

## A Novel $\delta$ -Subdivision Proteobacterial Lineage from the Lower Ocean Surface Layer

TERAH D. WRIGHT,<sup>1</sup> KEVIN L. VERGIN,<sup>1</sup> PHILIP W. BOYD,<sup>2</sup> AND STEPHEN J. GIOVANNONI<sup>1\*</sup>

Department of Microbiology, Oregon State University, Corvallis, Oregon 97331,<sup>1</sup> and Department of Oceanography, University of British Columbia, Vancouver, British Columbia, Canada V6T 1Z4<sup>2</sup>

Received 26 September 1996/Accepted 5 January 1997

**A small-subunit ribosomal RNA (16S rRNA) gene lineage (SAR324) affiliated with the  $\delta$  subdivision of the class *Proteobacteria* (DP) was discovered in a 16S rRNA gene clone library prepared from a water sample collected from 250 m in the western Sargasso Sea. This clone library of nearly full-length amplicons of bacterial 16S rRNA genes has been the subject of previous studies aimed at identifying bacteria that inhabit the lower ocean surface layer. The novel lineage was identified by randomly sequencing clones that did not hybridize to oligonucleotide probes specific for several abundant bacterioplankton groups identified in previous studies. Phylogenetic analysis indicated that SAR324 was most closely affiliated with the DP, although it showed no specific relationship to any DP 16S rRNA genes in databases. Eight of the clones in the library of 148 clones were identified as members of the SAR324 lineage by hybridization to an oligonucleotide probe specific for SAR324. Subsequent hybridizations showed that the SAR324 group is stratified in the lower surface layer of both the Atlantic and Pacific Oceans, with maxima between 160 and 500 m. The repeated discovery of sequences belonging to different gene clusters with similar distributions in this region of the water column suggests that microbial communities in the lower surface layer may be functionally specialized.**

The sequencing of 16S rRNA genes and the application of group-specific oligonucleotide probes has provided much new information on bacterioplankton diversity and has revealed a previously unseen structure in the spatial and temporal distributions of microorganisms in marine systems (17, 18, 34). An interesting and unforeseen conclusion of these studies is that a majority of the genes recovered from seawater belong to a limited number of phylogenetic groups. In a recent review, we reported that 86% of all bacterial genes ( $n = 440$ ) recovered from seawater fall within eight phylogenetic groups (16). Elsewhere it has been reported that all archaeal rRNA genes recovered so far from seawater belong to two phylogenetic groups (6, 11). These conclusions are important to microbial ecologists seeking to understand the relationship between microbial diversity and functional specialization within microbial communities because they suggest that a relatively limited array of molecular probes may be sufficient for monitoring the population dynamics of the majority of bacterioplankton. Equally important has been the observation of similar stratified patterns in the distributions of some of the major bacterial groups in different oceans (17, 18, 26). Collectively, these observations support the view that common principles may underlie the organization of bacterioplankton communities in temperate oceans.

Although common themes are emerging from investigations of bacterioplankton diversity, novel genes of potential ecological significance continue to be discovered (17, 18, 26). Bacterioplankton 16S ribosomal DNAs (rDNAs) have been recovered from surface and 100-, 200-, and 500-m samples from subtropical and temperate regions of the oceans, as well as Antarctica, and the eastern and western continental shelves of the United States (2, 6, 7, 11, 12, 14, 28). The majority of clones belong to the cyanobacterium (SAR6 and SAR7) and proteobacterium (SAR11, SAR86, SAR83, and SAR116 clus-

ters) divisions. Many genes from novel lineages related to the genus *Fibrobacter* (marine group A and SAR406), the green non-sulfur bacteria (SAR202), and the gram-positive bacteria (NH16-9 and BDA1-5), as well as numerous, unique clones that are rarely encountered have also been found (12, 14, 17, 18).

The vertical stratification of photosynthetic bacterioplankton populations is well-known, but only recently have investigations with oligonucleotide probes shown that many of the most abundant bacterioplankton lineages—species of unknown physiology—are also highly stratified (4, 32). Lee and Fuhrman (23) showed that community DNAs from the Pacific Ocean varied significantly among samples from 25, 100, 500, and 1,000 m. Early hybridization studies with phylogenetic group-specific oligonucleotide probes indicated that bacterial genes cloned from surface samples were often dominant only in the upper surface layer (2, 14, 30). More recently, we have found that four uncultured microbial groups (SAR202, SAR406, SAR11G1 subcluster, and marine archaea group I) form stratified populations at Atlantic and Pacific Ocean sites (references 17 and 18, and unpublished data). These results suggest that stratification may be an important property of community structure in marine systems and that unique communities might occur in the lower surface layer.

The work presented here is part of an ongoing effort to characterize the bacterioplankton of the lower surface layer at the Bermuda Atlantic Time-Series study site (BATS) in the western Sargasso Sea. Previous analyses of this clone library from 250 m resulted in the discovery of a novel gene lineage, SAR202, and a deep-water phylogenetic subgroup of the SAR11 cluster (14, 17). Here we describe another novel gene clone lineage (SAR324), which is most closely affiliated with the  $\delta$  subdivision of the class *Proteobacteria* (DP), and show that it also forms stratified populations in the lower surface layers of both the Atlantic and Pacific oceans.

### MATERIALS AND METHODS

**Sampling and nucleic acid extraction.** Water samples were collected from BATS (31°50'N, 64°10'W) and from ocean station PAPA in the subarctic north

\* Corresponding author. Phone: (541)737-1835. Fax: (541)737-0496. E-mail: giovanni@bcc.orst.edu.

