Published in Environmental Microbiology 11, issue 1, 47-55, 2008
which should be used for any reference to this work

Pan-oceanic distribution of new highly diverse clades of deep-sea diplonemids

Enrique Lara,[1] David Moreira,[1] Alexander Vereshchaka[2] and Purificación López-García[1]*

[1] *Unité d'Ecologie, Systématique et Evolution, UMR CNRS 8079, Université Paris-Sud, bâtiment 360, 91405 Orsay Cedex, France.*
[2] *Institute of Oceanology of the Russian Academy of Sciences, 117997 Moscow, Russia.*

Summary

Molecular rRNA gene surveys reveal a considerable diversity of microbial eukaryotes in different environments. Even within a single clade, the number of distinct phylotypes retrieved often goes beyond previous expectations. Here, we have used specific 18S rRNA PCR primers to investigate the diversity of diplonemids, a poorly known group of flagellates with only a few described species. We analysed surface and deep-sea plankton samples from different oceanic regions, including the water-column in the Marmara Sea. We retrieved a large diversity of diplonemid phylotypes, most of which formed two novel distinct clades without cultured representatives. Although most marine diplonemid phylotypes appeared to be cosmopolitan, they showed a marked stratified distribution through the water column, being very scarce or absent in surface waters. The small and specific diplonemid diversity found in surface samples and the fact that most sequences of uncultured diplonemids found in other studies came from deep-sea environments suggest that the two major uncultured diplonemid clades group species preferentially inhabit the deep ocean.

Introduction

Environmental surveys of eukaryotic diversity based on the amplification, cloning and sequencing of the small subunit ribosomal RNA gene (18S rRNA) have revealed considerable protist diversity during the last 10 years. Since the first exploratory work of Van Hannen and colleagues in experimental continuous flow systems (Van Hannen *et al*., 1999), new studies regularly increase the database of environmental eukaryotic DNA sequences. In certain cases, sequences which had no affinity with any previously identified organism have been found, and could thus potentially derive from novel high-level taxonomic eukaryotic subdivisions (López-García *et al*., 2003; Berney *et al*., 2004). Even within supposedly well-sampled groups such as the ciliates, 18S rRNA gene sequences have revealed an unexpectedly high diversity (Šlapeta *et al*., 2005). Moreover, if the DNA extracted from the environment is amplified with PCR primers targeting specifically a particular eukaryotic group, this diversity turns out to be even much higher. In the case of the Cercozoa, an assemblage of mainly phagotrophic flagellated and amoeboid protists, the application of such an approach multiplied the number of known cercozoan sequences and revealed the existence of many previously undetected clades (Bass and Cavalier Smith, 2004). This result is remarkable, as sequences from this group were never found to dominate any environmental eukaryotic clone library built using general eukaryotic-specific primers. It is thus reasonable to think that the situation could possibly be similar in other eukaryotic clades, for which the small amount of sequences retrieved from identified cultures or from the environment would represent only a small portion of the total biodiversity.

Diplonemea (the diplonemids) are a clade of heterotrophic flagellates. They have recently drawn much attention because of their highly unusual mitochondrial genome divided in more than a hundred chromosomes (Marande *et al*., 2005), and also because of the peculiar way their mitochondrial RNA is edited (Marande and Burger, 2007). Molecular phylogeny places them as sister group of the Kinetoplastida; together with the Euglenida, they form the taxon Euglenozoa (Maslov *et al*., 1999; Simpson and Roger, 2004). To date, only two genera of diplonemids have cultured representatives; these are *Diplonema* and *Rhynchopus*. They are small, marine free-living phagotrophic forms, although some *Rhynchopus* strains appear to have at least one life stage as parasites of crustaceans (Von der Heyden *et al*., 2004). It has been suggested that another genus, *Hemistasia*, could possibly belong to this taxon, because it shares some ultrastructural features with *Diplonema* and *Rhynchopus* (Simpson, 1997), but so far no DNA sequence is available for members of this genus. In addition to these described

Table 1. List of the marine samples analysed in this study.

Sample	Location	Coordinates		Depth (m)	Total clones	Number of phylotypes[a]	Number of singletons
Ma110	Marmara Sea	40°52′N	28°09′E	1250	21	6	2
Ma115	Marmara Sea	40°52′N	28°09′E	1000	43	16	11
Ma121	Marmara Sea	40°52′N	28°09′E	500	36	15	7
Ma126	Marmara Sea	40°52′N	28°09′E	100	43	20	10
Ma131	Marmara Sea	40°52′N	28°09′E	25	46	12	8
Ma136	Marmara Sea	40°52′N	28°09′E	15	26	3	2
KM4	Ionian Sea	36°20′N	16°00′E	3000	43	18	9
DH17	South Atlantic	62°23′S	53°36′W	5	0	0	0
DH20	South Atlantic	62°23′S	53°36′W	20	0	0	0
DH113	South Atlantic	54°59′S	58°22′W	5	0	0	0
DH116	South Atlantic	54°59′S	58°22′W	100	0	0	0
DH117	South Atlantic	54°59′S	58°22′W	1000	45	19	16
DH134	South Atlantic	57°31′S	56°52′W	100	0	0	0
DH136	South Atlantic	57°31′S	56°52′W	500	44	25	14
DH182	South Atlantic	64°18′S	61°55′W	25	0	0	0
R26	North Atlantic	33°13′N	33°54′W	2280	45	20	11
BS16	North Atlantic	29°08′N	43°13′W	3000	46	18	8
Bi2	North Atlantic	48°10′N	16°12′E	3000	29	18	11
Ti7	North Atlantic	41°46′N	50°14′E	3000	17	9	5
Ti11	North Atlantic	41°46′N	50°14′E	3500	19	11	5
Ti12	North Atlantic	41°46′N	50°14′E	3000	16	9	5
PHC3	East Pacific Rise	12°49′N	103°56′W	1695	35	3	0

