6. Fingerprinting Viral Assemblages by Pulsed Field Gel Electrophoresis (PFGE)

Grieg F. Steward
Monterey Bay Aquarium Research Institute, Moss Landing, California
Present Address: University of Hawaii at Manoa, Department of Oceanography

CONTENTS

Introduction
Principle
Application
Conclusion

INTRODUCTION

Viruses are the most abundant microorganisms in marine and freshwater environments and perhaps the most genetically diverse (Fuhrman and Suttle, 1993). Counting viruses in aquatic samples is now a routine matter, but assessing the diversity and dynamics within complex assemblages is still a challenge. DNA-based fingerprinting approaches, which rely on amplification of rRNA gene fragments by PCR, have facilitated analyses of bacterial community composition. These approaches have more restricted application when analyzing viral assemblages, because of the extreme genetic diversity among viruses. Unlike in bacteria, there are no gene sequences conserved in all viruses that can serve as universal primer sites for PCR amplification. PCR-based analyses of viral assemblages must therefore target specific subsets of the total viral assemblage. For example, PCR amplification of specific genes has recently been used to examine the genetic diversity among cyanophages (Fuller et al., 1998) and among phycodnaviridae (Chen et al., 1996; Short and Suttle, 1999). A more general fingerprinting approach, which encompasses the total viral assemblage, is a valuable complement to these more specific, higher resolution analyses. The approach described here uses variation in genome size as the basis for obtaining a fingerprint of a viral assemblage (Klieve and Swain, 1993). A whole genome fingerprinting approach is possible, because viral genomes can vary greatly in length (a few thousand to hundreds of thousands of base pairs) yet they fall within a range that is easily resolved using pulsed field gel electrophoresis (PFGE). The PFGE fingerprinting technique provides a quick and relatively simple means of visualizing differences in the composition of viral assemblages (Swain et al., 1996; Wommack et al., 1999a; Steward et al., 2000). As a supplement to the more specific treatment of PFGE provided in this chapter, the reader is encouraged to consult the excellent introductory text to PFGE by Birren and Lai (1993).

PRINCIPLE

Viruses are harvested from a water sample by ultrafiltration and the capsids are destabilized to release the viral DNA. Intact viral genomes are separated by size in an agarose gel by PFGE. After separation, the DNA banding pattern is revealed with a fluorescent DNA stain. The banding pattern provides a visual record of the genome size distribution, which can be used for qualitative and quantitative comparisons among samples. Using image analysis, the molecular weight and mass of DNA in each band are determined by comparison with the migration rate and fluorescence intensity of DNA standards. Molecular weight and mass are then used to calculate the genome copy number in each discrete band or molecular weight size range.
Equipment and reagents

• *Tangential flow ultrafiltration system:* A tangential flow system is necessary if viruses are to be harvested from a liter or more of water. Choose a system with the smallest minimum recirculation volume which will still provide a reasonable processing time (i.e., less than an hour or two). A system with smaller tubing and a smaller membrane surface area will process more slowly, but will also have a smaller minimum recirculation volume. This means the concentration factor achievable for a given initial sample volume will be greater. Spiral wound membranes—designed for processing low viscosity, particle-free fluids—are well suited to harvesting viruses from prefiltered (0.2 µm) water. Other configurations, however, such as membrane cassettes or hollow fiber cartridges, and other systems such as vortex flow filtration, can also be used. A 30,000 molecular weight cut-off (MWCO) is recommended to ensure retention of smaller viruses, but 100,000 MWCO membranes may be adequate. Membranes rated at < 30,000 MWCO will filter more slowly and retain more unwanted low molecular weight material.

• *Sterile filters (0.2 µm pore size):* Sterivex GV filters (Millipore, Bedford, Massachusetts, USA) or their equivalent are adequate for filtering 10 ml to 10 l of water depending on the particle load. Filters with larger surface area (e.g., pleated capsule filters) are required for larger volumes (ten to hundreds of liters).

• *Centrifugal ultrafiltration units (100,000 MWCO):* Large capacity (20 to 80 ml) units are useful as a secondary concentration step. They can also be used for primary concentration of small volume samples (≤ 1 liter). Small capacity (0.5 ml) centrifugal ultrafiltration units are used for final virus concentration and preparation of viral DNA.

• *Centrifuge:* Moderate speed (capable of 4,000 g) with swinging buckets and adapters to hold large capacity centrifugal ultrafiltration units. A centrifuge with refrigeration is recommended.

• *Microcentrifuge:* Speed must be adjustable. A unit with refrigeration is recommended.

• *Pulsed field gel electrophoresis system:* the system should be capable of providing resolution of DNA up to several hundred-thousand base pairs. The procedures described in this chapter are for units using a clamped homogeneous electric field (CHEF) configuration which are available commercially from Bio-Rad.

