Limnol. Oceanogr., 45(8), 2000, 1697–1706
© 2000, by the American Society of Limnology and Oceanography, Inc.

Genome size distributions indicate variability and similarities among marine viral assemblages from diverse environments

Grieg F. Steward[1], *Janine L. Montiel*[2], *and Farooq Azam*
Scripps Institution of Oceanography, University of California–San Diego, La Jolla, California 92093-0202

Abstract

Pulsed field gel electrophoresis (PFGE) was used to determine the size distributions of virus-like DNA in seawater from diverse environments (Arctic Ocean, Ross Sea, coastal Pacific Ocean, and northern Adriatic Sea). Changes in DNA banding patterns indicated that shifts in the viral assemblage composition occurred on the order of $\leq$2 d during an intense dinoflagellate bloom in coastal Pacific waters. Different DNA banding patterns from diverse locations also indicated spatial variability in composition, but all of the samples analyzed had similar features. Size frequency distributions for virus-like genomes (VLGs) were multimodal, with major peaks occurring around 31–36 kilobases (kb) and 58–63 kb. The smallest discrete band resolved was 26 kb, the largest was $>$200 kb, and the overall mean VLG size was 50 $\pm$ 4 kb (mean $\pm$ SD, n = 30). On average in surface seawater, $>$90% of the VLGs occurred in the 26–69 kb size range, and at least half were between 28 and 45 kb. This first extensive survey of viral genome sizes in seawater indicates that most marine viruses have physical properties similar to other known viruses. The distributions revealed that the vast majority of the detected VLGs had sizes typical of bacteriophages, whereas only a few percent were in the size range of known algal viruses.

The abundance, distribution, and dynamics of viruses in various aquatic environments have been extensively investigated at the assemblage level (Wommack and Colwell 2000). In contrast, the diversity within natural viral assemblages and how assemblage composition changes over space and time are less well known. Studies of host range, morphology, and genetic variability using traditional culture-based and new PCR-based techniques all suggest that viral diversity in seawater is quite high (reviewed by Wommack and Colwell 2000). However, since many potential host organisms have not yet been cultivated and the range of genetic variability among aquatic viruses is still poorly known, it has been argued that many viruses of potential ecological importance could be missed by these techniques (Steward and Azam 2000).

Pulsed field gel electrophoresis (PFGE) was introduced as another useful technique for resolving different viruses in complex assemblages based simply on their genome size (Klieve and Swain 1993). Such an approach allows an overview of the viral diversity that is faster and less expensive than transmission electron microscopy (TEM) and that avoids the various biases inherent in culture- and PCR-based assays. Since the technique uses intact, unamplified DNA, it does not require a priori knowledge about the genetic diversity within a viral assemblage. It also directly provides quantitative data on a fundamental physical characteristic of the viruses comprising an assemblage; that is, the genome size. This whole-genome fingerprinting approach is thus complementary to the more specific PCR-based fingerprinting approaches and to morphological descriptions obtained by TEM.

So far there have been few reports on genome size distributions within aquatic viral assemblages using PFGE, and the performance of the method has remained relatively untested. In the few aquatic environments that have been investigated, virus-like DNA was found to span a size range from about 15 kb to $>$300 kb, with most of the DNA present in the range of 30 to 70 kb (Steward and Azam 2000; Wommack and Colwell 2000). In this study, we present some experiments designed to test whether our procedure reliably separated genomes in a mixture of known viruses and could distinguish viral from truly dissolved DNA. We then used the method to examine the viral genome size distributions in seawater from diverse locations as well as to document changes in the distribution over time at one location. Our results support, though do not prove, the assumption that the DNA we harvested and analyzed was indeed viral. The results also indicate that viruses in natural marine assemblages were predominantly within the ranges of genome sizes and buoyant densities reported for other known bacterial and algal viruses. The DNA banding patterns revealed that, despite obvious spatial and temporal variability, a few size classes of virus-like genomes are consistently most abundant in seawater from widely varying environments.

Materials and methods

Tests of sample processing and PFGE conditions—Initial tests of sample processing and separation conditions were

[1] Present address: Ocean Sciences Department, A451 Earth and Marine Science Building, University of California, 1156 High Street, Santa Cruz, California 95064 (gsteward@cats.ucsc.edu).

[2] Present address: Roche Diagnostics Corporation, Chief Technology Office, 2929 7th Street, Suite 100, Berkeley, California 94710.

Acknowledgments

We gratefully acknowledge R. Sambrotto for collecting samples during the SCICEX 96 cruise in the Arctic ocean and the support of the captains and crews of the USS *Pogy* and the USS *Archerfish* for support during SCICEX 96 and 97. We also thank S. Fonda-Umani and the scientists at the Laboratorio di Biologia Marina in Trieste, Italy, for accommodating sample collections in the northern Adriatic. Sample collection and processing was supported by NSF (OPP 96-23130, OPP 96-17045, and OPP 95-30851) grants to F.A. Additional support for G.F.S. was provided by a postdoctoral Fellowship at the Monterey Bay Aquarium Research Institute funded by the David and Lucille Packard Foundation.

Table 1. Sampling locations and dates. Letter codes correspond to sample lane designations for the gels in Figs. 5–7.

Region (Cruise or specific site)	Code	Latitude	Longitude	Depth (m)	Date
Adriatic (Gulf of Trieste)	A	45°43′N	013°34′E	1	1 Apr 97
	B	45°43′N	013°34′E	1	9 Apr 97
Antarctic (Ross Sea)	A	76°30′S	169°00′E	4	21 Jan 97
	B	76°30′S	169°00′E	550	21 Jan 97
Arctic (SCICEX96)	A	88°06′N	044°00′W	5	20 Sep 96
	B	79°54′N	141°32′W	5	30 Sep 96
	C	74°17′N	145°33′W	5	4 Oct 96
	D	71°08′N	147°05′W	5	7 Oct 96
(SCICEX97)	E	84°15′N	026°16′E	55	3 Sep 97
	F	85°02′N	144°58′W	131	5 Sep 97
	G	81°00′N	148°08′W	119	9 Sep 97
	H	78°60′N	148°25′W	55	9 Sep 97
	I	70°54′N	141°50′W	55	12 Sep 97
	J	75°31′N	179°35′W	55	16 Sep 97
	K	80°27′N	148°57′E	55	20 Sep 97
	L	78°55′N	127°58′E	55	21 Sep 97
	M	85°17′N	151°54′E	55	24 Sep 79
	N	84°08′N	177°77′E	55	24 Sep 97
	O	80°56′N	175°51′W	55	25 Sep 97
	P	77°38′N	158°42′W	55	27 Sep 97
	Q	75°03′N	166°48′W	55	2 Oct 97
Pacific Coast (Mission Bay)	A	32°47′N	117°13′W	1	26 Jul 96
(Scripps Pier)	B	32°53′N	117°15′W	1	5 May 97
		32°53′N	117°15′W	1	7 May 97
		32°53′N	117°15′W	1	9 May 97
		32°53′N	117°15′W	1	13 May 97
		32°53′N	117°15′W	1	30 May 97
		32°53′N	117°15′W	1	9 Jun 97
		32°53′N	117°15′W	1	25 Jun 97
		32°53′N	117°15′W	1	27 Oct 97

performed with bacteriophage isolates. Isolates used in the tests were bacteriophage lambda and three bacteriophages isolated by one of the authors (G.F.S.) from coastal seawater in La Jolla, California: (1) Roseophage SIO1, which infects the Roseobacter isolate SIO1 (Rohwer et al. 2000); (2) a phage designated Ø21/50, and (3) a phage designated Ø21/43. According to 16S rRNA gene sequences, the hosts for the latter two phages are most closely related to members of the Gram Positive and Flexibacter–Cytophaga–Bacterioides bacterial groups, respectively (R. Long pers. comm.).

