APPLIED AND ENVIRONMENTAL MICROBIOLOGY, Apr. 1996, p. 1171–1177 Vol. 62, No. 4
0099-2240/96/$04.00+0

Detection of Stratified Microbial Populations Related to *Chlorobium* and *Fibrobacter* Species in the Atlantic and Pacific Oceans

D. A. GORDON AND S. J. GIOVANNONI*

Department of Microbiology, Oregon State University, Corvallis, Oregon 97331

Received 26 June 1995/Accepted 31 December 1995

A gene lineage (SAR406) related to *Chlorobium* and *Fibrobacter* species was found in 16S rRNA gene clone libraries prepared from samples from two oceans. The clone libraries were constructed from total picoplankton genomic DNA to assess bacterial diversity in the lower surface layer. The samples were collected by filtration from a depth of 80 m at a site in the western Sargasso Sea and from a depth of 120 m at a site in the Pacific Ocean, approximately 70 km from the Oregon coast. The PCR and primers which amplified nearly full-length 16S rRNA genes were used to prepare the clone libraries. Among the diverse gene clones in these libraries were two related clones (SAR406 and OCS307) which could not be assigned to any of the major bacterial phyla. Phylogenetic analyses demonstrated that these genes were distant relatives of the genus *Fibrobacter* and the green sulfur bacterial phylum, which includes the genus *Chlorobium*. The inclusion of SAR406 in phylogenetic trees inferred by several methods resulted in support from bootstrap replicates for the conclusion that *Fibrobacter* and *Chlorobium* species and SAR406 are a monophyletic group. An oligonucleotide probe that selectively hybridized to clone SAR406 was used to examine the distribution of this gene lineage in vertical profiles from the Atlantic and Pacific Oceans and in monthly time series at 0 and 200 m in the Atlantic Ocean. During stratified periods, the genes were most abundant slightly below the deep chlorophyll layer. Seasonal changes in the surface abundance of SAR406 rDNA were highly correlated with chlorophyll *a* levels ($r = 0.75$).

Recent studies have used molecular cloning methods to explore bacterial diversity without the biases imposed by cultivation on microbiological media (1). The application of molecular techniques, including the cloning of environmental 16S rRNA genes, has provided researchers with new tools to address questions of microbial ecology as well as biodiversity. Comparisons of novel environmental 16S rRNA genes with databases of characterized genes allow unidentified sequences to be placed in a phylogenetic context. The sequence information can also be used to design oligonucleotide probes for use in studying the spatial and temporal dynamics of uncharacterized organisms in environmental samples (7). Molecular techniques thus augment microbial cultivation as a means to detect and study new organisms. Evidence of many novel, and as-yet-uncultivated, organisms has been found in this manner (3, 13, 15, 29).

In marine environments, bacterioplankton constitute a major component of the biomass, but the abundance and physiological activities of specific species are largely unknown (2, 8). The work presented here is part of a study examining the dynamics of microbial communities at a specific location, the Bermuda Atlantic Time Series Station (BATS), in an unusually oligotrophic region in the western Sargasso Sea (4). It has been suggested that long-term changes in the stratification of planktonic communities in oligotrophic subtropical gyre ecosystems might influence ocean surface layer geochemical cycles (19). Our purpose was to determine whether Sargasso Sea bacterioplankton populations are structured temporally and spatially and whether fluctuations in the density and activity of these populations are correlated with specific physical and chemical variables.

We report here the preliminary characterization of a 16S rRNA gene lineage cloned originally from water samples collected at BATS, from a depth of 80 m, in August of 1991. This gene clone, SAR406, could not be assigned to any of the major bacterial phyla but was closely related to a fragmentary sequence recovered previously from seawater by Fuhrman and coworkers (12). Analysis of the full-length gene sequences indicates that they constitute a deeply diverging lineage but are most closely related to green sulfur bacteria and *Fibrobacter* spp. The SAR406 lineage was found to have a stratified distribution, contributing proportionately more to bacterioplankton abundance in the lower surface layer than at the surface.

MATERIALS AND METHODS

Sampling. Water samples were collected with Niskin bottles attached to a CTD (conductivity, temperature, and depth) rosette from BATS (31°50′N, 64°10′W) and at a site in the Pacific Ocean located approximately 70 km from the Oregon coast (46°45′N, 126°2′W). Monthly time series samples spanning the period from August 1991 to February 1994 were collected at the BATS by the same method. Samples were collected from at least two depths (0 and 200 m) each month. Additional depths were sometimes sampled. Twelve to 24 liters of seawater from each depth was filtered onto 0.2-μm-pore-size polysulfone filters (SUPOR-200; Gelman Sciences Inc., Ann Arbor, Mich.) under vacuum at 7 lb/in^2. The filters were then placed in heat-sealing plastic bags with 5 ml of lysis buffer (20 mM EDTA, 400 mM NaCl, 0.75 M sucrose, and 50 mM Tris HCl [pH 9.0]) and stored at −20°C or in liquid N_2 for later processing on shore. A Sea-Bird CTD was used to measure continuous profiles of temperature and fluorescence.