a. Phylotypes being considered as groups of sequences having more than 99% identity at the 18S rRNA gene.
Geographic location, depth, total number of diplonemid clones sequenced and the number of different phylotypes retrieved per sample are indicated as well as the number of phylotypes appearing only once in each clone library (singletons).

bona fide diplonemids, there is evidence of the existence of uncultured organisms branching in 18S rRNA gene phylogenies as a distant sister group to *Diplonema* and *Rhynchopus*. First, an environmental sequence related to the classical diplonemids was retrieved from deep-sea (3000 m) plankton at the Antarctic Polar Front (López-García *et al*., 2001). Subsequently, other sequences branching with it and defining a diplonemid-related group came from the fluid–seawater interface at the Lost City hydrothermal vents, another pelagic deep-sea environment (López-García *et al*., 2007).

In this study, we have investigated the extent of marine diplonemid diversity by amplifying environmental 18S rRNA genes with diplonemid-specific primers. We studied samples from different oceanic regions to test whether their geographic origin could influence the distribution of the phylotypes encountered. Moreover, we studied the repartition of the phylotypes at different depths through the water column in the Marmara Sea (from 15 m down to 1250 m) to get insight about the potential stratification of different diplonemid phylotypes.

Results and discussion

Detection of new diplonemid clades in different oceanic regions

Previous 18S rRNA gene surveys of deep-sea environments had revealed sequences that formed a distant, though related, group to the sequences of cultivated diplonemid species. These sequences constituted only a small fraction of the total eukaryotic diversity retrieved (López-García *et al*., 2007). In order to confirm the existence of this group as sister to the classical diplonemids, and to explore the extent of its diversity and its distribution in oceans, we designed a pair of specific primers targeting both the classical diplonemids and the recently identified group sister to them (see *Experimental procedures*). We looked for the presence of diplonemids in marine samples including plankton from different deep-sea and surface waters (Marmara Sea, Ionian Sea, South and North Atlantic), as well as a sample from an artificial colonization substrate (East Pacific Rise). We also tested freshwater samples, including a suboxic pond and a peat bog. PCR products were obtained for all deep-sea samples (including the colonization substrate), and also some surface samples (15, 25 and 100 m from the Marmara Sea; see Table 1). However, no amplicon could be retrieved for two South Atlantic euphotic zone samples (Table 1), as well as for freshwater samples.

Initially, we obtained partial sequences from a total of 554 clones (Table 1). All of them belonged to diplonemids, confirming the specificity of the primers used. Partial sequences were trimmed for ambiguities and used to construct a local database. Clones exhibiting more than four differences between themselves in the 400 first nucleotides (which comprise the most variable

