Available online at www.sciencedirect.com
SCIENCE DIRECT®
European Journal of Protistology 40 (2004) 119–135
European Journal of PROTISTOLOGY
www.elsevier.de/ejop

Phylogeny of *Heteromita*, *Cercomonas* and *Thaumatomonas* based on SSU rDNA sequences, including the description of *Neocercomonas jutlandica* sp. nov., gen. nov.

Flemming Ekelund[a,*], Niels Daugbjerg[b], Line Fredslund[a]

[a] *Department of Terrestrial Ecology, Biological Institute, University of Copenhagen, Universitetsparken 15, Copenhagen DK-2100, Denmark*
[b] *Department of Phycology, Biological Institute, University of Copenhagen, Øster Farimagsgade 2D, Copenhagen K DK-1353, Denmark*

Received 6 November 2003; accepted 11 December 2003

Abstract

We present six new small subunit (SSU) ribosomal DNA sequences from the heterotrophic flagellate genera *Cercomonas*, *Heteromita*, and *Thaumatomonas*. We aligned the sequences against 42 SSU rDNA sequences from GenBank, including organisms which previously have been demonstrated to group with *Cercomonas*, *Heteromita*, and *Thaumatomonas* and some additional outgroup taxa. The ribosomal secondary structure was taken into account when aligning the sequences. To compensate for unequal evolutionary rates, we subjected 1959 homologue positions (including gaps) to the process of "substitution rate calibration", resulting in a reduced dataset of evolutionarily informative positions (979 bp). We analysed the data by three different algorithms: neighbour-joining, parsimony, and maximum likelihood. The three algorithms all supported monophyly of a clade consisting of *Cercomonas*, *Heteromita*, and *Thaumatomonas*, chlorarachniophytes, the haplosporids, the plasmodiophorids, the euglyphids, *Massisteria and Cryothecomonas.* Likewise, the three algorithms strongly supported the separation of *Cercomonas* into two clades. Hence we suggest that some of the members of the present genus *Cercomonas* should be placed in a new genus *Neocercomonas* gen. nov.

Keywords: *Cercomonas*; *Heteromita*; *Neocercomonas* gen. nov.; Phylogeny; SSU rDNA sequences; *Thaumatomonas*

Introduction

Heterotrophic flagellates in the genera *Cercomonas* and *Heteromita* are extremely common in soil environments, where they often make up more than 50% of the total protozoan biomass (Sandon, 1927; Foissner, 1991; Ekelund and Patterson, 1997; Ekelund et al., 2001). However, they are often ignored both in ecological research and in textbooks, probably because they are small and without conspicuous and distinctive specific characters. The application of various molecular methods, e.g. specific probes or in situ PCR, would be a good alternative for qualitative as well as quantitative studies of *Cercomonas* and *Heteromita* in environmental samples. Still, before we can apply molecular tools on environmental samples, we need basic knowledge of their genetic and morphological diversity; and, in order to design probes and primers for specific taxonomic groupings, we need information about their phylogeny as well.

Previous SSU rDNA analyses have suggested that *Cercomonas* and *Heteromita* are phylogenetically closely related to the heterotrophic flagellate genus *Thaumatomonas* and to the euglyphid testate amoebae

*Corresponding author. Tel.: +45-3532-1275; fax: +45-3532-1250.
E-mail address: fekelund@zi.ku.dk (F. Ekelund).

0932-4739/$ - see front matter © 2004 Elsevier GmbH. All rights reserved.
doi:10.1016/j.ejop.2003.12.002

(Cavalier-Smith and Chao, 1996/97; Wylezich et al., 2002). All these organisms seem related to the chlorarachniophyte algae and the intracellular phytoparasitic plasmodiophorids (McFadden et al., 1994; Bhattacharya et al., 1995; Cavalier-Smith and Chao, 1996/97; Keeling et al., 1998). Cavalier-Smith (1996/97, 1998b) suggested the name "Cercozoa" for this whole assemblage of organisms. Based on shared morphological characters, it has been suggested that various additional flagellates and amoebae belong to this assemblage (Cavalier-Smith, 1998b, 1999; Patterson, 1999). The "Cercozoa" as a taxon is ill defined in terms of morphological synapomorphies, and further studies are needed to confirm whether this assemblage has any taxonomic relevance.

Gene sequence comparisons of nuclear-encoded SSU rDNA are commonly used to infer higher level eukaryote phylogenetic relationships. SSU rDNAs are probably evolutionarily homologous and functionally equivalent in all organisms. They are not transferred laterally between organisms; they do not recombine, and are thus widely accepted as consistent with the evolutionary history of the organisms concerned (Sogin and Hinkle, 1997). Sequences of conserved and variable positions are inter-spaced in the SSU rDNA, which allows analysis of both closely related and evolutionarily distant taxa on an extended evolutionary time scale (Sogin, 1989).

However, pair-wise distance methods of inferring relationships, such as neighbour-joining (Saitou and Nei, 1987) are heavily biased if they are not corrected for differences in substitution rates among sites (Yang and Roberts, 1995). This is only possible if the variable areas can be clearly distinguished from the conserved areas and if quantitative information on the relative rate of fixation of substitutions is available for the areas of different variability (Van de Peer et al., 1993b). For example, Hillis et al. (1994) demonstrated that unequal substitution rates diminish the ability of most phylogenetic reconstruction methods to recover the correct tree. Van de Peer and colleagues developed a method for measuring the relative substitution rate of individual sites in a nucleotide sequence alignment on the basis of a distance approach (Van de Peer et al., 1993b, 1996a, b). The equation that describes the actual relationship between sequence dissimilarities and evolutionary distances is similar to the general formula for computing gamma distances as proposed by Rzhetsky and Nei (1994) (Van de Peer et al., 1997). The main advantage of the "substitution rate calibration" approach, according to its developers, is that nucleotide variability estimates do not depend upon a given tree topology, contrary to estimates inferred from maximum parsimony or likelihood methods (Sullivan et al., 1996; Yang, 1996).

Molecular data on the organisms treated here are still relatively sparse, although papers treating part of the group have been published recently (Burki et al., 2002; Vickerman et al., 2002; Wylezich et al., 2002; Cavalier-Smith and Chao, 2003, see later). In this paper we present six new rDNA sequences of flagellates from *Cercomonas*, *Heteromita*, *Thaumatomonas* and *Neocercomonas* gen. nov. We have attempted to compute an improved phylogeny of this group, constructing the SSU rDNA sequence alignment according to the alignments based on a secondary structure model provided in the European Small Subunit Ribosomal RNA database of Van de Peer et al. (2000).

We have also provided morphological descriptions of the newly sequenced organisms, and named a new genus, because existing information on many of these small flagellates is inadequate for purposes of identification, and we wish to convey an understanding of how we have come to recognise them.

During the final revision of the present paper, Cavalier-Smith and Chao (2003) published a revised classification of "Cercozoa", which included several new SSU rDNA sequences. Because of the recent appearance of this paper, we have only used it to a limited extent. Because Cavalier-Smith and Chao (2003) focus on a higher level classification than we do, our paper would probably not have been substantially different, had we known the study by Cavalier-Smith and Chao (2003) at an earlier stage.

Material and methods

Cultures

One strain of *Heteromita globosa* (Stein, 1878) Kent 1880, and two strains with "*Cercomonas* morphology" were isolated from an agricultural soil in Foulum, Jutland, Denmark by F. Ekelund. Prof. K. Vickerman kindly provided cultures of *Thaumatomonas lauterborni* De Saedeleer (1931), *T. seravini* Mylnikov and Karpov (1993), and an unidentified *Thaumatomonas* sp. soil isolate. The organisms were kept in culture at 10°C in 50 ml Nunc-flasks by subsequent re-inoculations in Neff's Amoebae Saline (Page, 1988) on either 0.1 g l^{-1} tryptic soy broth (Difco, Detroit, MI) or a sterilised wheat grain to sustain bacterial growth.

To avoid eukaryotic contamination of the flagellates, each species was passed through three sequential dilution cultures in 96-well microtitre plates at 2-week intervals. We used the highest dilutions producing healthy cultures in a well for subsequent re-inoculations. During this procedure we fed the flagellates a suspension of *Pseudomonas fluorescens* DR54, which is known to produce fungicidal substances (Nielsen et al., 1999). Careful examination by light microscopy for protozoa, yeast cells and hyphae ensured the absence of

contaminating eukaryotes. To reduce the amount of co-extracted bacterial DNA, the flagellates were allowed to prey heavily on their bacterial food before harvesting for DNA extraction.

DNA extraction and sequencing

DNA was extracted from 1.0 ml culture-samples with the FastDNA SPIN Kit (for Soil) (Cat ♯6560-200, Bio101, Vista, CA) using a FastPrep FP120 (Bio101) machine for the bead-beating step of the manual. Primers ND1F and ND6R (Table 1) were used to amplify the full SSU rDNA gene using the PCR-protocol of Johnsen et al. (1999), with some modifications. Each PCR-tube contained a total volume of 25 µl PCR reaction mixture: 1 × Ampli *Taq* PCR buffer (PE Biosystems, Norwalk, Connecticut.), 0.1% DNAse-free bovine serum albumin (Amersham Pharmacia, Uppsala, Sweden), 0.4 mM dNTP mixture (PE Biosystems), 0.5 µm of each primer of desalted purity (Life Technologies, Roskilde, Denmark), 0.12 µl Ampli *Taq* Gold polymerase (PE Biosystems) and 1 µl template DNA. The PCR-amplification conditions were as follows: 10 min at 95°C; 35 cycles of 30 s at 95°C, 30 s at 54°C and 2 min at 72°C; 6 min at 72°C; and final hold at 4°C. PCR was performed on a Techne (Genius) thermal cycler.

