

PROF. DAVID BASS (Orcid ID : 0000-0002-9883-7823)

MR. JAN JANOUŠKOVEC (Orcid ID : 0000-0001-6547-749X)

DR. CÉDRIC BERNEY (Orcid ID : 0000-0001-8689-9907)

Article type : Original Article

Bass et al.---*Aquavolon* are Freshwater and Soil Rhizarian Protists

Rhizarian ‘Novel Clade 10’ Revealed as Abundant and Diverse Planktonic and Terrestrial Flagellates, including *Aquavolon* n. gen.

David Bass^{a,b}, Tikhonenkov DV^{c,d}, Foster R^a, Dyal P^a, Janouskovec J^{d,e}, Keeling PJ^d, Gardner M^a, Neuhauser S^f, Hartikainen H^a, Mylnikov AP^c, Berney C^{a,1}

^a Department of Life Sciences, The Natural History Museum, Cromwell Road, London SW7 5BD, UK

^b Centre for Environment. Fisheries and Aquaculture Science (Cefas), Barrack Road, The Nothe, Weymouth DT4 8UB, UK

^c Institute for Biology of Inland Waters, Russian Academy of Sciences, Borok 152742, Russia

^d University of British Columbia, Botany Department, Vancouver V6T1Z4, BC, Canada

^e University College London, Department of Genetics, Evolution and Environment, Gower Street, London, WC1E 6BT, UK

^f Institute of Microbiology, University of Innsbruck, Technikerstraße 25, 6020 Innsbruck, Austria

¹ present address: Sorbonne Université & CNRS, UMR AD2M, Station Biologique de Roscoff, Place Georges Teissier, 29680 Roscoff, France

Correspondence: David Bass, Department of Life Sciences, The Natural History Museum, Cromwell Road, London SW7 5BD, UK

Telephone number: +44 207 942 5387; e-mail d.bass@nhm.ac.uk

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi:

10.1111/jeu.12524

This article is protected by copyright. All rights reserved.

ABSTRACT

Rhizarian 'Novel Clade 10' (NC10) is frequently detected by 18S rRNA gene sequencing studies in freshwater planktonic samples. We describe a new genus and two species of eukaryovorous biflagellate protists, *Aquavolon hoantrani* n. gen. n. sp., and *A. dientrani* n. gen. n. sp., which represent the first morphologically characterized members of NC10, here named Aquavolonida ord. nov.. The slightly metabolic cells possess naked heterodynamic flagella, whose kinetosomes lie at a right angle to each other and are connected by at least one fibril. Unlike their closest known relative *Tremula longifila*, they rotate around their longitudinal axis when swimming and only very rarely glide on surfaces. Screening of a wide range of environmental DNA extractions with lineage-specific PCR primers reveals that Aquavolonida consists of a diverse radiation of protists, which are most diversified in freshwater planktonic habitats and as yet undetected in marine environments. Earlier branching lineages in Aquavolonida include less frequently detected organisms from soils and freshwater sediments. The 18S rRNA gene phylogeny suggests that Aquavolonida forms a common evolutionary lineage with tremulids and uncharacterized 'Novel Clade 12', which likely represents one of the deepest lineages in the Rhizaria, separate from Cercozoa (Filosa), Endomyxa and Retaria.

Key words

Aquavolon hoantrani, *Aquavolon dientrani*, Cercozoa, *Tremula*, 18S rRNA, environmental sequencing, Aquavolonida, Rhizaria

SEQUENCING of the 18S rRNA gene from environmental samples has revealed a large diversity of undescribed protists in the vast majority of habitat types sampled (Richards & Bass 2005). Some of these are regularly detected and acquire their own quasi-taxonomic status, for example the various MAST (marine stramenopiles; Massana et al. 2004) and MALV (marine alveolates; López García et al. 2001) lineages, and the LKM11 group related to fungi (Lara et al. 2010; Jones et al. 2011). Likewise, the first environmental sequencing survey of the phylum Cercozoa (Bass & Cavalier-Smith 2004) revealed novel lineages that were too distantly related to known taxa to infer anything about their biology; they have since been referred to as cercozoan Novel Clades (NCs) and numbered 1 to 9.

Since 2004 several of these initial NCs have been identified by cell isolation or culturing efforts: the order Marimonadida (Howe et al. 2011) accounts for most of NC1, including genera *Pseudopirsonia* (Kühn et al. 2004), *Auranticordis* (Chantangsi et al. 2008), and *Abollifer* (Shiratori et al. 2014); the first member of NC4 was shown to be *Trachyrhizium* (Shiratori et al. 2016), which forms a clade with *Lecythium* and *Diaphoropodon* (Dumack et al. 2017a); NC6 belongs to a group including *Metopion*, *Metromonas*, and *Micrometromonas* (Howe et al. 2011; this study); NC7 includes Pansomonadida (*Aurigamonas* and *Agitata*; Bass et al. 2009b); and NC8 is part of Vampyrellida (Bass et al. 2009a; Hess et al. 2012; Berney et al. 2013). NCs 2, 3 and 5 remain

uncharacterized, although NC2 groups phylogenetically with *Clautriavia* (Chantangsi & Leander 2010a), and Bass et al. (2009a) showed that NC5 is a subclade of Granofilosea, a class of filose amoebae. More data availability and improved phylogenetic methods have also shown that 'Basal Group T' in Bass & Cavalier-Smith (2004) is more morphologically and phylogenetically diverse than that label suggests, including the flagellate *Discomonas* (Chantangsi & Leander 2010b) and the testate amoeba *Penardeugenia* (Dumack et al. 2017b).

Further NCs were defined in Bass et al. (2009a), the most relevant here being NCs 10, 11, and 12. These lineages sometimes group together with moderate support. Their exact phylogenetic position within Rhizaria remains unresolved, but they typically appear distinct from both core Cercozoa (previously Filosa) and Endomyxa (Bass et al. 2009a; Howe et al. 2011). On the basis of recent phylogenomic analyses placing the root of the rhizarian radiation between Cercozoa (Filosa) and Endomyxa+Retaria (Sierra et al. 2016; Krabberød et al. 2017), this group may therefore form one of the deepest branching rhizarian lineages. NC11 was shown by Howe et al. (2011) to contain *Tremula longifila*, a biflagellate gliding cell (ATCC 50530), which along with its direct relatives inhabits freshwater sediments and soil. Cellular information for NC12 is still lacking, despite its frequency and diversity in a wide range of environments. NC12 was originally defined on the basis of sequences analysed in Bass et al. 2009a. Since then a larger diversity of related environmental sequences has been revealed, analysis of which sometimes suggests the original NC12 may correspond to two separate, strongly supported clades plus some unaffiliated, single-sequence lineages.

Likewise, NC10 has so far remained uncharacterized. It is however remarkable for being very frequently and almost exclusively detected in freshwater planktonic samples, from Antarctic to subtropical lakes. Supplementary Table S1 provides a list of all NC10 environmental clones we identified in GenBank at the onset of this work, with information about their provenance. Key examples of these are: LG21-01 and LG01-12 in Richards et al. (2005), A50 in Lefranc et al. (2005) (not A51, which is a vampyrellid; Berney et al. 2013), 'Novel clade I' in Lefèvre et al. (2007, 2008) (except PCC4AU2004, which belongs to NC12), all of the cercozoan sequences in Fig. 4 of Lepère et al. (2008) (except *Cercomonas* and PCC4AU2004), the 'Cercozoa freshwater clade' in Taib et al. (2013) (except again PCC4AU2004), the OTUs identified as belonging to NC10 in Simon et al. (2015) and Yi et al. (2017), and in Lara et al (2015), in which NC10 lineages are shown to be well but variably represented in peat bogs in Argentinian Tierra del Fuego.

In this study we investigated the environmental diversity and ecology of NC10 by using clade-specific PCR primers to screen freshwater, soil, and marine environmental DNA samples, as well as fluorescent in situ hybridization (FISH). Further, phylogenetic analysis of 18S rRNA gene sequences of flagellates isolated from a reservoir and wetland in Vietnam showed them to belong to NC10. These were described using light and transmission electron microscopy (TEM) as two new species within the new genus *Aquavolon*. In the light of 18S rRNA gene sequence signatures shared by all NC10 lineages uncovered in this study, morphological observations on our isolates, and the

habitat and likely phenotype of other NC10 members, we formally rename the group Aquavolonida, ord. nov.

Materials and Methods

Environmental DNA screening and NC10 clone library sequencing

DNA samples from filtered lake water from the English Lake District (<2 µm and 2-20 µm size fractions), river biofilms, and soils that were used for the study of Grossman et al. (2016) were also used for the present study. DNA was extracted from filtered water using the Qiagen Blood & Tissue DNA Extraction Kit (Qiagen, Hilden, Germany), from biofilms newly formed on ceramic tiles placed in the River Lambourn, UK by a standard CTAB/phenol-chloroform extraction method, and from soils using the FastDNA SPIN Kit for Soil (MP Biomedicals LLC, UK). Marine environmental DNA extractions previously used in other studies (Bass & Cavalier-Smith 2004; Bass et al. 2009a, b; Berney et al. 2013; Hartikainen et al. 2014; Logares et al. 2014) were also used for screening.

Two PCR strategies with distinct levels of specificity to NC10 were used for environmental DNA screening, both based on a two-step, nested PCR approach using distinct sets of PCR primers (see Results for the rationale and specificity of these PCR strategies). Strategy I: first PCR with primers C7f-NC10 and sB2n, nested PCR with primers V8f-NC10 and EndoR1. Strategy II: first PCR with primers V2f-NC and C9r-NC, nested PCR with primers C5f-NC and V8r-NC. Construction of Endomyxa-enriched clone libraries was achieved using a two-step, semi-nested PCR approach: first PCR with primers s1259F and sB2n, semi-nested PCR with primers s1259F and EndoR1. PCR primers corresponding to the NC10-specific FISH probes we designed were used to verify their strict specificity to a subset of NC10 using the following semi-nested PCR approach: first PCR with primers V4f-NC10 and C7r-NC10, semi-nested PCR with primers V4f-NC10 and C6r-NC10. Supplementary Table S2 provides the sequence and position of these primers.

PCR amplifications were done in a total volume of 25 µl with an amplification profile typically consisting of 2 min. at 95°C, followed by 30 cycles (first PCR) or 25 cycles (nested PCR) of 30 sec. at 95°C, 30 sec. at 65°C, and 2 min. at 72°C, followed by 5 min. at 72°C for the final extension. Purified amplicons of the appropriate length were cloned into TOPO-TA vectors and screened using primers M13for (5'-CGT TGT AAA ACG ACG GCC AGT-3') and M13rev (5'-CAC AGG AAA CAG CTA TGA CCA-3'). Positive PCR products were cleaned using a polyethylene glycol (PEG) protocol and sequenced using the same primers on an ABI-377 sequencer. Sequences were edited in FinchTV (<http://www.geospiza.com/finchtv/>) and aligned in BioEdit (Hall 1999).

