ELSEVIER

Comparative Biochemistry and Physiology Part B 133 (2002) 463–476

CBP

www.elsevier.com/locate/cbpb

Review

Marine phage genomics☆

John H. Paul[a,*], Matthew B. Sullivan[b], Anca M. Segall[c], Forest Rohwer[c]

[a]*College of Marine Sciences, University of South Florida, 140 Seventh Ave S., St. Petersburg, FL, 33701 USA*
[b]*Massachusetts Institute of Technology/Woods Hole Oceanographic Institution Joint Program in Biological Oceanography, Cambridge, MA, USA*
[c]*Biology Department, San Diego State University, San Diego, CA, USA*

Received 7 May 2002; received in revised form 3 September 2002; accepted 4 September 2002

Abstract

Marine phages are the most abundant biological entities in the oceans. They play important roles in carbon cycling through marine food webs, gene transfer by transduction and conversion of hosts by lysogeny. The handful of marine phage genomes that have been sequenced to date, along with prophages in marine bacterial genomes, and partial sequencing of uncultivated phages are yielding glimpses of the tremendous diversity and physiological potential of the marine phage community. Common gene modules in diverse phages are providing the information necessary to make evolutionary comparisons. Finally, deciphering phage genomes is providing clues about the adaptive response of phages and their hosts to environmental cues.
© 2002 Elsevier Science Inc. All rights reserved.

Keywords: Phages; Genomics; Lysogeny; Roseophage; *Synechococcus*; *Prochlorococcus*

1. Introduction

Direct counts show that there are ~3–10 virus-like particles for every cell in the marine environment (Bergh et al., 1989; reviewed in Wommack and Colwell, 2000 and Fuhrman, 1999). Bacteria and Archaea are the most common cells in seawater, and it is believed that most of the viral-like particles are phages that prey upon these prokaryotes. Since the oceans are the world's largest biosphere, marine phages are probably the most abundant biological entities on the planet. Through their lytic activities, phages modulate carbon flow through microbial food webs by attacking both autotrophic and heterotrophic microbes (reviewed in Fuhrman, 1999). As prophages, marine phages may also confer a wide range of traits to their hosts including: immunity to superinfection (Hershey, 1971); toxin production (Waldor and Mekalanos, 1996); and the capability to transfer modular blocks of genes (Jiang and Paul, 1998a; Paul, 1999).

Even though 'The Age of Genomics' was heralded by the sequencing of phage *Escherichia coli* phi 174 in 1977 (Sanger et al., 1977), there are only three completed marine phage genomes currently in GenBank. This number will undoubtedly increase over the next decade. Here, we review the current state of the field of marine phage genomics and argue that these genomes, because of their small size, offer unprecedented opportunities for exploring eco-genomics, testing evolu-

☆ Contribution to a special issue of CBP on Comparative Functional Genomics.

*Corresponding author. Tel.: +1-727-553-1168; fax: +1-727-553-1189.
E-mail address: jpaul@seas.marine.usf.edu (J.H. Paul).

1096-4959/02/$ - see front matter © 2002 Elsevier Science Inc. All rights reserved.
PII: S1096-4959(02)00168-9

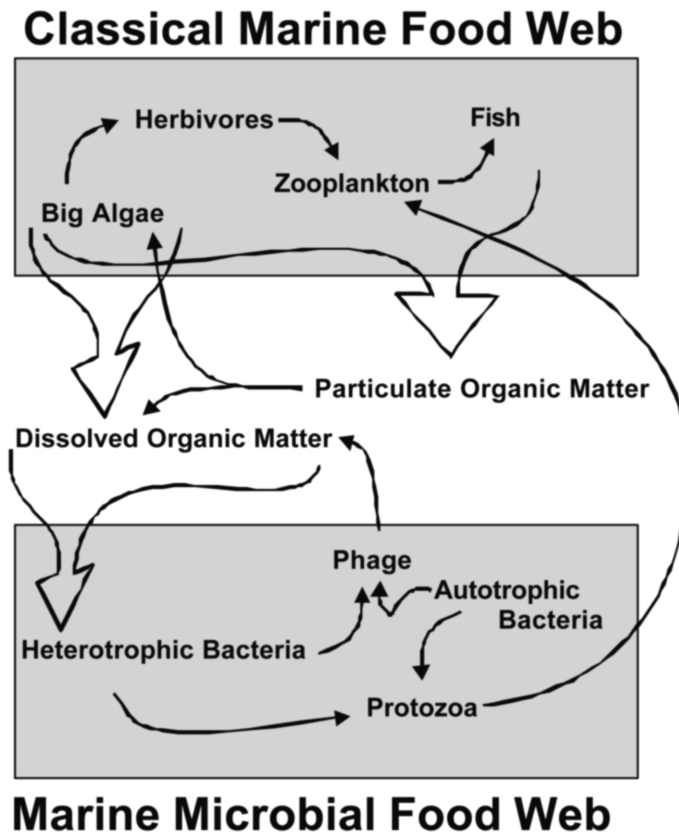

Fig. 1. Phage and the cycling of organic carbon matter through marine food webs. Arrows indicate direction of organic carbon flow.

tionary models and understanding genetic transduction within the environment.

2. Overview of marine phage ecology

2.1. Phage effects on carbon flow

The influences of phages on ecosystem dynamics are best understood in the marine environment. The marine microbial food web (MMFW) is the consortium of heterotrophic and autotrophic prokaryotes, as well as their predators, that inhabit the Earth's oceans and seas (Azam, 1998). The MMFW regulates the transfer of energy and nutrients to higher trophic levels and greatly influences global carbon (C) and nutrient cycles (Pomeroy, 1974; Azam et al., 1983). Dissolved organic matter (DOM) is the largest biogenic sink of carbon in the ocean (Kennish, 2001). Because the DOM pool is so large, heterotrophic bacterial populations are not resource limited; instead, they are controlled by predation (Fuhrman and Noble, 1995). The two predator guilds responsible for top-down control of the MMFW are the protozoa and phages (Fuhrman and Noble, 1995). In near-shore waters each of these predator guilds accounts for 50% of the microbial mortality each day (Fuhrman and Noble, 1995). To put the effects of these two bacterial predator guilds into perspective, ~49.3 Gt of C is fixed by phytoplankton per year in the world's oceans (Field et al., 1998), while global marine bacterial production is estimated to be 26–70 Gt of C per year (Wilhelm and Suttle, 1999). Therefore, the majority of the marine-biotic C is cycled into microbes and most of these microbes are killed by protozoa and phage predators.

When bacteria are eaten by protozoa, there is a possibility that the carbon can be transferred to the larger members of the marine food web (Fig. 1; Fukami et al., 1999). In contrast, when a bacterium is killed by a lytic phage, both the lysed host cell and the phage become part of the DOM pool (Middelboe et al., 1992, 1996). Since DOM is only utilized by other heterotrophic bacteria that

are also susceptible to lytic phages, this carbon never leaves the MMFW. The more rapidly this cycle repeats itself, the greater the amount of respiratory CO_2 that is produced, leaving less organic carbon stored in the world's oceans (Fig. 1). Thus phage activity may prove troublesome to proposed efforts to fertilize the oceans for increased carbon sequestration and fisheries yields.

In the open ocean, *Prochlorococcus* and *Synechococcus* are the numerically dominant autotrophs (reviewed in Waterbury et al., 1986; Partensky et al., 1999). Phages that infect these important primary producers have been isolated (Suttle, 1993; Suttle and Chan, 1993; Waterbury and Valois, 1993; Wilson et al., 1993; Sullivan, 2001). Although ecological studies of the impact of cyanophage on *Synechococcus* communities suggest that direct mortality is low relative to that observed for heterotrophic bacteria (reviewed in Suttle, 2000), the impact of cyanophage pressures on population structure and diversity of these systems may be significant (Waterbury and Valois, 1993). In near-shore communities, viruses have been isolated that infect the major 'large' phytoplankton species (e.g. eukaryotic diatoms and dinoflagellates). While viruses are not thought to be the major agent of cyanobacterial mortality, virus-induced mortality may be responsible for the 'sudden crashes' that terminate many blooms of eukaryotic algae (Sieburth et al., 1988; Bratbak, 1993; Nagasaki et al., 1993; Bratbak et al., 1995, 1998).