Pacific Ocean (approximately 50°N, 145°W) with Niskin bottles attached to a conductivity, temperature, and depth (CTD) rosette. The samples from PAPA have not been studied previously; however, the BATS samples used in this study have been described elsewhere in previous studies of other bacterioplankton groups (17). Monthly time series samples (30) were collected from BATS at two depths (0 and 200 m) from August 1991 to February 1994. In addition to monthly samples, samples were collected at BATS from 40, 80, 120, 160, and 250 m 10 times during the same period. The samples from ocean station PAPA were collected in September 1995 from depths ranging from 0 to 3,300 m. From 24 to 48 liters of seawater was filtered from each depth. A Sea-Bird CTD rosette was used to measure continuous profiles of temperature.

Total cellular nucleic acids were extracted from the filters by procedures optimized for small sample sizes, as described elsewhere (17).

**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 (27F, AGA GTT GAT CMT GGC TCA G; 1522R, AAG GAG GTG ATC CAN CCR CA) as described previously (13, 17). The clone library was constructed by using the plasmid vector pCRII (Invitrogen, San Diego, Calif.) as described in the manufacturer's instructions. Transformants were screened for full-length insertions by *EcoRI* restriction digestion. Clones were numbered discontinuously from 177 to 325 and stored in LB (10 g of tryptone per liter, 5 g of NaCl per liter, 5 g of yeast extract per liter, 50 µg of kanamycin per ml)–7.0% dimethyl sulfoxide at –80°C.

**Gene sequencing and phylogenetic analysis.** Plasmid DNAs were purified for sequencing from clones grown overnight at 37°C in Luria-Bertani broth by using a Prep-A-Gene DNA Purification Kit (Bio-Rad Laboratories, Hercules, Calif.) or a QIAprep Spin Plasmid Miniprep Kit (Qiagen, Inc., Chatsworth, Calif.) according to the manufacturer's instructions. Plasmid DNAs were sequenced bidirectionally with universal and bacterial primers by using an Applied Biosystems 373A automated sequencer as described previously (2, 13, 15, 22). DNA sequence data was manually aligned to bacterial sequences obtained from the Ribosomal Database Project (RDP) by using the program GDE, supplied by Steve Smith (Millipore Corporation, Bedford, Mass.) (25). Sequences were evaluated by the program CHECK\_CHIMERA, also provided by the RDP, to aid in the identification of chimeric gene artifacts. Phylogenetic relationships were inferred by the neighbor-joining method and by parsimony by using the Phylogeny Inference Package (PHYMLIP), version 3.4 (9, 29). Regions of ambiguous alignment and hypervariability were excluded from the analysis. Secondary structure analysis of the 16S rRNA gene was performed with the program gRNAid, supplied by Shannon Whitmore (Mentor Graphics, Wilsonville, Oreg.).

**Hybridization.** Vertical profiles of SAR324 rRNA and rDNA amplicons were measured by hybridizations to dot blots as described previously (17, 18). For the rDNA replicates used to generate error bars, bacterial rDNAs were amplified in three separate reactions from seawater by using bacterial 16S rDNA primers (27F; 1492R, GGT TAC CTT GTT ACG ACT T) (13). For hybridizations to environmental high-molecular-weight RNA, 100, 50, 20, and 10 ng of each RNA sample were blotted, and the slopes of the lines were determined by linear regressions. Nucleic acids were adsorbed onto Zetaprobe membranes (Bio-Rad Laboratories, Inc., Carson City, Calif.), cross-linked by UV radiation and baking, and stored desiccated at –20°C before probing.

An oligonucleotide probe specific for the SAR324 lineage (SAR324R; CGA AAG ACC CTC CGG) was designed to complement positions 625 to 639 of the 16S rRNA gene (*Escherichia coli* numbering system). The probe was prescreened for potential cross-reactivity with the program CHECK\_PROBE, provided by the RDP (25). T4 polynucleotide kinase was used to label the 5' terminus of the oligonucleotide probe with [ $\gamma$ -<sup>32</sup>P]ATP as described previously (31). The empirical melting temperature ( $T_m$ ) of the probe was determined by quantifying the amount of probe hybridized to dot blots of SAR324 rDNA after 15-min washes at temperatures from 30 to 55°C.

The rDNA and RNA blots were hybridized in Z-Hyb buffer (1 mM EDTA, 0.25 M Na<sub>2</sub>HPO<sub>4</sub>, 7% sodium dodecyl sulfate, pH 7.2) containing ca. 50 ng of radiolabeled oligonucleotide probe as described previously (17, 18). Following hybridization, the blots were exposed to PhosphorImager plates (Molecular Dynamics, Sunnyvale, Calif.), followed by quantification with a Molecular Dynamics PhosphorImager SI and IMAGEQUANT software. Data were analyzed as described previously, with the hybridization of the bacterial probe 338R used as a denominator so that variation in the hybridization of the specific probe (SAR324R) is expressed in relative units that are proportional to bacterial RNA; changes in the plotted values represent variation in the proportion of bacterial RNA contributed by the SAR324 group. SAR324 rDNA hybridization values are expressed as percentages, since SAR324 genes (amplicons) were available for use as standards for normalization, as described previously (17, 18). Amplicons were not used to normalize the rRNA hybridization data, since this would have assumed that free energy of binding for the probe to RNA targets was the same as for DNA targets, an assumption that is unlikely to be true.

**Nucleotide sequence accession numbers.** Nucleotide sequences were filed in GenBank under the following accession numbers: SAR324, U65908; SAR257, U65909; SAR237, U65910; SAR214, U65911; SAR248, U65912; SAR308, U65913; SAR218, U65914; and SAR276, U65915.

