APPLIED AND ENVIRONMENTAL MICROBIOLOGY, Apr. 2002, p. 1576–1584 Vol. 68, No. 4
0099-2240/02/$04.00+0 DOI: 10.1128/AEM.68.4.1576–1584.2002

Phylogenetic Diversity of Marine Cyanophage Isolates and Natural Virus Communities as Revealed by Sequences of Viral Capsid Assembly Protein Gene g20†

Yan Zhong,[1] Feng Chen,[2]* Steven W. Wilhelm,[3] Leo Poorvin,[3] and Robert E. Hodson[1]

Department of Marine Sciences, University of Georgia, Athens, Georgia 30602[1]; Center of Marine Biotechnology, University of Maryland Biotechnology Institute, Baltimore, Maryland 21202[2]; and Department of Microbiology, The University of Tennessee, Knoxville, Tennessee 37996[3]

Received 21 November 2001/Accepted 25 January 2002

In order to characterize the genetic diversity and phylogenetic affiliations of marine cyanophage isolates and natural cyanophage assemblages, oligonucleotide primers CPS1 and CPS8 were designed to specifically amplify ca. 592-bp fragments of the gene for viral capsid assembly protein g20. Phylogenetic analysis of isolated cyanophages revealed that the marine cyanophages were highly diverse yet more closely related to each other than to enteric coliphage T4. Genetically related marine cyanophage isolates were widely distributed without significant geographic segregation (i.e., no correlation between genetic variation and geographic distance). Cloning and sequencing analysis of six natural virus concentrates from estuarine and oligotrophic offshore environments revealed nine phylogenetic groups in a total of 114 different g20 homologs, with up to six clusters and 29 genotypes encountered in a single sample. The composition and structure of natural cyanophage communities in the estuary and open-ocean samples were different from each other, with unique phylogenetic clusters found for each environment. Changes in clonal diversity were also observed from the surface waters to the deep chlorophyll maximum layer in the open ocean. Only three clusters contained known cyanophage isolates, while the identities of the other six clusters remain unknown. Whether or not these unidentified groups are composed of bacteriophages that infect different *Synechococcus* groups or other closely related cyanobacteria remains to be determined. The high genetic diversity of marine cyanophage assemblages revealed by the g20 sequences suggests that marine viruses can potentially play important roles in regulating microbial genetic diversity.

Viruses are known to be an abundant and dynamic component in the marine microbial communities that can regulate the biomass production and species composition of bacteria and phytoplankton, influence biogeochemical cycling, and mediate gene transfer between microorganisms in aquatic ecosystems (1, 6, 18, 22, 24, 29, 33). Unicellular cyanobacteria of the genus *Synechococcus* are among the most abundant forms of marine picoplankton and reportedly contribute up to 25% of the primary production in the open ocean (28). Viruses that infect specific strains of marine *Synechococcus* spp. (known as cyanophages) typically range from 10^3 to 10^5 ml^{-1} in near-shore and offshore waters and sometimes can reach concentrations in excess of 10^6 ml^{-1} (11, 19, 21, 27). Proctor and Fuhrman (18) found that 1 to 3% of *Synechococcus* spp. from a variety of marine environments contained visible phage particles. Estimates done by using different approaches have indicated that viruses may infect 2 to 8% of the *Synechococcus* population daily (8, 21).

Viruses also potentially influence the genetic diversity of microbial communities, mainly via virus selection pressure and virus-mediated genetic exchanges such as transduction, transformation, and lysogenic conversion (6, 9, 27). Cyanophage isolates are diverse in terms both of their morphology (20, 27) and their genetic fingerprints (11, 32). However, little is known about the genetic diversity and phyletic linkages of cyanophages in natural marine environments. Recently, the genes for viral capsid assembly g20 proteins from three marine cyanophages that infect marine *Synechococcus* spp. were sequenced and their sequences were similar to that of the g20-encoding gene of bacteriophage T4, a gene responsible for the initiation of head assembly, DNA packing, and head-tail junction (7). Several conserved regions were found to be unique among the g20 sequences of three marine cyanomyoviruses (7). A set of PCR primers (CPS1 and CPS2) was designed based on the conserved regions of these cyanophages and used to specifically amplify a 165-bp g20 fragment from cyanomyoviruses in natural seawater (7). This primer set was subsequently applied in PCR-DGGE (denaturing gradient gel electrophoresis) to study cyanophage diversity along a transect in the Atlantic Ocean (30, 31). High genetic diversity of cyanophage g20 sequences was found in both surface waters and the deep layer of the euphotic zone. Natural cyanophage community structure changed significantly from surface to depth, and maximum diversity was always coincident with maximum *Synechococcus* sp. abundance (31).

It is difficult to resolve the complex interactions between cyanophage and *Synechococcus* populations without knowing their phylogenetic distribution and phyletic variation in different environments. While natural marine *Synechococcus* populations have been well characterized phylogenetically (13, 14,

* Corresponding author. Present address: Center of Marine Biotechnology, University of Maryland Biotechnology Institute, 701 East Pratt St., Suite 236, Baltimore, MD 21202. Phone: (410) 234-8866. Fax: (410) 234-8896. E-mail: chenf@umbi.umd.edu.

† Contribution no. 568 from the Center of Marine Biotechnology, University of Maryland Biotechnology Institute.

