A Multiple PCR-primer Approach to Access the Microeukaryotic Diversity in Environmental Samples

Thorsten Stoeck, Brett Hayward, Gordon T. Taylor, Ramon Varela, and Slava S. Epstein

Department of Ecology, Technical University Kaiserslautern, Kaiserslautern, Germany
Marine Science Center, Northeastern University, Nahant, MA 01908, USA
Marine Sciences Research Center, Stony Brook University, Stony Brook, NY 11794, USA
Estacion de Investigaciones Marinas de Margarita, Fundacion la Salle de Ciencias Naturales, Nueva Esparta, Venezuela
Biology Department, Northeastern University, Boston, MA 02115, USA

Submitted June 28, 2005; Accepted October 30, 2005
Monitoring Editor: Michael Melkonian

The Cariaco Basin off the Venezuelan coast in the Caribbean Sea is the world's largest truly marine body of anoxic water. The first rRNA survey of microbial eukaryotes in this environment revealed a number of novel lineages, but sampled only a fraction of the entire diversity. The goal of this study was to significantly improve recovery of protistan rRNA from the Basin. This was achieved by a systematic application of multiple PCR primer sets and substantially larger sequencing efforts. We focused on the most diverse habitat in the basin, anoxic waters roughly 100 m below the oxic—anoxic interface, and detected novel lineages that escaped the single PCR primer approach. All clones obtained proved unique. A 99% sequence similarity cut-off value combined these clones into operational taxonomic units (OTUs), over 75% of which proved novel. Some of these OTUs form deep branches within established protistan groups. Others signify discovery of novel protistan lineages that appear unrelated to any known microeukaryote. Surprisingly, even this large-scale multi-primer rRNA approach still missed a substantial part of the samples’ rRNA diversity. The overlap between the species lists obtained with different primers is low, with only 4% of OTUs shared by all three libraries, and the number of species detected only once is large (55%). This strongly indicates that, at least in anoxic environments, protistan diversity may be much larger than is commonly thought. A single sample appears to contain thousands of largely novel protistan species. Multiple PCR primer combinations may be needed to capture these species.

Introduction

The extent of microeukaryotic diversity on our planet is unknown (Patterson 1999), and so is the true extent of its ecological and evolutionary importance (Countway et al. 2005; Dawson and Pace 2002). The prevailing view is that it far exceeds the diversity known from classical works, and that molecular techniques represent our best tool for detecting the “missing” part of
microeukaryotic diversity (Baldauf 2003). Even though light microscopy approaches for documenting the diversity and distribution of protists have contributed significantly to our current knowledge, these techniques are mainly restricted to fairly large (generally >20 μm) organisms with a distinctive morphology. Additionally, the traditional approaches are hampered by the fact that most protists cannot be presently cultivated. This uncultivated majority may have escaped microscopic analyses (Dawson and Pace 2002; Moreira and López-García 2002). It was not until recently that first applications of an rRNA approach to microeukaryotic discovery provided a strong indication that an enormous reservoir of protists may indeed remain to be discovered (Amaral-Zettler et al. 2002; Countway et al. 2005; Dawson and Pace 2002; Díez et al. 2001; Edgcomb et al. 2002; López-García et al. 2001, 2003; Massana et al. 2004a,b; Moon-van der Staay et al. 2001; Romani and Vaulot 2004; Stoeck and Epstein 2003; Stoeck et al. 2003a). All these studies reported a high number of novel phylotypes that could not be linked to any morphologically described taxa. However, the extent of molecular studies of microeukaryotic diversity is still dwarfed by that in prokaryotic microbiology, and the 18S rRNA of microbial eukaryotes has been sampled only in a handful of natural environments. In parallel, the studies of 16S rRNA gene diversity abundantly showed the shortcomings of this approach, including PCR biases (Acinas et al. 2004; Caron et al. 2004). It is very reasonable to assume that a single PCR primer set does not amplify equally well the 18S rRNA gene from all protistan taxa (Caron et al. 2004), but so far only one study has considered the ramifications of such a bias (Dawson and Pace 2002), and none systematically applied multiple primer sets. Therefore, even in those few environments that have been investigated with respect to their 18S rRNA gene diversity, the sequence recovery was likely to be biased, and the degree of undersampling may be even higher than suggested by the rarefaction curves (Chao et al. 2005; Colwell et al. 2004; Heck et al. 1975). In fact, it may be so high that the results of the earlier studies, including ours, may or may not be viewed as informative inventories, and may or may not be meaningfully compared.

