

## MINI-REVIEW

# Prophages in marine bacteria: dangerous molecular time bombs or the key to survival in the seas?

John H Paul

College of Marine Science, University of South Florida, St Petersburg, FL, USA

**Bacteriophages are realized to be numerous and important components of oceanic food webs principally because of their lytic capabilities. The subtle changes that temperate phages impart to their hosts in the oceans are far less understood. Occurrences of lysogeny in the oceans correlate well with conditions unfavorable for rapid host growth. In coliphage  $\lambda$ , phage encoded repressors have been shown to modulate host metabolic gene expression and phenotype, resulting in economizing host energy expenditure. Comparison of lysogenized marine bacteria to the uninfected hosts indicated that prophage acquisition is correlated with host metabolic gene suppression. Screening 113 marine bacterial genomes for prophages yielded 64 prophage-like elements, 21 of which strongly resembled gene transfer agents (GTAs). The remaining 39 putative prophages had a relatively high incidence of transcriptional regulatory and repressor-like proteins (~2/40 kb prophage sequence) compared to lytic marine phages (~0.25/40 kb phage sequence). Here, it has been hypothesized that marine prophages directly contribute to host survival in unfavorable environments by suppression of unneeded metabolic activities. It has been further suggested that such metabolic downshifts are the result of phage-encoded repressors and transcriptional regulators acting directly on host genes. Finally, the widespread occurrence of GTAs may be an efficient mechanism for horizontal gene transfer in the oceans.**

*The ISME Journal* (2008) 2, 579–589; doi:10.1038/ismej.2008.35

**Keywords:** gene transfer agents; marine bacteria; microbial genomes; prophages; repressors

## Introduction

Many phages can develop a symbiotic relationship with their hosts which is termed lysogeny (Ackermann and Dubow, 1987). This usually involves integration of the temperate phage genome (termed prophage) into one of the host's replicons. Such silent viral infections are passed to daughter cells and can be 'activated' (or induced) to cause host lysis by chemical or physical manipulations and presumably natural environmental factors that are poorly understood. Thus, prophages can be viewed as 'dangerous molecular time bombs' capable of bursting their hosts in response to environmental perturbations and spewing progeny capable of infecting other cells (Figure 1). Evidence is accumulating that the lysogenic symbiosis is a highly refined relationship resulting from the coevolution of phage and host genomes (Chen *et al.*, 2005). However, nearly all that is known about the biology and the molecular basis of this interaction is derived

from coliphages ( $\lambda$  and its close relatives). Much less is known of the details of lysogeny in marine bacterial isolates, and even less concerning the ecology of lysogeny in the oceans.

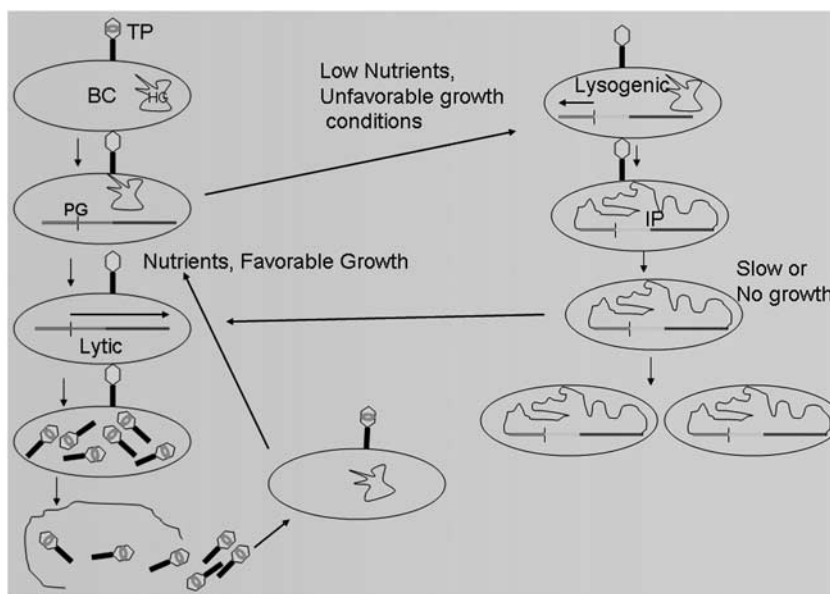
Here, what is known about lysogeny in the seas has been consolidated, including its occurrence, the ecology of lysogeny, some recent genomic information on putative marine prophages and the emergence of gene transfer agents (GTAs), and concluded with a unifying hypothesis of the importance of lysogeny in bacterial survival in the oceans.

### *Occurrence of lysogeny among isolates*

The capability to detect prophages and therefore lysogeny in the ocean is only as good as our methods. Typically the occurrence of lysogeny is based upon the use of a chemical or physical treatment followed by the detection of phage particles produced. However, this procedure ignores the basic biological principal of lysogeny, the capability to establish a stable symbiosis with a host, which is far more difficult to demonstrate experimentally. Thus, defective (that is, remnant or non-infective) prophages, bacteriocins and potentially GTAs have undoubtedly been detected by this

Correspondence: JH Paul, College of Marine Science, University of South Florida, 140 7th Avenue S, St Petersburg, FL 33701, USA.

E-mail: jpaul@marine.usf.edu



**Figure 1** Schematic of the environmental control of lytic/lysogenic decision in a temperate marine phage (TP). A temperate phage attaches to a host bacterial cell (BC), injecting its phage genome (PG). If high nutrients or other environmental conditions that favor host growth exist (green background), the lytic phage genes are activated (green and blue part of the phage genome) resulting in phage production, host lysis and infection of more rapidly growing host cells. If the infection occurs during conditions unfavorable for host growth (light blue background), the lysogenic genes are activated and the phage genome integrates with the host genome (HG) to form an integrated prophage (IP). The integrated prophage can become activated to a lytic lifestyle when favorable growth conditions are restored.

procedure (Ackermann and DuBow, 1987). Additionally, not all prophages are induced by the most common chemical treatment, mitomycin C.

Lysogeny is extremely common among cultivated bacteria; 47% of 1200 diverse strains (of medical or non-marine origin) were lysogenic (Ackermann and DuBow, 1987). It has been estimated that 60–70% of the bacterial genomes sequenced to date contain prophages (Rob Edwards, Personal communication; Casjens, 2003; Casjens and Hendrix, 2005) and two-third of all sequenced  $\gamma$ -proteobacterial and low GC Gram-positive bacteria harbor prophages (Brussow *et al.*, 2004). Bacterial genomes generally lacking prophages fall into two groups, the archaea and intracellular pathogens. The latter group is under tremendous pressure for genome reduction (Canchaya *et al.*, 2004). Temperate phages of archaea have been described for *Thermoproteus tenax* and *Sulfolobus sulfotarius* (Ackermann and DuBow, 1987), but there is scant literature on other archaeal lysogens.