TABLE 1. Specificity test of PCR amplification with primers CPS1 and CPS2

Sample(s)	Family[a]	PCR product	Host(s)	Area of isolation	Reference
Cyanophage					
P3	M	+	WH7805	Sapelo Island, Ga.	11
P5	M	+	WH7803	Sapelo Island, Ga.	11
P6	M	+	WH7805	Dauphin Island, Ala.	11
P8	M	+	WH8101	Sapelo Island, Ga.	11
P12	M	+	WH8101	Sayll Estuary, Ala.	11
P16	M	+	WH7803	Savannah River Estuary, Ga.	11
P17	M	+	WH7803	Qingdao Coast, China	11
P39	M	+	WH7805	Savannah River Estuary, Ga.	11
P61	M	+	WH7803	Satilla River Estuary, Ga.	11
P66	M	+	WH7803	Satilla River Estuary, Ga.	11
P73	M	+	WH7805	Satilla River Estuary, Ga.	11
P76	M	+	WH8007	Altamaha River Estuary, Ga.	11
P77	M	+	WH8007	Altamaha River Estuary, Ga.	11
P79	M	+	WH7805	Satilla River Estuary, Ga.	11
P81	M	+	WH7805	Altamaha River Estuary, Ga.	11
ϕ9	M	+	WH7803	Woods Hole, Mass.	27
S-PWM1	M	+	WH7803	Gulf of Mexico	20
S-BnM1	M	+	WH7803	Bergen, Norway	32
S-PM2	M	+	WH7803	Plymouth, United Kingdom	32
S-WHM1	M	+	WH7803	Woods Hole, Mass.	32
P1	S	−	WH7803	Sayll Estuary, Ala.	11
SBP1	P	−	WH7803	Bermuda	32
ϕ12	P	−	WH7805	Gulf Stream	27
Other phages					
IH6-ϕ1	M	−	IH6[b]	Inner Harbor, Baltimore, Md.	This study
IH6-ϕ7	P	−	IH6[b]	Inner Harbor, Baltimore, Md.	This study
IH11-ϕ2	M	−	*Alteromonas* sp.	Inner Harbor, Baltimore, Md.	This study
IH11-ϕ5	P	−	*Alteromonas* sp.	Inner Harbor, Baltimore, Md.	This study
IH-ϕ1	M	−	*Vibrio natriegens*	Inner Harbor, Baltimore, Md.	This study
CB8-ϕ2	M	−	CB8[b]	Chesapeake Bay, Md.	This study
CB8-ϕ6	M	−	CB8[b]	Chesapeake Bay, Md.	This study
CB8-ϕ7	P	−	CB8[b]	Chesapeake Bay, Md.	This study
CB-ϕ8	M	−	*Vibrio alginolyticus*	Chesapeake Bay, Md.	This study
CB-ϕ9	M	−	*Aeromonas* sp.	Chesapeake Bay, Md.	This study
T4	M	−	*Escherichia coli*	Sigma	
T5	S	−	*Escherichia coli*	Sigma	
T7	P	−	*Escherichia coli*	Sigma	
Viral lysates					
27A		+	WH7803	Sargasso Sea	This study
31B		+	WH7803	Sargasso Sea	This study
32A		+	WH7803	Sargasso Sea	This study
44B		+	WH7803	Sargasso Sea	This study
Viral community					
SE		+		Skidaway, Ga. (31°59′N, 81°01′W)	This study
GS26		+		Gulf Stream (37°19′N, 71°37′W)	This study
GS27		+		Gulf Stream (36°24′N, 71°20′W)	This study
SS40		+		Sargasso Sea (28°53′N, 65°04′W)	This study
SS47		+		Sargasso Sea (34°43′N, 68°07′W)	This study
SS48		+		Sargasso Sea (36°47′N, 71°03′W)	This study

[a] M, P, and S represent virus families *Myoviridae*, *Podoviridae*, and *Siphoviridae*, respectively.
[b] Strain not characterized.
[c] +, positive PCR amplification; −, no desired PCR product.

25), little is known about the phyletic relatedness and diversity within natural cyanophage communities in different marine environments. Our earlier study has shown that the short g20 gene sequence flanked by primers CPS1 and CPS2 (7) yielded less robust phylogenetic inference compared to that flanked by primers CPS1 and CPS8 (Y. Zhong, F. Chen, and R. E. Hodson, Abstr. 100th Gen. Meet. Am. Soc. Microbiol. abstr. 100, p. 465, 2000).

In this study, primers CPS1 and CPS8 were used to amplify the g20 gene fragment (ca. 592 bp) from cyanophage isolates and natural viral assemblages. The phylogenetic diversity of cyanophage communities in various marine environments was investigated based on the partial g20 gene sequences recovered from natural viral communities.

MATERIALS AND METHODS

Phage strains. Twenty-three strains of cyanophage isolates and 13 strains of bacteriophages that infect heterotrophic bacteria were used in this study (Table 1). Cyanophages were isolated by the dilution method (11), while bacteriophages were isolated by plaque assay (23). Phage DNA was extracted by using the protocols described by Lu et al. (11).

TABLE 2. Stations where VCs were collected

VC name	Date	Concn factor	Depth (m)	Latitude	Longitude	Temp (°C)	Salinity (‰)
SE1	8/29/99	14,630	Surface	31°59′N	81°01′W	26.5	33.0
GS26	5/27/00	27,077	Surface	37°19′N	71°37′W	26.3	36.2
GS27	5/28/00	9,908	83	36°24′N	71°20′W	21.9	36.4
SS40	6/5/00	11,046	134	28°53′N	65°4′W	25.4	36.4
SS47	6/12/00	13,831	100–148	34°43′N	68°7′W	23.9	36.7
SS48	6/13/00	21,417	Surface	36°47′N	71°3′W	23.8	36.7

Sampling and ultrafiltration. Water samples (50 to 100 liters) were collected from the river estuary and the open ocean (Table 2). SE1 was a surface sample collected from the pier of the Skidaway Institute of Oceanography in Savannah, Ga. GS26 and GS27 were from the surface and deep chlorophyll maximum (DCM) of the Gulf Stream along the edge of the Sargasso Sea. SS48, SS47, and SS40 were from a surface station and two DCM stations in the Sargasso Sea. Viral particles in the water samples were concentrated by following the protocols described by Chen et al. (3). Briefly, 40 to 80 liters of water was filtered through 0.22- or 0.45-μm-pore-size low-protein-binding Durapore membranes (Millipore). The filtrate was concentrated by ultrafiltration through an Amicon S10Y30 spiral cartridge (Millipore) with a molecular weight cutoff of 30,000 in a ProFlux M-12 system (Millipore) at 30% of the maximum pump speed and 16 to 18 kPa of back pressure. Virus concentrates (VCs) were stored in the dark at 4°C. In order to further concentrate viral particles, an aliquot of 30 ml of VC was centrifuged with an ultracentrifuge (Sorvall Discovery 100S) with a SURESPIN 630 swinging-bucket rotor (Sorvall Inc., Newtown, Conn.) at 40,000 × *g* and 4°C for 3 h. The virus pellet was resuspended in 200 μl of distilled water and kept frozen at −20°C.

Primer design and PCR amplification. Primers CPS1 and CPS2 were designed by Fuller et al. (7). In order to include more nucleotide sequences of g20 for phylogenetic analysis, three other cyanophage-specific oligonucleotide primers (CPS3, CPS4, and CPS8) were designed and used to amplify the g20 gene from cyanophages and their natural assemblages. The sequences of primers CPS3, CPS4, and CPS8 are 5′-TGGTA(T/C)GT(T/C)GATGG(A/C)AGA-3′ (nucleotides 388 to 405), 5′-CAT(A/T)TC(A/T)TCCCA(A/T/C)TCTTC-3′ (nucleotides 1237 to 1253), and 5′-AAATA(C/T)TT(G/A/T)CCAACA(A/T)ATGGA-3′ (nucleotides 1354 to 1376), respectively. One microliter of either extracted DNA (0.1 μg) or VC was used as a DNA template for PCR amplification. The reaction mixture (total volume, 25 μl) contained 1 μl of template DNA, 20 pmol of CPS1 or CPS3, 16 pmol of CPS8 or CPS4, 1× PCR buffer (50 mM Tris-HCl, 100 mM NaCl, 0.15 mM $MgCl_2$), 250 μM each deoxyribonucleotide triphosphate, and 0.75 U of Expand High Fidelity DNA polymerase (Roche, Indianapolis, Ind.). PCR amplification was carried out with a PTC-200 DNA Engine thermocycler (MJ Research). Thermal cycling consisted of an initial denaturation step of 94°C for 3 min, followed by 35 cycles of denaturation at 94°C for 15s, annealing at 36°C for 15s, ramping at 0.3°C/s, and elongation at 73°C for 1 min, with a final elongation step of 73°C for 4 min. A 6-μl aliquot of PCR product was analyzed by electrophoresis in a 1.5% agarose gel in 0.5× Tris-borate-EDTA buffer (5.4 g of Tris base, 2.75 g of boric acid, 0.29 g of EDTA, 1 liter of distilled water) and stained with ethidium bromide for 15 min. The gel image was captured and analyzed with a gel documentation system (Alpha Innotech Corp., San Leandro, Calif.). Replicate PCR amplifications were combined to decrease the bias caused by PCR drift (17, 32).

