APPLIED AND ENVIRONMENTAL MICROBIOLOGY, Nov. 1998, p. 4522–4529 Vol. 64, No. 11
0099-2240/98/$04.00+0

Kinetic Bias in Estimates of Coastal Picoplankton Community Structure Obtained by Measurements of Small-Subunit rRNA Gene PCR Amplicon Length Heterogeneity

MARCELINO SUZUKI,[1,2]† MICHAEL S. RAPPÉ,[2]‡ AND STEPHEN J. GIOVANNONI[2]*

College of Oceanic and Atmospheric Sciences[1] and Department of Microbiology,[2] Oregon State University, Corvallis, Oregon 97331

Received 26 March 1998/Accepted 9 September 1998

Marine bacterioplankton diversity was examined by quantifying natural length variation in the 5′ domain of small-subunit (SSU) rRNA genes (rDNA) amplified by PCR from a DNA sample from the Oregon coast. This new technique, length heterogeneity analysis by PCR (LH-PCR), determines the relative proportions of amplicons originating from different organisms by measuring the fluorescence emission of a labeled primer used in the amplification reaction. Relationships between the sizes of amplicons and gene phylogeny were predicted by an analysis of 366 SSU rDNA sequences from cultivated marine bacteria and from bacterial genes cloned directly from environmental samples. LH-PCR was used to compare the distribution of bacterioplankton SSU rDNAs from a coastal water sample with that of an SSU rDNA clone library prepared from the same sample and also to examine the distribution of genes in the PCR products from which the clone library was prepared. The analysis revealed that the relative frequencies of genes amplified from natural communities are highly reproducible for replicate sets of PCRs but that a bias possibly caused by the reannealing kinetics of product molecules can skew gene frequencies when PCR product concentrations exceed threshold values.

Libraries of small-subunit rRNA gene (SSU rDNA) clones prepared by PCR are widely applied to study the microbial diversity of natural ecosystems. These studies have provided dramatic evidence that the majority of microbial communities are dominated by previously unknown organisms (4, 8, 18). However, quantitative comparisons using clone libraries to assess microbial community structure have been limited by several factors, including (i) undersampling of diversity and (ii) uncertainty about sources of bias in the cloning process, in particular bias by the PCR. Undersampling, often estimated by coverage values or by rarefaction curves, results from the difficulty of processing a large number of clones (11). The lack of an alternative means to quantitatively assess the composition of complex mixtures of rDNAs from in situ communities has made it difficult to evaluate methodological sources of bias by the cloning process. The method we describe here, length heterogeneity analysis by PCR (LH-PCR), overcomes some of these problems by quickly providing a profile of amplicon diversity in complex mixtures of PCR products.

LH-PCR is similar to the approach used in our earlier study of bias in the PCR (16) and the recently published terminal restriction fragment length polymorphism (9) and fluorescent restriction fragment length polymorphism (2) techniques. In both LH-PCR and these methods, the proportions of PCR amplicons originating from different genes are estimated from the fluorescence emission of labeled PCR primers. However, instead of identifying PCR amplicons based on restriction endonuclease sites, in LH-PCR the discrimination of amplicons originating from different organisms is based on natural variation in the lengths of SSU rDNAs.

In a previous study, we investigated biases introduced during the amplification of rDNAs by PCR (16). In that study, the templates consisted of pairwise mixtures of SSU rDNAs from bacteria belonging to three different phylogenetic groups. To estimate bias, we compared the proportions of genes in the PCR products with their proportions in the starting template mixtures. We observed that, above threshold product concentrations, PCR dramatically biased the frequency distribution among gene homologs relative to the original mixture. A kinetic model based on competition between primers and products which successfully explained the experimental results was developed. These results indicated that this type of bias by PCR might lead to an increase in net diversity estimates among amplicons relative to the gene diversity of the native DNA mixture. Evidence also indicated that artifacts resulting from this phenomenon could be controlled by limiting the number of replication cycles to maintain product levels below threshold values. However, the effect of this type of bias on the composition of PCR products amplified from natural community DNA was uncertain since, in a complicated mixture of genes, a single gene might not reach threshold concentrations at which competition between product and primer reannealing would have a pronounced effect.

Here we present the results of a study in which LH-PCR was used to estimate the community composition of bacterioplankton from a water sample collected off the Oregon coast. In order to trace the phylogenetic origin of the domains amplified by LH-PCR, we performed an analysis of the length variability in SSU rDNAs of bacterial strains cultivated or directly cloned from the same seawater sample, as well as sequences retrieved from gene sequence databases. We found that the relative gene frequencies obtained from natural communities by LH-PCR were highly reproducible when PCR product concentrations were limited to relatively low values but that, at high concentrations, the kinetic bias caused by template reanneal-

* Corresponding author. Mailing address: Department of Microbiology, Oregon State University, Corvallis, OR 97331. Phone: (541) 737-1835. Fax: (541) 737-0496. E-mail: giovanns@bcc.orst.edu.

† Present address: Monterey Bay Aquarium Research Institute, Moss Landing, CA 95010.

‡ Present address: Station Biologique, CNRS, INSU et Université Pierre et Marie Curie, Roscoff Cedex, France.

ing significantly skewed gene frequencies. SSU rDNA amplicons with sizes corresponding to the alpha subdivision of the class *Proteobacteria* (alpha-*Proteobacteria*) represented the largest fraction of the bacterial rDNA amplicons (ca. 65%), while no other size class of SSU rDNA amplicons represented greater than 10% of the bacterial rDNA amplicons. Overall, the results suggest that LH-PCR is an effective tool for assessing microbial community structure and that clone libraries may often overrepresent bacterioplankton diversity because the relative frequencies of dominant species have been reduced by a systematic bias.

MATERIALS AND METHODS

Sample collection, nucleic acid isolation, and clone library construction. On 28 April 1993, a subsurface (10-m) water sample was collected by Niskin bottles at a station located 8 km off the mouth of Yaquina Bay, Oreg. (44°39.1′N, 124°10.6′W). The water was prescreened through 10-μm-pore-size Nitex mesh and transported in autoclaved polyethylene carboys to the laboratory for the remaining analyses. Picoplankton from 4-liter (subsample 1) and 16-liter (subsample 2) subsamples were collected by filtration onto 0.2-μm-pore-size polysulfone filters (Supor-200; Gelman Sciences Inc., Ann Arbor, Mich.). Total cellular nucleic acids were isolated from the picoplankton samples by lysis with proteinase K and sodium dodecyl sulfate, followed by phenol-chloroform extraction as previously described (6). A portion of the DNA sample isolated from subsample 2 was used as template in the amplification of nearly full-length SSU rDNAs by PCR and subsequently cloned into a plasmid vector as described elsewhere (12, 17). SSU rDNA clones recovered in this library have been partially described elsewhere (12, 17).

