Benthic eukaryotic diversity in the Guaymas Basin hydrothermal vent environment

Virginia P. Edgcomb[†‡], David T. Kysela[†], Andreas Teske[‡], Alvin de Vera Gomez[‡], and Mitchell L. Sogin[†§]

[†]The Josephine Bay Paul Center for Comparative Molecular Biology and Evolution, Marine Biological Laboratory, and [‡]Biology Department, Woods Hole Oceanographic Institution, Woods Hole, MA 02543

Communicated by Norman R. Pace, University of Colorado, Boulder, CO, March 30, 2002 (received for review August 3, 2001)

Molecular microbial ecology studies have revealed remarkable prokaryotic diversity in extreme hydrothermal marine environments. There are no comparable reports of culture-independent surveys of eukaryotic life in warm, anoxic marine sediments. By using sequence comparisons of PCR-amplified small subunit ribosomal RNAs, we characterized eukaryotic diversity in hydrothermal vent environments of Guaymas Basin in the Gulf of California. Many sequences from these anoxic sediments and the overlaying seawater represent previously uncharacterized protists, including early branching eukaryotic lineages or extended diversity within described taxa. At least two mechanisms, with overlapping consequences, account for the eukaryotic community structure of this environment. The adaptation to anoxic environments is evidenced by specific affinity of environmental sequences to aerotolerant anaerobic species in molecular trees. This pattern is superimposed against a background of widely distributed aerophilic and aerotolerant protists, some of which may migrate into and survive in the sediment whereas others (e.g., phototrophs) are simply deposited by sedimentary processes. In contrast, bacterial populations in these sediments are primarily characteristic of anoxic, reduced, hydrocarbon-rich sedimentary habitats.

Phylogenetic studies of ribosomal RNAs from cultured species imply that large evolutionary distances separate major groups of protists. Certain aerotolerant anaerobic parasites seem to represent early diverging lineages in the eukaryotic line of descent (1). Because they do not require atmospheric levels of oxygen for growth, these organisms lack mitochondria and peroxisomes. The absence of these features in the most divergent eukaryotic lineages and the thermophilic phenotype of the deepest prokaryotic branches suggest that warm, anoxic environments surrounding deep-sea hydrothermal vents might support novel and diverse eukaryotic communities. Little information about the phenotypic and evolutionary diversity of free-living protists from anoxic marine environments is available. Most studies of protist diversity rely upon morphological characters to differentiate between genera. However, culture-based studies and microscopical criteria cannot provide quantitative measures of genetic diversity. More significantly, these methods frequently do not provide comprehensive profiles of community composition. Cultivation-independent molecular surveys of eukaryotic microbial diversity based on comparisons of PCR amplicons of small subunit (SSU) rRNA genes from marine pelagic environments reveal a microbial world with few taxonomically assignable protists (2, 3). There are no analogous molecular surveys of eukaryotic microbial diversity in anoxic, deep-sea sediments.

Here, we report the results of a molecular survey of eukaryotic community SSU rRNA gene composition in sediments and the seawater interface proximal to a deep-sea hydrothermal vent. The study site is Guaymas Basin, Gulf of California, which is a hydrothermally active environment that includes vent plumes, seeps, and anoxic sediments, each exhibiting a wide range of temperatures. This environment supports surface-attached microbial mats, diverse prokaryotic and eukaryotic microbes, and symbiont-harboring invertebrates. Warm, sulfide- and hydrocarbon-rich, anaerobic sediments accumulate at a rate of 1–2 mm/year because of high biological productivity in the water column and a large terrigenous input from Baja and the Mexican mainland (4). The approximately 100–500 m thick sediments release vast quantities of petroleum, short-chain fatty acids, and ammonia by means of pyrolysis of complex organic substrates (5, 6). The vent fluids that percolate through the organic-rich sediments have an unusual chemistry based on high-carbonate content and near-neutral pH and release large amounts of methane (7).