The amplified SSU rDNA PCR products were purified using the QIAquick PCR Purification Kit (QIAGEN GmbH, Hilden, Germany). We sequenced on an ABI model 377 sequencer, using the primers of Table 1 with the ABI PRISM BigDye Terminator Cycle Sequencing protocol (PE Applied Biosystems, 1998).

Alignment and tree construction

From the database of small ribosomal subunit RNA (Van de Peer et al., 2000), we retrieved pre-aligned sequences to which we added the six novel sequences we determined as well as the sequence of *Cercomonas* "*longicauda*" (Zaman et al., 1999) and *Massisteria marina* (Atkins et al., 2000). All organisms used in our analyses, including those obtained from GenBank, are listed in Table 2. Each alignment includes all known positions in the eukaryotic SSU rDNAs, and is based on similarities in both the primary and secondary structures of the presumed closest relatives to the organism. We further processed the retrieved sequences with the BioEdit Sequence Alignment Editor ver. 5.0.9 (Hall, 1999).

From this first alignment we assembled 53 eukaryotic taxa which represent major eukaryotic 'crown' lineages, in order to identify the nearest sister taxa to the "cercozoan clade". This was done by "substitution rate calibration" with subsequent pair-wise distance calculations and tree construction by neighbour-joining, as follows. The 53-taxon alignment was carefully revised for obvious alignment ambiguities, and regions of substantial gaps and uncertain positional homology. Alignment columns corresponding to terminal length variation were omitted as recommended by Swofford et al. (1996). The result was a final alignment of 2921 positions (including gaps). We used the TREECON software package (Van de Peer and De Wachter, 1994) to perform the nucleotide "substitution rate calibration" with three iterations of the calibration procedure as in Van de Peer et al. (1993b). This procedure takes into account all alignment positions that contain nucleotides in more than 25% of the aligned sequences and are not absolutely conserved. The program excludes the remaining positions from the output alignment file, which is used for further computations.

Dissimilarities according to the derived relative substitution rates for each position of the reduced dataset (1227 bp) were converted to pair-wise evolutionary distances using a value of $p = 0.52$ in the Van de Peer et al. (1993b) correction equation given as an option in TREECON. The TREECON package was also used to perform the neighbour-joining analysis

Table 1. List of SSU rDNA primers used in the present study

Primer	Use[a]	Nucleotide sequence	Positions[b]
ND1F	P + S	5′ AACCTGGTTGATCCTGCCAG 3′	(107–132)
ND2F	S	5′ GATTCCGGAGAGGGAGCCTG 3′	(1540–1565)
ND3F	S	5′ GCCGCGGTAATTCCAGCTCC 3′	(2075–2104)
ND4F	S	5′ AGAGGTGAAATTCTCAGA 3′	(3791–3819)
ND5F	S	5′ GGTGGTGCATGGCCGTTC 3′	(4786–4810)
ND6R	P + S	5′ GATCCTTCTGCAGGTTCACC 3′	(6373–6354)
ND7R	S	5′ GAACGGCCATGCACCACC 3′	(4810–4786)
ND8R	S	5′ TCTGAGAATTTCACCTCT 3′	(3819–3791)
ND9R	S	5′ CCGCGGCTGCTGGCACCAGAC 3′	(2082–2051)
ND10R	S	5′ CTCAGGCTCCCTCTCCGG 3′	(1568–1544)

[a] P = PCR amplification, S = Sequencing.
[b] Numbers refer to positions in the alignment of Van de Peer et al. (1997).

Table 2. Organisms used, with strain and GenBank accession numbers

Organism	Strain	Acc.-No.
Acanthamoeba lenticulata	ATCC50704	U94732
Acanthocoepsis unguiculata	—	L10823
Anemonia sulcata	—	X53498
Apusomonas proboscidea	ATCC 50315, CCAP 1905/1	L37037
Assulina muscorum	—	AJ418791
Bacillaria paxillifer	—	M87325
Bigelowiella natans	CCMP 621	AF054832
Blepharisma americanum	—	M97909
Cafeteria roenbergensis	—	L27633
Ceramium macilentum	ATCC30927	U32562
Cercomonas ''longicauda''	ATCC 50344	AF101052
Cercomonas longicauda	SCCAP C 1	AY496047
Cercomonas sp.	ATCC 50316	U42448
Cercomonas sp.	ATCC 50317	U42449
Cercomonas sp.	ATCC 50318	U42450
Cercomonas sp.	ATCC 50319	U42451
Chilomonas paramecium	Strain CCAP 977/2a	L28811
Chlorarachnion reptans	CCMP 238	U03477
Chlorarachnion reptans	Grell, Tunisia	X70809
Chlorarachnion sp.	CCMP1408	U02075
Chlorarachnion sp.	CCMP 242	U03479
Chlorarachniophyte, unidentified	BC 52	AF076172
Chlorarachniophyte, unidentified	CCMP 1239	AF076173
Chlorarachniophyte, unidentified	CCMP 1258	AF076174
Chrysonephele palustris	—	U71196
Chytridium confervae	81-1,U.C. Microgarden, Berkeley	M59758
Cryothecomonas aestivalis	—	AF290539
Cryothecomonas longipes	—	AF290540
Crypthecodinium cohnii	—	M64245
Cyanophora paradoxa	Kies	X68483
Diaphanoeca grandis	—	L10824
Dictyostelium discoideum	—	K02641
Emiliania huxleyi	CCMP 374	L04957
Entamoeba gingivalis	ATCC30927	D28490
Erythrotrichia carnea	—	L26189
Euglypha acanthophora	Millstream	AJ418788
Euglypha filifera	Costa Rica	AJ418785
Euglypha rotunda	CCAP 1520/1	X77692
Euglypha tuberculata	Millstream	AJ418787
Gloeochaete wittrockiana	SAG B 46.84	X81901
Guillardia theta	—	X57162.1
Gymnochlora stellata	—	AF076171
Haplosporidium nelsoni	—	U19538
Hartmannella vermiformis	CCAP#1534 /7B	X75515
Heteromita globosa	CAP 1961/2	U42447
Heteromita globosa	SCCAP H 251	AY496043
Lotharella globosa	—	AF076169
Lotharella sp.	CCMP 240	AF076168
Lotharella vacuolata	CCMP 240	AF054890
Massisteria marina	ATCC50266	AF174369
Microciona prolifera	—	L10825
Minchinia teredinis	Isolate 1993, Wachaprague, VA	U20320
Neocercomonas jutlandica	SCCAP C 161	AY496048
Ochromonas danica	—	M32704
Oxytricha nova	—	X03948
Paulinella chromatophora	M0880 (Melkonian Cult. Coll.)	X81811
Pavlova salina	strain Plymouth 486	L34669

Table 2 (*continued*)

Organism	Strain	Acc.-No.
Pelagomonas calceolata	CCMP 1214	U14389
Perkinsus marinus	—	X75762
Phreatamoeba balamuthi	—	L23799
Physarum polycephalum	CL	X13160
Plasmodiophora brassicae	—	U18981
Porphyra umbilicalis	—	Ll26202
Prorocentrum micans	(Strain Ehrenberg) DNA, clone pr10	M14649
Proteromonas lacertae	ATCC 30270	U37108
Prymnesium patelliferum	Strain Plymouth 527	L34670
Rhodella maculata	CCMP 736	U21217
Saccharomyces cerevisiae	—	J01353
Salpingoeca infusionum	ATCC50559	AF100941
Spongospora subterranea	f. sp. nasturtii	AF245217
Teleaulax amphioxeia	K-0434	AJ007287
Thaumatomonas lauterborni	SCCAP T 2	AY496045
Thaumatomonas seravini	SCCAP T 3	AY496044
Thaumatomonas sp.	—	U42446
Thaumatomonas sp.	SCCAP T 1	AY496046
Theileria parva	Clone 7344/G5/F5/E8	AF013418
Trachelocorythion pulchellum	—	AJ418789
Trinema enchelys	—	AJ418792
Volvox carteri, f. *nagariensis*,	Female strain HK10 (UTEX 1885)	X53904
Zamia pumila	Leaf tissue DNA, clone Zfr-1042	M20017

(Saitou and Nei, 1987), including bootstrap analysis (Felsenstein, 1985) with computation of 1000 replicate trees from re-sampled data, and to draw the corresponding evolutionary consensus tree. Although rooting the tree of life is not a trivial matter, we decided to use *Dictyostelium discoideum* to polarise the distance tree (see for example: Van der Auwera and De Wachter, 1997). The two haplosporidians, *Minchinia teredinis* and *Haplosporidium nelsoni*, and the plasmodiophorids appear associated with the "cercozoan clade" in this analysis and were consequently included in the subsequent analyses.

We prepared a 48-taxon alignment of all "cercozoan" SSU rDNA sequences listed in Table 2. The in-group included *Heteromita* (two sequences), *Cercomonas* (seven sequences), *Thaumatomonas* (four sequences), eight euglyphid testate amoebae, *Cryothecomonas* (two sequences), *Massisteria marina*, and sequences of 12 chlorarachniophyte algae and two plasmodiophorids. The two haplosporid sequences were also included for the reasons mentioned above. Sequences of five heterokont algae and three alveolates constituted the outgroup taxa.

This 48-taxon alignment was subjected to the "substitution rate calibration" with three iterations as previously described. The calculated value of p was 0.53. The reduced TREECON output alignment file (979 positions) formed the basis of analyses using three methods of phylogenetic tree-construction: neighbour-joining, parsimony, and maximum likelihood. The tree topology of the phylogeny inferred by neighbour-joining was examined by performing 1000 bootstrap replicates. Parsimony and maximum likelihood (Felsenstein, 1981) trees were computed with PAUP* by heuristic searches (Swofford, 2003). A total of 1000 random additions of sequences resulted in three equally parsimonious trees each of 3917 steps. For parsimony bootstrap analysis 1000 replicates were performed. For maximum likelihood analyses 10 random additions of sequences were performed and bootstrap analysis was based on 100 replicates. The tree topology inferred from parsimony and maximum likelihood methods was almost identical.