LH-PCR. Ten nanograms of purified genomic DNA from each subsample was used as template for LH-PCR. The forward primer, 27F (5′-AGA GTT TGA TCM TGG CTC AG-3′) (3), was 5′ end labeled with the phosphoramidite dye 6-FAM (graciously supplied by Applied Biosystems Inc., Foster City, Calif.) or purchased from Genset (San Diego, Calif.). The reverse primers used were 355R (5′-GCT GCC TCC CGT AGG AGT-3′) (1) for domain A and 536R (5′-GWA TTA CCG CGG CKG CTG-3′) (5) for domain B, synthesized at the Central Services Laboratory, Center for Gene Research and Biotechnology, Oregon State University. In a final volume of 100 μl, reaction mixtures contained 0.2 mM premixed deoxynucleoside triphosphates (Stratagene, La Jolla, Calif.), 1.5 mM $MgCl_2$, 5% acetamide, 0.5 μM forward primer, 0.5 μM (one) reverse primer, and 2.5 U of *Taq* DNA polymerase (Promega, Madison, Wis.). All reactions used the Ampliwax hot-start protocol (Perkin-Elmer Cetus, Norwalk, Conn.) in a PTC100 thermal cycler (MJ Research Inc., Watertown, Mass.) programmed to 16 cycles for primer 355R (except for the reactions evaluating PCR bias) or 21 cycles for primer 536R, each consisting of 96°C denaturation for 1 min, 55°C annealing for 1 min, and 72°C extension for 3 min.

The concentration of labeled PCR products was measured in a Shimadzu UV160U spectrophotometer (Shimadzu Co., Kyoto, Japan) or estimated after electrophoresis in an agarose minigel stained with ethidium bromide (50 μg/ml) and compared with mass standards. The PCR products were purified with Qiaquick spin columns (Qiagen, Chatsworth, Calif.). Approximately 10 fmol of the LH-PCR products was discriminated by Long Ranger (FMC, Rockland, Maine) polyacrylamide gel electrophoresis in a model 377 automated DNA sequencer (Applied Biosystems Inc.) with Genescan (Applied Biosystems Inc.), a software package that estimates the sizes of bands in the gel and their integrated fluorescence emission. The output of the software is electropherograms in which the bands are represented by peaks and the integrated fluorescence of each band is the area under the peaks (see Fig. 1). The integrated fluorescence increased linearly with concentrations of up to 50 fmol of PCR products, indicating that the relative proportion of the integrated fluorescence of each peak corresponded to the proportion of each amplicon in the PCR products (data not shown). The relative abundance of amplicons was estimated as the ratio between the integrated fluorescence of each of the peaks and the total integrated fluorescence of all peaks.

Length heterogeneity analysis of published sequences. In LH-PCR, amplicons originating from different templates are identified by length heterogeneity in hypervariable regions of the SSU rDNA. Three such regions occur in the 5′ end of the gene, around locations homologous to *Escherichia coli* positions 90, 190, and 450. In order to verify the phylogenetic coherence of length heterogeneity contained in these variable regions, we compared the length heterogeneity of domains homologous to the domain between *E. coli* positions 8 and 355 (domain A) and positions 8 and 536 (domain B). The analysis included previously published sequences of bacterial species isolated from the same water sample as that used for the LH-PCR analysis or directly cloned from DNA extracted from subsample 2 (17), as well as SSU rDNA sequences of bacterial species isolated from seawater or directly cloned from DNA extracted from seawater, retrieved from the GenBank, Ribosomal Database Project (10), and ARB (14) sequence databases.

Bias by PCR. Two experiments were performed to evaluate the introduction of bias by PCR. In order to evaluate the bias described by Suzuki and Giovannoni (16) in the amplification of domain A (16) and to optimize the number of cycles for LH-PCR, we performed a time course experiment in which PCRs of domain A, with DNA purified from subsamples 1 and 2 as templates, were stopped by freezing at 10 (only subsample 1), 12, 14, 16, 18, 20, and 25 cycles. Concentrations of LH-PCR products from subsample 1 were measured spectrophotometrically as described above. Concentrations of LH-PCR products from subsample 2 were estimated from the agarose minigel as described above, except for the products of reactions with subsample 2, and stopped after 12 cycles, which were calculated assuming an amplification efficiency of 85% per cycle (13).

To evaluate the introduction of reannealing bias by PCR in the amplification of full-length rDNAs from mixed populations of bacteria, we used the optimized LH-PCR protocol for domain A as described above to compare the genotypic bacterioplankton community structure of (i) a genomic DNA sample from the Oregon coast (subsample 2) and (ii) the nearly full-length PCR products from subsample 2 after 35 cycles of amplification, used to prepare the SSU rDNA clone library, as described previously (12, 17). Triplicate LH-PCRs were performed as described above with 10 ng of genomic DNA or 60 pg of nearly full-length SSU rDNA PCR amplicons as templates, calculated so that the reactions using genomic DNA and full-length PCR amplicons contained approximately the same numbers of copies of SSU rDNAs. For this calculation, we assumed a bacterial origin for 50% of the DNA, an average chromosome size of 2 Mbp, and an average of two copies of the ribosomal operon per chromosome.

Finally, to estimate the introduction of bias by the cloning per se, we compared the community structure estimated by LH-PCR from full-length PCR products amplified from subsample 2 to that inferred from the relative proportion of SSU rDNA clones recovered in the clone library, grouped according to the sizes of domain A obtained directly from their SSU rDNA sequences.

RESULTS

Predicted length heterogeneity in the 5′ region of SSU rDNAs. Three regions at the 5′ end of the SSU rDNA (V1, *E. coli* SSU rDNA positions 72 to 101; V2, *E. coli* SSU rDNA positions 176 to 221; and V3, *E. coli* SSU rDNA positions 451 to 481) are variable between different phylogenetic groups of bacteria. Insertions and deletions in these variable regions cause natural variability in the nucleotide lengths of molecules amplified with the 27F and 355R primer pair (domain A, ca. 312 to 363 bp) and the 27F and 536R primer pair (domain B, ca. 472 to 574 bp).

The lengths of domains A and B of bacteria isolated from seawater or SSU rDNAs directly cloned from seawater DNA are shown in Table 1 and are generally coherent with phylogenetic relationships. Many discrete fragment lengths are monophyletic but are shared by multiple species (e.g., 316 bp). Alpha-*Proteobacteria* and cyanobacteria have the shortest lengths for both domains. Beta-, gamma-, and delta-*Proteobacteria* and the *Flexibacter-Bacteroides-Cytophaga* group have intermediate lengths, and the longest domains are those from genes of low- and high-G+C gram-positive bacteria and members of the *Vibrio fischeri* subgroup of the gamma-*Proteobacteria*. Most phylogenetic groups have a unique combination of lengths for domains A and B (i.e., alpha-*Proteobacteria* have a domain A length of 315 bp and domain B lengths between 470 and 472 bp). The lengths of domains A and B of genes with plastid origins were not included in this study and are described elsewhere (12).