Materials and Methods

Sample Collection. Sediment cores A and C were obtained from the Everest Mound area in the southern Guaymas vent field (Alvin dives 3202 and 3207; 1998), with concurrent *in situ* temperature profiles using Alvin's thermoprobe. The temperature gradient for the top 3 cm of core A ranged from 3°C to 65°C and from 3°C to 45°C in core C. A thick (up to several cm) layer of flocculent mat of the sulfur-oxidizing bacteria *Beggiatoa* overlaid the core A sediments, whereas only a very thin layer (≈2 mm) was evident at the surface of core C. Sediment cores were transported to the ocean surface in sealed coring devices that captured the *in situ* supernatant waters of the sediments. Cores were sectioned anaerobically into 1-cm layers immediately after return to the surface and stored at −80°C for later molecular analysis. The 1-cm segments of cores A and C were designated A1-A3 and C1-C3, respectively, and the surface of core C was designated CS.

DNA Extraction and Sequencing of SSU rDNA Genes. Two sediment cores, A and C, and the overlaying seawater of core C were used in molecular analyses. Eukaryotic SSU rRNA sequences were obtained from the sediment–seawater interface layer of core C and the first 3 cm of sediment of cores A and C. Samples were initially homogenized by bead-beating for 15 s. Nucleic acids were extracted by SDS/proteinase K incubation and purified by phenol:chloroform extraction and subsequent ethanol precipitation, as described (8). Samples for which PCR amplification was initially unsuccessful were purified further by agarose gel electrophoresis and extraction. Partial (*Escherichia coli* positions 528-1505) SSU rDNA was amplified by the PCR method by using a nested amplification with the universal primers U514F (5′-GTGCCAGCMGCCGCGG-3′) and U1492R (5′-ACCTTGTTACGACTT-3′) followed by the eukaryotic primer E528F (5′-CGGTAATTCCAGCTCC-3′) and U1492R. A single, full-length eukaryotic clone (CS_R003) was obtained from a concurrent study using archaeal primers A8F (5′-TCCGGTTGATCCTGCC-3′) and A1492R (5′-GGCTACCTTGTTACGACTT-3′). Clone libraries were prepared from PCR products by using a TOPO-XL cloning kit (Invitrogen). Nearly full-length, double-stranded DNA sequences were determined by using

Abbreviation: SSU, small subunit.

Data deposition: The sequences reported in this paper have been deposited in the GenBank database (accession nos. AY046599–AY046873).

[§]To whom reprint requests should be addressed. E-mail: sogin@mbl.edu.

