Vol. 150: 275–285, 1997 | MARINE ECOLOGY PROGRESS SERIES
Mar Ecol Prog Ser | Published April 30

Widespread Archaea and novel Bacteria from the deep sea as shown by 16S rRNA gene sequences

Jed A. Fuhrman*, Alison A. Davis

Department of Biological Sciences, University of Southern California, Los Angeles, California 90089-0371, USA

ABSTRACT: Marine microbial diversity is important yet poorly-known, due to low culturability and undersampling. However, 16S rRNA gene sequences cloned directly from biomass allow us to know what microbial types are present, irrespective of culturing, and to create probes suitable for biodiversity studies. Many sequences are needed for good probe design. Here we report on sequences from 57 deep sea clones, obtained by the polymerase chain reaction with 'universal' primers, from 500 m and 3000 m depths in the northeast Pacific and 1000 m depth in the subtropical Atlantic. The most common group, with 19 of the new sequences (10 Atlantic), was a recently reported crenarchaeal cluster, Group I. We also found 6 sequences in 2 other archaeal groups in the broad methanogen-halophile lineage; 2 of these were in a distinct lineage not previously reported. The bacterial sequences included 22 dispersed among the α and γ Proteobacteria (8 related to SAR 11), 5 related to a previously reported broad group (Group A) of marine clones poorly affiliated with known (cultured and sequenced) major bacterial divisions, 6 in a second group with little affiliation to any previously reported division (we call this Group B), 1 in a third possible major novel group, 2 deeply branched within the 'Green Nonsulfur' lineage, and 1 branching with a soil clone. In contrast to the vast majority of the sequences, a cluster of 5 sequences was very close to a cultured marine proteobacterium, *Alteromonas macleodii*. It appeared that 5 of the clones were chimeric, although this label is difficult to apply when sequences are only distantly related to those in the database, as was common. We conclude that the deep sea contains numerous novel and widespread major prokaryotic lineages. Given the huge volume of this habitat and typical bacterial abundances, it appears that the previously unknown archaeal and bacterial groups may be the most abundant organisms on Earth.

KEY WORDS: Archea · Bacteria · 16S rRNA · Deep sea · Marine · Phylogeny · Clone

INTRODUCTION

Marine microbiological studies have advanced to the point where much is known about the 'bulk' properties, such as total biomass, growth rates, nutrient utilization, etc. However, we know almost nothing about what kinds of prokaryotic microorganisms exist in the sea and how they are distributed in space and time. This means that we are treating the microbial system as a huge 'black box,' and any variations in the composition of that box are unknown. The potential diversity of microbes is immense (Woese 1987). Such diversity is interesting in its own right, and it also is almost certain to affect the 'bulk' properties of the system, such as conversion factors for rate or biomass measurements, the propagation of viral infection (largely species-specific), effects on grazers, etc. Thus there are several reasons for investigating this diversity.

In recent years, new techniques based upon molecular biological methods have permitted investigation of such diversity in ways that avoid the difficulties and problems associated with 'classical' culture-based approaches (Olsen et al. 1986, Pace et al. 1986, Amann et al. 1995). Among the most powerful of these techniques is the ability to obtain nucleotide sequences of 16S rRNA (or the genes encoding for that RNA) directly from biomass. These sequences can be analyzed phylogenetically in relation to a huge existing database (containing sequences from thousands of different organisms) in order to indicate what kind of organisms are present in a given sample. Sequences

*E-mail: fuhrman@usc.edu

that do not match the database directly can still be placed on a phylogenetic tree to indicate the relationship of the unknown organisms to cultured ones and other unknowns. One of the earliest habitats to be investigated by this method is the marine plankton, and several novel microbial groups, including archaea and bacteria, have been discovered there (Giovannoni et al. 1990, 1996, Britschgi & Giovannoni 1991, Schmidt et al. 1991, Fuhrman et al. 1992, 1993, DeLong et al. 1993, 1994, Mullins et al. 1995, Gordon & Giovannoni 1996). As extensive as these studies are, reporting on a few hundred sequences, they still just barely scratch the surface of potential diversity over space and time in the world ocean. Extant sequences are from a handful of samples, mostly from the euphotic zone. Even though a few groups have been found repeatedly in similar kinds of samples, there are undoubtedly many groups that have yet to be found. This report extends the coverage of such cloning and sequencing investigations in the deep sea, with samples from both Pacific and Atlantic basins.

The immediate goal of this study is to develop 'lists' of the types of organisms (via 16S rRNA sequences) found in marine plankton. Although this may appear to be simply exploratory or descriptive, such lists with many sequences are necessary before we can go on to the next step of creating probes for quantitative studies of biomass and activities. This is because even for the most common groups, one should have several sequences, some closely related to each other and others more distantly related, in order to design probes with a known specificity. For example, if a novel major group is only known from 1 or 2 sequences, creating a probe to match some unique region on those sequences might match only those 1 or 2 species, or might match a whole phylum or even a random subset of one or more phyla—one cannot know which. On the other hand, a large sequence database, especially from the habitat to be studied, allows selection of probe sequences that match narrow or broad groups as desired.

Here we report a 16S rRNA cloning and sequencing analysis from 3 deep sea samples collected from the Atlantic and Pacific Oceans at depths ranging from 500 to 3000 m. The deep sea is probably the largest habitat on earth and one of the least known biologically. Its microbial diversity is virtually unknown. The results suggest that these environments include numerous previously undescribed microbial lineages, both archaeal and bacterial, and that specific lineages are widespread.