Nucleic acid extraction. Total cellular nucleic acids were extracted from the filters by use of a combination of procedures optimized for small sample sizes. To lyse cells, the samples were thawed on ice, sodium dodecyl sulfate (SDS) was added to 1%, and proteinase K was added to 100 μg/ml. The bags were then resealed and incubated in an HB-1 hybridization oven (Techne, Cambridge, United Kingdom) at 37°C for 30 min and then at 55°C for 10 min. Lysate was transferred to polypropylene centrifuge tubes, extracted with an equal volume of fresh phenol-$CHCl_3$-isoamyl alcohol (25:24:1), and again extracted with an equal volume of $CHCl_3$-isoamyl alcohol (24:1). The purified nucleic acids were then precipitated with 2 volumes of 100% ethanol and 0.1 volume of 2 M sodium acetate (pH 5.2), collected by centrifugation, and resuspended in aqueous solution. Cellular RNA and DNA were fractionated by isopycnic centrifugation in cesium trifluoroacetate (1.6 g/cm^3) under the following conditions: 1.4-ml volume for 48 h at 40,000 rpm, 15°C, in a TLS-55 rotor (Beckman Inc., Palo Alto, Calif.). DNA and RNA were precipitated from the appropriate fractions by the

* Corresponding author. Phone: (541) 737-1835. Fax: (541) 737-0496. Electronic mail address: giovanns@bcc.orst.edu.

addition of 1 volume of isopropanol and washed in 70% ethanol. The resulting pellets were then resuspended in Tris-EDTA buffer.

Cloning. Prokaryotic 16S rRNAs were amplified for cloning from the mixed population genomic DNAs by PCR with *Taq* polymerase (Promega, Madison, Wis.) and bacterial 16S primers for DNA coding for rRNA (rDNA primers 27F [AGA GTT TGA TCM TGG CTC AG] and 1518R [AAG GAG GTG ATC CAN CCR CA]) (14). The reaction volumes were 100 μl, containing 10 ng of template, 5% acetamide, 200 nM each primer, 1.5 mM $MgCl_2$, 200 μM total deoxynucleoside triphosphate (dNTP), standard *Taq* buffer, and 2.5 U of *Taq* polymerase. Thirty-five cycles were performed with a Coy thermocycler under the following conditions: annealing for 1 min at 50°C, elongation at 72°C for 3 min, and denaturation at 94°C for 1 min. The clone library was constructed by ligating the amplification products into the plasmid vector pCRII (Invitrogen, San Diego, Calif.) as described in the manufacturer's instructions. Double-stranded plasmid DNA sequencing was performed with an Applied Biosystems 373A automated sequencer. The sequences were determined bidirectionally.

Sequence analysis. Sequence analysis was performed with the programs GDE, supplied by Steve Smith (Millipore, Bedford, Mass.), and gRNAid, supplied by Shannon Whitmore (Mentor Graphics, Wilsonville, Oreg.). Regions of ambiguous alignment were excluded from the phylogenetic analysis. The clone sequences were evaluated by the program CHECK_CHIMERA, provided by the Ribosomal Database Project (RDP) (22). This program has been designed to aid in the identification of chimeric gene artifacts. The programs Phylogeny Inference Package (Phylip) version 3.5 (11) and fastDNAml (10, 24) were used to infer phylogenetic relationships.

Accession numbers. Nucleotide sequences were filed in GenBank under the accession numbers U41450 for SAR406 and U34043 for OCS307.

DNA vertical profiles. Bacterial rDNAs were amplified from seawater for dot blots with the primers 27F and 1492R (GGT TAC CTT GTT ACG ACT T). Unreacted oligonucleotides and dNTPs were removed with a Qiaquick-spin PCR purification kit (Qiagen, Chatsworth, Calif.) prior to quantifying yield by spectrophotometry. The reaction products were resuspended in 180 μl of Tris-EDTA buffer, denatured by the addition of 20 μl of 2.0 N NaOH, incubated for 10 min at room temperature, and blotted on a Zetaprobe membrane (Bio-Rad Laboratories, Richmond, Calif.) with a dot-blot manifold (Minifold I SRC 096/90; Schleicher & Schuell, Inc., Keene, N.H.). Each well was washed with 200 μl of 2× SSPE (1× SSPE is 0.18 M NaCl, 10 mM NaH_2PO_4, and 1 mM EDTA [pH 7.7]) to neutralize the NaOH. The amounts of amplified DNA blotted were 100 ng per well for natural populations and 30 ng per well for controls. The blots were dried in vacuo at 70 to 80°C for 15 min, exposed to 120 mJ of short wavelength UV radiation per m^2 with a UVC-515 Multilinker (Ultra-Lum, Inc., Carson, Calif.), and stored desiccated before probing.

RNA vertical profiles. RNA concentrations were determined by measuring optical A_{260}. One hundred, 50, 20, and 10 ng of each RNA sample were heated to 37°C for 2 min, dot blotted, and fixed as described above. Data were analyzed as described previously (15).

Oligonucleotide probes. An oligonucleotide probe specific for the SAR406 lineage (SAR406R; AAG GAT CCG CTG CAT TAT TCG) was designed to complement the 16S rRNA domain corresponding to positions 176 to 188 of the *Escherichia coli* 16S rRNA. T4 polynucleotide kinase was used to label the 5′ terminus of the oligonucleotide probe with ^{32}P (26). The probe was prescreened for potential cross-reactivity by use of the program CHECK_PROBE provided by the RDP.

