in marine interstitial habitats, except for *Botuliforma benthica* gen. et sp. nov., which showed a

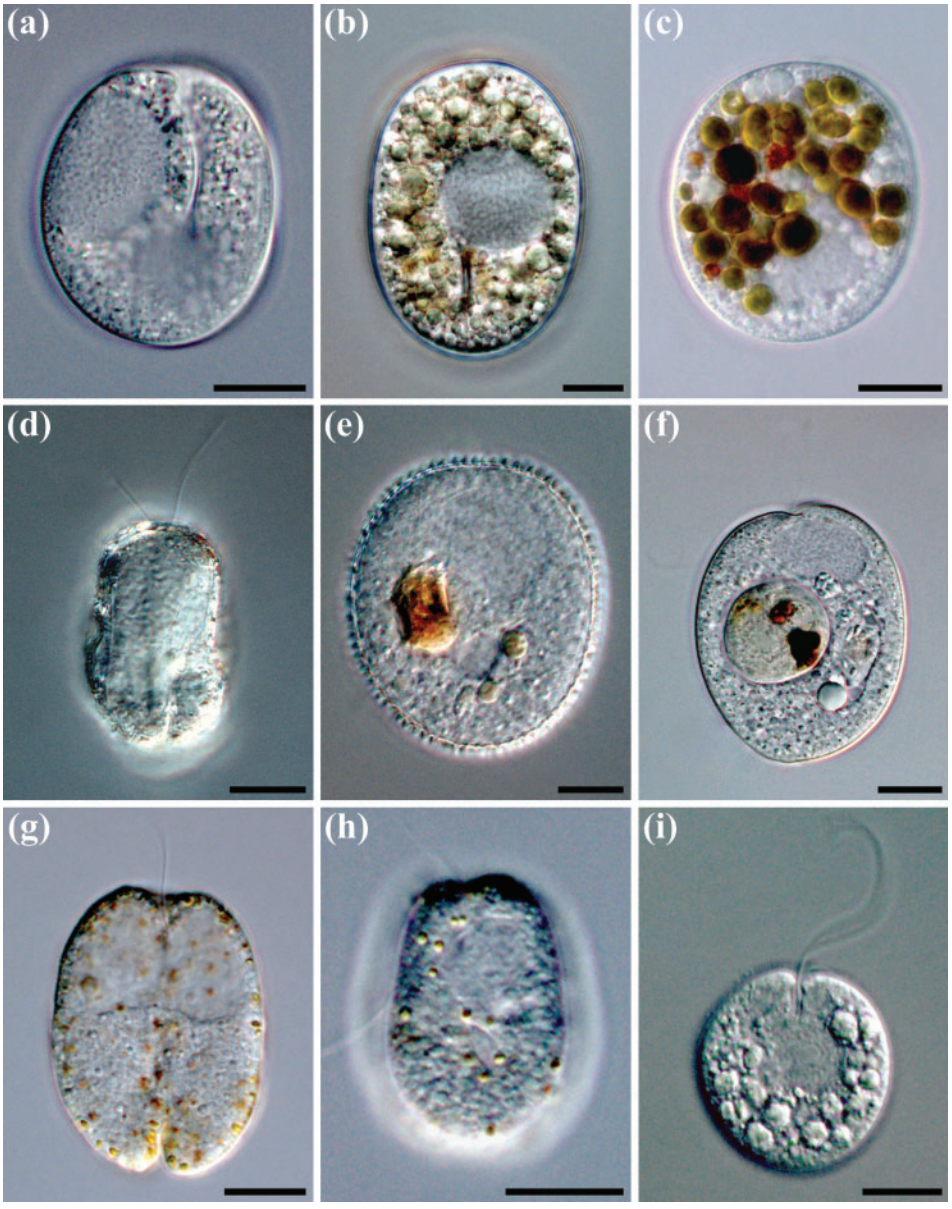

Fig. 1. Light micrographs of nine novel cercozoans found in this study. (a) *Protaspis rotunda* sp. nov. showing a ventral view with an ovoid nucleus at the anterior right side of the cell and a curved slit in the middle of the cell. (b) *Protaspis obaniformis* sp. nov. showing granulated cytoplasm, spherical nucleus in the middle of the cell and a posterior ventral slit. (c) *Protaspis oviformis* sp. nov. showing food particles with varying colours within its cytoplasm. (d) *Botuliforma benthica* gen. et sp. nov. showing two anterior flagella, a large anterior nucleus and a thick and rough cell wall. (e) *Ventrifissura artocarpoidea* gen. et sp. nov. showing a cell with numerous pointed warts and an anterior spherical nucleus. (f) *Ventrifissura foliiformis* gen. et sp. nov. showing a cell with a smooth cell surface, large food particle and a small anterior nucleus. (g) *Verrucomonas bifida* gen. et sp. nov. showing numerous warts of varying colours on the cell surface, a ventral furrow, a bilobed nucleus and a notch at its posterior end. (h) *Verrucomonas longifila* gen. et sp. nov. showing two flagella, an ovoid nucleus at the anterior of the cell and several yellowish warts on cell surface. (i) *Discomonas retusa* gen. et sp. nov. showing a prominent long anterior flagellum, a discoidal nucleus at the anterior of the cell and numerous spherical granules distributed around the periphery of the cell. Bars, 10 μm.

rotational pattern of swimming. The DNA sequence data demonstrated that the newly discovered cercozoan flagellates clustered with a diverse assortment of environmental DNA sequences, forming one previously established clade and four novel clades, as described below (Figs 4–5). These molecular data served as our primary taxonomic guides and enabled us to establish nine novel species that are challenging to distinguish at the morphological level.