Morphological characterisation of cultures

Heteromita globosa and the two *Cercomonas* strains isolated from Danish soil as well as the unidentified *Thaumatomonas* sp. soil isolate provided by Prof. K. Vickerman were examined using an Olympus BX50 microscope equipped with phase and interference contrast at 1000× magnification. Photos were taken with an Olympus OM4 camera using a flash. All organisms characterised by us are named according to ICZN. All cultures sequenced by us are kept in the Scandinavian Culture Collection of Algae and Protozoa (SCCAP located at the Biological Institute, University of Copenhagen). Strain accession numbers are given in Table 2.

Results

Sequence analysis

Complete double-stranded sequences were obtained for the *Thaumatomonas* sp. soil isolate (strain SCCAP T 1), *Thaumatomonas seravini*, and *Cercomonas longicauda* strain SCCAP C 1). Partial double-stranded sequences were obtained for the *Thaumatomonas lauterborni* (1410 bp), *Heteromita globosa* (strain SCCAP H 251) (1611 bp), and *Neocercomonas jutlandica* (strain SCCAP C 161) (1352 bp). The six SSU rDNA nucleotide sequences have been submitted to GenBank; accession numbers are listed in Table 2.

Our *Heteromita globosa* sequence showed 99% SSU rDNA homology with a strain available in GenBank (U42447), but differed from this by having a large ~600 bp insertion near the 5′-end. The three new thaumatomonad sequences suggest a low genetic variability within *Thaumatomonas*, as we found 98% homology among the four sequences. The two new *Cercomonas* sequences of this study differed in base composition by 11%. The new *Cercomonas longicauda* strain (SCCAP C 1) shared 98% homology with the sequence of the ATCC 50344 *Cercomonas* "*longicauda*" strain, and 95% homology with the ATCC 50316 *Cercomonas* sp. strain. The *Neocercomonas jutlandica* (SCCAP C 161) sequence shared 95% homology with *Cercomonas* spp. strains ATCC 50317 and ATCC 50318, and a 92% sequence homology with the ATCC 50319 *Cercomonas* sp. strain.

Neighbour-joining analysis of the 53 eukaryote taxa

The tree constructed by neighbour-joining following pair-wise distance calculations of the weighted assembly of the 53 eukaryote taxa is shown in Fig. 1. The branching of the major eukaryotic lineages is not well resolved, as most of the deep nodes have bootstrap support values of less than 50%. Subject to this shortcoming, the sister groups closest to the "cercozoan clade" appear to be the plasmodiophorids, heterokonts and apicomplexa, but the reliability of this clustering is weakly supported in terms of bootstrap values. It is thus not possible to safely identify the closest relatives to the group. It was therefore decided to include both heterokonts and alveolates as outgroup taxa for the subsequent analyses of the phylogenetic relationships within the group.

In the 53-eukaryote taxon distance tree (Fig. 1), which is designed to obtain an equal taxon sampling among major lineages, the "cercozoan clade" is monophyletic but only supported by a bootstrap value of 53%. *Massisteria marina* forms a sister taxon to a clade comprising *Euglypha*/*Thaumatomonas*/*Heteromita*/*Cercomonas*. Unexpectedly the two haplosporidian organisms, *Minchinia teredinis* and *Haplosporidium nelsoni*, clustered as the sister group to the "cercozoan clade". This branching pattern was supported by a bootstrap value of 71%. The neighbour-joining tree suggests that the haplosporids emerge after the plasmodiophorids, but prior to the other groups, which led us to include the two haplosporidian sequences in the further phylogenetic analyses. The position of several individual taxa with long branches is anomalous, but these do not interfere with the general conclusions about the principal groups.

Phylogenetic analysis of the "Cercozoa"

The tree derived from the neighbour-joining analysis of 48 taxa is shown in Fig. 2, and that from parsimony and maximum likelihood methods in Fig. 3. In these more detailed analyses monophyly of the "cercozoan clade" is supported with bootstrap values of 98% in maximum likelihood analysis and 100% in parsimony analysis (Fig. 3). In the "substitution rate calibration", *Massisteria marina* diverged before the plasmodiophorids and the rest of the "Cercozoa" (Fig. 2), which does not support the monophyly of the group if it is accepted that *Massisteria* should be included as a member of the "Cercozoa". All three methods suggest that the *Cercomonas* strains are separated into two very distinct groups, labelled "type 1" and "type 2", respectively. Monophyly of these two groups of strains of *Cercomonas* is supported by very high bootstrap values (≥99%). "Type 1" includes *Cercomonas longicauda* (SCCAP C 1), *Cercomonas "longicauda"* (ATCC 50344), and *Cercomonas* sp. (ATCC 50316). The "type 2" group includes the three remaining ATCC *Cercomonas* strains as well as our strain SCCAP C 161, which we transfer to the new genus *Neocercomonas* (see below). The four *Thaumatomonas* sequences and the two *Heteromita globosa* strains also constitute separate and well-supported clades, each with ≥99% bootstrap support. Except for suggesting an early divergence of "type 1" *Cercomonas* strains, the order of emergence of the remaining three groupings (*Cercomonas* "type 2", *Heteromita* and *Thaumatomonas*) is not well indicated in the present analyses due to an incongruent branching pattern. The position of the euglyphid testate amoebae included in the analysis varies according to the method chosen. The neighbour-joining analysis (Fig. 2) suggests that *Paulinella chromatophora* is more closely related to the chlorarachniophytes than to the other euglyphid testate amoebae, although this relationship is not supported in terms of bootstrap values. The remaining seven taxa of testate amoebae included form one clade supported by a bootstrap value of 99%. In parsimony and maximum likelihood analyses the testate amoebae are monophyletic within the clade comprising

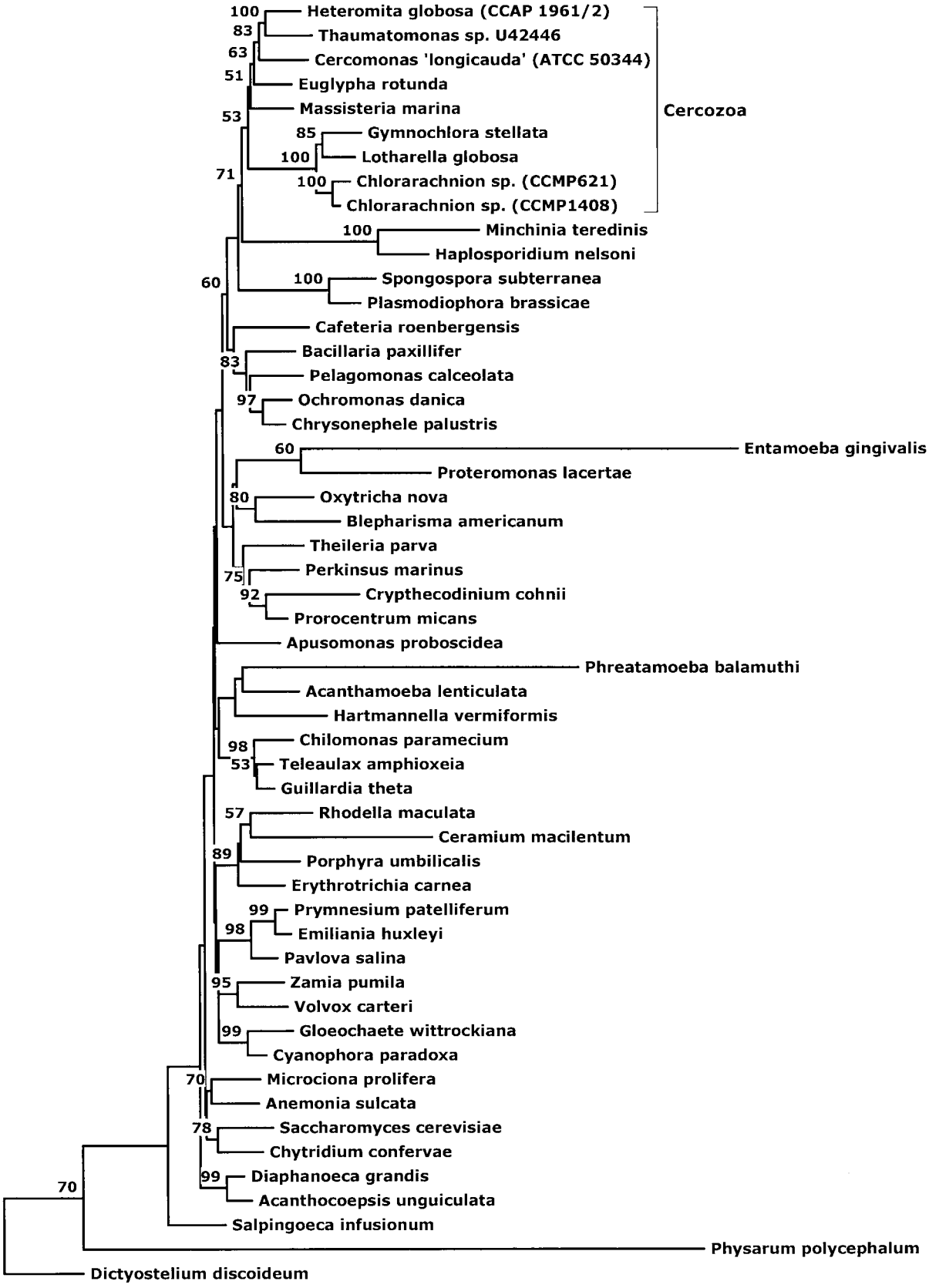


Fig. 1. Neighbour-joining tree from a distance matrix of 53 eukaryotic SSU rDNA sequences and based on the "substitution rate calibration" ($p = 0.52$). The reduced data matrix comprised 1227 base pairs. The slime mould *Dictyostelium discoideum* was used to root the tree. Bootstrap values $\geq 50\%$ are shown at internal nodes (1000 replicates were performed).