Hybridization. Blots were prehybridized in 10 ml of Z-hyb buffer (1 mM EDTA, 0.25 M Na_2HPO_4 [pH 7.2], 7% SDS) for 10 to 30 min at room temperature, the prehybridization buffer was decanted, and 6 ml of fresh Z-hyb buffer containing 5 to 50 ng of oligonucleotide was added. Blots were hybridized at 30°C for 5 h to overnight, washed three times for 15 min at room temperature in 25 ml of 0.2× SSPE–0.1% SDS, and one time for 15 min at 45°C. The empirical melting temperature (T_m) of the probe was determined by quantifying the amount of probe hybridized to dot blots of SAR406 rDNA after 15-min washes at a range of increasing temperatures (25 to 65°C). Blots were placed in Seal-a-Meal bags on top of wet 3 MM blotter paper (Whatman, Hillsboro, Oreg.). Hybridization was quantified with an AMBIS Mark II radioanalytic system (Automated Microbiological Systems, San Diego, Calif.) and a Phosphorimager SI (Molecular Dynamics, Sunnyvale, Calif.). The following equation was used to calculate the specific hybridization of the SAR406 oligonucleotide probe to the depth profiles:

$$\Phi x = \left[\left(\frac{Px}{Pc} \right) \cdot \left(\frac{P'x}{P'c} \right)^{-1} \right] \times 100$$

where Φx is the specific hybridization (percentage of PCR products hybridizing to probe x), P is the hybridization of probe to sample rDNA, P' is the hybridization of probe to pure, known control rDNA, x is the specific probe, and c is the bacterial probe (338R). Bound probe was removed from the blots by washing twice in 0.2× SSPE–0.1% SDS buffer at 70°C for 10 min.

RESULTS

A library of 97 bacterial 16S rRNA gene clones from the Sargasso Sea 80-m sample was screened by hybridization with radiolabelled oligonucleotide probes specific for some bacterial gene lineages (SAR11, SAR202, SAR6, SAR7, and SAR83) that had been encountered previously in other clone libraries (4, 15, 16). Fifty-five of the clones hybridized to these taxon-specific probes in this initial screening. Unidentified clones were chosen at random for phylogenetic identification by sequencing of the 5′ and/or 3′ regions. These domains of the gene contain both highly conserved and highly variable regions suitable for preliminary phylogenetic analysis and probe design. Separate phylogenetic analyses of the 5′ and 3′ domains of SAR406 did not reveal specific affiliations to sequences representative of the major bacterial phyla. The nucleotide sequence was completed and then evaluated with the RDP program SIM_RANK, yielding an S_{AB} of 0.799 with the environmental gene clone NH16-12 as the highest value obtained (12). NH16-12 is part of a gene cluster (marine group A) first described by Fuhrman and colleagues on the basis of partial rRNA gene sequences. This 430-base gene fragment differs in 17 positions from SAR406 and is thus 96% similar over the 430 bases available for comparison. The two other members of this cluster, NH49-10 and NH16-16, are 81 and 79% similar, respectively, to SAR406. This similarity was not immediately recognized because the 5′ and 3′ sequences of SAR406 determined initially did not overlap with the 430-base NH16-12 fragment.

An analysis of signature nucleotides and secondary structural models supported the conclusion that SAR406 and OCS307 are closely related representatives of a gene lineage naturally occurring in seawater. A secondary structural model of the predicted SAR406 and OCS307 16S rRNA molecules is shown in Fig. 1. The two sequences are consistent with a single secondary structure in which a majority of the substitutions differentiating the two clones are located in hypervariable regions distributed throughout the molecule. Substitutions were found in variable regions V1, V2, V3, V5, V6, V7, V8, and V9, i.e., all of the variable regions defined for prokaryotes (6).

Clones SAR406 and OCS307 exhibited several unusual base substitutions at conserved positions (Table 1). Some of these substitutions are shared with some members of other phylogenetic groups, including *Bacteroides* spp. (30). However, the overall pattern of substitutions in SAR406 and OCS307 is unique, and the two clones could not be associated with any of the taxa listed in Table 1 by signature nucleotides alone.

A variety of phylogenetic methods were employed to determine the evolutionary position of SAR406. In separate phylogenetic analyses, the 5′ and 3′ regions of the SAR406 gene behaved similarly; neither domain alone showed a significant affiliation with other phylogenetic groups. This result was consistent with the observations of Fuhrman and colleagues, who were unable to assign a phylogenetic position to clone NH16-12 by analysis of a partial sequence.

The phylogenetic relationships of the complete SAR406 gene are shown in Fig. 2. This phylogenetic tree was inferred by the neighbor-joining method from the full SAR406 sequence and selected 16S rRNA sequences representing the bacterial phyla. A total of 1,297 positions were included in the mask, which excluded those positions for which there were no homologs among every sequence in the analysis. The gene sequences of two archaea (*Methanococcus voltae* and *Haloferax volcanii*) were used as an outgroup. Fast DNA Maximum Likelihood (fastDNAml) (10, 24) and parsimony methods were used to derive phylogenetic trees from the same data set. Bootstrap resampling (100 replicates) of the data was used to provide statistical support for the phylogenetic relationships obtained with each method of inference. Bootstrap values for parsimony and neighbor-joining methods are shown in Fig. 2,