Recently, we applied the rRNA approach to study the diversity of microbial eukaryotes in the anoxic waters of the Cariaco Basin off the coast of Venezuela in the Caribbean Sea with a single primer set (Stoeck et al. 2003a). The findings included new clades unrelated to any known eukaryotes, as well as deep branches within established protistan groups, but this study did not escape the limitations identified above. The dominance of lineages that occurred only once in our 18S rRNA environmental clone libraries, and the biases associated with the single primer set used, made it extremely likely that we captured only an insignificant portion of the entire protistan diversity in the Cariaco Basin (Stoeck et al. 2003a). To address this shortcoming, we returned to the Cariaco Basin and re-sampled its richest habitat (anoxic waters below the oxic/anoxic interface). We doubled the amount of 18S rRNA sequence information by employing a multi-primer PCR approach together with an amplification artifact-reducing protocol to minimize the PCR biases (Acinas et al. 2004; Caron et al. 2004). As expected, we discovered a number of novel protistan lineages. However, unexpectedly, we did not come any closer to a more complete inventory. The multi-primer approach tripled the number of species detected in the environment, but it also showed that the target community was significantly more diverse than that which emerged from the traditional single primer pair-based data sets. It seems that as the detection approaches improve, the estimates of the overall diversity grow, with the diversity of protists in at least anoxic environments appearing to be truly spectacular.

Results and Discussion

The three clone libraries, constructed using the primer sets E528F/Univ1391, E528F/Univ1492, and E528F/Univ1517, contained 241, 119, and 137 clones respectively. The clustering, based on a 99% sequence identity criterion, produced 24, 31, and 36 OTUs unique to the respective library, with most of them (35—65%) detected only once. Some of these OTUs were shared between the individual libraries, decreasing the study-wide number of unique OTUs to 72. Collectively, this almost triples the number of protistan OTUs reported previously from the same depth in the Cariaco Basin (Stoeck et al. 2003a). Taken together, with over 500 clones obtained in the above study, the total of more than 1000 18S rRNA gene clones represents one of the largest protistan clone libraries obtained to date from a single environment.

Do the multiple-primer PCR and more traditional single primer pair-based PCR approaches qualitatively recover different uncultured protists or does
the use of different PCR primers simply reveal “more of the same”? Direct comparison is only possible when the respective clone libraries contain all recoverable rRNA species, but this condition is not met in this or any other studies to date. However, the two rates of recovery can be compared indirectly, e.g., by considering the proportion of species detected only once by each primer set (“singletons”). The logic is that as the size of the clone library grows, so does the number of species captured and rare species start getting recaptured, consequently decreasing the proportion of singletons. It follows that if PCR primers exhibited little selection, then the three libraries constructed here would represent different subsamples from essentially a single pool of PCR products. If so, the combined data set should contain fewer singletons than the individual (smaller) libraries. However, this is not what we observed. Figure 1 illustrates the OTU frequencies in three individual clone libraries and in the combined data set. The proportion of singletons in the former vs. the latter is essentially the same (0.53 ± 0.16 and 0.55 respectively). This means that the three libraries are likely to be subsamples from different pools of PCR products, and the likeliest explanation for this is a PCR bias. This in turn means that the three libraries obtained here, and the fourth obtained earlier (Stoeck et al. 2003a), recovered different parts of the target community. It follows that each primer set, and thus each library, bring 18S rRNA gene sequences not recoverable with other primer pairs. Incidentally, this tentative conclusion makes one specific prediction: the lists of species recovered by the individual libraries should overlap little, if at all. Indeed, only three OTUs were shared between all three clone libraries (Fig. 1).