To test the relative recoveries of viral vs. dissolved DNA by the ultracentrifugation procedure, a seed stock was prepared containing all three intact viral isolates. Half of the seed stock was heated to 65°C for 30 min to release viral DNA from the capsids. Four aliquots (60 μl) each from the heated and unheated portions were diluted into 8- by 12-ml aliquots of 0.2-μm filtered seawater (collected from Scripps Pier). Two of the heated and two of the unheated stock dilutions were then treated with nucleases (deoxyribonuclease I and ribonuclease A [DNase I and RNase A], each at 1 U ml^{-1} final concentration). The samples receiving no nucleases were immediately centrifuged in an SW41 rotor (Beckman) at 40,000 rpm (274,000 $\times$ g) for 2 hr at 20°C. Supernatants were carefully decanted, and residual liquid on the walls of the tube were removed with a sterile cotton swab. Pellets were resuspended in 60 μl of 10$\times$ TEGED (100 mM Tris, 10 mM EGTA [egtazic acid], 10 mM EDTA [edetic acid]) while warming to 60°C for 0.5 h with occasional gentle agitation. Nuclease-treated samples were incubated for about 5 h at room temperature (~20–23°C) then processed by the same procedure. Equal amounts of each resuspended pellet were run by PFGE using the conditions described below. Duplicate aliquots of the undiluted seed stock representing 100% recovery were run as controls.

Collection and concentration of viruses from seawater—Names and coordinates of sample locations and sampling dates are summarized in Table 1. Coastal Pacific Ocean samples were collected using a hand-operated diaphragm pump or a weighted polycarbonate flask on a line from a public dock in Mission Bay (San Diego, California) and from the Pier at Scripps Institution of Oceanography (La Jolla, California). The Scripps Pier samples were collected between 1300 and 1500 h on eight occasions from 5 May to 30 October 1997. Samples from the northern Adriatic Sea were collected with Niskin bottles from a small fishing boat. Samples from the Ross Sea were collected in January 1997 with Niskin bottles during the Process 2 cruise of the United

States Joint Global Ocean Flux Study in the Southern Ocean. The above samples (12 ml to 10 liters) were pressure filtered (≤10 psi) through 0.2-μm Gelman Acrodisc syringe tip filters (12-ml volumes) or Gelman Culture Capsule filters (volumes ≥1 liter). Arctic Ocean samples were collected during two cruises aboard U.S. Navy nuclear submarines as part of the SCICEX program administered by the United States National Science Foundation Office of Polar Programs in 1996 and 1997. SCICEX water samples were collected via a through-the-hull sampling system, which tapped into one of the ships' seawater intake lines as close as possible to the hull. The sampling line consisted of teflon tubing reinforced with an outer stainless steel braided sheath. Flow was controlled by a stainless steel needle valve. Sampling depth was determined by the ship's position in the water column. A pleated capsule filter (0.2-μm pore size, Gelman) was attached to the sampling line, and 50 liters of filtrate was collected in a polypropylene carboy for immediate processing.

Except for samples from the Arctic Ocean, viruses in the 0.2-μm filtrates were concentrated to 75–100 ml using a 100,000 nominal molecular weight cutoff (MWCO) membrane (200 cm^2) in a Benchmark® vortex flow filtration system (Membrex). Aliquots of each concentrate were treated with a mixture of DNase I and RNase A, 1 U ml^{-1} each final concentration) for 0.5 h at 20°C to degrade any dissolved nucleic acids. Aliquots (4 or 12 ml) of each nuclease-treated sample were transferred to polyallomer centrifuge tubes and centrifuged at 20°C in an SW60 or SW41 rotor (Beckman). Centrifugation conditions were chosen that would pellet particles having a sedimentation coefficient of ≥70 S with 100% efficiency. Conditions were 45,000 rpm (273,000 × g) for 1.6 h in the SW60 Ti rotor or as described above for the SW41 rotor. Pellets were resuspended and viral DNA was released by the addition of 30–60 μl of 5× TBE (0.45 M Tris base, 0.44 M boric acid, 0.01 M EDTA) or 10× TEGED and warming to 60°C for 0.5 h with occasional gentle agitation. Samples were pipetted as infrequently and as slowly as possible to avoid shearing the high–molecular weight nucleic acids liberated from the viruses. Any samples not used right away were stored at 4°C.

Arctic Ocean samples were concentrated to approximately 200 ml on 100,000 (SCICEX 96) or 30,000 (SCICEX 97) MWCO spiral-wound cartridges using a ProFlux M12 ultrafiltration apparatus (Amicon). Concentrates were preserved with sodium azide and stored at 4°C. Because sodium azide was not approved for use on the first cruise (SCICEX 96), concentrates were initially preserved by addition of NaCl (1 M final), with azide added postcruise. To remove any residual bacterial contamination, 100 ml of each viral concentrate was filtered through 0.2-μm syringe-tip filters (Gelman Acrodisc), with the filter changed after each 20–40 ml to minimize clogging. Filtrates were then further concentrated using Centricon Plus-20 centrifugal ultrafiltration units (Millipore). The units have a reservoir capacity of 20 ml and so were centrifuged then refilled five times to obtain the concentrate from an entire 100-ml sample in one unit. Centrifugation was at 2,000 rpm and 10°C in a refrigerated centrifuge (Beckman GPR) with a swinging bucket rotor. Concentration times for each round ranged from about 10 min for the first 20 ml to approximately 30 min for the last. Final concentrate volumes ranged from 100 to 175 μl. Aliquots (75 μl) of each final concentrate were transferred to polyallomer microcentrifuge tubes, and viruses were pelleted by centrifuging at 40,000 × g, for 20 min at 5°C using a TLA rotor in a benchtop ultracentrifuge (Beckman). Supernatants were aspirated and 25 μl of 10× TEGED was added to each pellet. Pellets were solubilized and DNA was released from the viruses by heating to 65°C for 5 min, followed by cooling to room temperature (~20–23°C).