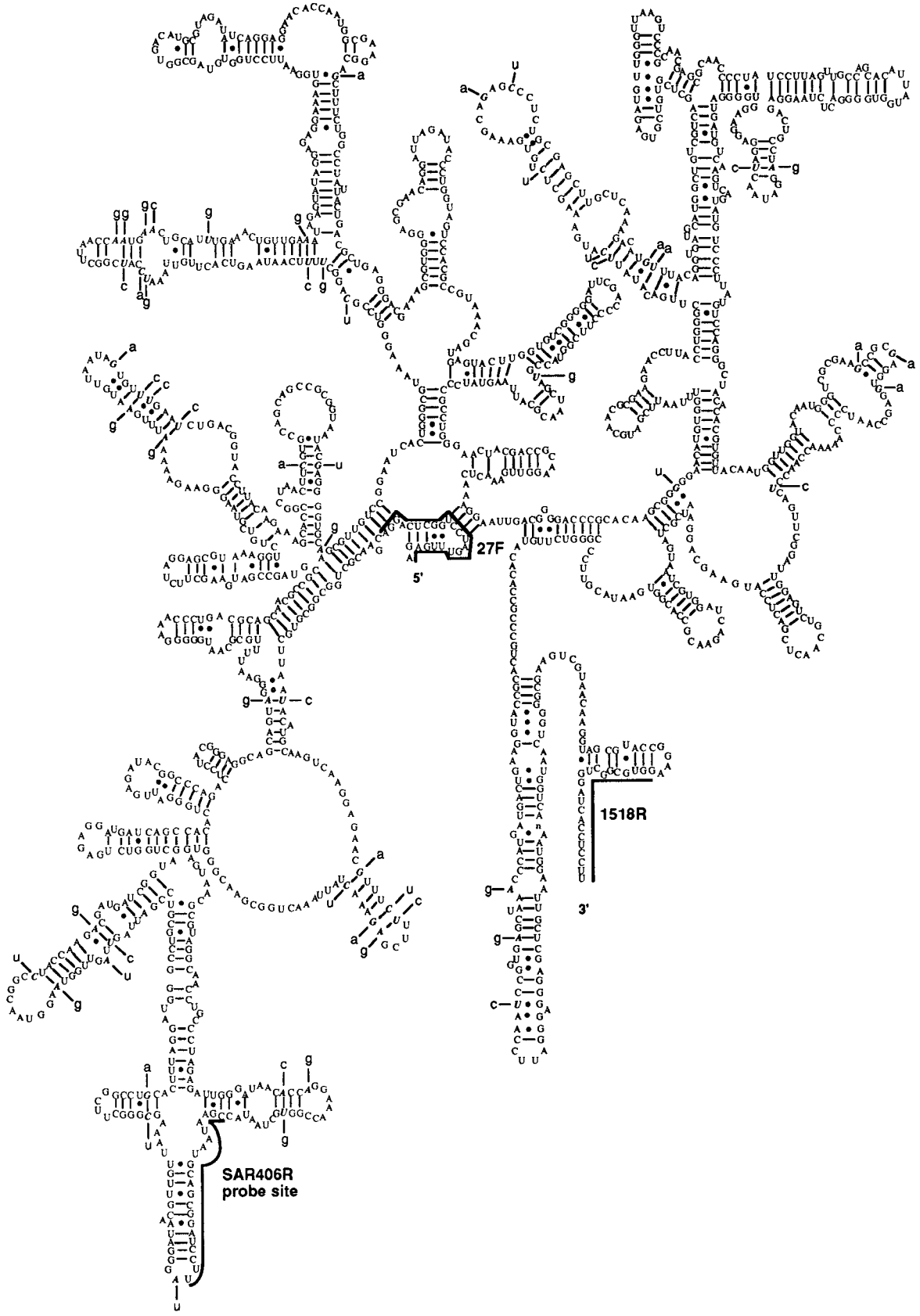

FIG. 1. Model of the secondary structure of the SAR406 and OCS307 16S rRNA molecules. Positions at which the sequence of OCS307 differs from that of SAR406 are indicated by lowercase letters. The secondary structures and signature nucleotides of the two molecules are identical, but they differ by substitutions in variable regions.

TABLE 1. Signature nucleotides for SAR406 and OCS307

Position	Consensus nucleotide	SAR406-OCS307 nucleotide	Taxa with SAR406 substitution
1202	U	G	*Bacteroides* spp., marine group A, green nonsulfur bacteria, *Chlamydia* sp., cyanobacteria, α-proteobacteria, gram-positive bacteria with high G + C content
1207	G	U	Marine group A
1234	C	U	*Bacteroides* spp.
1410	A	G	*Bacteroides* sp., *Fibrobacter* spp., green sulfur bacteria, marine group A, green nonsulfur bacteria, *Planctomyces* spp., α-proteobacteria

but similar results were obtained with the maximum likelihood method, with minor exceptions described below. The gene clone NH16-12 was not included in this analysis because it would necessitate a truncated mask, reducing the number of phylogenetically informative positions available for comparison.

The phylogenetic analyses indicated an origin for the SAR406 gene lineage which lies within the main radiation of the bacterial phyla but also provided significant support for a monophyletic group encompassing *Fibrobacter* spp., the green sulfur bacteria (represented here by *Chlorobium limicola*), and SAR406. Bootstrap values for this clade were 85% for the neighbor-joining method, 73% for the parsimony method, and 65% for fastDNAml. In the absence of the SAR406 sequence, for the same data set, the values were 21, 10, and 31%, respectively (data not shown). Thus, the inclusion of the SAR406 gene in phylogenetic trees caused a rearrangement of relationships among some deeply branching bacterial lineages, which were associated with significant changes in bootstrap values.

An oligonucleotide probe, designed to specifically hybridize to the SAR406 gene, was evaluated with the RDP program CHECK_PROBE. The results of CHECK_PROBE indicated that there were a minimum of three mismatches with any known 16S rRNA sequence. The empirical T_m was determined to be between 45 and 50°C. This supported the selection of 45°C as the stringent wash temperature. The radiolabelled oligonucleotide probe was used to screen the 80-m Sargasso Sea library, but no further clones of this type were found.