www.pnas.org/cgi/doi/10.1073/pnas.062186399

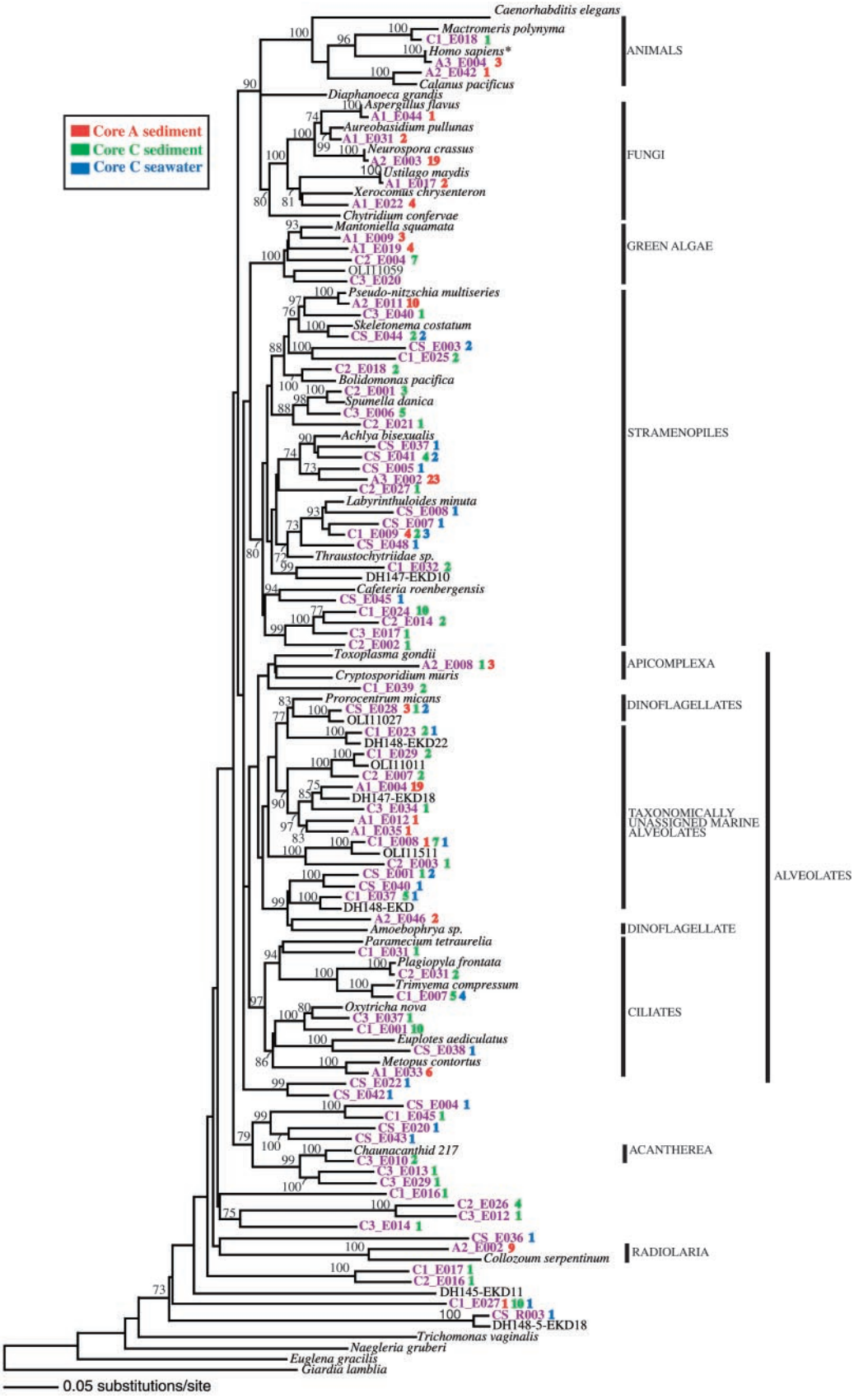

Fig. 1. Phylogenetic tree of partial SSU rRNA genes obtained from Guaymas sediment cores. Sequences in purple were obtained in the current study. Color-coded numbers to the right of sequence names indicate counts of similar sequences (< 0.06 substitutions per site in a distance analysis of complete sequences) represented from specific core samples. Tree topology was obtained by a heuristic search of 10 random-addition starting trees under a Tamura

LI-COR's simultaneous, bi-directional sequencing protocols resolved on a 4200 scanner (LI-COR, Lincoln, NE). Approximately 45 randomly selected clones/clone library were sequenced from each core horizon. This selection represents a small fraction of cloned amplicons from each library. Low-quality sequence reads and those that did not correspond to full-length amplicons were discarded, yielding 276 clones for phylogenetic analysis.

Phylogenetic Analysis. We compiled the Guaymas clone SSU rDNA sequence data in ARB (9) and aligned these data with sequences obtained from the GenBank database by using the ARB FastAligner utility. By using phylogenetically conserved secondary structures, we refined the ARB alignment. Clustering patterns in a neighbor-joining analysis identified representative clones for more complex phylogenetic analyses based upon comparisons of 1,021 aligned positions. Red algae showed no specific affinity to any environmental sequences and were excluded from the final analysis because of their destabilizing effect on the eukaryotic "crown." A heuristic search for the tree topology under a Tamura Nei minimum evolution distance model used 10 random-addition starting trees, and tree bisection reconstruction (TBR) branch-swapping (PAUP* V.4.0B; ref. 10). We assessed the relative stability of topological elements by using 100 bootstrap replicates under minimum evolution. Heuristic searches for bootstrap analysis used neighbor-joining starting trees and TBR branch-swapping. SSU rRNA sequences are available under GenBank accession nos. AY046599–AY046873.