• *Gel documentation equipment:* A variety of system configurations are possible. Laser gel scanners are preferred as they provide the greatest sensitivity and resolution and direct acquisition of digital images. For systems using illumination by UV lamps, epi-illumination is reported to provide higher sensitivity than transillumination for gels stained with SYBR Green I (Molecular Probes, Inc.). Regardless of illumination method, documentation with a digital camera is preferred over film since the images can be directly imported into gel analysis programs. If film must be used, a scanner can be used to convert the photographs into digital images, but the dynamic range and resolution will be lower than for the direct, digital image acquisition.

• *Gel analysis software:* the software should have capabilities for band recognition, calculation of integrated intensity, and molecular weight determination

• *Fluorescent DNA stain* for DNA quantification such as PicoGreen (Molecular Probes)

• *Fluorometer* capable of measuring fluorescence of fluorescently stained DNA (e.g., 502 nm excitation, 523 nm emission peaks for PicoGreen).

• *Purified Bacteriophage lambda DNA*

• *Pipets, micropipettors*

• *Agarose:* A PFGE-grade agarose with a standard gelling temperature such as SeaKem Gold (FMC Bioproducts) can be used for routine application. If DNA is to be recovered from the gel following electrophoresis, then a low-melting-point agarose such as SeaPlaque GTG (BioWhittaker) should be used instead. A low-melting-point agarose is also recommended for embedding samples in agarose plugs prior to electrophoresis.
• **DNA molecular weight standards:** standards should cover the range from about ten- to several hundred-thousand base pairs. Low- and Midrange PFG Markers (New England BioLabs) are convenient as they each provide full range coverage with a single marker. Other useful markers are a 5 kb ladder (concatemers of a 4.8 kb plasmid) and a lambda ladder (concatemers of lambda phage genomes) which can be used in combination (available from various suppliers).

• **DNA mass standards:** DNA mass ladder (GIBCO/BRL) or dilutions of lambda phage DNA

• **Running Buffer** (10 M TBE stock contains per liter: 108 g Tris base, 55 g Boric acid and 40 ml of 0.5 M EDTA, pH 8)

• **Loading Buffer** (10 stock contains 25% ficoll, 0.25% Bromophenol blue or xylene cyanol)

• **SYBR Green I** (10 000 stock; Molecular Probes) or ethidium bromide (5 mg ml⁻¹ stock)

• **Sodium Azide:** 10% stock solution in water, filtered (0.2 µm)

• **SM or Marine SM (MSM):** These are storage buffers for non-marine or marine bacteriophages. SM contains 100 mM NaCl, 10 mM MgSO₄, 50 mM Tris (pH 7.5), and 0.01% gelatin (Sambrook et al., 1989). MSM is a modification of the original SM recipe which more closely matches the ionic composition and pH of seawater and contains 450 mM NaCl, 50 mM MgSO₄, 50 mM Tris (pH 8.0), and 0.01% gelatin. Sterilize by autoclaving and store at room temperature. These buffers can also be prepared without the gelatin for situations where adding additional, high molecular weight protein to the sample is undesirable.

*Alternative:* The final virus concentration step may be accomplished by ultracentrifugation instead of centrifugal ultrafiltration. In this case an ultracentrifuge and appropriate tubes are required and could replace the centrifuge, microcentrifuge and centrifugal ultrafiltration devices.

**Assay**

In addition to the following detailed description, an overview of the viral fingerprinting procedure is presented as a flow chart (Figure 6.1).

**Virus concentration and storage**

The sample volume required can vary greatly depending on the initial concentration of viruses, losses during processing, sensitivity of detection, and whether extra DNA is desired for multiple gel runs, archiving, or other analyses. For a typical seawater sample, at least 10⁹ viruses or 50 ng of viral DNA are needed in the end to obtain a single fingerprint in a 10 mm wide well. For a typical surface seawater concentration of 10¹⁰ viruses l⁻¹, this translates into a minimum sample volume for a single fingerprint of roughly 100 ml. A larger volume is recommend, however, to account for losses and to have extra material. In practice, process volumes may range from around 10 ml to > 50 l, with a 1 to 3 liters being sufficient in most cases.

**Processing large volumes (> 1 liter)**

• Filter sample (0.2 µm pore size) to remove bacteria. For water with a high particle load, prefiltration through a larger pore filter may be useful to avoid clogging the 0.2 µm filter.

• Concentrate viruses in the filtrate by tangential flow ultrafiltration to ≤ 400 ml then proceed with small volume concentration (below).

**Processing small volume samples or primary concentrates (< 400 ml)**

• For primary viral concentrates skip to the next step. For small samples, filter through a 0.2 micron Sterivex filter (Millipore) via syringe or peristaltic pump to remove bacteria.