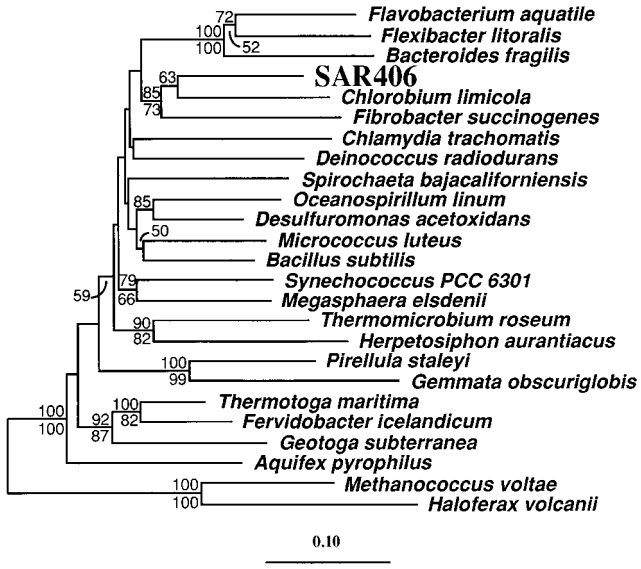

FIG. 2. Phylogenetic tree showing relationships of SAR406 and representative bacterial and archaeal 16S rRNA genes. This tree was inferred from nearly complete sequences by the neighbor-joining and parsimony methods. The number of bootstrap replicates which supported each branch, from a total of 100 replicates, are shown above (neighbor-joining) and below (parsimony) each node. Bootstrap values below 50% are not shown.

Profiles of hybridization of the SAR406 probe to amplified rDNA from samples from various depths from the Sargasso Sea and Pacific Ocean (Oregon coast) suggested that genes of the SAR406 lineage were proportionately most abundant at depths below 80 m (Fig. 3). An 80-m water sample had served as the source of the water used to make the clone library, where the gene was first encountered. Since the SAR406 gene lineage was found to be present in the Pacific Ocean samples, a second clone library was constructed from samples collected at a depth of 120 m off the Oregon coast. This library was screened with the SAR406 probe, clone OCS307 was identified, and a complete nucleotide sequence was determined. OCS307 was found to be 96% similar to SAR406 over the 1,492 positions the two sequences have in common. It was not included in the phylogenetic analyses because it was very similar to SAR406 and had no effect on the branching order of the trees.

Monthly time series data from BATS revealed a strong trend in the distribution of the SAR406 gene cluster with depth. Average values for SAR406 rDNA proportional abundance at 200 m were 9.4%, i.e., about five times higher than the average values at 0 m over the same interval ($P = <0.001$). The rRNA hybridization values for the SAR406 probe at the surface were uniformly low; fluctuations such as those seen with rDNA hybridization targets could not be resolved. However, comparisons of rRNA hybridization time series data indicated that the gene was proportionately four times more abundant at 200 m than at 0 m ($P = 0.04$), which was similar to the result obtained with rDNA hybridization targets. The time series data are consistent with the data obtained from the vertical profiles and support the view that the gene cluster is situated primarily in the lower regions of the surface layer throughout much of the year. As shown in Fig. 4, rDNA hybridization data indicated that the SAR406 gene cluster was highly correlated with chlorophyll *a* concentrations ($r = 0.75$) in surface samples.

DISCUSSION

The SAR406 gene lineage was found to be a deep phylogenetic branch of the bacteria. These genes could not be associated with high probability with any of the known bacterial phyla, although they most often formed a sister group with either the *Fibrobacter* or the *Chlorobium* lineages. The chlorobia are green sulfur bacteria, a monophyletic group of obligate anaerobic photoautotrophs, requiring exogenous reduced electron sources for photosynthesis and carbon dioxide fixation (31). The genus *Fibrobacter* is a dominant cellulolytic genus of the rumen of cattle and sheep (23). While both *Fibrobacter* and *Chlorobium* spp. are obligate anaerobes, *Fibrobacter* spp. use a fermentative metabolism, and the two groups appear to have few obvious features in common. This is consistent with the

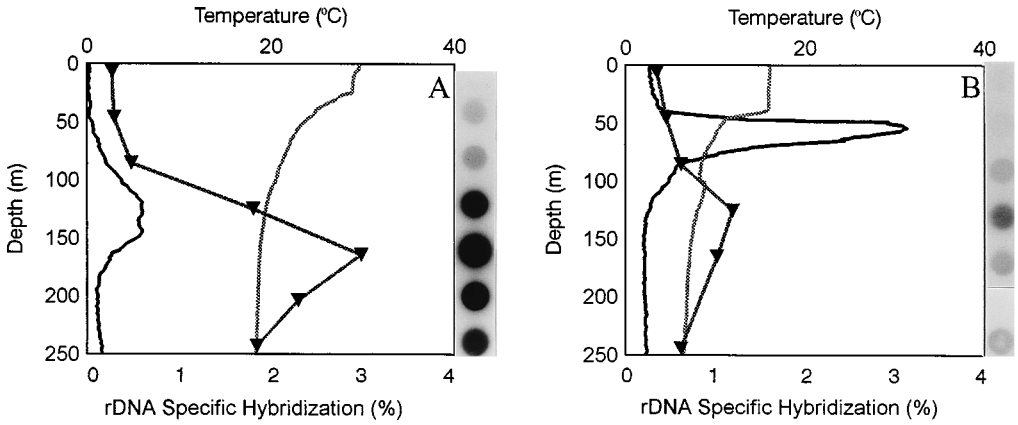

FIG. 3. The distribution of SAR406 rDNAs in the upper 250 m of the oceanic water column evaluated by dot blot hybridization. (A) Upper 250 m at BATS during a stratified period in August 1991; (B) water samples collected in October 1993 at a station located on the continental shelf (ca. 800-m water depth) off the Oregon coast. Hybridization of the SAR406 probe to total bacterial rDNA is shown to the right of temperature and fluorescence data collected with a CTD. There was no sample from 200 m for the dot blot of samples from the Oregon coast. Symbols: ‾, temperature; ‾, fluorescence (relative units), ▼, rDNA specific hybridization.

substantial evolutionary distance between the two groups. Thus, although the SAR406 lineage, *Fibrobacter* spp., and the green sulfur bacteria form a phylogenetic cluster, the distances among the three are too great to allow for inferences regarding the physiology of SAR406.