Isolation of protist strains

Colp-16, Kat-1, and Colp-11 were isolated as single cells by a glass micropipette. Cells were picked from Petri dishes and washed in a series of 20µm droplets of Pratt's

medium. Clone Colp-16 was isolated from a near-shore water sample with bottom detritus from Suoi Da Reservoir (Hồ Suối Đá, viet.), Bình Thuận Province, S.R. Vietnam (11.122526N, 108.183158E) collected on May 10th 2013 (Water Temp. 33.08°C, pH 8.42, conductivity 96 μ S/cm, TDS 48 ppm, DO 6.30 ppm). Clone Kat-1 was isolated from a water sample containing decaying plant material from a wetland near farmland (11.509471N, 107.365586E), Đắc Lua Village, Dong Nai Province, S.R. Vietnam collected on May 15th 2012 (Water Temp. 27.5°C, pH 5.77, TDS 36 ppm, DO 1.4 ppm). Clone Colp-11 was isolated from a water sample with bottom detritus in a small boggy pond surrounded by pine trees within white sand dunes (11.066331N, 108.426602E) near Bau Trang lake, Bình Thuận Province, S.R. Vietnam. Sample was collected on May 9th 2013 (Water Temp. 36.72°C, pH 6.17, conductivity 30 μ S/cm, TDS 15 ppm, DO 4.36 ppm). Hydrochemistry characteristics were measured using the multiparameter meter HI 9828 (Hanna instruments). Field studies in Vietnam were conducted under permits issued by the administration of Cát Tiên National Park, Vietnam, and authorized by Vietnam-Russian Tropical Centre, Coastal Branch (Nha Trang, Vietnam).

Culturing and microscopy

All strains were propagated on the bodonid *Parabodo caudatus* strain BAS-1 or the chrysophyte *Spumella* sp. strain OF-40 (both freshwater) grown in Pratt's medium by using the bacterium *Pseudomonas fluorescens* as food (Tikhonenkov et al. 2014). Cells grown in laboratory cultures were harvested following peak abundance after eating most of the prey and collected by centrifugation (10000 x g for 10 minutes, room temperature). Genomic DNA was extracted from fresh cells using the Epicentre DNA extraction kit (Cat. No. MC85200). All three cultures died after one year of continuous passaging.

Light microscopy observations of Colp-16 were made by using a Zeiss AxioScope A.1 equipped with a DIC and phase contrast water immersion objectives (63x) and the analog video camera AVT HORN MC-1009/S. Light microscopy observations of Kat-1 and Colp-11 were made by using a Zeiss AxioPlan2 Imaging microscope equipped with a DIC contrast objective (40x) and the video camcorder Canon XL H1s.

For transmission electron microscopy (TEM), cells were centrifuged, fixed in 0.6% glutaraldehyde and 2% OsO₄ (final concentration) prepared using a 0.1 M cacodylate buffer (pH 7.2) at 1°C for 30-60 min, and dehydrated in an alcohol and acetone series (30, 50, 70, 96, and 100%; 20 minutes in each step). Finally, cells were embedded in a mixture of araldite and epon (Luft 1961). Ultrathin sections were obtained using a LKB ultramicrotome. TEM observations were done by using the JEM-1011 (Japan) electron microscope.

Full 18S rRNA gene sequencing & phylogenetic analyses

The full 18S rRNA gene sequence of isolates Colp-16 and Kat-1 was amplified using general eukaryotic primers 18SFU and 18SRU (Tikhonenkov et al. 2016), while that of isolate Colp-11 was obtained with general eukaryotic primers PF1 and FAD4

(Tikhonenkov et al. 2014). Products were then cloned and sequenced by Sanger dideoxy sequencing. The first half of the 18S rRNA gene of environmental clone Wsoil (one of the NC10 phylotypes identified in Endomyxa-enriched clone libraries from soil samples) was amplified from the corresponding soil DNA extraction and sequenced using an Endomyxa-biased forward primer sA4-en2 (5'-CWG TGA AAC TGC RGA TGG-3') and a phylotype-specific reverse primer in hypervariable region V7, V7r-nc14 (5'- CCT CTA AGA AGT CAT TCT CCA ACG-3'). The 18S rRNA gene sequences generated in this study were deposited in the NCBI GenBank database under accession numbers MG760572, MG818163 - MG818165, and MG818835 - MG818856.

Sequence alignments including our new cloned and isolate-derived sequences and previously published sequences representative of the known diversity of Cercozoa and Endomyxa were generated using MAFFT v. 7 (e-ins-i algorithm) (Kato and Standley 2013) and masked to omit ambiguously aligned positions. Phylogenetic analyses were performed on the CIPRES Science Gateway (Miller et al. 2010). Maximum likelihood analyses were performed with the program RaxML v 8 (Stamatakis et al 2008; Stamatakis 2014) and Bayesian analyses using MrBayes v 3.2.5 (Ronquist et al. 2012). Two separate MC³ runs with randomly generated starting trees were carried out for 5 million generations each with one cold and three heated chains. The evolutionary model applied included a GTR substitution matrix, a four-category autocorrelated gamma correction and the covarion model. All parameters were estimated from the data. Trees were sampled every 1,000 generations. The first 1.25 M generations were discarded as burn-in (trees sampled before the likelihood plots reached stationarity) and a consensus tree was constructed from the remaining sample.

TSA-FISH and flagellar staining

Newly collected freshwater plankton samples from the English Lake District in which NC10 had been frequently detected by PCR were prepared for TSA-FISH and flagella antibody staining. 200 ml samples were fixed with formaldehyde and filtered through 0.6 µm pore-size polycarbonate filters. The filters were stored at -80 °C in the dark until being processed. Prior to hybridisation the filters were sliced into small wedges and equilibrated in 50µl PBS and 0.3% triton in humid chamber for 1 hour at 40°C. Then to minimize endogenous peroxidases the filters were incubated for 10 mins with 10µl of 3% H₂O₂. The filters were washed 3 times with fresh wash buffer (0.9M NaCl, 5mM EDTA, 20mM Tris-HCl (pH 7.5), 0.01% sodium dodecyl sulfate).

Hybridization was carried out by covering filter pieces with 20 µl of hybridization buffer (35% deionized formamide, 0.9M NaCl, 20 mM Tris-HCl (pH 7.5), 0.01% sodium dodecyl sulfate and 10% blocking reagent) containing 2 µl of horseradish peroxidase (HRP)-labeled probe (stock at 50 ng/µl) and incubating in a humid chamber at 40°C overnight. Two oligonucleotide probes specific to the planktonic clade of NC10 were designed: NC10-V4 (5'-GTATTCACAGCAATAAGCCTGCTT-3') and more specific NC10-C6 (5'-CTCCACTTCTTGGGTGCC-3'). The probe EUK516 (5'-ACCAGACTTGCCCTCC-3') (Amann et al. 1990) was used as a positive control for each batch of hybridisations. To test for any false positives caused by endogenous

peroxidases further procedures were undertaken: 1) a control without probe (no fluorescence were detected), 2) incubation of filters with 3% hydrogen peroxide to remove peroxidases. The formamide concentration was also reduced in test hybridisations to determine whether the patchy specific probe signal was caused by too stringent hybridization conditions.

After hybridisation filters had three five-minute wash steps with wash buffer (as before) at 42°C. Filters were equilibrated in TNT buffer (0.1M Tris-HCl (pH 7.5), 0.15M NaCl, 0.05% Tween 20) for 15 minutes at room temperature. TSA Fluorescein was made up as per kit instructions (Perkin Elmer) and 50µl of fluorophore working solution was added to each filter and incubated for 30 minutes at room temperature in the dark. The filters were then washed in pre-warmed TNT buffer and incubated for 15 mins at 55°C replacing with fresh buffer every 5 minutes. The filters were removed and air dried (still in the dark).

Immediately after the TSA-FISH protocol the filters were fixed in 4% formaldehyde in PBS for 10 minutes. Samples were rinsed in PBS and incubated in blocking solution (1x PBS with 4% BSA, 0.3% Triton X, 0.02% sodium azide) for 1 hour at 4°C. Primary antibody, monoclonal anti-acylated tubulin antibody produced in mouse (Sigma) was diluted 1:200 with blocking solution and incubated for 3 hours at 4°C. Samples were washed 5 times with PBS and 0.3% triton X. Secondary antibody goat anti-mouse igG TRITC (abcam) was diluted 1:400 in blocking solution and sample was incubated for 1 hour at 4°C. The filters were washed 5 times in PBS and 0.3% triton X and air dried on slides for 5 minutes. They were then mounted with Vectasheld containing DAPI and the coverslips sealed, and examined under a Nikon A1 confocal microscope.

Results

Diversity and habitat range of Aquavolonida (= NC10)

We set out to explore the diversity and ecology of Aquavolonida by screening environmental DNA extractions using two different PCR strategies (Table 1 and Supplementary Table S2). First, we designed PCR primers targeting 'core' Aquavolonida, i.e. the diversity of NC10 environmental clones that was publicly available at the onset of this study (PCR strategy I; see Supplementary Table S1). Applying PCR strategy I to a range of terrestrial and marine samples showed that Aquavolonida could only be identified in freshwater habitats, although not exclusively in plankton samples: we also detected Aquavolonida phylotypes in freshwater sediments and in association with river biofilms (Table 1). However, the PCR primers used in PCR strategy I could have so high a specificity to previously known, 'core' Aquavolonida sequences that they may not allow detection of more divergent Aquavolonida lineages, if any.

To address this issue, we relied on the results of preliminary surveys of general endomyxan diversity in various environments that were conducted at the NHM as part

of a research project on Endomyxa (data not shown). This screening effort was designed to explore as widely as possible the diversity of Endomyxa and NCs 10 to 12 (using a broad Cercozoa-specific forward primer), while excluding the overwhelming diversity of core Cercozoa (Filosa) using 'anti-Filosa' reverse primers (see Supplementary Table S2). In clone libraries generated with these primers from soils samples, we identified (at low frequency compared to other soil endomyxans like Vampyrellida) two very distinct, novel phylotypes (clones Dsoil and Wsoil) that showed specific phylogenetic affinities to 'core' Aquavolonida. We designed a Wsoil-specific reverse primer to obtain its full 18S rRNA gene sequence from the corresponding soil DNA extraction. In parallel, preliminary phylogenetic analysis of the full 18S rRNA gene sequences of flagellates isolated from a reservoir and wetland in Vietnam (morphologically described below) showed them to also be related to 'core' Aquavolonida, some of them with sequences highly similar to clone Dsoil.

All of these relatives to 'core' Aquavolonida would be missed by the primers used in PCR strategy I. We therefore designed PCR strategy II to target more broadly the diversity of Aquavolonida, and applied these primers to the same habitat types tested with PCR strategy I (Table 1). Figure 1 shows a Bayesian phylogenetic tree of the diversity of Aquavolonida 18S rRNA gene phylotypes identified with PCR strategy II, using tremulids and NC12 as outgroups. 'Core' Aquavolonida, which comprises all previously identified sequences from GenBank, all our cloned sequences from freshwater planktonic samples, and a few cloned sequences from freshwater sediments and river biofilm, form the bulk of Aquavolonida diversity and are here labelled lineage Aquavolonida NC10-A. The NC10-A radiation is subdivided into three phylogenetically distinct subgroups Ai to Aiii; Ai is the most diverse both among GenBank sequences and our clone libraries. No morphological information is available for any of them; however Ai is the subgroup targeted by our FISH probes (see below).