Results and Discussion

From our molecular study of ribosomal RNA sequences (Fig. 1) we recognized three kinds of previously uncharacterized eukaryotic diversity. The first corresponds to sequences that are unrelated to those of any other eukaryotes, and they seem to represent early branches in the eukaryotic tree, e.g., C1_E016, the clade C2_E026 plus C3_E012, the lineages C3_E014, CS_E036, the clade C1_E017 plus C2_E016, the lineage C1_E027, and the clade CS_R003 plus DH148–5-EKD18. The second type of diversity corresponds to sequences that represent deep branches within well described eukaryotic clades including the stramenopiles, apicomplexa, dinoflagellates, ciliates, acantharea, and radiolaria. Their depth of branching is oftentimes comparable to class level differences. For example, the well supported clade comprising the sequences C1_E024, C2_E014, C3_E017, and C2_E002 is basal to other stramenopiles and likely represents organisms that are nonphotosynthetic heterotrophs. We also include within this category clades represented by C1_E023, C1_E029, A1_E004, C1_E008, and C1_E037, which are closely related to sequences from taxonomically unassignable alveolates ("OLI-" and "DH-" types) obtained in molecular surveys of marine pelagic environments (2, 3). Because they are bracketed by sequences from dinoflagellates, we predict that these sequences represent similar phototrophic species. Finally, we observed clades that are neither deep in the tree nor specifically related to any recognized eukaryotic clade (e.g., CS_E022 and CS_E042).

Comparable numbers of PCR amplicons were cloned from core horizons at similar depths in Cores A and C, but only 45 were selected at random for detailed sequence analysis. Sequences of core A clustered to form a smaller number of clades than did sequences from core C. In Fig. 1, the larger number of lineages from core C relative to core A reflects this bias. Because culture-independent, molecular surveys do not provide reliable indicators of viability, we cannot absolutely determine whether the sequences in our survey reflect active members of protist communities or merely recent deposition of protists into the sediment. However, we can take advantage of phylogenetic affinity with well studied species to make predictions about the phenotypes and lifestyles of organisms represented in our environmental sequence database. Some of our recovered sequences are from organisms related to specialists of anaerobic environments. For example, within the ciliates, we recovered sequences (C1_E007 and A1_E033) that are closely related to sequences from the anaerobes *Metopus contortus* and *Trimyema compressum*, respectively. In addition, the detection of rRNA sequences from diverse flagellates is consistent with published accounts of viable flagellate protists from several hydrothermal vent sites, including Guaymas Basin (11). Rapid rates of reproduction allow these small flagellates to colonize temporally and spatially variable habitats such as these fine-particle, hydrothermally heated sediments.

Other recovered sequences are closely related to those from eukaryotes that are more generally distributed in marine environments. These include certain fungi, green algae, stramenopiles, alveolates, acanthareans, and radiolarians (12). For example, in the CS layer of core C, we recovered the sequence of a specific relative of the ciliate *Euplotes aediculatus* (a surface-associated filter feeder), described from a variety of habitats including marine and freshwater sediments and pelagic environments (13). Many of our detected rRNA sequences are unlikely to be from active members of the benthic population. Sequences of phytoplanktonic taxa, such as green algae and diatoms (14), are from organisms that require solar radiation for active photosynthesis and are, therefore, not suited to this aphotic environment. Other sequences may represent dead or encysted forms; for example, dinoflagellate cysts have been detected in marine sediments (15). Some sequences (A2_E003, A2_E011, A3_E002, C1_E024, A1_E004, C1_E001, and C1_E027) were more frequently recovered (nearly identical sequences from $\geq$10 clones) from our libraries. It is impossible however, to differentiate greater representation within the microbial community from biased recovery/amplification (16).