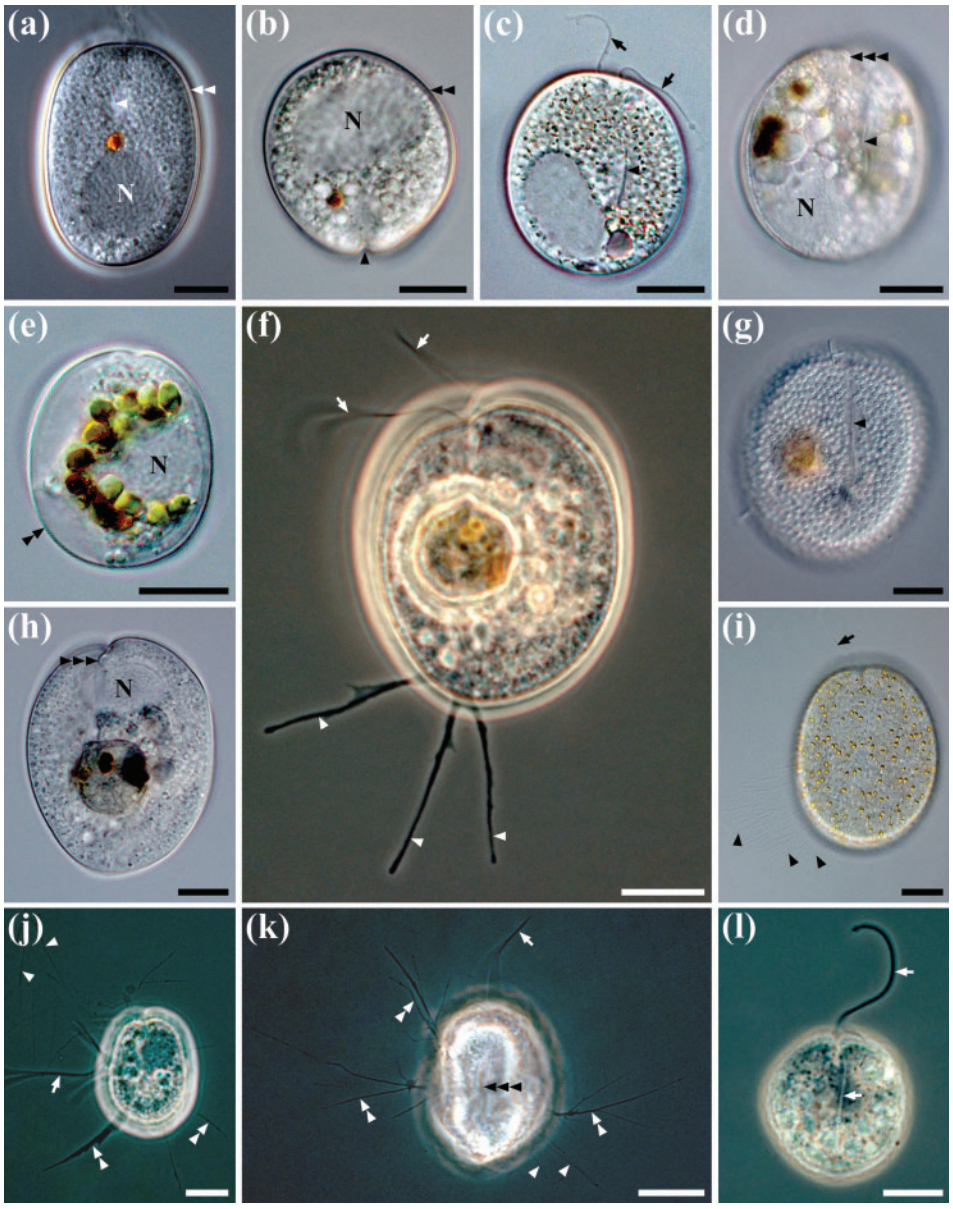

Fig. 2. Light micrographs of the marine benthic cercozoans examined in this study. (a) *Protaspis grandis* showing an ovoid nucleus (N) at the posterior of the cell, a ventral slit (arrowhead) and a thick cell wall (double arrowhead). (b) *Protaspis obliqua* showing an anterior nucleus (N), a posterior notch (arrowhead) and a thick cell wall (double arrowhead). (c) *Protaspis rotunda* sp. nov. showing two flagella (arrows) and a curved slit in the middle of the cell (arrowhead). (d) *Protaspis oviformis* sp. nov. showing a posterior nucleus (N), a ventral slit (arrowhead) and an anterior protrusion (triple arrowhead). (e) *Protaspis oviformis* sp. nov. showing a nucleus (N) in the middle of the cell, a thick cell wall (double arrowhead) and numerous ingested green food bodies within the cytoplasm. (f) *Ventrifissura foliiformis* gen. et sp. nov. showing two flagella (arrows) and branched pseudopodia (arrowheads). (g) *Ventrifissura artocarpoidea* gen. et sp. nov. showing a cell with numerous pointed warts and a ventral slit (arrowhead). (h) *Ventrifissura foliiformis* gen. et sp. nov. showing a cell with a smooth cell surface, a large food particle, an anterior nucleus (N) and an anterior protrusion (triple arrowhead). This image was taken through the dorsal side and focused at the ventral side of the cell. (i) *Verrucomonas bifida* gen. et sp. nov. showing an anterior flagellum (arrow), numerous orange warts on the cell surface and ejected extrusomes (arrowheads). (j) *Verrucomonas longifila* gen. et sp. nov. showing the posterior flagellum (arrow), fine pseudopodia (double arrowheads) and ejected extrusomes (arrowheads). (k) *Botuliforma benthica* gen. et sp. nov. showing one of two anterior flagella (arrow), ejected extrusomes (arrowheads), very fine and branched pseudopodia (double arrowheads) and a ventral groove (triple arrowhead). (l) *Discomonas retusa* gen. et sp. nov. showing a prominent long anterior flagellum and a shorter posterior flagellum (arrows). Bars, 10 μm.

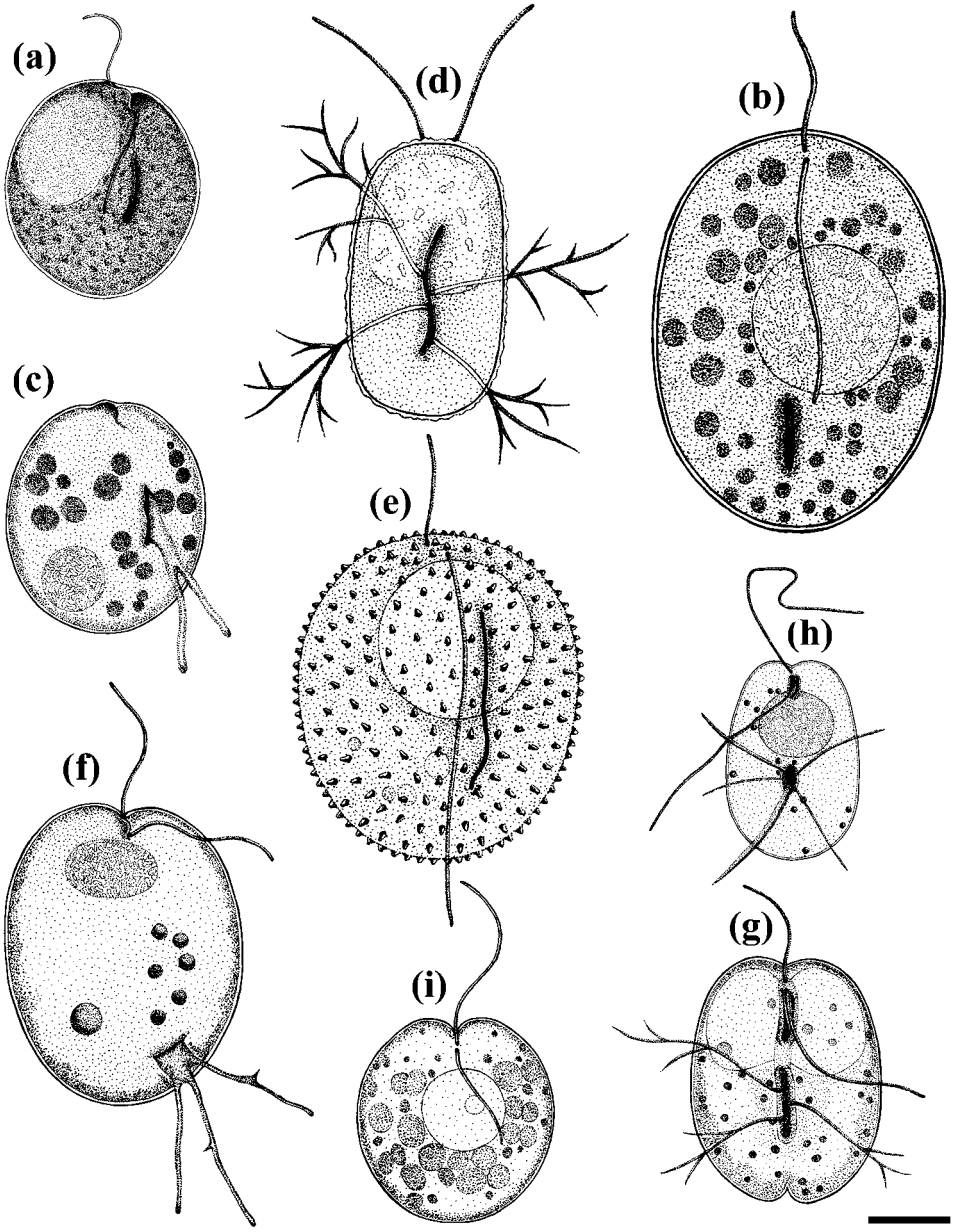

Fig. 3. Diagrammatic line drawings of the nine novel cercozoans found in this study. (a) *Protaspis rotunda* sp. nov. (b) *Protaspis obaniformis* sp. nov. (c) *Protaspis oviformis* sp. nov. (d) *Botuliforma benthica* gen. et sp. nov. (e) *Ventrifissura artocarpoidea* gen. et sp. nov. (f) *Ventrifissura foliiformis* gen. et sp. nov. (g) *Verrucomonas bifida* gen. et sp. nov. (h) *Verrucomonas longifila* gen. et sp. nov. (i) *Discomonas retusa* gen. et sp. nov. Bar, 10 μm.

The Cryomonadida clade

Phylogenies inferred from SSU rDNA sequences demonstrated that one previously described species, *Protaspis obliqua*, and the following three novel species of the genus *Protaspis* clustered within the Cryomonadida clade, which also contained *Protaspis grandis* and several recognized species of the genus *Cryothecomonas* (Figs 4–5). The genus *Protaspis* currently contains 11 recognized species: *P. gemmifera*, *P. glans*, *P. grandis*, *P. maior*, *P. metarhiza*, *P. obliqua*, *P. obovata*, *P. simplex*, *P. tanyopsis*, *P. tegere* and *P. verrucosa*.

Protaspis rotunda sp. nov. Chantangsi and Leander, 2009

Size. Cells 25–27 μm wide and 30–31 μm long (number of cells observed >100).

Diagnosis. Cells dorsoventrally flattened and slightly oval to round in outline with smooth surface; uninucleate biflagellate; flagella inserted subapically separated by an anterior protrusion; nucleus is 11 μm wide and 15 μm long; nucleus is ovoid and located at anterior to the right of

Table 2. Upper triangular matrix showing the number of nucleotide differences; lower triangular matrix showing the percentage of pairwise sequence divergences between small subunit rRNA genes based on Kimura's two-parameter model

Data are from 17 cercozoans, 11 of which were generated in this study. The sequences were 618 bp in length. Taxa: 1, *Cryothecomonas aestivalis* (GenBank accession no. AF290541); 2, *Cryothecomonas aestivalis* (AF290539); 3, *Protaspis* (ex. *Cryothecomonas*) *longipes* (AF290540); 4, *Protaspis grandis* (DQ303924); 5, *Protaspis obliqua* isolate 1 (FJ824121); 6, *Protaspis obliqua* isolate 2 (FJ824122); 7, *Protaspis rotunda* (FJ824123); 8, *Protaspis obaniformis* (FJ824124); 9, *Protaspis oviformis* (FJ824125); 10, *Ebria tripartita* (DQ303922); 11, *Ebria tripartita* (DQ303923); 12, *Botuliforma benthica* (FJ824126); 13, *Ventrifissura artocarpoidea* (FJ824127); 14, *Ventrifissura foliiformis* (FJ824128); 15, *Verrucomonas bifida* (FJ824129); 16, *Verrucomonas longifila* (FJ824130); 17, *Discomonas retusa* (FJ824131).