Below 'core' Aquavolonida, we identified a grade of three earlier-branching lineages (B-D) that by contrast were not identified in freshwater planktonic samples. Aquavolonida NC10-B comprises most of our isolated strains, and the two new species described in this paper (see below). This lineage was detected in soils, freshwater sediments, and river biofilm. The single, divergent sequence in Aquavolonida NC10-C is another isolated strain, a gliding flagellate with a similar gross morphology and feeding mode to the new species, to be formally described elsewhere. The earliest branching lineage (Aquavolonida NC10-D) is composed entirely of environmental sequences (including clone Wsoil), and was so far exclusively detected in soil samples. Even though PCR strategy II targets a much broader diversity of organisms than PCR strategy I, it did not detect any lineage in addition to those that were used to design the primers.

Morphology of the novel isolates

The three novel isolates Colp-16, Kat-1, and Colp-11 from freshwater biotopes of Southern Vietnam all belong to lineage Aquavolonida NC10-B. At present this is the only Aquavolonida lineage for which we have data on both cell morphology and habitat range. We create the new genus *Aquavolon* (see the Taxonomic Summary below) to

accommodate this lineage detected in soils, freshwater sediments, and river biofilm samples. Isolate Colp-16 is distinct from the other two isolates both morphologically and in its 18S rRNA gene sequence and is here described as *Aquavolon hoantrani* n. gen. n. sp. (type species of the genus). The other two isolates are morphologically undistinguishable and have identical 18S rRNA gene sequences; Kat-1 is described as *Aquavolon dientrani* n. sp. and Colp-11 is regarded as a co-specific isolate.

Aquavolon hoantrani n. gen. n. sp. (Fig. 2A-G) (clone Colp-16)

Light microscopy

Cells are elongated-oval, slightly metabolic, not flattened, 7-10 μm long and 3.5–5 μm wide. The anterior end of the cell is usually wider than the posterior one and both are roundish. No anterior rostrum. A remarkable lateral depression is situated in the middle lateral point of the cell body (Fig. 2A-D). Pseudopodia were not observed.

Two long heterodynamic flagella originate near the anterior cell end (Fig. 2E). The anterior flagellum is 1.5-2 times longer its cell body length, makes weak waving movements and directs forward at cell swimming. The posterior flagellum is about 2-3 times the cell length, lies along the cell body at the anterior part of the cell and stands apart from the cell at the level of lateral depression, directs backward and trails behind the cell (Fig. 2B-D). A contractile vacuole (not shown) and nucleus (Fig. 2A,F) are located subapically. The cytoplasm contains light-refracting granules (Fig. 2C,D). Cells usually swim rapidly near the bottom of Petri dishes and very rarely glide on the surface. The anterior flagellum does not wrap around the anterior part of the body.

A. hoantrani is a predator, feeding on other flagellates (e.g. bodonids) and perished in the absence of eukaryotic prey. However, *A. hoantrani* did not feed very actively and never captured all available prey cells. Cannibalism was not observed. A large food vacuole is formed at the posterior end of cell-body following the feeding and cells become wider (Fig. 2F). Division was not observed. Reproduction or resting cysts were not seen in the culture.

Ultrastructure

The cell surface is covered only with a plasma membrane. Two smooth heterodynamic flagella with typical axonemes (9+2) insert subapically (Fig. 3A-D). The flagella basal bodies lie approximately at a right angle to each other and are connected by at least one fibril (Fig. 3A). Two central microtubules contact with the axosome (Fig. 3C). Kinetosome 1 (basal body of posterior flagellum) is very long (more than 1 μm) (Fig. 3D). The distal end of each kinetosome is connected with 9 transition fibers to the plasma membrane (Fig. 3C-F). Some microtubules (mrt) extend from the kinetosome to the plasma membrane (Fig. 3D, F, G). The cell possesses two wide microtubular bands. The microtubular band 1 begins from the kinetosome 1 and contains up to 14 microtubules (Fig. 3C,G, H). The microtubular band 2 starts from kinetosome 2 and consists of 5-7 microtubules (Fig. 3D,G,I). The microtubules of this band are surrounded

by dark amorphous material. The details of arrangement of microtubular bands have not been established.

A nucleus with nucleolus is situated close to the kinetosomes (Fig. 3A, C, D, G, 4A). The narrowed end of the nucleus directs to the kinetosomes (Fig. 3A). Golgi apparatus of typical structure is located near the nucleus (Fig. 3D). Several mitochondria were observed to have tubular cristae (Fig. 3A, 4A, B). The contractile vacuole, surrounded by small vesicles, lies close to the nucleus (Fig. 3G, 4C). Food vacuoles contain the engulfed eukaryote prey on different stages of digesting. The remnants of *Parabodo caudatus* (see flat mitochondrial cristae and putative kinetoplast) are visible inside the food vacuole (Fig. 4D). *A. hoantrani* phagocytoses the cells of eukaryotic prey intact. The cytostome and cytopharynx or longitudinal groove for prey engulf are absent. Small spherical osmiophilic dark-stained granules are embedded inside the vesicles (Fig. 4B, E). These structures likely represent extrusive organelles, which resemble muciferous bodies of some cercomonads and Pseudosporida (Incertae sedis). Spherical granules of reserved substance were found in the cytoplasm (Fig. 4F).

Aquavolon dientrani n. sp. (Fig. 2H-P) (clones Kat-1 and Colp-11)

The general morphology, body plan, and feeding are the same as in *A. hoantrani*. *A. dientrani* is also characterized by elongated-oval, slightly metabolic, not flattened cells with wider anterior end and lateral depression (Fig. 2H, I). Cells are slightly larger than in the other species and reach 8.5-11 μm . In contrast with *A. hoantrani*, the anterior flagellum of *A. dientrani* is very short, about the half of cell length or slightly longer (Fig. 2G-L), makes very fast flapping movements at swimming and often invisible. Sometimes anterior flagellum wraps around the anterior part of the cell. The posterior flagellum is 2-2.5 times longer than the cell body, also stands apart from the cell at the level of lateral depression and trails behind the cell.

A contractile vacuole and nucleus situated close to the anterior end (Fig. 2H,N). Light-refracting granules were observed in the cytoplasm. Cells usually swim with rotation around their axes, only rarely gliding on the substrate (Supplementary materials, Video 1). Sometimes after contact with the substrate, the cells can form short, relatively wide pseudopodia posteriorly (Fig. 2N, O), or produce a single 'tail' pseudopodium (Fig. 2P). Like the previous species, *A. dientrani* captures other flagellates and perishes in the absence of eukaryotic prey. Division was not observed. Reproduction or resting cysts were not seen in the culture.

TSA-FISH of freshwater planktonic samples

Two TSA-FISH probes designed to target specific motifs in the variable region V4 and in the conserved region between V5 and V6 of the 18S rRNA (Supplementary Table 2) were used to determine the morphology of the planktonic 'core' Aquavolonida (lineage NC10-Ai; Fig. 1). Cells showing the green fluorophore of the specific probe were very infrequently seen on all filters examined. In contrast, the broadly targeted eukaryotic probe illuminated many cells on the same filters. Both Aquavolonida-specific probes

independently generated the same signal, i.e. flagellated cells of the same morphological type and within the size range given for the two *Aquavolon* spp., above. Unusually (and in contrast to the pan-eukaryotic control), the ribosomal probe was in all cases localized to distinct areas in the posterior half of the cell (Fig. 5). Possible causes for this were tested in various ways (see Methods). Endogenous peroxidases were limited by hydrogen peroxide treatment, and tested for in a control hybridisation without the probe; neither of these suggested that the patchy signal was a false positive caused by the presence of peroxidases. Further, the patchy signal was consistent across the two independent probes tested and a range of formamide concentrations moderating the hybridization stringency.

Improved phylogeny of Aquavolonida ord. nov.

Using the new full-length 18S rRNA gene sequences generated in this study, we reassessed the phylogenetic position of Aquavolonida within Rhizaria. Figure 6 shows a Bayesian phylogenetic analysis in which a strongly supported NC10 (Aquavolonida) forms a clade with Tremulida, the two main lineages comprising NC12, and two single sequence lineages from moss pillars in an Antarctic freshwater lake (clone MPE2-26; Nakai et al. 2012) and an oxygen-depleted marine environment (clone CCW46; Stoeck & Epstein, 2003). This whole clade is sister to Endomyxa, albeit without strong support, but is clearly distinct from the rest of Cercozoa (= Filosa). Relationships within Endomyxa and Filosa are consistent with previous studies (e.g. Bass et al. 2009a).

Sequence signatures in the 18S rRNA gene (Table 2) clearly support the monophyly of Aquavolonida as defined here, i.e. comprising four main lineages NC10-A (= 'core' Aquavolonida), NC10-B (= *Aquavolon* n. gen.), NC10-C (= isolate Lapot), and NC10-D (environmental lineage so far exclusively found in soils). These signatures are detailed below following the helix numbering in Wuyts et al. (2000) and were reflected in the primers we designed for PCR strategy II (see Table 1 and Supplementary Table S2). They consist of two complementary substitutions from U-A to G-C and from G-C to C(U)-G(A) in helix 11 (variable region V2, primer V2f-NC), one complementary substitution from A-T to T(C)-G(A) in helix 48 (conserved region between V8 and V9, primer C9r-NC), and a specific motif of two adjacent substitutions (AU instead of YG) in a non-binding part of the 3' stem of helix 25 (conserved region between V4 and V5, primer C5f-NC). Some of these signatures can be found in part in other eukaryotic lineages, but their combination is uniquely found only in, and shared by, all NC10 lineages. Closely related groups (tremulids, NC12, and environmental clones MPE2-26 and CCW46) all lack these signatures. The clade consisting of NC10, NC12, Tremulida, and environmental clones MPE2-26 and CCW46 is supported by a complementary substitution of A-U to C-G in helix 29 (variable region V5) and a complementary substitution of A-U to U-A in helix 45 (variable region V8, primer V8r-NC). Finally, a unique substitution from U to A in the loop between helices 20 and 21 characteristic of all Endomyxa clades (e.g. Vampyrellida, Phytomyxea, *Filoreta*, *Gromia*, *Ascetosporea*) is also found in the clade comprising NC10, NC12 and Tremulida, supporting a closer relationship to Endomyxa than core Cercozoa (= Filosa).