METHODS

Water samples were collected by multiple Niskin bottles on a Rosette sampler (General Oceanics), with total volumes ranging from 140 to 200 l. Temperatures and salinities were measured with an *in situ* conductivity/temperature/depth sensor (Sea Bird Electronics). Locations and background data are shown in Table 1. Organisms were immediately collected by pressure filtration onto 142 mm diameter 0.22 µm pore size Durapore filters (Millipore Corp.) after prefiltration through 142 mm diameter Gelman AE glass fiber filters to remove eukaryotes. The Durapore filters were frozen, taken to the laboratory for further storage at −80°C, and subsequently extracted in hot SDS; DNA was purified by ethanol precipitation and phenol/chloroform extraction (Fuhrman et al. 1988). To amplify a portion of the 16S rRNA gene by PCR with a Perkin Elmer Cetus GeneAmp kit, a small portion of the extracted DNA, approximately 50 ng, was mixed with 1 µM each of 'universal primers' 537F (TTGAGCTCAAGCTTCAGCMGCCGCGGTAATWC) and 1492 R (TTTTGGATCCTCTAGAACGGGCGGTGTGTRC) (Fuhrman et al. 1992) (note: these primers contain linkers at the 5' ends that were not needed in this study), plus other PCR reagents (10 mM Tris, 50 mM KCl, 1.5 mM $MgCl_2$, 0.001% gelatin, 200 µM each dNTP, 5 units of AmpliTaq polymerase per 100 µl reaction) and amplified under the following conditions: 1 min, 92°C, 1 min, 55°C, and 1.5 min, 72°C for 30 cycles. Negative controls with water instead of DNA showed no products. Bands of ca 900 bp were cut from a 1.2% agarose gel (Seakem LE agarose from FMC Bioproducts, Rockland, Maine; run with Tris/EDTA/acetate buffer), extracted and purified with a Geneclean Kit

Table 1 Data for water samples taken from the northeast Pacific (p712 and pN1) and subtropical Atlantic (pB1) and used in this study. Dates given as mo/d/yr

Sample	Date	Depth (m)	Location	Bacterial abundance (10^7 cells l^{-1})	Temperature (°C)	Salinity (ppt)	Volume (l)
p712	Sep 14, 1990	500	32° 51.6' N, 118° 55.1' W	34	6.5	34.3	160
pN1	Sep 19, 1990	3000	32° 32.27' N, 121° 21.53' W	1.6	1.8[a]	34.6[a]	200
pB1	Sep 30, 1989	1000	32° 17.38' N, 64° 35.65' W	21	7	35.1	140

[a]Temperature and salinity from historical data for the region (sample was too deep for the sensors that were available)

(Bio 101 Inc., Vista, California), and ligated to vector PCR 1000 (TA Cloning Kit; InVitrogen Corp., San Diego, California), which was used to transform *Escherichia coli* according to the manufacturers' instructions. White colonies were checked for the presence of inserts in the plasmid by agarose gel analysis of restriction-enzyme-cut minipreps of the plasmids, prepared by the alkaline lysis method and phenol extractions (Maniatis et al. 1982). Plasmids with inserts of the proper size (not prescreened) were sequenced by Sequenase 2 (Amersham USB) initially, but later by the cycle sequencing DeltaTaq method (Amersham USB, using 30 cycles for labeling and 40 cycles for termination steps), and separated by electrophoresis with Long Ranger modified acrylamide (J.T. Baker, Phillipsburg, NJ, USA), all according to manufacturer instructions and with 'universal' sequencing primers from the kits. With the DeltaTaq method and a long electrophoresis gel apparatus (Base Runner, Kodak IBI), we were able to read about 400 to 500 bases from each reaction in a single set of 4. Sequences were aligned to the Ribosomal Database Project alignment (RDP) (Maidak et al. 1994) and analyzed by maximum likelihood, parsimony, and neighbor joining methods (Kimura 2-parameter model) from PHYLIP 3.5 or fastDNAml (Felsenstein 1981, 1993, Olsen et al. 1994), with the aid of the GDE computer program package (Smith et al. 1994; obtained from RDP, Maidak et al. 1994). Clones were first partially sequenced (at least 200 bases from both the ends) early in the project, and most of those that did not show clear affiliations with other sequences in the RDP database or other fully sequenced clones were fully sequenced (about 860 bases between the 'universal' primers). A few clones were somehow truncated, so we have only partial sequences. Due to the fragmentary nature of related clones from the database as well as some of our own, we had to create multiple trees with different masks to make all the suitable comparisons. Note that what we call 'full length' clones represent the full PCR product and plasmid insert length between the 'universal' primers, and do not represent the full 16S rRNA gene. We checked for chimeric sequences primarily by examining phylogenetic trees made separately from the 200 bases at the 5' and 3' ends of the clones, looking for significant shifts in affiliations. Initially we used the RDP program CHECK_CHIMERA, which was often useful, but we found that we needed to use caution in interpreting the results; they were difficult to interpret when the clones were distantly related to all but a few short sequence fragments in the database (as was often the case). Sometimes the CHECK_CHIMERA results appeared to indicate that a sequence may be chimeric even though we found other full-length clones from different samples that matched very closely along the entire length. The fact that they came from different samples strongly suggests they are not chimeras.

DNA sequences have been submitted to GenBank, with accession numbers U81525 to U81549 (p712), U86455 to U86487 (pN1), and U86488 to U86519 (pB1).

RESULTS

Samples were collected from depths of 500 and 3000 m in the Pacific, and 1000 m in the Atlantic. Temperature profiles, indicating the physical structure of the water column, showed well-mixed water over the top 40 m (at 20°C in Pacific, 26°C in Atlantic), a steep thermocline and rapid temperature reduction to about 100 m (where temperatures were about 9°C in the Pacific and 18°C in the Atlantic), and a more gradual temperature decline to sample depth temperatures (Table 1).

Archaea

Almost half (25 of 57) of the randomly selected clones from these 3 deep sea samples were archaeal, and these fell into 3 distinct groups (Fig. 1). Two of these groups were related to previously reported marine archaeal clones: one such group (Group I), containing 19 new clones reported here, is within the Crenarchaeota (formerly thought to contain only extreme thermophiles; Woese 1987) and includes several marine planktonic and abyssal or symbiotic archaeal clones previously reported (DeLong 1992, Fuhrman et al. 1992, 1993, DeLong et al. 1994, McInerney et al. 1995a, b, Preston et al. 1996). Relatives also include 2 soil archaeal clones (Ueda et al. 1995; see also RDP, Maidak et al. 1994) as well as clones recently found in freshwater lake sediments (Hershberger et al. 1996). The second group falls among the Euryarchaeota (containing methanogens, halophiles, and some thermophiles; Woese 1987), is distantly related (<80% sequence similarity) to *Thermoplasma*, and has also previously been reported from coastal and offshore marine plankton (DeLong 1992, Fuhrman et al. 1993). The third group is only distantly related (<80% sequence identity) to previously reported sequences, and the closest known sequences are from *Thermoplasma* and from Marine Group II (Fig. 1A).