2.2. Transduction in marine environments

Besides their enormous influence on marine biogeochemistry, phages have important effects on genetic exchange in the marine environment (Jiang and Paul, 1998a). Phages can mediate DNA exchange between different bacteria by transduction, which occurs when host DNA is accidentally packaged into the phage during assembly (Masters, 1996; Weisberg, 1996). When the mispackaged phage infects another bacterium, instead of injecting phage DNA, it transfers DNA from its former host. Jiang and Paul (1998a) estimated that 1.3×10^{14} transduction events occur per year in the Tampa Bay Estuary, Florida. Extrapolation suggests that marine phages transduce 10^{28} base pairs of DNA per year in the world's oceans.

2.3. Lysogeny in marine environments

Not all phage-host encounters lead to host cell lysis, many rather result in lysogeny or pseudolysogeny (Ackermann and DuBow, 1987). Through meticulous work at the single cell level with *Bacillus*, lysogeny was first described as 'the hereditary power to produce bacteriophage' (Lwoff, 1953). This 'hereditary power' is due to the integration of invading phage DNA into the host cell genome (now termed a prophage) rather than proceeding through the lytic pathway. The prophage will remain integrated in the host cell genome until it is induced to 'abandon ship' and proceed through the lytic pathway. The molecular mechanism underlying prophage integration and excision are well understood in model systems (Hershey, 1971). In contrast, pseudolysogeny is a poorly understood phenomenon. It is often invoked to describe conditions where constant phage production occurs in the presence of a high abundance of host cells, thus allowing large numbers of host cells and their phage to coexist. Two mechanisms that might explain such observations are the following: (1) a mixture of sensitive and resistant host cells; or (2) a mixture of temperate and virulent phages (Williamson et al., 2001).

Lysogeny has been shown to improve the general fitness of the host (Edlin et al., 1975), largely from lysogenic conversion, or the expression of prophage-encoded genes. A common lysogenic conversion phenotype is immunity to superinfection (Hershey, 1971), but lysogenic conversion can also result in altered structural characteristics (Pruzzo and Satta, 1988; Vaca-Pacheco et al., 1999; Mirold et al., 2001), as well as resistance to antibiotics (Mlynarczyk et al., 1997) and reactive oxygen species (Figueroa-Bossi and Bossi, 1999). Of particular importance and global significance is the spread of toxin/virulence genes (often termed 'pathogenicity islands') by lysogenic conversion. Diphtheria, botulinum, cholera, pertussis, shiga and many other exotoxins are prophage encoded (reviewed in Davis and Waldor, 2002). Recent evidence suggests that these toxin genes can be transferred by transduction and other lateral gene transfer mechanisms (Boyd and Waldor, 1999; Faruque et al., 1999; Yaron et al., 2000).

Lysogeny is a common phenomenon in the marine environment. A recent study suggests that lysogeny in oligotrophic waters is common amongst cultivated bacteria, as 40% of 110 marine

bacterial isolates produced phage or bacteriocin-like particles upon treatment with an inducing agent (Jiang and Paul, 1998b). Efforts to quantify lysogeny in natural populations have resulted in a wide range of values for the proportion of the population lysogenized. For example, Weinbauer and Suttle (1996) found that in the Gulf of Mexico, 1.5–11.4% of the microbial population was lysogenized, whereas a detailed seasonal study of lysogeny in Tampa Bay indicated that the lysogenic fraction could range from 0 to 100% (average 27.6±37.1%; Williamson et al., 2002). During the seasonal study, lysogeny was primarily detected in winter months, consistent with the theory that lysogeny is favored in times of low host cell density. The environmental factors that lead to the control of lysogeny in the marine environment are largely unknown although links to nutrients such as phosphate have been suggested (Tuomi et al., 1995; Wilson et al., 1996, 1997). The molecular control of lysogeny in many phage host systems is complex (Ptashne, 1992; Friedman and Court, 2001) and usually involves genomic elements termed lysogeny modules (Lucchini et al., 1999). Essentially, nothing is known of the molecular or environmental control of these genomic elements in marine phages. In addition to investigations of bacterial lysogens in the environment, two recent studies have suggested that natural populations of the cyanobacterium *Synechococcus* can be lysogenized (McDaniel et al., 2002; Ortmann et al., 2002).

2.4. Microbial and phage diversity in the marine environment

Through their role as species-specific predators, phages may also help maintain microbial diversity. In the absence of host cell resistance and providing that contact rates remain high, lytic phages could potentially lyse all individuals of a species—thus phage attack can result in a rapid succession of microbial species (Thingstad and Lignell, 1997; Wommack and Colwell, 2000). Experimental evidence that phage exert a strong selective pressure on microbial populations comes from host-range analysis of phage isolates and the observation that very closely related bacterial species and even strains of the same species are infected by different phages (Moebus, 1991; Suttle and Chan, 1993, 1994; Waterbury and Valois, 1993).

The biodiversity of marine phage is essentially unknown. The few studies that have addressed this question suggest that diversity is high (reviewed in Borsheim, 1993). Moebus and colleagues (Moebus, 1991, 1992a,b Moebus and Nattkemper, 1991) screened over 900 isolates of culturable marine bacteria and found that approximately one-third were susceptible to at least one, and often multiple, lytic phages. Over the course of these studies, the authors concluded that: (1) the majority of bacterial strains were probably susceptible to phage infection; and (2) the phages isolated in these studies were specific to single hosts. Further work characterized a subset of these phages using DNA–DNA hybridizations, %GC, genome size estimations and host-range analysis and showed that they were genetically diverse (Wichels et al., 1998). Kellogg et al. (1995) isolated 60 phages from Florida and Hawaii that infected *Vibrio parahaemolyticus* and analyzed them by restriction fragment length polymorphism (RFLP), host specificity and Southern blotting. RFLP analysis separated the 60 phage isolates into six distinct groups that were then further genetically characterized through Southern blotting using a 1.5-kb DNA probe cloned from one of the isolates. Host range analysis showed that these phages were host-specific as none of the 60 isolates were able to infect closely related *Vibrio* species. Together these findings strongly suggest that phage diversity is at least as high, and probably higher, than the diversity of bacteria. That is, for each bacterium there is probably at least one, and often multiple, phages capable of infecting it and each phage is usually specific to a single microbial species or strain.

3. The current state of marine phage genomics

3.1. Cultured marine phage genomes

The first marine phage genome to be completely sequenced, *Pseudoaltermonas espejiana* BAL-31 phi PM2, was isolated off the coast of Chile in the 1960s (Espejo and Canelo, 1968). Phage PM2 was also the first lipid-containing phage ever isolated and it serves as the type phage for the International Committee of Viral Taxonomy (ICTV) family *Corticoviridae* (Murphy et al., 1995). The genome of phi PM2 is circular and 10 079 bp long that is replicated in a rolling circle fashion (Mannisto et al., 1999). The genes that encode structural and replication proteins have