• Concentrate viruses by centrifugal ultrafiltration to ≤ 250 µl. Samples exceeding the capacity of the filter reservoir are concentrated by repeated rounds of centrifugation. After each round, the filtrate is discarded and the upper reservoir is refilled with additional sample.

• Recover sample by inverting the filter cup into a collection tube and centrifuging at 250 g for 1 min. Rinse membrane with 250 µl of MSM, 0.2 µm-filtered sample, or ultrafiltrate, recover and pool.

• Add sodium azide to 0.1% final concentration and store at 4 °C in the dark.
Figure 6.1. Flow chart illustrating the major steps in generating and analyzing whole-genome fingerprints of viral assemblages. Bulleted details within each box are those recommended and presented in the text, but alternatives and variations are possible at each step. Dashed arrows represent alternative pathways which include temporary storage of samples.

Alternatively, viruses can be concentrated from small volumes by ultracentrifugation. The centrifugation time and speed required will vary depending on the rotor. Run conditions are calculated from the k factor of the rotor and an assumed sedimentation coefficient of about 100S in order to pellet the smallest viruses. The formula is \( t = \frac{k}{S} \), where \( t \) is time in hours, \( k \) is the k-factor and \( S \) is the sedimentation coefficient in Svedbergs. The k factor assumes sedimentation of particles in pure water at 25°C. The time may need to be corrected for salinity and temperature effects on the density and viscosity of water. For example, if viruses are pelleted from seawater vs freshwater centrifugation time should be increased by about 12%. Temperature effects on water viscosity become pronounced below 20°C, so centrifugation time should also be increased by an additional 25% for each 5°C decrease below 20°C (Suttle, 1993).
Preparation of viral DNA for electrophoresis

There are two approaches to DNA preparation, in agarose plugs or in solution. Traditionally, samples for PFGE are embedded in agarose plugs and treated with a chelator (EDTA), detergent, and proteinase K (Klieve and Swain, 1993; Wommack et al., 1999a). However, if the samples are handled carefully, embedding is not necessary for DNA in the size range of typical viral genomes. In addition, proteinase K digestion is not strictly necessary since viral capsids can be sufficiently destabilized and the DNA released by exposure to EDTA and moderate heating.

DNA in solution

1. Transfer an aliquot of viral concentrate containing 40 - 100 ng of DNA (ca. 1 to 2 × 10⁹ viruses) to a Microcon 100 and concentrate to near dryness at 1000 × g. Required centrifugation time is typically 5 to 20 min.
2. Rinse the Microcon by gently adding 50 microliters of 1× TE then concentrating to near dryness (ca. 5 microliters) as above. Centrifugation time is usually 5 to 10 min.
3. Repeat rinse twice more.
4. Add 10 µl TE to the Microcon (be sure it wets the membrane) and recover by inverting the filter cup in a collection tube and centrifuging at ca 250 × g for 1 min. Repeat the recovery step once or twice more pooling all material into the same collection tube.
5. Heat the recovered sample to 60° C for 10 min.
6. Cool the sample on ice and spin down any condensation in the tube by brief centrifugation.
7. For precise control on the amount of DNA loaded per lane, determine the DNA concentration in the viral concentrate using the PicoGreen DNA quantification assay (Molecular Probes) and a fluorometer. For the assay, dilute 1 - 5 µl of the recovered DNA into 1× TE and proceed according to the manufacturer’s protocol.
8. Dilute samples as necessary with an appropriate volume of 1× TE then add 1/10 volume of 10× Loading Buffer. For maximum band sharpness, the final sample volume should close to the well capacity.
9. Gently mix sample and loading buffer. An effective, gentle mixing method is to roll the tube back and forth between thumb and forefinger while simultaneously inverting the tube five to ten times.

DNA in agarose plugs

1. Prepare a viral concentrate of about 1-5 × 10¹¹ viruses ml⁻¹ (5-25 µg viral DNA ml⁻¹). Concentrates may be prepared by centrifugal ultrafiltration or by ultracentrifugation. Recover concentrated viruses in a small volume of SM or MSM (without gelatin). If using ultracentrifugation, add the buffer to the viral pellet, cover the tube with Parafilm, and let sit several hours to overnight at 4° C. Pipet in and out several times to mix, then quickly transfer to a microcentrifuge tube. If using centrifugal ultrafiltration, concentrate and recover as described in Step 4 of the previous section substituting SM or MSM (without gelatin) for the TE.
2. Melt a stock of 1.5% (w:v) low-melting-point agarose in SM or MSM (without gelatin) using a microwave oven or hot water bath.
3. Transfer aliquots of agarose to microtubes in a heating block or water bath at 50 to 60° C to keep them molten.
4. Combine equal volumes of viral concentrate (at room temperature) and molten agarose, vortex or pipet in and out several times to mix, then quickly transfer to plug molds and let solidify.
5. Push the plugs into a microcentrifuge tube containing about 3 to 5 volumes of freshly prepared extraction buffer (100 mM EDTA, pH 8.0, 1% SDS and 1 mg ml⁻¹ proteinase K).
6. Incubate overnight at room temperature with gentle agitation.
7. Rinse the plugs 3 × 30 min with 1× TE.
8. Plugs can be stored at 4° C in 1× TE.
9. To prepare a plug for electrophoresis, place it on a piece of Parafilm, blot away excess liquid and trim with a razor blade so it will fit within a gel well.