Chimeric genes created by in vitro recombination during PCR are a potential source of artifact in investigations of phylogenetic diversity (15, 27, 29). Secondary structural models can be used to detect some chimeric genes by revealing mispaired bases or signature features representing genes of different origins. CHECK_CHIMERA, provided by the RDP, is a useful aid in chimera identification, particularly when the recombined genes are substantially different. It has been suggested that neither the CHECK_CHIMERA program nor secondary structural models are sufficiently rigorous methods of detecting chimeric genes (25).

Independent phylogenetic comparisons of 5′ and 3′ domains are likely to be among the most sensitive qualitative tests for chimeras for the same reasons that phylogenetic analyses are more sensitive than pairwise sequence similarities as a means of detecting microbial affiliations: phylogenetic inference simultaneously weighs relationships in a set of many sequences. In this particular case, however, the relationship between SAR406 and OCS307 and their nearest (but still distant) relatives did not emerge until full-length sequences were obtained, and there was no specific association revealed by the separate analyses of the 5′ and 3′ regions of the molecules.

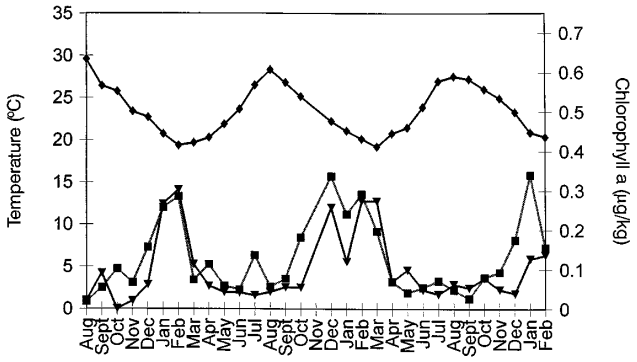

FIG. 4. Time series data comparing relative SAR406 hybridization bacterial rDNA amplified from surface water samples with measured chlorophyll *a* concentrations during the period from August 1991 to December 1993. Bacterial rDNAs were amplified from total bacterioplankton genomic DNA by PCR. During the time period shown, the average SAR406 rDNA hybridization value at 200 m was five times higher than that at the surface. The hybridization values at 200 m (data not shown) fluctuated with no discernible pattern. Symbols: ◆, temperature; ▼, chlorophyll *a*; ■, SAR406 DNA (relative units).

Perhaps the best evidence that a gene lineage is not chimeric is the confirmation of multiple genes of a common type, particularly from different sources, such as those we describe here. Figure 1 provides evidence against SAR406 and OCS307 being chimeras. SAR406 and OCS307 are 96% similar. The pattern of conserved secondary structure and base substitutions throughout the variable regions of the molecule is consistent with natural variation created by random evolutionary processes. The possibility that two genes with unique similarities in conserved regions could independently result from in vitro recombination events in complex gene mixtures is unlikely because (i) the conserved regions and structural features of the molecules are essentially identical, indicating that if these are chimeras, they must have been formed independently from members of the same clade, and (ii) only very complex chimeras involving multiple exchanges could explain the variability in Fig. 1.

16S rRNA sequence analyses done previously by Woese and coworkers, which predated the application of bootstrap analysis as an estimate of reliability, indicated a phylogenetic relationship between the green sulfur bacteria and the *Flavobacterium-Bacteroides* lineage (31). Although the phylogenetic analyses presented in Fig. 2 do not support this relationship with high bootstrap values, we note that some *Bacteroides* and *Flavobacterium* species share signature nucleotides with SAR406 and OCS307 (Table 1). A specific relationship between green sulfur bacteria and the *Flavobacterium-Bacteroides* phylum is not supported by significant bootstrap values in Fig. 2 because several lineages, including *Deinococcus radiodurans*, *Spirochaeta bajacaliforniensis*, and *Chlamydia trachomatis*, frequently swapped positions with the *Flavobacterium-Bacteroides* group in the analysis.

The influence of a single sequence on the ordering of phylogenetic topologies and bootstrap confidence estimates, such as that described here for SAR406, is not an unusual feature of molecular phylogeny. It is particularly likely to occur in cases such as this, which involve relatively early events in microbial evolution which involve many short internode distances in phy-

logenetic trees, e.g., a radiation of lineages. The results presented here might best be viewed as a further refinement of the phylogenetic picture, which may well be tested again by the future discovery of further novel gene lineages.

Our results with rRNAs and amplified rDNAs from natural communities indicate that the SAR406 gene cluster is vertically stratified in the water column and undergoes regular seasonal oscillations in relative abundance at the surface. We present the hybridization data as relative rRNA and rDNA abundances, which are measures of microbial community activity and composition. The cellular abundance of bacterioplankton, as measured by direct counting methods, varies with depth and sample and is, on average, ca. 20% less at 200 m than at 0 m at our study site. Thus, relative gene abundance, expressed as a percentage, and cellular abundance are different measures which should not be confused. The former may be more useful for evaluating the ecology of species because it provides a reasonable measure of the contribution of that species to the community, whereas the latter is of greater interest where biomass is the concern. As with other uncultured species, the absolute abundance of SAR406 rRNA cannot be accurately estimated from the hybridization of oligonucleotides because no pure SAR406 RNA is available for the standardization of hybridization results. The case is different with rDNA, where cloned genes are available as standards.