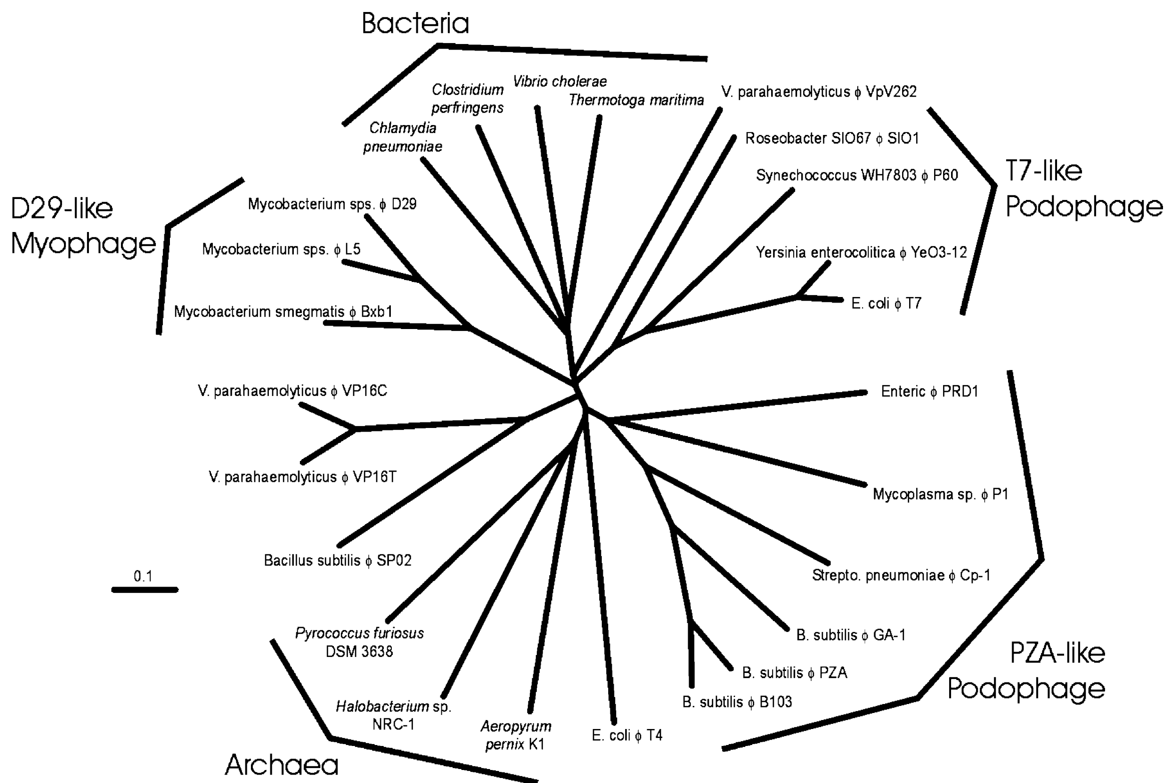

Fig. 2. Phylogenetic trees of DNA polymerase genes found in bacteria, Archaea and various phages of marine and terrestrial origin. Marine phage DNA polymerase sequences include *V. parahaemolyticus* ϕVpV262, ϕVP16C and ϕVP16T; *Roseobacter* SIO67 ϕSIO1 and Synechococcus WH7803 ϕP60.

been identified amongst the 42 potential open reading frames (ORFs) (Mannisto et al., 1999). The phi PM2 genome also contains a plasmid-like maintenance region, suggesting that the genome may be transferred among bacteria either as a plasmid or as a free phage (Mannisto et al., 1999). Essentially nothing is known about the ecology of phi PM2.

The second marine phage genome to be sequenced, *Roseobacter* SIO67 phi SIO1, was isolated off the Scripps Institution of Oceanography pier in 1990 (Rohwer et al., 2000). Phage SIO1 has a 39 906 bp dsDNA genome with 34 predicted ORFs (Rohwer et al., 2000). The phi SIO1 primase/helicase, DNA polymerase and endodeoxyribonuclease 1 share significant similarity to presumed homologs in *Escherichia coli* phi T3 and T7 (Rohwer et al., 2000). The *Synechococcus* WH7803 phi P60 genome (Chen and Lu, in press), discussed below also contains ORFs with significant BLAST hits to these genes. This suggests that there is a group of marine phage that uses a DNA replication mechanism much like phi T3 and T7 (Fig. 2). The ORFs that probably encode the phi SIO1 structural proteins, as suggested by their position in the genome, do not share significant similarity to other phage proteins currently in GenBank, but may be related to sequences found in the marine phages *Vibrio parahaemolyticus* phi VpV 262 (Hardies et al., 2002). The phi SIO1 genome also contains a number of tantalizing hints about its ecology. In particular, it appears the phage has linked its lifecycle to the phosphate metabolism of its host cell (see below).

The first cyanophage genome to be completely sequenced, *Synechococcus* WH7803 phi P60, has a 47 872 bp dsDNA genome that contains 80 potential ORFs (Chen and Lu, in press). The DNA replication genes are related to those of phi SIO1 and phi T7. Interestingly, the phi P60 DNA polymerase gene appears to be more closely related to those encoded by the non-marine phage T3, T7 and phi-YeO3-12 than to the marine phi SIO1

Table 1
The cultured marine phage that have completed genome sequences

Phage	Genome Size (kb) and morphology	Genome access and refs.
Pseudoalteromonas	10.1	NC 000867
espejiana phi PM2	Corticovirus	(Mannisto et al. 1999)
Roseobacter SIO67 phi	39.9	NC 002519
SIO1	Podovirus	(Rohwer et al. 2000)
Synechococcus WH7803	47.8	AF338467
phi P60	Podovirus	(Chen and Lu, in press)
Vibrio parahaemolyticus	49.7	
phi TB16T	Myovirus	
Vibrio parahaemolyticus	47.5	
phi TB16C	Myovirus	
Vibrio parahaemolyticus	45.9	http://biochem.uthscsa.edu/~hs_lab/phage.html
phi VpV 262	Podovirus	(Hardies et al., 2002)

Currently there are only three marine phage genomes available in GenBank. Three other genomes have been sequenced and are currently being annotated.

gene (Chen and Lu, in press). Phage P60 and phi SIO1 also encode a ribonucleotide reductase that is probably involved in recycling host rNTPs to dNTPs that can be incorporated into nascent phage genomes. Phage P60 encodes a RNA polymerase, which appears to be absent from the phi SIO1 genome. Since the RNA polymerase is essential to invasion and transcription of T7 (Calendar, 1988), phi P60 and phi SIO1 must have very different lifecycles.

Three marine phages that infect two strains of the human pathogen *Vibrio parahaemolyticus* have also been completely sequenced (Table 1). *Vibrio parahaemolyticus* phi VpV 262 is a Podovirus isolated from the Strait of Georgia, BC, Canada (Hardies and Serwer, 2002). The genome is 45 874 bp linear dsDNA genome with 73 predicted ORFs. The genome contains DNA polymerase, primase and helicase genes that appear to be more closely related to their bacterial homologues than to the phi T7 or phi SIO1 genes.

Two other Vibriophages, phi TB16T and phi TB16C, were separated from each other from a lysate of phage VP16, isolated from Tampa Bay (Kellogg et al., 1995). Although the original phage was thought to be a myovirus based on its contractile tail (Kellogg et al., 1995), the DNA sequence similarities and presence of a cos site suggest that the viruses are more likely to be siphoviruses (Segall and Rohwer, unpublished data). The two viruses are closely related (73–91%) over roughly 80% of their genomes, but differ by multiple unique insertions ranging between approximately 200 and 5000 bp in size. Each virus encodes 63-64 ORFs. The structural genes are clustered on the left side of the genome near the cos site and include a gene related to phage lambda's large terminase subunit, as well as genes related to those encoding tail and tail fiber proteins, portal, sheath and tape measure proteins of various phages and prophages. The rest of the genome includes ORFs similar to a DNA polymerase, a helicase, a polypeptide deformylase, and two genes weakly similar to transcriptional regulators. The two Vibriophages also contain numerous other genes similar to ORFs encoding hypothetical proteins from other phage, prophage and bacterial genomes. Although the original phage VP16 gave clear plaques, TB16T and TB16C were separated from each other based on their plaque morphology—TB16T gave turbid plaques with clear centers, whereas TB16C gave entirely clear plaques. Although lysogeny has not been proven, there is significant evidence from the genome sequence predicting that these phages have a temperate lifestyle—including the putative regulatory elements and the ability to circularize. Of great interest is that the sequences of the mixture of two genomes formed separate contigs during sequencing and assembly, despite substantial regions of identity between the two genomes (Rohwer and Segall, unpublished data). This suggests that even very closely related viruses can be sequenced from a mixed lysate, even when, as indicated by restriction digests, one genome makes up less than 5% of the DNA in the lysate. Subsequent sequencing of libraries made of DNA from the isolated phages showed that we did not

obtain any chimeric sequences between the two genomes.