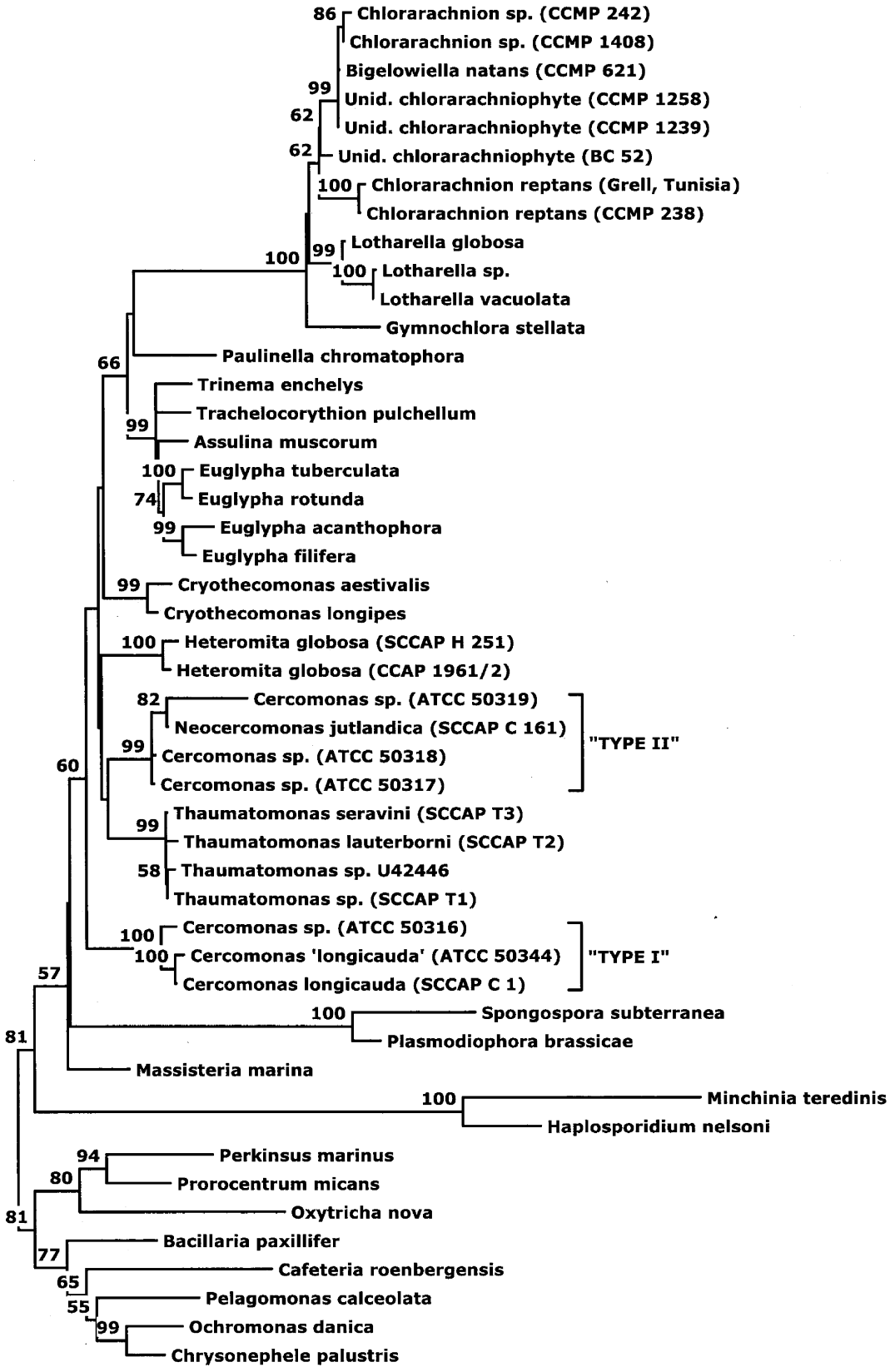


Fig. 2. Neighbour-joining tree from a distance matrix of 48 eukaryotic SSU rDNA sequences and based on the "substitution rate calibration" ($p = 0.53$). The reduced data matrix comprised 979 base pairs. Three alveolates and five heterokonts were used as outgroup taxa. Bootstrap values $\geq 50\%$ are shown at internal nodes (1000 replicates were performed).

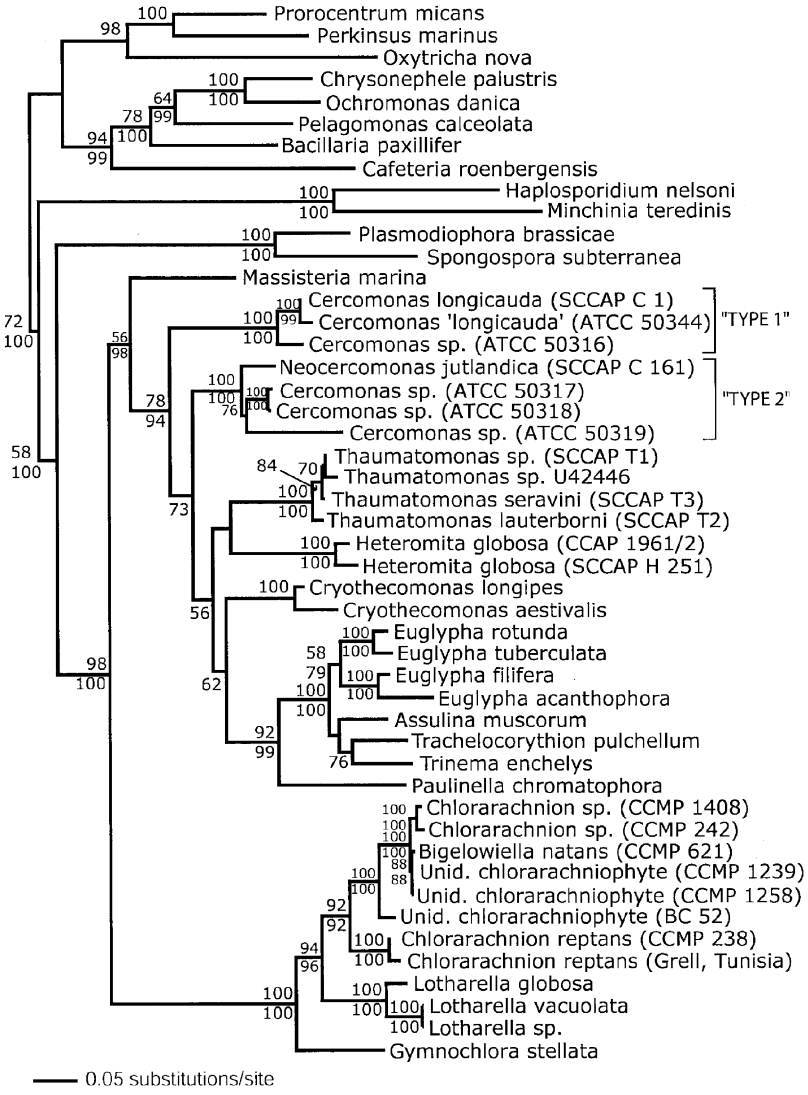


Fig. 3. Maximum likelihood tree based on the reduced data matrix comprising 979 base pairs from 48 eukaryotic SSU rDNA sequences. The best ln likelihood score was −18177.098. Three alveolates and five heterokonts were used as outgroup taxa. Bootstrap percentages above branches are calculated from ML analyses (100 replicates) and numbers below branches are calculated from parsimony analyses (1000 replicates). Parsimony analysis of the same data matrix with 1000 random additions of sequences produced three equally parsimonious trees each 3917 steps (CI = 0.462; RI = 0.669).

Cercomonas (“type 2”), *Heteromita*, *Thaumatomonas* and *Cryothecomonas* (≥92%). The chlorarachniophyte clade is strongly supported (100% bootstrap support in all analyses), and is separated from the remaining “Cercozoa” by a long evolutionary branch (Figs. 2 and 3). The early divergence of the chlorarachniophytes within the “Cercozoa” is favoured by high bootstrap support in parsimony and maximum likelihood analyses. All methods show five distinct groupings within the Chlorarachniophyta, corresponding to the five major clades of Ishida et al. (1999).

In the neighbour-joining tree (Fig. 2) the haplosporids diverge separately and prior to the plasmodiophorids. The only strong indication of a sister group relationship between haplosporids and plasmodiophorids is given by the parsimony analysis, with a bootstrap support of 82% (data not shown). Although the plasmodiophorids and haplosporids do form sister groups to the other “cercozoan” taxa, they are clearly separated from these by long evolutionary branches. This either indicates early diversification or increased evolutionary rates within these lineages (Figs. 1–3) or a slowdown in the evolutionary rates of the “cercozoan” lineages.

Characterisation of cultures

In the sections below we characterise morphologically the cultures that we have sequenced. Where there are old, and often inadequate, descriptions of species

lacking type material, it has become a practice with some people over the last 20 years or so to re-describe the species and designate neotype material, which we have done. See Foissner (2002) for details on neotypification of protists.