Discussion

At the onset of this study, NC10 (Bass et al. 2009a) was only known from environmental sequences derived from samples (mostly planktonic) collected in various freshwater lakes worldwide. Our screening of environmental DNA extractions from a selection of terrestrial and marine habitats using two complementary PCR strategies with distinct clade-specific primers (see Fig. 1 and Table 1) reveals that although these 'core' Aquavolonida (lineage NC10-A) seem to represent the bulk of the diversity of the clade, many more phylotypes basal to them can be found in other terrestrial habitats, from soils to freshwater sediments. To date however, members of Aquavolonida remain undetected in marine environments, even in large-scale, broadly targeted surveys such as BioMarkS (Logares et al. 2014) and *Tara*-Oceans (de Vargas et al. 2015), which readily detect NC10 when applied to freshwater samples (Lara et al 2015; Grossman et al. 2016; Yi et al. 2017).

The results of our environmental DNA screening also provide information about the habitat range and relative diversity of each Aquavolonida lineage. First, both PCR strategies show that NC10-A are not exclusively planktonic, but can also be found in freshwater sediments and associated to biofilms. However, PCR strategy I (targeting NC10-A only) generated significantly more positives (and uncovered a much higher diversity of 18S types) from planktonic samples than from sediments and river biofilm, and most of that diversity belongs to the subclade targeted by our TSA-FISH probes (NC10-Ai). Even recently flooded soil/grassland sites produced a high frequency of positive PCR results. Therefore, although our TSA-FISH results indicate that NC10-A cells do not necessarily occur in high density in planktonic samples, they appear to be genetically diverse, very widely distributed, and present in most freshwater habitats (predominantly in the water column as genuinely planktonic protists).

By contrast, we couldn't detect other Aquavolonida lineages in planktonic samples, suggesting that only NC10-A have adapted to a planktonic lifestyle (see the results obtained with PCR strategy II in Table 1). The most basal lineage we detected (NC10-D) seems restricted to soils, and was detected at relatively high frequency in such samples. The lineage including the new genus *Aquavolon* (NC10-B) was found in soils, freshwater sediments, and river biofilm. However, in soils it was detected at lower frequency than NC10-D, and in freshwater sediments and river biofilm at lower frequency than NC10-A. NC10-B and NC10-D also both seem to comprise a much lower diversity of 18S types than NC10-A (see Fig. 1).

By providing for the first time phenotypic data on members of Aquavolonida, we show that it contains both swimming (new genus *Aquavolon*) and gliding (isolate Lapot, which will be formally described elsewhere) biflagellates with morphological affinities to *Tremula longifila*, the first described taxon from NC11 (Howe et al. 2011). At present, phenotypic data are available only for two (B & C) of four Aquavolonida lineages.

However, (1) their morphological similarity to the more distantly related *Tremula*, (2) the specific locomotion of *Aquavolon*, (3) the inference from environmental sequencing that the most diverse, later-diverging NC10-A are predominantly found in planktonic habitats and probably absent from soils, (4) TSA-FISH results concordant with morphology of *Aquavolon* as shown by light microscopy (Fig. 2), and (5) the strongly supported sister-relationship between NC10-A and NC10-B, together suggest that these two lineages at least probably comprise more or less similar, eukaryovorous swimming flagellates. As NC10-A sequences are frequently detected in freshwater plankton environmental sequencing surveys, our findings will enable more informed interpretation of such studies.

Unlike Aquavolonida and Tremulida, previous environmental sequencing results show that members of NC12 are found in a broader range of both marine and non-marine habitat types including soils, sediments, and water column, which suggests a wider diversity of cell types and perhaps lifestyles. It is interesting that so little of the large diversity in the broader clade comprising NCs 10 to 12 has been isolated into culture. This is perhaps because they have a general requirement for eukaryotic prey, which are relatively infrequently used in isolation attempts. Single cell-sorting surveys of freshwater habitats should be more informative about their true abundance, which based on detection frequency representing rich lineage diversity, appears to be high.

Phylogenetically, *Aquavolon* can be most directly compared morphologically with *Tremula*, up to now the only identified protist within NCs 10 to 12 of known phenotype. *Tremula longifila* is similarly a non-amoeboid biflagellate heterotrophic protist with a rounded or elongated or spindle-shaped cell and, like *Aquavolon*, has variably pronounced medium notch (Howe et al. 2011). In contrast to *Aquavolon*, the flagella of *T. longifila* do not appear to insert close together at the anterior end of the cell; however, although unconfirmed by ultrastructural studies, the proximal part of the posterior axoneme is likely intracellular (Howe et al. 2011). *Aquavolon* can be distinguished from Cercomonadida, Glissomonadida, and Granofilosea (*Massisteria*) by the absence of a microbody (or paranuclear body) with amorphous contents, which usually joins the posterior part of the nucleus.

Tremula is so far unique in Rhizaria with respect to its gliding on both exceptionally long flagella, which are simultaneously in contact with the substratum (Howe et al. 2011). All known metromonads, glissomonads, pansomonads, cercomonads, thecofiloseans, and imbricateans glide on their posterior flagellum only. Very unusually for rhizarian flagellates, but similar to the putative glissomonad *Katabia gromovi* (Karpov et al. 2003), *Aquavolon* swims almost perpetually, members of both taxa rotating along their longitudinal axes. They also share subapical insertion of flagella, the same kinetosome orientations, narrow posterior cell end, and the shape of their mitochondria cristae, but *Aquavolon* lacks microbodies, mushroom-like bodies (refractile granules), kinetocysts and cysts.

Eukaryovory is also relatively rare in known rhizarian flagellates. Other examples include the metromonads *Metromonas simplex* and *Metopion fluens*, and the

pansomonad *Aurigomonas solis* (Vickerman et al. 2005), which capture whole prey cells, usually bodonids, chrysomonads, or euglenids. *Aquavolon* differs ultrastructurally from metromonads and *Aurigamonas* in lacking large extrusive organelles (or haptosomes: *Aurigamonas*), lacking flagellar hairs, and in having a simple plasma membrane rather than a complex envelope of a thin bilayer membrane and an outer layer of thin short fibers (metromonads) or axopodia-like appendages (haptopodia) (*Aurigamonas*). Interestingly, both *Aquavolon* and *Aurigamonas* have long kinetosomes, but *Aquavolon* also possesses two wide microtubular bands.

Our Bayesian phylogenetic analysis places NC10 (including *Aquavolon*), Tremulida (NC11) and NC12 as a sister clade to Endomyxa, which includes many amoeboid organisms producing filopodia, rhizopodia, and reticulopodia. Many are obligate endoparasites of land plants, stramenopile algae (e.g. Neuhauser et al 2014) and animals (e.g. Hartikainen et al 2014; Ward et al 2016, 2018), but others are free-living protists, e.g. the reticulopodial amoeba *Filoreta marina* (Bass et al 2009b), which also has mitochondria with tubular cristae (Myl'nikov & Myl'nikov 2011). Flagellate cells are apparently relatively rare in Endomyxa. *Gromia* produces flagellated dispersal cells or gametes (Hedley & Bertaud, 1962), and plasmodiophorids (Phytomyxea) and paradinids (Ascetosporea) form flagellated zoospores. Other ascetosporeans produce non-flagellated spores (e.g. haplosporidians). *Aquavolon* shares some morphological similarity with the biflagellate zoospores of Plasmodiophorida, in which the kinetosomes are also usually very long: 1.6 – 1.8 μm in *Polymyxa betae* (Barr & Allan, 1982) and *Sorosphaera veronicae* (Talley et al., 1978). However, the kinetosomes of the closely related *Plasmodiophora brassicae* are appreciably shorter (0.5 μm) (Buczacki & Clay, 1984). Interestingly, the angle between two kinetosomes in different plasmodiophorids ranges from 150° in *Polymyxa graminis* to 30° in *Woronina pythii* and *Plasmodiophora brassicae* (Barr & Allan, 1982; Miller & Dylewski, 1983; Buczacki & Clay, 1984). The zoospores of *Spongospora subterranea* show a lateral insertion of the flagella at an angle of about 180° to each other and a bulging ring at the points of insertion (Merz, 1992).

Recent phylogenomic analyses focusing on Rhizaria show a basal split between two major assemblages, Cercozoa (= Filosa) on the one side, and Endomyxa plus Retaria on the other side (e.g. Sierra et al. 2016; Krabberød et al. 2017). In this context, the phylogenetic position of the clade comprising Aquavolonida, Tremulida and NC12 (if they are confirmed to form a clade) becomes of particular significance. 18S rRNA gene phylogenies place it between these two major assemblages of Rhizaria, which makes it a potential third lineage at the deepest rhizarian divergence. On the basis of a shared 18S rRNA sequence signature (Table 2), we hypothesize that the Aquavolonida-Tremulida-NC12 clade may be sister to the highly diverse radiation of Endomyxa + Retaria, and that the ancestral rhizarian phenotype was likely a unicellular heterotrophic biflagellate. Phylogenomic and comparative genomic analyses including members of Aquavolonida and Tremulida are now required to resolve the deepest relationships within Rhizaria.

Taxonomic summary

We describe the new order Aquavolonida to include all members of NC10 on the basis of its phylogenetic (high statistical support, unifying sequence signatures) and ecological (seemingly exclusively non-marine) coherence. This name, based on the only described genus, reflects the tendency of the clade to adapt to an aquatic habitat, but is not intended to imply that all lineages are actual swimmers.

Eukaryota; Rhizaria; Aquavolonida ord. nov. Bass and Berney

Diagnosis. Defined as the least inclusive clade containing the last common ancestor of the rhizarian lineages sharing a unique combination of 18S rRNA sequence signatures consisting of two complementary substitutions from U-A to G-C and from G-C to C(U)-G(A) in helix 11, one complementary substitution from A-T to C(T)-G(A) in helix 48, and a specific motif of two adjacent substitutions (AU instead of YG) in a non-binding part of the 3' stem of helix 25 (all bold and underlined in Table 2), and all their descendants.

Remarks. This is a node-based definition intended to apply to a crown clade; qualifying clause: the name does not apply if any of the following fall within the specified clade: *Tremula longifila* Howe and Cavalier-Smith 2011, or the organisms matching environmental clones MPE2-26, CCW-46, NAMA KO-20, and CCA33. Aquavolonida corresponds to the rhizarian clade previously referred to as Novel Clade 10 as identified in this study, i.e. consisting of at least four lineages, one of which contains *Aquavolon* n. gen. The sequence signatures defining the order are found in these four lineages, but we cannot exclude that they were secondarily lost or reversed in as yet unidentified members of the order with a more divergent 18S rRNA sequence. Based on existing evidence, other members of the order will likely but not necessarily all be biflagellated, eukaryovorous, unicellular protists and are expected to be restricted to terrestrial habitats. Lineages within Aquavolonida are here hypothesised to show a gradual morphological and behavioural adaptation to an aquatic lifestyle with a swimming locomotion, hence the name chosen for the order.

Etymology. Derived from *Aquavolon*, first described genus in the order, and likely to represent its most common phenotype.

Aquavolon n. gen. Tikhonenkov, Mylnikov, and Bass

Diagnosis. Unicellular protist with two smooth subapical heterodynamic flagella. The flagella kinetosomes lie approximately at a right angle to each other and are connected by at least one fibril. The posterior kinetosome is very long ($> 1 \mu\text{m}$). Cell slightly metabolic, not flattened, with remarkable lateral depression in the middle lateral point of the cell body. Single contractile vacuole and nucleus located anteriorly. Several mitochondria have tubular cristae. Rapidly swimming and rarely gliding protist. Eukaryovorous.