An assay which relied on PCR produced an estimate of 9.4% for the average abundance of the SAR406-OCS307 gene group at 200 m in 30 consecutive time series samples. The SAR406 probe corresponds to a relatively variable region of 16S rRNAs; hence, it may underestimate abundance if the SAR406-OCS307 group forms a diverse gene cluster including members with sequence variations in the probe hybridization target region. A particular concern of molecular ecologists is that studies of 16S rDNA genes amplified from environmental samples with the PCR may be subject to biases in relative gene abundance introduced by PCR (9). Recent studies suggest that there are conditions under which both biases and variability are small enough to permit comparisons between samples (28). As pointed out by Farrelly and coworkers (9), the relative abundance of a gene in a population will be influenced by factors such as the number of rRNA genes in genomes and the size of genomes among population members. However, all measures of microbial abundance have particular properties of meaning and empirical accuracy which impose limitations on interpretations (e.g., biomass versus cellular activity). In this light, we suggest that measures of relative gene abundance may prove useful for ecological studies.

Previous studies have shown that spatial and temporal variability occurs in bacterioplankton communities (18, 20, 21). However, with the exception of photosynthetic bacterioplankton, the extent of bacterial community stratification in the open ocean is not well understood. Our results suggest that the bacteria harboring the SAR406 gene may be adapted to grow best in a discrete region of the water column. The evidence of a stratified distribution for the SAR406 gene cluster does not give a clear indication of its ecological role. The hybridization data in Fig. 4 indicate a strong positive correlation between the relative abundance of SAR406 at the sea surface and surface chlorophyll *a* values. Much of this correlation appears to result from increases in SAR406 surface abundance which occur in concert with spring phytoplankton blooms. These phytoplankton blooms are brought about by the strong seasonal variations of mixing and stratification which are characteristic of the western Sargasso Sea (5). The reasons for the temporal correlation between chlorophyll *a* and SAR406 at the surface are unknown. One hypothesis is that SAR406 is a phototroph. If so, its seasonal and depth-specific patterns of abundance could simply be a reflection of its physiological properties as a photosynthetic organism. More likely, SAR406 may be a heterotrophic organism that depends on phytoplankton, possibly using the exudates of photosynthetic organisms as carbon sources.

16S rRNA has been widely used for ecological studies (16, 17, 29), sparking occasional debate among researchers. Some of this discussion has focused on the limitations of rRNA, noting that this molecule rarely gives clues as to the physiology and biogeochemical roles of microorganisms. This limitation is due primarily to the observation that physiology and phylogeny are not congruent properties among most of the bacterial phyla. For example, proteobacteria are a cohesive phylogenetic group but contain organisms that vary greatly in their physiological capabilities. As we demonstrate, SAR406, *Fibrobacter* spp., and the green sulfur bacteria possess some degree of phylogenetic affiliation. However, this affiliation alone is not sufficient to allow predictions regarding the physiology of this uncharacterized organism.

In summary, the SAR406 gene lineage represents a phylogenetically ancient branch of the bacteria which appears to be most closely associated with *Fibrobacter* spp. and the green sulfur bacteria. Moreover, this unusual group exhibits a stratified distribution in the western Sargasso Sea, where its maximum abundance occurs below the deep chlorophyll layer throughout most of the year. A striking correlation between the surface abundance of this gene and surface indicators of phytoplankton biomass suggests a possible link between blooms of this organism and biological or physical factors. Although little is known about the impact of microbial stratification on the cycling of nutrients in the open oceans, recent data suggest that stratification might be an important factor in the cycling of carbon in the surface layer (5). The study of one widespread component of these stratified communities, the species corresponding to the SAR406 gene, may provide insights into the role of bacterioplankton community structure in marine food webs.

ACKNOWLEDGMENTS

This work was supported by National Science Foundation grant OCE 9016373 for the study of microbial diversity at BATS and Department of Energy grant FG 0693ER61697 from the ocean margins program for the study of bacterioplankton population dynamics over continental shelves.

We are grateful to BATS group for sampling and to Kevin Vergin and Nanci Adair for processing the nucleic acid samples. We would also like to thank Kate Field, Mike Rappé, Marcelino Suzuki, and Ena Urbach for useful comments.

REFERENCES

1. **Amann, R., L. Ludwig, and K. Schleifer.** 1995. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. Microbiol. Rev. **59:**143–169.
2. **Azam, F., T. Fenchel, J. G. Field, J. S. Gray, L. A. Meyer-Reil, and F. Thingstad.** 1983. The ecological role of water-column microbes in the sea. Mar. Ecol. Prog. Ser. **10:**257–263.
3. **Barns, S. M., R. E. Fundyga, M. W. Jeffries, and N. R. Pace.** 1994. Remarkable archaeal diversity detected in a Yellowstone National Park hot spring environment. Proc. Natl. Acad. Sci. USA **91:**1609–1613.
4. **Britschgi, T. B., and S. J. Giovannoni.** 1991. Phylogenetic analysis of a natural marine bacterioplankton population by rRNA gene cloning and sequencing. Appl. Environ. Microbiol. **57:**1313–1318.
5. **Carlson, C. A., H. W. Ducklow, and A. F. Michaels.** 1994. Annual flux of dissolved organic carbon from the euphotic zone in the northwestern Sargasso Sea. Nature (London) **371:**405–408.
6. **Dams, E., L. Hendricks, Y. Van de Peer, J. Neefs, G. Smits, I. Vandenbempt, and R. De Wachter.** 1988. Compilation of small ribosomal subunit RNA sequences. Nucleic Acids Res. **16:**87–183.