As of this writing, the genomic sequencing of cyanophage S-PM2 is being finished (Nicholas H. Mann, personal communication). The genome is ~170 kb and includes several very large genes, including several that encode proteins >3000 amino acids in length. There are some rather surprising genes such as one that is homologous to a gene involved in complementary chromatic adaptation. When finished, this genome should be particularly exciting because it will be the first marine representative of the Myoviruses related to coliphage T4.

In addition to marine phages, a number of groups are starting to sequence marine viruses. *Emiliania huxleyi* is a marine coccolithophorid, with a world-wide distribution that forms vast coastal and mid-oceanic blooms (Holligan et al., 1993). Double stranded DNA viruses that infect *E. huxleyi* (EhV) have recently been isolated (Wilson et al., in press). Phylogenetic analysis of DNA polymerase gene fragments of these viruses suggests that EhVs belong to a new genus with the proposed name *Coccolithovirus*, within the family of algal viruses Phycodnaviridae (Schroeder et al., in press). Work is currently underway (shotgun sequencing completed, finishing begun) to sequence the 410-kb genome of one of these viruses with a completion date due in summer 2002 (William Wilson, personal communication). Tai et al. (2002) have also reported the complete sequence of a virus associated with lysis of the eukaryotic fish-killer *Heterosigma akashiwo*.

3.2. Prophage in completed marine bacterial genomes

In the near future, prophage contained in marine bacterial genomes will be an important data source when studying marine phage. The best studied prophage system in a marine bacterial genome is CTX-phi, the lysogenic phage that encodes the cholera toxin in pathogenic strains of *Vibrio cholerae* (Waldor and Mekalanos, 1996). CTX-phi is a temperate, filamentous phage that is secreted by the same extracellular protein secretion system used for cholera toxin production (Davis et al., 2000a). The phage uses a type IV pilus as a receptor that is encoded by an adjacent genetic element, the TCP gene cluster, itself a putative prophage (Karaolis et al., 1999). Three strains of CTX-phi have been characterized from classical, El Tor and Calcutta isolates of *V. cholerae*. Interestingly, the former strain encodes functional phage genes that are expressed, but lacks the ability to produce infectious virions due to their occurrence as single prophage elements (Davis et al., 2000b). In contrast, the latter two phage strains occur in tandem arrays of multiple prophages or single prophage plus prophage parts and these strains are capable of generating infectious virions (Davis and Waldor, 2000). The increasing numbers of publicly available microbial genomes have led to the discovery that prophage elements may exist in nearly every microbial genome (S. Casjens; personal communication). Putative prophage have been identified by inspecting regions of microbial chromosomes for the following characteristics: (1) genes possessing homology to known phage genes; (2) a contiguous group of genes containing few, if any, obviously non-phage genes; (3) the phage genes organized in a phage-like manner (e.g. integrase near the end, structural genes clustered, correctly ordered, etc.); and (4) 'unknown' genes in the putative prophage not obviously organized in a non-phage-like manner. Using this approach, we have identified putative prophages in the complete marine genomes of *Prochlorococcus* MED4, *Prochlorococcus* MIT 9313 and *Synechococcus* WH 8102 (http://www.jgi.doe.gov/). However, these regions were later shown not to be prophage regions as the phage genes were too few, were not organized in a phage-like manner and were interupted by obvious non-phage genes.

There is significant evidence that some prophages integrate into select regions of a host cell genome. Thirty-four of 58 cases of the integration of genetic elements (including prophage) were found to occur in attB sites within tRNA or tmRNA genes (Williams, 2002). Prophage integration at these sites might be beneficial because: (1) tRNA and tmRNA genes have ~four- to ninefold lower mutation rates than other protein encoding regions; and (2) these genes are small thus requiring a smaller region to be mimicked by the phage attP. Of particular interest to phage ecologists is the fact that tRNA promoters are known to be regulated by growth rate (Swenson et al., 1994)—in effect, allowing a prophage integrating into a tRNA gene to monitor the physiological state of its host through transcriptional coupling to the tRNA gene. It is unknown how integrase genes of genetic elements are able to recognize tRNA-like

elements, but Williams (2002) suggests that the definitive secondary structure of the DNA molecule may be involved.

To further aid in the identification of prophage in microbial genomes, we assessed the usefulness of DNA Structural Atlases (http://www.cbs.dtu.dk/services/GenomeAtlas/; Pedersen et al., 2000) as a diagnostic tool. These atlases display DNA structural characters, such as DNA curvature, DNA flexibility and DNA stability, in the form of a color-coded wheel that is useful for visually revealing interesting structural features of a genomic sequence (Pedersen et al., 2000). The five models used to predict these structural characters are based on either empirical data (e.g. DNase I sensitivity, X-ray crystallography data, trinucleotide preferences and gel mobility) or quantum mechanical calculations. To evaluate this tool for predicting prophage, we examined the DNA structural atlases of genomes containing well-characterized prophage elements to qualitatively look for diagnostic trends. Although there often appears to be significant structure within the prophage regions, this structure is not associated with all known prophage and often occurs throughout the genome in known non-prophage regions. However, factors such as the type of prophage and the length of time since its last activity might greatly affect these DNA structural properties and require a more intensive, quantitative analysis to identify diagnostic trends.

3.3. Uncultured marine phage genomes

Based upon pulsed-field gel electrophoresis of natural marine phage communities, we know that marine viral genomes fall into three size ranges: 35–40, 50–65 and 120–140 kb (Wommack et al., 1999; Steward et al., 2000). It has long been known that there is little similarity between the marine bacteria that have been cultivated and those phylotypes that are known to be prevalent as determined by 16S rDNA analyses (Fuhrman and Campbell, 1998). Therefore, cultured marine phage will most probably not be representative of the community. To circumvent the limitations imposed by culturing, a number of laboratories have started to sequence the genomes of total marine viral communities.

Breitbart et al. (in press) have constructed a shotgun library from an uncultured, near-shore marine viral community. Only 30% of the sequences from this library possessed appreciable similarity to those in the GenBank. Of the significant hits, 32% were phage in origin and 3% were most closely related to eukaryotic viruses. Among the phage genes showing similarity, Podovirus genes were most common (43%). Representatives of the Sipho- and Myoviridae were also found. These broad trends have now been observed in a second, near shore library (Breitbart and Rohwer, unpublished data). Another research group has partially sequenced a shotgun library made from a phage community from 70 m in Monterey Bay (Steward and Preston, unpublished data). As with the near-shore phage community, most of the sequences show no similarity to Genbank entries. Both libraries contained sequences with significant similarity to phage DNA polymerase genes, RNA polymerase, integrases, transposases and reverse transcriptases.

Breitbart et al. (in press) proposed a mathematical model that uses the number of observed contigs to predict phage richness and diversity in the sample. According to their calculations, the most abundant phage in the sample made up 4% of the population and phage diversity is very high (e.g. Shannon–Weaver index value of 7–8; Shannon and Weaver, 1963). This model also predicts that it is technically possible to sequence an entire marine viral community.

4. Uses of marine phage genomes

4.1. Classification of marine phages

A major goal of phage genomic sequencing projects should be to provide the information necessary to classify marine phage into guilds that reflect their biology. Current phage taxonomy relies on the morphological characteristics of the free phage particle as established by the International Committee on Taxonomy of Viruses (ICTV) (Murphy et al., 1995). The ICTV classification, however, provides very little information about the ecological niches or lifestyles of phage. Additionally, the ICTV system does not have sufficient resolution to address phage biodiversity questions, nor will it be useful for analyzing uncultured marine phage or prophage genomes. In response to these shortcomings, numerous groups are actively constructing phage taxonomical systems based on completed genomic sequences (Lawrence et al., 2002; Rohwer and Edwards, 2002). These

systems will help classify marine phages into families that provide information about their life-cycles and ecological roles, as well as identify phage types that deserve more detailed analyses.