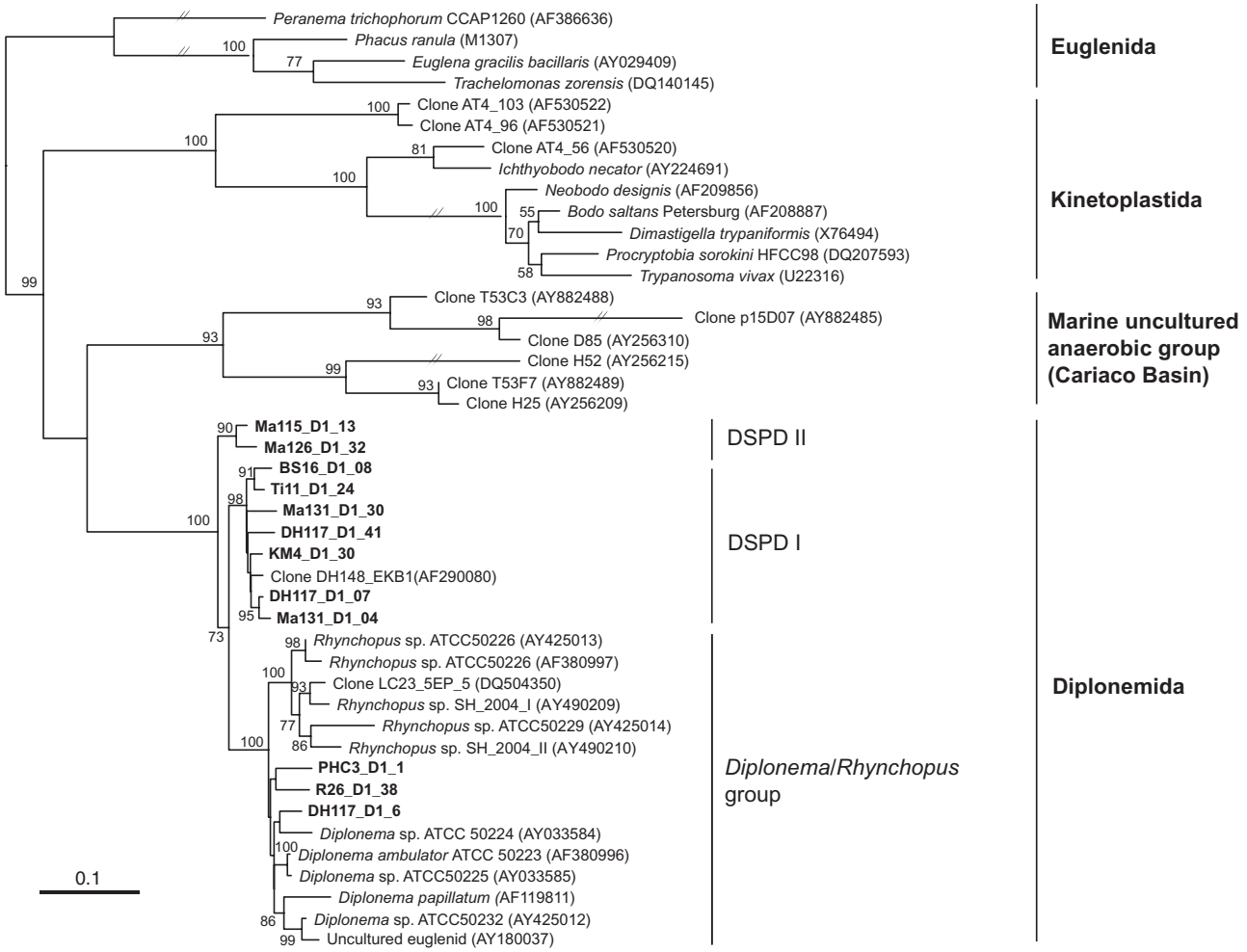

Fig. 1. 18S rRNA maximum likelihood phylogenetic tree illustrating the position of the novel deep-sea clades related to the *Diplonemida* within the *Euglenozoa*. Clone names in bold represent sequences obtained in this study. The numbers at the nodes indicate bootstrap support (values under 70% were omitted). The length of branches indicated with a double barred sign was reduced by half.

part of the 18S rRNA gene) were considered as different and treated as separate phylotypes. Thus, we considered as independent phylotypes groups of sequences sharing more than 99% sequence identity, resulting in a total of 95 different phylotypes that were fully sequenced. The total length of the sequenced 18S rRNA fragment was generally of about 1200 bp, more in the presence of insertions. These ranged from 10 (Bi_D1_30) to 464 nucleotides (Ma121_D1_12). In addition, a class I intron was encountered in clone DH117_D1_35.

A first phylogenetic analysis was performed in order to place the new diplonemid sequences in a global euglenozoan tree. In addition to 12 representative new diplonemid sequences from this study, it comprised selected sequences from other Diplonemida, Euglenida and Kinetoplastida, as well as some environmental sequences from the anoxic deep-sea Cariaco Basin, which had been tentatively assigned to diplonemids (Stoeck *et al*., 2003; 2006); the tree was rooted with the Euglenida. A total of 44 sequences and 825 characters were analysed (Fig. 1). The maximum likelihood (ML) phylogenetic analysis of this data set supported the monophyly of diplonemids, including the marine environmental groups, with maximum bootstrap support (BS 100%). However, our analysis did not retrieve a significant statistical support for the sisterhood of the sequences from Cariaco basin and the diplonemids. Given the particularly long branches displayed by the Cariaco basin environmental sequences, it is highly probable that their phylogenetic position is influenced by a long-branch attraction artefact, so that their position with respect to the *bona fide* diplonemids remains unclear (Epstein and López-García, 2008). Within diplonemids, three distinct, well-supported clades were observed: one comprising sequences of cultured repre-

sentatives from the described genera *Diplonema* and *Rhynchopus* (as well as some environmental sequences) plus two other ones, comprising only environmental sequences. These were designated here DSPD I and II (deep-sea pelagic Diplonemids I and II; see Fig. 1). In this analysis, DSPD II branched at the base of a group composed by DSPD I and the *Diplonema*/*Rhynchopus* group (BS 73%). To improve the resolution of the internal phylogeny of the diplonemids, including the two new groups DSPD I and II, we carried out a second ML phylogenetic analysis that comprised all distinct phylotypes obtained in this work, plus all available full-length 18S rRNA gene sequences belonging to Diplonemida in the database. To maximize the number of sites retained for the analysis, we used as outgroup five slow-evolving kinetoplastid sequences (kinetoplastids being the closest relatives to diplonemids). A total of 1150 positions were taken into account (Fig. 2). In contrast with the previous analysis, DSPD I and II branched together (BS 72%). As the number of sites analysed here is higher than in the first analysis (1150 versus 825) and as the long-branch attraction effect was alleviated by the use of the closest possible outgroup available, the topology of this phylogenetic tree is likely more reliable. Within the *Diplonema*/*Rhynchopus* group, the genus *Rhynchopus* appeared monophyletic by contrast to *Diplonema*.