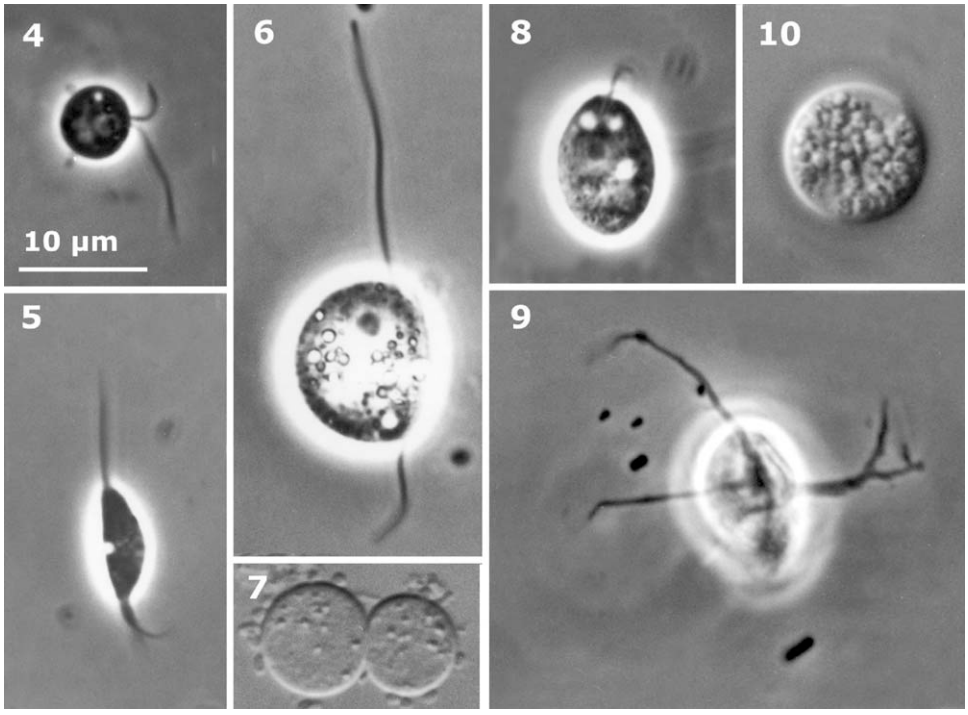
Heteromita globosa (Stein, 1878) Kent, 1880 (Figs. 4 and 11)

Description: Cells, 4–6 μm long, globular, ovate or obovate in outline, not compressed. Anterior flagellum about cell length, posterior flagellum 1.5–2.5 times cell length. Nucleus central to anterior, contractile vacuole anterior or posterior to nucleus. Occasionally with pseudopodia, mostly as a posterior tail-like extension, more rarely from other parts of the cell. Cysts readily formed in cultures, globular, 3–4 μm in diameter.

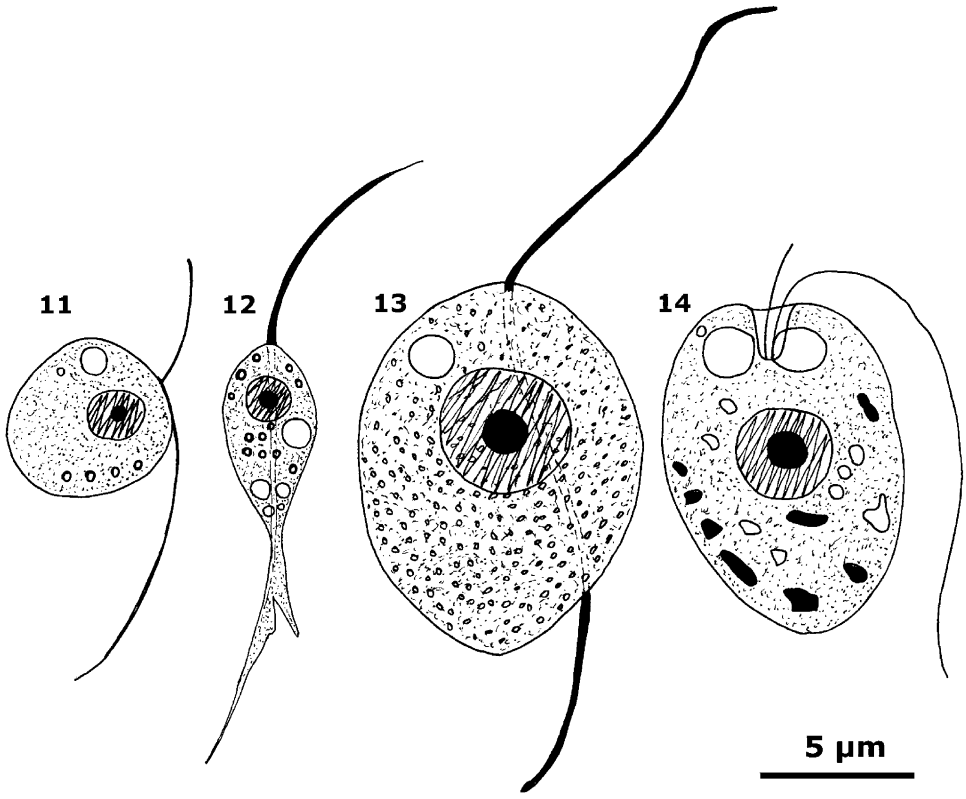
Movement: A slow wobbling, posterior flagellum trails, anterior flagellum moves in an irregular manner in front of the cell. Cells which lose contact with the substrate move haphazardly. Creeping cells never observed.

Comments: *Heteromita globosa* as conceived here agrees with earlier descriptions (Sandon, 1927; Robertson, 1928; MacDonald et al., 1977), except that the contractile vacuole is not always posterior to the nucleus as described by MacDonald et al. (1977). The main criterion used to distinguish between *Heteromita* and *Cercomonas* is motility; organisms which always move in a wobbling manner are assigned to *Heteromita* while forms which mainly move by creeping are assigned to *Cercomonas*.

Different clones of *Heteromita globosa* may differ in size, shape and tendency to form pseudopodia (Ekelund et al., 2001; F. Ekelund personal observation). Some isolates only form pseudopodia occasionally; other forms nearly always carry a posterior pseudopodial tail. Because of the continuous variation between these extremes it has not been possible to delimit separate species. Additional morphological and molecular work is needed to clarify whether all forms are best conceived as one species or should be split into several species. Still, in order to define the taxon *Heteromita globosa* in an unambiguous manner we suggest the diagnosis below. The diagnosis is slightly broader than the description given above, as various isolates differ somewhat. The four short SSU rDNA signature sequences (H1–H4) which we have included in the description are shared by the CCAP 1961/2 strain, the sequence we present in this paper from SCCAP H 251, and another isolate, with a more amoeboid outline, that



Figs. 4–10. Light micrographs of live cells. Fig. 4. *Heteromita globosa*, wobbling motile cell (phase contrast). Fig. 5. *Cercomonas longicauda*, creeping cell (phase contrast). Fig. 6. *Neocercomonas jutlandica*, creeping cell (phase contrast). Fig. 7. *Neocercomonas jutlandica*, resting cysts (DIC). Fig. 8. *Thaumatomonas* sp., SCCAP T 1), gliding cell (phase contrast). Fig. 9. *Thaumatomonas* sp. (SCCAP T 1), cell producing ventral pseudopodia for food uptake (phase contrast). Fig. 10. *Thaumatomonas* sp. (SCCAP T 1), resting cyst (DIC). All same scale.



Figs. 11–14. Drawings based on live cells. Fig. 11. *Heteromita globosa*, wobbling motile cell. Fig. 12. *Cercomonas longicauda*, creeping cell. Fig. 13. *Neocercomonas jutlandica*, creeping cell. Fig. 14. *Thaumatomonas* sp. (SCCAP T 1), gliding cell.

we have also sequenced (data not shown). Apart from the *Heteromita globosa* sequence in GenBank (U42447), which matched all four signature sequences, seven other organisms in GenBank matched one or more of the signature sequences that we suggest for *Heteromita globosa*. *Cryothecomonas longipes* (AF290540) matched one sequence, but as *Cryothecomonas* is clearly morphologically different from *Heteromita globosa*, we do not regard one match and three mismatches as seriously damaging to the diagnosis. The other six sequences, which matched one or more of the signature sequences, were environmental clones. Three of these were short partial (132 bp) fossil ice clones (Willerslev et al., 1999), and the other three were partial (1275–1278bp) cloned sequences from anoxic habitats, one of which matched all four signature sequences. All six environmental sequences showed a high similarity to *Heteromita globosa* within the sequenced region, with the highest similarity found for the sequence which matched all four signature sequences. As we do not have access to morphological data on the clones, and as more studies are needed to finally reveal the variation of the *Heteromita globosa* SSU rDNA, the following diagnosis simply defines *Heteromita globosa* as organisms with an external "*Heteromita globosa* morphology" that possess all four signature sequences.

Heteromita globosa (Stein, 1878) Kent, 1880 (neotype micrograph: Fig. 4; neotype culture: SCCAP H251)

Diagnosis: Cells, 4–10 µm long, globular, ovate or obovate in outline, with two heterodynamic flagella. Occasionally with pseudopodia, mostly as a posterior tail-like extension, more rarely from other parts of cell. No cytostome. Movement is slow and wobbling, posterior flagellum trails, anterior flagellum moves in an irregular manner in front of the cell. With the four *SSU rDNA signature sequences*:

H1: (5′TGAGTGAACGTTGCATTTTC3′),
H2: (5′CTAAATAGTTCGACAAAGGA3′),
H3: (5′TGTGTAGCCGACGGAAGTTC3′) and
H4: (5′TGATGCATTCATCAAGTTTA3′).

Neotype location: Agricultural soil, Foulum, Jutland, Denmark.

Neocercomonas jutlandica gen. nov. et sp. nov. (Figs. 6, 7 and 13)

Description: Actively creeping cells moderately metabolic, nearly globular to droplet shaped, rarely ovoid–oblong. Length, not including pseudopodia, 10–16 µm. Pseudopodia normally formed from the posterior end,

rarely from sides of the cell. Two heterodynamic, not visibly acronematic, flagella arise at the front end. Flagella usually about 1.5 times cell length, occasionally shorter, anterior flagellum sometimes nearly absent. Nucleus central in anterior end. Contractile vacuole lateral, anterior to nucleus. Cell rich in granules; refractile granules posterior to nucleus; non-refractile scattered in the cytoplasm, mostly anterior to the nucleus. Cell content, in particular contractile vacuole, difficult to see because of the granules. Globular cysts, 5–8 μm in diameter, readily formed in culture.