Clone library construction. The PCR amplicon from each VC was purified with the Wizard PCR prep DNA purification system (Promega). The purified products were ligated into the pGEM-T Easy cloning vector (Promega) and then transformed into JM109 competent cells (Promega) in accordance with the manufacturer's instructions. Positive clones (white colonies) were picked randomly and transferred onto a new agar plate. Clones were numbered 1 to 80 and assigned the prefix SE (for Skidaway Estuarine), GS (for Gulf Stream), or SS (for Sargasso Sea).

Sequencing. About 35 clones from each clone library were randomly picked, and the plasmid inserts were PCR amplified with vector-specific primers T7 and SP6 along with Expand High Fidelity DNA polymerase (Roche). PCR amplification involved a 3-min initial denaturation, followed by 35 cycles of 94°C for 15s, 45°C for 15s, 73°C for 1 min, and a final 4-min extension at 73°C. PCR products of the expected sizes were purified as described above and sequenced bidirectionally with primers T7 and SP6 using the BigDye Terminator Cycle Sequencing Kit (Applied Biosystems Inc.) on an ABI310 or ABI3700 automated DNA sequencer (Applied Biosystems Inc.). For cyanophage isolates, the PCR products were purified and sequenced directly on both strands with CPS1 and CPS8, respectively.

Phylogenetic analysis. Pairs of sequences were aligned and corrected manually with Sequencher 3.0 software (Gene Codes Corporation, Ann Arbor, Mich.), and the consensus sequences were further analyzed. Sequence alignment and similarity analysis were performed with the Genetics Computer Group (Madison, Wis.) package and checked visually for chimeric artifacts. Phylogenetic reconstruction was accomplished with the phylogeny inference package (PHYLIP version 3.5 [5]). Evolutionary distances were calculated by the Jukes-Cantor method with the DNADIST program. Phylogenetic trees were constructed from distance matrices by the neighbor-joining method, which was implemented with the NEIGHBOR program. The node reproducibility for tree topology was estimated by bootstrap analysis, which included 100 replicate data sets.

Nucleotide sequence accession numbers. The nucleotide sequences determined in this study have been submitted to GenBank and assigned accession numbers AY027938 to AY028078.

RESULTS

Primer specificity and PCR amplification. Two sets of primers were designed to detect cyanophage isolates and natural cyanophage concentrates. Primers CPS3 and CPS4, designed within a conserved region of the g20 gene, could amplify a ca. 860-bp fragment from some of the cyanophage isolates but not from the natural VCs. Primers CPS1 and CPS8 efficiently amplified g20 gene fragments from both cyanophage isolates and VCs. Therefore, they were used to investigate cyanophage diversity and phylogenetic relationships. Among a wide range of cyanophages and other bacteriophages tested for primer specificity, the primers CPS1 and CPS8 could amplify only cyanophages belonging to the family *Myoviridae*. No cyanophages of the family *Podoviridae* or *Siphoviridae* or other bacteriophages could be amplified (Table 1). For all of the 20 cyanomyovirus strains (cyanophage of the family *Myoviridae*) tested, an expected product of ca. 592 bp was amplified successfully. Amplification with VCs yielded results similar to those obtained with cyanophage isolates.

Phylogenetic diversity of cyanophage isolates. The sequence alignment of the g20 gene fragments from 11 cyanophage isolates demonstrated that the region amplified by primers CPS1 and CPS8 was highly conserved and therefore suitable for inference of genetic relatedness among cyanophages (data not shown). Nucleotide similarity analysis revealed that cyanophage isolates shared 64.7 to 96.1% sequence similarity with each other, while enteric coliphage T4 was only 50.4 to 55.2% similar to any cyanophage isolates (Table 3). Based on the phylogenetic affiliation among cyanophage isolates and clones from six natural virus communities (shown later; refer to Fig. 3), these isolates were assigned to three clusters, with most of the isolates falling into clusters I and III and only one isolate (31B) falling into cluster II. S-BnM1 and P17 were distant from

TABLE 3. Percent nucleotide sequence similarity between cyanophage isolates and phage T4

Isolate or phage	% Nucleotide sequence similarity													
	31B	27A	32A	P12	P6	S-PM2	S-WHM1	P77	P79	P81	44B	S-PWM1	P17	T4
S-BnM1	70.9	69.8	69.1	69.1	68.3	67.8	70.4	69.6	69.1	68.0	69.5	68.5	69.3	55.1
31B		71.4	69.7	68.2	68.4	68.2	65.0	64.7	68.4	66.6	66.2	66.2	68.3	51.6
27A			90.4	79.6	80.1	76.8	69.4	70.0	70.2	68.6	70.2	68.9	69.9	53.7
32A				81.2	80.5	75.0	68.5	68.3	68.2	66.0	68.3	68.0	69.7	52.6
P12					90.0	83.5	70.6	70.9	70.0	68.2	68.5	68.7	72.8	53.1
P6						84.0	69.8	69.8	69.8	68.0	68.5	67.8	71.5	53.9
S-PM2							67.2	67.6	68.5	66.7	67.0	67.4	70.8	53.5
S-WHM1								96.1	79.9	78.1	79.7	77.6	71.0	55.2
P77									81.5	79.8	81.9	78.9	71.5	54.6
P79										96.1	93.9	89.6	73.6	53.0
P81											90.2	87.9	70.5	52.2
44B												87.6	72.8	53.9
S-PWM1													71.9	51.9
P17														50.4
T4														

any isolates or clones; therefore, they were not grouped into any clusters (Fig. 1).

Phylogenetic diversity of cyanophages in natural virus communities. Cyanophage diversity in estuarine water, the open ocean surface, and the DCM was investigated by examining cyanophage g20 gene diversity in concentrated virus communities (Fig. 2a to f). For each community, about 35 clones were sequenced, among which identical clones were encountered, suggesting that the diversity of cyanophage populations had been adequately sampled (13). A comprehensive phylogenetic tree consisting of isolated cyanophages and representative clones from natural virus communities was constructed to show their genetic relatedness (Fig. 3).