DNA size and mass markers

A number of size markers for pulsed field gel electrophoresis are commercially available. The commonly available 5 kb and lambda ladders can be used in combination to cover the range of viral genome sizes. The lambda ladder is supplied in agarose plugs while the 5 kb ladder is in solution.
The lambda ladder is at a sufficiently high concentration that only a small sliver of the plug is needed (< 1 mm). This leaves enough room that the 5 kb ladder can be loaded into the same well. Full size range coverage can then be achieved in a single lane. Specially designed ladders which span the entire range of viral genome sizes are also available (New England BioLabs). At least two size-marker lanes should be run, one at either end of the gel. For precise molecular weight determinations, a marker lane in the middle of the gel may be needed in some cases. For quantifying DNA based on fluorescence intensity, DNA mass standards should also be run to compensate for any non-linearity in fluorescent signal detection. Mass ladders are commercially available which allow calibration to be carried out using only a single gel lane. However, these mass ladders are usually comprised of small DNA fragments (e.g., 1 to 10 kb for the High DNA Mass™ Ladder; Life Technologies Inc.). Most of the small fragments will run off the gel using the PFGE conditions typical for viral fingerprinting. This can be prevented by loading the mass ladder in the middle or near the end of the run. Alternatively, the mass ladder or a series of dilutions of lambda DNA could be run separately by conventional electrophoresis in a gel of the same thickness and percent agarose. The gels should then be stained simultaneously in the same batch of stain and photographed together.

Running the gel

1. Cast a 1% agarose gel in 0.5 TBE. A comb with 5 to 10 mm-wide wells is recommended. Wider wells make it easier to distinguish faint bands from background. First load any samples which are embedded in plugs. Before inserting a plug, fill the well with running buffer to ensure no air bubbles become trapped in the well. For the lambda ladder marker, cut off a thin sliver with a razor blade then place the sliver (about 0.5 mm x 0.5 mm x 5-10 mm) against the side of the well. Small spatulas or plastic inoculating loops are convenient for handling gel plugs and slices.
2. Place the gel in the electrophoresis chamber, add running buffer slowly to avoid dislodging the gel from the casting plate, and start recirculating and chilling the buffer. The chiller should be adjusted so that the buffer temperature in the chamber is 14°C. Recirculate the buffer at a rate of about 0.5 to 1 l min⁻¹ (ca. 1 tank volume every 2 to 3 minutes). Note: To save time, the chamber can be pre-filled and the buffer chilled while the gel is solidifying. When lowering the gel into the buffer, care must be taken not to dislodge the gel from casting plate.
3. Load liquid samples and markers, pipetting slowly to avoid shearing the DNA.
4. Run the gel at 6 V cm⁻¹ for 18 h with pulses ramping from 1 to 10 s at an included angle of 120°. These conditions provide good separation from 10 to ca. 200 kb. Pulse and run time may be varied to improve separation in specific size ranges. Others have used a pulse ramp of 1 to 15 s over a run time of 22 h (Wommack et al., 1999a). Longer pulse and run times improves separation of higher molecular weight material. If pulse and run times are increased too much, however, resolution in the lower molecular weight regions will be sacrificed. Consult Birren and Lai (1993) for a more detailed discussion of the effects of switch and run times on DNA separations.
5. Stain the gel in the dark for 0.5 to 1 h in about 150 to 200 ml of running buffer containing SYBR Green I (1 µ) or ethidium bromide (0.05 µg ml⁻¹). If using ethidium bromide, then destain the gel for about 0.5 h in 5 to 10 gel-volumes of water.
6. Acquire an image of the gel with a gel documentation system. If using film, several exposures may be useful to optimize the number of bands within the dynamic range of the film. Laser scanners should be set to high resolution to improve band discrimination.

Analysis of banding patterns

**Description of an individual sample**

Quantitative analyses of banding patterns are possible with a variety of commercially available software programs. The program should have the capability to 1) identify bands (with adjustable detection parameters and manual editing capability), 2) integrate the fluorescence intensity of individual and partly overlapping bands, and 3) calculate molecular weight and mass for each band relative to migration distance and intensity of standards on the same gel. This information along with some simple spreadsheet calculations will provide a statistical description of a given sample in terms of the range, mean, median, and mode of the detectable viral genome sizes. The data can also
be graphically displayed as a size-frequency distribution. The calculations for determining the fraction of virus-like genomes within a given band or within size-range bins are presented below. Note that DNA size standards are usually reported in terms of length (base pairs) rather than molecular weight, which is used in the equations here. DNA size in base pairs must first be converted to molecular weight \((M)\) using:

\[
M = S \cdot M_{bp}
\]

where \(S\) is the size in base pairs and \(M_{bp}\) is average molecular weight of a base pair (\(= 660\)).