Pulsed field gel electrophoresis—DNA concentration of each sample was measured by fluorometry using PicoGreen (Molecular Probes). An aliquot containing 60 ng of DNA from each resuspended pellet or of virus mixture was diluted to 45 μl with sterile H_2O, then 5 μl of 10× loading buffer (25% ficoll, 0.25% Bromophenol blue) was added. Molecular weight markers run on the gels included a lambda DNA ladder, a 5-kb ladder (both from Sigma), and an 8–48 kb ladder (BioRad). Agarose gels (1%) were run in 0.5× TBE at 14°C, 6 V cm^{-1} and a 120° included angle for 16 hr with a 1–6-s switch time. To visualize the nucleic acids, gels were stained with SYBR Green I (Molecular Probes), illuminated on a UV transilluminator (330 nm peak wavelength), and photographed with Polaroid 667 (positive) or 55 (positive/negative) film.

Image and data analysis—Digital images were obtained from gel photographs using a flatbed reflectance scanner (DuoScan, AGFA). Analysis of digital images was performed using RFLPScan software (Scanalytics) to calculate molecular weight and relative intensity of bands. The molecular weight of each integrated band was calculated by the software by comparing its migration position relative to DNA size standards. In some cases, individual bands were detected and quantified using Gaussian integration. Since resolution of the final images did not always permit clear discrimination of individual bands, the distribution of genome sizes in natural samples was also calculated from a semicontinuous integration of fluorescence intensity down the lane. This was achieved by using dropline integration with the smoothing factor set to zero and the baseline set to the minimum fluorescence value in the lane. The fluorescence threshold for detection (and inclusion in the integration) was made equal for all lanes on a gel and was adjusted to be just below the faintest band detectable by eye. The maximum number of bands detected was set between 50 to 90 to achieve high resolution coverage over the integrated areas. The width of each individual integration band was typically 1–3% of the molecular weight at that position in the gel. For example, integration was performed in roughly 0.5-kb steps in the 50-kb region of the gel.

For each band, the integrated fluorescence was divided by the molecular weight to obtain a proxy for genome copy number in arbitrary, but proportionally correct, units. The genome copy number in a given band was then expressed as percentage of the total genomes detected within the sample. We assumed that the relationship between DNA mass and fluorescent signal is linear, but some deviation from linearity can occur at higher DNA concentrations (Schneeberger et al. 1995; Steward in press). Thus, the genome copy

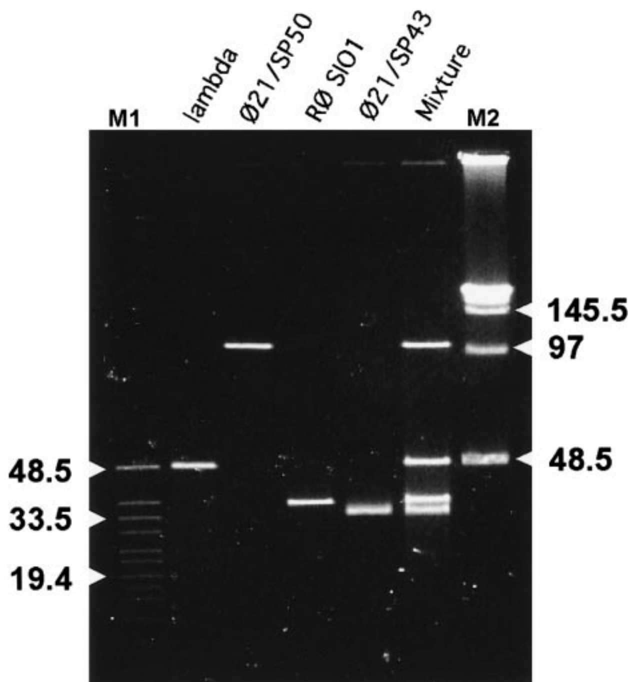

Fig. 1. Separation by PFGE of the genomes of four different bacterial viruses run separately and as a mixture. Marker lanes contain an 8–48-kb ladder (M1) or a lambda ladder (M2). Sizes (kb) of selected marker bands are indicated on the left and right.

number in the most intense bands may be somewhat underestimated. Histograms for individual samples were prepared by summing percentages in 3-kb size bins (bands $\leq$ 75 kb) or 25-kb size bins (bands $>$ 75 kb). To create an average histogram from multiple individual histograms, a mean percentage for each bin was calculated by averaging the individual percentages within a bin across all samples.

Results

We confirmed that PFGE was able to discriminate viruses based on differences in genome size using a simple artificial assemblage. Each virus produced a single sharp band on the gel at a characteristic position, whether run separately or in a mixture of all four viruses (Fig. 1).

To test whether our sample processing procedure selected for intact viruses rather than dissolved DNA, we compared the recovery of viral DNA from the three marine isolates under different conditions (Fig. 2). Viral DNA was recovered from pellets when intact viruses were subjected to ultracentrifugation either with or without DNase added to the sample, although recovery was not 100% in either case (range 20–60%). Heating of the viruses to 65°C for 30 min prior to centrifugation resulted in little to no recovery of viral DNA in the pellet. DNase treatment after heating the viruses and prior to centrifugation resulted in no detectable recovery of viruses in the pellet.

We tested the effect of ultrafiltration on banding pattern using a seawater sample from Mission Bay, San Diego. We were able to detect a banding pattern by direct pelleting of viruses from as little as 11.5 ml of 0.2-μm filtered seawater (Fig. 3). Essentially the same pattern was obtained when viruses were first concentrated by ultrafiltration prior to pelleting. The only detectable differences were band distortion and some smearing due to overloading the gel with DNA. We observed no differences in banding pattern between samples incubated with or without DNase and RNase prior to pelleting the viruses (Fig. 3).