The diversity of the diplonemid sequences retrieved from our samples appeared to be extremely high when compared with the relatively low number of sequences from this group identified in previous published surveys using eukaryote-general primers from similar deep-sea pelagic environments. The clone libraries DH117 and DH136 (South Atlantic) and Bi2 (North Atlantic) presented a relatively high amount of singletons (i.e. phylotypes appearing only once in a given clone library; see Table 1), accounting for 36%, 32% and 38% respectively; their rarefaction curves did not reach a plateau either (data not shown). This indicates that the sequence diversity captured in these libraries is likely to be even higher than our present estimations. Most of the sequences obtained in this study clustered within the group DSPD I (81 different phylotypes, including several from previous environmental surveys), suggesting that this group is relatively abundant in various deep-sea locations. DSPD II has been detected in this study for the first time and can be considered as a novel deep-branching diplonemid clade. This illustrates the fact that the use of taxon-specific primers can reveal the occurrence of a previously unsuspected diversity, confirming similar observations within the Cercozoa for which the use of taxon-specific primers revealed also an unexpected diversity, including several novel clades (Bass and Cavalier Smith, 2004). Furthermore, these studies also suggest that the extent of protist diversity, as evaluated by molecular tools, is far larger than that estimated from the already available 18S rRNA gene surveys carried out with general eukaryotic primers (Massana *et al*., 2004; Countway *et al*., 2005; Zuendorf *et al*., 2006).

By contrast to the planktonic samples, where groups DSPD I and DSPD II dominated (altogether 98.7% of all the sequences obtained), the phylotype composition of an environmental library derived from a colonization substrate deposited on the sea bottom (PHC3, 1695 m) differed strongly, with 83% of the total number of clones belonging to the *Diplonema*/*Rhynchopus* group (Table 1). Moreover, no phylotype from this library was found in those from planktonic samples. This biased composition towards the *Diplonema*/*Rhynchopus* group suggest that deep-sea members of this clade are most likely benthic, in contrast to the DSPD I and II, which are most likely planktonic.

In spite that we used a very narrow definition of phylotype (18S rRNA gene sequences having >99% identity), diplonemid sequences belonging to the same phylotype were frequently found in clone libraries derived from several samples originated from distant geographical regions (Fig. 2). For instance, some phylotypes, such as Ma115_D1_13 were encountered in deep-sea environments from the Ionian and the Marmara Seas as well as in the North and South Atlantic, despite important differences in the physico-chemical parameters of the corresponding water masses. Whereas Mediterranean and Marmara deep waters have warm temperatures (*c*. 14°C) and high salinity (38–39‰), the open deep ocean has low average temperatures (2°C) and lower salinity (*c*. 35‰). The presence of identical phylotypes in clone libraries derived from samples collected in geographically distant environments and very different water masses suggests a pan-oceanic distribution of diplonemids. This would be consistent with the general idea that many marine pelagic protists are cosmopolitan due to the lack of geographic barriers in oceans, possibly due to the existence of a global oceanic circulation facilitating dispersal.

Furthermore, our results show that a high diversity of diplonemid phylotypes coexist within a single plankton sample. Several hypotheses can be invoked to explain the coexistence of various phylogenetically related organisms in an apparently homogeneous environment. If the resources used by pelagic diplonemids are different (e.g. different preys or hosts), that coexistence can be the result of a narrow specialization of different ecotypes which would therefore not be in competition among themselves. By contrast, if the resources used by diplonemids are overlapping, we would be in a 'paradox of the plankton' situation, where a limited range of resources supports a much wider range of planktonic organisms (Hutchinson, 1961). The way generally proposed to resolve this paradox is to state that fluctuating physicochemical and biological

