Protist, Vol. 162, 762–773, November 2011
http://www.elsevier.de/protis
Published online date 1 July 2011

Protist

ORIGINAL PAPER

Subulatomonas tetraspora nov. gen. nov. sp. is a Member of a Previously Unrecognized Major Clade of Eukaryotes

Laura A. Katz[a,b,1], Jessica Grant[a], Laura Wegener Parfrey[a,b,2], Anastasia Gant[a], Charles J. O'Kelly[c], O. Roger Anderson[d], Robert E. Molestina[e], and Thomas Nerad[f]

[a]Department of Biological Sciences, Smith College, 44 College Lane, Northampton, Massachusetts 01063, USA
[b]Program in Organismic and Evolutionary Biology, University of Massachusetts, 611 North Pleasant Street, Amherst, Massachusetts 01003, USA
[c]Friday Harbor Laboratories, University of Washington, 620 University Road, Friday Harbor, Washington 98250, USA
[d]Biology, Lamont-Doherty Earth Observatory of Columbia University, Palisades, New York 10964, USA
[e]American Type Culture Collection, Protistology Collection, 10801 University Blvd., Manassas, Virginia 20110, USA
[f]George Mason University, PWII, Manassas, Virginia 10801, USA

Submitted October 22, 2010; Accepted April 18, 2011
Monitoring Editor: Hervé Philippe.

While a large number of aerobic free-living protists have been described within the last decade, the number of new anaerobic or microaerophilic microbial eukaryotic taxa has lagged behind. Here we describe a microaerophilic genus and species of amoeboflagellate isolated from a near-shore marine site off the coast at Plymouth, Massachusetts: *Subulatomonas tetraspora* nov. gen. nov. sp. This taxon is closely related to *Breviata anathema* based on both microscopical features and phylogenetic analyses of sequences of three genes: SSU-rDNA, actin, and alpha-tubulin. However, *Subulatomonas tetraspora* nov. gen. nov. sp. and *B. anathema* are morphologically distinctive, differ by 14.9% at their SSU-rDNA locus, and were isolated from marine and 'slightly brackish' environments, respectively. Phylogenetic analyses of these two taxa plus closely related sequences from environmental surveys provide support for a novel clade of eukaryotes that is distinct from the major clades including the Opisthokonta, Excavata, Amoebozoa and 'SAR' (Stramenopile, Alveolate, Rhizaria).

Key words: Novel eukaryotic lineage; microaerophilic eukaryotes; orphan lineage; *Breviata.*

Introduction

Analyses of molecular data from a growing number of taxa indicate that eukaryotic diversity exists largely within a few major clades, or 'supergroups' (Adl et al. 2005; Burki et al. 2007; Hampl et al. 2009; Keeling et al. 2005; Parfrey et al. 2010; Simpson and Roger 2004). These include the well-supported

[1]Corresponding author; fax +1-413-585-3786
[2]Current address: Department of Chemistry & Biochemistry, University of Colorado, 215 UCB, Boulder, CO 80309, USA
e-mail Lkatz@smith.edu (L.A. Katz).

doi:10.1016/j.protis.2011.05.002

Opisthokonta (animals, fungi and their microbial relatives; Burki et al. 2007; Rodriguez-Ezpeleta et al. 2007; Steenkamp et al. 2006) and several clades that have emerged from molecular analyses – the Amoebozoa (Cavalier-Smith 1996/97), 'SAR' (Stramenopiles, Alveolates and Rhizaria; Burki et al. 2007; Rodriguez-Ezpeleta et al. 2007), and the Excavata, which were originally defined based on an 'excavate' feeding grove in the last common ancestor of the group (Simpson and Patterson 1999). Beyond these major lineages, there are numerous 'orphan' lineages: lineages with no characterized close relatives whose phylogenetic position is often unstable, including *Telonema,* the apusomonads, Centrohelizoa and *Breviata anathema* (e.g. Parfrey et al. 2010; Simpson and Patterson 2006). While phylogenomic analyses have generated hypotheses for the placement of several of these lineages (e.g. Burki et al. 2009; Minge et al. 2009), these relationships should be further tested as these analyses are vulnerable to systematic error due to the low taxon sampling in the clades of interest and large amount of character data (Heath et al. 2008; Parfrey et al. 2010; Zwickl and Hillis 2002). Hence, increasing taxonomic sampling is critical to elaborating the eukaryotic tree of life.

A further factor contributing to our understanding of eukaryotic diversity is the use of culture-independent molecular analyses. Characterization of DNA from environmental samples have provided evidence of previously-unrecognized major clades of eukaryotes at the deepest levels (Dawson and Pace 2002; Takishita et al. 2007a, b, 2010) and within lineages such as alveolates, stramenopiles (Massana and Pedros-Alio 2008), and Foraminifera (Habura et al. 2008). One such clade of eukaryotes identified from environmental sequences was later found to include the enigmatic amoeboflagellate *Breviata anathema* (Walker et al. 2006). Yet a limitation of culture-independent approaches is that they provide only limited details about the nature of biological diversity, often only a single gene sequence.

Here we describe a genus and species of marine amoeboflagellate, *Subulatomonas tetraspora* nov. gen. nov. sp., and assess the phylogenetic position of this taxon using molecular tools. Both morphologic and multigene molecular analyses indicate that *Subulatomonas tetraspora* is closely related to *Breviata anathema*. Thus, we use multigene concatenated analyses to test three hypotheses on the position of this clade: H1) *S+B* (*Subulatomonas* + *Breviata*) branches within the Amoebozoa (Minge et al. 2009); H2) *S+B* is sister to apusomonads (Walker et al. 2006); H3) *S+B* belongs to a novel clade of eukaryotes including SSU-rDNA sequences from culture-independent surveys (Dawson and Pace 2002; Takishita et al. 2007a, b, 2010; Walker et al. 2006).

Results

Light Microscopy

Subulatomonas tetraspora is a small microaerophilic marine amoeboflagellate with a distinctive awl-shaped body and dynamic neck that both extends along a substantial portion of the single flagellum and reappears when the flagellum moves to new location (Fig. 1). The prominent neck of *S. tetraspora* is unique to this taxon, causing the cell body to become awl-shaped when gliding and reappearing quickly as the flagellum moves across the body (Fig. 1; http://www.science.smith.edu/departments/Biology/lkatz/submovies.html). Flagellates appear somewhat elongate with the main cell body 5-10 μm long (mean = 7, STD = 1.2, n = 40) and 3-5 μm wide (mean = 3.5, STD = 0.5, n = 35; Fig. 1). Cell motility occurs by swimming and gliding. Cells have a single visible flagellum [6-12 μm long (mean = 10, STD = 2.5, n = 20)] that emerges from an extension of the body termed 'neck'. The neck of the cell is approximately 6 μm long (Fig. 1; mean = 7, STD = 2, n = 25). During gliding motility small pseudopodia extend from the neck (Fig. 1C and 1D) and there is often a long pseudopod trailing from the posterior of the cell (Fig. 1B-D). When the cells attach to a substrate, long filose pseudopodia emerge from all areas of the body and the flagellum appears to be resorbed (Fig. 1A, C, D). These pseudopodia are up to 30 μm long (mean = 12 μm, STD = 10, n = 25). Cysts are 4-6 μm in diameter and almost always appear in clusters of four. Unfortunately, micrographs of the clusters of four cysts were not available for inclusion in this manuscript because cultivation of *S. tetraspora* became problematic.