With successively deeper sections through the cores, recovery of eukaryotic SSU rRNA amplicons diminished. Below 3 cm depth in cores A and C (approximately 65°C and 45°C, respectively), we only recovered prokaryotic sequences. The decline in efficiency of recovering eukaryotic SSU rRNA amplicons deeper in the sediments likely reflects temperature and redox limitations of the eukaryotic community, with attendant degradation of nucleic acids in metabolically inactive organisms. In contrast, recovery of prokaryotic rRNA gene sequences continues well into the sediment (>12 cm). Archaeal methanogens that can grow to 110°C have been identified (17), and sulfate reduction has been measured at temperatures near 100° in these sediments (18, 19).

Eukaryotic rRNA sequences recovered from the sediment encompass the majority of described lineages in the eukaryotic domain. In addition to anoxic specialists, there exists a diverse collection of eukaryotic sequences that is comparable in composition to sequences from benthic, pelagic, and near-surface water populations. In contrast, a parallel molecular survey of the same sedimentary cores describes sequences from a prokaryotic

Nei minimum evolution distance model using 1,021 aligned positions. Distance bootstrap values over 70% from an analysis of 100 bootstrap replicates are given at respective nodes. (Bar = 0.05 substitutions per site.) *Marine mammal rRNA sequences corresponding to the region amplified in this study are not available in public databases.

Table 1. Bacterial 16S rRNA clone profile of Guaymas sediments

Bacterial phylogenetic groups	Core A 0–2 cm (n = 68), % clones	Core C 0–1 cm (n = 45), % clones	Core C sediment supernatant (n = 24), % clones	Oxic vs. anoxic habitat of related clones or isolates
Epsilon-*Proteobacteria*	32.4	28.9	25.0	Most likely sulfur-oxidizing interfaces
Gamma-*Proteobacteria*	—	22.2	20.8	
Total	32.4	51.1	45.8	
Delta-*Proteobacteria*	11.8	13.4	12.5	
Green non-sulfur bacteria	14.7	8.9	—	
Subdivision OP11	8.8	8.9	8.3	
Subdivision OP8	5.9	2.2	—	Anoxic
Subdivision OP9	—	2.2	4.2	
Subdivision OP5	—	—	4.2	
Subdivision OP3	—	2.2	—	
Subdivision OP1	1.4	—	—	
Total	42.7	37.8	29.2	
Cytophaga/*Flavobact*	—	2.2	20.8	
Planctomycetales	2.9	2.2	—	
Verrucomicrobium	4.4	—	—	Oxic
Myxobacteria (delta)	—	4.5	—	
Chloroplasts	—	2.2	—	
Total	7.3	11.1	20.8	
Guaymas new bacterial group*	13.2	—	—	
High G + C grampositives	4.4	—	—	Unknown
Unidentified	—	—	4.2	
Total	17.6	—	4.2	

*Bacterial phylotype not affiliated with other sequences so far; found only in Guaymas.

community (20) where most lineages are characteristic of anoxic, reduced, hydrocarbon-rich sedimentary habitats (Table 1). This microbial community includes relatives of the green non-sulfur bacteria and members of the candidate subdivisions OP1, OP3, OP8, OP9, and OP11. Similar community compositions occur in sulfidic geothermal springs (21), deep subsurface soil (22), a methanogenic aquifer contaminated with aromatic and chlorinated hydrocarbons (23), and from enrichments on aromatic and chlorinated hydrocarbons (24). Delta-proteobacterial sequences were primarily from members of the sulfate-reducing family *Desulfobulbaceae*, which specialize in propionate oxidation (25). Propionate and acetate are the predominant, naturally occurring short-chain fatty acids in the Guaymas Basin hydrothermally altered sediments (26). ε- and γ-proteobacterial sequences from Guaymas are related to those of sulfur-oxidizing bacterial symbionts of marine invertebrates. These symbionts require the simultaneous presence of sulfide and chemical oxidants, a characteristic of the sediment surface (27). However, evidence of nonsymbiotic filamentous ε-proteobacteria also has been reported in microbial mats from hydrothermal vents (28).