Marine phage genomes are already helping to differentiate phages into operational taxonomic units (OTUs) that predict biological properties. For example, total genome analyses and individual DNA polymerase sequences show that Podoviruses belong to two OTUs with fundamentally different DNA replication mechanisms (Pecenkova and Paces, 1999; Chen and Lu, in press, Rohwer and Edwards, 2002). The first group, the phi T7-like Podophages, includes: phi T7; phi T3; phi P60; phi SIO1; and *Yersinia enterocolitica* phi Ye03-12 (Fig. 2; Rohwer and Edwards, 2002). This group of T7-like phages replicates their genomes by a primase/DNA polymerase mechanism (Ackermann and DuBow, 1987). A second group, the PZA-like Podophage includes: phi PRD1; *Bacillus subtilis* phi PZA; *B. subtilis* phi GA-1; *B. subtilis* phi B103; *Streptococcus pneumoniae* phi Cp-1; and *Mycoplasma* sp phi P1 (Rohwer and Edwards, 2002). The PZA-like Podophage replicate their DNA using a covalently linked 5′ terminal protein primer (Salas, 1991).

Completed phage genomes are also beginning to help identify conserved sequences to facilitate studies of phage evolutionary history, biodiversity and biogeography. Rohwer and Edwards (2002) have suggested that conserved sequences within phage groups be called 'signature genes'. It should be noted that as is often the case with studies of natural diversity, a growing database of more sequences and genomes will greatly facilitate the identification of novel phage taxa and signature genes.

4.2. Evolution of marine phage

Genomic data enable determination of evolutionary relationships between marine and non-marine phages (Fuller et al., 1998; Rohwer et al., 2000; Hambly et al., 2001). Additionally, since the marine environment probably represents the largest and oldest biosphere on the planet, vital clues to the origin of phages may reside in the sequences of marine phage. Fuller et al. (1998) first proposed specific evolutionary relationships between marine and non-marine phages by sequencing regions of structural proteins in *Escherichia coli* phi T4-like phage. Hambly et al. (2001) has also sequenced the entire region homologous to gp18-23 in phi T4-like phage and showed that it is conserved in the marine phage *Synechococcus cyanophage* S-PM2. The podophage phi SIO1 and phi P60, as well as data from the uncultured phage libraries, suggest that there is a large group of marine phage encoding DNA replication machinery closely related to that of T7 (Fig. 2; Rohwer et al., 2000; Rohwer and Edwards, 2002; Chen and Lu, in press). Breitbart et al. (in press) have proposed that this group of phages is numerically dominant in the world's oceans.

Using analyses similar to those employed for enteric and dairy phages, Hardies et al. (2002) proposed that the phi VpV262 genome contains an identifiable moron ('more DNA'; Hendrix et al., 2000; Juhala et al., 2000; Hardies et al., 2002). By comparing codon usage preferences, these researchers have suggested that phi VpV262 genes are in equilibrium with each other, but not with the host (Hardies et al., 2002). This observation suggests that this phage may have a broader host range than expected, extending beyond *V. parahaemolyticus*. This type of analysis will be useful for determining phylogenetic relationships among marine and non-marine phage (Blaisdell et al., 1996).

4.3. Biogeography of marine phages

Genomic sequence information enables the construction of primers and probes to detect specific phage in the environment. The *Roseobacter* SIO67 phi SIO1 genomic sequence, for example, was used to design specific primers that could detect ~10 phage in a sample. These primers have been used to show that phi SIO1 is present in the waters around Scripps Pier most of the year and that the phage population rapidly increases during *Lingulodinium polyhedrum* blooms (Breitbart, Deyanat-Yazdi, Rohwer, unpublished data). These specific primers have also been used to rapidly differentiate between phages that infect *Roseobacter* SIO67 (Rohwer et al., 2000).

To examine *Synechococcus* cyanophage, Zhong et al. (2002) redesigned the PCR primers of Fuller et al. (1998) to specifically amplify a larger region (592 bp) of the g20 homologue of marine cyanophage for use in phylogenetic and biogeographic studies. Analysis of g20 sequences from cyanophage isolates revealed that: (1) the isolates were highly diverse yet more closely related to each

other than to enteric coliphage T4; and (2) there was no correlation between genetic variation of the clones and geographic location. Analysis of g20 sequences from six clone libraries of natural virus concentrates revealed that: (1) six of nine phylogenetic clusters represented novel uncultured g20 sequences; and (2) the phylogenetic composition of the cloned sequences from varying environments and depths were different from each other. All of these results indicate a high genetic diversity of marine cyanophage assemblages.

In an attempt to understand the diversity and biogeography of viruses infecting eukaryotic algae Chen et al. (1996) designed PCR primers to selectively amplify part of the DNA polymerase genes from viruses that infect two eukaryotic algae, an endosymbiont *Chlorella*-like alga and *Micromonas pusilla*. These primers were used to amplify sequences from environmental samples for phylogenetic analyses and to examine biodiversity using denaturing gradient gel electrophoresis (DGGE) (Short and Suttle, 1999, 2000, 2002; Short et al., 2000).

4.4. Prediction of the ecological niches of marine phages

Phages are 50% DNA by weight, which means they require a high proportion of phosphate. Since phosphate is often limiting in the marine environment (Bjorkman et al., 2000; Cavender-Bares et al., 2001), phosphate concentrations may limit phage production. If we consider that the average burst size in the marine environment is $\sim$50 virions (Borsheim, 1993; Wommack and Colwell, 2000) and that the average phage genome is $\sim$50 kb (Steward et al., 2000), then a typical lytic cycle requires the production of $\sim$2.5 Mb of phage DNA. This is roughly equivalent to the genome size of one marine bacterium (calculated to be 2.3 Mb from Simon and Azam, 1989). Tantalizing hints of the importance of phosphate in the life of marine phages from the available genomes include the occurrence of genes involved in recycling or scavenging more phosphate (e.g. ribonucleotide reductases, phoH, thymidine synthetases, endo- and exo-nucleases) (Wikner et al., 1993).

5. Technical challenges associated with sequencing marine phage genomes

The key to any genomic sequencing project is a high coverage shotgun library. This can be more problematic for marine phage than for 'typical' prokaryotic sequencing projects. The foremost challenge is obtaining a sufficient amount of DNA. Because the majority of marine bacterial hosts grow slowly and to lower densities than non-marine bacteria, typical yields of phage DNA are in the ng range, compared to the micro-gram quantities of DNA used in typical shotgun library protocols.

A second problem that may be encountered when sequencing marine phage is unclonable DNA. Phages often modify their genomic DNA to avoid host restriction systems or to target their DNA for activity by specialized phage-encoded enzymes. *V. parahaemolyticus* phi TB16T and phi TB16C could not be cloned using standard approaches (e.g. enzymatic digestion and cloning) (Rohwer et al., 2001). The extent of this phenomenon in other marine phage is unknown.

To circumvent problems associated with both limiting amounts and modified DNA, alternative shotgun cloning protocols have been developed. Random amplified shotgun libraries (RASLs) are constructed by first amplifying the DNA with random 10-mer oligonucleotides as primers (Rohwer et al., 2001). The resulting products are then blunt end digested and cloned. Using the RASL method, shotgun libraries sufficient to sequence phage-sized genomes can be constructed from $\sim$20 ng of initial DNA. Recently, Breitbart et al. (in press) have used a second method for constructing high coverage shotgun libraries, called Linker amplified shotgun libraries (LASLs), from uncultured marine viral communities. To make a LASL, the DNA is physically broken into 2-kb fragments using a Hydroshear. The fragments are then end-repaired and asymmetrical linkers are ligated to the fragment ends. Primers to the linkers are then used to PCR-amplify the products before they are cloned. Using the LASL protocol, it is possible to construct libraries containing a million clones from $<$10 ng of initial DNA. Both RASL and LASL protocols have been shown to generate essentially random coverage without evidence of chimeric molecules.