## RESULTS

SAR324 and seven related genes were identified in a library of 148 bacterial 16S rRNA gene clones from a Sargasso Sea sample from a 250-m depth. Because so many bacterioplankton genes had already been identified and probes were available, the library was screened by hybridization with radiolabeled oligonucleotide probes that are specific for the bacterioplankton lineages (SAR11, SAR83, SAR406, and SAR202) that had previously been shown to be numerically significant in this and other rRNA clone libraries (2, 17, 18). Clones that did not hybridize to these probes were selected at random for phylogenetic analyses, and the 5' and/or 3' regions of the 16S rRNA genes were sequenced. Of the 37 clones that were randomly sequenced, three (SAR248, SAR276, and SAR324) appeared to be loosely affiliated with the DP in preliminary phylogenetic analyses. Subsequently, complete bidirectional sequences were determined for these clones. The SAR248, SAR276, and SAR324 gene sequences were evaluated with the RDP programs SIM\_RANK and CHECK\_CHIMERA (25). CHECK\_CHIMERA provided results which supported the conclusion that the genes were not chimeric artifacts, but no inferences regarding their phylogeny could be drawn from the low  $S_{AB}$  values (0.4 to 0.5) that were obtained by the SIM\_RANK analysis. SIM\_RANK results are expressed as  $S_{AB}$  values, the number of shared oligomers of seven bases, divided by the number of unique oligomers in either the submitted sequence or the database sequence, whichever is lower.

Phylogenetic analyses indicated that the novel genes formed a monophyletic group that included no cultured representatives, thus conforming to the definition of an environmental gene cluster (14, 26). In separate phylogenetic comparisons of 5' and 3' domains, the genes behaved similarly; neither domain alone showed a significant affiliation with any phylogenetic group other than the DP (data not shown). A phylogenetic tree inferred by the neighbor-joining method from full-length sequences of SAR324, SAR248, and SAR276 and other 16S rRNA sequences representing the proteobacteria is shown in Fig. 1 (29). Bootstrap resampling (100 replicates) of the data was used to provide statistical support for the phylogenetic position of the SAR324 lineage (8, 10). SAR324 and related gene clones always formed a monophyletic cluster within the DP; however, bootstrap values supporting the DP as a monophyletic clade were low (64%). The branching orders within the clade were not well supported by bootstrap replicates; hence, the deepest branches are shown here as a polytomy. Bootstrap support for the DP group was improved considerably (from 64 to 88%) by omitting *Desulfovibrio desulfuricans* and *Bdellovibrio bacteriovorus* from the analysis (data not shown). The inclusion of the SAR324 lineage in this analysis caused no rearrangement of relationships or significant changes in bootstrap values of previously sequenced genes.

Primary sequence similarities and signature sequence analyses confirmed the loose association of the SAR324 gene lineage with the DP. Sequence similarities among and within the four subdivisions of the proteobacteria and SAR324, SAR248, and SAR276 were calculated from ca. 1,020 nucleotide positions (Table 1). Only one member each of the  $\alpha$ ,  $\beta$ , and  $\gamma$  subdivisions is shown; however, similarity matrices involving larger data sets gave similar results (data not shown). Although the similarities between SAR324 and members of the DP were low (0.856 to 0.907), they are not unusual given the range of similarity values among characterized DP (0.867 to 0.987). Among the DP, *D. desulfuricans* had the lowest similarity value (0.854) when compared to SAR324, consistent with the observation that removal of this sequence from the phylogenetic

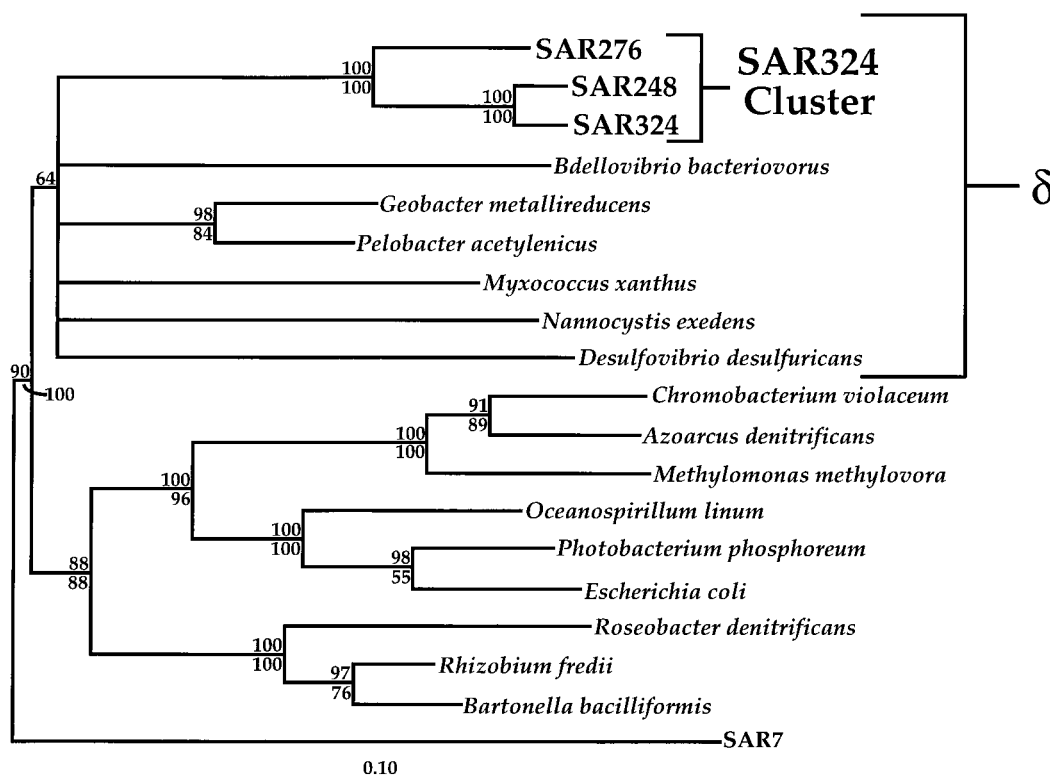


FIG. 1. Phylogenetic tree showing relationships of the SAR324 cluster and representative bacterial 16S rRNA genes. This tree was inferred by the neighbor-joining method and included ca. 1020 nucleotide positions in the analysis. The number of bootstrap replicates out of 100 that supported each branch is shown above (neighbor joining) and below (parsimony) the nodes. Values less than 50% are not shown. The DP species are shown as a polytomy because the branching order was not well resolved.

analysis significantly increased the bootstrap values supporting the DP as a monophyletic clade.