Electron Microscopy

Our knowledge of the fine structure of this isolate is limited to the major cellular organelles and the cell surface, which is uncoated. Trophic cells, both flagellates and amoebae, are "naked," lacking a detectable glycocalyx or other cell covering (Fig. 2A). The single nucleus contains a centrally located, spherical nucleolus (Figs 1G, 2B). A small Golgi body is present in the vicinity of the nucleus, but no microtubules were observed in this region (Fig. 2A). Mitochondria were not observed,

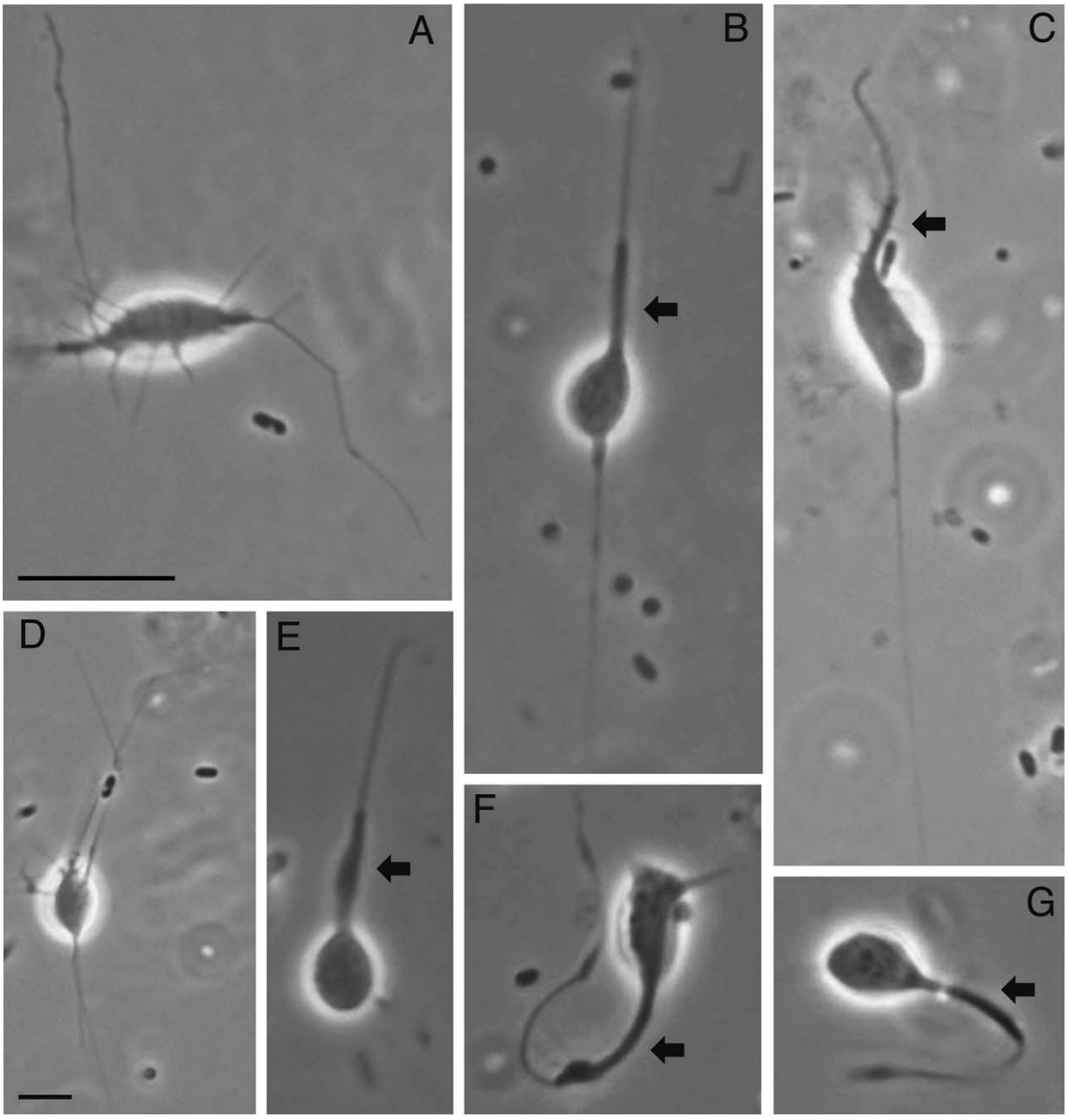

Figure 1. Light micrographs of *Subulatomonas tetraspora*. Arrow points to neck. **A.** amoeboid form with many pseudopodia **B.** gliding form with single anterior flagellum and trailing pseudopodium; **C.** gliding form with pseudopodia extending from the neck; **D.** amoeboid form; **E.** swimming form; **F.** settling form; and **G.** swimming form. Scale bar 10 μm for A-C and E-G, 5 μm for D.

but long, slender, double-membrane-bound bodies with textured contents, possibly hydrogenosomes, are present throughout the cytoplasm and especially in the vicinity of the nucleus (Fig. 2A, B, D). Bacteria are present in food vacuoles and in the cytoplasm (Fig. 2C). Finer details of the uncoated cell surface and one of the putative hydrogenosomes are shown in Figure 2D.

Cytoskeletal elements were not well preserved and further analyses are needed to resolve the fine structure of the flagellar apparatus, particularly the organization of the basal body.