(i) Skidaway Estuarine surface sample (SE1). A total of 29 different g20 sequences were assigned to five clusters, with the majority of the sequences belonging to clusters A and I (Fig. 2a). Nine different nucleotide sequences in cluster A shared greater than 92.9% similarity, while the similarity among the 15 different nucleotides in cluster I varied from 79.3 to 98.2%. According to the phylogenetic affiliation revealed in Fig. 3, three minor clusters (E, F, and III) were assigned to this community, with only one or two clones in each cluster. Clone SE38 could not be assigned to any cluster due to the lack of any closely related neighbors.

(ii) Gulf Stream surface sample (GS26). The Gulf Stream surface sample (GS26) community exhibited much less diversity in the cyanophage g20 gene pool than the Skidaway Estuarine surface community did. Among the 37 clones sequenced, 13 different g20 sequences were found and assigned to three clusters. GS2608 and GS2624 were in clusters III and B, respectively, while all of the other sequences fell into cluster E with 95.8 to 99.8% pairwise nucleotide similarity (Fig. 2b).

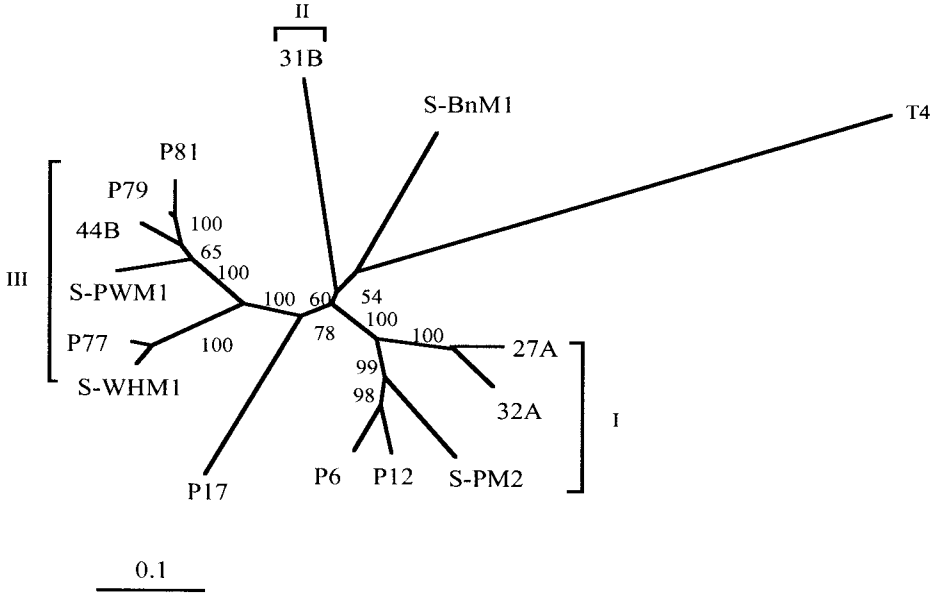

FIG. 1. Phylogenetic affiliation of cyanophage isolates based on 552-bp nucleotide sequences between primers CPS1 and CPS8. Coliphage T4 was used as an outgroup. Values at tree branches are bootstrap values with 100 replicates. The scale bar is equivalent to 0.1 replacement per site. Only bootstrap values larger than 50 are shown.

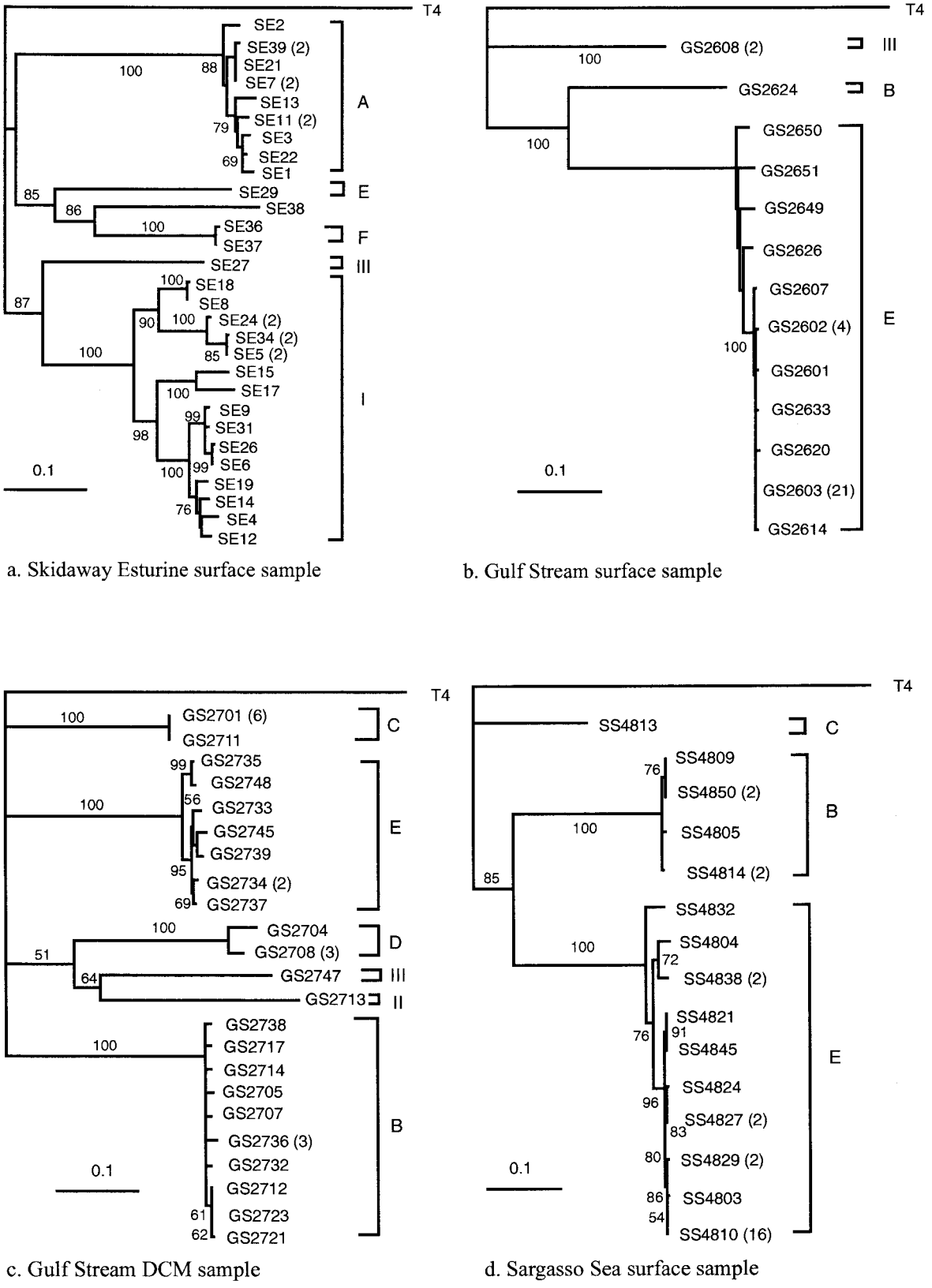

FIG. 2. Phylogenetic relationships among the clones amplified with primers CPS1 and CPS8 from concentrated natural virus communities at Skidaway Estuarine surface SE1 (a), Gulf Stream surface GS26 (b), Gulf Stream DCM GS27 (c), Sargasso Sea surface SS48 (d), Sargasso Sea DCM SS47 (e), and Sargasso Sea DCM SS40 (f). The neighbor-joining tree was constructed based on 552-bp nucleotide sequence alignment with T4 as the outgroup. Only bootstrap values of greater than 50 are shown. Each value in parentheses is the number of clones bearing identical nucleotide sequences in a community. The clusters were assigned on the basis of Fig. 3.