Marine bacterial isolates have been found to contain a high proportion (43%) of inducible prophage-like particles (Jiang and Paul, 1994; Jiang *et al.*, 1998). Stopar *et al.* (2004) found that 71% of the isolates from the Gulf of Trieste were lysogenic. The overall frequency of occurrence of inducible prophage-like particles in 130 Baltic Sea isolates was 28% (Leitet *et al.*, 2006). Samples collected in the spring had a higher incidence of prophage than those collected in summer or fall (~35%). Collectively these results indicate that among marine

bacterial isolates, roughly half contain prophages as detected by chemical induction.

Lysogeny has also been described in algae and cyanobacteria. The occurrence of lysogeny in freshwater filamentous cyanobacteria has been known for more than 35 years (Padan *et al.*, 1972). Lysogeny has been described in marine cyanobacteria, including a filamentous strain (Ohki and Fujita, 1996) and a *Synechococcus* strain (Sode *et al.*, 1994, 1997).

#### *Lysogeny in ambient marine microbial populations*

The importance of lysogeny in the marine environment has been a topic of considerable debate. Early work with natural populations suggested that a relatively low proportion of marine bacteria were lysogens (Wilcox and Fuhrman, 1994). Weinbauer and Suttle (1996, 1999) also concluded that a small proportion (1.5–11.4%) of the bacterial populations in the Gulf of Mexico were lysogenized. Eight of 10 eutrophic environments were found to contain inducible lysogens (bacteria whose prophage can be induced to grow lytically), whereas only 3 of 11 offshore environments contained inducible lysogens (Jiang and Paul, 1996). A series of environmentally relevant pollutants (polynuclear aromatic hydrocarbons, polychlorinated biphenyls and pesticides) could all cause prophage induction in natural populations (Jiang and Paul, 1996; Cochran and Paul, 1998). Estimates of the fraction of the bacterial population lysogenized in a seasonal study of Tampa Bay varied from 0 to approximately 100%

(average  $27.6 \pm 24.7\%$ ; Williamson *et al.*, 2002). Further studies in Tampa Bay and British Columbia indicated that at certain times the entire population was lysogenized (Cochran and Paul, 1998; Ortmann *et al.*, 2002). The reason for the high viral abundance in estuaries may be because of prophage induction, resulting from natural environmental factors and anthropogenic impacts.

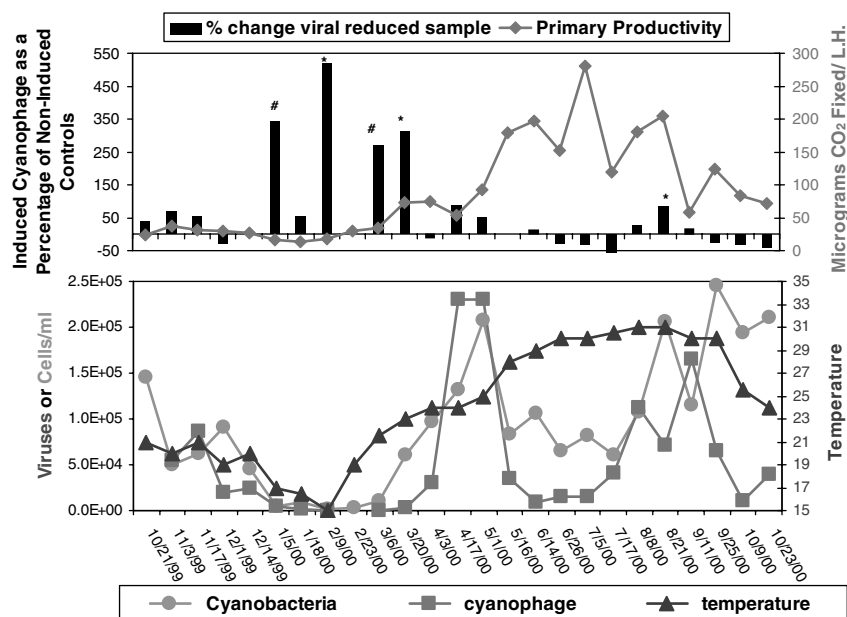
#### Lysogeny in natural populations of cyanobacteria

There has been tremendous interest in the potential for lysogeny in natural populations of *Synechococcus* but very little data. Hewson *et al.* (2001) implied that a *Lyngbia* bloom may have been terminated by prophage induction. A seasonal study of the distribution of cyanobacterial lysogens, determined as mitomycin C-inducible viruses infective for *Synechococcus* WH7803, was conducted in Tampa Bay (Figure 2; McDaniel *et al.*, 2002). As for the heterotrophic microbes, the majority of the samples induced in the winter, with one positive in August. Primary production and the incidence of lysogeny were inversely correlated, suggesting that conditions unfavorable for autotrophic growth favored lysogeny (Figure 2). This is the first definitive evidence for lysogeny among natural populations of marine *Synechococcus*. Because lysogeny confers immunity to infection by related viruses, it is suggested that this process accounts for the resistance to viral infection often observed in *Synechococcus* (Waterbury and Valois, 1993).

Wilson *et al.* (1996, 1998) suggested that phosphate levels could provide a control of the lysogenic switch in *Synechococcus*. When phosphate-depleted

cultures of WH7803 were infected with cyanophage S-PM2, there was an 18 h delay in lysis, along with an 80% decrease in burst size. These researchers concluded that S-PM2 lysogenized the *Synechococcus* strain under low P conditions. An alternate interpretation of the data is that lytic phage production was simply P-limited. It is well known that viruses salvage host nucleotides for nucleic acid synthesis, yet synthesis requires nucleotides to be energized to the NTP form, which requires energy and phosphate. In a microcosm study using waters in a Norwegian Fjord, a nitrogen enrichment experiment led to a large *Synechococcus* bloom ( $5 \times 10^5$  cells ml<sup>-1</sup>), which was P-limited (N:P ratio was 75:1). Addition of P to the microcosm resulted in bloom collapse concomitant with a dramatic increase in phage production (Wilson *et al.*, 1998). These workers suggested that P concentrations controlled the shift from lysogeny to lytic existence, although P-limitation of lytic phage production could also account for these findings. A study of the potential effect of phosphate on prophage induction in natural heterotrophic bacterial populations of the Gulf of Mexico indicated that phosphate additions stimulated lytic phage production, but had no effect on prophage induction (Williamson and Paul, 2004). In this same environment, phosphate amendment actually caused a decrease in lytic phage production in natural populations of marine *Synechococcus*, without stimulating mitC-induced prophage production (McDaniel and Paul, 2005).