Movement: With different types of behaviour. Floating forms with a sac-like appearance are observed in newly fed cultures. Recently settled cells are usually globular to weakly droplet-shaped, without pseudopodia, and move in a wobbling manner much like *Heteromita globosa*, although more slowly; the posterior flagellum has contact with the substrate but does not adhere to the cell. The anterior flagellum beats in a sweeping way. Settled cells gradually assume a creeping, more amoeboid, and strongly droplet-shaped form as they gain contact with the substrate for most of their length. Pseudopodia are formed from the posterior end, which is occasionally drawn out as a long tail. The anterior flagellum in creeping cells moves in an erratic snake-like manner, the anterior half being more active than the posterior. In old cultures, strongly amoeboid cells, which move very slowly, sometimes appear.

Comments: Cells with a morphology corresponding to this description of *Neocercomonas jutlandica* can be found in almost any soil sample (F. Ekelund unpublished results). We suggest that a unique signature sequence, which is not shared by any other organism, should be used to define *N. jutlandica*. Since both the phylogenetic analysis of Cavalier-Smith and Chao (1996/97) and our analysis strongly suggest that *Cercomonas* is not a monophyletic group we find it best to divide the genus into two separate groups. The signature sequences we have used to characterise *Neocercomonas* are shared by *N. jutlandica* and the "Type 2 *Cercomonas* strains" included in GenBank. Hopefully, future research will provide morphological criteria for the distinction between *Cercomonas* and *Neocercomonas*.

Neocercomonas jutlandica sp. nov. gen. nov. (type micrograph: Fig. 6; type culture: SCCAP C161)

Diagnosis: Cells, 10–16 μm long, nearly globular to droplet-shaped, rarely ovoid–oblong. Usually with pseudopodia from the posterior end, more rarely from sides of the cell. Two heterodynamic flagella, usually c. 1.5 times cell length but occasionally shorter, arise at the front end. Nucleus central and in anterior end. Contractile vacuole lateral and anterior to nucleus. Cell rich in granules; refractile granules posterior to nucleus; non-refractile granules scattered in the cytoplasm, mostly anterior to the nucleus. Cysts globular 5–8 μm in diameter, readily formed. Movement usually a slow creeping with the anterior flagellum making erratic snake-like motions and the posterior trailing. *SSU rDNA signature sequence*:

NJ: (5′ TCCACGAATTCCTCGGGATT 3′)

Type location: Agricultural soil, Foulum, Jutland, Denmark.

Neocercomonas gen. nov

Diagnosis: Cells amoeboid, usually with pseudopodia. No cytostome. Two heterodynamic flagella. Movement usually creeping with the anterior flagellum directed forwards and the posterior trailing. *SSU rDNA signature sequences*:

N1: (5′GAGGGACTATCGGTGATTTA3′) and

N2: (5′TCGAGCTTACAACCTTGGTT3′).

Type species: *Neocercomonas jutlandica* Ekelund, Fredslund & Daugbjerg.

Cercomonas longicauda Dujardin 1841 (Figs. 5, 12)

Description: Length, not including pseudopodia, 4–10 μm. Cell highly metabolic with strong amoeboid properties. Two heterodynamic flagella arise at the front end, anterior flagellum of about cell length, posterior flagellum 1–2 times the cell length. Nucleus central and in anterior cell end. Contractile vacuole lateral, close to and usually behind, but occasionally anterior to, the nucleus. Cytoplasm rich in granules, some granules in posterior end strongly refractile. Several, differently sized, vacuoles in the posterior end. Globular cysts, 3–4 μm in diameter, readily formed.

Movement: Actively creeping cells are droplet to spindle-shaped with anterior flagellum directed forward, stroke of the anterior flagellum covers a cone in front of the cell with a top angle of about 30°, actively creeping cells sometimes display a weak wobbling. During active movement strands of cytoplasm are drawn after the cell, posterior flagellum mostly not visible because it is covered by the cytoplasm. In newly fed cultures cells assume a floating condition with a spherical to shortly droplet-shaped outline. The cells are usually attached to the substratum by the posterior flagellum and move erratically in the water. In older cultures where food is depleted cells become strongly amoeboid and move very slowly; in this condition pseudopodia may be formed from any part of the cell, although mostly from the posterior end. These pseudopodia look different and are sometimes strongly branched.

Comments: *Cercomonas longicauda* resembles *Cercomonas minimus* Mylnikov (1985), but is larger than the

3.3–6 μm reported by Mylnikov (1985). *Cercomonas longicauda*, and forms with similar morphology, are extremely common in almost all types of soil (F. Ekelund unpublished results). Sandon (1927) reported "*Cercomonas longicauda* Dujardin 1841" to be a common soil inhabitant. According to Sandon's brief description the cell he saw looked much like *C. longicauda* as conceived here, although it was reported with a longer anterior flagellum. We stress that we have not studied the morphology of the strain, which Zaman et al. (1999) provisionally named *Cercomonas* "*longicauda*", but the SSU rDNA sequence of the Zaman et al. strain differs from that of this *C. longicauda*. The genus *Cercomonas* was erected by Dujardin and the type species (*C. longicauda*) was lectotypified by Fromentel. The signature sequence we suggest here for *C. longicauda* is unique for the organism that we have investigated, and the signature sequences that we suggest for the genus *Cercomonas* are unique for the *Cercomonas* type 1 strains.

Cercomonas longicauda (neotype micrograph: Fig. 5; neotype culture: SCCAP C1)

Diagnosis: Cells, 4–10 μm long, highly metabolic with strong amoeboid properties. Two heterodynamic flagella at the front end, anterior flagellum of about the cell length, posterior flagellum 1–2 times the cell length. Nucleus central in anterior cell end. Contractile vacuole lateral, close to and usually behind, occasionally anterior to, the nucleus. Cytoplasm rich in granules, some granules in posterior end strongly refractile. Several, differently sized, vacuoles in the posterior end. Cysts globular, 3–4 μm in diameter, readily formed. Movement creeping with anterior flagellum directed forward, stroke of the anterior flagellum covers a cone in front of the cell with a top angle of about 30°, with cytoplasmic strands from posterior end that cover the trailing flagellum. *SSU rDNA signature sequence*:

CL: 5′ACAACGTAACCCTTGGTTAT3′.

Neotype location: Agricultural soil, Foulum, Jutland, Denmark.

Cercomonas Dujardin 1841

Diagnosis: Cells amoeboid, usually with pseudopodia. No cytostome. Two heterodynamic flagella. Movement usually creeping with the anterior flagellum directed forwards and the posterior trailing. *SSU rDNA signature sequences*:

C1: (5′TGGATTCGATGAAGACTAAC3′),
C2: (5′ACTAGGGATTAGTGGAAGTT3′),
C3: (5′TTTTCGACTCCATTAGCACC3′) and
C4: (5′GAAGTCTTTGGGTTCTGGGG3′).

Type species: *Cercomonas longicauda* Dujardin 1841

Thaumatomonas sp. (soil strain provided by Keith Vickerman, Figs. 8–10, 14)

Description: Trophic cells 8–13 μm long, outline usually oval, occasionally droplet-shaped, anterior end obliquely truncate. Two heterodynamic flagella emerge from a flagellar pocket, about 1/10 of cell length deep. Anterior flagellum of 1/3 to 1/2 cell length, posterior flagellum about 1.5 times cell length. Nucleus central or somewhat anterior. Two conspicuous anterior contractile vacuoles, one on each side of the flagellar pocket, into which they empty their contents. Cytoplasm filled with granules of different size and shape. A ventral furrow extends from anterior to posterior end. Extensive branched pseudopodia sometimes formed from furrow. Cysts, globular, thick walled, readily formed in culture, variable in size, diameter 7–12 μm.

Movement: A steady non-wobbling gliding. The anterior flagellum beats to the left, the posterior flagellum trails passively. Sometimes cells stop and form pseudopodia.

Comments: The two described species of *Thaumatomonas*, *T. lauterborni* De Saedeleer (1931) and *T. seravini*, Mylnikov and Karpov (1993), differ only in minor details. *Thaumatomonas seravini* has a less elongate cell outline than *T. lauterborni*; *T. seravini* possesses endosymbiotic bacteria, which are absent from *T. lauterborni*, and the surface scales of the two species differ slightly (De Saedeleer, 1931; Mylnikov and Karpov, 1993). The SSU rDNA sequences suggest that the four *Thaumatomonas* strains studied are very close relatives. If surface scales were not usually considered an important character for species distinction among flagellates, we would be tempted to consider *T. lauterborni*, *T. seravini* and the two unnamed forms conspecific. However, if *T. lauterborni* and *T. seravini* are to be considered separate species, then we cannot assign a specific name to the examined isolate, until we have examined its ultrastructure.

Discussion

Position of the investigated organisms among other eukaryotes

The neighbour-joining tree (Fig. 1) illustrates the common problem of resolving the order of divergence of the major eukaryotic lineages by SSU rDNA comparisons (Van de Peer et al., 1993a; Kumar and Rzhetsky, 1996; Van de Peer and De Wachter, 1997). Bootstrap values above 70% are likely to define stable groupings (Hillis and Bull, 1993; for a detailed discussion, see Wainwright et al., 1993; Felsenstein and Kishino, 1993), but every deep node in Fig. 1 is supported by less than 50%. A 'big bang' in eukaryotic evolution in the time

following mitochondrial endosymbiosis (700–1000 million years ago) has been suggested to explain the rapid radiation of the eukaryotic lineages (Philippe and Adoutte, 1998). This energetic advantage could have led to the extinction of all amitochondrial ancestors. The proceeding diversification of mitochondrial eukaryotes may then have occurred during a relatively short period of time, which is why we fail to find either morphological or molecular synapomorphies sufficiently distinctive to determine the higher order dichotomies in the eukaryote tree (Philippe and Adoutte, 1998). The multifurcation of eukaryotes, the so-called crown (Knoll, 1992), due to short branch lengths in phylogenies based on SSU rDNA, may on the other hand be explained by lack of resolving power or limited amount of data. A more promising approach to elucidate protistan phylogeny is to obtain additional gene sequences, use several genes concatenated, better algorithms to model sequence evolution for reconstructing phylogenies, and ultimately information from complete genome studies.