Analyses of coastal bacterioplankton diversity. An example of an LH-PCR electropherogram is shown in Fig. 1. It shows the length heterogeneity of domain A for PCR products obtained directly from DNA extracted from seawater subsample 2. The 23 peaks are labeled A through W and correspond to amplicons with varying lengths in domain A. Organisms that produce amplicons corresponding in size to these peaks were identified by reference to a clone library prepared from the same seawater sample. These clones are indicated in Table 1 by the prefix "env.OCS."

We were able to assign an OCS gene clone or R2A cellular clone to peaks from natural community DNA (Table 1). The domain A peaks of sizes 317 bp (E), 318 bp (F), 319 bp (G),

TABLE 1. Length in nucleotides between positions homologous to *E. coli* SSU rDNA positions 8 through 355 (domain A) or 8 through 536 (domain B) for a variety of marine bacteria[a]

Accession no.	Cellular strain or environmental gene clone	Taxonomic affiliation	Size of domain (bp)	
			A	B
	env.OCS28 [A]	α	312	468
U70681	env.OM55	α	312	468
U75264	env.SAR490	α	312	ND
X52170	env SAR12	c	313	ND
U75260	env.SAR418	α	313	ND
U75262	env.SAR440	α	313	ND
U75263	env.SAR466	α	313	ND
X52169	env.SAR6	c	313	470
	Synechococcus sp. strain WH8101	c	313	471
	Synechococcus sp. strain WH8103	c	313	472
U78945	env.OCS122 [B]	α	314	470
U75259	env.SAR414	α	314	ND
	env.SAR420	α	314	472
X52171	env.SAR7	c	314	471
U64002	*Rhizobium* sp. strain BAL25	α	314	470
U63957	*Zoogloea* sp. strain BAL43	α	314	470
L10934	env.FL1	α	315	470
L10935	env.FL11	α	315	471
	env.OCS24 [C]	α	315	471
U70678	env.OM25	α	315	471
U75258	env.SAR241	α	315	471
U75254	env.SAR464	α	315	471
X78315	*Roseobacter algicola*	α	315	471
L15345	Strain LFR	α	315	472
U63935	*Caulobacter* sp. strain BAL3	α	316	472
U75252	env.OCS12 [D]	α	316	472
	env.OCS126 [D]	α	316	472
U78942	env.OCS19 [D]	α	316	ND
U78943	env.OCS84 [D]	α	316	ND
U70684	env.OM136	α	316	472
U70585	env.OM143	α	316	471
U70686	env.OM155	α	316	472
U70692	env.OM299	α	316	ND
U70679	env.OM38	α	316	472
U70680	env.OM42	α	316	472
U70682	env.OM65	α	316	473
X52280	env.SAR1	α	316	473
X52172	env.SAR11	α	316	473
U75255	env.SAR203	α	316	472
U75256	env.SAR211	α	316	472
U75257	env.SAR220	α	316	472
	env.SAR402	α	316	472
U64003	*Erythrobacter* sp. strain BAL26	α	316	472
U64005	*Erythrobacter* sp. strain BAL28	α	316	472
U64011	*Erythrobacter* sp. strain SCB34	α	316	472
U64025	*Erythrobacter* sp. strain SCB48	α	316	472
U63952	*Erythromicrobium* sp. strain BAL34	α	316	472
U63958	*Flavobacterium* sp. strain BAL44	α	316	ND
U63939	*Rhizomonas* sp. strain BAL11	α	316	472
U63934	*Rhodobacter* sp. strain BAL2	α	316	472
U63949	*Rhodobacter* sp. strain BAL27	α	316	472
U63956	*Sphingomonas* sp. strain BAL40	α	316	472
U63959	*Sphingomonas* sp. strain BAL45	α	316	472
U63960	*Sphingomonas* sp. strain BAL46	α	316	472
U63962	*Sphingomonas* sp. strain BAL48	α	316	485
U63937	*Sphingomonas* sp. strain BAL5	α	316	470
U63998	*Sphingomonas* sp. strain SCB21	α	316	472
U78913	Strain R2A114 [D]	α	316	ND
U78918	Strain R2A163 [D]	α	316	ND
U78919	Strain R2A166 [D]	α	316	ND
U78910	Strain R2A62 [D]	α	316	ND
U78912	Strain R2A84 [D]	α	316	ND
	env.OCS138 [E]	α	317	473
	env.OCS154 [E]	α	317	473
	env.OCS180 [E]	α	317	473
	env.OCS53 [E]	α	317	473
U70687	env.OM188	α	317	473
U70689	env.OM242	α	317	ND
U75649	env.SAR193	α	317	473
	env.SAR222	α	317	ND
	env.SAR239	α	317	ND
	env.SAR258	α	317	ND
U75253	env.SAR407	α	317	473
U14583	Strain 307	α	318	474
U64009	Strain BAL32	α	318	474