Bacteria

Slightly over half of the clones are bacterial, and they fall into several groups (Fig. 2A, B, Table 2). Some are

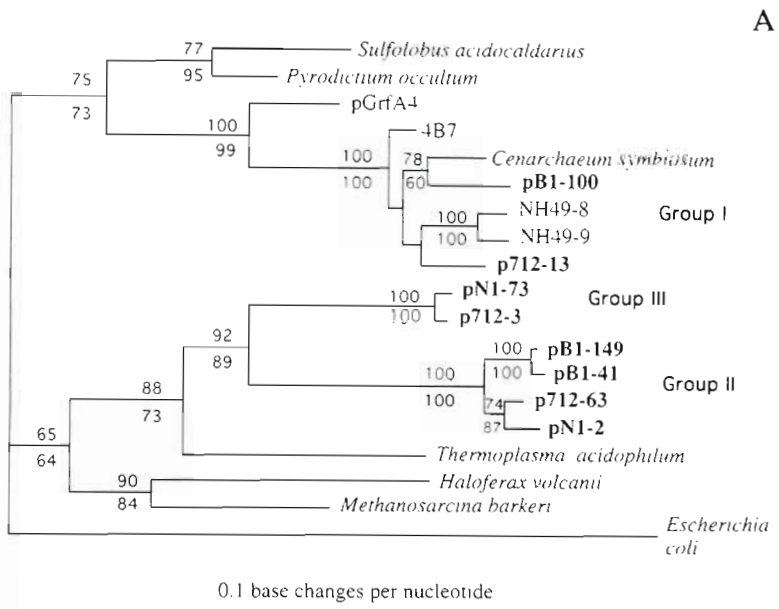


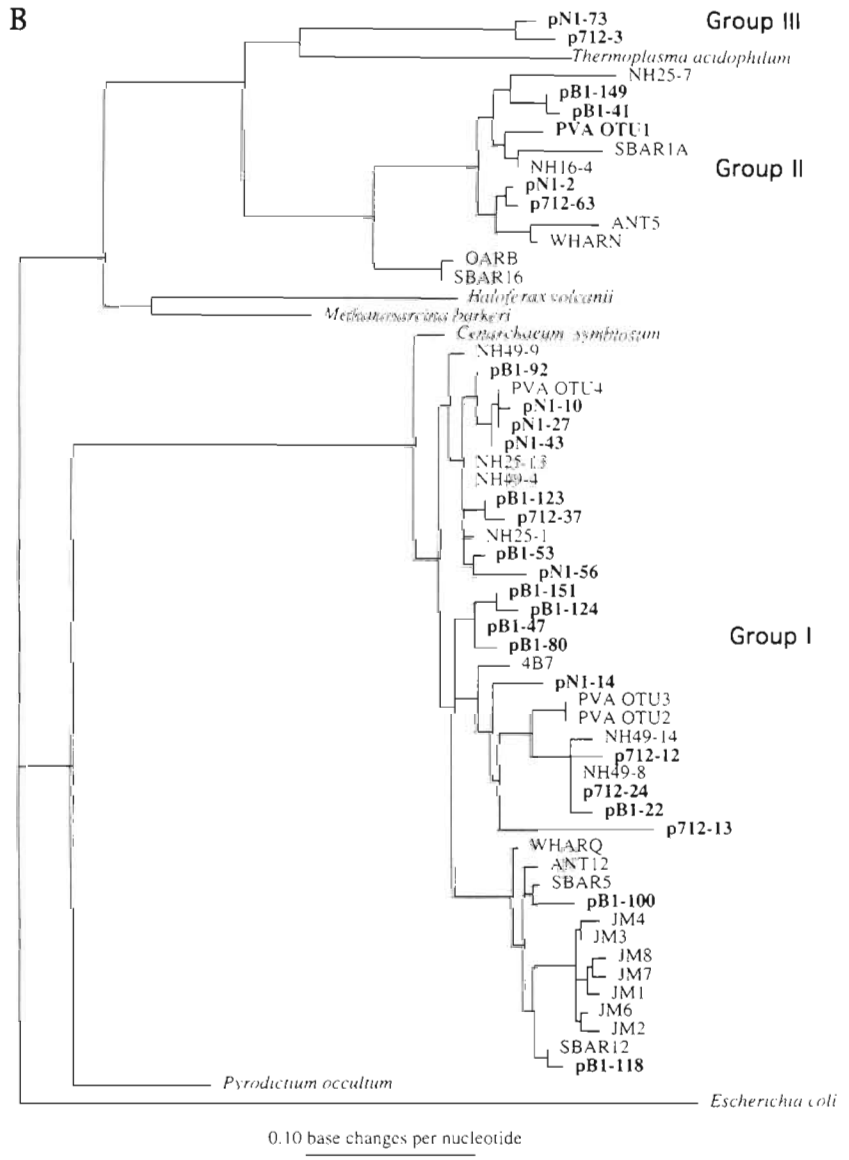
Fig. 1 (above and facing page). Archaeal clone analysis, by maximum likelihood (Olsen et al. 1994), with new clones from this report appearing in bold type and other sequences from RDP or GenBank (aligned manually to the RDP alignment). (A) Analysis from clones fully sequenced between the 'universal' primers used for amplification prior to cloning (*Escherichia coli* positions 537–1392). Numbers above the nodes represent bootstrap replicate percentages, using 100 replicates, by maximum likelihood, and those below the nodes by parsimony (PHYLIP DNAPARS, Felsenstein 1985): the bootstrap values indicate the percentage of replicate trees in which the sequences to the right of the node all group together, and bootstrap values above 50% are shown. (B) Analysis including partial sequences from this study as well as previously reported sequences. The mask (region of the molecule in which all these clones overlap) corresponds to *E. coli* positions 1232–1392. NH clones are from the same general region as the p712 and pN1 Pacific clones reported here, with NH49 from the 500 m depth and NH25 from 100 m (Fuhrman et al. 1992, 1993); WHAR clones are from Woods Hole, SBAR from Santa Barbara California, OAR from Oregon, and ANT from Antarctica (DeLong 1992, DeLong et al. 1994); JM clones are from an abyssal holothurian midgut (McInerney et al. 1995b); 4B7 is from the U.S. Pacific coast plankton (Stein et al. 1996); *C. symbiosum* is from a sponge symbiont (Preston et al. 1996); pGrfA4 is from freshwater lake sediment (Hershberger et al. 1996); and PVA clones are from a volcanic seamount near Hawaii, obtained from GenBank (submitted by C. Moyer, F. Dobbs, and D. Karl). Because shorter sequences yield less accurate phylogenies, refer to (A) for the best information on a given lineage

related specifically to previously reported clones found in subsurface waters, but others are not.