Hewson and co-workers studied the potential for virulent phage attack of natural populations and cultures of the nitrogen-fixing pelagic cyanobacterium *Trichodesmium* spp. While incubating



**Figure 2** Level of prophage induction of *Synechococcus* populations (closed bar) compared to primary productivity (green diamonds), *Synechococcus* abundance (pink circles), cyanophage counts (brown  $\square$ ) and temperature (blue  $\triangle$ ). Results are expressed as percentage change in treatment compared to control (\* and # indicate statistical significance at the 95 and 90% confidence interval, respectively).

*Trichodesmium* colonies in containers, these workers noted the spontaneous induction of virus-like particles and trichome lysis, believed to be the result of prophage induction. Cultivated *Trichodesmium* trichomes also showed production of phage when incubated with mitC (Ohki, 1999; Hewson *et al.*, 2004).

An explanation of the variation in occurrence of lysogeny in the environment comes from the combined work of Weinbauer *et al.* (2003), McDaniel *et al.* (2002, 2006) and Williamson *et al.* (2002). The seasonal distribution of lysogens in Tampa Bay inversely reflected temperature, host abundance and bacterial and primary production (McDaniel *et al.*, 2002; Williamson *et al.*, 2002). The greatest occurrence of lysogeny was in the winter months, when bacterial and primary production was lowest. Similar observations have been made for Antarctic oligotrophic lakes, where lysogeny was detected during Austral winter and early spring samplings (Laybourn-Parry *et al.*, 2007; S  wstr  m *et al.*, 2007). Lysogeny was not observed in the highly productive waters of the Mississippi River plume, whereas surface and subsurface oligotrophic waters of the Gulf of Mexico contained lysogens (McDaniel *et al.*, 2006). Finally, analysis of Baltic and Mediterranean Sea bacterial populations indicated that subsurface populations had a higher incidence of lysogeny, including anoxic waters (Weinbauer *et al.*, 2003). An inverse log relationship was observed between lytic infection (frequency of visibly infected cells; FVIC) and lysogeny. These results indicate that lysogeny is favored in waters of low host abundance and productivity, as a survival strategy for viruses during episodes of low energy sources or unfavorable conditions in general (Figure 1).

#### *Why be temperate?*

The advantages of lysogeny from the perspective of a phage are obvious, most notably the capability to survive times of low host abundance. This includes protection from proteolytic digestion, grazing to some degree and UV inactivation. Yet the advantages to the host are probably more subtle—the possession of extra DNA must have a metabolic cost (hence the tendency to lose non-essential plasmids through segregation). Indeed, a prophage is also a dangerous molecular time bomb that often results in the lysis of at least some of the hosts, and under certain environmental conditions, perhaps the entire host population.

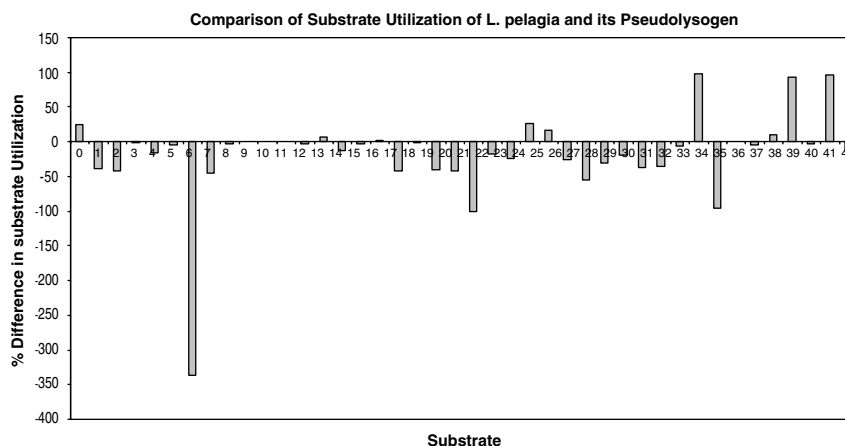
Although concept that prophages enhance host fitness has been known for some time, (Edlin *et al.*, 1975) yet how this occurs is only now being understood by a combination of physiological analyses coupled with whole genome expression array analyses. Prophage genes that promote host fitness are termed fitness factors (Brussow *et al.*, 2004) and provide five mechanisms for accomplishing this: (1) prophages serve as anchor points for

gene rearrangements, often functioning as transposons, (2) prophages can interrupt genes through insertion, producing metabolic losses through silencing non-essential gene functions, (3) prophages usually confer hom immunity to related phage infections (resistance to superinfection), (4) lysis of closely related nearby strains ('kill the relatives') and (5) introduction of new fitness factors by conversion or transduction Jiang and Paul (1998). A sixth mechanism has been proposed here; the silencing of host genes by phage repressors.

Fitness factors are genes not essential for the phage life cycle, and are termed lysogenic conversion factors or morons (for more DNA, or genes that encode functions beyond that needed to complete the lytic infection cycle; Brussow *et al.*, 2004). No greater attention has been given to any lysogenic conversion factors than those of pathogenic bacteria, sometimes referred to as virulence factors. Pathogenicity islands, or blocks of foreign genes encoding virulence factors have often turned out to be prophages. Well known prophage-encoded virulence factors include the cholera toxin genes of *Vibrio cholera* (Waldor and Mekalanos, 1996), virulence genes of *Staphylococcus aureus* (Sumbly and Waldor, 2003), scarlet fever genes of *Sterptococcus pyogenes*, shiga toxin genes (Tyler *et al.*, 2004) and many more. In fact, nearly all pathogenic bacteria contain at least one prophage (and many several), whereas pathogens whose genomes lack prophages may be the result of the strain selected for genomic sequencing (Brussow *et al.*, 2004). Diseases of marine organisms have been shown to be the result of prophage conversion as well (Oakey *et al.*, 2002; Munro *et al.*, 2003).