Closing viral genomes after the initial shotgun sequencing phase is also complicated by limited amounts of DNA. One way to use less DNA than direct sequencing is to make primers to the ends of all the available contigs and perform PCR with the mixture. This approach was used when closing phi TB16C, phi TB16T, phi VpV 265 and phi

SIO1, as well as with bacterial genomes (Rohwer et al., 2000; Hardies et al., 2002; Tettelin et al., 1999). Limiting DNA also makes it almost impossible to directly sequence the ends of linear phage genomes. Cloning large pieces of phage genomes to help with closure will probably not be successful because phage genes like holins and lysozymes are usually lethal to *E. coli.*

6. The future

The field of marine phage genomics is in its infancy. Many more marine phage genomes are 'in the pipeline' for sequencing. Due to their small size, 100 phage genomes can be sequenced for the same cost as one large bacterial genome. Thus, the study of phage genomes is particularly economical. Moreover, phage genomes are easier to understand—their small size makes it practical to model phage lifecycles in the mind or on a computer. Combined with our increasingly detailed knowledge of the marine microbial food web, marine phage should be leading the effort to understand how a community of organisms and the environment interact with each other at the genomic level. Just as a reductionist's view of phage biology led to significant advances in the field of molecular biology, it is reasonable to expect that a reductionist's view will prove invaluable to our understanding of complex natural microbial systems. The ability to produce a large number of genomes and extract a wealth of useful and predictive information from them is a compelling reason for the phage field to be leading the way into massive comparative genomics.

Acknowledgments

The authors would like to thank phage and virus researchers who provided unpublished data for this review—Grieg Steward and Christina M. Preston (uncultured phage library), Curtis Suttle and Stephen C. Hardies (*Vibrio parahaemolyticus* phi VpV 262), William Wilson (*Emiliania huxleyi* coccolithovirus) and Nicholas H. Mann (cyanomyovirus S-PM2). Part of this work was sponsored by: NSF SGER OCE01-16900 to FR and Farooq Azam; an NSF grant to JHP; and an NSF Biocomplexity award to JHP and AMS.

References

Ackermann, H.-W., DuBow, M.S., 1987. Viruses of prokaryotes. General Properties of Bacteriophages, Vol. Volume I. CRC Press Inc, Boca Raton.

Azam, F., 1998. Microbial control of oceanic carbon flux: the plot thickens. Science 280, 694–696.

Azam, F., Fenchel, T., Field, J.G., Gray, J.S., Meyer-Reil, L.A., Thingstad, F., 1983. The ecological role of water-column microbes in the sea. Mar. Ecol. Prog. Ser. 10, 257–263.

Bergh, Ø., Borsheim, K.Y., Bratbak, G., Heldal, M., 1989. High abundance of viruses found in aquatic environments. Nature 340, 467–468.

Bjorkman, K., Thomson-Bulldis, A.L., Karl, D.M., 2000. Phosphorus dynamics in the North Pacific subtropical gyre. Aquat. Microb. Ecol. 22, 185–198.

Blaisdell, B.E., Campbell, A.M., Karlin, S., 1996. Similarities and dissimilarities of phage genomes. Proc. Nat. Acad. Sci. USA 93, 5854–5859.

Borsheim, K., 1993. Native marine bacteriophages. FEMS Microbiol. Ecol. 102, 141–159.

Boyd, E.F., Waldor, M.K., 1999. Alternative mechanism of cholera toxin acquisition by *Vibrio* cholerae: Generalized transduction of CTXPHI by bacteriophage CP-T1. Infec. Immun. 67, 5898–5905.

Bratbak, G., 1993. Viral mortality of the marine alga *Emiliania huxleyi* (Haptophyceae) and termination of algal blooms. Mar. Ecol. Prog. Ser. 93, 39–48.

Bratbak, G., Jacobsen, A., Heldal, M., 1998. Viral lysis of *Phaeocystis pouchetii* and bacterial secondary production. Aquat. Microb. Ecol. 16, 11–16.

Bratbak, G., Levasseur, M., Michaud, S., et al., 1995. Viral activity in relation to *Emiliania huxleyi* blooms: A mechanism of DMSP release? Mar. Ecol. Prog. Ser. 128, 133–142.

Breitbart, M., Segall, A., Mead, D., Azam, F., Rohwer, F., in press. Genomic analysis of an uncultivated viral community.

Calendar, R. (Ed.), 1988. The Bacteriophages. Plenum Press, New York.

Cavender-Bares, K.K., Karl, D.M., Chisholm, S.W., 2001. Nutrient gradients in the western North Atlantic Ocean: Relationship to microbial community structure and comparison to patterns in the Pacific Ocean. Deep-Sea Res. Part I Oceanog. Res. Papers 48, 2373–2395.

Chen, F., Lu, J., in press. Genomic sequence and evolution of marine cyanophage P60: a new insight on lytic and lysogenic phages. Appl. Env. Microbiol. 68.

Chen, F., Suttle, C.A., Short, S.M., 1996. Genetic diversity in marine algal virus communities as revealed by sequence analysis of DNA polymerase genes. Appl. Env. Microbiol. 62, 2869–2874.

Davis, B.M., Lawson, E.H., Sandkvist, M., Ali, A., Sozhamannan, S., Waldor, M.K., 2000a. Convergence of the secretory pathways for cholera toxin and the filamentous phage, CTXphi. Science 288, 333–335.

Davis, B.M., Moyer, K.E., Boyd, E.F., Waldor, M.K., 2000b. CTX prophages in classical biotype Vibrio cholerae: functional phage genes but dysfunctional phage genomes. J. Bacteriol. 182, 6992–6998.

Davis, B.M., Waldor, M.K., 2000. CTXphi contains a hybrid genome derived from tandemly integrated elements. Proc. Natl. Acad. Sci. USA 97, 8572–8577.

Davis, B.M., Waldor, M.K., 2002. Mobile genetic elements and bacterial pathogenesis. In: Craig, N.L. (Ed.), Mobile DNA II. ASM Press, Washington DC.

Edlin, G., Lin, L., Bitner, R., 1975. Lambda lysogens of *Escherichia coli* reproduce more rapidly than non-lysogens. Nature 255, 735–737.

Espejo, R.T., Canelo, E.S., 1968. Properties of bacteriophage PM2: a lipid-containing bacterial virus. Virology 34, 738–747.

Faruque, S.M., Rahman, M.M., Asadulghani, Islam, K.M.N., Mekalanos, J.J., 1999. Lysogenic conversion of environmental *Vibrio mimicus* strains by CTXPHI. Infect. Immun. 67, 5723–5729.

Field, C.B., Behrenfeld, M.J., Randerson, J.T., Falkowski, P., 1998. Primary production of the biosphere: integrating terrestrial and oceanic components. Science 281, 237–240.

Figueroa-Bossi, N., Bossi, L., 1999. Inducible prophages contribute to *Salmonella* virulence in mice. Mol. Microbiol. 33, 167–176.

Friedman, D.I., Court, D.L., 2001. Bacteriophage lambda: Alive and well and still doing its thing. Curr. Opin. Microbiol. 4, 201–207.

Fuhrman, J.A., 1999. Marine viruses: biogeochemical and ecological effects. Nature 399, 541–548.

Fuhrman, J.A., Campbell, L., 1998. Microbial microdiversity. Nature 393, 410–411.

Fuhrman, J.A., Noble, R.T., 1995. Viruses and protists cause similar bacterial mortality in coastal seawater. Limnol. Oceanog. 40, 1236–1242.

Fukami, K., Watanabe, A., Fujita, S., Yamaoka, K., Nishijima, T., 1999. Predation on naked protozoan microzooplankton by fish larvae. Mar. Ecol. Prog. Ser. 197, 285–291.

Fuller, N.J., Wilson, W.H., Joint, I.R., Mann, N.H., 1998. Occurence of a sequence in marine cyanophages similar to that of T4 g20 and its application to PCR-based detection and quantification techniques. Appl. Env. Microbiol. 64, 2051–2060.

Hambly, E., Tetart, F., Desplats, C., Wilson, W.H., Krisch, H.M., Mann, N.H., 2001. A conserved genetic module that encodes the major virion components in both the coliphage T4 and the marine cyanophage S-PM2. Proc. Nat. Acad. Sci. USA 98, 11411–11416.

Hardies, S.C., Comeau, A.M., Serwer, P., Suttle, C.A., 2002. Genome sequence of a novel marine phage infecting *Vibrio parahaemolyticus*. American Society of Limnology and Oceanography, Honolulu.

Hardies, S.C., Serwer, P., 2002. Bacteriophage sequence and evolution <http://biochem.uthscsa.edu/~hs_lab/phage.html>.

Hendrix, R.W., Lawrence, J.G., Hatfull, G.F., Casjens, S., 2000. The origins and ongoing evolution of viruses. Trends Microbiol. 8, 499–500.