Remarkably, even the single primer set-based approach revealed so much of uncultured protistan diversity that the existing inventories were called “the tip of the iceberg” (Moreira and López-García 2002). The level of diversity emerging from the multiple-primer approach is even larger, and points to a truly spectacular richness of protists. It now seems very likely that all 18S rRNA surveys conducted to date, most of them based on a single primer set amplification, have missed even more substantial portions of protistan biodiversity of the respective environments than are deduced from rarefaction curves.

Not surprisingly, all rRNA gene sequences obtained here are unique and do not match any sequences in the public databases, including those detected previously in the same environment (Stoeck et al. 2003a). Grouped on the basis of the 99% cut-off value, these sequences form 72 OTUs, of which 61 proved novel, i.e. they do not include any previously known rRNA sequence. Below, we describe the OTUs detected and discuss their significance.

Generally, the Cariaco OTUs can be divided into two subgroups. The first is represented by OTUs whose closest relatives are named organisms from established taxa. The most divergent OTU in this group is formed by the clone T37E6 (96.0% similarity to Strombidium sp. SNB99-2 [Strueder-Kypke and Lynn 2003]) (Fig. 2). The least divergent OTUs in this group are over 99% similar to one or another named organism, and thus are mere variants within previously described OTUs (or species). An example of such a sequence is our clone P13F2 that is 99.35% similar to Pentapharsodinium tyrrhenicum (Saunders et al. 1997) (Fig. 2).

Overall, the OTUs in this subset are likely to be new strains within established species, or new species within the described genera and families. Their characteristics are likely to be predictable by extension of what is known about their described close relatives. The number of such OTUs is rather low (<25%). Examples include five diatom OTUs and one unclassified flagellate (Fig. 3), five fungal OTUs (Fig. 4), three ciliate OTUs, one dinoflagellate (Fig. 2), one bicosoecid, and one novel OTU within the Cercozoa (Fig. 5).

The second subgroup of OTUs, while falling into established clades, has its closest relatives from among environmental sequences rather than named organisms. This group encompasses the

---

**Figure 1.** Operational taxonomic units (OTUs) shared among the three 18S rRNA gene libraries constructed with different primer sets (see Methods). The area of the ovals area is proportional to the size of the respective clone library and the overlap area is proportional to the number of OTUs shared by the relevant libraries.
majority (75%) of the newly detected OTUs. Most of them belong to the clades known exclusively by 18S environmental signatures, with some representing deep branches within, or sister groups of, the established clades. Selected examples include the following:

**Alveolates**

A group of eight OTUs falls within the ubiquitous uncultured marine alveolates Group I (Fig. 2). This group was first described by López-García et al. (2001) and has since been detected in, for example, the Guaymas Basin (Edgcomb et al. 2002), the Mid-Atlantic Ridge (López-García et al. 2003), the coastal and equatorial Pacific (Moon-van der Staay et al. 2001; Worden, unpubl. sequence data from GenBank), as well as in our earlier study of the Cariaco Basin (Stoeck et al. 2003a). This phylogenetically diverse and cosmopolitan group of organisms contains no named species and represents an important target for cultivation efforts.

We detected a group of five OTUs that fall into the candidate ciliate class CAR-H, a clade observed first in our previous rRNA survey of the Cariaco area (Stoeck et al. 2003a) (Fig. 2). The novelty of the CAR-H clade and degree of separation from known ciliate classes suggest a divergence at the class level. This is remarkable because the Phylum Ciliophora is among the best studied Protozoa, and claims have been made that most species of ciliates have been largely discovered by now (Finlay and Fenchel 1999; for criticism, see Foissner 1999; Richards and Bass 2005). The proposed clade, if confirmed by cultivation studies and z-taxonomy, would represent the first new ciliate class-level group discovered in decades. An additional interesting feature of this clade is that its occurrence seems to be limited to anoxic environments.