#### *Prophages as modulators of phenotypic diversity*

Prophage infection can dramatically alter the host phenotype to the point that the organism's role in the ecosystem may be changed. Although the role of the expression of prophage genes (morons) in the process of lysogenic conversion is coming to light, the way that temperate phages affect host gene expression is largely unknown. For example, phage *Vibrio harveyi* myovirus-like (VHML) infection of five *V. harveyi* strains resulted in altered phenotypes as determined by substrate utilization (Vidgen *et al.*, 2006). Upon prophage infection, one *V. harveyi* was so altered phenotypically as to be profiled as a *V. campbelli* (Vidgen *et al.*, 2006). A major observation was that there was a general suppression of substrate utilization in the prophage-infected hosts compared to their uninfected counterparts. Because different strains had differing reductions in the types of substrates utilized, it was proposed that VHML may have chromosomally integrated into different sites in different strains (Vidgen *et al.*, 2006). However, based upon the structure of the VHML phage genome, it is doubtful that it chromosomally integrates at all in any of its hosts, but rather



**Figure 3** Relative substrate utilization between *Listonella pelagia* and its pseudolysogen. Values are expressed as percentage difference in substrate utilization, and are calculated as  $((G_{\text{PSEUD}} - G_{\text{LP}}) / G_{\text{PSEUD}}) \times 100$  where  $G_{\text{PSEUD}}$  is the growth of the pseudolysogen and  $G_{\text{LP}}$  is the growth of the uninfected *L. pelagia*. Substrate identity appears in Table 1 (supplement; Long, 2006).

exists as a linear plasmid as for coliphage N15 (see below).

We performed Biolog substrate utilization analysis for uninfected *Listonella pelagia* and its pseudolysogen infected with  $\phi$ HSIC (Figure 3; Long, 2006). Pseudolysogeny is the process by which cells grow to a high cell density while producing a high titer of phages, often resulting in sigmoid growth kinetics (Ackermann and DuBow, 1987).

Positive values in Figure 3 indicate that the pseudolysogen grew better than the uninfected strain on the particular substrate, whereas negative values indicate less growth or lack of substrate utilization by the infected strain. These results indicate that prophage maintenance resulted principally in a reduction in substrate utilization capability.

How could suppression of metabolic capabilities be an advantage to a lysogen? A seminal paper on the improved fitness through phage regulation of host genes found that  $\lambda$  infection resulted in the complete suppression of the host *pckA* gene, which encodes phosphoenolpyruvate carboxykinase, a gene critical for gluconeogenesis (Chen *et al.*, 2005). It was found that the  $\lambda$  cI repressor, which shuts down lytic phage gene expression, also bound to the operator of the *pckA* gene. There were at least seven potential cI binding sites in the  $-10$  to  $-35$  RNA polymerase-binding site of the operator of the *pckA* gene. Additionally, there were distinct cI homolog binding sites for four other temperate coliphages in this operator site. These findings argue strongly for selective pressure to acquire cI-binding sites in the operators of certain host genes, and reinforces the concept of coevolution of the phage and host in lysogeny. Chen *et al.* (2005) argue that the increase in fitness coming from the shut-down of the *pckA* gene is derived from the slower growth in a glucose-free environment, thereby providing survival in an unfavorable environment. cI repressor-like open reading frames (ORFs) have

been found in marine prophage genomes (see below), including VHML (Oakey *et al.*, 2002; Scott *et al.*, 2006). Perhaps lysogenic repressors serve to 'throttle down' marine bacterial metabolism, conserving energy by shutting down unnecessary metabolic pathways and ensuring survival under unfavorable conditions.

## Marine prophage genomics

### Detection of prophages in marine bacterial genomes

One hundred thirteen marine bacterial genomes were investigated for the presence of prophages based upon existing annotation (Supplementary Table 2). Protein-coding features were scanned for the term 'phage'. When a phage term was encountered, the surrounding genes were examined for other known phage genes. When clusters of phage genes were found (cluster defined as at least 3 phage genes in a 10 consecutive ORF window), the prophage-like fragment was imported into KODON (Applied Maths, Inc., Austin, TX, USA) for further analysis (checking annotation and graphical representation of prophages). Although the current annotation was generally maintained, certain ORFs were re-analysed using BLASTP and the annotation updated.

A screening of 113 marine bacterial genomes resulted in detection of 64 prophage-like elements based on our criteria (Supplementary Table 3 and Figure 1). Because of polylysogeny (a lysogen containing more than one prophage), the prophage-like elements found were spread amongst 49 bacterial strains, for an incidence of 43%. Of the 64, 21 (33%) were GTA-like elements with an average size of approximately 15 kb (Table 1). These were found almost exclusively in marine  $\alpha$ -proteobacteria, the exceptions being the  $\gamma$ -proteobacteria *Photobacterium profundum* and *Geobacillus kaustophilus* (Table 1). In the 1970s Barry Marrs and co-workers described

**Table 1** GTA-like elements in marine bacterial genomes

Strain	Accession no.	GTA (bp)	Other prophages
<i>Aurantimonas</i> sp. SI85-9A1	NZ_AAPJ000000000	14 997	
<i>Erythrobacter</i> sp. NAP1	NZ_AAMW000000000	13 519	
<i>Erythrobacter litoralis</i> HTCC2594	NC_007722	15 643	
<i>Loktanella vestfoldensis</i> SKA53	NZ_AAMS000000000	14 202	
<i>Maricaulis maris</i> MCS10	NC_008347	16 193	
<i>Nitrobacter</i> sp. Nb-311A	NZ_AAMY000000000	18 669	Yes
<i>Oceanicaulis alexandrii</i> HTCC2633	NZ_AAMQ000000000	18 172	
<i>Oceanicola granulosus</i> HTCC2516	NZ_AAOT000000000	10 681	
<i>Photobacterium profundum</i> 3TCK	NZ_AAPH000000000	11 590	
<i>Rhodobacterales bacterium</i> HTCC2150	NZ_AAXZ000000000	14 569	
<i>Rhodobacterales bacterium</i> HTCC2654	NZ_AAMT000000000	16 708	Yes
<i>Roseobacter</i> CCS2	NZ_AAYB000000000	14 109	
<i>Roseobacter denitrificans</i> Och114	NC_008209	13 783	
<i>Roseobacter</i> MED 193	NZ_AANB000000000	14 668	Yes
<i>Roseovarius rubinhibens</i> ISM	NZ_AALY000000000	13 917	
<i>Roseovarius</i> sp. 217	NZ_AAMV000000000	17 038	
<i>Roseovarius</i> sp. HTCC2601	NZ_AATQ000000000	15 873	Yes
<i>Sagittula stellata</i> E-37	NZ_AAYA000000000	14 778	
<i>Silicibacter pomeroyi</i> DSS-3	NC_003911	15 208	
<i>Silicibacter</i> TMS1040 Prophage 3	NC_008044	16 000	
<i>Sulfitobacter</i> sp. NAS-14.1	NZ_AALZ000000000	14 257	Yes