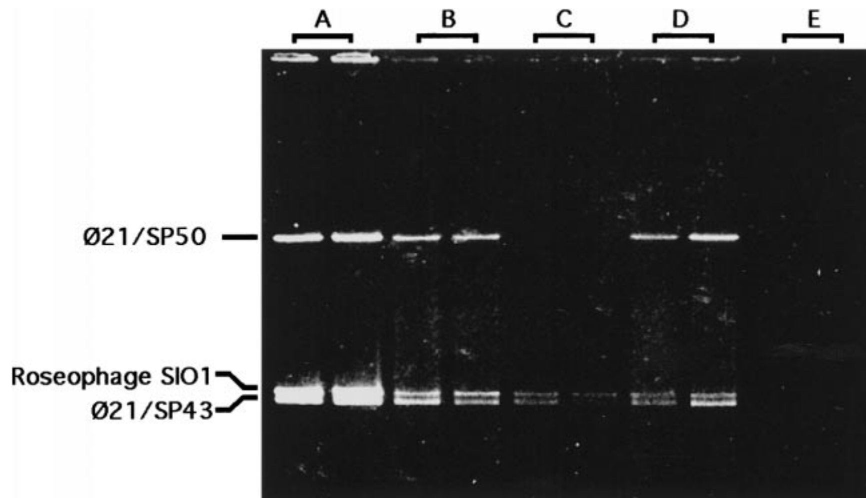

Fig. 2. Recovery test using three marine bacteriophages. Each treatment was performed in duplicate. (A) Control lanes in which the virus mixture was diluted in TBE and loaded directly on the gel. (B) Viruses were diluted into seawater and pelleted by ultracentrifugation, and the DNA was recovered in 5× TBE or 10× TEGED with heating to 60°C for 30 min. (C) Same as in B, but viruses were heated to release viral DNA prior to dilution in seawater. (D) Same as in B, but viruses were treated with DNase I after dilution in seawater and prior to ultracentrifugation. (E) Same as in D, but viruses were heated to 65°C to release viral DNA prior to dilution in seawater and treatment with DNase I.

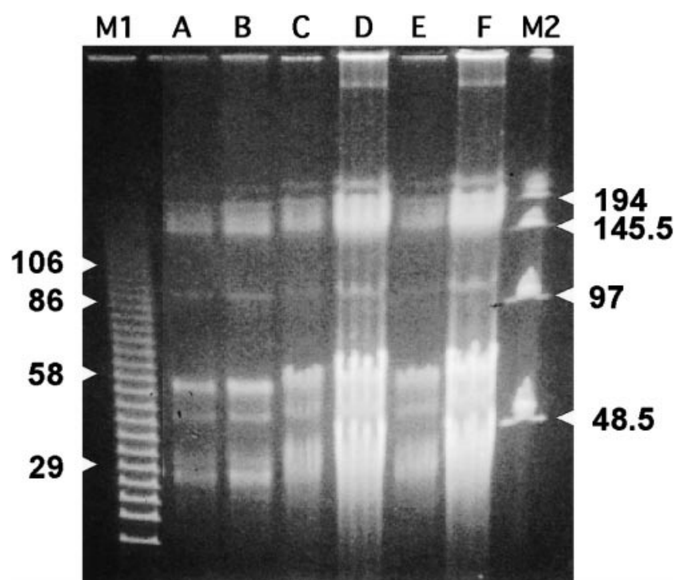

Fig. 3. Banding patterns for virus-like DNA in 0.2-μm filtered Mission Bay seawater. DNA was obtained by pelleting viruses from 11.5 ml (lane A) or 23 ml (lane B) of unconcentrated seawater or by pelleting 0.5 ml (lanes C, E) or 4.5 ml (lanes D, F) of a viral concentrate prepared by tangential flow ultrafiltration. Samples were treated with nucleases prior to pelleting (lanes A–D) or left untreated (lanes E, F). Marker lanes contain a 5-kb ladder (M1) or a lambda ladder (M2). Sizes (kb) of selected marker bands are indicated on the left and right.

To further test whether the recovered DNA had virus-like characteristics, we fractionated a viral concentrate from coastal seawater by equilibrium density gradient centrifugation in CsCl (Fig. 4). Most of the DNA in the viral concentrate was in the fractions with densities between 1.46 and 1.39, with distinct banding patterns occurring in the different fractions. No DNA was detected in fractions with densities of dissolved, naked DNA (i.e., around 1.6).

Among natural seawater samples, we observed different banding patterns of virus-like genomes (VLGs) from diverse environments (Fig. 5). Samples with the greatest number of distinct bands and the largest detectable VLGs were from the temperate coastal waters of the Pacific Ocean and the Adriatic Sea. Deeper samples produced less distinct banding patterns, with the two major peaks shifted to somewhat higher sizes relative to the major peaks observed in surface waters. Samples from across the Arctic Ocean basin showed only minor variability in banding pattern (Fig. 6). The most notable difference in banding pattern among the Arctic Ocean samples is the presence of a 45-kb band in all four 1996 samples that is not detected in any of the 13 samples from 1997. Because the sample from 1996 and 1997 were collected from different depths, it is not known whether the differences resulted from spatial or temporal variability.

Changes in the VLG size distribution are also evident among samples collected over a 6-month period in coastal Pacific waters (Fig. 7). The number of discrete bands resolvable by image analysis varied from 16 (on 30 May 1998) to 8 (on 30 October 1998).

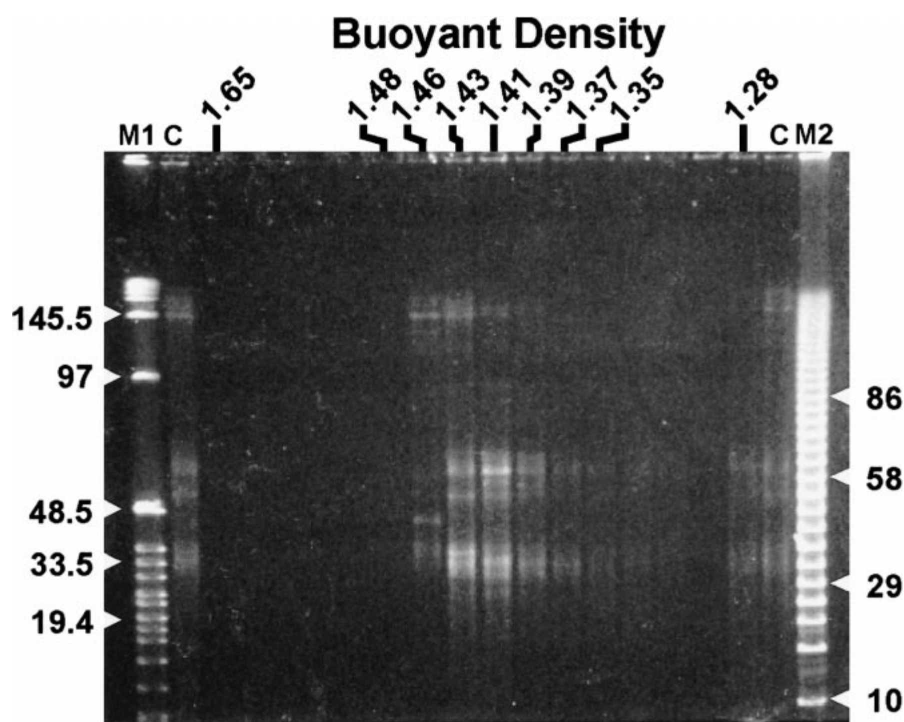

Fig. 4. PFGE banding pattern for a viral concentrate following fractionation of the intact viruses on a CsCl equilibrium density gradient. The densities of selected fractions are indicated across the top. Unfractionated material was also run for comparison (lanes C). Marker lanes contain a mixture of an 8–48-kb ladder and a lambda ladder (M1) or a 5-kb ladder (M2). Sizes (kb) of selected marker bands are indicated on the left and right.