Etymology. From aqua = water, Latin, and volo = to fly, Latin.

Zoobank Registration. urn:lsid:zoobank.org:act:635F34CC-265E-4F6C-88A4-

C0EC7E84E778

Type species. *Aquavolon hoantrani*.

Aquavolon hoantrani n. sp. Tikhonenkov, Mylnikov, and Bass

Diagnosis. Cells are elongated-oval with lateral depression, 7-10 μm long and 3.5–5 μm wide. The anterior end is wider than posterior one and both are roundish. The anterior flagellum is 1.5-2 times longer its cell body length, makes flapping movements and directs forward at cell swimming. The posterior flagellum is about 2-3 times the cell length, lies along the cell body at the anterior part of the cell and stands apart from the cell at the level of lateral depression, directs backward and trails behind the cell. Reproduction or resting cysts have not been found in culture.

Type material. A block of chemically fixed resin-embedded cells of the type strain, Colp-16, is deposited in Marine Invertebrate Collection, Beaty Biodiversity Museum, University of British Columbia as MI-PR210. This constitutes the name-bearing type of the new species (a hapantotype).

Type Figure. Figure 2A illustrates a live cell of strain Colp-16.

Type locality. Water with bottom detritus from freshwater Suoi Da Reservoir (Hồ Suối Đá, viet.), Bình Thuận Province, S.R. Vietnam.

Etymology. Named after Dr. Hoan Tran, Vietnamese biologist and ecologist, who significantly contributed to sample collection and field trips organization and management in S.R. Vietnam.

Gene sequence. The 18S rRNA gene sequence has the GenBank Accession Number MG760572.

Zoobank Registration. urn:lsid:zoobank.org:act:3E585A8A-9459-4E1F-956B-FE547535EE1F

Aquavolon dientrani n. sp. Tikhonenkov, Mylnikov, and Bass

Diagnosis. Elongated-oval cells 8.5-11 μm long with wider anterior end and lateral depression. Anterior flagellum is about the half of cell length or slightly longer, makes fast flapping movements at swimming and often invisible, may wraps around the anterior part of the cell. The posterior flagellum is 2-2.5 times longer than the cell body, stands apart from the cell at the level of lateral depression and trails behind the cell. Cells can form short, relatively wide pseudopodia posteriorly or produce single tail pseudopodium. Reproduction or resting cysts have not been found in culture.

Type Figure. Figure 2I illustrates a live cell of *A. dientrani*.

Type locality. Wetland water with decaying plant material, Đắc Lua Village, Dong Nai Province, S.R. Vietnam.

Etymology. Named after Mr. Tran Duc Dien, Vietnamese biologist and ecologist, who significantly contributed to sample collection and field trips organization and management in S.R. Vietnam.

Gene sequence. The 18S rRNA gene sequence has the GenBank Accession Number

MG818163.

Zoobank Registration. urn:lsid:zoobank.org:act:C99A3C2D-AF23-4A48-ACE0-377340713E8C

Acknowledgments

This work was supported by grants from the Russian Foundation for Basic Research (no. 17-04-00899, 17-04-00565) to DT and APM and in the framework of the state assignment of FASO Russia (theme no. AAAA-A18-118012690098-5), from the Natural Sciences and Engineering Research Council of Canada (227301) to PJK, and NERC grants NE/H009426/1 (to DB and CB) and NE/H000887/1 (to DB). RF was supported by Grant-in-Aid funding to The Natural History Museum. CB is grateful to the International Society of Protistologists and the Gordon and Betty Moore Foundation for current funding (grant GBMF5257 / UniEuk project, www.unieuk.org), and was also supported by the 'Investissements d'Avenir' programme OCEANOMICS (ANR-11-BTBR-0008) awarded by the French Government via Agence Nationale de la Recherche. SN was supported by a Schroedinger Fellowship hosted at the NHM.

We thank Dr. Hoan Q. Tran and Tran Duc Dien for help with sample collection in Vietnam, as well as the staff of the Vietnam-Russian Tropical Centre, Coastal Branch (Nha Trang, Vietnam), especially Nguyen Thi Hai Thanh and Tran Thanh Quang, for their assistance with trip management and sampling. We also thank Prof. Stephen Maberly for providing samples and assistance from the English Lake District.

LITERATURE CITED

- Amann, R.I., Binder, B.J., Olson, R.J., Chisholm, S.W., Devereux, R. & Stahl, D.A. 1990. Combination of 16S rRNA-targeted oligonucleotide probes with flow cytometry for analyzing mixed microbial populations. *Appl. Environ. Microbiol.*, 56:1919-1925.
- Barr, D.J.S. & Allan, P.M.E. 1982. Zoospore ultrastructure of *Polymyxa graminis* (Plasmodiophoromycetes). *Can. J. Bot.* 60:2496-2504.
- 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.
- Bass, D., Chao, E.E.-Y., Nikolaev, S., Yabuki, A., Ishida, K.-I., Berney, C., Pakzad, U., Wylezich, C. & Cavalier-Smith, T. 2009b. Phylogeny of novel naked filose and reticulose Cercozoa: Granofilosea cl. n. and Proteomyxidea revised. *Protist*, 160:75-109.
- Bass, D., Howe, A.T., Mylnikov, A.P., Vickerman, K., Chao, E.E.-Y., Edwards Smallbone, J., Snell, J., Cabral Jr, C. & Cavalier-Smith, T. 2009a. Phylogeny and classification of Cercomonadida: *Cercomonas*, *Eocercomonas*, *Paracercomonas*, and *Cavernomonas* gen. n. *Protist*, 160:483-521.

Berney, C., Romac, S., Mahé, F., Santini, S., Siano, R., & Bass, D. 2013. Vampires in the oceans: predatory cercozoan amoebae in marine habitats. *ISME J.*, 7:2387-99.

Buczacki, S.T. & Clay, C.M. 1984. Some observations on secondary zoospore development in *Plasmodiophora brassicae*. *Trans. Br. Mycol. Soc.* 82: 339-344.

de Vargas, C., Audic, S., Henry, N., Decelle, J., Mahé, F., Logares, R., Lara, E., Berney, C., Le Bescot, N., Probert, I., Carmichael, M., Poulain, J., Romac, S., Colin, S., Aury, J.M., Bittner, L., Chaffron, S., Dunthorn, M., Engelen, S., Flegontova, O., Guidi, L., Horák, A., Jaillon, O., Lima-Mendez, G., Lukeš, J., Malviya, S., Morard, R., Mullet, M., Scalco, E., Siano, R., Vincent, F., Zingone, A., Dimier, C., Picheral, M., Searson, S., Kandels-Lewis, S.; Tara Oceans Coordinators, Acinas, S.G., Bork, P., Bowler, C., Gorsky, G., Grimsley, N., Hingamp, P., Iudicone, D., Not, F., Ogata, H., Pesant, S., Raes, J., Sieracki, M.E., Speich, S., Stemmann, L., Sunagawa, S., Weissenbach, J., Wincker, P., & Karsenti, E. 2015. Eukaryotic plankton diversity in the sunlit ocean. *Science* 348:1261605. doi: 10.1126/science.1261605.

Chantangsi, C., Esson, H. J. & Leander, B. S. 2008. Morphology and molecular phylogeny of a marine interstitial tetraflagellate with putative endosymbionts: *Auranticordis quadrivirberis* n. gen. et sp. (Cercozoa). *BMC Microbiol.*, 8:123. doi: 10.1186/1471-2180-8-123.

Chantangsi, C. & Leander, B. S. 2010a. Ultrastructure, life cycle and molecular phylogenetic position of a novel marine sand-dwelling cercozoan: *Clautriavia biflagellata* n. sp. *Protist*, 161:133-47.

Chantangsi, C. & Leander, B. S. 2010b. An SSU rDNA barcoding approach to the diversity of marine interstitial cercozoans, including descriptions of four novel genera and nine novel species. *Int. J. Syst. Evol. Microbiol.*, 60:1962-1977.

Dumack, K., Siemensma, F. & Bonkowski, M. 2017b. Rediscovery of the testate amoeba genus *Penardeugenia* (Thaumatomonadida, Imbricatea). *Protist*, in press.

Dumack, K., Bonkowski, M., Clauß, S. & Völcker, E. 2017a. Phylogeny and Redescription of the Testate Amoeba *Diaphoropodon archeri* (Chlamydephryidae, Thecofilosea, Cercozoa), De Saedeleer 1934, and Annotations on the Polyphyly of Testate Amoebae with Agglutinated Tests in the Cercozoa. *J. Euk. Microbiol.*, in press doi: 10.1111/jeu.12474

Hedley, R.H., & Bertaud, W.S. 1962. Electron-Microscopic Observations of *Gromia oviformis* (Sarcodina). *J. Protozool.* 9:79-87.

Hess, S., Sausen, N. & Melkonian, M. 2012. Shedding light on vampires: the phylogeny of vampyrellid amoebae revisited. *PLoS One*, 7:e31165. doi: 10.1371/journal.pone.0031165.

Grossmann, L., Jensen, M., Heider, D., Jost, S., Glücksman, E., Hartikainen, H., Mahamdallie, S.S., Gardner, M., Hoffmann, D., Bass, D. & Boenigk, J. 2016. Protistan community analysis: key findings of a large-scale molecular sampling. *ISME J.*, 10: 2269-2279.

Howe, A.T., Bass, D., Scoble, J., Lewis, L., Vickerman, K., Arndt, H. & Cavalier-Smith, T. 2011. Novel cultured protists identify deep-branching environmental DNA clades of Cercozoa: new genera *Tremula*, *Micrometopion*, *Minimassisteria*, *Nudifila*, *Peregrinia*. *Protist*, 162:332-372.

Hall, T.A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl. Acids Symp. Ser.*, 41:95-98.

Hartikainen, H., Ashford, O.S., Berney, C., Okamura, B., Feist, S.W., Baker-Austin, C., Stentiford, G.D. & Bass, D. 2014. Lineage-specific molecular probing reveals novel diversity and ecological partitioning of haplosporidians *ISME J.*, 8:177-186.

Jones, M.D.M., Forn, I., Gadhela, C., Egan, M.J., Bass, D., Massana, R. & Richards, T.A. 2011. Discovery of novel intermediate forms redefines the fungal tree of life. *Nature*, 474:200-203.

Karpov, S.A., Ekelund, F., & Moestrup, Ø. 2003. *Katabia gromovi* nov. gen., nov. sp. – a new soil flagellate with affinities to *Heteromita* (Cercomonadida). *Protistology*, 3:30-41.

Katoh, K. & Standley, D. 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.*, 30:772-780.

Krabberød, A.K., Orr, R.J.S., Bråte, J., Kristensen, T., Bjørklund, K.R. & Shalchian-Tabrizi, K. 2017. Single cell transcriptomics, mega-phylogeny, and the genetic basis of morphological innovations in Rhizaria. *Mol. Biol. Evol.*, 34:1557-1573.