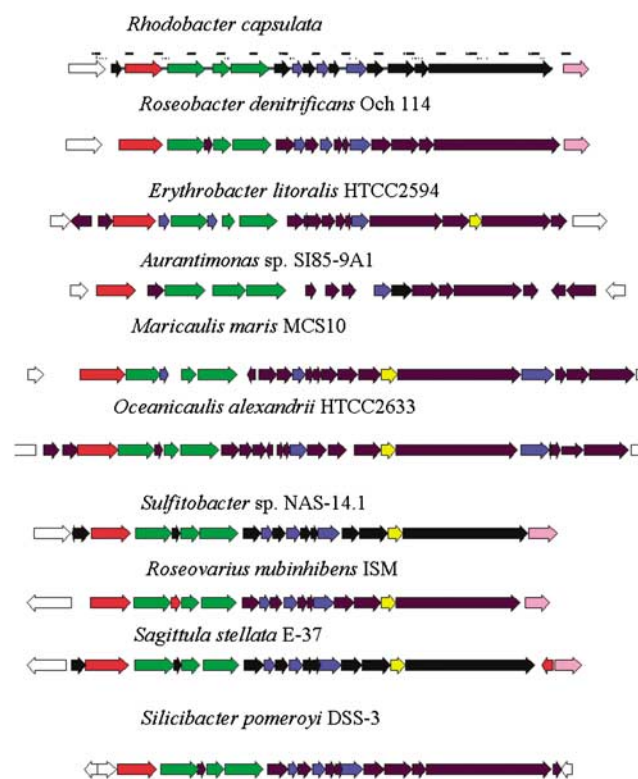
Abbreviation: GTA, gene transfer agent.

defective phage-like particles in *Rhodobacter capsulata* that packaged host DNA with high efficiency catalyzing effective gene transfer and named these GTAs (Marrs, 1974). GTAs are virtually little genetic escape pods that can ensure gene survival in dying host cells. They encode for the process of 'constitutive transduction', which is one of the few procaryotic gene transfer mechanisms to be heritable and chromosomally encoded (Lang and Beatty, 2001). Although these observations were viewed as highly novel at the time, little research followed until two independent findings came to light—the report of a GTA-like sequence in the genome of *Silicibacter* TMS1040 (Chen *et al.*, 2006) and the widespread occurrence of GTAs in the genome sequences of  $\alpha$ -proteobacteria (Lang and Beatty, 2007).

#### Genomic organization of GTAs

Gene transfer agents are usually smaller than prophages (13–16 kb) and have a distinctive genomic organization, normally arranged in 15 ORFs (Figure 4). The first ORF is a small one that is often missing, followed by the terminase large subunit gene. In many phages, the terminase small subunit gene precedes the large subunit, and these two genes work in unison to package the phage DNA in the capsid. Usually the small subunit possesses DNA recognition specificity and the large subunit possesses nucleolytic and ATPase activities (Lin *et al.*, 1999). In GTAs this process has been altered to digest and package the host DNA.

It is hypothesized that the lack of a small terminase subunit is responsible for this aberrant packaging as GTAs have lost the DNA specificity function. ORFs 3–5 are the portal, prohead protease



**Figure 4** Genomic organization of gene transfer agents (GTAs) detected in nine marine bacterial genomes. Red-terminase large subunit; Green-portal, capsid protease and major capsid proteins; Blue-tail associated; Black, unknown or conserved hypothetical; yellow-lytic enzyme; Pink, serine O-acetyl transferase (host); white-host.

and major capsid proteins. These are usually similar to coliphage HK97. In ORFs 6–11 there are usually some tail genes, including a head to tail adapter



protein, a major tail protein and sometimes a tail tape measure protein (these appear as blue arrows in Figure 4). Absent from the prototypical *R. capsulata* GTA, but present in many marine GTAs is a cell wall hydrolase (usually ORF 14, a yellow arrow in Figure 4). A diagnostic ORF for GTAs is the large ORF 15, which is sometimes annotated as a 'host specificity' protein, although no clear function has been assigned (Lang and Beatty, 2007). Interestingly, the first host protein following the GTA is often a serine *O*-acetyltransferase, suggesting that GTAs themselves are vertically transmitted, and are not inserting themselves in random locations as integrating prophages often do. The phylogeny of prophages seldom correlates with the phylogeny of the host, whereas the  $\alpha$ -proteobacteria contain a very evolutionarily conserved GTA organization (Lang and Beatty, 2007).

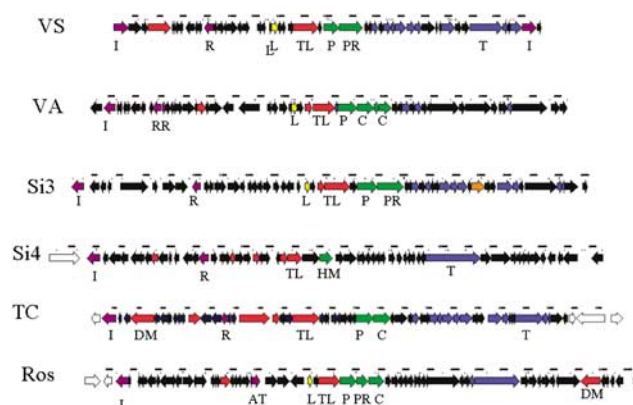
We have searched marine microbial metagenomes in the CAMERA database for the conserved GTA terminase using BLASTX and found 249 good hits. To determine the relative abundance of GTAs in the cells from which the metagenomes were derived, we sought to establish the ratio of the GTAs to the total number of genomes in each metagenome. Adapting the same strategy that Moran and colleagues (Howard *et al.*, 2006) used to estimate the total genome copies in each metagenome based upon occurrence of the single copy *recA* homologs, GTAs were present in approximately 2.9% of the total bacterial genomes.

## Prophage genomic organization in marine bacteria

A diversity of genomic organization exists for the non-GTA-like prophage elements found in marine bacterial genomes. These have been categorized based on similarities to non-marine prophage genomes in Supplementary Table 3. Our type A prophages have a Sfi11-like organization (Canchaya *et al.*, 2004) containing a lysogeny module (coliphage  $\lambda$ -like integrase, and at least one phage repressor) followed by replicative genes that include a terminase, portal, capsid and tail genes (exemplified by prophages detected in *Thiomicrospira crunogena* and *V. alginolyticus* 12G01; Figure 5). These phages are 37–41 kb in length, and have genomic organizations shared by other cultivated phages.