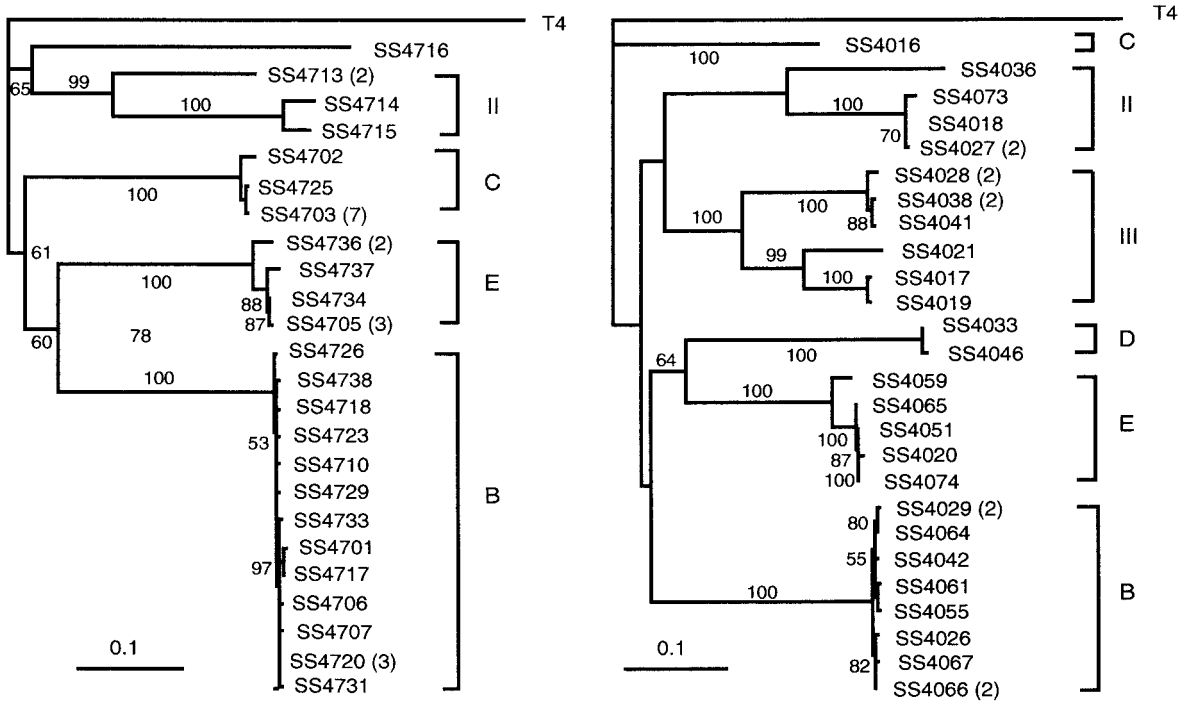

FIG. 2—*Continued.*

(iii) Gulf Stream DCM sample (GS27). The cyanophage community at the DCM of the Gulf Stream appeared to be more diverse than that at the surface (GS26). Among the 33 clones analyzed, 23 different g20 gene sequences were encountered. The DCM layer contained the same three clusters as did the surface sample (clusters III, B, and E) plus three other clusters (C, D, and II) that were not present on the surface (Fig. 2c). However, unlike the surface sample (GS26), where cluster E was dominant, clusters B and E were the two major clusters at the DCM of the Gulf Stream.

(iv) Sargasso Sea surface sample (SS48). The g20 gene diversity and phylogenetic structure of the Sargasso Sea surface sample (SS48) were similar to those of the surface sample in the Gulf Stream (GS26). Fifteen different g20 sequences were found in the 35 clones analyzed. Of the three phyletic clusters in this community, clusters B and E were both present and E was the dominant cluster (Fig. 2d). The higher number of repeat sequences and fewer phyletic groups in GS26 and SS48 indicated a less diversified g20 gene pool in these oceanic surface waters.

(v) Sargasso Sea DCM sample (SS47 and SS40). Two samples in the DCM layer of the Sargasso Sea were analyzed and exhibited similar phylogenetic patterns. Among the 36 and 31 clones studied in samples SS47 and SS40, 24 and 26 different g20 sequences were identified, respectively. Four common phyletic clusters (II, C, E, and B) were shared and the tree topologies within clusters II, E, and B were very similar between the two communities (Fig. 2e and f). Compared to Sargasso Sea surface sample SS48, the two DCM samples (SS47 and SS40) both contained more clusters. The dominant cluster shifted from cluster E in the surface to cluster B in the DCM layer, as was observed in the Gulf Stream. However, the two DCM samples in the Sargasso Sea were also differed from each other. For example, clone SS4716, found in the SS47 sample, did not cluster with any other clones or cyanophage isolates whereas SS40 had two more clusters than SS47.

Comprehensive results. Among 207 g20 clones randomly picked from six natural viral communities, 114 different g20 homologs were encountered. No clone had a nucleotide sequence identical to that of any of the cyanophage isolates in this study, indicating that cyanophages are very diverse in natural environments. The distribution of clones and g20 sequences within each community is summarized in Table 4. Nine distinct clusters are revealed in the comprehensive phylogenetic tree in Fig. 3 based on the amino acid sequence alignment of g20. Clusters I, II, and III comprised all of the 11 known isolates and 32 different natural clones (nucleotide sequences). Cluster I included five isolates and 15 different sequences that were exclusively from the Skidaway Estuarine sample. The pairwise sequence similarity levels in cluster I ranged from 75.2 to 98.0%. Cluster II included an isolate from the Sargasso Sea and eight different sequences from three oceanic DCM communities in the Gulf Stream and the Sargasso Sea. The sequence identity levels in this cluster ranged from 73.2 to 99.5%. Cluster III consisted of six isolates and nine different clones from the Skidaway Estuary, the Gulf Stream, and the Sargasso Sea, and there was 67.0 to 98.8% sequence similarity in this cluster.