The SAR324 sequences were compared to 16S rRNA signature sequences for the bacterial phyla and their subdivisions, which were previously published by Woese (35) and Haddad et al. (20). The highest observed percentage of shared signature positions for SAR324 was to the DP (87 to 90%) (Table 2). Among 51 characterized DP sequences that were similarly analyzed, the correspondence of nucleotide identities at signa-

ture sequence positions for the DP was 86 to 100% (data not shown).

The proposed secondary structure for members of the SAR324 cluster is structurally unique and conserved within the group (Fig. 2). Seven of the nine signature sequence mismatches between SAR324 and the DP are compensatory base changes (changes in variable nucleotides that preserve the secondary structure of the 16S rRNA molecule) across regions of double-stranded pairing. The inset illustrates variable region

TABLE 1. Sequence similarities (based on ca. 1,020 nucleotide positions) among the SAR324 cluster and 16S rRNA gene sequences from representatives within the proteobacteria

Bacterium <sup>a</sup>	Similarity with:											
	SAR276	SAR278	SAR324	<i>B. bacteriovorus</i>	<i>G. metallireducens</i>	<i>P. acetylenicus</i>	<i>M. xanthus</i>	<i>N. exedens</i>	<i>D. desulfuricans</i>	<i>C. violaceum</i>	<i>O. linum</i>	<i>R. denitrificans</i>
SAR276 ( $\delta$ )												
SAR248 ( $\delta$ )	0.953											
SAR324 ( $\delta$ )	0.947	0.985										
<i>Bdellovibrio bacteriovorus</i> ( $\delta$ )	0.873	0.873	0.873									
<i>Geobacter metallireducens</i> ( $\delta$ )	0.906	0.888	0.890	0.896								
<i>Pelobacter acetylenicus</i> ( $\delta$ )	0.907	0.889	0.892	0.898	0.957							
<i>Myxococcus xanthus</i> ( $\delta$ )	0.882	0.878	0.881	0.871	0.921	0.925						
<i>Nannocystis exedens</i> ( $\delta$ )	0.879	0.867	0.868	0.873	0.895	0.895	0.987					
<i>Desulfovibrio desulfuricans</i> ( $\delta$ )	0.856	0.851	0.854	0.867	0.895	0.898	0.876	0.881				
<i>Chromobacterium violaceum</i> ( $\beta$ )	0.872	0.873	0.871	0.848	0.870	0.873	0.858	0.841	0.849			
<i>Oceanospirillum linum</i> ( $\gamma$ )	0.863	0.865	0.864	0.868	0.890	0.883	0.883	0.872	0.869	0.893		
<i>Roseobacter denitrificans</i> ( $\alpha$ )	0.855	0.857	0.857	0.860	0.890	0.885	0.868	0.845	0.864	0.852	0.882	

<sup>a</sup> The symbols in parentheses indicate the subdivisions of proteobacteria.

TABLE 2. Signature nucleotides relating SAR324 and SAR276 to the DP

Position <sup>a</sup>	DP <sup>b</sup>	SAR324/276	+/- <sup>c</sup>	Position	DP	SAR324/276	+/-
44	G <sup>d</sup>	A	-	875	U:c	U	+
50	A	A/U	+/-	877	Y:a	C	+
107	G	G	+	878	Y:a	U/C	+
108	C	C	+	906	A:g	A	+
124	G	G	+	916	G	G	+
129	U	C/U	+/-	929	G	G	+
129:1	A	A	+	947	G:u	A	-
199	R	C/A	+/-	948	Y	C	+
233	Y	U	+	976	G	G	+
236	G	G	+	1015	A:g	G	+
237	C	C	+	1024	G:c	U	-
242	C:g <sup>e</sup>	C	+	1026	G	C	-
284	G:c	G	+	1116	Y	U	+
370	C	C	+	1120	Y:G	C	+
371	G:a	A/G	+	1219	A	A	+
390	C:u	U/C	+	1233	R	G	+
391	G	G	+	1234	C:a	U	-
398	C	U	-	1246	G:u	U	+
438	G:u	U	+	1252	A	A	+
449	A	A	+	1260	G:Y	U/G	+
485	G	G	+	1291	C:g	G	+
496	G:a	A	+	1297	Y	U	+
502	A:g	A	+	1298	C:a	C	+
513	C	A	-	1325	C	C	+
543	U:c	U	+	1421	Y	U	+
554	U:a	U	+	1426	U:R	A/G	+
564	U:c	C	+	1431	Y:a	A	+
640	A	A/G	+	1437	C	C	+
689	R:u	A	+	1441	G:u	G	+
690	G:a	G	+	1443	G	G	+
698	Y:a	U	+	1460	A	A	+
722	G:a	G	+	1464	G	G	+
760	G	G	+	1465	C:u	C/U	+
812	G	G	+	1467	Y	C	+
822	R:u	A/G	+	1469	A:u:c	A	+
823	R:u	G	+	1481	U:c	C/U	+
825	A:g	A	+	1520	G:c	C	+
871	U	U	+				