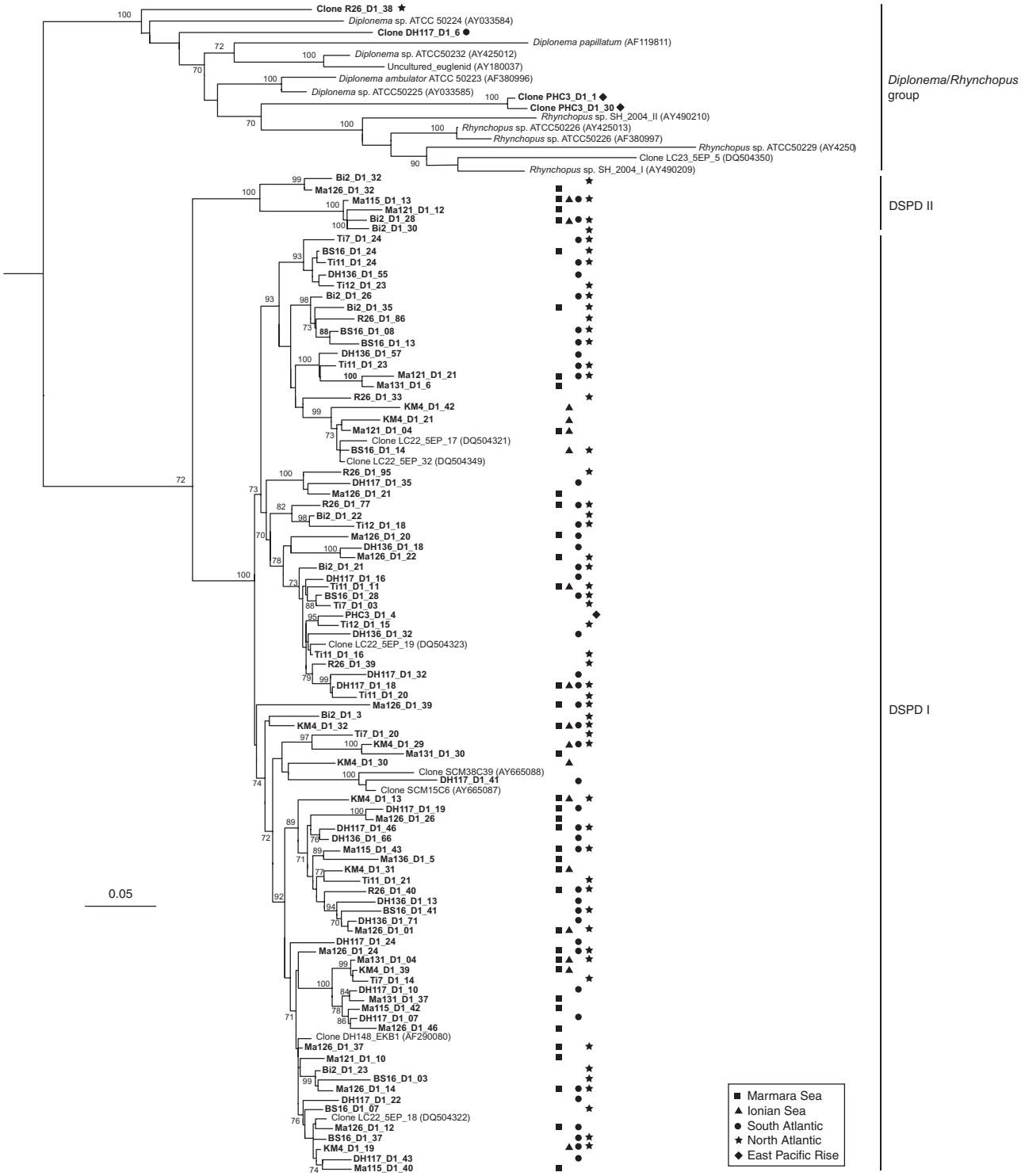

Fig. 2. 18S rRNA maximum likelihood phylogenetic tree representing the full phylogenetic diversity of the sequences obtained in this study within the diplonemids. The tree is rooted with five slow-evolving kinetoplastid sequences (not shown). Clone names in bold represent sequences obtained in this study. The numbers at the nodes indicate bootstrap support (values under 70% were omitted).

parameters (i.e. predation, parasitism) induce a chaotic situation in planktonic population dynamics which maintains high diversity (Scheffer *et al*., 2003). Whether this explanation applies to the deep-ocean, where conditions are thought to be much less fluctuating than surface waters, remains to be tested.

Stratification of diplonemid phylotypes in the water column

As sequences from the marine diplonemid clades sister to the classical *Diplonema/Rhynchopus* group had been previously nearly exclusively identified in deep-sea waters and as taxon-specific primers revealed a larger diversity of diplonemids in various deep oceanic samples, we wanted to test whether diplonemids could be also detected in surface waters using specific primers. For that purpose, we tried to amplify diplonemid genes from different surface samples, including seven different South Atlantic samples from depths ranging from 5 to 100 m and also samples from the water column at the Central Basin of the Marmara Sea (15–1250 m depth) (Table 1). Despite the use of various PCR conditions, we were unable to amplify diplonemid sequences from any surface sample from the South Atlantic. Only the surface samples from the Marmara Sea yielded an amplification product that, after cloning and sequencing, revealed much lower phylotype diversity than the aphotic part of the water column (Table 1). These results strongly suggest that most marine pelagic diplonemids are actual inhabitants of deep oceanic layers.

Diplonemid phylotypes showed a strong stratification through the Marmara water column, each detected phylotype exhibiting a preference by a particular depth range. Thus, Ma115_D1_13 was found at and below 500 m, Ma126_D1_21 between 25 m and 1000 m with a peak at 100 m and Ma131_D1_30 only at 25 m and 15 m, above the chemocline, these clones being the most common ones in the samples from Marmara and representing a high proportion of the total number of sequences retrieved from these clone libraries (Fig. 3). A strong chemocline was present at approximately 25 m depth in the Marmara Sea, with surface waters entering from the Black Sea having lower temperature (11.8°C) and salinity (26.9%) than the water mass below the chemocline (14.9°C and 38.7‰ respectively) of Aegean origin (Stashchuk and Hutter, 2001). The overall diversity of diplonemid phylotypes also appeared to support a distinction between surface and deep marine samples, as deduced from a cluster analysis based on the frequency of the different diplonemid phylotypes found in our samples (Fig. 4). In fact, the shallowest samples (15 and 25 m Marmara Sea) clustered together with a strong statistical support (95%) and a long basal branch, indicating that their diplonemid