Clusters A through F, however, contained only clones whose sequences did not match any of the known cyanophage isolates. Sequence similarity within these clusters was greater than 92.9%, except in cluster E, in which the lowest similarity level was only 79.3%. Clusters A and F contained clones exclusively from the Skidaway Estuary, while clusters B, C, and D were

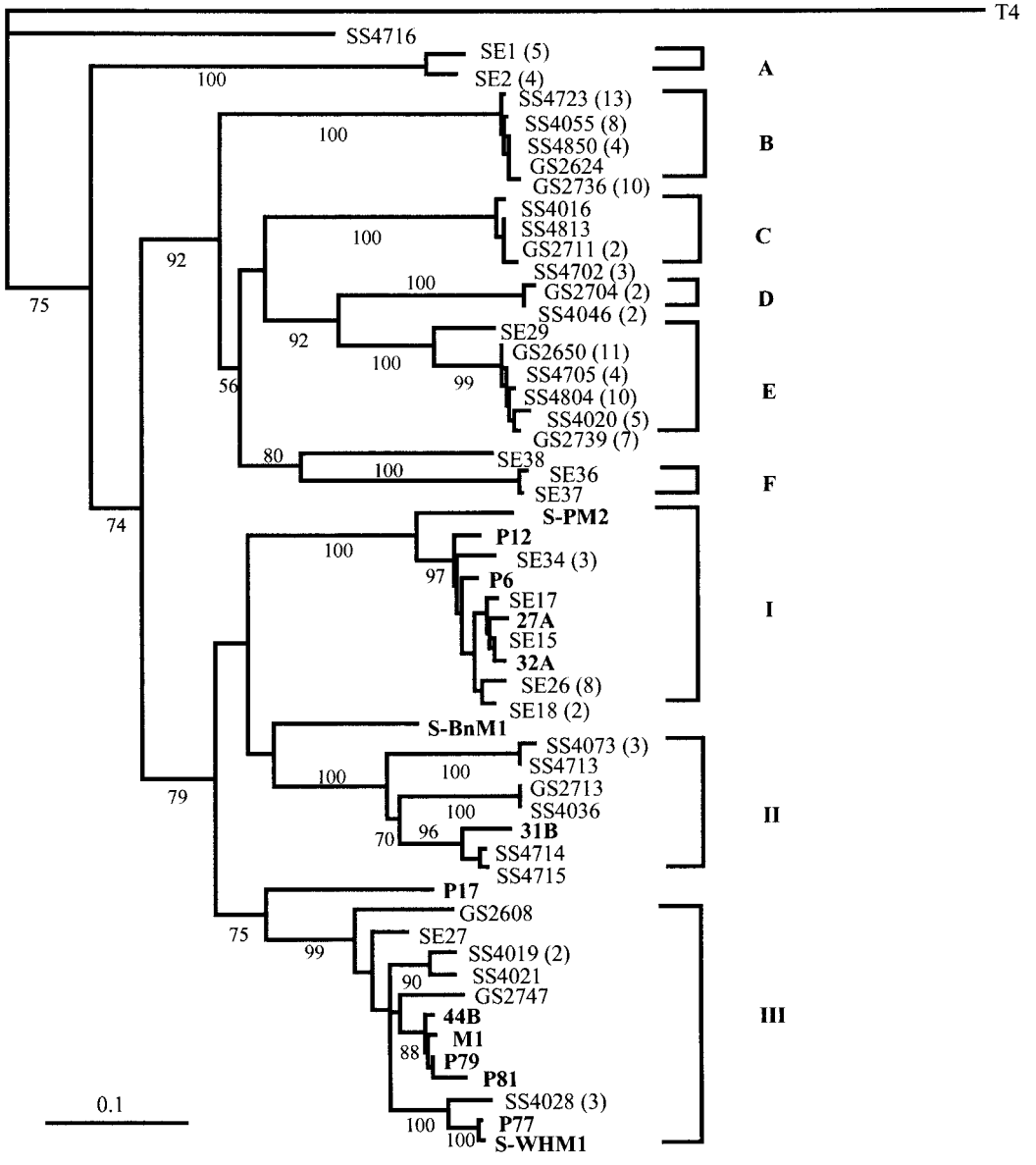

FIG. 3. Neighbor-joining tree showing the phylogenetic affiliation of cyanophage isolates (boldface type) and representative clones from all six of the natural virus communities studied. The tree was constructed on the basis of a 176-amino-acid sequence alignment with T4 as the outgroup. Each value in parentheses is the number of different nucleotide sequences in the same cluster and same community as the representative clone. Clusters A through F and I through III were assigned on the basis of phylogenetic relatedness. Bootstrap values of less than 50 were not shown. The scale bar indicates 0.1 substitution per site.

formed exclusively by the clones from the open ocean. Additionally, cluster D consisted only of the clones from the oceanic DCM communities, as did cluster II, but there was no unique cluster for surface communities in the ocean samples. Clusters B and E were the two largest clusters in the oceanic samples, while E is the only cluster that included clones from all six communities (Table 4).

DISCUSSION

Viruses are very diverse in the ocean. Since their genomes do not contain universal genetic markers, such as rRNAs, it is necessary to develop virus-specific primers in order to investigate their genetic diversity in nature. Chen et al. (3) used algal virus-specific primers for PCR cloning analysis of natural marine VCs and identified a diverse community of genetically related viruses of the family *Phycodnaviridae*. In this study, we employed the gene for cyanophage capsid assembly protein g20 to investigate the genetic diversity of cyanophage isolates and natural cyanophage communities. Although the g20 gene has been used in DGGE analyses of natural cyanophage communities (30, 31), this is the first study to reveal phylogenetic affiliations among cyanophage isolates and natural cyanophage communities based on suitable regions in the g20 gene. The primers we designed, CPS1 and CPS8, amplify a ca. 592-bp conserved region that is adequate for phylogenetic inference.

TABLE 4. Distribution of g20 gene clones and nucleotide sequences in the six virus communities studied

VC	No. of clones[a]/no. in each cluster[b]										Totals
	I	A	F	II	D	B	C	E	III	Ungrouped[c]	
SE1	18/15	12/9	2/2					1	1	1	35/29
GS26						1		34/11	2/1		37/13
SS48						6/4	1	28/10			35/15
GS27				1	4/2	12/10	7/2	8/7	1		33/23
SS47				4/3		15/13	9/3	7/4		1	36/24
SS40				5/4	2/2	10/8	1	5/5	8/6		31/26
Total	18/15	12/9	2/2	10/8	6/4	44/31	18/5	83/29	12/9	2/2	207/114[d]

[a] Number of clones in the cluster.
[b] Number of different nucleotide sequences found in the cluster.
[c] Ungrouped, single clone that did not cluster with others.
[d] Identical sequences found in more than one community were only counted once.

The target of these primers, cyanomyovirus, is the dominant form of cyanophages in marine environments (11, 27). Therefore, application of primers CPS1 and CPS8 allows us to investigate the phylogenetic diversity of the vast majority of cyanophages in natural marine habitats.