Kühn, S., Medlin, L.K., & Eller, G. 2004. Phylogenetic position of the parasitoid nanoflagellate *Pirsonia* inferred from nuclear-encoded small subunit ribosomal DNA and a description of *Pseudopirsonia* n. gen. and *Pseudopirsonia mucosa* (Drebes) comb. nov. *Protist*, 155:143-56.

Lara, E., Seppey, C.V.W., González Garraza, G., Singer, D., Victoria Quiroga, M., & Mataloni, G. 2015. Planktonic eukaryote molecular diversity: discrimination of minerotrophic and ombrotrophic peatland pools in Tierra del Fuego (Argentina), *J. Plankt. Res.*, 37:645–655.

Lefèvre, E., Bardot, C., Noël, C., Carrias, J.F., Viscogliosi, E., Amblard, C. & Sime-Ngando, T. 2007. Unveiling fungal zooflagellates as members of freshwater picoeukaryotes: evidence from a molecular diversity study in a deep meromictic lake. *Environ. Microbiol.*, 9:61-71.

Lefèvre, E., Roussel, B., Amblard, C. & Sime-Ngando, T. 2008. The molecular diversity

of freshwater picoeukaryotes reveals high occurrence of putative parasitoids in the plankton. *PLoS One* 3: e2324. doi: 10.1371/journal.pone.0002324.

Lefranc, M., Thenot, A., Lepère, C. & Debroas, D. 2005. Genetic diversity of small eukaryotes in lakes differing by their trophic status. *Appl. Environ. Microbiol.*, 71:5935-5942.

Lepère, C., Domaizon, I. & Debroas, D. 2008. Unexpected importance of potential parasites in the composition of the freshwater small-eukaryote community. *Appl. Environ. Microbiol.*, 74:2940-2949.

Logares, R., Audic, S., Bass, D., Bittner, L., Boutte, C., Christen, R., Claverie, J.M., Decelle, J., Dolan, J.R., Dunthorn, M., Edvardsen, B., Gobet, A., Kooistra, W.H., Mahé, F., Not, F., Ogata, H., Pawlowski, J., Pernice, M.C., Romac, S., Shalchian-Tabrizi, K., Simon, N., Stoeck, T., Santini, S., Siano, R., Wincker, P., Zingone, A., Richards, T.A., de Vargas, C. & Massana R. 2014. Patterns of rare and abundant marine microbial eukaryotes. *Curr. Biol.* 24:813-821.

Hartikainen, H., Stentiford, G.D., Bateman, K.S., Berney, C., Feist, S.W., Longshaw, M., Okamura, B., Stone, D., Ward, G.M., Wood, C., & Bass, D. 2014. Mikrocytids are a broadly distributed and divergent radiation of parasites in aquatic invertebrates. *Curr. Biol.* 24:807-812.

Lara, E., Moreira, D. & Lopez-García, P. 2010. The environmental clade LKM11 and *Rozella* form the deepest branching clade of fungi. *Protist*, 161:116-121.

López García, P., Rodríguez Valera, F., Pedrós-Alió, C., & Moreira, D. 2001. Unexpected diversity of small eukaryotes in deep-sea Antarctic plankton. *Nature*, 409:603-607.

Massana, R., Castresana, J., Balagué, V., Guillou, L., Romari K, Groisillier, A., Valentin, K, & Pedrós-Alió, C. 2004. Phylogenetic and ecological analysis of novel marine stramenopiles. *Appl. Environ. Microbiol.*, 70:3528-3534.

Merz, U. 1992. Observations on swimming pattern and morphology of secondary zoospores of *Spongospora subterranea*. *Plant Pathol.* 41:490-494.

Miller, M.A., Pfeiffer, W. & Schwartz, T. 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. In Proceedings of the Gateway Computing Environments Workshop (GCE), 14 Nov. 2010, New Orleans, LA, pp 1-8.

Miller, C.E., & Dylewski, D.P. 1983. Zoosporic plant pathogens of lower plants – What can be learned from the likes of *Woronina*? In: Zoosporic plant pathogens a modern perspective, (ed. Buczacki, S. T.), pp. 249-283. Academic Press, London.

Myl'nikov, A. P., & Myl'nikov, A. A. 2011. The cell structure of the reticulopodial amoeba *Filoreta marina* Bass et Cavalier-Smith (Cercozoa). *Inland Water Biology*, 4:440–445.

Nakai, R., Abe, T., Baba, T., Imura, S., Kagoshima, H., Kanda, H., Kanekiyo, A., Kohara, Y., Koi, A., Nakamura, K., Narita, T., Niki, H., Yanagihara, K., & Naganuma, T. 2012. Microflorae of aquatic moss pillars in a freshwater lake, East Antarctica, based on fatty acid and 16S rRNA gene analyses. *Polar Biol.* 35:425-433.

Neuhauser, S., Kirchmair, M., Bulman, S., & Bass, D. 2014. Cross-kingdom host shifts of phytomyxid parasites. *BMC Evol. Biol.* 14:33.

Richards, T.A. & Bass, D. 2005. Molecular screening of free-living microbial eukaryotes: diversity and distribution using a meta-analysis. *Curr. Op. Microbiol.*, 8:240-252.

Richards, T.A., Vepritskiy, A.A., Gouliamova, D. & Nierzwicki-Bauer, S.A. 2005. The molecular diversity of freshwater picoeukaryotes from an oligotrophic lake reveals diverse, distinctive and globally dispersed lineages. *Environ. Microbiol.*, 7:1413-1425.

Ronquist, F., Teslenko, M., Van Der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A. & Huelsenbeck, J.P. 2012. Mrbayes 3.2: Efficient bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* 61:539-542.

Shiratori, T., Yokoyama, A. & Ishida, K.-I. 2014. Phylogeny, Ultrastructure, and Flagellar Apparatus of a New Marimonad Flagellate *Abollifer globosa* sp. nov. (Imbricatea, Cercozoa). *Protist*, 165:808-824.

Shiratori, T. & Ishida, K.-I. 2016. *Trachyrhizium urniformis* n. g., n. sp., a Novel Marine Filose Thecate Amoeba Related to a Cercozoan Environmental Clade (Novel Clade 4). *J. Euk. Microbiol.* 63:722-731.

Sierra, R., Cañas-Duarte, S.J., Burki, F., Schwelm, A., Fogelqvist, J., Dixelius, C., González-García, L.N., Gile, G.H., Slamovits, C.H., Klopp, C., Restrepo, S., Arzul, I., & Pawlowski, J. 2016. Evolutionary Origins of Rhizarian Parasites. *Mol. Biol. Evol.* 33:980-983.

Simon, N., Jardillier, L., Deschamps, P., Moreira, D., Restoux, G., Bertolino, P. & López-García, P. 2015. Complex communities of small protists and expected occurrence of typical marine lineages in shallow freshwater systems. *Environ. Microbiol.* 17:3610-27.

Stamatakis, A. 2014. RAxML Version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30:1312-1313.

Stamatakis, A., Hoover, P. & Rougemont, J. 2008. A rapid bootstrap algorithm for the

RAxML web servers. *Syst. Biol.* 57:758-771.

Stoeck T, Epstein S. 2003. Novel eukaryotic lineages inferred from small-subunit rRNA analyses of oxygen-depleted marine environments. *Appl. Environ. Microbiol.* 69:2657-63.

Taib, N., Mangot, J.-F., Domaizon, I., Bronner, G. & Debroas, D. 2013. Phylogenetic affiliation of SSU rRNA genes generated by massively parallel sequencing: new insights into the freshwater protist diversity. *PLoS ONE* 8:e58950. doi: 10.1371/journal.pone.0058950.

Talley, M.R., Miller, C.E. & Braselton, J.P. 1978. Notes on the ultrastructure of zoospores of *Sorosphaera veronicae*. *Mycologia* 70:1241-1247.

Tikhonenkov, D.V., Janouškovec, J., Mylnikov, A.P., Mikhailov, K.V., Simdyanov, T.G., Aleoshin, V.V., & Keeling, P.J. 2014. Description of *Colponema vietnamica* sp.n. and *Acavomonas peruviana* n. gen. n. sp., two new alveolate phyla (Colponemidia nom. nov. and Acavomonidia nom. nov.) and their contributions to reconstructing the ancestral state of alveolates and eukaryotes. *PLoS ONE* 9:e95467. doi: 10.1371/journal.pone.0095467.

Tikhonenkov D.V., Janouskovec J., Keeling P.J. & Mylnikov A.P. 2016. The Morphology, Ultrastructure and SSU rRNA Gene Sequence of a New Freshwater Flagellate, *Neobodo borokensis* n. sp. (Kinetoplastea, Excavata). *J. Eukaryot. Microbiol.* 63:220-232.

Vickerman, K., Appleton, P.L., Clarke, K.J. & Moreira, D. 2005. *Aurigamonas solis* n. gen., n. sp., a soil-dwelling predator with unusual helioflagellate organisation and belonging to a novel clade within the Cercozoa. *Protist* 156:335-354.

Ward, G.M., Bennett, M., Bateman, K., Stentiford, G.D., Feist, S.W., Kerr, R., Williams, S., Berney, C., & Bass, D. 2016. A new phylogeny and eDNA insight into paramyxids: an increasingly important but enigmatic clade of protistan parasites of marine invertebrates. *Int. J. Parasitol.* 46:605-19.

Ward, G.M., Neuhauser, S., Groben, R., Ciaghi, S., Berney, C., Romac, S., & Bass, D. 2018. Environmental sequencing fills the gap between parasitic haplosporidians and free-living giant amoebae. *J Euk Microbiol*, in press.

Wuyts, J., De Rijk, P., Van de Peer, Y., Pison, G., Rousseeuw, P. & De Wachter, R. 2000. Comparative analysis of more than 3000 sequences reveals the existence of two pseudoknots in area V4 of eukaryotic small subunit ribosomal RNA. *Nucleic Acids Res.* 28:4698-4708.

Yi, Z., Berney, C., Hartikainen, H., Mahamdallie, S., Gardner, M., Boenigk, J., Cavalier-

Smith, T. & Bass, D. 2017. High throughput sequencing of microbial eukaryotes in Lake Baikal reveals ecologically differentiated communities and novel evolutionary radiations. *FEMS Microbiol. Ecol.*, 93(8). doi: 10.1093/femsec/fix073.

Figure legends

Figure 1. Bayesian SSU rRNA gene phylogeny of Aquavolonida (NC10) diversity, including new taxa *Aquavolon hoantrani* and *A. dientrani*, with tremulids and NC12 used as outgroups. Bayesian Posterior Probabilities (BPP) and Maximum Likelihood (ML) bootstrap values are indicated on nodes; black filled circles indicate values of BPP \geq 0.95 and ML bootstrap \geq 95%. Sequences generated in this study are shown in bold text.