<sup>a</sup> *E. coli* numbering system.<sup>b</sup> DP signature nucleotides.<sup>c</sup> Match (+) or mismatch (-) between the SAR324 lineage and the DP.<sup>d</sup> Composition uppercase, major base; accounts for >90% of cases assayed.<sup>e</sup> Composition lowercase, minor base; found in <15% of cases assayed.

two in SAR276, which shows a deletion, relative to SAR324, that is the only topological secondary structure variation found so far within the SAR324 lineage (19). Variable region two contains two insertions or deletions (indels) relative to *E. coli*, the first from positions 183 to 194 and the second from positions 203 to 218. The first indel contains 25 and 11 nucleotides for SAR324 and SAR276, respectively. A comparison of 51 other characterized DP sequences revealed variations of 14 to 23 nucleotides in this same region. For the second indel, two of the 51 DP sequences analyzed had the same deletion (*B. bacteriovorus* and *Nannocystis exedens*). The structural variation of the 16S rRNA gene within the SAR324 sequence cluster is consistent with variation seen in other clades previously encountered in this 250-m clone library (reference 17 and unpublished data). We call attention to this variation because it

underscores the substantial variability within this group, which suggests the presence of multiple bacterial species.

The discovery of multiple genes of a common type, like those presented here, is evidence that members of the SAR324 gene lineage form a novel, diverse cluster and are not chimeric artifacts. For example, SAR248 and SAR324 differ by only 22 nucleotides, which are distributed throughout the gene. The nucleotide differences are confined mostly to hypervariable loops and compensatory base changes in stem regions; hence, they introduce no incongruities to secondary structural models. The possibility that two genes with unique similarities in conserved and variable regions result from in vitro recombination in a complex gene mixture is unlikely.

An oligonucleotide probe (SAR324R), designed to specifically hybridize to the SAR324 lineage (Fig. 2), was evaluated

FIG. 2. Proposed secondary structure for the SAR324 16S rRNA gene. The DP signature sequences shared by SAR324 are marked with asterisks. The lowercase letters represent DP signature nucleotides, which are different than the SAR324 nucleotide at that position. The target site for the SAR324 probe is shown between *E. coli* positions 625 and 639. The dashed box indicates variable region two of the molecule, the region of variation between SAR324 and SAR276. The inset shows variable region two of SAR276.



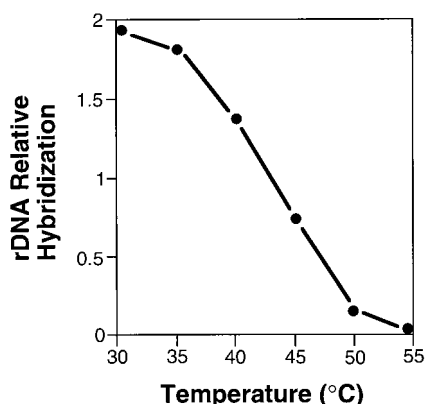


FIG. 3. The thermal stability of the SAR324 probe. The empirical  $T_m$  of the probe was determined by quantifying the amount of probe hybridized to dot blots of SAR324 rDNA after 15-min washes at a range of increasing temperatures (30 to 55°C).

with the RDP program CHECK\_PROBE (25). There were a minimum of three mismatches with any known 16S rRNA gene sequence and a minimum of five mismatches with all other sequenced clones from the 250-m clone library. The specificity of the probe was also shown empirically by using blotted arrays of 16S rDNA genes from several cultured and uncultured bacterioplankton under stringent hybridization conditions (data not shown). The empirical  $T_m$  for this oligonucleotide probe was determined to be between 40 and 45°C, which supported the selection of 40°C as the stringent wash temperature (Fig. 3). No cross-hybridization to unrelated genes was encountered. The radiolabeled oligonucleotide was used to screen the 250-m Sargasso Sea library, and six additional SAR324-related clones were detected by strong hybridization signals.

The sequencing of the additional clones provided further evidence for the highly diverse nature of the SAR324 lineage and further supported the specificity of the probe as a marker for a monophyletic microbial group. One of the clones was a chimera (SAR206), which was detected after partial sequencing of the 3' and 5' ends of the gene (data not shown). The remaining five clones were phylogenetically related to SAR324, although the genes encompassed substantial variation in the form of nucleotide substitutions similar to those described above (Fig. 4). A similarity matrix based on ca. 300 nucleotide positions revealed that five (SAR324, SAR257, SAR237, SAR214, and SAR248) of the eight clones in the SAR324 lineage were 97 to 99% similar to each other in an analysis that excluded hypervariable regions. Seven of the eight clones had secondary structures comparable to SAR324 in variable region two (Fig. 2).

The results of the hybridization analyses indicate that SAR324 is vertically stratified in the water columns of both the Atlantic and Pacific oceans (Fig. 5). In hybridizations to 30 consecutive time series samples from two depths (0 and 200 m) in the Sargasso Sea, SAR324 was always found to be more abundant at 200 m than at 0 m (data not shown). A one-tailed  $t$  test assuming unequal variances indicated that the SAR324 lineage was proportionately three times more abundant at 200 m than at 0 m ( $P = 1.0 \times 10^{-4}$ ). The time series data are consistent with the data obtained from the rDNA and rRNA vertical profiles and support the hypothesis that the SAR324 lineage is located in the lower surface layer and mesopelagic.