Phylogenetic analysis with isolated cyanophages indicated that marine cyanophages were genetically divergent yet more closely related to each other than to bacteriophage T4 (Fig. 1). Similar results were reported for marine microalgal viruses, which formed a unique group distant from other double-stranded DNA viruses infecting eukaryotes (2). Similarity analysis of g20 sequences further suggested that no correlation was present between genetic variation and geographic distance (Tables 1 and 3). Isolates from the same or adjacent water bodies could share a high level of g20 identity, e.g., 90.4% between 27A and 32A and 96.1% between P79 and P81. However, isolates from different water bodies could also share a high level of similarity, e.g., 96.1% between S-WHM1 and P77. On the other hand, cyanophage isolates from the same region can be as diverse as those from different oceans. For example, there is 66.2% similarity between 31B and 44B and 68.3% similarity between 31B and P17. Our results obtained with cyanophage isolates support the suggestion that genetically related marine viruses are widely distributed in the ocean without significant geographic segregation (10).

Natural cyanophage communities from both estuarine and oceanic environments exhibited even greater genetic diversity than did the cyanophage isolates we studied. Among the nine phylogenetic clusters identified from the g20 sequences, three clusters contained known cyanophage isolates while the other six clusters contained only natural clones. Even within a single cluster (e.g., cluster I, II, or III), clones demonstrated their divergent nature by the wide range of sequence similarity (73.2 to 99.5%). Furthermore, except for the two surface oceanic samples, all of the other communities contained five or six distinct genetic groups, with up to 29 different genotypes found in a single sample. The cyanophage diversity reported here is much greater than that reported from previous studies (11, 31). Wilson et al. (30) used g20 gene-based cyanomyovirus-specific primers and DGGE to investigate cyanophage diversity in the Atlantic Ocean. Only between 2 and 10 different genotypes were found by DGGE analysis in each concentrated virus community. The difference between our results and those of Wilson et al. (31) may result from the much higher sensitivity of our cloning-sequencing analysis and the degeneracy in our primers that was not included in the DGGE-oriented PCR primers. Many of the cyanophages used in this study were isolated from Georgia estuaries and exhibited diverse transmission electron microscopy morphology and restriction fragment length polymorphism patterns (11). Our results not only revealed much greater diversity among the cyanophages in marine environments but also demonstrated their phylogenetic affiliation via nucleic acid sequence analysis.

High phage diversity in natural marine environments could be due to either phage-host genetic exchange through transduction or genetic exchange between coinfecting phages. The g20 gene is within a mobile genetic module that is exchangeable between viruses during coinfection of bacterial hosts (12). The diverse cyanophage assemblages and dynamic phage-host interactions observed suggest that marine phages may be important in shaping the genetic diversity and composition of marine microbial communities.

It is likely that clusters I, II, and III represent lytic phages that infect WH7803-like *Synechococcus* spp. The wide geographical origins of isolates in clusters I and III further suggest that cyanophages in these clusters may infect *Synechococcus* spp. that are adapted to widely differing marine environments, whereas cluster II includes cyanophages whose hosts are adapted solely to oligotrophic environments. The identities of clusters A through F were enigmatic due to the absence of any isolates in them. Possibly, some clusters may represent cyanophages infecting *Synechococcus* cells that are genetically distant from the WH7803-like strains, from which most of our cyanophages were isolated. It is known that considerable genetic diversity exists within natural marine *Synechococcus* assemblages (13, 25). Isolation of more cyanophages that infect different marine *Synechococcus* spp. will shed light on the origins of the unknown clusters. We also cannot rule out the possibility that some g20 clones in clusters A to F are other bacteriophages whose hosts are closely related to *Synechococcus* spp. For example, marine *Prochlorococcus* spp. are abundant (4, 16) and genetically similar to *Synechococcus* spp. in terms of 16S rRNA (26) and RNA polymerase gene phylogeny (15). It is possible that our primers may amplify *Prochlorococcus* phages. This hypothesis remains to be tested when the phages infecting *Prochlorococcus* spp. become available.

The genetic composition and structure of natural cyanophage communities in estuaries and the open ocean were dif-

ferent from each other. While 56% of the estuarine g20 clones studied were within clusters I, II, and III, the vast majority (87% average) of the oceanic clones fell into clusters B to F. Moreover, unique phylogenetic clusters were found in each environment. For example, clusters A, F, and I were amplified from the estuarine sample whereas clusters B, C, D, and II were only present in the oceanic samples (Table 4). We also observed consistently higher cyanophage diversity in the oceanic DCM samples than in the surface samples, even though different waters were studied (i.e., the Gulf Stream versus the Sargasso Sea). Finally, it is evident that the dominant g20 clusters shifted from cluster E at the oceanic surface to cluster B at the DCM (Table 4). The distinct cyanophage population structures in estuarine water versus offshore water and in surface water versus the DCM layer suggested the presence of different host populations in response to different light, nutrient, and other physical conditions and a dynamic interaction between cyanophage and cyanobacterial populations. Little is known about the influence of environmental conditions on the genetic structure of natural *Synechococcus* assemblages. Presumably, the increasing cyanophage diversity from the surface to the DCM layer is related to the more diverse host populations at depth in stratified water columns.

The gene for viral capsid assembly protein g20 was successfully used in this study as a marker molecule to investigate genetic diversity and phylogenetic affiliation among cyanophage isolates and natural cyanophage assemblages in estuarine and oligotrophic environments. Phylogenetic analysis of cyanophage g20 gene sequences revealed that natural cyanophage populations are strikingly diverse and that their genetic structures vary greatly in different marine environments. This study suggests that marine viruses are under significant selection pressure and are able to maintain their diversity in response to host resistance. Further studies on the covariation of the genetic diversity of virus and host populations with changing environmental variables would provide new insights into the ecological roles of marine viruses.

ACKNOWLEDGMENTS

This work was supported by grants from The National Science Foundation (OCE-973060, OCE-0049098, and OCE-9977040) and the Department of Energy (DE-FG02-97ER62451).

We thank Jingrang Lu at the University of Georgia for providing the cyanophage isolates from Georgia estuaries. We also thank Ju Sheng for testing PCR primers on other bacteriophages isolated from the Baltimore Inner Harbor. Curtis Suttle at the University of British Columbia, John Waterbury at Woods Hole Oceanographic Institution, and William Wilson at the Plymouth Marine Laboratory, Plymouth, United Kingdom, kindly provided cyanophage isolates for this study.