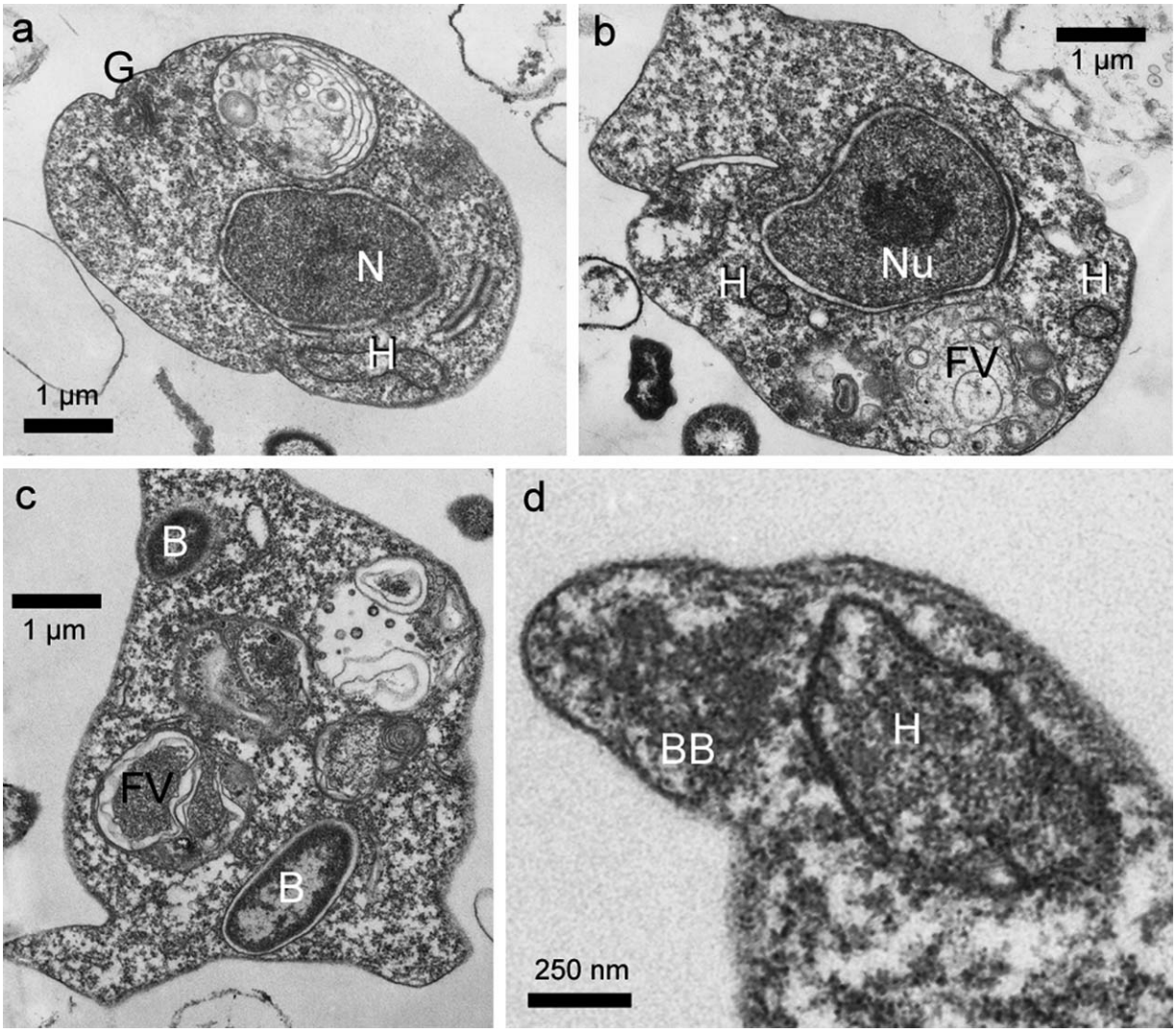

Figure 2. TEM analysis of *S. tetraspora.* **A.** Cross-section of a trophic cell at the level of the nucleus (N), with putative hydrogenosomes (H), and the Golgi body (G). The cell surface is "naked", without visible structures. Scale bar = 1 μm. **B.** Nucleus with centrally-located, spherical nucleolus. Scale bar = 1 μm. **C.** Tangential section of a cell near the anterior end, with food vacuoles (FV) containing bacteria, bacteria (B) outside of food vacuoles, and a putative hydrogenosome (H). Scale bar = 1 μm. **D.** An enlarged image of the uncoated cell surface with a putative hydrogenosome (H). Protruding region is possibly the site of flagellum emergence with a basal body (BB). Scale bar = 250 nm.

Concatenated Analysis

Phylogenetic analysis of the concatenated data (SSU-rDNA plus amino acid sequences from 15 genes) with a total of 6904 characters inferred in RAxML (Fig. 3) places *Subulatomonas tetraspora* sister to the enigmatic taxon *Breviata anathema* with 100% bootstrap support (BS). Although only three of these 16 genes are available for *S. tetraspora*, our previous analyses demonstrate that taxa can be placed accurately onto phylogenies with this level of missing data (Parfrey et al. 2010). The multigene analysis includes 89 broadly sampled lineages that span the eukaryotic tree and represent all major groups (Adl et al. 2005; Patterson 1999) as analyzed in Parfrey et al. (2010). We also include additional isolates of apusomonads and *Ancyromonas*, because these lineages were previously hypothesized to be sister to *Breviata* (Walker et al. 2006). Together, *S. tetraspora* and *B. anathema* fall in a well-supported clade with the apusomonads (89% BS), and *Ancyromonas* is

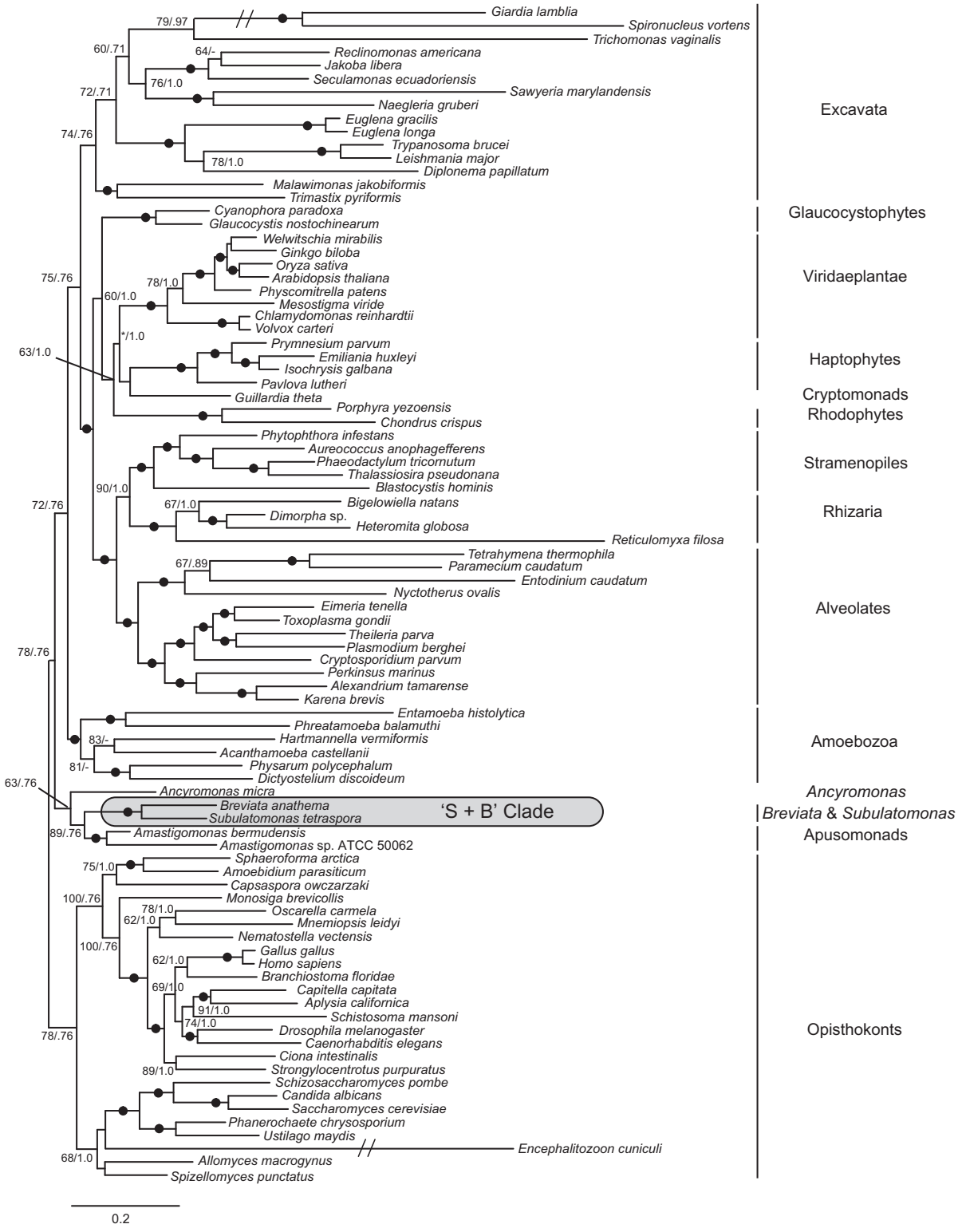

Figure 3. Multigene analysis reveals a close relationship between *Subulatomonas tetraspora* and *Breviata anathema.* Most likely tree from RAxML analysis of 16 genes with 89 diverse eukaryotes. '*S+B'* refers to clade used in hypothesis testing. Support from RAxML and Bayesian analyses are indicated for nodes with greater than 50% support. A * indicates less than 50% support and - indicates node not present in Bayesian analysis. Black dots indicate support values of 95 of greater BS and .95 or greater posterior probabilities.

sister to this clade (63% BS; Fig. 3). The topology here is generally concordant with previously published phylogenies of eukaryotes (e.g. Hampl et al. 2009; Parfrey et al. 2010; Yoon et al. 2008). We recover many of the proposed major clades of eukaryotes in this analysis with moderate to high support: Opisthokonta (78% BS support), Excavata (74% BS support), Amoebozoa (97% BS support), Rhizaria (99% BS support) and the SAR group (100% BS support; Fig. 3). Moreover, most of the groups with ultrastructural identities (e.g. alveolates, stramenopiles, Euglenozoa, haptophytes) are recovered with full support. Bayesian analysis of this dataset yields a topology that is consistent with the likelihood analysis (Fig. 3). Relationships within some 'shallower' nodes, such as animals, are inconsistent with current views as is expected given our limited taxon sampling at this scale.