Genome copy number in each of the detected bands is then calculated as:

\[
C_i = \frac{D_i}{M_i} \cdot A_v
\]

where \(C_i\) is the genome copy number, \(D_i\) is the mass (in grams) and \(M_i\) is molecular weight (in Da) of the DNA in band \(i\), and \(A_v\) is Avogadro's number \(\left(6.022 \times 10^{23}\right)\).

The absolute number of genomes is directly dependent on the amount of sample loaded on the gel. Therefore, to facilitate comparisons among samples, copy number is better expressed as a relative abundance. The fraction of the total viral genomes present in each of the bands is calculated as:

\[
F_i = \sum_{i=1}^{n} \frac{C_i}{C_i}
\]

where \(F_i\) is the fraction of genomes and \(C_i\) the genome copy number in band \(i\), and \(n\) is the total number of bands.

These data can be graphically presented as size-frequency plots (Steward and Azam, 2000). In many cases, there are regions on the gel where individual bands cannot be resolved. To deal with this, as well as to standardize comparison among different samples, it may be preferable in some instances to bin the frequency data into defined size ranges and plot as a histogram (Steward et al. 2000).

The average genome size for an assemblage can be derived from these data by first calculating an abundance-weighted molecular weight \((M^*)\) for each band \(i\) as:

\[
M_i^* = M_i \cdot C_i
\]

The average genome size \((G_{avg})\) can then be calculated as

\[
G_{avg} = \frac{\sum_{i=1}^{n} M_i^*}{\sum_{i=1}^{n} C_i}
\]

The mode of the genome size distribution is readily identified as the genome size or (size bin) with the maximum \(F_i\). The median genome size can also be identified by calculating a running sum of \(F_i\) from the largest to the smallest genome (or vice versa). The genome size or size class at which the sum is at, or closest to, 50% is the median.
Comparisons among samples

A similarity coefficient can be calculated for any two samples from their genome size-frequency distributions. In the case of viral fingerprints, frequency data are expressed as proportions rather than absolute abundance (from Equation 3). Two measures of similarity designed to handle proportional data are the Renkonen index (percentage similarity) and the simplified Morisita index (Krebs, 1999). Binary similarity coefficients can also be calculated by considering only the presence or absence of different sized genomes (e.g., Wommack et al., 1999a). Since information on relative abundance is ignored, this provides a much cruder estimate of similarity. However, it may be preferred in situations where the relative recoveries of different viruses is suspected or known to vary significantly from sample to sample. In this case, changes in the relative brightness of different bands could be due to artifacts of sample processing and are best not considered.

Relationships among three or more samples can be analyzed by cluster analysis. The most frequently used strategy for clustering is the unweighted pair-group method using arithmetic averages (UPGMA). The algorithm makes use a matrix of similarity coefficients obtained from pairwise comparisons of all samples. This is an hierarchical, agglomerative approach that is used to create dendrograms illustrating the relative similarities among groups of samples. The details of the calculation of similarity coefficients and of cluster analysis are available from many sources and are not reproduced here. For one recent, lucid presentation and discussion of these methods the reader is referred to Krebs (1999).

Troubleshooting

• Migration is too fast or too slow. The recommended electrophoresis parameters apply for ca. 2 liters of 0.5% TBE in a CHEF unit (Bio-Rad) run at 14°C. If other buffer formulations, volume or temperatures are used, the parameters may need adjusting. For example, DNA will migrate faster in Tris-acetate EDTA (TAE) buffer than in TBE, but TBE is often used because it has a greater buffering capacity. Running gels at lower temperatures improves band resolution, but dramatically reduces the migration rate. Migration rate will also decrease with increasing buffer volume and increasing ionic strength. For a comprehensive discussion of these and other factors influencing migration rate of DNA in pulsed field gels the reader is referred to Birren and Lai (1993).

• Migration is uneven across the gel. Uneven migration is most likely the result of an improperly leveled gel box. The bottoms of some electrophoresis tanks can have some curvature making small, bubble levels less accurate. For best results use a bubble level which is short enough to fit in the tank, but long enough to rest on electrodes at either side of the tank. Alternatively, the tank level can be checked by adding a small amount of buffer to the chamber then tipping the chamber up, first on the front two feet then the back two feet. With the chamber tipped, adjust the feet so that the buffer level is at the same position relative to the electrodes on opposite sides of the tank.