Taxa	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1.	–	3	21	21	30	29	25	27	19	38	38	33	43	44	74	65	55
2.	0.5	–	18	18	27	26	22	24	16	35	35	30	42	41	71	64	52
3.	3.55	3.03	–	3	21	20	4	7	2	32	32	28	41	38	70	65	50
4.	3.55	3.03	0.5	–	21	20	7	7	4	32	32	28	42	39	70	65	50
5.	5.13	4.61	3.54	3.54	–	1	25	25	19	37	37	34	46	45	69	66	48
6.	4.95	4.43	3.37	3.37	0.16	–	24	24	18	36	36	33	45	44	70	67	49
7.	4.25	3.73	0.66	1.16	4.24	4.06	–	11	6	34	34	30	45	42	72	69	54
8.	4.61	4.08	1.16	1.16	4.24	4.06	1.83	–	8	38	38	32	46	43	75	70	53
9.	3.2	2.69	0.33	0.66	3.2	3.02	0.99	1.33	–	32	32	28	39	36	69	64	48
10.	6.55	6.01	5.47	5.47	6.38	6.2	5.82	6.55	5.47	–	0	16	49	47	67	64	42
11.	6.55	6.01	5.47	5.47	6.38	6.2	5.82	6.55	5.47	0	–	16	49	47	67	64	42
12.	5.65	5.12	4.76	4.76	5.83	5.65	5.11	5.46	4.76	2.68	2.68	–	40	41	70	69	48
13.	7.46	7.28	7.09	7.27	8.02	7.83	7.82	8.02	6.72	8.6	8.6	6.92	–	27	78	73	61
14.	7.62	7.08	6.53	6.71	7.81	7.63	7.25	7.44	6.17	8.18	8.18	7.08	4.59	–	76	72	54
15.	13.36	12.77	12.55	12.55	12.38	12.58	12.94	13.57	12.36	11.98	11.98	12.56	14.2	13.74	–	18	80
16.	11.59	11.41	11.59	11.59	11.81	12.01	12.38	12.6	11.4	11.42	11.42	12.38	13.19	12.95	3.03	–	77
17.	9.7	9.14	8.75	8.75	8.38	8.57	9.5	9.33	8.38	7.28	7.28	8.38	10.85	9.48	14.56	13.98	–

the cell; clear and homogeneous cytoplasm; curved furrow at the subanterior ventral side; locomotion by gliding; found living in a marine interstitial sand habitat. Small-subunit rRNA gene sequence (GenBank accession no. FJ824123). *Protaspis rotunda* differs from the 11 previously described species of the genus *Protaspis* in that this species shows a smooth cell surface. Although similar to *P. oviformis* in shape and cell appearance, the nuclear position of *P. rotunda* is at anterior to the right of the cell as opposed to at posterior or sometimes the middle right of the cell in *P. oviformis*. Cell shapes of *P. grandis* and *P. obaniformis* are oblong rather than the oval shape of *P. rotunda*.

Type locality. Tidal sand-flat at Pachena Beach (48° 47′ N 125° 07′ W), Vancouver Island, British Columbia, Canada. The organisms were collected in September 2006. This species was also found in June 2007 and June 2008 from the same location.

Iconotype. Figs 1(a), 2(c) and 3(a).

Etymology. The etymology for the specific epithet, L. fem. adj. *rotunda* round. The specific epithet reflects the round shape of this organism.

Protaspis obaniformis sp. nov. Chantangsi and Leander, 2009

Size. Cells 30–50 μm wide and 50–65 μm long (number of cells observed=19).

Diagnosis. Cells are broadly elliptical and dorsoventrally flattened; thick wall with smooth surface; uninucleate biflagellate; flagella inserted subapically; nucleus is 18 μm wide and 17 μm long; circular nucleus with granular (permanently condensed) chromosomes; nucleus is located in the middle of the cell, sometimes towards the anterior of the cell; cytoplasm contains numerous spherical light brownish granules; prominent vertically straight slit at the posterior ventral side; locomotion by gliding; found living in marine interstitial sand habitat. Small-subunit rRNA gene sequence (GenBank accession no. FJ824124). *Protaspis obaniformis* differs from the 11 previously described species of the genus in that this species possesses an 1/5 cell length posterior slit. Although *Protaspis obliqua* also shows a ventral furrow/groove at the posterior half of the cell, the posterior notch of *P. obliqua* has never been observed in *P. obaniformis*. In addition, cell shape and cell size of the former (i.e. 10–27 μm wide and 8–32 μm long) are different from the latter.

Type locality. Tidal sand-flat at Boundary Bay (49° 00′ N 123° 02′ W), Vancouver, British Columbia, Canada. The organisms were collected in May 2007.

Iconotype. Figs 1(b) and 3(b).

Etymology. The etymology for the specific epithet, Japanese *oban* is a kind of ancient Japanese coin with a

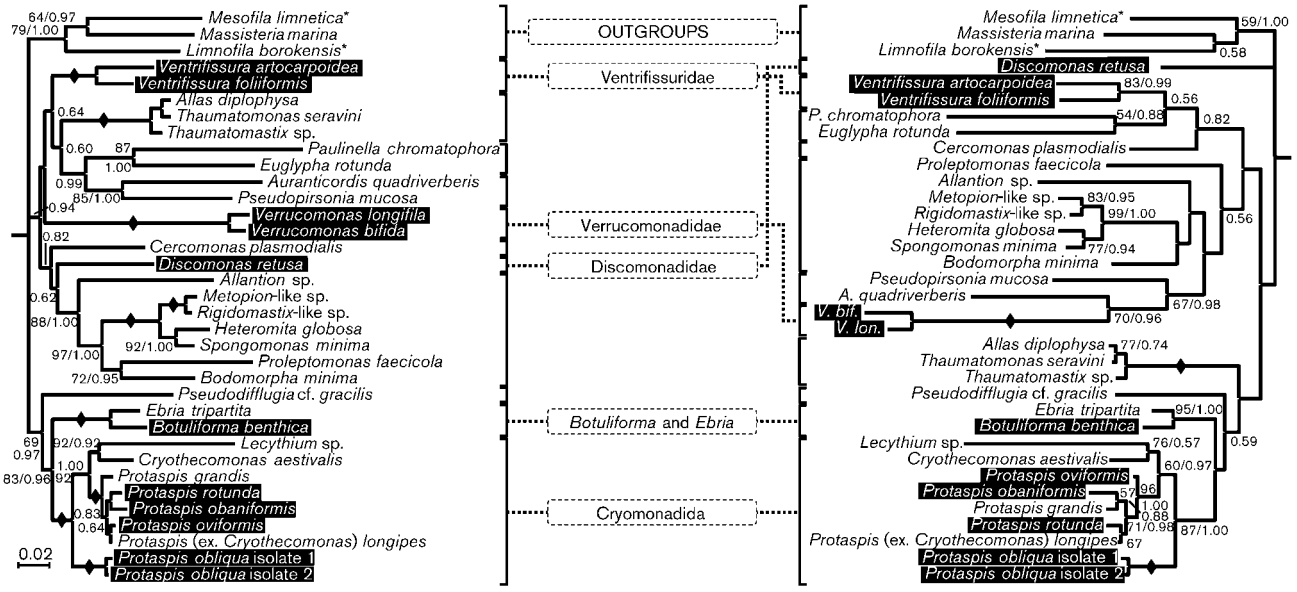

Fig. 4. Comparison between Bayesian phylogenies inferred from the 1617 bp full-length (left; mean ln $L=-11\,672.98$) and 583 bp 5′-half barcoding region (right; mean ln $L=-4072.00$) SSU rDNA sequence alignments of 35 cercozoan taxa; the phylogenetic positions of marine benthic cercozoans studied in this study are highlighted in black boxes. Each tree is a consensus of 36 002 trees with the GTR+I+G using four rate categories implemented. Two isolates of *Protaspis obliqua* are labelled as isolates 1 and 2. Numbers >0.50 at the nodes indicate Bayesian posterior probabilities and PhyML bootstrap percentages >50 %. Diamonds represent Bayesian posterior probability of 1.00 and PhyML bootstrap value of 100 %. Bar, 0.02 substitutions per site. **Limnofila borokensis* was previously misidentified as *Gymnophrys cometa* (GenBank accession no. AF411284) and *Mesofila limnetica* was previously referred to as *Dimorpha*-like sp. (AF411283).

broad oval shape; L. fem. *forma* shape. The specific epithet, N.L. fem. *obaniformis*, depicts the shape of this organism, which is similar to an ancient Japanese coin.