We also detected a pair of OTUs clustering within the marine alveolate Group II (López-García et al. 2001). This group is well supported by 100% bootstrap (Fig. 2), but its relationship to other alveolates, including dinoflagellates, is unclear. The clade is ubiquitously distributed and may be quite important ecologically, but little can be said about the biology of this group as it is composed almost exclusively of environmental sequences. The only cultured representative is *Amoebophrya* sp. Phylogenetic analyses placed this organism at the base of the dinoflagellates (Gunderson et al. 1999). However, based on morphological, structural, and life cycle characteristics, Grassé (1952) classifies *Amoebophrya* as a flagellate incertae sedis. The author does not assign *Amoebophrya* a precise systematic position and cannot even affirm that *Amoebophrya* is a flagellate at all. This is supported by the phylogenetic studies of Moreira and López-García (2002) showing that *Amoebophrya* might be the first described member of a new group of alveolates. However, more sequence data of the order Syndiniales and cultured representatives of the uncultured marine alveolate Group II will be required to clarify this issue. The phylogenetic position of clone T53B4, closely related to another clone (H67) previously found in Cariaco, is uncertain (bootstrap 51%). Using a slightly modified alignment and different z and I, these sequences appeared branching clearly within the uncultured marine alveolate Group II.

**Stramenopiles**

We identified 14 diatom OTUs, out of which eight contain environmental sequences detected previously (Stoeck and Epstein 2003; Stoeck et al. 2003a) and six are novel. Such abundance of diatom 18S rRNA sequences in an anoxic and anoxic environment is remarkable. While it is possible that all these sequences belong to sinking and dying (or dead) cells, we cannot exclude the possibility that some of them in fact represent genuinely anaerobic heterotrophic diatoms. This interpretation is consistent with a high abundance of live diatoms in marine anoxic environments.

---

**Figure 2.** Minimum evolution phylogenetic tree of eukaryotic small-subunit rRNA showing the position of alveolate clones detected in the Cariaco basin. The tree was constructed under maximum-likelihood criteria by using a TIM+I+G DNA substitution model with the variable-site gamma distribution shape parameter (z) at 0.7613, the proportion of invariable sites at 0.2213 and base frequencies and a rate matrix for the substitution model as suggested by Modeltest (see Methods), based on 727 unambiguously aligned positions. Distance bootstrap values over 50% from an analysis of 1000 bootstrap replicates are given at the respective nodes. Three acantharean sequences were chosen as outgroup. Sequences retrieved in this study are color-coded. Sequences marked in red refer to library 1 (primer set Euk 528F/U1517), sequences in green to library 2 (primer set Euk528F/U1492R), and sequences in blue to library 3 (primer set Euk528F/U1391R). Colored sequences in brackets indicate clones from different libraries that shared a sequence similarity of at least 99% to the sequence included in the phylogeny (representing an operational taxonomic unit, OTU).
ARTICLE IN PRESS

E233 (Cariaco)
T60H4 (T45D10)
T60A7 (p15D9)
CCA67 (Sippewisset)
BOLA250 (Bolinas tidal flat)

Nanofrustulum shiloi
Thalassionema nitzschioides

E171 (Cariaco)
T37A12 (p14C2, p15B03)
H65 (Cariaco)

Thalassiosira rotula
E29 (Cariaco)

Nannochloropsis limnetica
Eustigmatos magna
Ochromonas danica
Paraphysomonas foraminifera
Costaria costata
Zonaria diesingiana

Hyphochytrium catenoides
Rhizidiomyces apophysatus
Thraustochytrium multirudimentale

ANT12-11 (Antarctica)

BAQ1 (coastal Pacific)
UEPACRg5 (coastal Pacific)
HE100427.21 (North Sea)
NA11-5 (North Atlantic)

Lagenidium giganteum
Apodachlya brachynema

D85 (Cariaco)

Blastocystis hominis

0.01 substitutions/site
sediments, and the knowledge of obligate and facultative heterotrophy for at least some diatoms (Admiraal and Peletier 1979; Hellebust and Lewin 1977; Li and Volcani 1987).

Four novel Cariaco clones (T53A1, p14A1, T41B10, T45C1), which branch among the heterotrophic stramenopiles, are not related to named species (Fig. 3). In our analysis, these sequences are affiliated with other environmental sequences from different oceanic regions, such as the uncultured marine stramenopile cluster MAST-3 (Massana et al. 2004b). This group consistently appears as a separate clade in several published studies, though sufficient statistical support for this clade is lacking (Fig. 3; Dawson and Pace 2002; Diez et al. 2001; Massana et al. 2004b; Stoeck et al. 2003a).