The SAR324 probe was hybridized to amplified rDNA prepared from vertical profiles of seawater samples and to high-molecular-weight RNA from the Atlantic Ocean and Pacific

Ocean to more accurately determine the position of SAR324 in the water column. In both cases, the SAR324 cluster was found to be most abundant in the aphotic zone, peaking at 200 m at BATS and at 500 m in the profiles from ocean station PAPA (Fig. 5), although this difference between sites may have been due to the different depth ranges sampled. Like that of other uncultured bacterioplankton, the absolute abundance of SAR324 rRNA could not be accurately estimated from hybridization of oligonucleotide probes to rRNA because no pure SAR324 RNA is available for standardization. However, rDNA amplicons from the target organism (SAR324) are available; therefore, the rDNA hybridization values in Fig. 5A are expressed as percentages. At the position of the maximum in its distribution in the Atlantic samples, SAR324 comprised 18% of bacterial rDNA amplicons, indicating that it is a very abundant group.

## DISCUSSION

The data presented here reveal the existence of a previously unknown bacterioplankton group, show that they have a wide biogeographical distribution, and provide insight into their ecological role by demonstrating that these organisms are most abundant in the lower ocean surface layer. Furthermore, the evidence shows that this phylotype is in fact a diverse but monophyletic gene cluster and therefore might be regarded as a collection of species.

The particular emphasis of this investigation was dictated by a long-term research strategy that will utilize fluorescent probes to identify single cells in future studies of environmental samples. From the beginning, it has been clear that strategies involving rRNA probes for uncultured bacterioplankton groups would be sound only if thorough sequence databases that explored the genetic diversity within gene clusters were available for probe design (14, 27). Recently, Amann and colleagues (1) obtained perplexing results when hybridizing fluorescent probes to natural populations of beta-1 proteobacteria in activated sludge. The data verified that the diversity of genes in environmental gene clusters indeed represented real diversity at the cellular level, but they also showed that the specificity of probes could not easily be extrapolated from the analysis of a limited data set of environmental sequences.

The physiology of the SAR324 gene cluster is unknown and cannot be deduced from its observed phylogenetic associations; however, the physiological variability of the DP provides a background for the construction of hypotheses regarding the

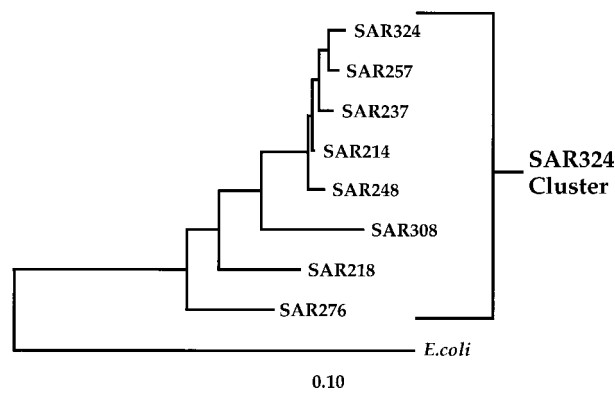


FIG. 4. Phylogenetic tree showing relationships among genes within the SAR324 cluster. This tree was inferred by the neighbor-joining method from ca. 300 nucleotides.