Continued

TABLE 1—*Continued*

Accession no.	Cellular strain or environmental gene clone	Taxonomic affiliation	Size of domain (bp)	
			A	B
U78909	Strain R2A57 [F]	α	318	ND
U64019	Strain SCB42	α	318	472
	env.OCS14 [J]	α	324	480
	env.OCS27 [J]	α	324	483
U70683	env.OM75 [J]	α	326	482
U78944	env.OCS116 [M]	α	328	484
U70714	env.OM110	γ	328	ND
U78917	Strain R2A153 [M]	α	328	ND
U65915	env.SAR276	δ	329	ND
	env.OCS124 [N]	α	330	486
M58793	*Microscilla marina*	f	330	505
U20797	env.SAR202	x	331	488
U20798	env.SAR307	x	331	488
	env.OCS2	γ	334	ND
U63945	*Aeromonas* sp. strain BAL19	γ	336	ND
U05570	*Methylobacterium pelagicum*	γ	338	518
U14585	Strain 34-p	α	338	520
U70702	env.OM241	γ	339	520
U70696	env.OM60	γ	339	520
L35470	env.SAR160	γ	339	520
U85887	*Flavobacterium* sp. strain A103	f	339	515
U85888	*Flavobacterium* sp. strain A265	f	339	515
D32219	Strain K189C	γ	339	ND
U64010	*Aeromonas* sp. strain SCB33	γ	340	ND
X82144	*Alteromonas luteoviolacea*	γ	340	522
X82147	*Alteromonas rubra*	γ	340	522
M93352	*Deleya aquamarina*	γ	340	521
	env.OCS111 [Q]	β	340	521
	env.OCS7 [Q]	β	340	521
L35469	env.SAR156	γ	340	521
U85873	*Halomonas variabilis*	γ	340	521
U85872	*Halomonas variabilis*	γ	340	521
U85871	*Halomonas variabilis*	γ	340	521
L42618	*Halomonas variabilis*	γ	340	521
L35540	*Methylobacterium pelagicum*	γ	340	521
X72775	*Methylomicrobium pelagicum*	γ	340	521
RDP	*Oceanospirillum kriegii*	γ	340	521
U63961	*Rhodoferax* sp. strain BAL47	β	340	461
U85854	Strain C079	γ	340	521
D32220	Strain K189B	γ	340	ND
D32221	Strain unid gamma-proteobacterium	γ	340	ND
U78920	Strain R2A9 [Q]	γ	340	ND
X82143	*Alteromonas espejiana*	γ	341	522
X67024	*Alteromonas haloplanktis*	γ	341	522
L10938	*Alteromonas macleodii*	γ	341	522
X82140	*Alteromonas undina*	γ	341	523
	env.OCS43 [R]	β	341	ND
L35471	env.SAR166	γ	341	519
	env.SAR470	γ	341	ND
	env.SAR471	γ	341	523
U63946	*Flavobacterium* sp. strain BAL22	f	341	ND
U63938	*Flavobacterium* sp. strain BAL9	f	341	ND
X87339	*Methylophaga thalassica*	γ	341	521
X98336	*Pseudoalteromonas antarctica*	γ	341	522
U85857	*Pseudoalteromonas* sp. strain MB6-03	γ	341	522
U85858	*Pseudoalteromonas* sp. strain MB8-02	γ	341	522
U64012	*Aeromonas* sp. strain SCB35	γ	342	ND
U64020	*Aeromonas* sp. strain SCB43	γ	342	523
U63953	*Alcaligenes* sp. strain BAL37	β	342	ND
X82141	*Alteromonas piscicida*	γ	342	524
U63943	*Cytophaga* sp. strain BAL17	f	342	ND
U78946	env.OCS181	γ	342	523
	env.OCS44	γ	342	523
	env.OCS5	γ	342	523
	env.OCS66	β	342	523
	env.OCS98	β	342	521
U70718	env.OM111	γ	342	ND
U70698	env.OM133	γ	342	ND
U70694	env.OM23	γ	342	523
U70697	env.OM93	γ	342	523
U63954	*Flavobacterium* sp. strain BAL38	f	342	ND
X82134	*Pseudoalteromonas atlantica*	γ	342	523
X82136	*Pseudoalteromonas carrageenovora*	γ	342	523
U85856	*Pseudoalteromonas* sp. strain IC006	γ	342	523
U85859	*Pseudoalteromonas* sp. strain IC013	γ	342	523
U85860	*Pseudoalteromonas* sp. strain MB6-05	γ	342	523
U85861	*Pseudoalteromonas* sp. strain SW08	γ	342	523