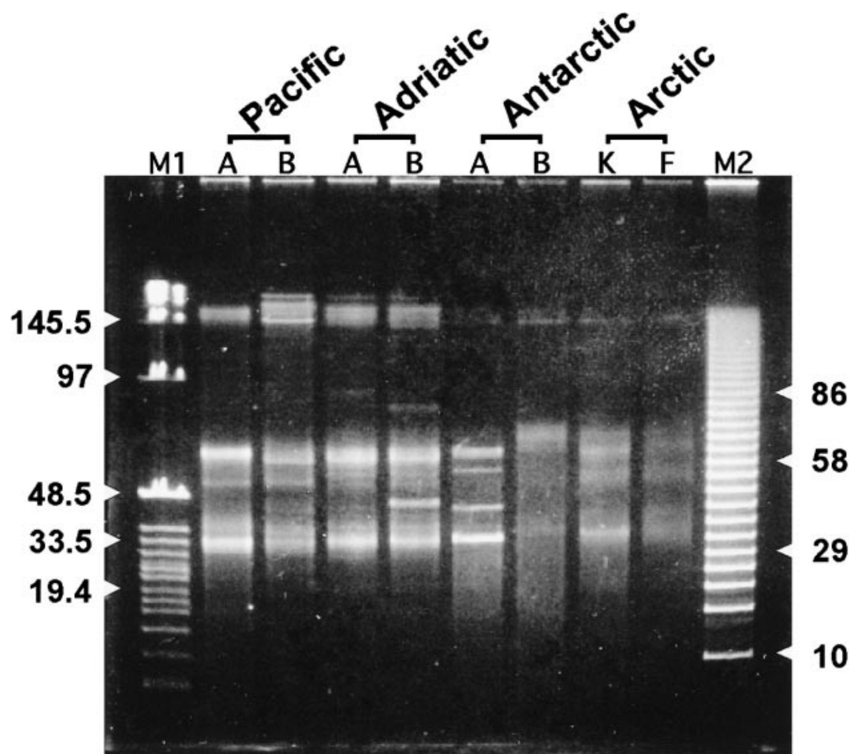

Fig. 5. PFGE banding patterns for virus concentrates obtained from different marine environments as follows: coastal Pacific waters of Mission Bay and Scripps Pier (Pacific A and B, respectively), a station in the northern Adriatic Sea sampled on 1 and 9 April 1997 (Adriatic A and B, respectively), a station in the Ross Sea, Antarctica sampled at 4- and 550-m depths (Antarctic A and B, respectively), and two stations in the Arctic Ocean collected at 55 and 131 m (Arctic K and F, respectively). Depths, dates, and exact locations for each sample are presented in Table 1. Marker lanes contain a mixture of an 8–48-kb ladder and a lambda ladder (M1) or a 5-kb ladder (M2). Sizes (kb) of selected marker bands are indicated on the left and right.

The smallest DNA size resolvable as a discrete band was about 26 kb, and the maximum discrete band size detected was estimated to be roughly 300 kb. Migration distance (molecular weight) and staining intensity (mass) were used to estimate the frequency distribution of putative viral genomes in different size classes for each of the surface water samples. The mean viral genome sizes calculated from the distributions ranged between 37 and 59, with an overall average of 50 $\pm$ 4 kb (mean $\pm$ SD, $n = 30$).

For regions where more than one sample was analyzed (Arctic Ocean, Scripps Pier, and Adriatic Sea), all surface water genome size histograms in the region were averaged as described (see *Image and data analysis*) to create a single representative size distribution for that region. These regional average histograms were then averaged with the remaining single distributions for Mission Bay and the Ross Sea to create one grand average distribution for surface seawater (Fig. 8). Two major peaks occurring at 31–33 kb and 61–63 kb comprised 15% and 9% of the total number of genomes, respectively. The shoulder to the right of the 31–33-kb size class indicates a poorly resolved third peak in or near the 37–39-kb size class, which contains approximately 8% of the total VLGs. A fourth peak ($\sim$6% of the VLGs) was resolved in the 52–54-kb size class. Among the higher molecular weight bins (>75 kb), a peak was observed in the 151–200-kb size class, which contained $\sim$2% of the total VLGs. In the grand average distribution, 93% of all VLGs were in the size range from 25 to 69 kb. Approximately half (51%) of the VLGs were between 28 and 45 kb.

Discussion

Methodological considerations—Our tests with viral isolates demonstrated the potential for genome size to serve as a discriminating characteristic in mixtures of viruses. One of the uncertainties inherent in the analysis of natural viral assemblages by PFGE is whether the DNA analyzed is actually comprised only of viral genomes or contains significant amounts of DNA from other sources. Previous studies have suggested that viruses are a minor component of total dissolved DNA in seawater (e.g., Jiang and Paul 1995). However, the procedures we have used for viral DNA isolation and analysis by PFGE are likely to be highly selective for viral DNA relative to that from other sources. Our tests with isolates showed that DNA in intact viruses was resistant to DNase and was pelleted using our ultracentrifugation conditions. In contrast, dissolved DNA in the virus size range (i.e., that which we released from capsids by heating) was inefficiently pelleted under those conditions. Treatment with DNase prior to centrifuging completely removed the residual

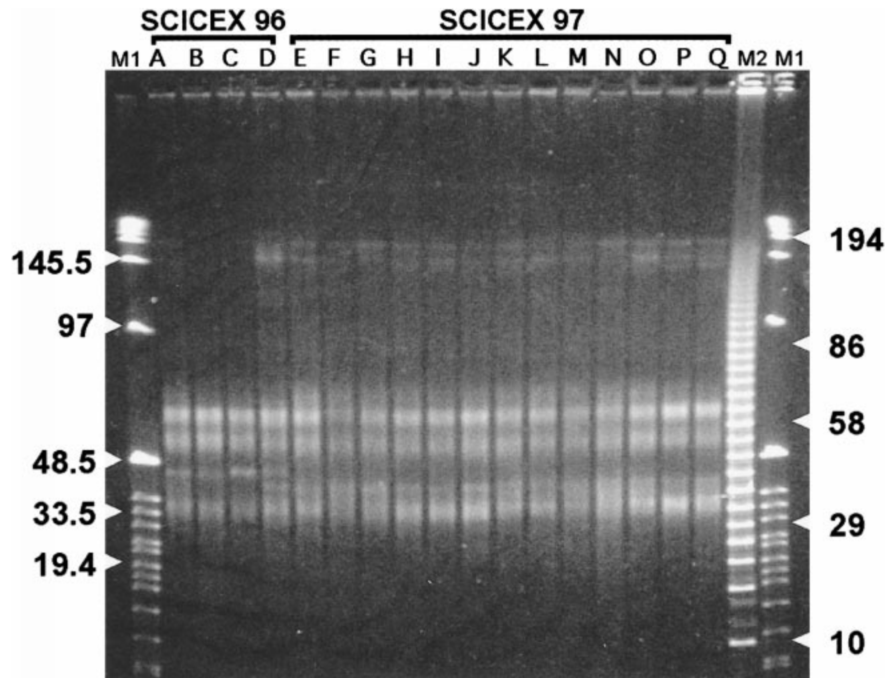

Fig. 6. PFGE banding patterns for samples collected across the Arctic Ocean. Sample from SCICEX 96 were collected at 5 m depth. Samples from SCICEX 97 were collected at a depth of 58 m, except F (131 m) and G (119 m). Station locations are presented in Table 1. Marker lanes contain a mixture of an 8–48-kb ladder and a lambda ladder (M1) or a 5-kb ladder (M2). Sizes (kb) of selected marker bands are indicated on the left and right.