Breviata anathema has been hypothesized to fall in numerous places across the eukaryotic tree of life, including sister to the Amoebozoa or apusomonads (Minge et al. 2009; Walker et al. 2006). We used the 16 gene dataset to test these and other hypotheses on the phylogenetic position of the *Subulatomonas* + *Breviata* (*S+B*) clade within eukaryotes (Table 1). Specifically, we tested whether the *S+B* clade might form a monophyletic clade with *Ancyromonas* + apusomonads, opisthokonts, or all of these lineages combined; these hypotheses are not rejected by the approximately unbiased test (au), Kishino–Hasegawa test (kh), Shimodaira–Hasegawa test (sh), weighted Kishino–Hasegawa test (wkh), and weighted Shimodaira–Hasegawa test (wsh; Table 1). In contrast, our analyses reject the placement of the *S+B* clade with other major eukaryotic groups (e.g. Amoebozoa, SAR, Excavata, and a clade of photosynthetic lineages; Table 1 – but see sh test for analyses including Amoebozoa for exceptions).

SSU-rDNA Genealogy

To assess further the phylogenetic position of *S. tetraspora*, we performed analyses using the well-sampled SSU-rDNA gene with 485 taxa, 451 that were analyzed previously (Parfrey et al. 2010) plus 34 sequences from taxa and environmental sequences of interest to this study (Fig. 4). The topology of the SSU-rDNA genealogy is generally concordant with previous multigene analyses, and many clades are recovered with high support (e.g. stramenopiles, haptophytes, green algae; Fig. 4). The extensive taxon sampling results in recovery of most of the major clades of eukaryotes, and thus allows the position of *Subulatomonas* to be assessed within the full context of eukaryotic diversity. This single gene analysis provides information at shallow nodes, as deep nodes are weakly supported and rapidly evolving taxa are misplaced due to LBA, which are common for analyses of SSU-rDNA alone. Some deep clades such as Opisthokonta (49% BS) and Rhizaria (37% BS) are recovered but with low support (Fig. 4). Bayesian analyses of the SSU-rDNA data never converged (max diff >.1) after 25 million generations, further demonstrating the limitation of this marker for inferring ancient relationships.

As in the analyses of the concatenated multigene dataset, *S. tetraspora* and *B. anathema* fall together in a fully supported clade, which now includes nine uncultured eukaryotes (Clade I, Fig. 4, Table 2). These uncultured eukaryotes were sampled from a variety of environments that are united in that they are from anoxic or reduced-oxygen marine sediments, with the exception of a historical sample of agricultural soils from the Netherlands (WIM30, Fig. 4, Table 2). In turn, the clade containing *S. tetraspora* and *B. anathema* (Clade I) plus two uncultured eukaryotes, RM1_SGM28 and RM2_SGM71, is weakly sister to another clade of uncultured sequences that were isolated from a deep sea cold seep (Fig. 4). Together, this clade is sister to the apusomonad clade, consistent with the concatenated analyses, though support is low for this clade and for much of the backbone (Fig. 4). Inclusion of a few shorter sequences hypothesized to fall near *B. anathema* (Quaiser et al. 2011) reveal evidence of additional diversity within the *Subulatomonas* + *Breviata* clade, though none of these sequences cluster specifically with either *Subulatomonas* or *Breviata* (grey boxes, Supplementary Fig. S1).

To explore this SSU-rDNA genealogy further, we performed analyses removing the fastest evolving sites as estimated in HyPhy using the topology of the most likely RaxML tree. When the fastest rate class and then the two fastest rate classes were removed, the clade including *S. tetraspora*, *B. anathema* and related environmental sequences remained stable with 98% and 96% BS, respectively (Supplementary Figs S2 and S3). Overall, these trees retained similar structure as compared to the full dataset, though deep nodes collapse.

Taxonomic Summary and Description

Subulatomonas nov. gen.

Small microaerophilic, awl-shaped marine amoeboflagellate distinguished by presence of a dynamic neck surrounding single flagellum that is equal or greater in length than the sub-spherical

Table 1. Topology tests comparing most likely tree to a variety of alternative hypotheses, with rejected hypotheses indicated in bold.

	au	kh	sh	wkh	wsh
(S,B)	0.481	0.401	0.84	0.401	0.803
(S,B,Ancyro,Apuso)	0.655	0.599	0.933	0.599	0.948
(S,B,Ancyro,Apuso,Opisth)	0.498	0.359	0.862	0.359	0.876
(S,B,Opisth)	0.241	0.238	0.704	0.238	0.727
(S,B,Ancyro,Apuso,Amoe)	**0.023**	**0.036**	0.149	**0.021**	0.122
(S,B,Amoe)	**0.02**	**0.039**	0.129	**0.012**	**0.07**
(S,B,SAR)	**0.007**	**0.008**	**0.019**	**0.008**	**0.046**
(S,B,Ancyro,Apuso,SAR)	**0.005**	**0.01**	**0.023**	**0.005**	**0.03**
(S,B,Excavata)	**0.002**	**0.003**	**0.011**	**0.002**	**0.013**
(S,B,Ancyro,Apuso,Excavata)	**0.001**	**0.006**	**0.017**	**0.005**	**0.027**
(S,B,Photo)	**0.001**	**0.003**	**0.006**	**0.001**	**0.008**
(S,B,Ancyro,Apuso,Photo)	**4.00E-04**	**0.002**	**0.004**	**0.001**	**0.005**

S = *Subulatomonas*, B = *Breviata,* Ancyro = *Ancyromonas*, Apuso = Apusomonas, Opisth = Opisthokonta, Amoe = Amoebozoa, SAR = stramenopile+alveolates+Rhizaria, photo = red and green algae, haptophytes, cryptophytes and glaucocystophytes. Topology tests p-values for approximately unbiased test (au), Kishino–Hasegawa test (kh), Shimodaira–Hasegawa test (sh), weighted Kishino–Hasegawa test (wkh), weighted Shimodaira–Hasegawa test (wsh). Results in bold significantly reject monophyly ($p < 0.05$). All analyses based on concatenated dataset and implement in Consel (Shimodaira and Hasegawa 2001).

Table 2. Details of environmental sequences that are closely related to *S. tetraspora* plus other potentially novel clades.