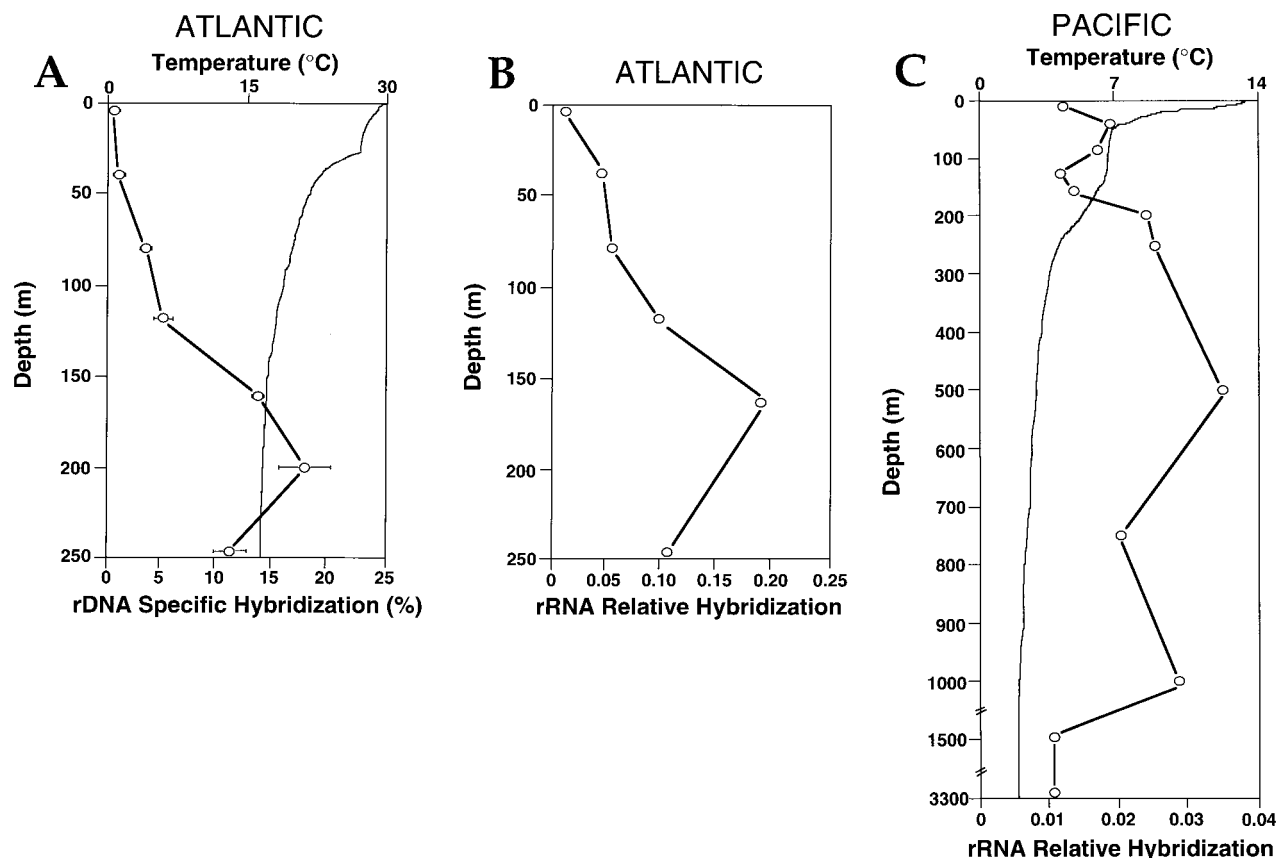


FIG. 5. Distribution of the SAR324 gene among 16S rDNA amplicons and high-molecular-weight RNA prepared from plankton samples. (A) The distribution in percentages of SAR324 rDNA, as a proportion of bacterial rDNA, in the upper 250 m at BATS. The means and standard deviations are shown for triplicate PCRs from a single nucleic acid sample. (B) Hybridization of the SAR324 probe to high-molecular-weight RNA from the upper 250 m at BATS, expressed in relative units. (C) Hybridization of the SAR324 probe to high-molecular-weight RNA from the upper 3,300 m at ocean station PAPA, expressed in relative units.

activity of the SAR324 group. Metabolically, the  $\delta$  subdivision is mainly divided into two groups: the aerobes (bdellovibrios and the myxobacteria) and the anaerobes, which use sulfate or other inorganic compounds as electron acceptors. Although it is unclear in this analysis, it is possible that the sulfide producers form the deepest branch within the DP and that the bdellovibrios and myxobacteria represent aerobic adaptations (35). Furthermore, previous studies characterized the bdellovibrios as a phylogenetically heterogeneous group composed of some "fast-clock" species, which has further complicated the resolution of the phylogenetic positions of the organisms within the DP (35).

The SAR324 lineage represents a unique cluster within the DP. Teske and colleagues (33) have recently obtained 16S rRNA sequences related to genera within the DP, but phylogenetic analyses similar to those described here failed to indicate any specific association between the genes reported in that study and the SAR324 lineage. The relationships reported here are the most significant relationships uncovered following a thorough search of public sequence databases.

The hybridization data are presented as SAR324 rDNA abundance among PCR amplicons, in percentages, and as relative rRNA abundances. As we have shown previously, these measures often lead to qualitatively similar conclusions where general trends in the ecological distributions of bacterioplankton are the subjects of interest (17, 18). Although relative gene frequencies are sensitive to the distribution of rDNA copy number and genome size, as well as cell numbers along envi-

ronmental transects, they nonetheless represent a type of information that is very informative, though it may not correspond directly to biomass or microbial activity (21). Likewise, relative rRNA abundance has its pitfalls, most notably in the fact that it measures protein synthesizing activity and so may underestimate the abundance of populations that are temporarily inactive. Notwithstanding the novelty of these measures, they are emerging as useful indicators of microbial distributions that complement other types of measurements, such as biomass, that often have their own limitations (24).

Although the physiology of the SAR324 cluster is unknown, the diversity and proportionally high abundance of this group in the aphotic zone suggest that this is a group of related species that are functionally specialized for life in abyssal regions of the ocean. Evidence of other microorganisms (SAR406, SAR202 and deep-water variants of SAR11) that specifically inhabit the aphotic zone has been described (17, 18). Collectively, these data suggest that the aphotic zone bacterioplankton community may be a more specialized microbial community than was generally thought previously. Organic carbon is exported to the aphotic zone by sinking particles and by injection of dissolved organic carbon during winter mixing (3). As a major constituent of the aphotic bacterioplankton community, it seems likely that the SAR324 cluster participates somehow in these processes. Although the overall activity of microbes in abyssal regions of the oceans is much reduced relative to microbial activity in the euphotic zone, the ocean depths nonetheless sustain a significant biomass of microorganisms that are

likely to be important in the ecology of the oceans (5). Further studies based on the data presented here will be aimed at elucidating the population genetics and ecological role(s) of these species.

#### ACKNOWLEDGMENTS

We are grateful to the Bermuda Biological Station for Research BATS group and Nanci Adair for collecting and processing nucleic acid samples from BATS, to Nelson Sherry and Michael Lipsen for sampling from ocean station PAPA, and to Frank Whitney, Institute of Ocean Sciences, Sidney, British Columbia, Canada, for the Pacific Ocean physical data. We also thank Douglas Gordon, Brian Lanoil, Michael Rappé, Ena Urbach, Marcelino Suzuki, and Kate Field for their many helpful suggestions.

This work was supported by NSF grant OCE 9016373 and DOE grant FG0693ER61697 to S.J.G. and by an N. L. Tartar Fellowship to T.D.W.