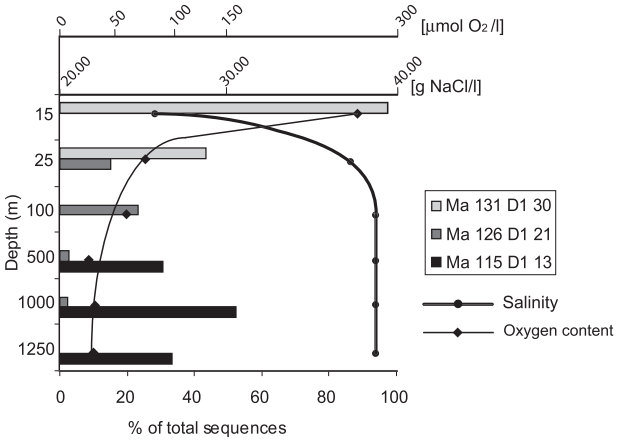

Fig. 3. Relative proportions of the three most common phylotypes identified in the Marmara Sea water column at the six different depths sampled. Salinity and oxygen content profiles as measured *in situ* during the sampling are indicated.

communities are very different from those of the other samples. The 100-m-deep Marmara Sea sample branched close to the shallow samples, in agreement with its intermediate depth between the shallow and deep (> 500 m) samples. The deep-sea samples from the Marmara water column (500, 1000 and 1250 m) clustered together with KM4 (3000 m Ionian Sea) with strong support (95%). Although we cannot completely exclude potential biases due to incomplete exploration of diplonemid diversity in our samples, all Marmara libraries, as well as KM4, contained a relatively low amount of singletons (8–26%; see Table 1). The corresponding rarefaction curves did also show a tendency to reach a plateau (data not shown). All this suggests that a large part of the diversity in our libraries was covered, and that undersampling did probably not influence much the results obtained by cluster analysis. Accordingly, though this method can also encompass some biases, the comparison of clone libraries using LIBSHUFF v 0.96 (Singleton *et al*., 2001) showed that the three libraries from deep sea environment from Marmara (Ma110, 115 and 121), as well as the two libraries from surface samples (Ma131 and 136) were not significantly different among them with a probability > 95%, in agreement with the clustering analysis (Fig. 4). This could correlate with the particular physico-chemical conditions found in the deep-sea zones of these two basins, the deep Marmara waters being of East Mediterranean origin. Nevertheless, a number of phylotypes found in deep Mediterranean and Marmara samples were also found in the North and South Atlantic (Fig. 2), which suggests that these physico-chemical parameters (particularly temperature and salinity) have a more limited effect on the diplonemid community composition than others that more generally influence stratification, such as light and pressure. As known diplonemids

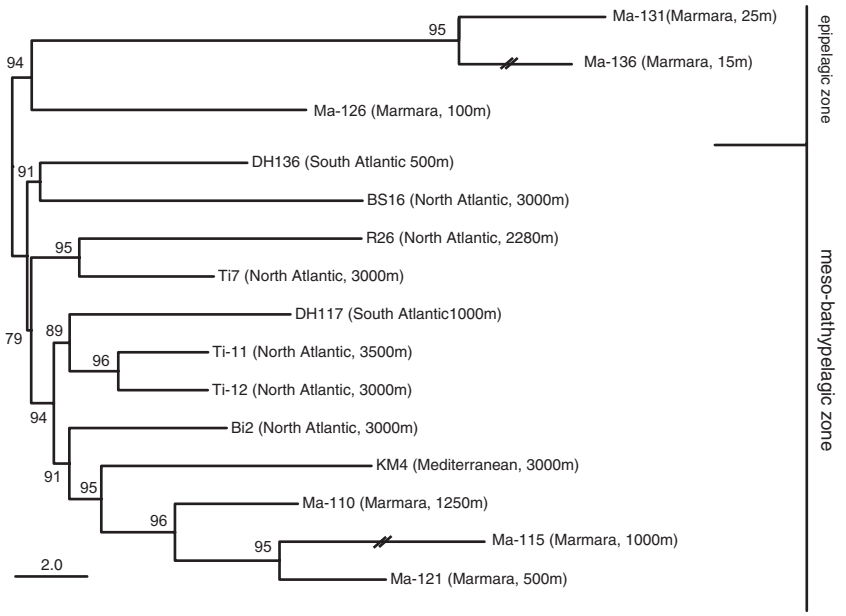

Fig. 4. Cluster analysis of the samples studied based on their diplonemid phylotype composition. The numbers at the nodes indicate jackknife values. The length of branches indicated with a double barred sign was reduced by half.

are not photosynthetic, but grazers or parasites, their stratification might be indeed dictated by the stratification of the prokaryotic community on which they may preferentially feed or the kind of hosts they parasitize to in case they were parasites.

In conclusion, our results demonstrate that, confirming previous studies (Bass and Cavalier Smith, 2004), the use of taxon-specific eukaryotic primers may reveal unsuspected levels of diversity within little explored protistan groups. By using this approach, we considerably extend the known diversity of the poorly explored diplonemids, revealing the existence of two novel phylogenetic clades that are globally distributed in deep-sea plankton. We also show that the diplonemid components of the protistan community exhibit a marked stratification in the water column, as has been observed for other protistan groups such as the Acantharea and the Polycystinea (Not *et al.*, 2007) and the Euglenozoa (Countway *et al.*, 2007). Finally, the spatial coexistence of various very closely related diplonemid phylotypes in the same planktonic samples may suggest that they are ecologically specialized, deriving resources from different preys, feeding behaviour or, perhaps, parasitized hosts.