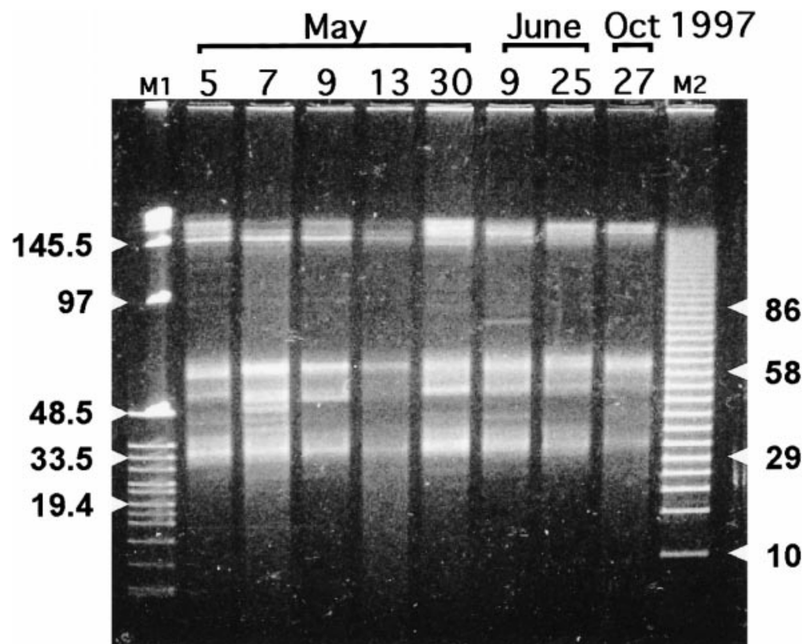

Fig. 7. PFGE banding patterns for viral concentrates collected during (5–9 May) and after a major dinoflagellate bloom in coastal waters of San Diego, California. Surface samples were collected between 1300 and 1500 h PST on each of the dates indicated across the top. Marker lanes contain a mixture of an 8–48-kb ladder and a lambda ladder (M1) or a 5-kb ladder (M2). Sizes (kb) of selected marker bands are indicated on the left and right.

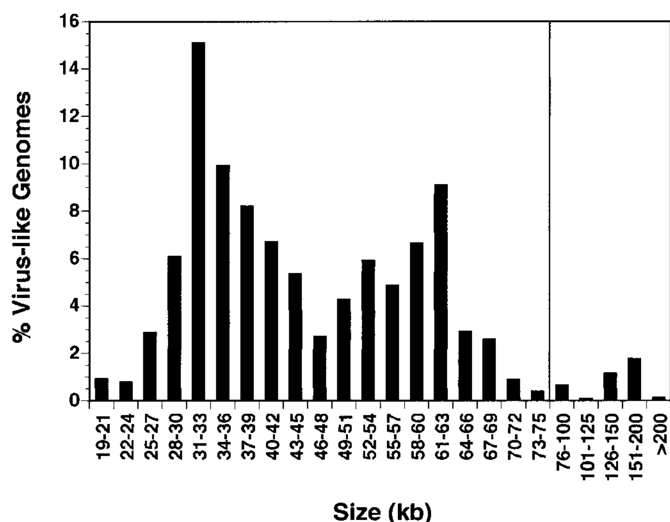

Fig. 8. Composite histogram showing the average estimated viral genome size distribution for five marine environments: Mission Bay, San Diego, coastal Pacific Ocean, northern Adriatic Sea, the Arctic Ocean, and the Ross Sea. Bin sizes to the left of the vertical line are 3 kb and to the right are 25 kb.

dissolved DNA contamination in the pellet. The indistinguishable banding patterns observed for a natural viral concentrate treated with or without DNase suggest that truly dissolved DNA was, at most, only a minor component of the operationally defined viral DNA. In general, most DNA appeared in more or less discrete bands or zones within the range of known viral genome sizes. In addition, for several samples tested by CsCl gradient fractionation (Fig. 4 and unpubl. obs.), the vast majority of DNA banded within a narrow range of buoyant densities typical for known intact bacteriophages (Table 2), rather than at the density of dissolved DNA which is around 1.6.

Despite the above circumstantial evidence, we cannot completely rule out the possibility that some nonviral DNA in seawater falls within our operational definition. For example, high molecular weight DNA from lysed organisms may be associated with lipids or proteins which could both protect it from degradation and alter its buoyant density. However, DNA from such sources would only appear as contamination if it were partially degraded (otherwise it would not band in the viral genome size range). Such partially degraded DNA is unlikely to occur in discrete sizes that would be recognizable as bands by PFGE. The distinctive banding patterns we observed for different buoyant density fractions also indicate that a given band represents a relatively precise stoichiometry of DNA and other macromolecules, as would be expected for viral particles, but is less likely for more haphazard complexes. It is possible that some nonviral, complexed form of DNA contributes to the background smearing that can be observed to some small degree in most samples. As noted above, however, that smearing could also be caused by mechanical or enzymatic degradation of viral DNA following release from capsids.

An issue not addressed by the present work is the possibility of selective recovery. It has been noted that filtration through 0.2-μm membranes can result in variable losses of viruses (Paul et al. 1991; Steward et al. 1992). Total recovery is not necessarily an issue, since equal amounts of DNA are loaded per lane and the fingerprints are interpreted only as relative intensities within a sample. However, selective losses of viruses could artificially alter the fingerprint. At present, there are no reports indicating whether filtration losses are biased toward some viral groups. Some very large virus-like particles (>0.2 μm) have been observed in seawater

Table 2. Comparison of genome sizes and buoyant densities of double-stranded DNA-containing viral isolates with data from natural assemblages by PFGE. Values of *n* refer to number of individual viral strains in the case of isolates and to number of samples analyzed in the case of natural assemblages. In cases where genome size was reported in Daltons, the size was calculated by dividing by the molecular weight of a base pair (660 Da).