These results suggest that several mechanisms account for the community composition of protists in the sediment. A significant, but relatively small component of recovered eukaryotic sequences are from organisms specialized to anaerobic sediments such as those found in Guaymas Basin. Selection for anaerobic specialists is evidenced by the recovery of sequences that are related to those of several groups of strict anaerobes not typical of the water column. The community structure of the oxic–anoxic interface likely includes both microaerophilic and facultative forms, some of which may migrate into and out of the anoxic sediments. This microbial community structure is augmented by contributions of microaerophilic and aerobic, particle-attached organisms from the water column. The active members of the eukaryotic community must feed primarily on some yet-unidentified fraction of the abundant bacteria sustained by hydrothermal activity and sedimentary input of organics (see Table 1).

Irrespective of events that shape eukaryotic population structures in anoxic sediments, the level of previously undescribed protist diversity is impressive. Because the sampling of clone libraries from this environment (and other marine environments) has not reached saturation (Fig. 2), and because of differences in PCR primers used in this and previously reported studies, we do not know the full extent of diversity or the overlap between different marine environments. Indeed, one of the basal environmental sequences from the overlaying seawater of core C, CS_R003, is nearly identical to sequence DH148–5-EKD18 described from deep-sea Antarctic plankton (3). Nonetheless, the largely specialized prokaryotic community present at this site

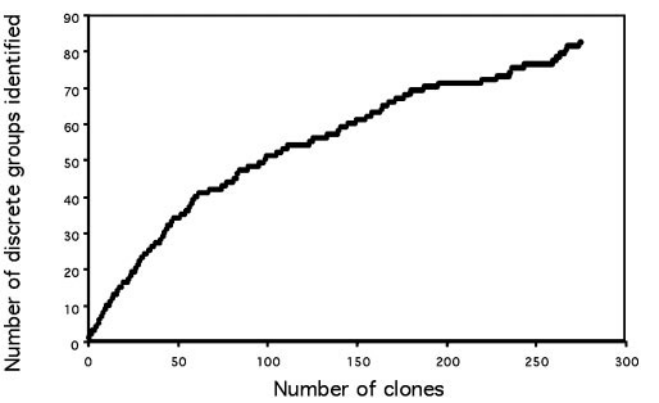

Fig. 2. Sampling saturation profile. Numbers of discrete groups are plotted as a function of numbers of clones sampled. Clone sequences were randomly resampled to completion without replacement to quantify coverage of phylotype diversity. Discrete groups are defined to encompass clones that shared less than 0.06 substitutions per site in a distance analysis based on an alignment of 1,021 nucleotide positions.

suggests that at least some of the many eukaryotic sequences identified here represent hydrothermal specialists of unique physiology and ecology. This remarkable diversity represents a small fraction of previously uncharacterized protist lineages that we are likely to discover in warm, anoxic, and other extreme environments.

We thank Tom Fenchel for helpful discussion on eukaryotic microbial distribution and ecology. We dedicate this study to Holger W. Jannasch, for whom this Guaymas Basin expedition was the final scientific adventure of his life. This project was supported by National Science Foundation Grant OCE 9714195 (to H. W. Jannasch and A.T.), National Aeronautics and Space Administration Exobiology Grant NRA97-OSS-11 and Astrobiology Cooperative Agreement NCC2-1054, and continuing support from the Unger G. Vetlesen Foundation (to M.L.S.). This paper is Woods Hole Oceanographic Institution contribution number 10481.