Name	Clade	Environment	Location	GB No.	Ref
WIM30	I	Historic agricultural soils, 0-25 cm below surface	Netherlands	AM114802	1
BOLA187	I	Anoxic marine intertidal sediments, ph7.8	USA	AF372745	2
BOLA366	I	Anoxic marine intertidal sediments, ph7.8	USA	AF372746	2
D4P07D12	I	Oxygen-depleted intertidal marine sediment	Greenland	EF100407	3
D1P01C8	I	Oxygen-depleted intertidal marine sediment	Greenland	EF100391	3
D1P02H06	I	Oxygen-depleted intertidal marine sediment	Greenland	EF100226	3
D3P05C09	I	Oxygen-depleted intertidal marine sediment	Greenland	EF100280	3
SA2_2C8	I	Anoxic waters, Framvaren Fjord	Norway	EF527166	4
SA2_2G5	I	Anoxic waters, Framvaren Fjord	Norway	EF527173	4
RM1-SGM28	I	Microbial mats in cold seep, Sagami Bay	Japan	AB505485	5
RM2-SGM71	I	Microbial mats in cold seep, Sagami Bay	Japan	AB505579	6
NAMAKO-1	II	anoxic sediment from lake Namako-ike	Japan	AB252741	6
RM1-SGM29	II	Microbial mats in cold seep, Sagami Bay	Japan	AB505486	5
RM2-SGM29	II	Microbial mats in cold seep, Sagami Bay	Japan	AB505537	5
CYSGM-24	III	methane cold seep sediment of Sagami Bay	Japan	AB275107	7
RM1-SGM30	III	Microbial mats in cold seep, Sagami Bay	Japan	AB505487	5
RM1-SGM31	III	Microbial mats in cold seep, Sagami Bay	Japan	AB505488	5
RM1-SGM32	III	Microbial mats in cold seep, Sagami Bay	Japan	AB505489	5
NAMAKO-2	IV	anoxic sediment from lake Namako-ike	Japan	AB252742	6
RM1-SGM33	IV	Microbial mats in cold seep, Sagami Bay	Japan	AB505490	5
RM1-SGM37	IV	Microbial mats in cold seep, Sagami Bay	Japan	AB505494	5
RM2-SGM45	IV	Microbial mats in cold seep, Sagami Bay	Japan	AB505553	5
RM2-SGM69	-	Microbial mats in cold seep, Sagami Bay	Japan	AB505577	5

Sampling depths are 35 meters for Framvaren Fjord, 1178m for Sagami Bay and 35m for Namako-ike lake. References: 1 = (Moon-van der Staay et al. 2006), 2 = (Dawson and Pace 2002); 3 = (Stoeck et al. 2007); 4 = (Behnke et al. 2010); 5 = (Takishita et al. 2010); 6 = (Takishita et al. 2007b); 7 = (Takishita et al. 2007a).

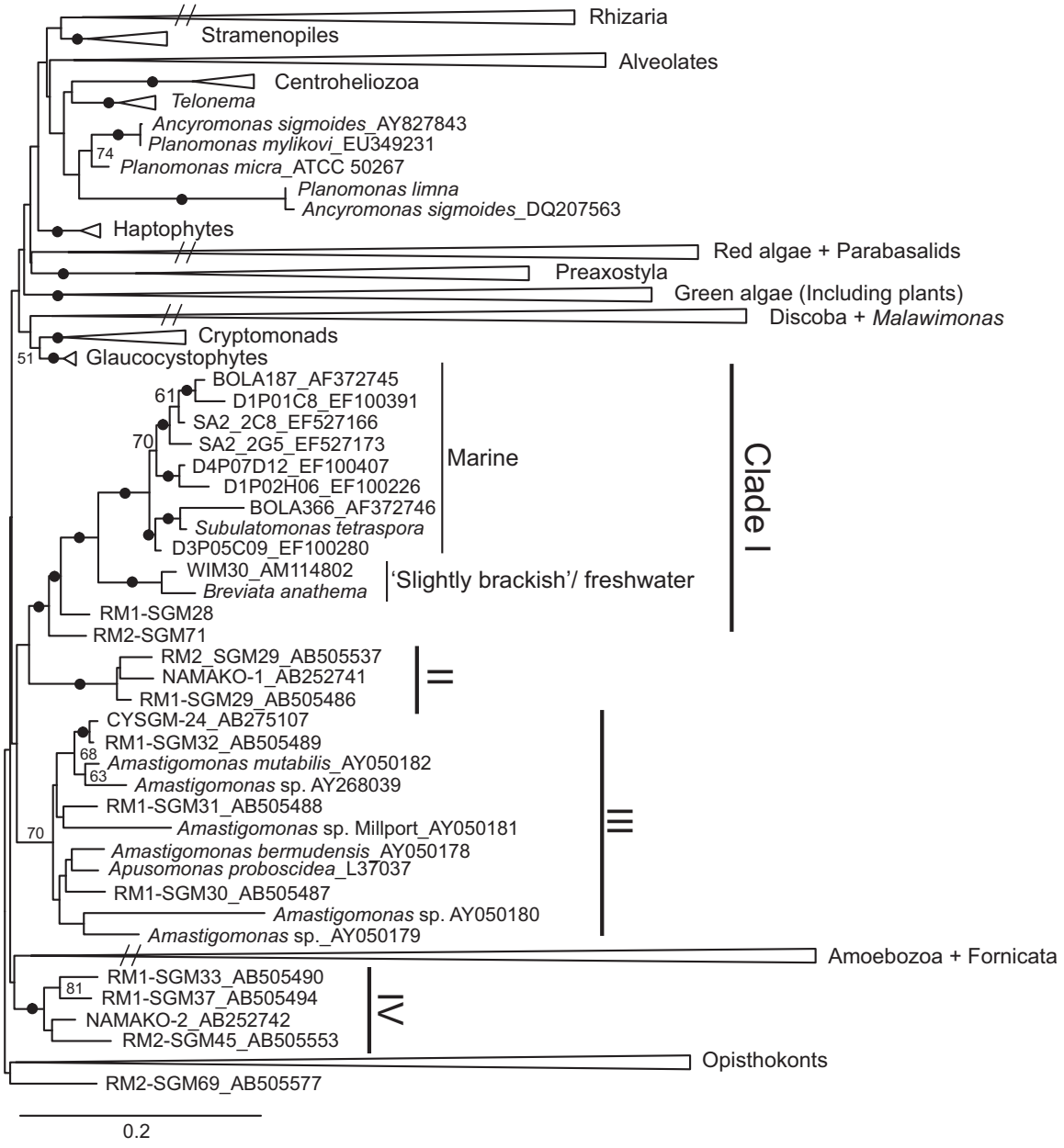

Figure 4. Most likely SSU-rDNA tree constructed with 485 broadly sampled eukaryotes as in Parfrey et al. (2010) and 1205 positions demonstrates that *Subulatomonas tetraspora* and *Breviata anathema* constitute a novel clade of eukaryotes along with environmental SSU-rDNA sequences. Within this clade, isolates are indicated as being either from marine environments or 'slightly brackish' to freshwater environments, as discussed in the text. Support values from RAxML analyses are shown for nodes with greater than 50%. Black dots indicate support greater than 85% BS. Nodes of major clades are collapsed so that the clade containing *Subulatomonas tetraspora* and *Breviata anathema* can be clearly seen. Clade numbers I-IV as in Table 2. See Figure S1 for a tree including recent partial sequences from Quaiser et al. (2011).

portion of the cell. Cells are uninucleate, bacterivorous and microaerophilic amoeboflagellates. Pseudopodia are present when cells are gliding and become abundant when cells are in contact with substrate, taking the form of branched filopodia that arise from all parts of the cell. Mitochondria absent, but hydrogenosome-like organelles present. Cysts are rounded, without obvious ornamentation or pore structures. Differentiated from *B. anathema* by the dynamic neck, the primary DNA sequences of several "marker" genes, and isolation from marine environment.