	Genome size (Da)		Buoyant density		
	Mean (range)	*n*	Mean (range)	*n*	Reference
Nonmarine phage isolates	48 (6.1–735)	116	1.48 (1.28–1.54)	116	1, 2
Marine phage isolates	53 (25–95)	23			3, 4
Siphoviridae	35 (25–41)	7			
Podoviridae	42 (36–48)	5			
Myoviridae	70 (33–95)	11			
Marine cyanophage isolates	91 (85–100)	5			5
Marine alga viruses	217 (110–350)	3			6, 7
Marine assemblages	(12–314)	23			8
	50 (26–350)	33	(1.39–1.46)	1	this study

1. Ackermann and DuBow 1987*a,b*.
2. GenBank database (www.ncbi.nlm.nih.gov).
3. Wichels et al. 1998.
4. Rohwer et al. 2000.
5. Wilson et al. 1993.
6. VanEtten et al. 1991.
7. Cottrell and Suttle 1991.
8. Wommack et al. 1999.

(Bratbak et al. 1992; Cochlan et al. 1993; Gowing 1993), and these would almost certainly be removed by filtration. However, they appear to be relatively rare, with most viruses having diameters of $\leq$100 nm (Børsheim 1993; Cochlan et al. 1993). One preliminary experiment has indicated that 0.2-μm filtration did not grossly alter the distribution of virus-like DNA in the sample tested (G. Steward unpubl. obs.). Further investigations are needed to determine the significance of this potential source of bias.

In previous studies, viruses have been embedded in agarose, then DNA was released from the viral capsids using SDS and proteinase K (Klieve and Swain 1993; Wommack et al. 1999). This procedure has been employed to prevent shearing of very high molecular weight DNA and is essential when analyzing chromosomal DNA of other organisms. We had good results, however, using a simpler procedure, in which DNA was released into EDTA-containing solutions by heating. In this procedure, one must take care that shear forces are minimized by avoiding vortex mixing or excessive pipetting. Clear banding patterns were usually obtained from viral DNA in solution that was pipetted slowly and no more than a few times. An advantage of this approach is greater band resolution, since band sharpness of embedded samples is limited by the thickness of the gel plug. On occasion, some samples did show evidence of degradation (e.g., Fig. 7 sample from May 13th), perhaps because of inadequate inactivation of DNase or because of mechanical shearing. This degradation may be avoided by using centrifugal ultrafiltration as the final virus concentration step rather than ultracentrifugation (Steward in press).

Variability in viral assemblages—Our PFGE data illustrate spatial and temporal variations in natural viral assemblages. Differences in banding patterns were detectable in samples from different locations, and the appearance and disappearance of distinct bands was evident on a scale of days to months. Since viruses are obligate parasites with some degree of host specificity (Dimmock and Primrose 1994), we expected that changes in the viral assemblage would be accompanied by shifts in the microbial community (e.g., Hennes et al. 1995). This is what we observed in the time series at Scripps Pier. The PFGE results reported here indicate shifts in the viral assemblage, and a parallel investigation (Fandino et al. 1998) using denaturing gradient gel electrophoresis (DGGE) revealed shifts in the bacterial community composition over the same time period. Without additional data, it would be difficult to use the PFGE and DGGE fingerprints to identify specific virus–host relationships. Nevertheless, the fingerprints are useful for revealing significant changes in certain members of the community, which can then be targeted for identifications and more detailed study.

Viral genome sizes in seawater—Perhaps even more striking than the variations among samples was the similarity in the general pattern of VLG size distributions across a wide range of seawater habitats. The time series study at Scripps Pier shows that the major features of the size distributions are also quite stable over time despite changes in the specific viral and bacterial community fingerprints. The similarity in the average viral genome size among diverse samples has some practical utility. In many previous studies, an average DNA content of marine viruses was assumed for a variety of reasons. These included calculating the potential influence of viral DNA synthesis on measurements of bacterial productivity (Børsheim et al. 1990; Heldal and Bratbak 1991), estimating the contribution of viruses to the dissolved DNA pool (Heldal and Bratbak 1991; Paul et al. 1991; Jiang and Paul 1995; Weinbauer et al. 1995), or interpreting the kinetics of viral productivity measurements in seawater (Steward et al. 1992; Wikner et al. 1993). Of necessity, these values were somewhat arbitrarily derived from reports of genome sizes for known bacteriophage isolates and ranged from 48 to 99 attograms DNA phage^{-1}. This translates into genome sizes of 44 to 90 kb, assuming nonsegmented double-stranded DNA genomes. Our empirical estimate of 50 $\pm$ 4 kb suggests that the assumed values were fairly accurate overall but may have been high by a factor of 1.8 in some cases. The consistency of the empirical average indicates that a mean DNA content of roughly 55 attograms virus^{-1} could be reasonably applied to total viral assemblages from a wide range of seawater environments.

We speculate that the prominent clusters of different-sized viral genomes may represent different families or morphological types of viruses. In a survey of bacterial and algal virus isolates, trends are evident in the mean genome sizes for different types of double-stranded DNA viruses (Table 2). Thus, the peaks around 31–33 and 34–39 kb may be enriched in DNA from members of the Sipho- and Podoviridae families, respectively, while the 61–63-kb peak may be comprised primarily of those in the family Myoviridae. Few marine cyanophages have been examined so far, but if those few are representative, then the very minor peak at 76–100 kb is likely to contain most of this virus type. Viruses of eukaryotic algae (Family Phycodnaviridae) tend to have large genomes (Reisser 1993). From the few marine algal viruses studied to date we expect them to be present in and around the peak observed at 126 to 200 kb. The relatively small percentage of these larger viral genomes in seawater tends to support previous arguments that most free viruses in seawater are produced by lysis of bacteria rather than phytoplankton (Murray and Jackson 1992; Cochlan et al. 1993). However, it should be kept in mind that the abundance of the larger genomes could be somewhat underestimated due to selective loss during filtration as discussed above.

We also stress that genome size alone is not necessarily a reliable indicator of virus morphology or host type. The range of genome sizes within any one group is quite large, and there is significant overlap among groups. In addition, the diversity within natural viral assemblages is still largely unexplored. Bacteriophage genomes are known to span at least two orders of magnitude (Ackermann and DuBow 1987*b*), with the largest being over 700 kb (*Bacillus subtilis* phage G). The very large genomes are therefore not necessarily unique to the algal viruses, and conversely, eukaryotic algal viruses having small genomes may yet be discovered. Further work is needed to determine whether genome size can be used to infer other virion properties in natural assemblages. If such relationships exist this would greatly enhance

the information content and improve the interpretation of viral genomic fingerprints.