Protaspis oviformis sp. nov. Chantangsi and Leander, 2009

Size. Cell is 25–40 μm wide and 30–45 μm long (number of cells observed=36).

Diagnosis. Cells are dorsoventrally flattened and slightly oval to roundish in outline with smooth surface; slightly thick wall; uninucleate cell with an anterior protrusion; nucleus is 8 μm wide and 8 μm long; nucleus with chromosome appearance is circular in outline and located at posterior or sometimes middle right of the cell; cytoplasm contains numerous spherical brownish, greyish and yellowish granules; vertical ventral slit positioned at the mid anterior end and towards the mid posterior end; finger-like pseudopodia present; locomotion by gliding; found living in marine interstitial sand habitat. Small-subunit rRNA gene sequence (GenBank accession no. FJ824125). *Protaspis oviformis* differs from the 11 previously described species of the genus in that this species possesses a very oval shape and smooth cell surface. Nuclear position of this species is located differently from its close relatives, *P. grandis* (i.e. posterior at the middle of the cell) and *P. rotunda* (i.e. anterior to the right of the cell). Unlike *P. obaniformis*, which possesses a posterior ventral slit, this structure of *P. oviformis* is located at the mid anterior end and towards the mid posterior end of the cell.

Type locality. Tidal sand-flat at Boundary Bay (49° 00′ N 123° 02′ W), Vancouver, British Columbia, Canada. The organisms were collected in March 2007.

Iconotype. Figs 1(c), 2(d–e) and 3(c).

Etymology. The etymology for the specific epithet, L. neut. n. *ovum* egg; L. fem. *forma* shape. The specific epithet reflects the shape of this organism.

Protaspis longipes (Schnepf & Kühn, 2000) Chantangsi and Leander comb. nov.

Basionym: *Cryothecomonas longipes* Schnepf and Kühn, 2000.

The Botuliformidae lineage

Phylogenies inferred from SSU rDNA sequences demonstrated that one new genus/species clustered strongly with the 'Ebriid' clade comprising an *Ebria tripartita* sequence and an environmental sequence (Figs 4–5).

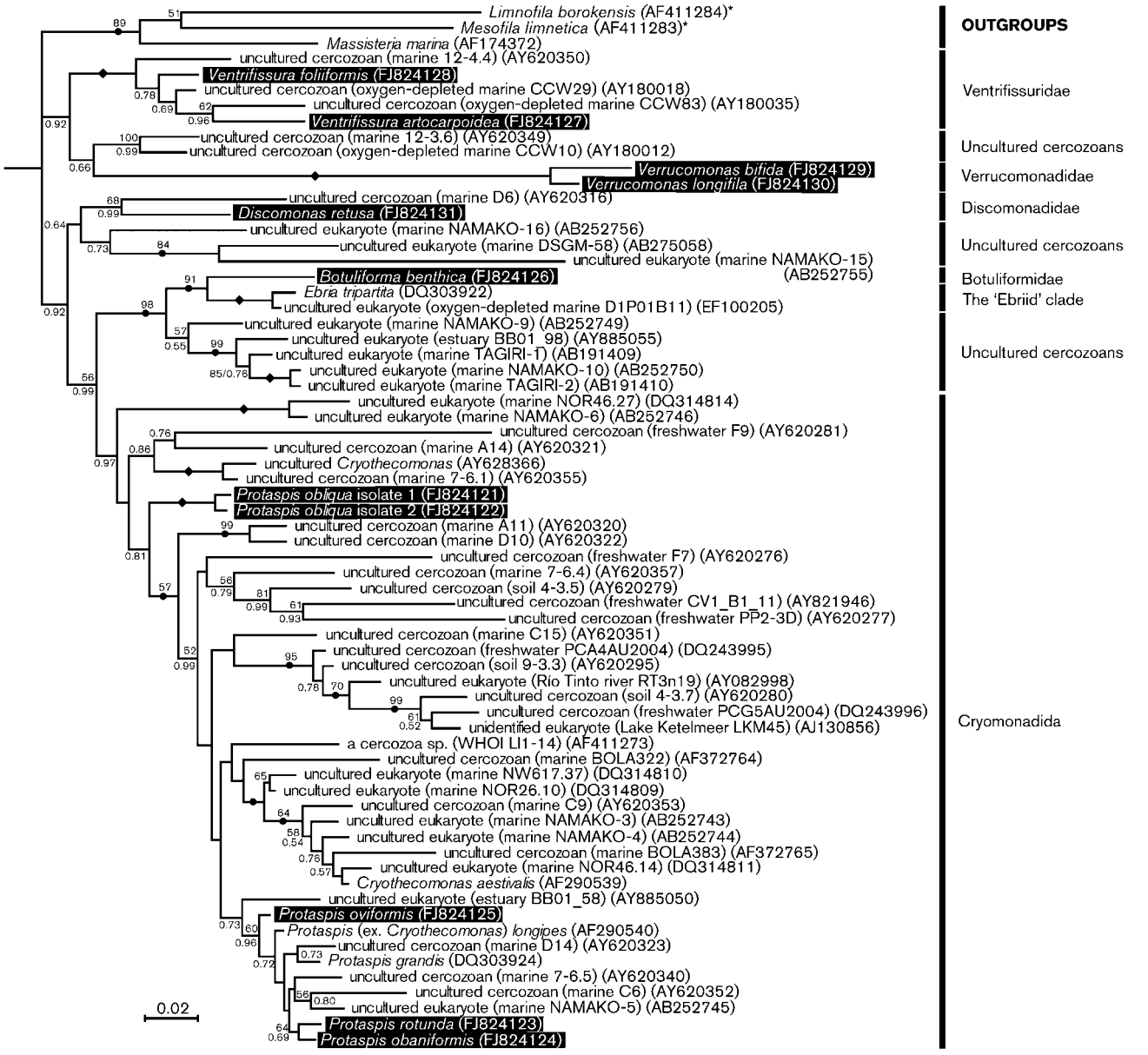

Fig. 5. Bayesian phylogeny deduced from 923 bp of SSU rDNA sequences of 67 cercozoan taxa including several sequences derived from environmental studies; the phylogenetic positions of marine benthic cercozoans examined in this study are highlighted in black boxes. The tree (mean ln $L=-7530.23$) is a consensus of 36 002 trees with the GTR+I+G using four rate categories implemented. Numbers >0.50 at the nodes indicate Bayesian posterior probabilities and PhyML bootstrap percentages >50 %. Black circles represent Bayesian posterior probability of 1.00. Black diamonds represent Bayesian posterior probability of 1.00 and PhyML bootstrap value of 100 %. Bar, 0.02 substitutions per site. **Limnofila borokensis* was previously misidentified as *Gymnophrys cometa* (AF411284) and *Mesofila limnetica* was previously referred to as *Dimorpha*-like sp. (AF411283).

Genus *Botuliforma* gen. nov. Chantangsi and Leander, 2009

Diagnosis. Cells are oblong; slightly thick wall with rough surface; uninucleate biflagellate; flagella inserted subapically; large nucleus with granular (permanently condensed) chromosomes located at anterior end of the cell; colourless cytoplasm; ventral furrow present; very fine filopodia observed; extrusomes present; locomotion by rotational swimming; found living in marine interstitial sand habitats.

Type species. *Botuliforma benthica.*

Etymology. The etymology for the generic name, L. masc. n. *botulus* sausage; L. fem. *forma* shape. The genus name reflects the sausage shape of the type species.

Botuliforma benthica sp. nov. Chantangsi and Leander, 2009

Size. Cell is about 20 μm wide and 35 μm long (number of cells observed=51).

Diagnosis. Structure as described for the genus; nucleus is 12 μm wide and 15 μm long; large nucleus is subspherical in outline and located at the anterior end of the cell; aggregation of cells forming a swimming spherical ball observed. Small-subunit rRNA gene sequence (GenBank accession no. FJ824126).

Type locality. Tidal sand-flat at Pachena Beach (48° 47′ N 125° 07′ W), Vancouver Island, British Columbia, Canada. The organisms were collected in June 2007 and June 2008.

Iconotype. Figs 1(d), 2(k) and 3(d).

Etymology. The etymology for the specific epithet, N.L. *benthica* of the benthos. The specific epithet reflects the natural habitat of this organism.

The Ventrifissuridae clade

Phylogenies inferred from SSU rDNA gene sequences demonstrated that one new genus and two new species clustered together with several environmental sequences and formed the Ventrifissuridae clade (Figs 4–5).