One clone from the 3000 m Pacific sample, pN1-52, branches deeply yet consistently (in 99 to 100% of the bootstrap replicates) with its closest cultured relative in the RDP database, *Thermomicrobium roseum*, and branches in all bootstrap replicates with a small cluster of clones recently reported to be found at the 250 m depth in the Sargasso Sea (Giovannoni et al. 1996) (Fig. 2A). The small yet broad phylogenetic group to which these sequences belong includes the genera *Chloroflexus*, *Herpetosiphon*, and others (Woese 1987, Giovannoni et al. 1996).

A set of clones from both oceans branch consistently, and in a high percentage of bootstrap replicates, with the group of organisms we previously called Marine Group A (Fuhrman et al. 1993) and with 2 marine clones recently reported that are very close to clone NH16-12 from Group A, OCS307 and SAR406 (Fig. 2A) (Gordon & Giovannoni 1996). Gordon & Giovannoni (1996) found that even with nearly complete 16S rRNA sequences of their 2 clones, the affiliations with bacterial groups known from cultures were remote and not fully clear, with the most likely relatives being the *Chlorobium* and *Fibrobacter* groups. All our full length clones group together in about 90% of bootstrap replicates whether analyzed with parsimony or maximum likelihood methods (Fig. 2A). Thus they appear to make up a reasonably coherent yet broad group. Two fragments from the 3' ends of apparently chimeric clones (p712-11 and pN1-16) also are included in this group (Fig. 3A, B: affiliations of fragments from chimeras, truncated clones, and those related to partial sequence fragments from the database).

These deep sea clones from both oceans also contain at least 1 other major group with little firm affiliation to known bacterial groups, and we have labeled this group Marine Group B. The 3 full-length clones (pN1-23, pB1-19, and pB1-87) branch together consistently and in all bootstrap replicates, and in some analyses branch together with the *Planctomyces*, but this affilia-



tion is not robust (Fig. 2A); some analyses place them with the Proteobacteria (not shown). Three other partial sequences also fit in this group, including the truncated clone p712-48 (3' end only) and the 5' end of the apparent chimera p712-79, whose 3' end appears proteobacterial (Fig. 3).

Clone pN1-33 may be distantly related (<80% sequence similarity) to *Planctomyces* (Fig. 2A), although the specific relationship is not robust to bootstrap analysis, and it branches deeply with Group B in some analyses. It may represent a novel major group, a deep affiliate of Group B, or possibly a chimera. In this case it is difficult to tell if it may be a chimera, given its distance from other known sequences, but analysis of the 200 bases at either end also shows no obvious relationship to anything in the database. Although we have the full length clone sequence, all we can say for sure is that pN1-33 is quite distant from any known sequence.

One clone, pB1-114, is most closely related to Japanese farm soil clone FIE18 (Ueda et al. 1995), with 90% sequence identity over the overlapping 280 base region (the 3' end of our clone, see Fig. 3B). Ueda et al. (1995) placed FIE18 in a cluster deeply branched with the Gram Positive, Low G+C Group, a placement generally confirmed by RDP (Maidak et al. 1994). In our analysis, the full length clone appeared remotely affiliated with a member of the *Fibrobacter* phylum, *Acidobacterium capsulatum*, but branches together with that sequence in only about 50 to 70% of the bootstrap replicates, depending on the type of analysis (Fig. 2A). Therefore, its affiliations could not be resolved with the data and methods we have used.

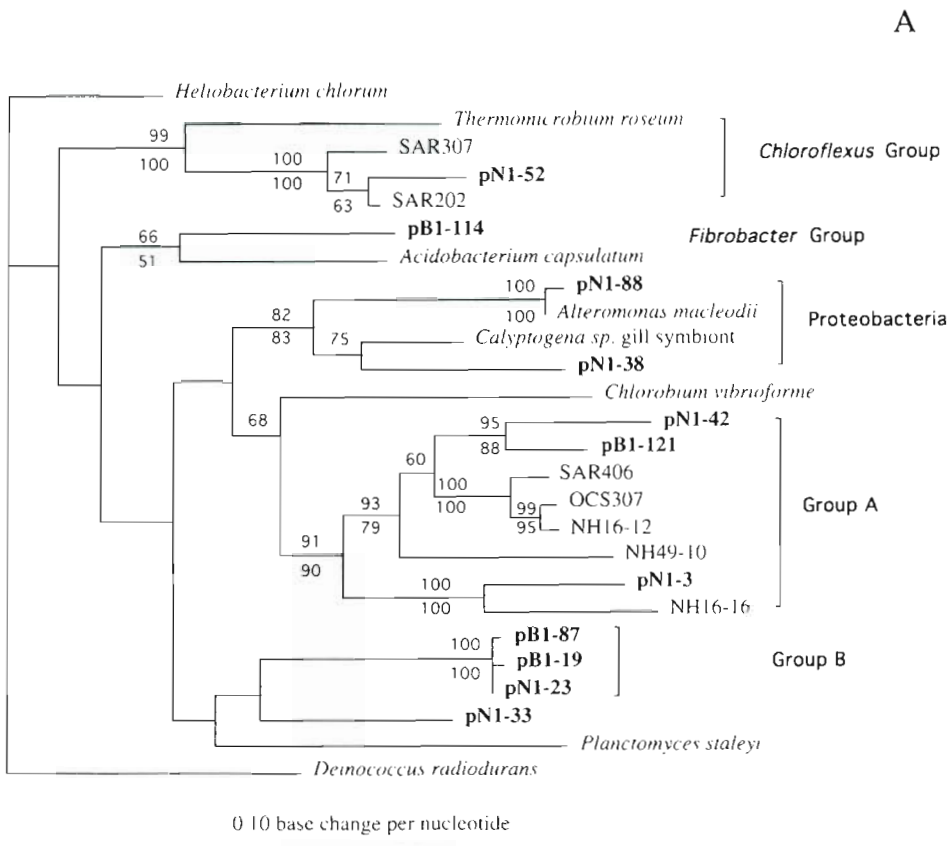
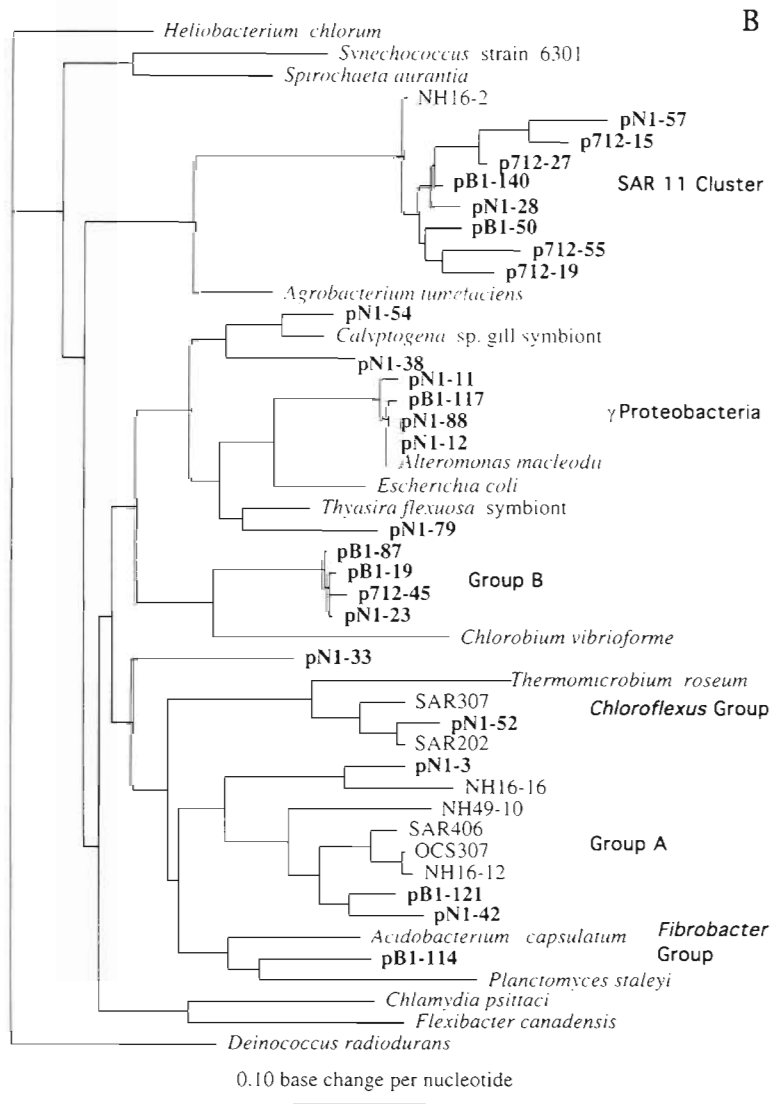