Phylogenetic analysis of investigated organisms

The sparse SSU rDNA sequence data do not effectively ensure a good representation of the morphological and molecular diversity thought to be present within "Cercozoa". This is especially true for the thaumatomonad flagellates, which are represented by only one genus out of seven. The bias in taxon sampling raises a potential risk of failing to infer the correct relationships among the "cercozoan" taxa. However, Rosenberg and Kumar (2001) demonstrated that improved taxon sampling will not necessarily change the position of a certain group.

Cavalier-Smith and Chao (1996/97) suggested that the genus *Cercomonas* is paraphyletic, as *Heteromita* and *Thaumatomonas* seemed to originate from within *Cercomonas*. Analysis of three additional *Cercomonas* SSU rDNA sequences included in this paper supports the splitting of *Cercomonas* into two separate genera. Consequently, we erect here a new genus *Neocercomonas*. Despite the genetic difference between *Neocercomonas* and *Cercomonas*, their morphologies are superficially similar, and we are not in a position to set up morphological criteria to separate them. Clearly, more morphological work, in combination with molecular work, is needed.

The genus *Cercomonas* is in severe need of revision. Despite more than 150 years of study, species distinction within the genus is still very difficult. There may be several reasons for this, most importantly, taxonomists have failed to properly distinguish new taxa from pre-existing ones, possibly due to the high phenotypic variability within the genus, and this has resulted in the absence of a sound and applicable species concept. Moreover, due to the plastic morphology it is difficult to point out good characters for species distinction using only light microscopy. Hence, different species within the group are probably often lumped together.

The same is true for strains of *Heteromita* that have been studied even less. From the abundance and wide distribution of *Heteromita* in different terrestrial habitats, the genus should be expected to harbour significant genetic variability, maybe as high as that in *Cercomonas*. A large insertion of ~600 bp was found in the sequence of our strain of *Heteromita globosa*. Such insertions are commonly encountered in protists. Many protozoa, fungi and green algae have introns in their SSU ribosomal genes (Gargas et al., 1995; Castlebury and Domier, 1998). It has been suggested that organisms that feed by phagocytosis may acquire self-splicing group 1 introns by horizontal transmission (Johansen et al., 1992), and that such transmissions can be relatively recent evolutionary events (Logsdon et al., 1995; Stoltzfus et al., 1997). Since, apart from the insert, the new strain of *H. globosa* shares 99% homology with the CCAP 1961/2 *H. globosa* sequence, we interpret the insertion as a result of a relatively recent evolutionary event.

The four SSU rDNA sequences from *Thaumatomonas* represent only 2% genetic variability, which strongly suggests that the four organisms make up a monophyletic branch. The synapomorphic character uniting the thaumatomonad flagellates is secretion of inorganic external scales over the body surface from vesicles associated with the mitochondria (Patterson, 1999). The scales can only be recognised in whole mount preparations using electron microscopy (EM). Currently seven flagellate genera are assigned to this group (Patterson and Zölffel, 1991), although many have not yet been studied by EM. Provided that the unique external scales are a true synapomorphic character, the six genera that have yet to be subjected to sequencing studies, namely *Thaumatomastix*, *Protaspis*, *Rhizaspis*, *Synoikomonas*, *Gyromitus*, and *Hyaloselene*, will also belong within the *Cercomonas*/*Heteromita*/*Thaumatomonas* clade. *Synoikomonas* differs from the remaining thaumatomonads in having plastids that are probably derived by primary endosymbiosis with a cyanobacterial organism (Patterson and Zölffel, 1991).

Several additional taxa, from which molecular data have not yet been obtained, have been suggested on the grounds of morphological similarity to be members of the "cercozoan" complex (Cavalier-Smith, 1998b, 1999; Patterson, 1999), for example, *Bodomorpha*, *Allas*, *Metromonas*, *Gymnophrys*, *Biomyxa*, *Spongomonas*, and *Rhipidodendron*. At least *Bodomorpha*, which resembles *Heteromita*, and *Allas* (resembling *Thaumatomonas*) are probably close relatives of members of the *Cercomonas*/*Heteromita*/*Thaumatomonas* clade. The

latter conclusion is supported by the findings of Cavalier-Smith and Chao (2003), which confirm the position of *Bodomorpha* and *Allas* near the *Cercomonas*/*Heteromita*/*Thaumatomonas* clade. However, their isolates of *Bodomorpha* and *Allas* have not been characterised morphologically, hence it is not possible finally to decide about this matter. Recent data strongly suggest that two flagellates with external morphologies quite different from "traditional" cercomonad flagellates are close relatives of *Heteromita*; these are *Proleptomonas faecicola* with an external appearance much like the trypanosomatid *Leptomonas* (Vickerman et al., 2002), and *Katabia gromovi* which bears a strong superficial resemblance to the genus *Katablepharis* (Karpov et al., 2004). It has been suggested that a vast number of flagellate and amoeboid taxa will reveal affinity to the "cercozoan clade" (Cavalier-Smith, 1998a; Vickerman, 1998; Patterson, 1999). While inclusion of many of these genera still needs to be validated by molecular data, the hypothesis was supported by the findings of Wylezich et al. (2002), who showed that a number of testate amoebae are closely related to the *Cercomonas*/*Heteromita*/*Thaumatomonas* clade. Unfortunately, the mutual relationships within the *Cercomonas*/*Heteromita*/*Thaumatomonas* clade were not resolved by any analysis in this paper, nor in the analyses by Vickerman et al. (2002). Hence, there is a clear need for additional sequencing studies of more organisms and more genes, as well as a need for more work which includes both morphological and molecular characters.

In this study, we attempted to overcome the problem of having different rates of substitution among sites in the SSU rDNA by application of the "substitution rate calibration". Even when the alignment was based on information from the secondary structure this method unfortunately did not provide better support in terms of bootstrap values for the topology for the deep branches. Hence, it is questionable if SSU rDNA gene sequences on their own are appropriate for phylogenetic studies of the deep branches in the tree of life.

Acknowledgements

We wish to thank Louise Buchholt Petersen for helpful supervision on the sequencing procedure, and Charlotte Hansen for technical assistance. The constructive criticism of David J. Patterson and the Editor substantially improved the manuscript. Part of this work was supported financially by the 'Center for Biological Processes in Contaminated Soil and Sediments' (BIOPRO), established under the Danish Environmental Research Program. F. Ekelund was supported by SNF grants 21-01-0464 and 21-02-0543.

References

Atkins, M.S., Teske, A.P., Anderson, O.R., 2000. A survey of flagellate diversity at four deep-sea hydrothermal vents in the Eastern Pacific Ocean using structural and molecular approaches. J. Eukaryot. Microbiol. 47, 400–411.

Bhattacharya, D., Helmchen, T., Melkonian, M., 1995. Molecular evolutionary analyses of nuclear-encoded small subunit ribosomal RNA identify an independent rhizopod lineage containing the Euglyphina and the Chlorarachniophyta. J. Eukaryot. Microbiol. 42, 65–69.

Burki, F., Berney, C., Pawlowski, J., 2002. Phylogenetic position of *Gromia oviformis* Dujardin inferred from nuclear-encoded small subunit ribosomal DNA. Protist 153, 251–260.

Castlebury, L.A., Domier, L.L., 1998. Small subunit ribosomal RNA gene phylogeny of *Plasmodiophora brassicae*. Mycologia 90, 102–107.

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.

Cavalier-Smith, T., 1998a. Neomonada and the origin of animals and fungi. In: Coombs, G.H., Vickerman, K., Sleigh, M.A., Warren, A. (Eds.), Evolutionary Relationships Among Protozoa. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 375–407.

Cavalier-Smith, T., 1998b. A revised six-kingdom system of life. Biol. Rev. 73, 203–266.

Cavalier-Smith, T., 1999. Zooflagellate phylogeny and the systematics of Protozoa. Biol. Bull. 196, 393–396.

Cavalier-Smith, T., Chao, E.E., 1996/97. Sarcomonad ribosomal RNA sequences, rhizopod phylogeny, and the origin of euglyphid amoebae. Arch. Protistenkd. 147, 227–236.

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

De Saedeleer, H., 1931. Notes de Protistologie IV.-*Thaumatomonas lauterborni* n.g., n. sp. flagellé nouveau. Recl Inst. zool. Torley-Rousseau. 3, 89–103.

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

Ekelund, F., Rønn, R., Griffiths, B.S., 2001. Quantitative estimation of flagellate community structure and diversity in soil samples. Protist 152, 301–314.

Felsenstein, J., 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. J. Mol. Evol. 17, 368–376.

Felsenstein, J., 1985. Confidence limits on phylogenies: an approach using the bootstrap. Evolution 39, 783–791.

Felsenstein, J., Kishino, H., 1993. Is there something wrong with the bootstrap on phylogenies? A reply. Syst. Biol. 42, 193–200.

Foissner, W., 1991. Diversity and ecology of soil flagellates. In: Patterson, D.J., Larsen, J. (Eds.), The Biology of Free-Living Heterotrophic Flagellates, Vol. 45. Systematics Association, Clarendon Press, Oxford, pp. 93–112.

Foissner, W., 2002. Neotypification of protists, especially ciliates (Protozoa, Ciliophora). Bull. Zool. Nom. 59, 165–169.

Gargas, A., Depriest, P.T., Tailor, J.W., 1995. Positions of multiple insertions in SSU rDNA of lichen-forming fungi. Mol. Biol. Evol. 12, 208–218.

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.