Type species: Subulatomonas tetraspora nov. sp.

Subulatomonas tetraspora nov. sp.

Marine species with characteristics of the genus. Length of flagellates 10-19 μm from the tip of the neck to the posterior end of the body; dimensions of sub-spherical portion of cell 5-10 μm long, 3-5 μm wide; neck 5-9 μm long, 0.7 μm wide. Flagella are 9-10 μm long. Fine extensive pseudopodia originate from all areas of the cell, particularly when the cells settle. When swimming cells change directions, the flagellum moves to a different point on the cell body where the neck is newly formed. Produces rounded cysts that lack ornamentation and typically appear in groups of four.

Habitat: The species was isolated from a sediment sample taken just off shore near Plymouth, Massachusetts (approximate coordinates: 41°58′30″N; 70°39′18″W).

Reference Material: The culture is maintained in the cryopreserved state in liquid nitrogen at the American Type Culture Collection (ATCC®) and has been accessioned as *Subulatomonas tetraspora* ATCC® 50623™.

Etymology: The genus name *Subulatomonas* is derived from the Latin *subula* meaning "awl" based on the awl-shaped morphology of the free-swimming cells. The species name, *tetraspora* refers to the tendency of the cysts to form in clusters of four.

Discussion

The genus described herein is a new microaerophilic marine amoeboflagellate that is closely related to *B. anathema* based on both morphological and molecular analyses. Although the morphology of the two genera is similar in some respects after a prolonged time under a coverslip, the morphology of *S. tetraspora* in freshly prepared slides is clearly distinct from that of *B. anathema*. The elongate and uniformly thick neck never appears in cells of *B. anathema*. In addition, the morphology of cysts in *S. tetraspora* is distinctive as this taxon forms clusters of four virtually uniformly spherical cysts. An additional distinguishing feature of these taxa is that *S. tetraspora* (and its uncultured relatives; Fig. 4; Table 2) was isolated from a marine environment while deposition notes at ATCC indicate that *B. anathema* was isolated from a 'slightly brackish' pond, Warwick Pond, in Bermuda and then cultured in freshwater media (Walker et al. 2006). While no salinity measurements were provided for *B. anathema* on isolation, Warwick Pond is described variably as freshwater and brackish (salinity ranges from 1-4 ppt in most months but can spike as high as 13 ppt in the driest months – personal communication, Dr. Jamie Bacon, Bermuda Zoological Society); significantly, the only environmental sequence to cluster specifically with *B. anathema* was isolated from agricultural soils in the Netherlands (Table 2), indicating that this clade is distinct from *S. tetraspora* in that its ecology is fresh to brackish water.

Multigene phylogenetic analyses place *S. tetraspora* sister to *B. anathema* (Fig. 3), though the two genera are divergent for molecular markers (e.g. 14.9% for SSU-rDNA) and cluster in separate clades in analyses of SSU-rDNA sequences with environmental sequences included (Fig. 4). Though there is no simple relationship between genetic distance and taxonomy, some comparisons of SSU-rDNA divergence provide a frame of reference: the distance between *S. tetraspora* and *B. anathema* is greater than the 12.2% between *Ancyromonas sigmoides* (GB#AY827844) and *Apusomonas proboscidea* (GB#L37037) and the 13.0% divergence between *Arabidopsis thaliana* (GB#NR_022452) and *Chlamydomonas reinhardtii* (GB# M32703).

Together *S. tetraspora* and *B. anathema* do not fall within any well-circumscribed clade of eukaryotes (e.g. Opisthokonta, SAR, Amoebozoa). *Breviata anathema* is an enigmatic taxon, originally misidentified as "*Mastigamoeba invertens",* that has been suggested to be a member of the Amoebozoa (Minge et al. 2009), sister to the Apusomonads (Walker et al. 2006), or an orphan lineage (Parfrey et al. 2010; Simpson and Patterson 2006). Our topology tests reject the placement of the *Subulatomonas* + *Breviata* (*S* + *B*) clade within the Amoebozoa and most other clades of eukaryotes (Table 1). However a sister relationship of *S* + *B* with apusomonads and or Opisthokonta is possible (Table 1).

Together, these data support the conclusion that *B. anathema, S. tetraspora,* plus some environmental sequences represent a novel clade of eukaryotes. Moreover, it is clear from recent sequence-based investigations of low-oxygen environments that the diversity of organisms in the clade containing the genera *Breviata* and *Subulatomonas* is greater than previously appreciated. To what extent this diversity represents variable body plans or is cryptic – that is, reflecting species with similar morphologies but different genetic signatures – cannot be assessed until additional close relatives have been sampled. This may require a focus on getting anaerobic/microaerophilc taxa into culture for detailed examination.

Methods

Isolation: A sample of sediment was taken in a Ziplock™ bag from a near-shore marine site in Plymouth, Massachusetts (approximate coordinates: 41°58′30″N; 70°39′18″W). The bag was pressed to expel any air and then sealed and transported back to the laboratory. Once at the laboratory, small portions of sediment were transferred to 16x125 mm borosilicate glass test tubes containing ATCC® medium 1873 (a seawater medium) bacterized with *Enterobacter aerogenes* (ATCC® 13048™). The tubes were filled to 2.5 cm of the top and incubated at 25 °C on a 15 degree horizontal slant with the caps screwed on tightly. Once a thriving culture was established, a small aliquot was added to the edge of a plate of non-nutrient marine agar ATCC® medium 994 that had been overlaid with a thin layer of ATCC® medium 1873 bacterized at least 24 hours previously with *Klebsiella pneumoniae subsp. pneumoniae* ATCC® 700831. The cells that had been inoculated at the edge of the plate were allowed to migrate away outward toward the opposite end of the agar plate. The agar plate was placed on an inverted microscope and single cells were picked with a finely drawn out Pasteur pipette. The single cells were then transferred to dram vials with bacterized ATCC® medium 1873. The vials were filled to within 1 cm of the lip of the vial and incubated horizontally on a 15 degree angle with the caps screwed on tightly at 25 °C. It was later found that the clonal cultures grew better in 16x125 mm borosilicate glass test tubes containing 15 ml of a 1:25 mixture of ATCC® medium 177 and ATCC® medium 1525.

Cultivation: Cultures of *Subulatomonas tetraspora* ATCC® 50623™ were maintained at ~25ºC in 16x125 mm borosilicate glass test tubes containing 15 ml of either ATCC® medium 1873 bacterized with *Klebsiella pneumoniae* subsp. *pneumoniae* ATCC® 700831 or a 1:25 mixture of ATCC® medium 177 and ATCC® medium 1525 bacterized with *Enterobacter aerogenes* (ATCC® 13048™). The cultures were sub-cultured biweekly.

Light microscopy: Light microscopy of live cells was performed using a Zeiss Axioskop compound microscope equipped with a Zeiss AxioCam digital camera. Photodocumentation and measurements of *Subulatomonas tetraspora* cells were performed using Zeiss Axiovision 4.6 software.

Transmission electron microscopy: Cells of *Subulatomonas* were fixed, sectioned, stained, and examined using the same procedures as those described for the microaerophilous amoebae *Monopylocystis* and *Sawyeria* (O'Kelly et al. 2003). Briefly, exponential-phase trophic cells (flagellates and amoebae) were fixed in a cacodylate-buffered glutaraldehyde-osmium tetroxide cocktail for 10 min at room temperature (ca. 25 °C). Fixed cells were collected on cellulose acetate/nitrate filters, washed in distilled water, enrobed in agar, dehydrated in a graded acetone series and embedded in epoxy resin (Eponbed 812-Araldite). Blocks were trimmed by hand and sectioned with a diamond knife. Sections were mounted on uncoated 600-mesh copper grids, stained with uranyl acetate and lead citrate, and examined with a Zeiss EM 902 transmission electron microscope operating at 80 kV.