Genus *Ventrifissura* gen. nov. Chantangsi and Leander, 2009

Diagnosis. Cells are broadly obovate and dorsoventrally flattened; cells either with smooth surfaces or with numerous pointed warts; uninucleate biflagellate; flagella inserted subapically with or without an anterior protrusion; circular to oblong-shaped nucleus located at the anterior end of the cell; colourless cytoplasm; ventral furrow present; filopodia observed; extrusomes present; locomotion by gliding; found living in marine interstitial sand habitat.

Type species. *Ventrifissura artocarpoidea.*

Etymology. The etymology for the generic name, L. masc. *ventris* belly; L. fem. *fissura* crack, cleft, chink. The genus name reflects the morphological feature of members within this genus of having a slit on the ventral side.

Ventrifissura artocarpoidea sp. nov. Chantangsi and Leander, 2009

Size. Cell is about 35–36 μm wide and 43–45 μm long (number of cells observed=three).

Diagnosis. Cells are broadly obovate and dorsoventrally slightly flattened; numerous pointed warts distributed evenly over the cell surface; uninucleate biflagellate; flagella inserted subapically without a protrusion; circular nucleus located at the anterior end of the cell, sometimes toward the left side; colourless cytoplasm, sometimes with food particles; ventral furrow present; locomotion by gliding; found living in marine interstitial sand habitat. Small-subunit rRNA gene sequence (GenBank accession no. FJ824127).

Type locality. Tidal sand-flat at Boundary Bay (49° 00′ N 123° 02′ W), Vancouver, British Columbia, Canada. The organisms were collected in May 2007.

Iconotype. Figs 1(e), 2(g) and 3(e).

Etymology. The etymology for the specific epithet, N.L. masc. *artocarpus* a genus of breadfruit; L. fem. *–oidea* suffix denoting resembling. The specific epithet reflects the shape of this organism, which is similar to a breadfruit.

Ventrifissura foliiformis sp. nov. Chantangsi and Leander, 2009

Size. Cell is about 30–35 μm wide and 40–47 μm long (number of cells observed >100).

Diagnosis. Cells are broadly obovate and (extremely) dorsoventrally flattened with a smooth cell surface; uninucleate biflagellate; flagella inserted subapically between an anterior protrusion; nucleus is 15 μm wide and 10 μm long; oblong nucleus with granular appearance located at the anterior end of the cell, sometimes toward the right side; colourless cytoplasm, sometimes with food particles; small clear vacuoles present; ventral furrow present; filopodia observed; extrusomes present; locomotion by gliding; found living in marine interstitial sand habitat. Small-subunit rRNA gene sequence (GenBank accession no. FJ824128).

Type locality. Tidal sand-flat at Boundary Bay (49° 00′ N 123° 02′ W), Vancouver, British Columbia, Canada. The organisms were collected in May 2007.

Iconotype. Figs 1(f), 2(f), 2(h) and 3(f).

Etymology. The etymology for the specific epithet, L. neut. *folium* leaf; L. fem. *forma* shape. The specific epithet reflects the shape of this organism, which is flattened like a leaf.

The Verrucomonadidae clade

Phylogenies inferred from SSU rDNA sequences demonstrated that one new genus and two novel species clustered together and formed the Verrucomonadidae clade (Figs 4–5).

Genus *Verrucomonas* gen. nov. Chantangsi and Leander, 2009

Diagnosis. Cells are dorsoventrally flattened and with a rough surface; coloured warts ranging from yellowish to

red, brownish, and golden observed on cell surface; uninucleate biflagellate; flagella inserted subapically; nucleus is located at the anterior end of the cell; colourless cytoplasm; ventral furrow present; anterior and posterior notches sometimes present; filopodia observed; extrusomes present; locomotion by gliding; found living in marine interstitial sand habitat.

Type species. *Verrucomonas bifida.*

Etymology. The etymology for the generic name, L. fem. *verruca* wart; L. fem. *monas* a unit (which refers to the flagellate). The genus name reflects the morphological feature of having coloured warts on the cell surface.

Verrucomonas bifida sp. nov. Chantangsi and Leander, 2009

Size. Cell is about 17–25 μm wide and 30–34 μm long (number of cells observed >100).

Diagnosis. Cells are dorsoventrally flattened with a rough surface; coloured warts ranging from yellowish to red, brownish, and golden observed on cell surface; uninucleate biflagellate; flagella inserted subapically; bilobed nucleus, 20 μm wide and 13 μm long, located at the anterior of the cell; colourless cytoplasm; ventral furrow present; anterior and posterior notches sometimes present; fine filopodia observed; extrusomes present; locomotion by gliding; found living in marine interstitial sand habitat. Small-subunit rRNA gene sequence (GenBank accession no. FJ824129).

Type locality. Tidal sand-flat at Boundary Bay (49° 00′ N 123° 02′ W), Vancouver, British Columbia, Canada. The organisms were collected in May 2007.

Iconotype. Figs 1(g), 2(i) and 3(g).

Etymology. The etymology for the specific epithet, L. fem. *bifida* two-cleft. The specific epithet reflects the shape of the organism's nucleus, which seems to be divided into two parts.

Verrucomonas longifila sp. nov. Chantangsi and Leander, 2009

Size. Cell is about 15–20 μm wide and 23–37 μm long (number of cells observed=11).

Diagnosis. Cells are dorsoventrally flattened with a rough surface; yellowish and reddish warts observed on cell surface; uninucleate biflagellate; flagella inserted subapically; elliptical nucleus with granular appearance, 8 μm wide and 5 μm long, located at anterior end of the cell; colourless cytoplasm; ventral furrow present; anterior and posterior notches sometimes present; fine filopodia observed; extrusomes present; locomotion by gliding; found living in marine interstitial sand habitat. Small-subunit rRNA gene sequence (GenBank accession no. FJ824130).

Type locality. Tidal sand-flat at Spanish Banks (49° 16′ N 123° 14′ W), Vancouver, British Columbia, Canada. The organisms were collected in April 2007.

Iconotype. Figs 1(h), 2(j) and 3(h).

Etymology. The etymology for the specific epithet, L. masc. *longus* long; L. fem. *fila* thread. The specific epithet reflects the longer flagella of this organism in comparison with those of its close relative, *Verrucomonas bifida.*

The Discomonadidae clade

Phylogenies inferred from SSU rDNA sequences demonstrated that one novel genus/species clustered with an environmental sequence and formed the Discomonadidae clade (Figs 4–5).

Genus *Discomonas* gen. nov. Chantangsi and Leander, 2009

Diagnosis. Cells are disc-shaped and dorsoventrally flattened; uninucleate biflagellate; flagella inserted subapically; nucleus is located in the middle-anterior of the cell; anterior notch present; colourless cytoplasm, sometimes with food particles at the posterior end of the cell; locomotion by gliding; found living in marine interstitial sand habitat.

Type species. *Discomonas retusa.*

Etymology. The etymology for the generic name, L. masc. *discus* flat, circular plate; L. fem. *monas* a unit (which refers to the flagellate). The genus name reflects the shape of this organism.

Discomonas retusa sp. nov. Chantangsi and Leander, 2009

Size. Cell is about 25 μm wide and 25 μm long (number of cells observed=five).

Diagnosis. Structure as described for the genus; nucleus is 9 μm wide and 9 μm long; discoidal nucleus with central nucleolus located at the middle-anterior of the cell. Small-subunit rRNA gene sequence (GenBank accession no. FJ824131).

Type locality. Tidal sand-flat at Boundary Bay (49° 00′ N 123° 02′ W), Vancouver, British Columbia, Canada. The organisms were collected in May 2007.

Iconotype. Figs 1(i), 2(l) and 3(i).

Etymology. The etymology for the specific epithet, L. fem. *retusa* notched at the apex. The specific epithet reflects the notched appearance at the anterior end of the organism.

DNA barcoding marine benthic cercozoans

DNA analyses based on the K2P model of 1798 bp full-length and 618 bp barcoding regions of SSU rDNA sequences of 17 benthic cercozoans showed relatively similar mean values of sequence divergences, 7.57 % and 6.98 %, respectively. Analyses of the barcoding region of isolates of three species, including *Cryothecomonas aestivalis*, *Ebria tripartita* and *Protaspis obliqua*, showed low intraspecific sequence variation and only a few nucleotide differences (Table 2). In general, intergeneric sequence divergences between examined genera were quite high (Table 2). It is also significant to note that *Cryothecomonas longipes* showed lower sequence divergences and fewer nucleotide differences from *Protaspis grandis* (Fig. 2a) than from *Cryothecomonas aestivalis*, which is consistent with the molecular phylogenetic data (Figs 4–5). Moreover, two isolates of *Protaspis obliqua* (Fig. 2b) showed high sequence divergences and large numbers of nucleotide differences from the other species/morphotypes currently recognized as '*Protaspis*'.