Fig. 2 (above and facing page). Bacterial clones (excluding likely chimeras) analyzed by maximum likelihood. Bold type indicates new clones from this study. (A) Full length clones (*Escherichia coli* positions 537–1392) representing the major bacterial groups found in this study. Bootstrap analyses as in Fig. 1A. (B) Full and partial clone sequences, analyzed from 200 bases from each end of the clones, corresponding to *E. coli* positions 537–737 and 1192–1392. NH clones are plankton from Fuhrman et al. (1993) and SAR and OCS clones are plankton from Gordon & Giovannoni (1996) or Giovannoni et al. (1996). Note that phylogenies from shorter sequences have reduced accuracy, so refer to (A) for the best information on a given lineage

Table 2. Summary of clones found in the 3 samples used in this study, including chimeric fragments and truncated clones

Group / relatives	p712 (Pacific, 500 m)	pN1 (Pacific, 3000 m)	pB1 (Atlantic, 1000 m)	Total[a]
Archaea Group I	4	5	10	19
Archaea Group II	1	1	2	4
Archaea Group III	1	1	0	2
Proteobacteria				
SAR 11 Relatives	4	2	2	8
Calyptogena or *Thyasira* symbiont relatives	4	4	0	8
Alteromonas macleodii relatives	0	3	2	5
NH16-18 relative	0	1	0	1
Group A	1	3	1	5
Group B	3	1	2	6
Thermomicrobium relatives	0	1	1	2
FIE18 relatives (*Fibrobacter*, Gram +?)	0	0	1	1
Unknown affiliation	0	1	0	1
Total	18	23	21	62

[a]The grand total of 62 sequences is greater than the number of clones (57) because 5 apparently chimeric clones are separated here into 2 fragments, each with different affiliations



About half of the bacterial sequences appear to be proteobacterial (Table 2). Starting with the most distant from known sequences, the clone pN1-38 is relatively distant from previously described groups, but is >90% similar to clone NH16-18 previously reported to be found by Fuhrman et al. (1993) at the 100 m depth in the same region from which this clone was collected at the 3000 m depth (Fig. 3A). The phylogenetic placement of this clone is uncertain, but about 80% of the bootstrap replicates show it to be deeply branched with the γ Proteobacteria, and all its closest relatives from RDP are in this group (Fig. 2A, B). Eight sequences (including 4 apparent chimeric fragments and 1 truncated clone) tend to branch with sulfur-oxidizing symbionts of *Calyptogena* or *Thyasira* (Figs. 2B & 3), but none are >95% similar It is interesting that 1 of the apparent chimeras (p712-1) consisted of fragments which were both related most closely to different S-oxidizing symbionts. The other apparent chimeras had fragments from widely different groups (Fig. 3).

Three clones from the 3000 m Pacific sample, and 1 from the Sargasso 1000 m sample, are extremely closely related (some almost identical) to the sequence from the cultured marine bacterium *Alteromonas macleodii*, an aerobic heterotroph (Fig. 2B). An additional clone, pB1-133 from the 1000 m Sargasso sample, is extremely close to this organism sequence over most of its 5' end, but the 160 bases at the 3' end diverge greatly from that sequence and instead align best with clone SAR307 that is related to *Thermomicrobium roseum* (Fig. 3).