Our type B prophages are similar to the type A in that they contain a recognizable integrase and are 31–49.5 kb in length but lack a discernable phage repressor protein (Supplementary Table 3). Although there is no consistent order in genomic organization among this group, five of the seven phages contain some types of DNA methylase.

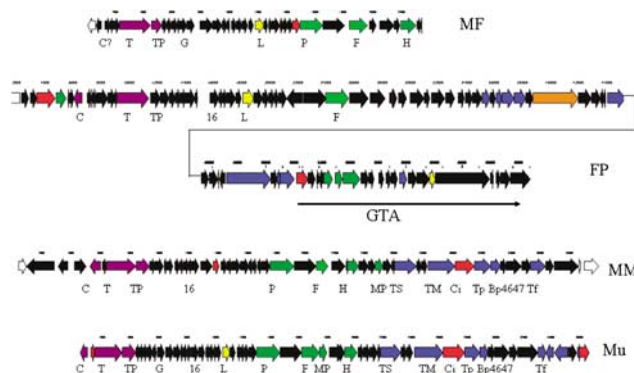
The largest group of prophages are the type C, which range in size from 12 to 29 kb, being shorter than many previously described temperate phages (Ackermann and DuBow, 1987; Canchaya *et al.*,



**Figure 5** Genomic organization of type A prophages (VS, *Vibrio splendidus* 12B1 prophage 1; VA, *V. alginolyticus* 12G01; Si3, *Silicibacter* TMS1040 prophage 3; Si4, *Silicibacter* TMS 1040 prophage 4; TC, *Thiomicrospira crunogena* XCL-2; Ros., *Roseobacter* SK209-2-6). Open reading frame (ORF) abbreviations are as follows: I, integrase; R, repressor; TS, terminase small subunit; TL, terminase large subunit; L, lysis; P, portal protein; PR, prohead protease; AT, antiterminator; C, capsid; DM, DNA methylase; HM, head morphogenesis protein. Colors of ORF relate to function groups as described in Figure 3.

2003; Casjens, 2003). Some of these may be prophage remnants, although at least one has had the sequence determined from induced prophages (*Bacillus* B14905 prophage 1). Ten of the type C prophages contain integrases, and four of these possess phage repressors or antiterminators, and thus could be considered shortened versions of the type A prophage elements discussed above. Of particular interest is a prophage element in *Synechococcus* sp 5701, the first complete prophage-like element to be recognized in a marine *Synechococcus*. This element contains three putative transposases and a terminal CP4-like integrase; however, no structural genes were annotated in this prophage.

Many marine prophages do not fit the criteria established for the A, B and C-type prophage groupings given above and are listed as 'other types' in (Supplementary Table 3). Three such marine prophages contained a genomic organization reminiscent of coliphage Mu, a prophage that lacks an integrase gene. Transposable phage Mu infects a wide range of Gram-negative bacteria and amplifies its genome by multiple rounds of replicative transposition (Toussaint *et al.*, 1994). Mu-like prophages are generally not inducible by mitomycin C. Control of lysogeny is effected through the action of the Rep repressor, that is encoded by the *c*-gene (the first ORF in Figure 5). A prophage element found in the genome of *Marinomonas* MED 121 is nearly identical in length (37 598 compared to 36 717) and genomic organization to that of enterobacteriophage Mu (Figure 6). A prophage element found in *Fulvimarina pelagi* HTCC2506 contained a Mu-like element linked to at least two other prophage fragments totaling 66 477 nt. This included a packaging and portal protein of unknown origin, followed by 41

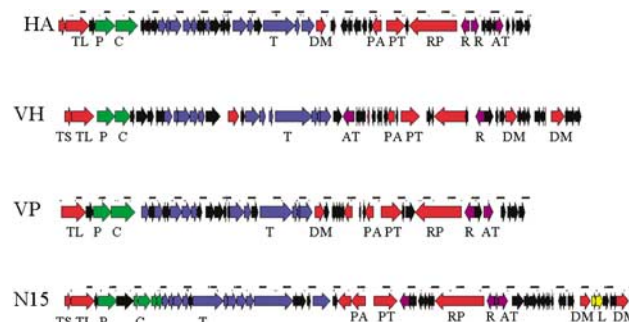


**Figure 6** Mu-like prophages found in marine bacterial genomes (MF, *Mariprofundus ferrooxydans* pV-1; FP, *Fulvimarina pelagi* HTCC2506; MM, *Marinomonas* sp. MED121; Mu, Enterobacteriophage Mu). Open reading frame (ORF) abbreviations are as follows: C, c-repressor; T, transposase; TP, transposition protein; G, gam; 16, gp16; F, virion morphogenesis F; MO, morphogenesis protein; H, major head subunit protein; TS, tail sheath; TM, tape measure; Ci, DNA circularization protein; Tp, tail protein; Bp, baseplate protein; 46,47, gp46, gp47; Tf, tail fiber protein. All other ORF abbreviations as in Figure 4, and ORF groupings by color as in Figure 3.

ORFs in the Mu-like region, a stretch of P2-like tail proteins that included a virulence associate protein, and ended with a 21 ORF segment highly identical to a GTA (Figure 6). It is not known if this element reflects one prophage or is a prophage and an autonomous GTA.

At least one prophage element resembled the organization and gene content of the defective prophage of *Bacillus subtilis* 168, PBSX. PBSX is a mitomycin C-inducible prophage-like particle that packages 13 kb portions of its host chromosome (Anderson and Bott, 1985), functioning like a GTA. A cluster of PBSX homologs was found in the genome of *G. kaustophilus* in a putative prophage of 33 152 nt. The structure was unique in that the first 38 ORFs contained all the PBSX genes, which were followed by an entire GTA-like sequence of 14 ORFs (Supplementary Figure 1). Indeed, the GTAs may not be limited to the  $\alpha$ -proteobacteria; in fact it is entirely possible that they originated from defective temperate phage of the  $\gamma$ -proteobacteria.