Furthermore, we discovered two labyrinthulid clones closely related to other environmental sequences from the Cariaco Basin and elsewhere (Edgcomb et al. 2002; Massana et al. 2004a; Stoeck et al. 2003a) (Fig. 3). Labyrinthulids are marine osmo-heterotrophs that occur globally typically prospering in oxygen-depleted deep-sea environments (Raghukumar 2002), and some are known to be anaerobes (Naqvi 1994).

**Rhizaria**

Within the Rhizaria (Radiolaria, Foraminifera, Cercozoa), we detected OTUs belonging to the Acantharea, Polycystinea and Cercozoa (Fig. 5). Nine acantharean OTUs obtained include two that branch within the Acantharea, and seven forming its sister clade. This sister clade is without a cultured representative, which indicates that the Acantharea as a group are likely to be undersampled and more diverse than currently thought. These organisms are free-floating planktonic heterotrophs with cosmopolitan distribution. They are rarely found in coastal waters and appear to occupy species-specific positions in the water column, often at considerable depth (Edgcomb et al. 2002; López-García et al. 2001; Stoeck et al. 2003a). The phylogenetic position of the acanthareans was considered rather enigmatic for a long time (Amaral-Zettler et al. 1997). Our analyses support the more recent findings that both Acantharea and Polycystinea (belonging to the Radiolaria) form monophyletic sister groups to the Cercozoa and branch within the radiation of the crown groups (Bass et al. 2005; López-García et al. 2002; Nikolaev et al. 2004; Polet et al. 2004).

**Euglenozoa**

The phylum Euglenozoa consists of three distinct protozoan groups: the euglenoids, diplonemids and kinetoplastids (Cavalier-Smith 1993; Simpson 1997). Two of the new Cariaco clones (T53C3, T53F7), together with two previously detected Cariaco clones (Stoeck et al. 2003a), show a close and well-supported (98% bootstrap) affiliation to *Rhynchopus* sp., a diplonemid flagellate (Fig. 6). This sequence clade is of special interest, as the phylogenetic position of diplonemids within the euglenozoa remains unsettled (Busse and Preisfeld 2002; Cavalier-Smith 2003; Maslov et al. 1999; Moreira et al. 2001; Simpson et al. 2002; Von der Heyden et al. 2004). Also in our analysis, the phylogenetic position of diplonemids is statistically not supported. Taxonomic sampling was shown to be a major factor affecting the topology of the SSU rRNA euglenozoan tree (Moreira et al. 2001). Since only few diplonemid sequences are available, concerted efforts are in order to enlarge the diplonemid sampling and retrieve novel diplonemid sequences is a high priority in euglenozoa phylogeny.

Several conclusions follow from the above observations.

First, the true extent of protistan diversity in our study site is unknown, and it is difficult to estimate with any certainty what portion of this diversity has been captured by our clone libraries. It appears however that this portion may be insignificant, and even with > 1000 clones from four clone libraries constructed with four different primer sets, the majority of protists inhabiting this anoxic environment might still have escaped detection. This may be typical of most, if not all, 18S rRNA surveys conducted to date.

Second, the protistan diversity in anoxic waters of the Cariaco Basin includes lineages of substantial phylogenetic novelty. The 18S rRNA gene sequence information obtained strongly support...
Figure 4. Minimum evolution phylogenetic tree of eukaryotic small-subunit rRNA showing the position of fungal clones detected in the Cariaco basin. The tree was constructed under a GTR+I+G DNA substitution model with the variable-site gamma distribution shape parameter ($\alpha$) at 0.5874, the proportion of invariable sites at 0.2632 and base frequencies and a rate matrix for the substitution model as suggested by Modeltest (see Methods), based on 936 unambiguously aligned positions. Distance bootstrap values over 50% from an analysis of 1000 bootstrap replicates are given at the respective nodes. Three alveolate sequences were chosen as outgroup. Only sequences of the closest relatives of the Cariaco clones, and selected representatives of other taxonomic groups are shown. The legend is the same as for Figure 2.