Molecular data: DNA samples of *Subulatomonas tetraspora* ATCC® 50623™, *Amastigomonas* sp. ATCC® 50062™, and *Amastigomonas bermudensis* ATCC® 50234™ were extracted using DNA Stat60™ (Tel- Test, Inc., Friendswood, Texas; Cat. No. TL-4220) per manufacture's instructions and with the addition of a phenol-chloroform-isoamyl step using Phase Lock Gel Heavy tubes (Eppendorf AG, Hamburg, Germany; Cat. No. 955154070).

We characterized the SSU-rDNA, actin and alpha-tubulin genes from *S. tetraspora, A. bermudensis and Amastigomonas* sp. (GenBank accession numbers HQ342676-HQ342683). Phusion DNA Polymerase (NEB# F-540L) was used to amplify these genes and resulting products were cloned in Invitrogen Zero Blunt Topo cloning kits. Primers for SSU-rDNA genes are from Medlin et al. (1988) and primers and protocols for actin and alpha-tubulin are from Tekle et al. (2007). Sequencing was performed with the BigDye terminator kit (Perkin Elmer) and sequences were run on an ABI 3100 automated sequencer. We fully sequenced 2-4 clones and surveyed up to 10 clones per gene to look for paralogs, though we found none.

Alignments: Inferred protein coding genes from this study were added to a previously published 15 gene amino acid alignment (Parfrey et al. 2010) and adjusted by eye in MacClade 4.08 (Maddison and Maddison 2005). The SSU-rDNA alignment was concatenated to this alignment and taxa containing data for ten or more of the genes were retained, creating an 89-taxon 6904 character matrix (Treebase #11326). Our analyses excluded ambiguously aligned regions.

Analysis of SSU-rDNA based on a large alignment that had taxa chosen to maximize taxonomic diversity across eukaryotes (Parfrey et al. 2010) with additional sequences added to test hypotheses on the phylogenetic position of *Subulatomonas* (Treebase #11326). To this alignment, we added sequences for all available *Ancyromonas* (syn. *Planomonas*; Heiss et al. 2010)) and apusomonads that were ≥1% divergent from one another. We also added environmental SSU-rDNA sequences from GenBank that were either inferred to be closely related to *S. tetraspora* by Blast analysis and/or through a search of the literature (Dawson and Pace 2002; Takishita et al. 2007a, b, 2010). Blast analyses were iterative. First we used sequences from key morphospecies to obtain a list of the top 30-40 uncultured sequences. These sequences were then added to our analyses and those that fell outside of our clades of interest were removed. Uncultured sequences were downloaded from GenBank in October 2009, with additional shorter SSU-rDNAs downloaded in October 2010 from a recent survey (Quaiser et al. 2011).

All sequences were aligned using ClustalW (Thompson et al. 1994) and then added to the unmasked SSU-rDNA alignment from Parfrey et al. (2010). The alignment was further edited and ambiguous characters masked manually in MacClade v4.08 (Maddison and Maddison 2005), resulting in a 485-taxon alignment with 1205 characters. Distances were calculated by first assessing the best model fit with datamonkey (http://www.datamonkey.org/) and then calculating distances under this model (TRN; (Tamura and Nei 1993) with DNAdist software (Felsenstein 1993) available at http://mobyle.pasteur.fr/cgi-bin/portal.py?#forms::dnadist).

To assess the effect of rate heterogeneity on the SSU-rDNA topologies, we partitioned the SSU-rDNA data matrix into eight rate classes using the GTR+I and rate variation among sites following a discrete gamma distribution, as implemented in HyPhy version .99b package (Kosakovsky Pond et al. 2005). Rate heterogeneity was computed on the topology shown in Figure 4. We then removed the fastest rate class, resulting in a 1134 character matrix and the two fastest rate classes, resulting in a 1019 character matrix, and analyzed these data sets with RAxML 7.0.4 (Stamatakis 2006; Stamatakis et al. 2008).

Phylogenetic analyses: Genealogies were constructed in RAxML and MrBayes. The MPI version of RAxML with rapid bootstrapping was used, with the GTRGAMMA model for nucleotide data and PROTGAMMA with matrix rtREV + F for amino acid data. A total of 200 rapid bootstrap iterations were used, followed by a maximum likelihood search as

implemented in RAxML 7.0.4. ProtTest (Abascal et al. 2005) was used to select the appropriate model of sequence evolution for the amino acid data using the concatenated 15 gene dataset. Bayesian analyses were performed with the parallel version of MrBayes 3.1.2 using the GTR+I+ Γ (nucleotide) and rtREV (amino acid) models of sequence evolution (Ronquist and Huelsenbeck 2003). Six simultaneous MCMCMC chains were run for 4,000,000 generations sampling every 1000 generations. For the concatenated dataset, the 50% majority-rule consensus tree was determined to calculate posterior probabilities for each node.

Topology testing: To test alternative placements of *S. tetraspora,* we performed a variety of tests including approximately unbiased (AU) test (Shimodaira 2002) as well as the more conventional Kishino-Hasegawa (KH) and Shimodaira-Hasegawa (SH) tests, as implemented in Consel 0.1j (Shimodaira and Hasegawa 2001). The most likely trees with these groups constrained to be monophyletic were built, and the site likelihood values for each constrained topology and the unconstrained topology were estimated using RAxML (Table 1).

Acknowledgements

Thanks to Daniel Lahr (Smith College/UMass Amherst) for helpful discussions on this manuscript. This is Lamont-Doherty Earth Observatory Contribution number 7475. The research was supported by an NSF Assembling the Tree of Life (043115) award to LAK and the ATCC, and an NSF RUI Systematics (0919152) award to LAK.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.protis.2011.05.002.

References

Abascal F, Zardoya R, Posada D (2005) ProtTest: selection of best-fit models of protein evolution. Bioinformatics **21**:2104–2105

Adl SM, Simpson AGB, Farmer MA, Andersen RA, Anderson OR, Barta JR, Bowser SS, Brugerolle G, Fensome RA, Fredericq S, James TY, Karpov S, Kugrens P, Krug J, Lane CE, Lewis LA, Lodge J, Lynn DH, Mann DG, McCourt RM, Mendoza L, Moestrup Ø, Mozley-Standridge SE, Nerad TA, Shearer CA, Smirnov AV, Spiegel FW, Taylor M (2005) The new higher level classification of eukaryotes with emphasis on the taxonomy of protists. J Eukaryot Microbiol **52**:399–451

Behnke A, Barger KJ, Bunge J, Stoeck T (2010) Spatio-temporal variations in protistan communities along an O-2/H2S gradient in the anoxic Framvaren Fjord (Norway). FEMS Microbiol Ecol **72**:89–102

Burki F, Shalchian-Tabrizi K, Minge M, Skjaeveland A, Nikolaev SI, Jakobsen KS, Pawlowski J (2007) Phylogenomics reshuffles the eukaryotic supergroups. PLOS ONE **2**:e790

Burki F, Inagaki Y, Brate J, Archibald JM, Keeling PJ, Cavalier-Smith T, Sakaguchi M, Hashimoto T, Horak A, Kumar S, Klaveness D, Jakobsen KS, Pawlowski J, Shalchian-Tabrizi K (2009) Large-scale phylogenomic analyses reveal that two enigmatic protist lineages, telonemia and centroheliozoa, are related to photosynthetic chromalveolates. Genome Biol Evol **1**:231–238

Cavalier-Smith T (1996/97) Amoeboflagellates and mitochondrial cristae in eukaryote evolution: Megasystematics of the new protozoan subkingdoms Eozoa and Neozoa. Arch Protistenkd **147**:237–258.