Experimental procedures

Sampling, DNA extraction and PCR amplification

The list of samples used in this study is shown in Table 1. Most of them were part of a collection of DNA previously extracted and available in the laboratory from small-size plankton samples collected in various deep-sea environments at different locations in the North Atlantic (Bi2, Ti7-12, R26 and BS16), South Atlantic (DH117, DH136) and in the Mediterranean (KM4), as well as a sample from a colonization substrate deposited in the East Pacific Rise. In addition, one sample from a suboxic freshwater pond (Šlapeta *et al.*, 2005) as well as two samples from a peat bog (E. Lara, D. Moreira, E.A.D. Mitchell, and P. López-García, unpublished) were tested in this study. DNA extraction was previously made as described (López-García *et al.*, 2001; 2007). Samples from the Marmara Sea were retrieved aboard the Atalante in May–June 2007 during the sampling cruise 'MARNAUT' using Niskin bottles mounted on a CTD rosette from various depths from 15 to 1250 m (Ma 110–136). Concomitantly, salinity, temperature (°C) and oxygen ($mmolO_2\ l^{-1}$) measurements for each sample were obtained. Water was pre-filtered through a 30-μm-mesh-size net (Nitex filters), and the biomass subsequently collected on a 0.22-μm-pore-diameter filter (GTTP, Millipore). Filters were trimmed into *c.* 1 mm^2 fragments with a sterile razorblade and DNA was then extracted using a MoBio Power Soil™ DNA extraction kit (Carlsbad, CA, USA) following the manufacturer's instructions. In order to design primers amplifying specifically the largest possible diversity of 18S rRNA genes from diplonemids, we retrieved all sequences of diplonemids available in GenBank (both from cultured species and from the environment) and included them in a general alignment of eukaryotic sequences using the programs within the MUST package (Philippe, 1993). The primer sequences are the following: DiploF (GATATCTAAACCTGTC) and DiploR (GCATTCCTCATTCAAGGA). Their positions are, respectively, 537–552 and 1792–1809 on the sequence of *Diplonema papillatum* (AF119811). PCR reactions were carried out in 25 μl of reaction buffer containing 1 μl DNA template (~1–5 ng), 1.5 mM $MgCl_2$, dNTPs (10 nmol each), 20 pmol of each primer and 1 U Taq DNA polymerase (Promega). PCR reactions were performed under the following conditions: 35 cycles (denaturation at 94°C for 15 s, annealing at 50°C for 30 s, extension at 72°C for 2 min) preceded by 2 min denaturation at 94°C, and followed by 8 min extension at 72°C. Amplicons were cloned into pCR2.1

Topo TA cloning vector (Invitrogen) and transformed into *Escherichia coli* TOP10′ One Shot cells (Invitrogen) according to the manufacturer's instructions. Cloned inserts were amplified and sequenced using the vector primers T7 and M13. Sequencing was performed by Genome-Express (Meylan, France). Sequences were deposited in GenBank under accession numbers EU635588 to EU635682.

Phylogenetic and statistical analyses

Sequences were aligned using ClustalX (Thompson *et al*., 1997). Ambiguously aligned regions and gaps were excluded in phylogenetic analyses. All phylogenetic analyses were performed using ML phylogenetic analysis with the program Treefinder (Jobb *et al*., 2004) applying a GTR + Γ + I model of nucleotide substitution (Rodríguez *et al*., 1990), taking into account a proportion of invariable sites, and a Γ-shaped distribution of the rates of substitution among variable sites, with four rate categories. All necessary parameters were estimated from the data sets. Bootstrap values were calculated from 1000 replicates. Rarefaction analyses were done using the software DOTUR (Schloss and Handelsman, 2005). The cluster analysis of the sample community composition was carried out using the Online Clustering Calculator at http://www2.biology.ualberta.ca/jbrzusto/cluster.php. A matrix of Euclidean distances was calculated for all samples and used to construct a neighbour-joining tree, whose robustness was evaluated by a jackknife analysis (100 replicates). The sample PHC3 was excluded from the analysis owing to its lack of common phylotypes with the rest of the samples. To test whether the diversity captured in the libraries grouping together in the cluster analyses was significantly similar or not, we used the software LIBSHUFF v 0.96 (Singleton *et al*., 2001).

Acknowledgements

We are grateful to P. Henry and N. Cagatay, chief scientists of the oceanographic cruise MARNAUT (Leg 2), for providing us the opportunity to participate, as well as to the shipboard scientific team. We also acknowledge Captain and crew of R/V L'Atalante and the help of the Turkish Navy to protect our ship in the zones of heavy ship traffic. We thank INSU for providing a CTD-rosette for water sampling and L. Fichen, who operated it. We also thank F. Rodríguez-Valera for providing planktonic samples. This work was supported by an ATIP Plus grant of the French Centre National de la Recherche Scientifique (CNRS), section 'Dynamique de la biodiversité' to P.L.G. E.L. was benefited from an ATIP-associated CNRS research assistant contract.

References

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

Berney, C., Fahrni, J., and Pawlowski, J. (2004) How many novel eukaryotic kingdoms? Pitfalls and limitations of environmental DNA surveys. *Biomed Central Biol* **2:** 1–13.

Countway, P.D., Gast, R.J., Dennett, M.R., Savai, P., Rose, J.M., and Caron, D.A. (2007) Distinct protistan assemblages characterize the euphotic zone and deep sea (2500 m) of the western North Atlantic (Sargasso Sea and Gulf Stream). *Environ Microbiol* **9:** 1219–1232.