A set of 8 clones are most closely related to the SAR11 cluster (Giovannoni et al. 1990), with several in a tight group very close to SAR11 itself and others a few percent distant, as is similar to the pattern previously reported for clones from shallower waters

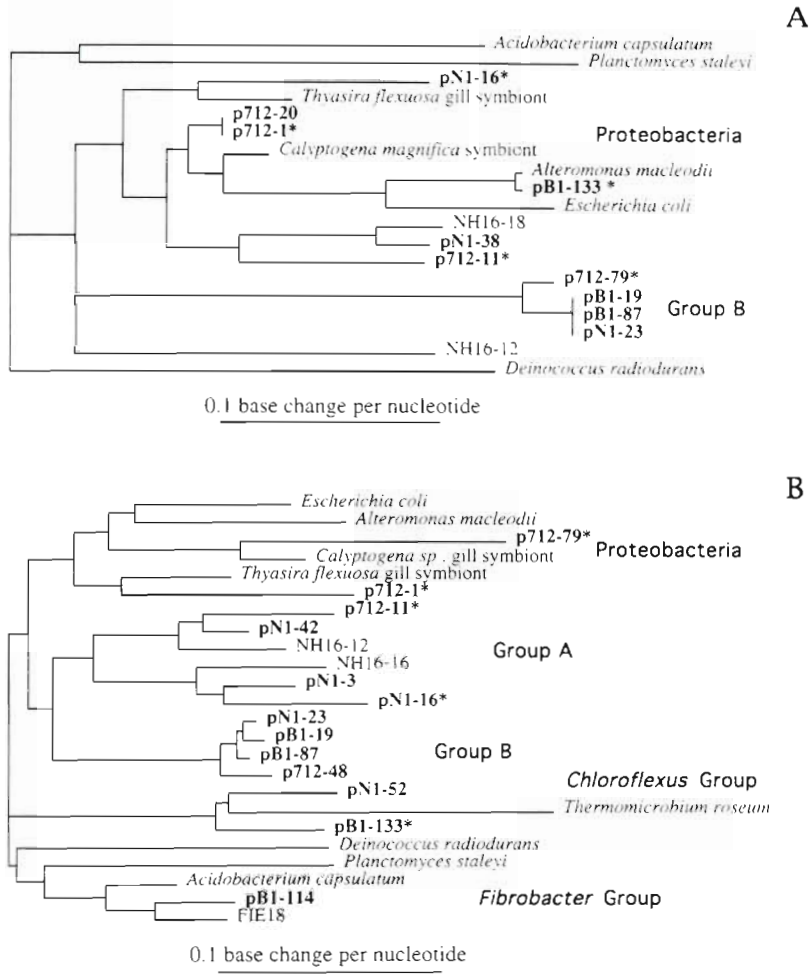


Fig. 3. Affiliations of sequence fragments from apparent chimeras (*), truncated clones, and those related to fragments from the database not amenable to analysis in Fig. 2. Bold type indicates new clones from this study. (A) Analysis from *Escherichia coli* positions 537–737 (i.e. 200 bases from the 5' end of the clones) by maximum likelihood. (B) Analysis from *E. coli* positions 1192–1392 (200 bases from the 3' end of the clones), by neighbor joining (with these data, maximum likelihood yielded an incorrect topology). NH clones are plankton from Fuhrman et al. (1993). Clone FIE 18 is from soil (Ueda et al. 1995). Note that phylogenies from shorter sequences have reduced accuracy, so refer to Fig. 2A for the best information on a particular lineage

(Fuhrman et al. 1993) (Fig. 4). Clones from this cluster reported by Schmidt et al. (1991) do not overlap with these, so cannot be included in the comparison.

DISCUSSION

It is striking that by far the most common group in these planktonic deep sea clones, with a total of 25 randomly isolated clones out of a total of (57 + 10) from the 500 to 3000 m samples reported here and by Fuhrman et al. (1992, 1993), is Archaeal Group I (the archaeal cluster related to clone NH49-9). Several other members of this cluster have been reported from surface waters of temperate and polar waters, sequenced by means of PCR with archaeal-specific primers, and from the guts of abyssal holothurians or within sponges (DeLong 1992, DeLong et al. 1994, McInerney et al. 1995a, b, Preston et al. 1996); relatives have also been found in soils (Ueda et al. 1995) and freshwater sediments (Hershberger et al. 1996). All of the marine clones, irrespective of their source, are much more closely related to each other than they are to any other known clone or culture sequence (Fig 1B). Among the marine clones, there may be some subgrouping within the cluster (Fig. 1B), and the data suggest that the clones from lake sediments branch outside the marine cluster (Fig. 1A and Hershberger et al. 1996) as do those from soils (Ueda et al. 1995; RDP, Maidak et al. 1994). Among the deep plankton, the groupings within the cluster are not geographic or depth-related.

Given that we used 'universal' primers rather than ones specific for Archaea, the preponderance of archaeal clones in our set suggests that the Archaea are probably relatively common in the deep sea locations studied. While we do not know if we can interpret the relative abundances of clones as approximating the relative abundance of organisms, a report by Suzuki & Giovannoni (1996) indicates that, with the same 'universal' primers we used, the proportion of amplified products was close to the proportion of original template concentrations of 2 different 16S rRNA gene templates in laboratory tests. Thus, the proportions of clones in our study may be a general indication of the proportions of genes in the DNA extracts. Assuming little discrimination in DNA extraction or purification, this may indicate that perhaps 10s of % of the deep sea prokaryotes are Archaea. However, as Farrelly et al. (1995) have pointed out, a quantitative conclusion depends on knowing the genome sizes and rRNA gene copy numbers in the various organisms, which are unknown and may not be uniform. Better estimates await the application of group-specific probes to deep sea samples, preferably by fluorescent *in situ* hybridization of individual bacteria (Amann et al. 1995). Along these lines,