Dawson SC, Pace NR (2002) Novel kingdom-level eukaryotic diversity in anoxic environments. Proc Natl Acad Sci USA **99**:8324–8329

Felsenstein J (1993) PHYLIP (Phylogeny Inference Package). Distributed by the author. Department of Genetics, University of Washington, Seattle

Habura A, Goldstein ST, Broderick S, Bowser SS (2008) A bush, not a tree: The extraordinary diversity of cold-water basal foraminiferans extends to warm-water environments. Limnol Oceanogr **53**:1339–1351

Hampl V, Hug L, Leigh JW, Dacks JB, Lang BF, Simpson AGB, Roger AJ (2009) Phylogenomic analyses support the monophyly of Excavata and resolve relationships among eukaryotic "supergroups". Proc Natl Acad Sci USA **106**:3859–3864

Heath TA, Hedtke SM, Hillis DM (2008) Taxon sampling and the accuracy of phylogenetic analyses. J Syst Evol **46**:239–257

Heiss AA, Walker G, Simpson AGB (2010) Clarifying the taxonomic identity of a phylogenetically important group of eukaryotes: *Planomonas* is a junior synonym of *Ancyromonas*. J Eukaryot Microbiol **57**:285–293

Keeling PJ, Burger G, Durnford DG, Lang BF, Lee RW, Pearlman RE, Roger AJ, Gray MW (2005) The tree of eukaryotes. Trends Ecol Evol **20**:670–676

Kosakovsky Pond SL, Frost SDW, Muse SV (2005) HyPhy: hypothesis testing using phylogenies. Bioinformatics **21**:676–679

Maddison DR, Maddison WP (2005) MacClade version 4. 08: An Analysis of Phylogeny and Character Evolution. Sinauer Associates, Sunderland, MA

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

Medlin L, Elwood HJ, Stickel S, Sogin ML (1988) The characterization of enzymatically amplified eukaryotes 16S-like ribosomal RNA coding regions. Gene **71**:491–500

Minge MA, Silberman JD, Orr RJS, Cavalier-Smith T, Shalchian-Tabrizi K, Burki F, Skjaeveland A, Jakobsen KS (2009) Evolutionary position of breviate amoebae and the primary eukaryote divergence. Proc R Soc Lond B Biol Sci **276**:597–604

Moon-van der Staay SY, Tzeneva VA, van der Staay GWM, de Vos WM, Smidt H, Hackstein JHP (2006) Eukaryotic diversity in historical soil samples. FEMS Microbiol Ecol **57**:420–428

O'Kelly CJ, Silberman JD, Zettler LAA, Nerad TA, Sogin ML (2003) *Monopylocystis visvesvarai* n. gen., n. sp and *Sawye-*

ria marylandensis n. gen., n. sp.: Two new amitochondrial heterolobosean amoebae from anoxic environments. Protist **154**:281–290

Parfrey LW, Grant J, Tekle YI, Lasek-Nesselquist E, Morrison HG, Sogin ML, Patterson DJ, Katz LA (2010) Broadly sampled multigene analyses yield a well-resolved eukaryotic tree of life. Syst Biol **59**:518–533

Patterson DJ (1999) The diversity of eukaryotes. Am Naturalist **154**:S96–S124

Quaiser A, Zivanovic Y, Moreira D, Lopez-Garcia P (2011) Comparative metagenomics of bathypelagic plankton and bottom sediment from the Sea of Marmara. ISME J **5**:285–304

Rodriguez-Ezpeleta N, Brinkmann H, Burger G, Roger AJ, Gray MW, Philippe H, Lang BF (2007) Toward resolving the eukaryotic tree: The phylogenetic positions of jakobids and cercozoans. Curr Biol **17**:1420–1425

Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics **19**:1572–1574

Shimodaira H, Hasegawa M (2001) CONSEL: for assessing the confidence of phylogenetic tree selection. Bioinformatics **17**:1246–1247

Simpson AGB, Patterson DJ (1999) The ultrastructure of *Carpediemonas membranifera* (Eukaryota) with reference to the "Excavate hypothesis". Eur J Protistol **35**:353–370

Simpson AGB, Roger AJ (2004) The real 'kingdoms' of eukaryotes. Curr Biol **14**:R693–R696

Simpson AGB, Patterson DJ (2006) Current Perspectives on High-level Groupings of Protists. In Katz LA, Bhattacharya D (eds) Genomics and Evolution of Microbial Eukaryotes. Oxford University Press, Oxford, pp 7–30

Stamatakis A (2006) RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics **22**:2688–2690

Stamatakis A, Hoover P, Rougemont J (2008) A rapid bootstrap algorithm for the RAxML Web-Servers. Syst Biol **57**:758–771

Steenkamp ET, Wright J, Baldauf SL (2006) The protistan origins of animals and fungi. Mol Biol Evol **23**:93–106

Stoeck T, Kasper J, Bunge J, Leslin C, Ilyin V, Epstein S (2007) Protistan diversity in the Arctic: a case of paleoclimate shaping modern biodiversity? PLos One **2**:, Article No.: e728

Takishita K, Kakizoe N, Yoshida T, Maruyama T (2010) Molecular evidence that phylogenetically diverged ciliates are active in microbial mats of deep-sea cold-seep sediment. J Eukaryot Microbiol **57**:76–86

Takishita K, Yubuki N, Kakizoe N, Inagaki Y, Maruyama T (2007a) Diversity of microbial eukaryotes in sediment at a deep-sea methane cold seep: surveys of ribosomal DNA libraries from raw sediment samples and two enrichment cultures. Extremophiles **11**:563–576

Takishita K, Tsuchiya M, KawatoM, Oguri K, Kitazato H, Maruyama T (2007b) Genetic diversity of microbial eukaryotes in anoxic sediment of the saline meromictic lake Namako-ike (Japan): On the detection of anaerobic or anoxic-tolerant lineages of eukaryotes. Protist **158**:51–64

Tamura K, Nei M (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. Mol Biol Evol **10**:512–526

Tekle YI, Grant J, Cole JC, Nerad TA, Anderson OR, Patterson DJ, Katz LA (2007) A multigene analysis of *Corallomyxa tenera* sp. nov. suggests its membership in a clade that includes *Gromia*, Haplosporidia and Foraminifera. Protist **158**:457–472

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

Walker G, Dacks JB, Embley TM (2006) Ultrastructural description of *Breviata anathema*, n. Gen., n. Sp., the organism previously studied as "*Mastigamoeba invertens*". J Eukaryot Microbiol **53**:65–78

Yoon HS, Grant J, Tekle YI, Wu M, Chaon BC, Cole JC, Logsdon JM, Patterson DJ, Bhattacharya D, Katz LA (2008) Broadly sampled multigene trees of eukaryotes. BMC Evol Biol **8**:, doi:10.1186/1471-2148-1188-1114

Zwickl DJ, Hillis DM (2002) Increased taxon sampling greatly reduces phylogenetic error. Syst Biol **51**:588–598