Countway, P.D., Gast, R.J., Savai, P., and Caron, D.A. (2005) Protistan diversity estimates based on 18S rDNA from seawater incubations in the western North Atlantic. *J Eukaryot Microbiol* **52:** 1–12.

Epstein, S.S., and López-García, P. (2008) 'Missing' protists: a molecular prospective. *Biodivers Conserv* **17:** 261–276.

Hutchinson, G.E. (1961) The paradox of the plankton. *Am Nat* **95:** 137–145.

Jobb, G., Von Haeseler, A., and Strimmer, K. (2004) TREEFINDER: a powerful graphical analysis environment for molecular phylogenetics. *BMC Evol Biol* **28:** 4–18.

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

López-García, P., Philippe, H., Gail, F., and Moreira, D. (2003) Autochthonous eukaryotic diversity in hydrothermal sediment and experimental microcolonizers at the Mid-Atlantic Ridge. *Proc Nat Acad Sci USA* **100:** 697–702.

López-García, P., Vereshchaka, A., and Moreira, D. (2007) Eukaryotic diversity associated with carbonates and fluid–seawater interface in Lost City hydrothermal field. *Environ Microbiol* **9:** 546–554.

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

Marande, W., Lukes, J., and Burger, G. (2005) Unique mitochondrial genome structure in diplonemids, the sister group of kinetoplastids. *Euk Cell* **4:** 1137–1146.

Maslov, D.A., Yasuhira, S., and Simpson, L. (1999) Phylogenetic affinities of *Diplonema* with the *Euglenozoa* as inferred from the *ssu* rRNA gene and partial COI protein sequences. *Protist* **150:** 33–42.

Massana, R., Balagué, V., Guillou, L., and Pedrós-Alió, C. (2004) Picoeukaryotic diversity in an oligotrophic coastal site studied by molecular and culturing approaches. *FEMS Microbiol Ecol* **50:** 231–243.

Not, F., Gausling, R., Azam, F., Heidelberg, J.F., and Worden, A.Z. (2007) Vertical distribution of picoeukaryotic diversity in the Sargasso Sea. *Environ Microbiol* **9:** 1233–1252.

Philippe, H. (1993) MUST, a computer package of Management Utilities for Sequences and Trees. *Nucleic Acids Res* **21:** 5264–5272.

Rodríguez, F., Oliver, J.L., Marin, A., and Medina, J.R. (1990) The general stochastic model of nucleotide substitution. *J Theor Biol* **142:** 485–501.

Scheffer, M., Rinaldi, S., Huisman, J., and Weissing, F.J. (2003) Why plankton communities have no equilibrium: solutions to the paradox. *Hydrobiologia* **491:** 9–18.

Schloss, P.D., and Handelsman, J. (2005) Introducing DOTUR, a computer program for defining operational taxonomic units and estimating species richness. *Appl Environ Microbiol* **71:** 1501–1506.

Simpson, A.G.B. (1997) The identity and composition of the *Euglenozoa*. *Arch Protistenkd* **148:** 318–328.

Simpson, A.G.B., and Roger, A.J. (2004) Protein phylogenies

robustly resolve the deep-level relationships within *Euglenozoa*. *Mol Phylogenet Evol* **30:** 201–212.

Singleton, D., Furlong, M.A., Rathbun, S.A., and Whitman, W.B. (2001) Quantitative comparisons of 16S rRNA gene sequence libraries from environmental samples. *Appl Environ Microbiol* **67:** 4374–4376.

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

Stashchuk, N., and Hutter, K. (2001) Modelling of water exchange through the Strait of the Dardanelles. *Continent Shelf Res* **21:** 1361–1382.

Stoeck, T., Taylor, G.T., and Epstein, S.S. (2003) Novel eukaryotes from the permanently anoxic Cariaco Basin (Caribbean Sea). *Appl Environ Microbiol* **69:** 5656–5663.

Stoeck, T., Hayward, B., Taylor, G.T., Varela, R., and Epstein, S.S. (2006) A multiple PCR-primer approach to access the microeukaryotic diversity in environmental samples. *Protist* **157:** 31–43.

Thompson, J.D., Gibson, T.J., Plewniak, F., Jeanmougin, F., and Higgins, D.G. (1997) The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* **24:** 4876–4882.

Van Hannen, E.J., Mooij, W.M., van Agterveld, M.P., Gons, H.J., and Laanbroek, H.J. (1999) Detritus-dependent development of the microbial community in an experimental system: qualitative analysis by denaturing gradient gel electrophoresis. *Appl Environ Microbiol* **65:** 2478–2484.

Von der Heyden, S., Chao, E.E., Vickerman, K., and Cavalier Smith, T. (2004) Ribosomal RNA phylogeny of bodonid and diplonemid flagellates and the evolution of euglenozoa. *J Euk Microbiol* **51:** 402–416.

Zuendorf, A., Bunge, J., Behnke, A., Barger, K.J.-L., and Stoeck, T. (2006) Diversity estimates of microeukaryotes below the chemocline of anoxic Mariager Fjord, Denmark. *FEMS Microb Ecol* **58:** 476–491.