One prophage-like element and two other marine phages had genomic arrangements resembling that of coliphage N15 (Figure 7). N15 is unusual in that it exists as a linear prophage with telomeric ends (Ravin and Shulga, 1970). Key components to this architecture are the protelomerase, the plasmid partitioning protein parA and a replication protein, as well as phage repressors to ensure a lysogenic lifestyle. The *Halomonas aquamarina* prophage HAP-1 contained these genes as well as protelomerase activity, and has been shown to exist as a linear prophage (Mobberley *et al.*, in press). Two other marine phages, the temperate VHML (Oakey *et al.*, 2002) of *V. harveyi* and VP882 of *V. parahaemolyticus* have nearly identical structure and are included in Figure 7.



**Figure 7** N15-like marine temperate phages (HA, *Halomonas aquamarina* prophage HAP-1; VH, *Vibrio harveyi* phage VHML; VP, *V. parahaemolyticus* phage VP882). Open reading frame (ORF) abbreviations are as follows: PA, parA protein; PT, protelomerase; RP, replicative protein. All other ORF abbreviations as in Figure 3.

The final group of prophage-like elements do not fit any pattern of genomic organization and are simply referred to as 'disorganized prophages'. Of particular interest is the large (>70 kb) prophage of *Silicibacter* TMS1040 (prophage 1; Chen *et al.*, 2006). This prophage has been experimentally detected by mitomycin C induction, PCR and pulsed field gel electrophoresis analysis (Chen *et al.*, 2006). In addition to containing two integrases, there is a GTA-like element internal to the genome. This GTA-like sequence may have been the result of co-infection of *Silicibacter* TMS 1040 with a GTA and temperate phage that recombined in an illegitimate recombination event. *Silicibacter* TMS 1040 has a GTA in addition to the one detected in the large prophage element. Either this prophage picked up a GTA or this large phage is the progenitor of the current GTA.

#### Transcriptional regulators and repressors

We have compared the occurrence of transcriptional regulators and repressors in 20 lytic marine phages (Supplementary Table 4). Repressors and transcriptional regulators were found on the basis of existing annotation. The 42 prophage elements and 2 temperate phage genomes have a relatively high occurrence of transcriptional regulators and gene repressors (averaging 2 per genome or 2.3/40 kb sequence) compared to lytic marine phages (0.60 per genome or 0.23/40 kb sequence). Therefore, in addition to providing advantages conferred by conversion (including DNA methylation systems and occasional virulence proteins), marine prophages may have their greatest role in the modulation of phage and host gene expression. The example given of host cell metabolic gene repression (Chen *et al.*, 2006) may only be a portion of the story. Certainly temperate phage genetic switches (for example,  $\lambda$ ) are highly evolved to respond to metabolic and nutritive status of the host cell. It may well be that the subtleties of existence in a primarily low nutrient environment require an



additional regulatory system encoded by the prophage that 'throttles back' metabolic activity ensuring both host and prophage survival. A higher metabolic rate might result in prophage induction under conditions that would not favor finding another host.

#### *Why are all marine bacteria not lysogenic?*

Being constantly surrounded by relatively high numbers of phages would intuitively lead to some type of phage infection (that is, lysis or lysogeny). A natural way to become resistant to infection (at least resistant to some phages) is to become lysogenized. Yet most studies of prophage occurrence in bacteria yield show that 40–60% are lysogenized (roughly half; Ackermann and DuBow, 1987; Jiang and Paul, 1994; this study). It may be that to keep lytic processes going, roughly half the bacteria need to be good hosts for lytic infection, thereby ensuring diversity through 'kill the winner'-type processes (Thingstad *et al.*, 1997). Lytic phage infection is generally regarded to cause nearly half the bacterial mortality in the ocean, and would not occur in a totally lysogenic microbial world. An equally valid question is—where do the lysogens go when conditions favor rapid bacterial growth? Does the population shift to genotypes that are less likely to be lysogenized? Or are there massive induction events at the end of the lysogeny season resulting from environmental cues that trigger these intragenomic molecular time bombs? Very little is known concerning the environmental control of the lytic/lysogeny switch in natural bacterial populations, partially because we lack the intensive studies of marine lysogens as have occurred for coliphage  $\lambda$ , and also because few studies have investigated the effects of environmental conditions or perturbations on natural populations of lysogens.

## Summary and conclusions

At least half of the cultivatable marine bacteria have been shown to contain temperate phage-like particles by chemically-mediated prophage induction. The occurrence of lysogeny in natural populations at first seems to be highly variable. However, systematic study of many diverse marine environments indicate that highly productive environments are dominated by lytic phage infection, whereas oligotrophic, anoxic or seasonally stressed environments favor lysogeny. An overview of a 100 recently published marine bacterial genomes indicated that roughly 40% contained putative prophage-like elements. Surprisingly, at least one-third of the putative prophages strongly resembled GTAs. The remaining two-thirds had diverse prophage genomic organizations, resembling coliphages N15 and Mu and the *Bacillus* defective prophage element PBSX. There was a high frequency of occurrence of

transcriptional regulators and repressors among the putative prophages that was not found by annotation analysis in 21 lytic marine phage genomes. Coupling these observations with the recent finding that coliphage  $\lambda$  repressors can actively repress and modulate host metabolic genes, we hypothesize that marine prophages serve to repress host growth in times of resource partitioning. That is, prophages repress not only their own lytic genes but unnecessary and wasteful host metabolic process genes. Such metabolic economization enables the host (and the prophage) to survive disfavorable conditions until environmental clues alert the cell that conditions of growth have changed. Thus, prophages are not solely dangerous molecular time bombs that can bring about cell mortality, but also serve as a key to bacterial survival in the oligotrophic oceans.

## Acknowledgements

This work was supported by NSF Biocomplexity Grant OCE-0221763.