Hershey, A.D. (Ed.), 1971. The Bacteriophage Lambda. Cold Spring Harbour Laboratory, Cold Spring Harbour, NY.

Holligan, P.M., Fernandez, E., Aiken, J., et al., 1993. A biogeochemical study of the Coccolithophore, *Emiliania huxleyi*, in the North-Atlantic. Global Biogeochem. Cycles 7, 879–900.

Jiang, S.C., Paul, J.H., 1998. Gene transfer by transduction in the marine environment. Appl. Env. Microbiol. 64, 2780–2787.

Jiang, S.C., Paul, J.H., 1998. Significance of lysogeny in the marine environment: studies with isolates and a model for viral production. Microbial Ecol. 35, 235–243.

Juhala, R.J., Ford, M.E., Duda, R.L., Youlton, A., Hatfull, G.F., Hendrix, R.W., 2000. Genomic sequences of bacteriophages HK97 and HK022: pervasive genetic mosaicism in the lambdoid bacteriophages. J. Mol. Biol. 299, 27–51.

Karaolis, D.K.R., Somara, S., Maneval, D.R., Johnson, J.A., Kaper, J.B., 1999. A bacteriophage encoding a pathogenicity island, a type-IV pilus and a phage receptor in cholera bacteria. Nature 399, 375–379.

Kellogg, C.A., Rose, J.B., Jiang, S.C., Turmond, J.M., Paul, J.H., 1995. Genetic diversity of related vibriophages isolated from marine environments around Florida and Hawaii, USA. Mar. Ecol. Prog. Ser. 120, 89–98.

Kennish, M.J., 2001. Practical Handbook of Marine Science. CRC Press, Boca Raton, FL.

Lawrence, J.G., Hatfull, G.F., Hendrix, R.W., 2002. The imbroglios of viral taxonomy: genetic exchange and the failings of phenetic approaches. J. Bacteriol. submitted for publication.

Lucchini, S., Desiere, F., Brussow, H., 1999. Similarly organized lysogeny modules in temperate Siphoviridae from low GC content gram-positive bacteria. Virology 263, 427–435.

Lwoff, A., 1953. Lysogeny Bacteriol. Rev. 17, 269–337.

Mannisto, R.H., Kivela, H.M., Paulin, L., Bamford, D.H., Bamford, J.K.H., 1999. The complete genome sequence of PM2, the first lipid-containing bacterial virus to be isolated. Virology 262, 355–363.

Masters, M., 1996. Generalized trasnduction. In: Neidhardt, F.H.C. (Ed.), *E. coli* and *Salmonella*: Cellular and Molecular Biology. ASM Press, Washington DC.

McDaniel, L., Houchin, L., Williamson, S., Paul, J.H., 2002. Lysogeny in natural *Synechococcus* populations. Nature. 415, 496.

Middelboe, M., Jørgensen, N.O.G., Kroer, N., 1996. Effects of viruses on nutrient turnover and growth efficiency of noninfected marine bacterioplankton. Appl. Environ. Microbiol. 62, 1991–1997.

Middelboe, M., Nielsen, B., Søndergaard, M., 1992. Bacterial utilization of dissolved organic carbon (DOC) in coastal waters—determination of growth yield. Arch. Hydrobiol. Ergeb. Limnol. 37, 51–61.

Mirold, S., Rabsch, W., Tschaepe, H., Hardt, W.-D., 2001. Transfer of the *Salmonella* type III effector sopE between unrelated phage families. J. Mol. Biol. 312, 7–16.

Mlynarczyk, G., Mlynarcyk, A., Zabicka, D., Jeljaszewicz, J., 1997. Lysogenic conversion as a factor influencing the vancomycin tolerance phenomenon in *Staphylococcus aureus*. J. Antimicrob. Chemother. 40, 136.

Moebus, K., 1991. Preliminary observations on the concentration of marine bacteriophages in the water around Helgoland. Helgoländer Meeresunters 45, 411–422.

Moebus, K., 1992. Further investigation on the concentration of marine bacteriophages in the water around Helgoland with reference to the phage–host systems encountered. Helgolaender Meeresunters 46, 275–292.

Moebus, K., 1992. Laboratory investigations on the survival of marine bacteriophages in raw and treated seawater. Helgoländer Meeresunters 46, 251–273.

Moebus, K., Nattkemper, H., 1991. Bacteriophage sensitivity patterns among bacteria isolated from marine waters. Helgolander Meeresunters 34, 375–385.

Murphy, F.A., Fauquet, C.M., Bishop, D.H.L.et al. (Eds.), 1995. Virus Taxonomy: Sixth Report of the International

Committee on Taxonomy of Viruses. Springer–Verlag, New York,. Supplement 10.

Nagasaki, K., Ando, M., Imai, I., Itakura, S., Ishida, Y., 1993. Virus-like particles in an apochlorotic flagellate in Hiroshima Bay, Japan. Mar. Ecol. Prog. Ser. 96, 307–310.

Ortmann, A.C., Lawrence, J.E., Suttle, C.A., in press. Lysogeny and lytic viral production during a bloom of the cyanobacterium *Synechococcus* spp. Microb. Ecol.

Partensky, F., Blanchot, J., Vaulot, D., 1999. Differential distribution and ecology of *Prochlorococcus* and *Synechococcus* in oceanic waters: A review. Bulletin de l'Institut Oceanographique (Monaco) 10, 457–475.

Paul, J.H., 1999. Microbial gene transfer: an ecological perspective. J. Mol. Microbiol. Biotechnol. 1, 45–50.

Pecenkova, T., Paces, V., 1999. Molecular phylogeny of phi29-like phages and their evolutionary relatedness to other protein-primed replicating phages and other phages hosted by gram-positive bacteria. J. Mol. Evol. 48, 197–208.

Pedersen, A.G., Jensen, L.J., Brunak, S., Staerfeldt, H.-H., Ussery, D.W., 2000. A DNA structural atlas for *Escherichia coli*. J. Mol. Biol. 299, 907–930.

Pomeroy, L.R., 1974. The ocean's food web, a changing paradigm. BioScience 24, 499–504.

Pruzzo, C., Satta, G., 1988. Capsular antigenic variations by lysogenic conversion in klebsiella–pneumoniae relationship with virulence. Curr. Microbiol. 16, 259–264.

Ptashne, M., 1992. A Genetic Switch. Blackwell Scientific Publications, Cambridge, MA.

Rohwer, F., Edwards, R., in press. The phage proteomic tree: a genome based taxonomy for phage. J. Bacteriol.

Rohwer, F., Segall, A., Steward, G., Seguritan, V., Breitbart, M., Wolven, F., Azam, F., 2000. The complete genomic sequence of the marine phage Roseophage SIO1 shares homology with non-marine phages. Limnol. Oceanog. 42, 408–418.

Rohwer, F., Seguritan, V., Choi, D.H., Segall, A.M., Azam, F., 2001. Production of shotgun libraries using random amplification. BioTechniques 31, 108–118.

Salas, M., 1991. Protein-Priming of DNA Replication. In: Salas, M. (Ed.). Ann. Rev. Biochem. 60, Annual Reviews Inc., Palo Alto, CA, 39–72.

Sanger, F., Air, G.M., Barrell, B.G., et al., 1977. Nucleotide sequence of bacteriophage phi X174 DNA. Nature 265, 687–695.

Schroeder, D., Oke, J., Malin, G., Wilson, W.H., in press. Coccolithovirus (Phycodnaviridae), characterisation of a new algal virus genus that infects *Emiliania huxleyi*. Arch. Virol.

Shannon, C.E., Weaver, W., 1963. The Mathematical Theory of Communication. University of Illinois Press, Urbana.

Short, S.M., Frederickson, C.M., Suttle, C.A., 2000. DGGE analysis of DNA fragments amplified from viruses that infect eukaryotic algae and cyanobacteria. Abstracts Gen. Meeting Am. Soc. Microbiol. 100, 465.

Short, S.M., Suttle, C.A., 1999. Use of the polymerase chain reaction and denaturing gradient gel electrophoresis to study diversity in natural virus communities. Hydrobiologia 401, 19–32.