Figure 2. External morphology of *Aquavolon hoantrani* n. gen., n. sp. isolate Colp-16 (A-G) and *A. dientrani* n. sp. isolate Kat-1 (H-P). af = anterior flagellum, cv = contractile vacuole, fv = food vacuole, ld = lateral depression, n = nucleus, pf = posterior flagellum, ps = pseudopodium, rg = light-refracting granules. Scale bar: 5 μ m for all figures.

Figure 3. The general view, anterior end of the cell and some flagellar root structures of *Aquavolon hoantrani*. A. The longitudinal section of the cell. B. Cross-section of the flagellum. C-D. Longitudinal sections of the anterior end of the cell. E-H. Relative arrangement of the kinetosomes. I. Some microtubular bands near the nucleus. as – axosome, cv = contractile vacuole, fv = food vacuole, ga = Golgi apparatus, k1 = kinetosome 1, k2 = kinetosome 2, mb1 = microtubular band 1, mb2 = microtubular band 2, mt = mitochondria, mrt = microtubules, n = nucleus, nu = nucleolus, pf = posterior flagellum, tf = transition fibers. Scale bar: (A, C, D, I) 1 μ m, (B, E – H) 0.2 μ m.

Figure 4. The nucleus and other cell structures of *Aquavolon hoantrani*. A. Nucleus. B. Mitochondria. C. Contractile vacuole. D. Food vacuole. E. Extrusome. F. Reserve substance. cr = mitochondrial cristae of *Parabodo caudatus*, cv = contractile vacuole, ex = extrusome, fv = food vacuole, kp = putative kinetoplast of *Parabodo caudatus*, mt = mitochondrion, n = nucleus, rs = reserve substance. Scale bar: (A) 1 μ m, (B-D, F) 0.5 μ m, (E) 0.2 μ m.

Figure 5. TSA-FISH of putative lineage NC10-Ai ('core') Aquavolonida cells from freshwater plankton (Derwent Water, Cumbria, UK). Nucleus = blue (DAPI), NC10-specific TSA-FISH signal = green, flagellar stain = orange. Confocal image focussed in the plane of the green signal. Scale bar = 10 μ m.

Figure 6. Bayesian SSU rRNA gene phylogeny showing the position of Aquavolonida, including new taxa *Aquavolon hoantrani* and *A. dientrani* (bold text), in relation to a broad selection of other rhizarians, including Cercozoa (= Filosa), Tremulida (= NC11), NC12, and Endomyxa, arbitrarily rooted on Radiolaria (shorter-branched Retaria). Bayesian Posterior Probabilities (BPP) and Maximum Likelihood (ML) bootstrap values are indicated on nodes; black filled circles indicate values of BPP \geq 0.95 and ML bootstrap \geq 95%.

SUPPORTING INFORMATION

Table S1. Aquavolonida sequences in GenBank.

Table S2. Sequences of the PCR primers used in this study.

Video S1. *Aquavolon dientrani* swimming. Video can be accessed at <https://yadi.sk/mail/?hash=jsf7XwEScPTOsgvjNRSf5UTcbJ9iAjh6mG6l4cGmALw%3D>

Table 1. Detection of Aquavolonida 18S rRNA phylotypes in various habitat types using two different PCR strategies. Dashes indicate that the primer strategy did not amplify Aquavolonida from that habitat type.

| Habitat type | PCR strategy I - 'core' Aquavolonida | PCR strategy II - general Aquavolonida diversity | |
|--------------------------|--------------------------------------|--|---|
| | positive samples | positive samples | Aquavolonida lineage(s) detected |
| Marine sediment | 0 out of 32 | 0 out of 32 | – |
| Marine plankton | 0 out of 32 | 0 out of 32 | – |
| Estuarine gradient | 1 out of 16 ^a | 0 out of 16 | – |
| Soils | 0 out of 16 ^b | 11 out of 16 | NC10-D (94% of clones) and NC10-B (6% of clones) |
| Freshwater sediment | 8 out of 16 | 14 out of 16 | NC10-A (63% of clones) and NC10-B (37% of clones) |
| Freshwater river biofilm | 7 out of 8 | 7 out of 8 | NC10-A (80% of clones) and NC10-B (20% of clones) |
| Freshwater plankton | 16 out of 16 | 16 out of 16 | NC10-A (67% of clones) and NC10-B (33% of clones) |

^a The only positive sample was from the freshwater end of the estuarine gradient.

^b On soil samples this PCR strategy can generate false positives (amplification of lobose amoebae with the primers used in the first PCR).

Table 2. Sequence signatures observed in the 18S rRNA gene supporting the monophyly of Aquavolonida ord. nov. and its relationship with Tremulida and NC12.

| Helix and positions ^a | Specific sequence motif ^b | Consensus sequence in other eukaryotes |
|-----------------------------------|--|--|
| Helix 11, 229-234 and 248-252 | [R(A) <u>G</u>C<u>Y</u>Y] - [R<u>R</u>RG<u>C</u>[^]Y] | [R(A)UCGC] - [GCGA [^] Y] |
| Between helices 20 & 21, 537-545 | GCGGUA <u>AU</u> A | GCGGUA <u>AU</u> |
| Helix 25, 866-878 and 885-900 | [UCHD <u>N</u>GG(CCA)GAG] - [UUC(U<u>A</u>UGAU)UCNBKGA] | [UYMRYKG(YCA)GAG] - [UUC(UYGGAU)YYRYKRA] |
| Helix 29, 1016-1028 and 1041-1053 | [CGR(GA)UBRGYSVW] - [UBBR <u>Y</u> YDG(CA)Y <u>C</u>G] | [AGG(GA)UYRGNRRR] - [YYNY <u>Y</u> RG(CA)CCU] |
| Helix 45, 1451-1456 and 1506-1511 | [GUUGCD] - [BGCAAC] | [GAYRVR] - [YBYRUC] |
| Helix 48, 1565-1573 and 1579-1587 | [AA(C)GCR <u>Y</u> KU] - [GM <u>R</u> YGC(A)UU] | [AR(C)GCRAGU] - [RCUYGC(R)YU] |

^a Helix numbering according to Wuyts et al. 2000, and positions given using NC10 environmental clone AY2009C14 (accession number HQ219347) as a reference.

^b Bold and underlined: three complementary and two single sequence signatures that together are uniquely found in Aquavolonida ord. nov. and are used as a diagnostic feature of this order; bold: two complementary sequence signatures supporting a clade comprising Aquavolonida, Tremulida, and NC12; underlined: single sequence signature shared by this clade and Endomyxa, but not core Cercozoa (= Filosa).