These results demonstrate the utility of PFGE as a fingerprinting technique for studying the diversity and dynamics of complex viral assemblages and provide new insights into the physical characteristics and diversity within marine viral assemblages. We have shown that the ranges of genome sizes and buoyant densities of operationally defined marine viruses fall well within the ranges reported for other known viruses. We also discovered that virus-like genomes show similar size distributions in widely varying marine environments, with the DNA clustering in the same few size classes. Despite these general similarities, we revealed differences in specific banding patterns among samples, indicating spatial and temporal variability in marine viral assemblages, with some significant shifts occurring within a period of 2 d or less.

References

ACKERMANN, H.-W., AND M. S. DUBOW. 1987*a.* Viruses of prokaryotes, v. 1: General properties of bacteriophages. CRC Press.

———, AND M. S. DUBOW. 1987*b.* Viruses of prokaryotes, v. 2: Natural groups of bacteriophages. CRC Press.

BØRSHEIM, K. 1993. Native marine bacteriophages. FEMS Microbiol. Ecol. **102:** 141–159.

———, G. BRATBAK, AND M. HELDAL. 1990. Enumeration and biomass estimation of planktonic bacteria and viruses by transmission electron microscopy. Appl. Environ. Microbiol. **56:** 352–356.

BRATBAK, G., O. H. HASLUND, M. HELDAL, A. NŒSS, AND RØEGGEN. 1992. Giant marine viruses? Mar. Ecol. Prog. Ser. **85:** 201–202.

COCHLAN, W. P., J. WIKNER, G. F. STEWARD, D. C. SMITH, AND F. AZAM. 1993. Spatial distribution of viruses, bacteria and chlorophyll-a in neritic, oceanic and estuarine environments. Mar. Ecol. Prog. Ser. **92:** 77–87.

COTTRELL, M. T., AND C. A. SUTTLE. 1991. Wide-spread occurrence and clonal variation in viruses which cause lysis of a cosmopolitan, eukaryotic marine phytoplankter, *Micromonas pusilla.* Mar. Ecol. Prog. Ser. **78:** 1–9.

DIMMOCK, N. J., AND S. B. PRIMROSE. 1994. Introduction to modern virology, 4th ed. Blackwell Science.

FANDINO, L. B., L. RIEMANN, G. F. STEWARD, R. A. LONG, AND F. AZAM. 1998. Bacterial succession during a dinoflagellate bloom analyzed by DGGE and rDNA sequencing. EOS Trans. (suppl) **79:** OS63.

GOWING, M. M. 1993. Large virus-like particles from vacuoles of phaeodarian radiolarians and from other marine samples. Mar. Ecol. Prog. Ser. **101:** 33–43.

HELDAL, M., AND G. BRATBAK. 1991. Production and decay of viruses in aquatic environments. Mar. Ecol. Prog. Ser. **72:** 205–212.

HENNES, K. P., C. A. SUTTLE, AND A. M. CHAN. 1995. Fluorescently labeled virus probes show that natural virus populations can control the structure of marine microbial communities. Appl. Environ. Microbiol. **61:** 3623–3627.

JIANG, S. C., AND J. H. PAUL. 1995. Viral contribution to dissolved DNA in the marine environment as determined by differential centrifugation and kingdom probing. Appl. Environ. Microbiol. **61:** 317–325.

KLIEVE, A. V., AND R. A. SWAIN. 1993. Estimation of ruminal bacteriophage numbers by pulsed-field gel electrophoresis and laser densitometry. Appl. Environ. Microbiol. **59:** 2299–2903.

MURRAY, A. G., AND G. S. JACKSON. 1992. Viral dynamics: A model of the effects of size, shape, motion and abundance of single-celled planktonic organisms and other particles. Mar. Ecol. Prog. Ser. **89:** 103–116.

PAUL, J. H., S. C. JIANG, AND J. B. ROSE. 1991. Concentration of viruses and dissolved DNA from aquatic environments by vortex flow filtration. Appl. Environ. Microbiol. **57:** 2197–2204.

REISSER, W. 1993. Viruses and virus-like particles of freshwater and marine eukaryotic algae—a review. Arch. Protistenkd. **143:** 257–265.

ROHWER, F., A. SEGALL, G. F. STEWARD, V. SEGURITAN, M. BREITBART, F. WOLVEN, AND F. AZAM. 2000. The complete genomic sequence of the marine phage Roseophage SIO1 indicates homology with non-marine phages. Limnol. Oceanogr. **45:** 408–418.

SCHNEEBERGER, C., P. SPEISER, F. KURY, AND R. ZEILLINGER. 1995. Quantitative detection of reverse transcriptase-PCR products by means of a novel and sensitive DNA stain. PCR Methods Appl. **4:** 234–238.

STEWARD, G. F. In press. Fingerprinting viral assemblages by pulsed field gel electrophoresis. *In* J. H. Paul [eds.], Methods in microbiology. Marine microbiology. Academic Press.

———, AND F. AZAM. 2000. Analysis of marine viral assemblages, p. 159–165. *In* C. R. Bell, M. Brylinski, and P. Johnson-Green [eds.], Microbial biosystems: New frontiers. Proceedings of the 8th International Symposium on Microbial Ecology. Atlantic Canada Society for Microbial Ecology.

———, J. WIKNER, D. C. SMITH, W. P. COCHLAN, AND F. AZAM. 1992. Estimation of virus production in the sea: I. Method development. Mar. Microb. Food Webs **6:** 57–78.

VAN ETTEN, J. L., L. C. LANE, AND R. H. MEINTS. 1991. Viruses and virus-like particles of eukaryotic algae. Microbiol. Rev. **55:** 586–620.

WEINBAUER, M. G., D. FUKS, S. PUSKARIC, AND P. PEDUZZI. 1995. Diel, seasonal, and depth-related variability of viruses and dissolved DNA in the Northern Adriatic Sea. Microb. Ecol. **30:** 25–41.

WICHELS, A., S. S. BIEL, H. R. GELDERBLOM, T. BRINKHOFF, G. MUYZER, AND C. SCHÜTT. 1998. Bacteriophage diversity in the North Sea. Appl. Environ. Microbiol. **64:** 4128–4133.

WIKNER, J., J. J. VALLINO, G. F. STEWARD, D. C. SMITH, AND F. AZAM. 1993. Nucleic acids from the host bacterium as a major source of nucleotides for three marine bacteriophages. FEMS Microbiol. Ecol. **12:** 237–248.

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.

WOMMACK, K. E., AND R. R. COLWELL. 2000. Virioplankton: Viruses in aquatic ecosystems. Microbiol. Mol. Biol. Rev. **64:** 69–114.

———, J. RAVEL, R. T. HILL, J. CHUN, AND R. R. COLWELL. 1999. Population dynamics of Chesapeake Bay virioplankton: Total-community analysis by pulsed-field gel electrophoresis. Appl. Environ. Microbiol. **65:** 231–240.

Received: 16 March 2000
Accepted: 31 August 2000
Amended: 7 September 2000