1. Sogin, M. L., Gunderson, J. H., Elwood, H. J., Alonso, R. A. & Peattie, D. A. (1989) *Science* **243,** 75–77.
2. Moon-van der Staay, S. Y., De Wachter, R. & Vaulot, D. (2001) *Nature (London)* **409,** 607–610.
3. Lopez-Garcia, P., Rodriguez-Valera, F., Pedros-Alio, C. & Moreira, D. (2001) *Nature (London)* **409,** 603–607.
4. Calvert, S. E. (1966) *J. Geol.* **76,** 546–565.
5. Bazylinski, D. A., Farrington, J. W. & Jannasch, H. W. (1988) *Org. Geochem.* **12,** 547–558.
6. Von Damm, K. L., Edmond, J. M., Measures, C. I. & Grant, B. (1985) *Geochim. Cosmochim. Acta* **49,** 2221–2237.
7. Welhan, J. A. (1988) *Chem. Geol.* **71,** 183–198.
8. Lin, C. & Stahl, D. (1995) *Appl. Environ. Microbiol.* **61,** 1348–1351.
9. Strunk, O., Gross, O., Reichel, B., May, M., Hermann, M., Stuckmann, S., Nonhoff, N., Lenke, M., Ginhart, A., Vilbig, A., *et al*. (1998) ARB (Department of Microbiology, Technical Univ. of Munich, Germany).
10. Swofford, D. L. (1998) PAUP*, Phylogenetic Analysis Using Parsimony (*and other methods) (Sinauer, Sunderland, MA), Version 4.0.
11. Atkins, M. S., Teske, A. P. & Anderson, O. R. (2000) *J. Eukaryotic Microbiol.* **47,** 400–411.
12. Lee, W. J. & Patterson, D. J. (1998) *Protist* **149,** 229–243.
13. Fenchel, T. (1987) *Ecology of Protozoa* (Springer, New York).
14. Platt, T. & Li, W. K. W. (1986) *Photosynthetic Picoplankton* (Department of Fisheries and Oceans, Ottawa).
15. Nehring, S. (1997) *Bot. Mar.* **40,** 307–324.
16. Picard, C., Ponsonnet, C., Paget, E., Nesme, X. & Simonet, P. (1992) *Appl. Environ. Microbiol.* **63,** 1028–1033.
17. Huber, R., Kurr, M., Jannasch, H. W. & Stetter, K. O. (1989) *Nature (London)* **342,** 833–834.
18. Jorgensen, B. B., Zawacki, L. X. & Jannasch, H. W. (1990) *Deep-Sea Res.* **37,** 695–710.
19. Jorgensen, B. B., Isaksen, M. F. & Jannasch, H. W. (1992) *Science* **258,** 1756–1757.
20. Teske, A., Hinrichs, K.-U., Edgcomb, V. P., Gomez, A., Kysela, D., Sylva, S. P., Sogin, M. L. & Jannasch, H. W. (2002) *Appl. Environ. Microbiol.* **68,** 1994–2007.
21. Hugenholtz, P., Pitulle, C., Hershberger, K. L. & Pace, N. R. (1998) *J. Bacteriol.* **180,** 366–376.
22. Chandler, D. P., Brockman, F. J., Bailey, T. J. & Fredrickson, J. K. (1998) *Microb. Ecol.* **36,** 37–50.
23. Dojka, M. A., Hugenholtz, P., Haack, S. K. & Pace, N. R. (1998) *Appl. Environ. Microbiol.* **64,** 3869–3877.
24. von Wintzingerode, F., Selent, B., Hegemann, W. & Gobel, U. B. (1999) *Appl. Environ. Microbiol.* **65,** 283–286.
25. Widdel, F. & Bak, F. (1992) in *The Prokaryotes*, eds. Balows, A., Trüper, H. G., Dworkin, M., Harder, W. & Schleifer, K.-H. (Springer, New York), 2nd Ed., Vol. 4, pp. 3352–3378.
26. Martens, C. S. (1990) *Appl. Geochem.* **5,** 71–76.
27. Nelson, D. C. & Fisher, C. R. (1995) in *The Microbiology of Deep-Sea Hydrothermal Vents*, ed. Karl, D. M. (CRC, Boca Raton, FL).
28. Longnecker, K. & Reysenbach, A.-L. (2001) *FEMS Microbiol. Ecol.* **35,** 287–293.