Molecular phylogenetic analyses of marine benthic cercozoans

Phylogenies deduced from the 1617 bp full-length and 583 bp barcoding region of SSU rDNA sequences of 35 cercozoan taxa, including representatives from several major cercozoan subgroups, showed very similar tree topologies (Fig. 4). *Protaspis obliqua* and three new species of *Protaspis*, namely *P. rotunda*, *P. obaniformis* and *P. oviformis*, clustered within the Cryomonadida clade, which currently contains three genera – *Cryothecomonas*, *Lecythium* and *Protaspis* (Fig. 4). However, *P. obliqua* was positioned distantly from the other species of the genus *Protaspis*, and *Cryothecomonas longipes* was placed within the main *Protaspis* lineage with very strong statistical support (Figs 4–5). The nearest sister group to the order Cryomonadida was a clade consisting of *Botuliforma benthica* gen. et sp. nov., *Ebria tripartita* and several undescribed cercozoans derived from environmental DNA surveys (Fig. 5). The remaining five species described here – namely *Ventrifissura artocarpoidea* gen. et sp. nov., *Ventrifissura foliiformis* gen. et sp. nov., *Verrucomonas bifida* gen. et sp. nov., *Verrucomonas longifila* gen. et sp. nov., and *Discomonas retusa* gen. et sp. nov. – branched with different undescribed cercozoans derived from environmental DNA surveys, forming three different clades with morphological features that were unknown prior to this study.

DISCUSSION

Hidden diversity of marine benthic cercozoans

Marine planktonic and benthic habitats house a large number of microeukaryotic lineages (Massana & Pedrós-Alió, 2008; Park *et al.*, 2008). The actual biodiversity in benthic environments, however, is not well understood because studies on these habitats are relatively infrequent compared with studies on planktonic environments (Bass & Cavalier-Smith, 2004; Chantangsi *et al.*, 2008; Fenchel, 1987; Hondeveld *et al.*, 1992; Hoppenrath & Leander, 2006a; Lee, 2008). Nonetheless, several studies based on environmental PCR analyses of SSU rDNA have shown that cercozoans, such as members of the genera *Auranticordis*, *Cercomonas*, *Massisteria*, *Metopion*, *Metromonas*, *Protaspis* and *Thaumatomonas*, are major components of benthic habitats (Al Qassab *et al.*, 2002; Bass & Cavalier-Smith, 2004; Chantangsi *et al.*, 2008; Hoppenrath & Leander, 2006a; Myl'nikov & Karpov, 2004; Park *et al.*, 2008). The genus *Protaspis*, in particular, has been reported from several aquatic and terrestrial environments worldwide (Auer & Arndt, 2001; Ekelund & Patterson, 1997; Hoppenrath & Leander, 2006a; Larsen & Patterson, 1990; Lee *et al.*, 2003, 2005; Lee & Patterson, 2000; Vørs, 1993).

The current composition of the genus *Protaspis*

The genus *Protaspis* was originally described by Skuja (1939) and currently contains 11 recognized species: *P. gemmifera*, *P. glans*, *P. grandis*, *P. maior*, *P. metarhiza*, *P. obliqua*, *P. obovata*, *P. simplex*, *P. tanyopsis*, *P. tegere* and *P. verrucosa*. However, our study shows the actual diversity of the group to be far greater. We have assigned three novel species – *P. rotunda*, *P. obaniformis* and *P. oviformis* – to this genus. These novel flagellates share several common generic features with the type species: (i) a rigid cell body with two heterodynamic flagella (at least in the first two species), (ii) both flagella insert subapically on the ventral side of the cell (at least in the first two species), and (iii) cells are dorsoventrally flattened and possess a ventral slit from which pseudopodia can be protruded (Hoppenrath & Leander, 2006a; Myl'nikov & Karpov, 2004; Skuja, 1939). However, the three novel members of the genus *Protaspis* described here did not possess the specific cell shapes and diagnostic features found in the previously recognized species.

Our molecular phylogenetic analyses demonstrated that the three novel species of the genus *Protaspis* grouped strongly with *P. grandis*. However, two isolates of *Protaspis obliqua*, whose SSU rDNA sequences were sequenced in this study, were positioned very distantly from the other members of the genus *Protaspis* in our molecular phylogenetic analyses; this result was consistent with the large number of nucleotide differences between *P. obliqua* and the other *Protaspis* sequences and led us to doubt the taxonomic status of *P. obliqua*. These data also suggest that a great deal of genetic diversity is hidden at the morphological level (Weisse, 2008). However, further comparative investigations of ultrastructural features in members of the genus *Protaspis* (especially the type species *P. glans*) might demonstrate differences that are consistent with the molecular phylogenetic data and lead to the future taxonomic reassignment of *P. obliqua*.

Cryothecomonas longipes is more closely related to the genus *Protaspis sensu stricto*

The genus *Cryothecomonas* was shown to be the closest relative of the genus *Protaspis* based on molecular

phylogenetic evidence and some ultrastructural features (Hoppenrath & Leander, 2006a; Thomsen *et al.*, 1991). Six species of the genus *Cryothecomonas*, namely *C. aestivalis*, *C. armigera*, *C. inermis*, *C. longipes*, *C. scybalophora* and *C. vesiculata*, have been described thus far and only two of them, *C. aestivalis* and *C. longipes*, have had their SSU rDNA sequenced (Drebes *et al.*, 1996; Kühn *et al.*, 2000; Schnepf & Kühn, 2000; Thomsen *et al.*, 1991). Our phylogenetic analyses demonstrated a closer relationship between *C. longipes* and the genus *Protaspis* (excluding *P. obliqua*) than to *C. aestivalis*, which was consistent with previous results (Hoppenrath & Leander, 2006a).

Thomsen *et al.* (1991) established *Cryothecomonas* as a new genus on the basis of differences between these species and the previously established genus *Protaspis*, such as the configuration of the flagellar apparatus and the location of the cell slit/groove (Thomsen *et al.*, 1991). Flagella are homodynamic and inserted apically in members of the genus *Cryothecomonas,* whereas members of the genus *Protaspis* have heterodynamic flagella inserted subapically (Skuja, 1939; Thomsen *et al.*, 1991). The cell slit/groove, where pseudopodia emerge, is located posterior-laterally in the genus *Cryothecomonas* and ventral-medially in the genus *Protaspis* (Skuja, 1939; Thomsen *et al.*, 1991). Ultrastructural studies on the latest member of the genus *Cryothecomonas*, namely *C. longipes*, demonstrated morphological differences between this species and the other four species in the genus *Cryothecomonas*, which were all described together when this genus was first established (Thomsen *et al.*, 1991). For instance, *C. longipes* possesses heterodynamic flagella; the anterior flagellum is inserted apically and the posterior flagellum is inserted subapically. Moreover, the ventral slit in *C. longipes* and members of the genus *Protaspis* is located on the ventral side of the cell. If flagellar orientation and the location of longitudinal groove were the key features that prompted Thomsen *et al.* (1991) to separate the genus *Cryothecomonas* from *Protaspis*, then *C. longipes* should not belong to the former genus. Several ultrastructural features shared by *C. longipes* and *P. grandis* have also been demonstrated, such as the presence of a multilayered cell wall, a nucleus with a prominent nucleolus and condensed chromosomes, structurally identical extrusomes and flagellar pits with distinctive funnels (Hoppenrath & Leander, 2006a; Schnepf & Kühn, 2000). All of these data are consistent with our molecular phylogenetic analyses that showed *C. longipes* within the main *Protaspis* clade. For these reasons, we have transferred *C. longipes* to the genus *Protaspis*: *P. longipes* comb. nov. (see Results section).

The benthic *Botuliforma benthica* gen. et sp. nov. is closely related to the planktonic *Ebria tripartita*

We have discovered a benthic swimming flagellate, namely *Botuliforma benthica* gen. et sp. nov., that is the nearest sister lineage to the clade consisting of *Ebria tripartita* and environmental sequence GenBank accession no. EF100205 as inferred from SSU rDNA sequences (Fig. 5). Interestingly, the environmental sequence was generated from the upper 2 cm of oxygen-depleted intertidal marine sediments, which is nearly identical to the habitat where we collected *B. benthica* gen. et sp. nov. The distinctively different DNA sequences, habitats and morphological features of *Botuliforma* and *Ebria* led us to establish a new genus for this novel benthic organism. The 'Ebriida' clade was previously shown to be a member of the Cercozoa and a close sister lineage to the Cryomonadida in SSU rDNA phylogenies (Hoppenrath & Leander, 2006b). Although morphological similarities between representatives of these two clades were not initially obvious, shared features have been demonstrated at the ultrastructural level (Hoppenrath & Leander, 2006b). Members of the genera *Ebria*, *Cryothecomonas* and *Protaspis* all share two unequal flagella, a nucleus with a prominent nucleolus and permanently condensed chromosomes, tubular mitochondrial cristae (at least in the last two genera) and feeding by means of pseudopodia (Hargraves, 2002; Hoppenrath & Leander, 2006a, b; Thomsen *et al.*, 1991). All current members of the Cryomonadida (genera *Cryothecomonas*, *Lecythium* and *Protaspis*) possess a test, layered wall or theca around the cells; in contrast, members of the genus *Ebria* possess a naked cell with a fine layer of fibrillar material lying outside the plasma membrane and an internal siliceous skeleton (Hargraves, 2002).