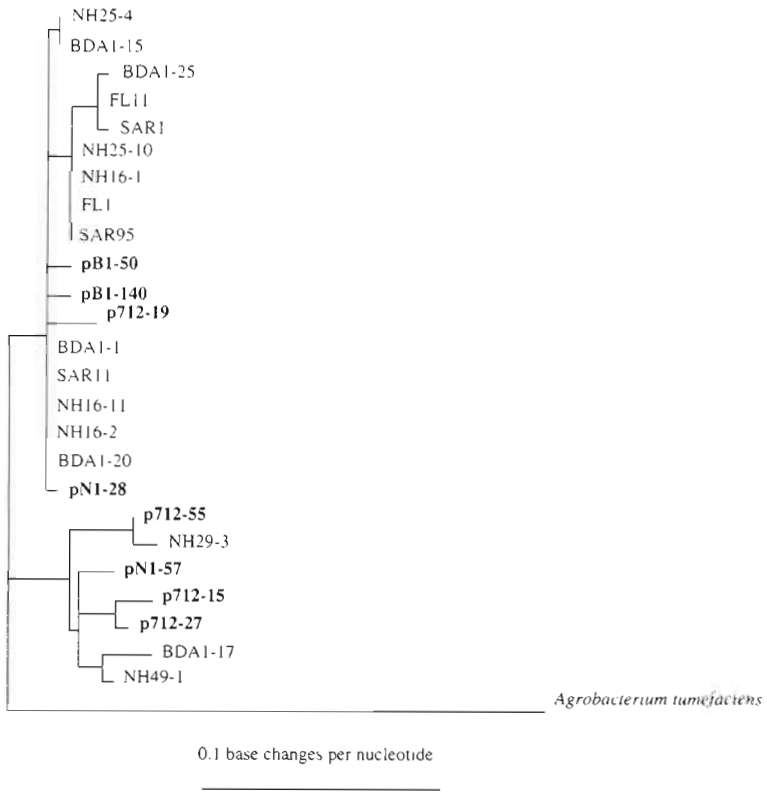


Fig. 4. Relationships among members of the SAR11 cluster. Analysis from *Escherichia coli* positions 537–737. Bold type indicates new clones from this study. SAR clones are from the Sargasso Sea near Bermuda (Giovannoni et al. 1990, Britschgi & Giovannoni 1991), NH from the California Current and BDA from the Sargasso Sea near Bermuda (Fuhrman et al. 1993), and FL from coastal California (DeLong et al. 1993)

we have been able to visualize marine Group I Archaea from seawater, with fluorescent probes and image-processed video microscopy (Fuhrman et al. 1994). Applying this approach to samples from the North Pacific, we have detected approximately 10 to 20% Archaea in several samples, and sometimes more (manuscript in preparation). This is generally consistent with the clone abundances.

Our report of a third, and previously undescribed, lineage of marine planktonic archaea further increases the known taxonomic breadth of the Archaea in general, and microbial plankton in particular. Until the reports of Fuhrman et al. (1992, 1993), DeLong (1992) and DeLong et al. (1994), Archaea were almost unknown from marine plankton. Now they appear to be readily detected, even with 'universal' PCR primers, at least in samples from the deeper waters.

We have little or no information on the physiology of these archaeal groups. Fuhrman et al. (1994) previously reported enrichment of an open-ocean near-surface seawater sample with Group I Archaea by addition of small amounts of peptone and yeast extract, simultaneous with antibiotic treatments to inhibit eubacteria. It seems quite likely that the Group I Archaea are psychrophilic or mesophilic aerobic heterotrophs, and, if so, this presents a strong departure from the concept that Archaea are either extreme thermophiles, extreme halophiles, or methanogens, all unable to compete well against eubacteria or eukaryotes. A similar conclusion was drawn by Preston et al. (1996) in their report of the crenarchaeotal symbiont of a marine sponge.

The 3 basic types of archaeal clones include 2 (Groups II and III) that place phylogenetically closest to the methanogens. While we do not yet know if these organisms are methanogenic, it is interesting to note that many parts of the aerobic ocean are saturated or supersaturated with methane from an apparent midwater source (DeAngelis & Lee 1994). The organisms responsible for the methane production are not known, but methanogenesis is apparently occurring within presumably anaerobic microzones in zooplankton guts (DeAngelis & Lee 1994). If so, it is possible that some of the Archaea we observed are methanogens that have been released from the guts or from decaying fecal matter, or possibly may live in anaerobic microzones, such as on organic particles, within the plankton.

The several clones in the newly reported Marine Group B appear to represent a new bacterial lineage that is not yet represented in the extensive 16S rRNA database. Our analyses do not allow us to place this group with confidence in relationship to others. The sequences within this group are all closely related to each other, yet together branch deeply with other sequences from the database. This is the sort of pattern observed with the near-surface representatives of the SAR11 cluster (e.g. Fig 4). We found that this group was present in both ocean basins, suggesting a broad, possibly global, distribution. The close relations within the group and distance to other known sequences allows for relatively easy probe development.

The relationships among what we call Marine Group A differ significantly in character from Group B in that its members are not closely related to each other (Figs. 2 & 3). It is interesting that there can be several members of this group so deeply branched, yet all are from marine plankton clones, and none from cultures or other habitats. The depth of the branches suggests a geologically early radiation of this lineage with signifi-

cant diversity, yet the lineage is not represented in culture databases.

One of the clones, pN1-52, plus a fragment of apparently-chimeric clone pB1-133 are deeply branched within the *Chloroflexus/Herpetosiphon/Thermomicrobium* lineage (Figs. 2 & 3) Clones from this group occurred in both ocean basins. Members of this group were reported to be cloned from 250 m depth in the Sargasso Sea, and a probe designed to bind to this group hybridized to RNA and PCR-amplified 16S rRNA genes from the top 250 m of Sargasso Sea and Oregon coast samples (Giovannoni et al. 1996). The probe binding was strongest at the lower portion of the subsurface chlorophyll maximum, with significant stratification. The authors concluded that the members of this group are adapted for growth in a discrete depth of the water column. It should be noted that our clones, from 1000 and 3000 m, were considerably below the euphotic zone and chlorophyll maximum layers. The phylogenetic position gives few physiologic clues, given the paucity of close relatives from the culture database and the physiologic variability of known members from this lineage. *Chloroflexus* is photosynthetic (*Herpetosiphon* and *Thermomicrobium* are not), while *Chloroflexus* and *Thermomicrobium* are thermophiles (*Herpetosiphon* is not). It does not seem possible that any organism could be actively photosynthetic in the dark waters at 3000 and 1000 m where these clones were found, although it is possible that such organisms originated in better lit waters and were transported to depth (e.g attached to sinking particles). Similarly, circumstances suggest that the organisms are not thermophiles. Results suggest that the organisms are not particularly rare and could well be adapted for life in the deep sea.

Overall, the results of this cloning and sequencing study point to the common occurrence of novel microbial lineages, some quite distant from any culture in the database, in the deep sea. Both Atlantic and Pacific Ocean basins have clones representing most of the same groups, which is not surprising given the ancient nature of this habitat and the geologically short time scales of global deep sea circulation. Although it was unusual for the clones to be close to any previously cultured organism, one cluster of 5, from both ocean basins, was related closely to the aerobic marine heterotroph *Alteromonas macleodii*. We cannot say at this time if the other clones represent organisms that are difficult to culture, or simply if the lack of cultured relatives reflects the gross undersampling of the deep sea. Given the size of this habitat and logistic difficulty in growing cultures that may require *in situ* temperature and pressure conditions, it may be some time before we learn the answer. However, the clones reported here may be used for the development of probes to investigate the *in situ* distribution of the organisms in the various groups we have reported.