REFERENCES

1. **Bergh, O., K. Y. Borsheim, G. Bratbak, and M. Heldal.** 1989. High abundance of viruses found in aquatic environments. Nature (London) **340:**467–468.
2. **Chen, F., and C. A. Suttle.** 1996. Evolutionary relationships among large double-stranded DNA viruses that infect microalgae and other organisms as inferred from DNA polymerase genes. Virology **219:**170–178.
3. **Chen, F., C. A. Suttle, and S. M. Short.** 1996. Genetic diversity in marine algal virus communities as revealed by sequence analysis of DNA polymerase genes. Appl. Environ. Microbiol. **62:**2869–2874.
4. **Chisholm, S. W., R. J. Olson, 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 (London) **334:**340–343.
5. **Felsenstein, J.** 1993. PHYLIP: phylogeny inference package (version 3.5). University of Washington, Seattle.
6. **Fuhrman, J. A.** 1999. Marine viruses and their biogeochemical and ecological effects. Nature (London) **399:**541–547.
7. **Fuller, N. J., W. H. Wilson, I. R. Joint, and N. H. Mann.** 1998. Occurrence of a sequence in marine cyanophages similar to that of T4 gp20 and its application to PCR-based detection and quantification techniques. Appl. Environ. Microbiol. **64:**2051–2060.
8. **Garza, D. R., and C. A. Suttle.** 1998. The effect of cyanophages on the mortality of *Synechococcus* spp. and selection for UV resistant viral communities. Microb. Ecol. **36:**281–292.
9. **Jiang, S. C., C. K. Kellogg, and J. H. Paul.** 1998. Characterization of marine temperate phage-host systems isolated from Mamala Bay, Oahu, Hawaii. Appl. Environ. Microbiol. **64:**535–542.
10. **Kellogg, C. A., J. B. Rose, S. C. Jiang, J. M. Thurmond, and J. H. Paul.** 1995. Genetic diversity of related vibriophages isolated from marine environments around Florida and Hawaii, USA. Mar. Ecol. Prog. Ser. **120:**89–98.
11. **Lu, J., F. Chen, and R. E. Hodson.** 2001. Distribution, isolation, host specificity, and diversity of cyanophages infecting marine *Synechococcus* spp. in river estuaries. Appl. Environ. Microbiol. **67:**3285–3290.
12. **Monod, C., F. Repoila, M. Kutateladze, F. Tetart, and H. M. Krisch.** 1997. The genome of the pseudo T-even bacteriophages, a diverse group that resembles T4. J. Mol. Biol. **267:**237–249.
13. **Palenik, B.** 1994. Cyanobacterial community structure as seen from RNA polymerase gene sequence analysis. Appl. Environ. Microbiol. **90:**3212–3219.
14. **Palenik, B., and H. Swift.** 1996. Cyanobacterial evolution and prochlorophyte diversity as seen in DNA-dependent RNA polymerase gene sequences. J. Phycol. **32:**638–646.
15. **Palenik, B., and R. Haselkorn.** 1992. Multiple evolutionary origins of prochlorophytes, the chlorophyll *b*-containing prokaryotes. Nature (London) **355:**265–267.
16. **Partensky, F., W. R. Hess, and D. Vaulot.** 1999. *Prochlorococcus*, a marine photosynthetic prokaryote of global significance. Microbiol. Mol. Biol. Rev. **63:**106–127.
17. **Polz, M. F., and C. M. Cavanaugh.** 1998. Bias in template-to-product ratios in multitemplate PCR. Appl. Environ. Microbiol. **64:**3724–3730.
18. **Proctor, L. M., and J. A. Fuhrman.** 1990. Viral mortality of marine bacteria and cyanobacteria. Nature (London) **343:**60–62.
19. **Suttle, C. A.** 2000. Cyanophages and their role in the ecology of cyanobacteria, p. 563–589. *In* B. A. Whitton and M. Potts (ed.), The ecology of cyanobacteria: their diversity in time and space. Kluwer Academic Publishers, Boston, Mass.
20. **Suttle, C. A., and A. M. Chan.** 1993. Marine cyanophages infecting oceanic and coastal strains of *Synechococcus*: abundance, morphology, cross-infectivity and growth characteristics. Mar. Ecol. Prog. Ser. **92:**99–109.
21. **Suttle, C. A., and A. M. Chan.** 1994. Dynamics and distribution of cyanophages and their effect on marine *Synechococcus* spp. Appl. Environ. Microbiol. **60:**3167–3174.
22. **Suttle, A. S., A. M. Chan, and M. T. Cottrell.** 1990. Infection of phytoplankton by viruses and reduction of primary productivity. Nature (London) **347:**467–469.
23. **Suttle, C. A., and F. Chen.** 1992. Mechanisms and rates of decay of marine viruses in seawater. Appl. Environ. Microbiol. **58:**3721–3729.
24. **Thingstad, T. F., M. Heldal, G. Bratbak, and I. Dundas.** 1993. Are viruses important partners in pelagic food webs? Trends Ecol. E **8:**209–212.
25. **Toledo, G., and B. Palenik.** 1997. *Synechococcus* diversity in the California Current as seen by RNA polymerase (rpoC1) gene sequences of isolated strains. Appl. Environ. Microbiol. **63:**4298–4303.
26. **Urbach, E., D. L. Robertson, and S. W. Chisholm.** 1992. Multiple evolutionary origins of prochlorophytes within the cyanobacterial radiation. Nature (London) **355:**267–270.
27. **Waterbury, J. B., and F. W. Valois.** 1993. Resistance to co-occurring phages enables marine *Synechococcus* communities to coexist with cyanophages abundant in seawater. Appl. Environ. Microbiol. **59:**3393–3399.
28. **Waterbury, J. B., S. W. Watson, F. W. Valois, and D. G. Franks.** 1986. Biological and ecological characterization of the marine unicellular cyanobacterium *Synechococcus*. Can. Bull. Fish. Aquat. Sci. **214:**71–120.
29. **Wilhelm, S. W., and C. A. Suttle.** 1999. Viruses and nutrient cycles in the sea: viruses play critical roles in the structure and function of aquatic food webs. Bioscience **49:**781–788.
30. **Wilson, W. H., J. F. Nicholas, I. R. Joint, and N. H. Mann.** 1999. Analysis of cyanophage diversity and population structure in a south-north transect of the Atlantic Ocean. Bull. Inst. Oceanogr. (Monaco) **19:**209–216.
31. **Wilson, W. H., J. F. Nicholas, I. R. Joint, and N. H. Mann.** 2000. Analysis of cyanophage diversity in the marine environment using denaturing gradient gel electrophoresis, p. 565–570. *In* C. R. Bell, M. Brylinsky, and P. Johnson-Green (ed.), Microbial biosystems: new frontier. Proceedings of the 8th International Symposium on Microbial Ecology. Halifax, Nova Scotia, Canada.
32. **Wilson, W. H., I. R. Joint, N. G. Carr, and N. H. Mann.** 1993. Isolation and molecular characterization of five marine cyanophages propagated on *Synechococcus* sp. strain WH7803. Appl. Environ. Microbiol. **59:**3736–3743.
33. **Wommack, K. E., and R. R. Colwell.** 2000. Virioplankton: viruses in aquatic ecosystems. Microbiol. Mol. Biol. Rev. **64:**69–114.