Continued on following page

TABLE 1—*Continued*

Accession no.	Cellular strain or environmental gene clone	Taxonomic affiliation	Size of domain (bp) A	Size of domain (bp) B
U85862	*Pseudoalteromonas* sp. strain SW29	γ	342	523
U85870	*Pseudomonas* sp. strain A177	γ	342	523
U85868	*Pseudomonas* sp. strain ACAM213	γ	342	523
U63942	*Pseudomonas* sp. strain BAL16	γ	342	ND
U63944	*Pseudomonas* sp. strain BAL18	γ	342	ND
U63947	*Pseudomonas* sp. strain BAL23	γ	342	ND
U64001	*Pseudomonas* sp. strain BAL24	γ	342	523
U85869	*Pseudomonas* sp. strain IC038	γ	342	523
U65012	*Pseudomonas stutzeri*	γ	342	523
U26420	*Pseudomonas stutzeri* ZoBell	γ	342	523
U78922	Strain R2A30	γ	342	ND
U63941	*Zoogloea* sp. strain BAL15	β	342	ND
U85897	*Arthrobacter* sp. strain MB6-07	h	343	504
U85896	*Arthrobacter* sp. strain MB8-13	h	343	516
U85893	*Arthrobacter* sp. strain MB90	h	343	503
U63940	*Cytophaga* sp. strain BAL13	f	343	ND
U70693	env.OM10	γ	343	523
U65912	env.SAR248	δ	343	499
U65908	env.SAR324	δ	343	499
U85890	*Flavobacterium* sp. strain ACAM123	f	343	519
U85889	*Flavobacterium* sp. strain IC001	f	343	519
U85891	Marine psychrophile IC025	f	343	519
X95640	*Methylophaga thalassica*	γ	343	524
U78924	Strain R2A103 [S]	f	343	ND
U64027	*Alteromonas* sp. strain SCB50	γ	344	525
X82061	*Corynebacterium glutamicum*	h	344	507
U70706	env.OM156	β	344	525
U70707	env.OM180	β	344	525
	env.SAR267	x	344	502
M63811	env.SAR92	γ	344	525
L35461	env.SAR86	γ	344	525
U85879	*Psychrobacter glacincola*	γ	344	525
U85878	*Psychrobacter glacincola*	γ	344	525
U85877	*Psychrobacter glacincola*	γ	344	525
U85876	*Psychrobacter glacincola*	γ	344	525
U85875	*Psychrobacter* sp. strain IC008	γ	344	525
U85874	*Psychrobacter* sp. strain MB6-21	γ	344	525
U78941	Strain R2A170	h	344	ND
U64026	Strain SCB49	f	344	ND
M58794	*Microscilla sericea*	f	345	520
U64021	*Antarcticum* sp. strain SCB44	f	345	521
X80629	*Corynebacterium glutamicum*	h	345	508
X84257	*Corynebacterium glutamicum*	h	345	508
Z46753	*Corynebacterium glutamicum*	h	345	508
M62796	*Cytophaga lytica*	f	345	521
L10948	env.AGG13	f	345	521
L10945	env.AGG41	f	345	520
L10946	env.AGG58	f	345	519
	env.SAR242	x	345	ND
M58775	*Flectobacillus glomeratus*	f	345	521
X67022	*Marinobacter hydrocarbonoclasticus*	γ	345	527
U63999	*Marinobacter* sp. strain SCB22	γ	345	526
X67025	*Marinomonas vaga*	γ	345	526
U14586	Strain 301	f	345	521
U85883	Strain IC054	f	345	ND
U85885	Strain IC063	f	345	521
U85884	Strain IC066	f	345	ND
U78933	Strain R2A10	f	345	ND
U78935	Strain R2A132	f	345	ND
U78940	Strain R2A160	h	345	ND
U78939	Strain R2A54	f	345	ND
U64015	Strain SCB38	f	345	462
M61002	*Vesiculatum antarcticum*	f	345	521
X82135	*Alteromonas aurantia*	γ	346	527
X82137	*Alteromonas citrea*	γ	346	527
X82137	*Alteromonas denitrificans*	γ	346	528
U85895	*Arthrobacter* sp. strain IC044	h	346	507
U64000	*Chromohalobacter* sp. strain SCB23	γ	346	ND
U85844	*Colwellia* sp. strain IC068	γ	346	527
U85845	*Colwellia* sp. strain ICP11	γ	346	527
L42615	*Deleya cupida*	γ	346	527
M93354	*Deleya marina*	γ	346	527
L42616	*Deleya pacifica*	γ	346	527
	env.OCS178	β	346	527
U70699	env.OM182	γ	346	527
U70704	env.OM43	β	346	527
U70705	env.OM58	β	346	ND
U70695	env.OM59	γ	346	527
	env.SAR226	x	346	504
	env.SAR250	x	346	503
	env.SAR259	x	346	501
	env.SAR269	x	346	504
U85863	*Marinobacter* sp. strain IC022	γ	346	527
U85864	*Marinobacter* sp. strain IC032	γ	346	527
RDP	*Marinomonas communis*	γ	346	527
RDP	*Marinomonas vaga*	γ	346	527
RDP	*Oceanospirillum beijerincki*	γ	346	527
U85880	*Psychrobacter immobilis*	γ	346	527
U78930	Strain R2A148	γ	346	ND
U85881	Strain IC051	f	346	522
U78931	Strain R2A173	γ	346	ND
U78924	Strain R2A44	γ	346	ND
U78927	Strain R2A86	γ	346	ND
U78928	Strain R2A88	γ	346	ND
U64017	Strain SCB40	f	346	489
U64022	Strain SCB45	γ	346	472
M62788	*Cyclobacterium marinus*	f	347	522
L10944	env.AGG32	f	347	522
U70703	env.OM252	γ	347	528
	env.SAR251	x	347	504
	env.SAR432	h	347	503
U63955	*Flavobacterium* sp. strain BAL39	f	347	ND
U64023	*Flexibacter* sp. strain SCB46	f	347	523
X87755	*Kytococcus sedentarius*	h	347	508
DEW	*Serratia rubidacea*	γ	347	527
U85906	*Shewanella frigidimarina*	γ	347	528
U85882	Strain IC076	f	347	523
Z25522	Strain purple	γ	347	524
U78929	Strain R2A113	γ	347	ND
U78932	Strain R2A5	f	347	ND
U85900	*Arthrobacter* sp. strain MB6-20	h	348	509
U63951	*Azospirillum* sp. strain BAL31	γ	348	ND
U85841	*Colwellia* sp. strain IC062	γ	348	529
U85842	*Colwellia* sp. strain IC064	γ	348	529
L10950	env.AGG53	γ	348	528
X54745	env.WHB461	γ	348	529
X54744	env.WHB462	γ	348	529
M58784	*Flexibacter litoralis*	f	348	525
RDP	*Oceanospirillum jannaschii*	γ	348	529
M22365	*Oceanospirillum linum*	γ	348	529
U85855	*Pseudoalteromonas* sp. strain MB8-11	γ	348	529
U85905	*Shewanella frigidimarina*	γ	348	529
U85904	*Shewanella frigidimarina*	γ	348	529
U85903	*Shewanella frigidimarina*	γ	348	529
U85902	*Shewanella frigidimarina*	γ	348	529
U85907	*Shewanella gelidimarina*	γ	348	529
X81623	*Shewanella putrefaciens*	γ	348	529
U85886	Strain ACAM210	f	348	524
X76334	*Vibrio vulnificus*	γ	348	529
X74727	*Vibrio vulnificus*	γ	348	529
X74726	*Vibrio vulnificus*	γ	348	529
X76333	*Vibrio vulnificus*	γ	348	529
Z22992	*Vibrio vulnificus*	γ	348	529
X56582	*Vibrio vulnificus*	γ	348	529
U64004	*Xanthomonas* sp. strain BAL27	γ	348	ND
U85846	*Colwellia* sp. strain ACAM179	γ	349	530
U85843	*Colwellia* sp. strain IC072	γ	349	530
M58770	*Cytophaga marinoflava*	f	349	525
U70708	env.OM271	f	349	525
U70709	env.OM273	f	349	525
	env.SAR196	n	349	520
L35504	*Nitrospina gracilis*	δ	349	508
RDP	*Oceanospirillum maris*	γ	349	530
U64014	Strain SCB37	f	349	480
X74685	*Photobacterium angustum*	γ	350	531
D25310	*Photobacterium phosphoreum*	γ	350	532
X74687	*Photobacterium phosphoreum*	γ	350	531
Z19107	*Photobacterium phosphoreum*	γ	350	531
U85908	*Shewanella hanedai*	γ	350	531
X82133	*Shewanella putrefaciens*	γ	350	532
U63948	*Shewanella* sp. strain BAL25	γ	350	ND
U14582	Strain 90-P(gv)1	γ	350	531
U85849	Strain IC004	γ	350	531

Continued on following page

TABLE 1—*Continued*

Accession no.	Cellular strain or environmental gene clone	Taxonomic affiliation	Size of domain (bp)	
			A	B
U85852	Strain IC085	γ	350	531
U78923	Strain R2A37 [U]	γ	350	ND
U14581	Strain S51-W(gv)1	γ	350	531
RDP	*Vibrio marinus*	γ	350	531
X82134	*Vibrio marinus*	γ	350	533
U85847	*Colwellia* sp. strain IC035	γ	351	532
U85853	Strain IC067	γ	351	532
U64018	Strain SCB41	f	351	ND
L35468	env.SAR145	γ	352	532
M96493	*Nitrococcus mobilis*	β	352	533
X82132	*Shewanella hanedai*	γ	352	533
	env.OCS155	h	353	509
U70710	env.OM1	h	353	509
U70711	env.OM231	h	353	ND
U41450	env.OCS307	fb	353	535
U85851	Strain IC059	γ	353	534
X56756	*Vibrio alginolyticus*	γ	353	535
U34043	env.SAR406	fb	354	536
D55729	*Planococcus okeanokoites*	l	354	534
U85899	*Planococcus* sp. strain IC024	l	354	534
X56578	*Vibrio harveyi*	γ	354	536
X56581	*Vibrio natriegens*	γ	354	536
D25308	*Photobacterium histaminum*	γ	355	537
D25309	*Photobacterium leiognathi*	γ	355	537
X62172	*Planococcus citreus*	l	355	535
U85898	*Planococcus* sp. strain MB6-16	l	355	535
D83367	*Staphylococcus halodurans*	l	355	534
Z26896	*Staphylococcus halodurans*	l	355	536
X66100	*Staphylococcus halodurans*	l	355	536
L37600	*Staphylococcus halodurans*	l	355	538
U78937	Strain R2A180 [V]	l	355	ND
X56575	*Vibrio campbellii*	γ	355	537
	env.SAR272	x	356	514
U14584	*Flectobacillus* sp. strain S38-W(gv)1	f	356	537
U78938	Strain R2A161	l	356	ND
	env.SAR256	x	357	514
X70642	*Listonella pelagia*	γ	357	ND
U85867	*Marinobacter* sp. strain IC065	γ	357	ND
X74722	*Listonella pelagia*	γ	358	539
X74686	*Photobacterium leiognathi*	γ	358	539
X74691	*Vibrio alginolyticus*	γ	358	539
X74690	*Vibrio alginolyticus*	γ	358	539
X74692	*Vibrio campbellii*	γ	358	539
X74702	*Vibrio fischeri*	γ	358	539
X70640	*Vibrio fischeri*	γ	358	538
X74706	*Vibrio harveyi*	γ	358	539
X74710	*Vibrio mediterranei*	γ	358	539
X74714	*Vibrio natriegens*	γ	358	539
X74716	*Vibrio nereis*	γ	358	539
X74717	*Vibrio nigripulchritudo*	γ	358	539
X74719	*Vibrio orientalis*	γ	358	539
U64016	*Vibrio* sp. strain SCB39	γ	358	524
U64024	*Vibrio* sp. strain SCB47	γ	358	523
Z31657	*Vibrio splendidus*	γ	358	538
X74724	*Vibrio splendidus*	γ	358	539
U64013	*Flexibacter* sp. strain SCB36	f	360	527
U64006	*Vibrio* sp. strain BAL29	γ	360	ND
U64007	*Vibrio* sp. strain BAL30	γ	360	526
X62171	*Marinococcus halophilus*	l	363	545
X90835	*Marinococcus halophilus*	l	374	556
U85901	*Halobacillus* sp. strain MB6-08	l	392	574