• Degraded DNA. Degraded DNA may be caused by nucleases or mechanical shearing and will appear as a smear on the gel. DNA released from viral capsids by EDTA and heat treatment tends to degrade with time in solution. Therefore, samples to be loaded in solution should be prepared fresh from the viral concentrates and should be mixed or pipetted gently to prevent mechanical shearing. If one wishes to store samples as DNA rather than as viral concentrates, the viruses should be embedded and the DNA extracted and stored in agarose plugs rather than in solution. Band resolution is somewhat lower for embedded DNA (see Applications, below), but the DNA will be better preserved. Remember to wear gloves whenever working with DNA to minimize the potential for DNase contamination. In addition, be sure the DNA storage buffer contains sufficient EDTA to chelate any magnesium ions as this will inhibit DNase.

On the other hand, if storing intact viruses, a buffer containing magnesium ions (like SM or MSM) is used in order to maintain stability of the viral capsids. A preservative such as 0.1% sodium azide should be added to prevent the growth of bacterial contaminants. Both DNA and viruses should be stored cold (4°C) and dark. Viral concentrates can also be stored frozen, but there are no reports yet on the relative stability of frozen vs. refrigerated concentrates of natural viral assemblages. Viral
counts do decline with time in refrigerated concentrates (unpubl. obs.) so samples should be analyzed as soon as possible.

**Artifactual banding patterns.** If a viral concentrate is stored improperly, the banding pattern could be compromised. Banding patterns from seawater samples are typically complex. If only one or two bands dominate a sample, it could be a result of an improperly stored concentrate. An example of this is shown in Figure 6.2. Two filtered (0.2 µm), unpreserved primary concentrates were stored at 4° C for about 3 weeks. Inspection of the samples by epifluorescence microscopy showed bacterial contamination in both. In addition, one of the samples showed a ≥ 12-fold increase in viral abundance. The PFGE banding pattern of that sample was dominated by one band indicating that one of the bacterial contaminants served as a host for viral replication during storage. Banding pattern could also be compromised by differential degradation of virus strains within a sample during storage. As mentioned above, viruses do decay with storage, but the degree to which this affects the viral fingerprint has not yet been tested.

Dominance of one or few bands is not necessarily an artifact as it could also occur as a result of natural processes, for example, by mass viral lysis of a monospecific bloom of some organism. In order to distinguish real dominance from artifact it is helpful to monitor recovery of viruses at each step of harvesting and concentration (e.g., by epifluorescence microscopy, Chapter 3). If recovery is unusually low, or impossibly high, or if growth of bacterial contaminant is detected during storage, the resulting fingerprint should be considered suspect.

![Figure 6.2](image.png)

**Figure 6.2.** Example of an artifactual banding pattern caused by storage of an unpreserved sample for three weeks at 4° C. Viruses were concentrated from surface waters of Monterey Bay, CA and separated by PFGE using a 1-6 s pulse ramp over 16 h. Other electrophoresis conditions are as described in the text. Lane 1 shows a typical banding pattern for one sample which had bacterial contamination, but showed no increase in viral abundance during storage. Lane 2 shows a sample which had bacterial contamination and showed a ≥ 12-fold increase in viral abundance during storage. Lane M is a marker lane containing a 5 kb ladder and a lambda ladder with the size of selected bands indicated on the left.
Genomic fingerprinting of viral assemblages by pulsed field gel electrophoresis may be applied to essentially any type of sample from which a sufficient number of viruses can be harvested. So far the technique has been used to analyze phage communities in the rumen of sheep (Klieve and Swain, 1993) and viral assemblages in estuarine (Wommack et al., 1999a; Wommack et al., 1999b) and oceanic waters (Steward and Azam, 2000; Steward et al., 2000). Preliminary data show that the technique can also be applied to viruses harvested from sea ice and sediment pore waters (unpubl. obs.). A comparison of various sample preparation procedures is illustrated using surface seawater from Monterey Bay (Figure 6.3A). Concentration by centrifugal ultrafiltration or by ultracentrifugation yielded essentially the same results, but the ultracentrifuged sample showed some degradation (smearing in the lower molecular weight range). This probably reflects limited mechanical shearing due to the gentle sample agitation, which was used when releasing DNA from the viral pellet. Banding patterns were also similar whether DNA was prepared in agarose plugs or in solution, although band resolution is somewhat poorer for embedded samples. This difference in resolution between samples loaded in solution or as gel plugs is also illustrated with lambda phage DNA (Figure 6.3A). Note that for embedded samples, band sharpness is limited by the thickness of the gel plug. Although concentration of viruses from 3 to 10 liters of water is recommended in order to have plenty of extra material, small-scale concentrations can also provide enough material for a fingerprint in some environments. For example only 80 ml of seawater from 20 m depth in Monterey Bay was sufficient to obtain a viral fingerprint (Figure 6.3B). This is a volume which can be easily processed using only centrifugal ultrafiltration.