The combination of features in *B. benthica* gen. et sp. nov. appears to be transitional between ebriids and cryomonads. For example, *B. benthica* has a thick and rough wall that might be homologous to the multilayered cell walls of members of the genera *Cryothecomonas* and *Protaspis*. In addition, a prominent nucleus with condensed chromosomes, fine pseudopodia and extrusomes were easily observed in *B. benthica* under LM (Figs 1d, 2k), and these features are also found in members of the genera *Cryothecomonas* and *Protaspis* (Drebes *et al.*, 1996; Hoppenrath & Leander, 2006a; Schnepf & Kühn, 2000; Thomsen *et al.*, 1991). Although ultrastructural data are currently unavailable for *B. benthica* gen. et sp. nov., the close relationship between this lineage and *E. tripartita* is very robust in molecular phylogenies inferred from SSU rDNA. Therefore, transmission electron microscopy of *B. benthica* gen. et sp. nov. might demonstrate the existence of inconspicuous siliceous skeletal elements that are homologous with ebriids.

The cellular identities of previously undescribed cercozoans

We established three new genera in addition to *Botuliforma* gen. nov., namely *Discomonas* gen. nov., *Ventrifissura* gen. nov. and *Verrucomonas* gen. nov., on the basis of their very distant molecular phylogenetic positions inferred from SSU rDNA sequences (Figs 4–5). Our phylogenetic analyses placed two novel species of the genus *Ventrifissura*, *V. artocarpoidea* and *V. foliiformis*, within a very distinct clade consisting of three environmental sequences derived from

marine environments: GenBank accession nos AY180018, AY180035 and AY620350 (Bass & Cavalier-Smith, 2004; Stoeck & Epstein, 2003). The two novel species of the genus *Verrucomonas* were separated from the other cercozoans by very long branch-lengths (Fig. 5). This Verrucomonadidae clade clustered weakly with two environmental sequences also obtained from marine environments, GenBank accession nos AY180012 and AY620349 (Bass & Cavalier-Smith, 2004; Stoeck & Epstein, 2003), and this more inclusive clade formed the nearest sister group to the Ventrifissuridae clade (Fig. 5). Analyses by Bass & Cavalier-Smith (2004) showed that several of these environmental sequences (e.g. uncultured cercozoan GenBank accession nos AY620349 and AY620350) are members of the Tectofilosida, and the morphological features we describe here help establish the cellular identities for some of the environmental sequence clades contained therein.

The SSU rDNA sequence of *Discomonas* gen. nov. formed the nearest sister lineage to uncultured cercozoan GenBank accession no. AY620316 (Fig. 5). Bass & Cavalier-Smith (2004) recognized this environmental sequence as an undescribed member of 'Basal Group T', which branched closely with thaumatomonads. Our analyses of two different datasets, however, showed incongruent phylogenetic positions for *Discomonas* gen. nov. Analyses of the 923 bp dataset showed *Discomonas* gen. nov. branching with environmental sequence AY620316, and analyses of the 1617 bp dataset showed *Discomonas* gen. nov. branching next to the Cercomonadida lineage (e.g. Cercomonas plasmodialis; Fig. 4). This incongruence was also observed by Bass & Cavalier-Smith (2004); they reported that AY620316 can sometimes be a sister lineage to the Metopiida and a few other sequences recognized by them as 'Novel Clade 6'. However, these relationships were recovered with only weak statistical support. Although these organisms need to be revisited in the future using alternative molecular markers, our study provides morphological information for *Discomonas* gen. nov. that sheds light on the probable cellular identity of AY620316 and 'Basal Group T' .

Our study has shown morphological evidence for at least two previously uncharacterized clades belonging to the Tectofilosida and helps substantiate the establishment of the Thecofilosea, which currently contains Tectofilosida and Cryomonadida (Bass & Cavalier-Smith, 2004). However, the original diagnosis for Tectofilosida by Cavalier-Smith & Chao (2003) [i.e. uninucleate cell surrounded by an organic flexible tectum or rigid test with one or two apertures for filopodia, sometimes including foreign mineral particles (agglutinated); cilia or silica scales absent; tubular mitochondrial cristae] should be amended by adding 'flagella present' as both the genera *Ventrifissura* and *Verrucomonas* possess flagella and are now shown to be members of the Tectofilosida.

Barcoding marine benthic cercozoans

DNA sequences can be effectively used as a diagnostic tool for the identification of species, an approach known as 'DNA barcoding' (Hebert *et al.*, 2003a, b). This general approach is starting to be used extensively for assessing the diversity of microeukaryotes having limited morphological details for species discrimination (Barth *et al.*, 2006; Chantangsi *et al.*, 2007; Lynn & Strüder-Kypke, 2006; Saunders, 2005; Scicluna *et al.*, 2006). Accordingly, one primary aim of our study was to help demonstrate the potential of the 618 bp SSU rDNA sequence as DNA barcode for identifying marine benthic cercozoans and possibly other cercozoans. Bass & Cavalier-Smith (2004) previously designed phylum-specific primers covering the 618 bp barcoding region for SSU rDNA, which facilitated PCR amplification of the gene sequences in the group of cercozoans we investigated here.

In our analyses, different isolates of *E. tripartita* and *P. obliqua* showed no and one nucleotide differences in their 618 bp barcoding regions, respectively. This result helped demonstrate the reliability of the DNA barcoding region for species identification. In contrast, two different isolates of *Cryothecomonas aestivalis* showed 0.5 % sequence divergence (three nucleotide differences), which was equal to the divergence value between *P. grandis* and *P. longipes* (ex. *Cryothecomonas longipes*) (Table 2). Therefore, if *P. grandis* and *P. longipes* constitute separate species as clearly shown by morphological and molecular evidence, then the two isolates of *C. aestivalis* could also be justifiably established as two (cryptic) species on the basis of this molecular marker. It turns out that the possibility of cryptic species within *C. aestivalis* has already been mentioned by Kühn *et al.* (2000) when they found a relatively high number of nucleotide differences between two morphologically indistinguishable strains of this 'species'. These results, therefore, also demonstrate the utility of the DNA barcoding region for discovering and delimiting cryptic species. Moreover, because the 618 bp barcoding region gene produced very similar phylogenetic tree topologies to those derived from the full length of the gene (Fig. 4), this short barcoding region can also help systematists infer the broader genealogical relationships of cercozoans.

The high copy number of the SSU rRNA gene in eukaryotic nuclear genomes facilitates the barcoding of isolated microeukaryotes even when very limited numbers of organisms can be obtained. Our study has shown that this approach is very useful for exploring the biodiversity of uncultured microeukaryotes; most of the SSU rDNA sequences reported here were derived from a small number of uncultured cells per DNA extraction, and in several cases, we acquired SSU rDNA sequences from only one uncultured cell, which was manually isolated from the ocean. Our study not only highlights the cryptic diversity of marine benthic cercozoans, which is extremely problematic for the non-specialist, but also demonstrates an alternative approach for species identification using the DNA barcoding principle. Moreover, we have provided morphological data for several cercozoan subgroups that were previously known only from environmental DNA surveys. Overall, these data provide insights into the cellular identities of

uncultured and undescribed cercozoans that help advance our understanding of cercozoan biodiversity in marine benthic ecosystems.

ACKNOWLEDGEMENTS

We thank K. Tangthongchaiwiriya for illustrating all nine novel cercozoans (Fig. 3). C. C. was supported by a national scholarship awarded by the Cooperative Research Network (CRN), Government of Thailand. This work was supported by grants to B. S. L. from the National Science and Engineering Research Council of Canada (NSERC 283091-04) and the Canadian Institute for Advanced Research, Program in Integrated Microbial Biodiversity.

REFERENCES

Al Qassab, S., Lee, W. J., Murray, S. & Patterson, D. J. (2002). Flagellates from stromatolites and surrounding sediments in Shark Bay, Western Australia. *Acta Protozool* **41**, 91–144.

Auer, B. & Arndt, H. (2001). Taxonomic composition and biomass of heterotrophic flagellates in relation to lake trophy and season. *Freshw Biol* **46**, 959–972.