[a] Lengths were calculated for bacteria and plastids isolated from seawater, as well as for SSU rDNAs cloned from community DNA. Published sequences were obtained from the RDP, GenBank, and ARB databases. We also included unpublished sequences of genes cloned from environmental DNA from the Oregon coast (12, 15) (prefix env.OCS) and the Sargasso Sea (prefix env.SAR). Letters in brackets correspond to the peaks in Fig. 1. α, β, γ, and δ, alpha, beta, gamma, and delta subdivisions, respectively, of the *Proteobacteria*; f, *Flexibacter*, *Bacteroides*, and *Cytophaga* phylum; l, low-G+C gram-positive phylum; h, high-G+C gram-positive phylum; x, *Chloroflexus* and *Herpetosiphon* phylum; c, cyanobacteria; fb, *Fibrobacter* phylum; n, *Nitrospira* phylum; ND, not determined.

and 321 bp (H) correspond to the sizes of SSU rDNAs of plastid origin (12). The domain A peak of 317 bp corresponds to sizes of both alpha-*Proteobacteria* and plastids (12). There is strong evidence that the 317-bp domain A peak corresponds mainly to alpha-*Proteobacteria*, since the ratio between the integrated fluorescence of peak E and all bacterial peaks is approximately the same for samples whether they were filtered through 0.8-μm-pore-size polycarbonate membranes, which removed the other plastid peaks, or not (12). Figure 2 presents electropherogram data in a histogram format. The data from Fig. 1 are shown in panel A, with error bars shown to represent the standard deviations for triplicate PCRs from the same DNA sample. Here, as in other measurements, we found the method to be highly reproducible.

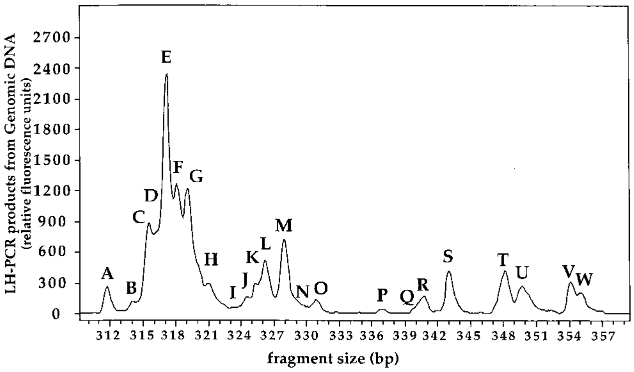

FIG. 1. Electropherogram of DNA fragments amplified by PCR with primer set A from genomic DNA isolated from seawater subsample 2. The letters A to W correspond to the peaks detected by the Genescan 2.1.2 software in at least one of triplicate reactions. The *x* axis represents the size of domains in base pairs estimated by comparison to the size standard GS2500 (Applied Biosystems Inc.). The *y* axis represents relative fluorescence units.

Bias by PCR. The possibility that a kinetic bias caused by template reannealing could occur during the amplification of domain A from bacterioplankton samples was investigated by examining the relationship between the final concentration of products obtained and the relative frequency of dominant genes in the population. A portion of this analysis is provided in Fig. 3, which shows the relative frequency of the 317-bp fragment (alpha-*Proteobacteria* and prymnesiophyte plastids) as a function of the total product molarity. The prediction for the kinetic bias effect is that the proportion of the dominant peak (the percentage of integrated fluorescence) should decrease with increasing product molarity, as observed in Fig. 3. This prediction assumes that the dominant peaks are composed primarily of genes of one or a few types. The final concentrations of product amplicons for the reactions used for Fig. 3 varied for the two samples of DNA isolated independently from the same water sample (subsamples 1 and 2) and also varied according to the number of cycles used for the amplification (12 to 25 cycles for subsample 1 and 12 to 18 cycles for subsample 2). We considered the possibility that biases caused by primer selection, which are dependent on the number of cycles, might have caused the bias. A plot of the same data shown in Fig. 3 with the number of amplification cycles replacing the final product concentrations on the abscissa revealed no relationship (data not shown). The observed bias is in accord with predictions for the kinetic bias effect.

Figure 1B shows the results of an optimized domain A LH-PCR analysis (i.e., the final product concentration was lower than 1.5 nM) of the full-length PCR products that were used to prepare the clone library. The template for this reaction was PCR amplicons obtained with the 27F-1542R primers from the same natural DNA sample that was used as template for the LH-PCR in Fig. 1A. There is a significant difference between the profiles in Fig. 1A and B. The comparison between panels