Figure 5. Minimum evolution phylogenetic tree of eukaryotic small-subunit rRNA showing the position of rhizarian clones recovered from the Cariaco Basin. The tree was constructed using a TrN+G DNA substitution model with the variable-site gamma distribution shape parameter ($\alpha$) at 0.5721 as suggested by Modeltest (see Methods), based on 713 unambiguously aligned positions. Distance bootstrap values over 50% from an analysis of 1000 bootstrap replicates are given at the respective nodes. Three ciliate sequences were chosen as outgroup. The legend is the same as for Figure 2.
Figure 6. Minimum evolution phylogenetic tree of eukaryotic small-subunit rRNA showing the position of euglenozoan clones recovered from the Cariaco Basin. The tree was constructed under maximum-likelihood criteria by using a GTR+I+G DNA substitution model with the variable-site gamma distribution shape parameter (α) at 0.8883, the proportion of invariable sites at 0.1861 and base frequencies and a rate matrix for the substitution model as suggested by Modeltest (see Methods), based on 670 unambiguously aligned positions. Distance bootstrap values over 50% from an analysis of 1000 bootstrap replicates are given at the respective nodes. Three archaean sequences were chosen as outgroup. The legend is the same as for Figure 2.
the reality of previously detected, deeply rooted protistan clades within the stramenopiles and alveolates, their sister, or ancestral groups. It would be particularly interesting to gain access to representatives of these clades to assess their morphological and physiological uniqueness as well as ecological and evolutionary importance. We are in the process of recovering such representatives using a novel methodology that combines fluorescence in situ hybridization and light and electron microscopy (Stoeck et al. 2003b).

**Methods**

**Study site and sampling:** The Cariaco Basin is the world’s largest truly marine body of anoxic water and represents a large natural sediment trap, collecting sinking debris and biota from surface waters. Detailed descriptions of this environment can be found elsewhere (Müller-Karger et al. 2001; Richards and Vaccaro 1956; Taylor et al. 2001). The sampling procedure is described in Stoeck et al. (2003a). In short, a 2.3 l seawater sample was collected just below the oxic/anoxic interface at a water depth of 340 m, and cells were collected on 47 mm Durapore membranes (0.65 μm pore-size) (Millipore, Billerica, MA) under gentle vacuum (~40 ml min⁻¹). Sampling and sample processing protocols were designed to prevent exposure to the atmosphere (Stoeck et al. 2003a). Membranes were placed in cryovials containing 1 ml of DNA extraction buffer (Stoeck et al. 2003a) to which 5 μl of proteinase K was added and then returned to the United States frozen in a block of ice.