Short, S.M., Suttle, C.A., 2000. Denaturing gradient gel electrophoresis resolves virus sequences amplified with degenerate primers. BioTechniques 28, 20–26.

Short, S.M., Suttle, C.A., 2002. Sequence analysis of marine cirus communities reveals that groups of related algal viruses are widely distributed in nature. Appl. Env. Microbiol. 68, 1290–1296.

Sieburth, J.M., Johnson, P.W., Hargraves, P.E., 1988. Ultrastructure and ecology of *Aureococcus anophagefferens* gen. et sp. nov. (Chrysophyceae) and the dominant picoplankter during a bloom in Narragansett Bay Rhode Island USA summer 1985. J. Phycol. 24, 416–425.

Simon, M., Azam, F., 1989. Protein content and protein synthesis rates of planktonic marine bacteria. Mar. Ecol. Prog. Ser. 51, 201–213.

Steward, G.F., Montiel, J.L., Azam, F., 2000. Genome size distributions indicate variability and similarities among marine viral assemblages from diverse environments. Limnol. Oceanog. 45, 1697–1706.

Sullivan, M.B., 2001. *Prochlorococcus* cyanophage: isolation, preliminary characterization and natural abundances. In: Sullivan, M.B. (Ed.), Biology. MAMIT/WHOI, Cambridge MA, p. 129.

Suttle, C.A., 1993. Enumeration and isolation of viruses. In: Suttle, C.A. (Ed.), Current Methods in Aquatic Microbiology. Lewis Publ, Chelsea.

Suttle, C.A., 2000. Cyanophages and their role in the ecology of cyanobacteria. In: Suttle, C.A. (Ed.), The Ecology of Cyanobacteria. Kluwer Academic Publishers, Netherlands, pp. 563–589.

Suttle, C.A., Chan, A.M., 1993. Marine cyanophages infecting oceanic and coastal strains of *Synechococcus*-abundance, morphology, cross-infectivity and growth characteristics. Mar. Ecol. Prog. Ser. 92, 99–109.

Suttle, C.A., Chan, A.M., 1994. Dynamics and distribution of cyanophages and their effect on marine *Synechoccus* spp. Appl. Env. Microbiol. 60, 3167–3174.

Swenson, D.L., Kim, K.J., Six, E.W., Clegg, S., 1994. The gene fimU affects expression of *Salmonella typhimurium* type 1 fimbriae and is related to the *Escherichia coli* tRNA gene argU. Mol. Gen. Genet. 244, 216–218.

Tai, V., Lang, A.S., Suttle, C.A., 2002. Genomic analysis of a novel marine virus associated with lysis of the toxic bloom former, *Heterosigma akashiwo*. American Society of Limnology and Oceanography, Honolulu, HI.Abst. OS11I-04

Tettelin, H., Radune, D., Kasif, S., Khouri, H., Salzberg, S.L., 1999. Optimized multiplex PCR: Efficiently closing a whole-genome shotgun sequencing project. Genomics 62, 500–507.

Thingstad, T.F., Lignell, R., 1997. Theoretical models for the control of bacterial growth rate, abundance, diversity and carbon demand. Aquat. Microb. Ecol. 13, 19–27.

Tuomi, P., Fagerbakke, K.M., Bratbak, G., Heldal, M., 1995. Nutritional enrichment of a microbial community: the effects on activity, elemental composition, community structure and virus production. FEMS Microb. Ecol. 16, 123–134.

Vaca-Pacheco, S., Paniagua-Contreras, G.L., Garcia-Gonzalez, O., de la Garza, M., 1999. The clinically isolated FIZ15 bacteriophage causes lysogenic conversion in Pseudomonas aeruginosa PAO1. Curr. Microbiol. 38, 239–243.

Waldor, M.K., Mekalanos, J.J., 1996. Lysogenic conversion by a filamentous phage encoding Cholera toxin. Science 272, 1910–1914.

Waterbury, J.B., Valois, F.W., 1993. Resistance to co-occurring phages enables marine *Synechococcus* communities to coex-

ist with cyanophages abundant in seawater. Appl. Environ. Microbiol. 59, 3393–3399.

Waterbury, J.B., Watson, S.W., Valois, F.W., Franks, D.G., 1986. Biological and ecological characterization of the marine unicellular cyanobacterium *Synechococcus*. Can. Bull. Fish. Aquat. Sci. 214, 71–120.

Weinbauer, M.G., Suttle, C.A., 1996. Potential significance of lysogeny to bacteriophage production and bacterial mortality in coastal waters of the Gulf of Mexico. Appl. Environ. Microbiol. 62, 4375.

Weisberg, R.A., 1996. Specialized transduction. In: Neidhardt, F.H.C. (Ed.), *E. Coli* and *Salmonella*: Cellular and Molecular Biology. ASM Press, Washington DC.

Wichels, A., Biel, S.S., Gelderblom, H.R., Brinkhoff, T., Muyzer, G., Schuett, C., 1998. Bacteriophage diversity in the North Sea. Appl. Env. Microbiol. 64, 4128–4133.

Wikner, J., Vallino, J.J., Steward, G.F., Smith, D.C., Azam, F., 1993. Nucleic acids from the host bacterium as a major source of nucleotides for three marine bacteriophages. FEMS Microbiol. Ecol. 12, 237–248.

Wilhelm, S.W., Suttle, C.A., 1999. Viruses and nutrient cycles in the sea. BioScience 49, 781–783.

Williams, K.P., 2002. Integration sites for genetic elements in prokaryotic tRNA and tmRNA genes: sublocation preference of integrase subfamilies. Nucl. Acids Res. 30, 866–875.

Williamson, S.J., McLaughlin, M.R., Paul, J.H., 2001. Interaction of the PhiHSIC virus with its host: lysogeny or pseudolysogeny? Appl. Environ. Microbiol. 67, 1682–1688.

Williamson, S.J., Houchin, L., McDaniel, L., Paul, J.H., 2002. Seasonal variation in lysogeny as depicted by phophage induction in Tampa Bay, Florida. Appl. Environ. Microbiol. 68, 4307–4314.

Wilson, W.H., Carr, N.G., Mann, N.H., 1996. The effect of phosphate status on the kinetics of cyanophage infection in the oceanic cyanobacterium *Synechococcus* sp. WH7803. J. Phycol. 32, 506–516.

Wilson, W.H., Joint, I.R., Carr, N.G., Mann, N.H., 1993. Isolation and molecular characterization of five marine cyanophages propagated on *Synechococcus* sp. strain WH 7803. Appl. Environ. Microbiol. 59, 3736–3743.

Wilson, W.H., Tarran, G.A., Schroeder, D., Cox, M., Oke, J., Malin, G., in press. Isolation of viruses responsible for the demise of an *Emiliania huxleyi* bloom in the English Channel. J. Marine Biol. Assoc. UK.

Wilson, W.H., Turner, S., Mann, N.H., 1997. Population dynamics of phytoplankton and viruses in a phosphate limited mesocosm and their effect on DMSP and DMS production. Estuarine Coastal Shelf Sci. 46, 49–59. (Suppl. A).

Wommack, K.E., Colwell, R.R., 2000. Virioplankton: viruses in aquatic ecosystems. Microbiol. Mol. Biol. Rev. 64, 69–114.

Wommack, K.E., Ravel, J., Hill, R.T., Chun, J.S., Colwell, R.R., 1999. Population dynamics of Chesapeake Bay virioplankton: total community analysis by pulsed-field gel electrophoresis. Appl. Environ. Microbiol. 65, 231–240.

Yaron, S., Kolling, G.L., Simon, L., Matthews, K.R., 2000. Vesicle-mediated transfer of virulence genes from *Escherichia coli* O157:H7 to other enteric bacteria. Appl. Env. Microbiol. 66, 4414–4420.

Zhong, Y., Chen, F., Wilhelm, S.W., Poorvin, L., Hodson, R.E., 2002. Phylogenetic diversity of marine cyanophage isolates and natural virus communities as revealed by sequences of viral capsid assembly protein gene g20. Appl. Env. Microbiol. 68, 1576–1584.