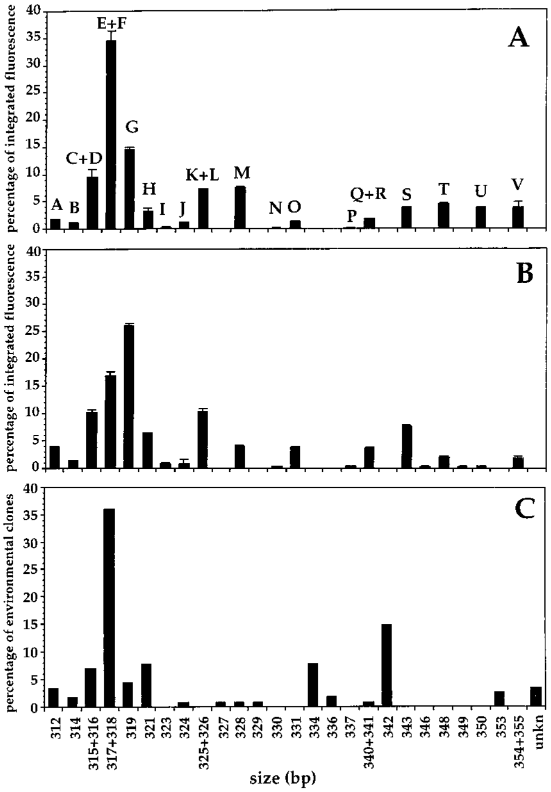

FIG. 2. Comparison between LH-PCR and SSU rDNA clone library. The figure shows the percentage of integrated fluorescence of each individual domain A produced by the optimized LH-PCR (final product concentration, <1.5 nM) from genomic DNA isolated from subsample 2 (A) or nearly full-length SSU rDNA PCR amplicons used to construct the clone library (B). The *x* axis represents the size of domains in base pairs, estimated by comparison to the size standard GS2500 (Applied Biosystems Inc.). The relative abundance of clones recovered in the OCS clone library classified by the length of domain A is shown in panel C. Error bars each represent one standard deviation from the mean of triplicate reactions.

A and B shows that the relative contributions of some peaks (A, F, G, K and L, N, P, and Q) increased when the full-length PCR amplicons were used as a template instead of genomic DNA, while other peaks (M, S, T, and U and V) decreased. We noted that the relative proportion of peak F (319 bp) is higher than that of peaks D and E (317 and 318 bp) in the PCR amplicons, although both are major peaks that contribute significantly to the total population of molecules.

It is evident that the reamplification reactions provided reproducible results and that therefore the difference between Fig. 1A and 1B could be attributed to the original PCR used to prepare amplicons for clone library construction. To explain these results, we reasoned that some of the differences between Fig. 1A and B (i.e., the decrease of the dominant peak at 317 bp) might be attributed to bias caused by the kinetic (template reannealing) effect, since 35 cycles of amplification were used for the PCR, and final product concentrations were greater than 10 nM.

Figure 1C shows the distribution of clones from the gene clone library, discriminated by the length of the domain A of their SSU rDNA, for comparisons to Fig. 1A and B. It is evident that the profiles in Fig. 1B and C are very different. The causes of these differences are uncertain but probably include a combination of three sources of error. The random error resulting from the sample size (number) of clones retrieved and identified in the library imposed limitations on the expected correspondence between LH-PCR results and clone library gene frequencies. Systematic biases by LH-PCR or by one of the steps of the cloning process are also plausible explanations for the observations.

Comparison of domain A with domain B. In general, there was good agreement between the community structure estimated by LH-PCR for domains A and B (data not shown). The main difference between LH-PCR for domains A and B was the resolution of different peaks by the Genescan software. Genescan resolved more peaks in the analysis of domain A and tended to merge domain B peaks, especially peaks for larger fragments (peaks > 520 bp). Some adjacent peaks of domain A were also merged in some of the electropherograms (peaks C and D, E and F, K and L, and Q and R).

Phylogenetic composition of the community. The relative proportions of LH-PCR peaks conform to previous observations that alpha-*Proteobacteria* dominate SSU rDNA clone libraries of surface samples. Peaks A to E and K to M, which correspond to alpha-*Proteobacteria* and plastids, respectively (Table 1), represent about 65% of the total fluorescence. Peaks F, G, and L correspond to plastid sequences (12). Most of the remaining peaks do not correspond to coherent phylogenetic groups when reference is made to all SSU rDNA sequences of bacteria isolated or environmental clones from seawater. However, most of the peaks had a corresponding isolate or environmental SSU rDNA clone from the same seawater sample. Peaks P and Q, which represented about 7% of the total fluorescence, corresponded to the sizes of beta-, gamma-, and delta-*Proteobacteria*, *Flexibacter-Bacteroides-Cytophaga*, and high-G+C gram-positive bacteria, many of which are cultivated strains. Peak S represented about 5% of the total fluorescence and corresponded to the sizes of previously cultivated members of the gamma-*Proteobacteria*. Finally, peaks T to V represented about 9% of the total fluorescence and corresponded to the sizes of several phylogenetic groups.

DISCUSSION

The adoption of molecular techniques for assessing microbial diversity has engendered a far-reaching appreciation of the importance of uncultured microbes but has also led to concerns about the limitations of the new methodologies. The ap-

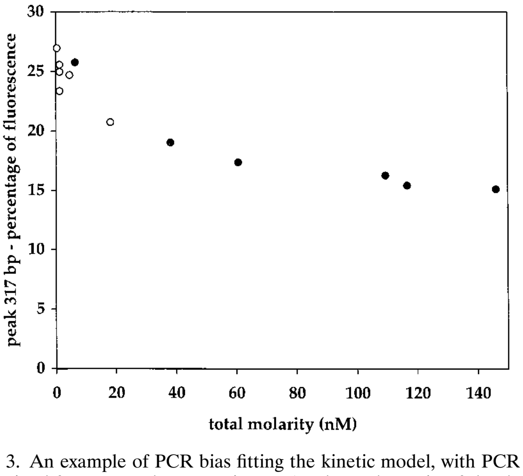

FIG. 3. An example of PCR bias fitting the kinetic model, with PCR amplicons obtained from natural community DNA. The molar ratio of the dominant fragment (317 bp) to total products is plotted as a function of the final product concentration. Primer set A was used for the amplification from environmental DNA subsamples 1 (solid circles) and 2 (open circles).

proach we employed here, LH-PCR, was designed to address some of these concerns in the context of the study of complex natural communities.

We found that gene frequencies measured by PCR amplification can be highly reproducible. We also observed a bias that selectively reduced the relative frequency of a dominant size class of fragments as a function of increasing final product concentrations, which can be explained by the template reannealing bias described by Suzuki and Giovannoni (16). This bias is caused by the fact that, as the reaction progresses, amplicons increase in concentration and primers decrease in concentration. At a certain point, amplicons should reanneal, inhibiting the primers from annealing and stalling the reaction. Assuming that different SSU rDNAs do not cross-reanneal, genes with higher initial concentrations in the original sample should experience template reannealing at lower combined product concentrations—which for reactions run at the same initial conditions should be dependent on the number of replication cycles—than genes with lower initial concentrations. Therefore, in reactions experiencing this bias, dominant species should be underrepresented and rarer species should be overrepresented. The fact that the observed trend was related to combined product concentration but not to the number of amplification cycles supports the idea that this bias is due to reannealing kinetics and not to some other form of bias, such as primer selection. This reannealing bias became significant above product concentrations of about 2 nM, and therefore, we recommend that amplification reactions be stopped before reaching this value.