**DNA isolation, PCR amplification, cloning, and sequencing:** The nucleic acids were obtained as described previously (Stoeck et al. 2003a). We amplified ~1000—1300 bp fragments of the 18S rRNA gene using three primer sets, consisting of a single forward primer E528F (5’-CGGTAATCCAGCTCC-3’) (Edgcomb et al. 2002) and three different universal reverse primers Univ1391RE (5’-GGGCGGTGTGTACAARGRG-3’) (Dawson and Pace 2002), Univ1492RE (5’-ACCTTGTACGRTT-3’) (Edgcomb et al. 2002), and Univ1517 (5’-ACGGCTACCTTGTTACACTTT-3’, modified from Univ1492RE). The first two primer sets successfully amplified the rRNA gene using the sample’s DNA extract as a template. The latter primer set was successful only with a nested PCR, with the template for the second PCR provided by the first reaction with primers Euk A (5’-AACCTGGTGATCCTGCGCACC-3’) and Euk B (5’-TGATCCTTCTGAGTATCCCTTAC-3’) (Medlin et al. 1988). The PCR protocol employed HotStart Taq DNA polymerase (Qiagen, Valencia, CA) in all cases; this protocol consisted of an initial hot-start incubation (15 min at 95°C) followed by 30 identical amplification cycles (denaturing at 95°C for 45 s, annealing at 55°C for 1 min, and extension at 72°C for 2.5 min), and final extension at 72°C for 7 min. The PCR products were cloned, separately for each primer set, using the TA cloning kit (Invitrogen, Carlsbad, CA) according to the manufacturer’s instructions. Plasmids were isolated from overnight cultures by using the Machery—Nagel NucleoSpin Robot-96 Plasmid Kit (Easton, PA) or amplified from plate colonies via the Templiphi 100 Amplification Kit (Amersham Biosciences, Piscataway, NJ) according to the manufacturers’ instructions. The presence of the target insert was confirmed by PCR reamplification as described above. Clones were initially grouped into operational taxonomic units (OTUs) based on partial sequences using M13F sequencing primer considering only sequences with ca. 600—800 Phred20 quality checked bp (sequence similarity cutoff for OUT definition = 99%). All in all, we partially sequenced nearly 600 clones, of which 497 proved to be protistan targets, which were considered for further analyses. The OTU grouping was then checked using Haelll (New England Biolabs, Beverly, MA) digestion patterns of all clones (ARDRA — amplified ribosomal DNA restriction analyses). At least one clone of each OTU (n = 91) was sequenced bidirectionally at the University of Maine DNA Sequencing Facility. Sequence variability within ARDRA patterns (OTUs) was also assessed by random complete bidirectional sequencing of multiple clones within ARDRA patterns (OTUs) (n = 7). Sequence comparison revealed that sequences within a single digestion pattern do not diverge by more than 1%.

**Phylogenetic analyses:** We followed the procedure described earlier (Stoeck et al. 2003a). In short, the 18S rRNA gene sequences were compared to those in GenBank and to >5000 prealigned eukaryotic small-subunit rRNA in the Ribosomal Database Project (Maidak et al. 2001). Multiple alignments were obtained and manually refined according to conserved secondary structures using ClustalX (Thompson et al. 1994), MacClade v. 4.06 (Maddison and Maddison 2000), and the secondary-structure prediction tool of the Vienna RNA Package (CANNONE et al. 2002; Hofacker et al. 1994). Further phylogenetic analyses included only conserved regions.
Minimum-evolutionary-distance trees and maximum-likelihood trees (where applicable) were constructed by using the PAUP* software package 4.0b10 (Swofford 2001). We used the program Modeltest (Posada and Crandall 1998) to choose the models of DNA substitution that best fit our data sets from among 56 possible models and assessed the relative stability of tree topologies by using 1000 bootstrap replicates.

The gene sequences from this study have been deposited in the GenBank database (accession numbers AY882442—AY882540).

Acknowledgements

We thank the captain and crew of the B/O Hermano Gines and the staff of the Estación de Investigaciones Marinas (EDIMAR) de Margarita (Fundación la Salle de Ciencias Naturales, Punta de Piedras, Isla de Margarita, Venezuela) for their help in the field. The authors are grateful to the scientists of the CARIACO time series program, particularly F. Muller-Karger and Y. Astor for logistical field support and ancillary data. We thank Dr. E. Jarroll (Northeastern University, Boston, MA, USA) for his untiring support of this project. Many thanks to Ted Maney (Marine Science Center, Northeastern University, Nahant, MA, USA) for effective management of logistical matters. The manuscript was substantially improved by two anonymous reviewers and by additional suggestions by the monitoring editor. This work was funded by the Deutsche Forschungsgemeinschaft Grants STO414/2-1 and STO414/2-2 to TS and US National Science Foundation Grants to GT (OCE9730278) and SE (MCB-0348341).

This is contribution # 255 of the Marine Science Center of Northeastern University, Nahant, MA, USA.

References


Busse I, Preisfeld A (2002) Phylogenetic position of Rhynchopus sp. and Diplonema ambulator as indicated by analyses of euglenozoan small subunit ribosomal DNA. Gene 284: 83—91


Naqvi SWA (1994) Denitrification processes in the Arabian Sea. Proc Ind Acad Sci (Earth Planet Sci) 103: 279—300


Available online at www.sciencedirect.com