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

Bass, D. & Cavalier-Smith, T. (2004). Phylum-specific environmental DNA analysis reveals remarkably high global biodiversity of Cercozoa (Protozoa). *Int J Syst Evol Microbiol* **54**, 2393–2404.

Berney, C., Fahrni, J. & Pawlowski, J. (2004). How many novel eukaryotic 'kingdoms'? Pitfalls and limitations of environmental DNA surveys. *BMC Biol* **2**, 13.

Cavalier-Smith, T. (1998a). A revised six-kingdom system of life. *Biol Rev Camb Philos Soc* **73**, 203–266.

Cavalier-Smith, T. (1998b). *Neomonada and the origin of animals and fungi.* In *Evolutionary Relationships Among Protozoa*, pp. 375–407. Edited by G. H. Coombs, K. Vickerman, M. A. Sleigh & A. Warren. London: Kluwer Academic Publishers.

Cavalier-Smith, T. & Chao, E. E. (2003). Phylogeny and classification of Phylum Cercozoa (Protozoa). *Protist* **154**, 341–358.

Chantangsi, C., Lynn, D. H., Brandl, M. T., Cole, J. C., Hetrick, N. & Ikonomi, P. (2007). Barcoding ciliates: a comprehensive study of 75 isolates of genus *Tetrahymena*. *Int J Syst Evol Microbiol* **57**, 2412–2425.

Chantangsi, C., Esson, H. J. & Leander, B. S. (2008). Morphology and molecular phylogeny of a marine interstitial tetraflagellate with putative endosymbionts: *Auranticordis quadriverberis* n. gen. et sp. (Cercozoa). *BMC Microbiol* **8**, 123.

Cummings, D. J. (1992). Mitochondrial genomes of the ciliates. *Int Rev Cytol* **141**, 1–64.

Drebes, G., Kühn, S. F., Gmelch, A. & Schnepf, E. (1996). *Cryothecomonas aestivalis* sp. nov., a colourless nanoflagellate feeding on the marine centric diatom *Guinardia delicatula* (Cleve) Hasle. *Helgol Meeresunters* **50**, 497–515.

Ekelund, F. & Patterson, D. J. (1997). Some heterotrophic flagellates from a cultivated garden soil in Australia. *Arch Protistenkd* **148**, 461–478.

Fenchel, T. (1987). *Ecology of Protozoa: the Biology of Free-Living Phagotrophic Protists.* Berlin: Springer-Verlag.

Godfray, H. C. J. (2002). Challenges for taxonomy. *Nature* **417**, 17–19.

Guindon, S. & Gascuel, O. (2003). PhyML - A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* **52**, 696–704.

Hargraves, P. E. (2002). The ebridian flagellates *Ebria* and *Hermesinum*. *Plankton Biol Ecol* **49**, 9–16.

Hebert, P. D. N., Ratnasingham, S. & deWaard, J. R. (2003a). Barcoding animal life: cytochrome *c* oxidase subunit 1 divergences among closely related species. *Proc Biol Sci* **270** (Suppl. 1), S96–S99.

Hebert, P. D. N., Cywinska, A., Ball, S. L. & deWaard, J. R. (2003b). Biological identifications through DNA barcodes. *Proc Biol Sci* **270**, 313–322.

Hondeveld, B. J. M., Bak, R. P. M. & van Duyl, F. C. (1992). Bacterivory by heterotrophic nanoflagellates in marine sediments measured by uptake of fluorescently labeled bacteria. *Mar Ecol Prog Ser* **89**, 63–71.

Hoppenrath, M. & Leander, B. S. (2006a). Dinoflagellate, euglenid or cercomonad? The ultrastructure and molecular phylogenetic position of *Protaspis grandis* n. sp. *J Eukaryot Microbiol* **53**, 327–342.

Hoppenrath, M. & Leander, B. S. (2006b). Ebriid phylogeny and the expansion of the Cercozoa. *Protist* **157**, 279–290.

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

Kimura, M. (1980). A simple method of estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol* **16**, 111–120.

Kühn, S. F., Lange, M. & Medlin, L. K. (2000). Phylogenetic position of *Cryothecomonas* inferred from nuclear-encoded small subunit ribosomal RNA. *Protist* **151**, 337–345.

Larsen, J. & Patterson, D. J. (1990). Some flagellates (Protista) from tropical marine sediments. *J Nat Hist* **24**, 801–937.

Lee, W. J. (2008). Free-living heterotrophic euglenids from marine sediments of the Gippsland Basin, southeastern Australia. *Mar Biol Res* **4**, 333–349.

Lee, W. J. & Patterson, D. J. (2000). Heterotrophic flagellates (Protista) from marine sediments of Botany Bay, Australia. *J Nat Hist* **34**, 483–562.

Lee, W. J., Brandt, S. M., Vørs, N. & Patterson, D. J. (2003). Darwin's heterotrophic flagellates. *Ophelia* **57**, 63–98.

Lee, W. J., Simpson, A. G. B. & Patterson, D. J. (2005). Free-living heterotrophic flagellates from freshwater sites in Tasmania (Australia), a field survey. *Acta Protozool* **44**, 321–350.

Long, E. O. & David, I. B. (1980). Repeated genes in eukaryotes. *Annu Rev Biochem* **49**, 727–764.

Lynn, D. H. & Strüder-Kypke, M. C. (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.

Marande, W. & Burger, G. (2007). Mitochondrial DNA as a genomic jigsaw puzzle. *Science* **318**, 415.

Massana, R. & Pedrós-Alió, C. (2008). Unveiling new microbial eukaryotes in the surface ocean. *Curr Opin Microbiol* **11**, 213–218.

Minelli, A. (1993). *Biological Systematics: The State of the Art.* London: Chapman & Hall.

Myl'nikov, A. P. & Karpov, S. A. (2004). Review of diversity and taxonomy of cercomonads. *Protistology* **3**, 201–217.

Norman, J. E. & Gray, M. W. (1997). The cytochrome oxidase subunit 1 gene (*cox1*) from the dinoflagellate, *Crypthecodinium cohnii*. *FEBS Lett* **413**, 333–338.

Park, S. J., Park, B. J., Pham, V. H., Yoon, D. N., Kim, S. K. & Rhee, S. K. (2008). Microeukaryotic diversity in marine environments, an analysis of surface layer sediments from the East Sea. *J Microbiol* **46**, 244–249.

Ronquist, F. & Huelsenbeck, J. P. (2003). MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**, 1572–1574.

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

Schnepf, E. & Kühn, S. F. (2000). Food uptake and fine structure of *Cryothecomonas longipes* sp. nov., a marine nanoflagellate *incertae sedis* feeding phagotrophically on large diatoms. *Helgol Mar Res* **54**, 18–32.

Scicluna, S. M., Tawari, B. & Clark, C. G. (2006). DNA barcoding of *Blastocystis*. *Protist* **157**, 77–85.

Skuja, H. (1939). Beitrag zur Algenflora Lettlands II. *Acta Horti Bot Univ Latviensis* **11/12**, 41–169.

Šlapeta, J., Moreira, D. & López-García, P. (2005). The extent of protist diversity: insights from molecular ecology of freshwater eukaryotes. *Proc Biol Sci* **272**, 2073–2081.

Sogin, M. L., Swanton, M. T., Gunderson, J. H. & Elwood, H. J. (1986). Sequence of the small subunit ribosomal RNA gene from the hypotrichous ciliate *Euplotes aediculatus*. *J Protozool* **33**, 26–29.

Stoeck, T. & Epstein, S. (2003). Novel eukaryotic lineages inferred from small-subunit rRNA analyses of oxygen-depleted marine environments. *Appl Environ Microbiol* **69**, 2657–2663.

Tamura, K., Dudley, J., Nei, M. & Kumar, S. (2007). MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Mol Biol Evol* **24**, 1596–1599.

Tautz, D., Arctander, P., Minelli, A., Thomas, R. H. & Vogler, A. P. (2002). DNA points the way ahead in taxonomy. *Nature* **418**, 479.

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

Thompson, J. D., Higgins, D. G. & Gibson, T. J. (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.

Thomsen, H. A., Buck, K. R., Bolt, P. A. & Garrison, D. L. (1991). Fine structure and biology of *Cryothecomonas* gen. nov. (Protista *incertae sedis*) from the ice biota. *Can J Zool* **69**, 1048–1070.

Uhlig, G. (1964). Eine einfache methode zur extraktion der vagilen, mesopsammalen Mikrofauna. *Helgol Wiss Meeresunters* **11**, 178–185.

Vørs, N. (1993). Heterotrophic amoebae, flagellates and heliozoa, from Arctic marine waters (North West Territories, Canada and West Greenland). *Polar Biol* **13**, 113–126.

Weisse, T. (2008). Distribution and diversity of aquatic protists: an evolutionary and ecological perspective. *Biodivers Conserv* **17**, 243–259.