Figure 6.3. Comparison of sample preparation procedures using surface seawater from Monterey Bay, CA. (A) An initial viral concentrate was prepared by tangential flow ultrafiltration then the viral DNA was processed for PFGE by one of three methods: Lane 1—viruses were further concentrated by centrifugal ultrafiltration, embedded in agarose, the DNA released with EDTA, SDS and proteinase K treatment, then loaded as an agarose plug. Lane 2—viruses were further concentrated by centrifugal ultrafiltration, the DNA released by EDTA and heat treatment, then loaded in solution. Lane 3—viruses were further concentrated by pelleting in an ultracentrifuge, viral DNA was released into solution by EDTA and heat treatment, then loaded in solution. Pulse conditions were a 1-10 s ramp over 18 h. Lanes labeled “M” contain a 5 kb ladder and a Saccharomyces cerevisiae chromosomal DNA marker (BioRad). Sizes of selected bands are indicated on the left. (B) Viruses were concentrated from 80 ml of 0.2 µm-filtered seawater (20 m depth) using a Centricon 20 (lane 1) or Centricon 80 (lane 2), all of each concentrate was then processed for electrophoresis on Microcon 100 centrifugal units and loaded on the gel as described in the text. Marker lane M contains 5 kb ladder and a lambda ladder (BioRad).
The banding patterns (fingerprints) reflect the diversity of genome sizes within a viral assemblage and comparisons among samples can reveal the spatial and temporal scales on which assemblage composition varies. It should be kept in mind, however, that not all of the viral diversity is necessarily resolved by a one dimensional separation. Different viruses can have the same genome size, so band number is only a minimum estimate of the total genetic diversity. In addition, much of the viral DNA from environmental samples is clustered in a few narrow size ranges. DNA in these regions appears as somewhat diffuse zones rather than discrete bands. The number of different genomes comprising these zones is not known, but could be quite large. Since bands are difficult to resolve in these regions, not all shifts in assemblage composition are necessarily detectable by PFGE alone. However, in practice, spatial and temporal variations in viral community composition have been demonstrated in a variety of environments (Swain et al., 1996; Wommack et al., 1999a; Steward et al., 2000).

Unlike the DNA fingerprinting techniques available for bacteria, viral assemblage fingerprinting by PFGE has the advantage of not involving PCR amplification, which can introduce a number of biases and artifacts. Since amplification biases are avoided with this method, band intensity is more directly related to the relative abundance of different sized genomes in the original sample. However, several points should be kept in mind. First, measured fluorescence intensity is not necessarily a linear function of DNA mass (Figure 6.4). Deviations from linearity may depend on the fluorescent stain used as well as the specifications of the system used to record and quantify fluorescence. Therefore, if one wishes to quantify genome copy numbers, the fluorescent signal must be calibrated using DNA mass standards. Second, dye binding could vary for different viral genomes depending on factors such as G+C content, presence of modified bases, and whether nucleic acid within the band is single or double stranded, RNA or DNA. To what extent these factors influence the analysis of marine viral assemblages is not well established. It is assumed, however, that most marine viruses contain double-stranded DNA (dsDNA) genomes which should stain with similar efficiencies. Third, there are potential biases in sample preparation. These could include selective losses of larger viruses during 0.2 µm filtration and of more labile viruses during concentration and storage. Again, the influence of these factors is not well known. However, comparison of a filtered and an unfiltered sample indicated that filtration did not dramatically affect the fingerprint (unpubl. obs.). To minimize the potential for degradation effects, samples should be kept cold and dark, harvested quickly and run on a gel as soon as possible.

Another advantage of the PFGE fingerprinting method is that the fingerprint is generated using intact viral DNA. This means that for any identified band of interest, the entire viral genome is accessible for further analysis by excising it from the gel. PFGE can thus serve as a starting point for more detailed analyses of viral assemblage composition. This capability is just beginning to be exploited for ecological studies (Wommack et al., 1999b).