Differences between the LH-PCR electropherograms obtained from natural community DNA and those obtained from the reamplification of full-length SSU rDNA amplicons indicate that the PCR may significantly bias the composition of clone libraries. However, such biases do not appear to occur randomly but rather are systematic. The shift in the dominance from the peak of 317 bp to the peak of 319 bp contradicts our previous expectation (16) that reannealing bias would lead amplicons originating from different templates to reach similar concentrations. A possible explanation for this discrepancy is the fact that each of the LH-PCR peaks represents SSU rDNAs originating from several different organisms. Reannealing between domains originating from different organisms could explain the observed shift in dominance between the two peaks. If the degree of similarity between the sequences with a domain A length of 317 bp were high enough for PCR amplicon cross-hybridization and kinetic inhibition, while the degree of similarity between sequences with a domain A length of 319 bp were low enough to not lead to cross-hybridization, one could envision that each of the 319-bp amplicons would experience lower levels of kinetic inhibition than the 317-bp amplicons. The average similarity among four SSU rDNA clones with a 317-bp domain A was 0.93 (0.89 to 1.00), while the degree of similarity between two SSU rDNA clones with a 319-bp domain A was 0.85, supporting this hypothesis. Another hypothesis which might explain the observed peak shift would be a large difference in the degree of diversity among the organisms corresponding to each of the peaks. Template reannealing inhibition should theoretically be lower for peaks with more gene types or peaks lacking a dominant gene type. This hypothesis cannot be tested with the data included in the current analysis.

Uncertainties about the numbers of ribosomal operons in different bacterioplankton species and differences in the relative sizes of genomes preclude the extrapolation from gene frequencies to cell abundance. Nevertheless, relative gene frequencies offer some advantages as a measurement for assessing the composition of natural microbial communities. In particular, rDNA frequency histograms (electropherograms) should be relatively insensitive to short-term variation in growth rates, which may affect rRNA abundance significantly under some circumstances (7).

LH-PCR is a promising method for the analysis of natural microbial populations. The main advantages of LH-PCR are that it surveys relative gene frequencies within complex mixtures of DNA, is reproducible, requires small sample sizes, and can be performed with many samples simultaneously. Furthermore, some of the size classes emerging from LH-PCR analyses can be related at the group level to environmental rDNA sequences. However, overlapping size classes leave ambiguities that require further analyses to resolve. The relative proportions of the electropherogram peaks from seawater are in agreement with previous findings that alpha-*Proteobacteria* members are dominant components of clone libraries constructed from DNA extracted from surface seawater. The observation that most SSU rDNAs of organisms cloned or cultivated from the same water sample have sizes that correspond to the peaks in the LH-PCR electropherograms also supports the validity of the method.

The main technical problem associated with LH-PCR is the accuracy of peak detection, especially when longer domains are used. Improvements in automated DNA sequencers and in Genescan software may increase the accuracy of the method for longer domains. This problem notwithstanding, domain B was useful to confirm the results obtained with domain A and, in some cases, to differentiate between phylogenetic groups with identical sizes in domain A, like the alpha-*Proteobacteria* and prymnesiophyte plastids.

The attributes of LH-PCR make it useful for quick assessments of the diversity of natural microbial communities for comparative purposes, for experimental designs that involve the manipulation of natural microbial communities (15), and for experiments, such as those we describe here, aimed at investigating the properties of PCR in applications employing complex mixtures of gene homologs.

ACKNOWLEDGMENTS

We are grateful to the staff of the Oregon State University Center for Gene Research Central Services Laboratory, and particularly to Anne-Marie Girard, for technical assistance.

This research was supported by grant OCE-9618530 from the National Science Foundation.

REFERENCES

1. **Amann, R. I., L. Krumholz, and D. A. Stahl.** 1990. Fluorescent-oligonucleotide probing of whole cells for determinative, phylogenetic, and environmental studies in microbiology. J. Bacteriol. **172:**762–770.
2. **Bruce, K.** 1997. Analysis of *mer* gene subclasses within bacterial communities in soils and sediments resolved by fluorescent-PCR-restriction fragment length polymorphism profiling. Appl. Environ. Microbiol. **63:**4914–4919.
3. **Giovannoni, S. J.** 1991. The polymerase chain reaction, p. 177–203. *In* E. Stackebrandt and M. Goodfellow (ed.), Nucleic acid techniques in bacterial systematics. John Wiley & Sons, Inc., New York, N.Y.
4. **Giovannoni, S. J., T. B. Britschgi, C. L. Moyer, and K. G. Field.** 1990. Genetic diversity in Sargasso Sea bacterioplankton. Nature **345:**60–63.
5. **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. (Erratum, **170:**2418.)
6. **Giovannoni, S. J., E. F. DeLong, T. M. Schmidt, and N. R. Pace.** 1990. Tangential flow filtration and preliminary phylogenetic analysis of marine picoplankton. Appl. Environ. Microbiol. **56:**2572–2575.
7. **Kramer, J. G., and F. L. Singleton.** 1993. Measurement of rRNA variations in natural communities of microorganisms on the southeastern U.S. continental shelf. Appl. Environ. Microbiol. **59:**2430–2436.
8. **Liesack, W., and E. Stackebrandt.** 1992. Occurrence of novel groups of the domain Bacteria as revealed by analysis of genetic material isolated from an Australian terrestrial environment. J. Bacteriol. **174:**5072–5078.

9. **Liu, W. T., T. L. Marsh, H. Cheng, and L. J. Forney.** 1997. Characterization of microbial diversity by determining terminal restriction fragment length polymorphisms of genes encoding 16S rRNA. Appl. Environ. Microbiol. **63:** 4516–4522.
10. **Maidak, B. L., G. J. Olsen, N. Larsen, R. Overbeek, M. J. McCaughey, and C. R. Woese.** 1997. The RDP (Ribosomal Database Project). Nucleic Acids Res. **25:**109–111.
11. **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.
12. **Rappé, M. S., M. T. Suzuki, K. L. Vergin, and S. J. Giovannoni.** 1998. Phylogenetic diversity of ultraplankton plastid small-subunit rRNA genes recovered in environmental nucleic acid samples from the Pacific and Atlantic coasts of the United States. Appl. Environ. Microbiol. **64:**294–303.
13. **Sardelli, A.** 1993. Plateau effect—understanding PCR limitations. Amplifications **9:**1,3,5,6.
14. **Strunk, O., O. Gross, B. Reichel, M. May, S. Hermann, N. Stuckman, B. Nonhoff, M. Lenke, A. Ginhart, A. Vilbig, T. Ludwig, A. Bode, K.-H. Schleifer, and W. Ludwig.** ARB: a software environment for sequence data. Submitted for publication.
15. **Suzuki, M. T.** The effect of protistan bacterivory on bacterioplankton diversity. Submitted for publication.
16. **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.
17. **Suzuki, M. T., M. S. Rappé, Z. W. Haimberger, H. Winfield, N. Adair, J. Ströbel, and S. J. Giovannoni.** 1997. Bacterial diversity among small-subunit rRNA gene clones and cellular isolates from the same seawater sample. Appl. Environ. Microbiol. **63:**983–989.
18. **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.