7. **DeLong, E. F.** 1992. Archaea in coastal marine bacterioplankton. Proc. Natl. Acad. Sci. USA **89:**5685–5689.
8. **Ducklow, H. W.** 1983. Production and fate of bacteria in the oceans. BioScience **33:**494–501.
9. **Farrelly, V., F. A. Rainey, and E. Stackebrandt.** 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.
10. **Felsenstein, J.** 1981. Evolutionary trees from DNA sequences: a maximum likelihood approach. J. Mol. Evol. **17:**368–376.
11. **Felsenstein, J.** Phylip 3.5. University of Washington, Seattle.
12. **Fuhrman, J., K. McCallum, and A. Davis.** 1993. Phylogenetic diversity of subsurface marine microbial communities from the Atlantic and Pacific oceans. Appl. Environ. Microbiol. **59:**1294–1302.
13. **Fuhrman, J., K. McCallum, and A. Davis.** 1992. Novel major archaebacterial group from marine plankton. Nature (London) **356:**148–149.
14. **Giovannoni, S. J.** 1991. The polymerase chain reaction, p. 177–201. *In* E. Stackebrandt and M. Goodfellow (ed.), Sequencing and hybridization techniques in bacterial systematics. John Wiley & Sons, Inc., New York.
15. **Giovannoni, S. J., T. B. Britschgi, C. L. Moyer, and K. G. Field.** 1990. Genetic diversity in Sargasso Sea bacterioplankton. Nature (London) **345:** 60–63.
16. **Giovannoni, S. J., and C. Cary.** 1993. Probing marine systems with ribosomal RNAs. Oceanography **6:**95–104.
17. **Giovannoni, S. J., T. D. Mullins, and K. G. Field.** 1995. Microbial diversity in oceanic systems: rRNA approaches to the study of unculturable microbes, p. 217–248. *In* I. Joint (ed.), Molecular ecology of aquatic microbes. Springer-Verlag, Berlin.
18. **Giovannoni, S. J., M. Rappé, and N. Adair.** 16S rRNA genes reveal stratified open ocean bacterioplankton populations related to the *Chloroflexus*/*Herpetosiphon* phylum. Submitted for publication.
19. **Karl, D. M., R. Leteller, D. Hebel, L. Tupas, J. Dore, J. Christian, and C. Win.** 1995. Ecosystem changes in the north pacific subtropical gyre attributed to the 1991–1992 El Niño. Nature (London) **373:**230–234.
20. **Lee, S., and J. A. Fuhrman.** 1990. DNA hybridization to compare species compositions of natural bacterioplankton assemblages. Appl. Environ. Microbiol. **56:**739–746.
21. **Lee, S., and J. A. Fuhrman.** 1991. Spatial and temporal variation of natural bacterioplankton assemblages studied by total genomic DNA cross-hybridization. Limnol. Oceanogr. **36:**1277–1287.
22. **Maidak, B. L., N. Larsen, M. McCaughey, R. Overbeek, G. Olsen, K. Fogel, J. Blandy, and C. Woese.** 1994. The ribosomal database project. Nucleic Acids Res. **22:**3485–3487.
23. **Montgomery, L., B. Flesher, and D. Stahl.** 1988. Transfer of *Bacteroides succinogenes* (Hungate) to *Fibrobacter* gen. nov. as *Fibrobacter succinogenes* comb. nov. and description of *Fibrobacter intestinalis* sp. nov. Int. J. Syst. Bacteriol. **38:**430–435.
24. **Olsen, G. J., H. Matsuda, R. Hagstrom, and R. Overbeek.** 1994. fastDNAml: a tool for construction of phylogenetic trees of DNA sequences using maximum likelihood. Comput. Appl. Biosci. **10:**41–48.
25. **Robinson-Cox, J. F., M. M. Bateson, and D. M. Ward.** 1995. Evaluation of nearest-neighbor methods for detection of chimeric small-subunit rRNA sequences. Appl. Environ. Microbiol. **6:**1240–1245.
26. **Sgaramella, V., and H. G. Khorana.** 1972. Total synthesis of the structural gene for an alanine transfer RNA from yeast: enzymatic joining of the chemically synthesized polynucleotides to form the DNA duplex representing nucleotide sequence 1-20. J. Mol. Biol. **72:**427–444.
27. **Shuldiner, A. R., A. Nirula, and J. Roth.** 1989. Hybrid DNA artifact from PCR of closely related target sequences. Nucleic Acids Res. **17:**4409.
28. **Suzuki, M. T., and S. J. Giovannoni.** 1996. Bias caused by template annealing in the amplification of mixtures of 16S rRNA genes by PCR. Appl. Environ. Microbiol. **62:**625–630.
29. **Ward, D. M., M. M. Bateson, R. Weller, and A. L. Ruff-Roberts.** 1992. Ribosomal RNA analysis of microorganisms as they occur in nature. Adv. Microb. Ecol. **12:**219–286.
30. **Woese, C. R.** 1987. Bacterial evolution. Microbiol. Rev. **51:**221–271.
31. **Woese, C. R., L. Mandelco, D. Yang, R. Gherna, and M. Madigan.** 1990. The case for the relationship of the flavobacteria and their relatives to the green sulfur bacteria. Syst. Appl. Microbiol. **13:**258–262.