Hillis, D.M., Bull, J.J., 1993. An empirical test of bootstrapping as a method for assessing confidence in phylogenetic analysis. Syst. Biol. 42, 182–192.

Hillis, D.M., Huelsenbeck, J.P., Cunningham, C.W., 1994. Application and accuracy of molecular phylogenies. Science 264, 671–677.

Ishida, K., Green, B.R., Cavalier-Smith, T., 1999. Diversification of a chimeric algal group, the chlorarachniophytes: phylogeny of nuclear and nucleomorph small-subunit rRNA genes. Mol. Biol. Evol. 16, 321–331.

Johnsen, K., Enger, Ø., Jacobsen, C.S., Thirup, L., Torsvik, V., 1999. Quantitative selective PCR of 16S ribosomal DNA correlates well with selective agar plating in describing population dynamics of indigenous *Pseudomonas* spp. in soil hot spots. Appl. Environ. Microbiol. 65, 1786–1789.

Johansen, S., Johansen, T., Haugli, F., 1992. Structure and evolution of myxomycete nuclear group I introns: a model for horizontal transfer by intron homing. Curr. Genet. 22, 297–304.

Karpov, S.A., Ekelund, F., Moestrup, Ø., 2004. *Katabia gromovi* nov. gen., nov sp.—a new soil flagellate with affinities to *Heteromita* (Cercomonadida). Protistology 3, 30–41.

Keeling, P.J., Deane, J.A., McFadden, G.I., 1998. The phylogenetic position of alpha- and beta-tubulins from the *Chlorarachnion* host and *Cercomonas* (Cercozoa). J. Eukaryot. Microbiol. 45, 561–570.

Knoll, A.H., 1992. The early evolution of eukaryotes: a geological perspective. Science 256, 622–628.

Kumar, S., Rzhetsky, A., 1996. Evolutionary relationships of eukaryotic kingdoms. J. Mol. Evol. 42, 183–193.

Logsdon Jr., J.M., Tyshenko, M.G., Dixon, C., Jafari, J.D., Walker, V.K., Palmer, J.D., 1995. Seven newly discovered intron positions in the triose-phosphate isomerase gene: evidence for the introns-late theory. Proc. Natl. Acad. Sci. USA 92, 8507–8511.

MacDonald, C.M., Darbyshire, J.F., Ogden, C.G., 1977. The morphology of a common soil flagellate, *Heteromita globosa* Stein (Mastigophorea: Protozoa). Bull. Br. Mus. Nat. Hist. (Zool.) 31, 255–264.

McFadden, G.I., Gilson, P., Hofmann, C.J.B., Adcock, G.A., Maier, U.G., 1994. Evidence that an amoeba acquired photosynthesis by retaining parts of an engulfed eukaryotic alga. Proc. Natl. Acad. Sci. USA 91, 3690–3694.

Mylnikov, A.P., 1985. New species of colourless flagellate *Cercomonas minimus*. Inf. Byull. Bot. Vnutrennich Vod 66, 17–18 (in Russian).

Mylnikov, A.P., Karpov, S.A., 1993. A new representative of colourless flagellates, *Thaumatomonas seravini* sp. n. (Thaumatomonadida, Protista). Zool. Zh. 73, 5–10 (in Russian).

Nielsen, T.H., Christophersen, C., Anthoni, U., Sørensen, J., 1999. Viscosinamide, a new cyclic depsipeptide with surfactant and antifungal properties produced by *Pseudomonas fluorescens* DR54. J. Appl. Microbiol. 86, 80–90.

Page, F.C., 1988. A New Key to Freshwater and Soil Gymnoamoebae. Freshwater Biological Association, Ambleside, UK.

Patterson, D.J., 1999. The diversity of eukaryotes. Am. Nat. 154, 96–124.

Patterson, D.J., Zölffel, M., 1991. Heterotrophic flagellates of uncertain taxonomic position. In: Patterson, D.J., Larsen, J. (Eds.), The Biology of Free-Living Heterotrophic Flagellates. Clarendon Press, Oxford, pp. 427–475.

Philippe, H., Adoutte, A., 1998. The molecular phylogeny of eukaryota: solid facts and uncertainties. In: Coombs, G.H., Vickerman, K., Sleigh, M.A., Warren, A. (Eds.), Evolutionary Relationships Among Protozoa. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 25–56.

Robertson, M., 1928. Notes on *Heteromita* (Protozoa, Flagellata). Parasitology 20, 10–24.

Rosenberg, M.S., Kumar, S., 2001. Incomplete taxon sampling is not a problem for phylogenetic inference. Proc. Natl. Acad. Sci. USA 98, 10751–10756.

Rzhetsky, A., Nei, M., 1994. Unbiased estimates of the number of nucleotide substitutions when substitution rate varies among different sites. J. Mol. Biol. 38, 295–299.

Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4, 406–425.

Sandon, H., 1927. The Composition and Distribution of the Protozoan Fauna of the Soil. Oliver and Boyd, Edinburgh and London.

Sogin, M.L., 1989. Evolution of eukaryotic microorganisms and their small subunit ribosomal RNAs. Am. Zool. 29, 487–499.

Sogin, M.L., Hinkle, G., 1997. Common measures for studies of biodiversity: molecular phylogeny in the eukaryotic microbial world. In: Reaka-Kudla, M.L., Wilson, D.E., Wilson, E.O. (Eds.), Biodiversity, II. Understanding and Protecting Our Biological Resources. Joseph Henry Press, Washington, DC, pp. 109–122.

Stoltzfus, A., Logsdon Jr., J.M., Palmer, J.D., Doolittle, W.F., 1997. Intron "sliding" and the diversity of intron positions. Proc. Natl. Acad. Sci. USA 94, 10739–10744.

Sullivan, J., Holsinger, K.E., Simon, C., 1996. The effect of topology on estimates of among-site rate variation. J. Mol. Evol. 42, 308–312.

Swofford, D.L., 2003. PAUP: Phylogenetic Analysis Using Parsimony (*and Other Methods). Version 4. Sinauer Associates, Sunderland, London.

Swofford, D.L., Olsen, G.J., Waddell, P.J., Hillis, D.M., 1996. Phylogenetic inference. In: Hillis, D.M., Moritz, C., Mable, B.K. (Eds.), Molecular Systematics. Sinauer, Sunderland, London, pp. 407–514.

Van de Peer, Y., De Wachter, R., 1994. TREECON for Windows: a software package for the construction and drawing of evolutionary trees for the Microsoft windows environment. Comput. Appl. Biosci. 10, 569–570.

Van de Peer, Y., De Wachter, R., 1997. Evolutionary relationships among the eukaryotic crown taxa taking into account site-to-site variation in 18S rRNA. J. Mol. Evol. 45, 619–630.

Van de Peer, Y., Neefs, J.-M., De Rijk, P., De Wachter, R., 1993a. Evolution of eukaryotes as deduced from small ribosomal subunit RNA sequences. Biochem. Syst. Ecol. 21, 43–55.

Van de Peer, Y., Neefs, J.-M., De Rijk, P., De Wachter, R., 1993b. Reconstructing evolution from eukaryotic small-ribosomal-subunit RNA sequences: calibration of the molecular clock. J. Mol. Evol. 37, 221–232.

Van de Peer, Y., Rensing, S.A., Maier, U.-G., De Wachter, R., 1996a. Substitution rate calibration of small subunit ribosomal RNA identifies chlorarachniophyte endosymbionts as remnants of green algae. Proc. Natl. Acad. Sci. 93, 7732–7736.

Van de Peer, Y., Van der Auwera, G., De Wachter, R., 1996b. The evolution of stramenopiles and alveolates as derived by "substitution rate calibration" of small ribosomal subunit RNA. J. Mol. Evol. 42, 201–210.

Van de Peer, Y., Jansen, J., De Rijk, P., De Wachter, R., 1997. Database on the structure of small ribosomal subunit RNA. Nucl. Acids Res. 25, 111–116.

Van de Peer, Y., De Rijk, P., Wuyts, J., Winkelmans, T., De Wachter, R., 2000. The European small subunit ribosomal RNA database. Nucl. Acids Res. 28, 175–176.

Van der Auwera, G., De Wachter, R., 1997. Complete large subunit ribosomal RNA sequences from the heterokont algae *Ochromonas danica*, *Nannochloropsis salina*, and *Tribonema aequale*, and phylogenetic analysis. J. Mol. Evol. 45, 84–90.

Vickerman, K., 1998. Revolution among the protozoa. In: Coombs, G.H., Vickerman, K., Sleigh, M.A., Warren, A. (Eds.), Evolutionary Relationships Among Protozoa. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 1–24.

Vickerman, K., Le Ray, D., Hoef-Emden, K., De Jonckheere, J., 2002. The soil flagellate *Proleptomonas faecicola*: cell organisation and phylogeny suggest that the only described free-living trypanosomatid is not a kinetoplastid but has cercomonad affinities. Protist 153, 9–24.

Wainwright, P.O., Hinkle, G., Sogin, M.L., Stickel, S.K., 1993. Monophyletic origins of the Metazoa: An evolutionary link with the fungi. Science 260, 340–342.

Willerslev, E., Hansen, A.J., Christensen, B., Steffensen, J.P., Arctander, P., 1999. Diversity of Holocene life forms in fossil glacier ice. Proc. Natl. Acad. Sci. USA 96, 8017–8021.

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

Yang, Z., 1996. Among-site rate variation and its impact on phylogenetic analyses. TREE 11, 367–372.

Yang, Z., Roberts, D., 1995. On the use of nucleic acid sequences to infer early branching in the tree of life. Mol. Biol. Evol. 12, 451–458.

Zaman, V., Zaki, M., Howe, J., Ng, M., Leipe, D.D., Sogin, M.L., Silberman, J.D., 1999. *Hyperamoeba* isolated from human feces: description and phylogenetic affinity. Eur. J. Protistol. 35, 197–207.