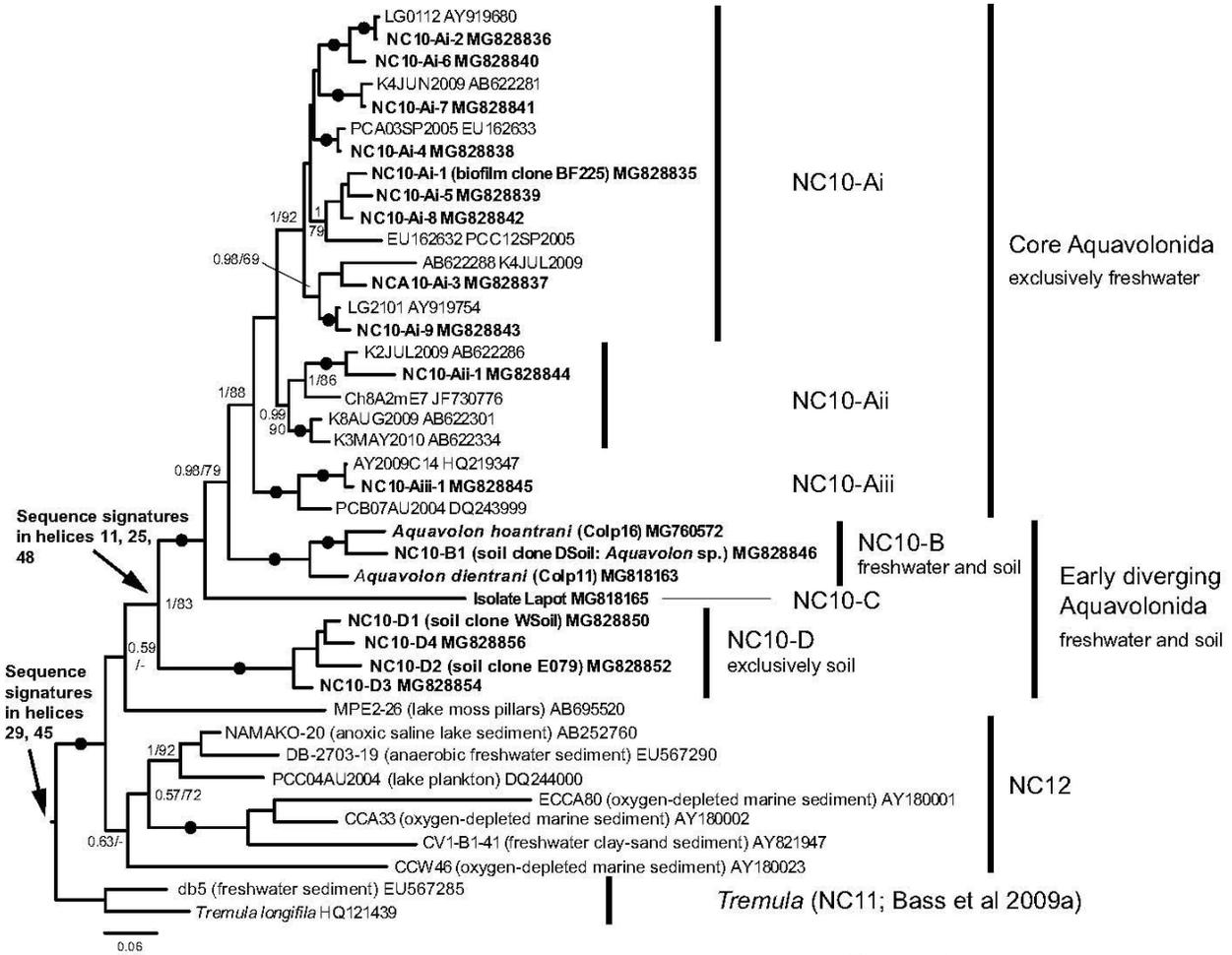
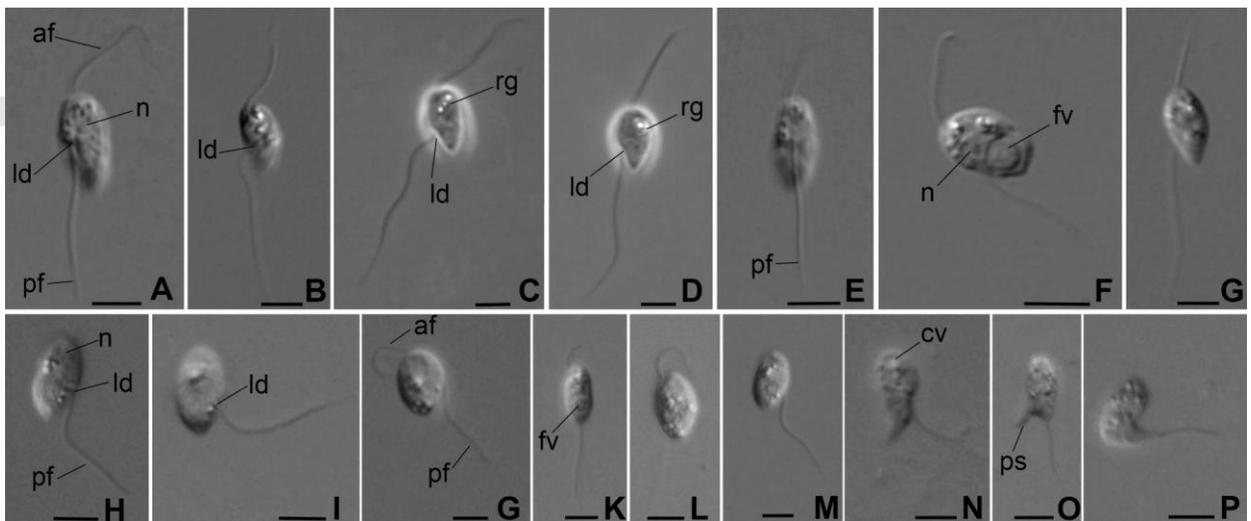
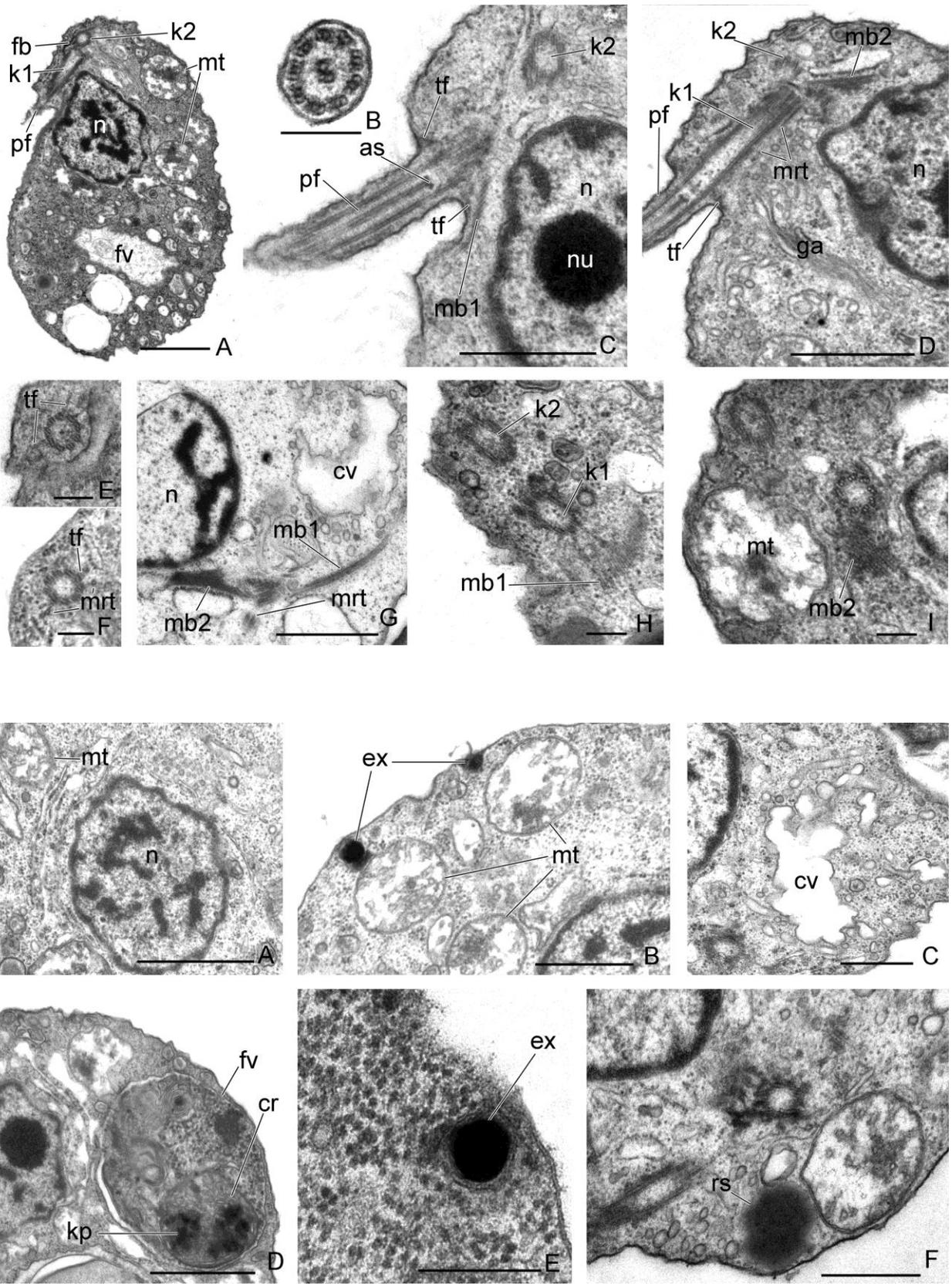


Figure 1





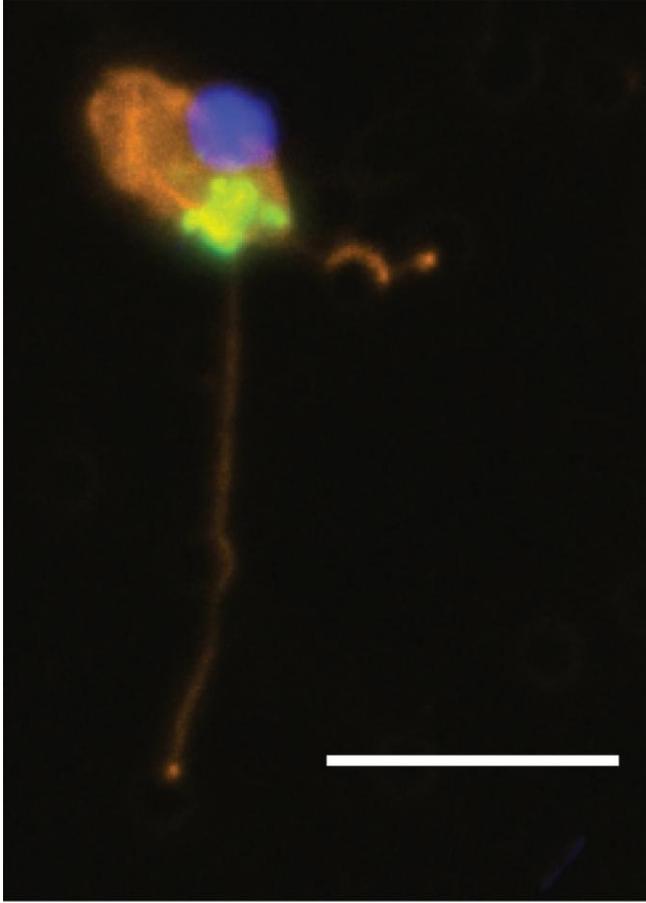


Figure 5

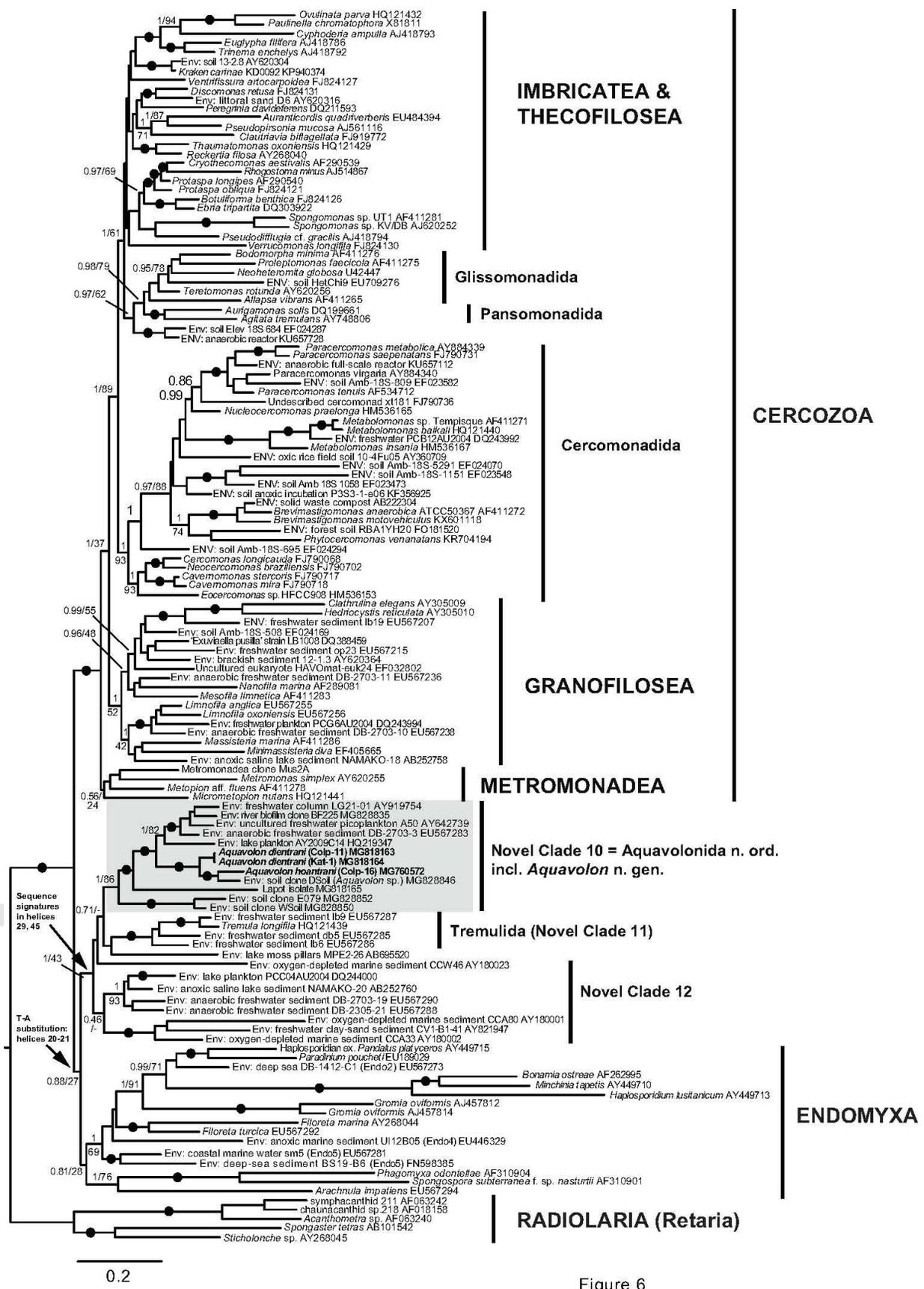


Figure 6