The detection limit of the assay can be estimated from the dilution series of lambda phage DNA in Figure 6.4. For this gel stained with SYBR Green I (Molecular Probes) and imaged with a laser gel scanner (Molecular Dynamics) the detection limit was ≤ 0.05 ng of DNA in a 5 mm well. Assuming a 30 kb dsDNA genome, about 10⁶ viruses would be detectable. For viruses with larger genomes, the detection limit is even lower, being about 10⁸ viruses for a genome size of 300 kb. Assuming that viruses from 100 ml of seawater (i.e., about 10⁹ viruses) are loaded in a lane, this method is theoretically capable of detecting a virus with a 300 kb genome which comprises < 0.1% of the community. Viruses with a more typical genome of ≥ 30 kb would be theoretically detectable even if <1% of the community. The practical detection limit is likely to be somewhat higher due to background fluorescence in sample lanes, but the method is clearly capable of detecting even relatively minor components of the viral assemblage.
CONCLUSION

Genomic fingerprinting by PFGE is a useful tool for exploring the diversity and dynamics of viruses in the environment. Fingerprints obtained with PFGE can be used to reveal variability in the composition of viral assemblage over space and time. Quantitative analysis of banding patterns also provides detailed information about viral diversity and genome size distributions in the environment. The ability to access intact viral genomes from the fingerprints means that PFGE is also a useful starting point for more detailed analyses. The generality of the method makes it a valuable complement to the more specific, PCR-based methods for analysis of viral diversity and dynamics. Together these techniques provide the means to identify and study the ecology of important viral groups in the ocean regardless of whether they can be cultivated.
References


List of suppliers

A/G Technology Corporation
101 Hampton Avenue
Needham, MA 02494-2628, USA
Tel: +1 781 449 5774 or +1 800 248 2535
Fax: +1 781 449 5786
E-mail: mail@agtech.com
Internet: www.agtech.com
Tangential flow ultrafiltration products
Amersham Pharmacia Biotech AB
SE-751 84 Uppsala, SWEDEN
Tel: +46 (0) 18 612 00 00
Fax: +46 (0) 18 612 12 00
Internet: www.apbiotech.com
Gel documentation/fluorescence DNA quantification system (Molecular Dynamics FluorImager), fluorometers

Bio-Rad Laboratories
Life Science Research Group
2000 Alfred Nobel Drive
Hercules, California 94547, USA
Tel: +1 800 424 6723
Fax: +1 510 741 5800
Internet: www.bio-rad.com
Pulsed field gel electrophoresis systems and supplies, Gel documentation systems, fluorescence DNA quantification equipment and reagents

BioWhittaker Molecular Applications
191 Thomaston Street
Rockland, ME 04841, USA
Internet: www.bioproducts.com
Tel: +1 800 341 1574
Fax: +1 207 594-3426
Agarose

Fisher Scientific
2000 Park Lane
Pittsburgh, PA 15275, USA
Tel: +1 800 766 7000
FAX: +1 800 926 1166
Internet: www1.fishersci.com
Centrifuges, general laboratory supplies and equipment, plasticware, chemicals, Polaroid film

GIBCO/BRL
Life Technologies, Inc.
9800 Medical Center Drive
Post Office Box 6482
Rockville, MD 20849-6482, USA
Tel: +1 800 338 5772
Fax: +1 800 3312286
Internet: www.lifetech.com
DNA mass ladder

Kendro Laboratory Products
31 Pecks Lane
Newtown, CT 06470-2337, USA
Tel: +1 800 522 7746
Fax: +1 203 270 2166, +1 203 270 2210 or +1 203 270-2115
E-mail: info@kendro.com
Internet: www.sorvall.com
Centrifuges and rotors

Millipore, Inc.
80 Ashby Road
Bedford, MA 01730-2271, USA
Tel: +1 781 533 6000
Fax: +1 781 533 3110
Internet: www.millipore.com

**Tangential flow and centrifugal ultrafiltration products (Millipore and Amicon brands), membrane filters**

**Molecular Probes, Inc.**
4849 Pitchford Ave.
Eugene, OR 97402-9165, USA
Tel: +1 541 465 8300
Fax: +1 541 344 6504
E-mail: order@probes.com
Internet: www.probes.com

**Fluorescent nucleic acid stains (SYBR Green I, PicoGreen)**

**New England BioLabs**
32 Tozer Road
Beverly, MA 01915, USA
Tel: +1 800 632 5227
Fax: +1 800 632 7440
E-mail: info@neb.com
Internet: www.neb.com or www.uk.neb.com

**DNA size markers**

**Phoretix International**
Tyne House, 26 Side
Newcastle upon Tyne
NE1 3JA UK
Tel: +44 (0)191 230 2121
Fax: +44 (0)191 230 2131
E-mail: web@phoretix.com
Internet: www.phoretix.com

**Gel analysis software**

**Scanalytics, Inc.**
8550 Lee Highway, Suite 400
Fairfax, VA 22031-1515
Tel: +1 703 208 2230
Fax: +1 703 208 1960
E-mail: info@scanalytics.com
Internet: www.scanalytics.com

**Gel analysis software**

**Sigma-Aldrich**
3050 Spruce Street
St. Louis, MO 63103, USA
Tel: +1 314 771 5765 or +1 800 325 8070
Fax: +1 314 771 5757
E-mail: custserv@sial.com
Internet: www.sigma-aldrich.com

Enzymes, chemicals, biochemicals, Polaroid film, DNA quantification standards and DNA size markers