#### REFERENCES

- Amann, R., J. Snaird, M. Wagner, W. Ludwig, and K.-H. Schleifer. 1996. In situ visualization of high genetic diversity in a natural microbial community. *J. Bacteriol.* **178**:3496–3500.
- 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**:1707–1713.
- 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* **371**:405–408.
- Chisholm, S. W., R. J. Olsen, E. R. Zettler, R. Goericke, J. B. Waterbury, and N. A. Welschmeyer. 1988. A novel free-living prochlorophyte abundant in the oceanic euphotic zone. *Nature* **334**:340–343.
- Cho, B. C., and F. Azam. 1988. Major role of bacteria in biogeochemical fluxes in the ocean's interior. *Nature* **332**:441–443.
- DeLong, E. F. 1992. Archaea in coastal marine bacterioplankton. *Proc. Natl. Acad. Sci. USA* **89**:5685–5689.
- DeLong, E. F., D. G. Franks, and A. L. Alldredge. 1993. Phylogenetic diversity of aggregate-attached vs. free-living marine bacterial assemblages. *Limnol. Oceanogr.* **38**:924–934.
- Felsenstein, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* **39**:783–791.
- Felsenstein, J. 1989. PHYLIP 3.5. University of Washington, Seattle.
- Felsenstein, J. 1988. Phylogenies from molecular sequences: inference and reliability. *Annu. Rev. Genet.* **22**:521–564.
- Fuhrman, J. A., K. McCallum, and A. A. Davis. 1992. Novel major archaeobacterial group from marine plankton. *Nature* **356**:148–149.
- Fuhrman, J. A., K. McCallum, and A. A. Davis. 1993. Phylogenetic diversity of subsurface marine microbial communities from the Atlantic and Pacific oceans. *Appl. Environ. Microbiol.* **59**:1294–1302.
- Giovannoni, S. J. 1991. The polymerase chain reaction, p. 177–203. *In* E. Stackebrandt and M. Goodfellow (ed.), *Sequencing and hybridization techniques in bacterial systematics*. John Wiley & Sons, Inc., New York, N.Y.
- Giovannoni, S. J., T. B. Britschgi, C. L. Moyer, and K. G. Field. 1990. Genetic diversity in Sargasso Sea bacterioplankton. *Nature* **345**:60–63.
- Giovannoni, S. J., E. F. DeLong, G. J. Olsen, and N. R. Pace. 1988. Phylogenetic group-specific oligodeoxynucleotide probes for identification of single microbial cells. *J. Bacteriol.* **170**:720–726.
- Giovannoni, S. J., M. S. Rappé, D. Gordon, E. Urbach, M. Suzuki, and K. G. Field. 1996. Ribosomal RNA and the evolution of bacterial diversity, p. 63–85. *In* D. M. Roberts, P. Sharp, G. Alderson, and M. Collins (ed.), *Evolution of microbial life*. Cambridge University Press, Cambridge, United Kingdom.
- Giovannoni, S. J., M. S. Rappé, K. L. Vergin, and N. L. Adair. 1996. 16S rRNA genes reveal stratified open ocean bacterioplankton populations related to the green non-sulfur bacteria. *Proc. Natl. Acad. Sci. USA* **93**:7979–7984.
- Gordon, D. A., and S. J. Giovannoni. 1996. Detection of stratified microbial populations related to *Chlorobium* and *Fibrobacter* species in the Atlantic and Pacific oceans. *Appl. Environ. Microbiol.* **62**:1171–1177.
- Gray, M. W., D. Sankoff, and R. J. Cedergren. 1984. On the evolutionary descent of organisms and organelles: a global phylogeny based on a highly conserved structural core in small subunit ribosomal RNA. *Nucleic Acids Res.* **2**:5837–5852.
- Haddad, H., F. Camacho, P. Durand, and S. C. Cary. 1995. Phylogenetic characterization of the epibiotic bacteria associated with the hydrothermal vent polychaete *Alvinella pompejana*. *Appl. Environ. Microbiol.* **61**:1679–1687.
- Kemp, P. F., S. Lee, and J. LaRoche. 1993. Estimating the growth rate of slowly growing marine bacteria from RNA content. *Appl. Environ. Microbiol.* **59**:2594–2601.
- Lane, D. J., K. G. Field, G. J. Olsen, and N. R. Pace. 1988. Reverse transcriptase sequencing of ribosomal RNA for phylogenetic analysis. *Methods Enzymol.* **167**:138–144.
- 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.
- Lee, S., C. Malone, and P. F. Kemp. 1993. Use of multiple 16S rRNA-targeted fluorescent probes to increase signal strength and measure cellular RNA from natural planktonic bacteria. *Mar. Ecol. Prog. Ser.* **101**:193–201.
- 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.
- Mullins, T. D., T. B. Britschgi, R. L. Krest, and S. J. Giovannoni. 1995. Genetic comparisons reveal the same unknown bacterial lineages in Atlantic and Pacific bacterioplankton communities. *Limnol. Oceanogr.* **40**:148–158.
- Olsen, G. J., D. J. Lane, S. J. Giovannoni, and N. R. Pace. 1986. Microbial ecology and evolution: a ribosomal RNA approach. *Annu. Rev. Microbiol.* **40**:337–365.
- Rappé, M. S., P. F. Kemp, and S. J. Giovannoni. 1995. Chromophyte plastid 16S ribosomal RNA genes found in a clone library from Atlantic Ocean seawater. *J. Phycol.* **31**:979–988.
- Saitou, N., and M. Nei. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**:406–425.
- Schmidt, T. M., E. F. DeLong, and N. R. Pace. 1991. Analysis of marine picoplankton community by 16S rRNA gene cloning and sequencing. *J. Bacteriol.* **173**:4371–4378.
- 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.
- Suzuki, K., N. Handa, H. Kiyosawa, and J. Ishizaka. 1995. Distribution of the prochlorophyte *Prochlorococcus* in the central Pacific Ocean as measured by HPLC. *Limnol. Oceanogr.* **40**:983–989.
- Teske, A., C. Wawer, G. Muyzer, and N. B. Ramsing. 1996. Distribution of sulfate-reducing bacteria in stratified fjord (Mariager Fjord, Denmark) as evaluated by most-probable-number counts and denaturing gradient gel electrophoresis of PCR-amplified ribosomal DNA fragments. *Appl. Environ. Microbiol.* **62**:1405–1415.
- Ward, D. M., R. Weller, and M. M. Bateson. 1990. 16S rRNA sequences reveal numerous uncultured microorganisms in a natural community. *Nature* **345**:63–65.
- Woese, C. R. 1987. Bacterial evolution. *Microbiol. Rev.* **51**:221–271.