Acknowledgements. This work was suported by U.S. NSF grant numbers DEB9200928 and DEB9401110. We thank Dr SangHoon Lee for sampling and DNA extraction, David Foley for providing temperature and salinity data for the region and depth of the pN1 clones, Ximena Hernandez, Joel Mefford, and Trista Wagoner for assisting lab work and analysis, and Dr Norman Pace for helpful comments on the manuscript.

LITERATURE CITED

Amann R, Ludwig L, Schleifer K (1995) Phylogenetic identification and *in situ* detection of individual microbial cells without cultivation. Microbiol Rev 59:143–169

Britschgi T, Giovannoni SJ (1991) Phylogenetic analysis of a natural marine bacterioplankton population by rRNA gene cloning and sequencing. Appl Environ Microbiol 57: 1707–1713

DeAngelis MA, Lee C (1994) Methane production during zooplankton grazing on marine phytoplankton. Limnol Oceanogr 39:1298–1308

DeLong EF (1992) Archaea in coastal marine environments. Proc Natl Acad Sci USA 89:5685–5689

DeLong EF, Franks DG, Alldredge AA (1993) Phylogenetic diversity of aggregate-attached vs. free-living marine bacterial assemblages. Limnol Oceanogr 38:924–934

DeLong EF, Wu KY, Prezelin BB, Jovine RVM (1994) High abundance of archaea in antarctic marine picoplankton. Nature 371:695–697

Farrelly V, Rainey FA, Stackebrandt E (1995) Effect of genome size and rrn gene copy number on PCR amplification of 16s rRNA genes from a mixture of bacterial species. Appl Environ Microbiol 61:2798–2801

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 (1993) Phylip 3.5. Phylogeny inference package. Distributed by author, University of Washington, Seattle

Fuhrman JA, Comeau DE, Hagstrom A, Chan AM (1988) Extraction of DNA suitable for molecular biological studies from natural planktonic microorganisms. Appl Environ Microbiol 54:1426–1429

Fuhrman JA, Lee SH, Masuchi Y, Davis AA, Wilcox RM (1994) Characterization of marine prokaryotic communities via DNA and RNA. Microb Ecol 28:133–145

Fuhrman JA, McCallum K, Davis AA (1992) Novel major archaebacterial group from marine plankton. Nature 356: 148–149

Fuhrman JA, McCallum K, Davis AA (1993) Phylogenetic diversity of marine subsurface microbial communities from the Atlantic and Pacific Oceans. Appl Environ Microbiol 59:1294–1302

Giovannoni SJ, Britschgi TB, Moyer CL, Field KG (1990) Genetic diversity in Sargasso Sea bacterioplankton. Nature 345:60–63

Giovannoni SJ, Rappe MS, Vergin KL, Adair NI (1996) 16S rRNA genes reveal stratified open ocean bacterioplankton populations related to the Green Nonsulfur bacteria. Proc Natl Acad Sci USA 93:7979–7984

Gordon DA, Giovannoni SJ (1996) Detection of stratified microbial population related to *Chlorobium* and *Fibro-*

bacter species in the Atlantic and Pacific Oceans. Appl Environ Microbiol 62:1171–1177
Hershberger KL, Barns SM, Reysenbach AL, Dawson SC, Pace NR (1996) Wide diversity of Crenarchaeota. Nature 384:420
Maidak BL, Larsen N, McCaughey MJ, Overbeek R, Olsen GJ, Fogel K, Blandy J, Woese CR (1994) The ribosomal database project. Nucleic Acids Res 22:3485–3487
Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning, a laboratory manual. Cold Springs Harbor Laboratory, Cold Springs Harbor
McInerney JO, Paskins L, Eardly D, Patching JW, Powell R (1995a) Extraction of prokaryotic genomic DNA from marine microbial communities suitable for amplification using the polymerase chain reaction. Int Rev ges Hydrobiol 80:351–360
McInerney JO, Wilkinson M, Patching JW, Embley TM, Powell R (1995b) Recovery and phylogenetic analysis of novel archaeal rRNA sequences from a deep-sea deposit feeder. Appl Environ Microbiol 61:1646–1648
Mullins TD, Britschgi TB, Krest RL, Giovannoni SJ (1995) Genetic comparisons reveal the same unknown bacterial lineages in Atlantic and Pacific bacterioplankton communities. Limnol Oceanogr 41:148–158
Olsen GJ, Lane DL, Giovannoni SJ, Pace NR (1986) Microbial ecology and evolution: a ribosomal RNA approach. Annu Rev Microbiol 40:337–365
Olsen GJ, Matsuda H, Hagstrom R, Overbeek R (1994) fastDNAml: a tool for construction of phylogenetic trees of DNA sequences using maximum likelihood. Comput Appl Biosci 10:41–48
Pace NR, Stahl DA, Lane DL, Olsen GJ (1986) The analysis of natural microbial populations by rRNA sequences. Adv Microbiol Ecol 9:1–55
Preston CM, Wu KY, Molinski TF, DeLong EF (1996) A psychrophilic crenarchaeon inhabits a marine sponge: *Cenarchaeum symbiosum* gen. nov., sp. nov. Proc Natl Acad Sci USA 93:6241–6246
Schmidt TM, DeLong EF, Pace NR (1991) Analysis of a marine picoplankton community by 16S rRNA gene cloning and sequencing. J Bacteriol 173:4371–4378
Smith SW, Overbeek R, Woese CR, Gilbert W, Gillevet PM (1994) The genetic data environment an expandable GUI for multiple sequence analysis. CABIOS 10:671–675
Stein JL, Marsh TL, Wu KY, Shizuya H, DeLong EF (1996) Characterization of uncultivated prokaryotes—isolation and analysis of a 40-kilobase-pair genome fragment from a planktonic marine archaeon. J Bacteriol 178:591–599
Suzuki MT, Giovannoni SJ (1996) Bias caused by template annealing in the amplification of mixtures of 16s rRNA genes by PCR. Appl Environ Microbiol 62:625–630
Ueda T, Suga Y, Matsuguchi T (1995) Molecular phylogenetic analysis of a soil microbial community in a soybean field. Europ J Soil Sci 46:415–421
Woese CR (1987) Bacterial evolution. Microbiol Rev 51: 221–271

This article was presented by F. Rassoulzadegan (Senior Editorial Advisor), Villefranche-sur-Mer, France

Manuscript first received: November 14, 1996
Revised version accepted: January 24, 1997