## References

- Ackermann HW, DuBow MS. (1987). *Viruses of Prokaryotes. Vol. I. General Properties of Bacteriophages*. CRC Press, Inc.: Boca Raton, FL.
- Anderson LM, Bott KF. (1985). DNA packaging by the *Bacillus subtilis* defective bacteriophage PBSX. *J Virol* **54**: 773–780.
- Brussow H, Canchaya C, Hardt W-D. (2004). Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. *Microbiol Mol Biol Rev* **68**: 560–602.
- Canchaya C, Fournous G, Brussow H. (2004). The impact of prophages on bacterial chromosomes. *Mol Microbiol* **53**: 9–18.
- Canchaya C, Proux C, Fournous G, Bruttin A, Brussow H. (2003). Prophage genomics. *Microbiol Mol Biol Rev* **67**: 238–276.
- Casjens S. (2003). Prophages and bacterial genomic: what have we learned so far? *Mol Microbiol* **49**: 277–300.
- Casjens S, Hendrix RW. (2005). Bacteriophages and the bacterial genome. In: Higgins PN (ed). *Bacterial Chromosome*. pp 39–53.
- Chen F, Wang K, Stewart J, Belas R. (2006). Induction of multiple prophages from a marine bacterium: a genomic approach. *Appl Environ Microbiol* **72**: 4995–5001.
- Chen Y, Golding I, Sawai S, Guo L, Cox EC. (2005). Population fitness and the regulation of *Escherichia coli* genes by bacterial viruses. *PLoS Biol* **3**: 1276–1282.
- Cochran PK, Paul JH. (1998). Seasonal abundance of lysogenic bacteria in a subtropical estuary. *Appl Environ Microbiol* **64**: 2308–2312.
- Edlin G, Lin L, Bitner R. (1975). Lambda lysogens of *Escherichia coli* reproduce more rapidly than non-lysogens. *Nature* **255**: 735–737.
- Hewson I, O'Neil JM, Dennison WC. (2001). Virus-like particles associated with *Lyngbya majuscula* (Cyanophyta; Oscillatoriaceae) bloom decline in Moreton Bay, Australia. *Aquat Microb Ecol* **25**: 207–213.

- Howard EC, Henriksen JR, Buchan A, Reisch CR, Bürgmann H, Welsh R *et al.* (2006). Bacterial taxa that limit sulfur flux from the ocean. *Science* **314**: 649–652.
- Jiang SC, Kellogg CA, Paul JH. (1998). Characterization of marine temperate phage-host systems isolated from Mamala Bay, Oahu, Hawaii. *Appl Environ Microbiol* **64**: 535–542.
- Jiang SC, Paul JH. (1994). Seasonal and diel abundance of viruses and occurrence of lysogeny/bacteriocinogeny in the marine environment. *Mar Ecol Progr Ser* **104**: 163–172.
- Jiang SC, Paul JH. (1996). The abundance of lysogenic bacteria in marine microbial communities as determined by prophage induction. *Mar Ecol Progr Ser* **142**: 27–38.
- Jiang SC, Paul JH. (1998). Gene transfer by transduction in the marine environment. *Appl Environ Microbiol* **64**: 2780–2787.
- Lang AS, Beatty JT. (2001). The gene transfer agent of *Rhodobacter capsulatus* and 'constitutive transduction' in prokaryotes. *Arch Microbiol* **175**: 241–249.
- Lang AS, Beatty JT. (2007). Importance of widespread gene transfer agent genes in  $\alpha$ -proteobacteria. *Trends in Microbiol* **15**: 54–62.
- Laybourn-Parry J, Marshall WA, Madan NJ. (2007). Viral dynamics and patterns of lysogeny in saline Antarctic lakes. *Polar Biol* **30**: 351–358.
- Leitet C, Riemann B, Hagstrom A. (2006). Plasmids and prophages in Baltic Sea bacterioplankton isolates. *J Mar Biol Ass UK* **86**: 567–575.
- Lin H, Rao VB, Black LW. (1999). Analysis of capsid portal protein and terminase functional domains: interaction sites required for DNA packaging in Bacteriophage T4. *J Mol Biol* **289**: 249–260.
- Long AK. (2006). Control of lysogeny in marine bacteria: studies with  $\phi$ HSIC and natural populations. MS Thesis, University of South Florida, Tampa.
- Marrs BL. (1974). Genetic recombination in *Rhodospirillum rubrum*. *Proc Natl Acad Sci USA* **71**: 971–973.
- McDaniel L, Houchin L, Williamson S, Paul JH. (2002). Lysogeny in natural populations of marine *Synechococcus*. *Nature* **415**: 496.
- McDaniel L, Paul JH. (2005). Effect of nutrient addition and environmental factors on prophage induction in natural populations of marine *Synechococcus* species. *Appl Environ Microbiol* **71**: 842–850.
- McDaniel L, Paul JH, de la Rosa MA. (2006). Temperate and lytic cyanophages from the Gulf of Mexico. *J Mar Biol Assoc UK* **86**: 517–527.
- Mobberley JM, Authement N, Paul JH, Segal AM. The temperate marine phage  $\Phi$ HAP-1 of *Halomonas aquamarina* possesses a linear plasmid-like prophage genome. *J Virol* (in press).
- Munro J, Oakey J, Bromage E, Owens L. (2003). Experimental bacteriophage-mediated virulence in strains of *Vibrio harveyi*. *Dis Aquat Org* **54**: 187–194.
- Oakey HJ, Cullen BR, Owens L. (2002). The complete nucleotide sequence of the *Vibrio harveyi* bacteriophage VHML. *J Appl Microbiol* **93**: 1089–1098.
- Ohki K. (1999). A possible role of temperate phage in the regulation of *Trichodesmium* biomass. *Bull Inst Ocean Monaco* **19**: 287–291.
- Ohki K, Fujita Y. (1996). Occurrence of a temperate cyanophage lysogenizing the marine cyanophyte *Phormidium persicinum*. *J Phycol* **32**: 365–370.
- Ortmann AC, Lawrence JE, Suttle CA. (2002). Lysogeny and lytic viral production during a bloom of the cyanobacterium *Synechococcus* spl. *Microbial Ecol* **43**: 225–231.
- Padan E, Shilo M, Oppenheim AB. (1972). Lysogeny of the blue-green alga *Plectonema boryanum* by LPP2-SPI cyanophage. *Virology* **47**: 525–526.
- Ravin VK, Shulga MG. (1970). Evidence for extrachromosomal location of prophage N15. *Virology* **40**: 800–807.
- Säwström M, Anesio A, Grane'li W, Laybourn-Parry J. (2007). Seasonal viral loop dynamics in two large ultraoligotrophic Antarctic freshwater lakes. *Microbial Ecol* **53**: 1–11.
- Scott KM, Sievert SM, Abril FN, Ball LA, Barrett CJ, Blake RA *et al.* (2006). The genome of deep-sea vent chemolithoautotroph *Thiomicrospira crunogena*. *PLoS Biol* **4**: 1–17.
- Sode K, Oonari R, Oozeki M. (1997). Induction of a temperate marine cyanophage by heavy metal. *J Mar Biotech* **5**: 178–180.
- Sode K, Oozeki M, Asakawa K, Burgess JG, Matsunaga T. (1994). Isolation of a marine cyanophage infecting the marine unicellular cyanobacterium, *Synechococcus* so NKBG 042902. *J Mar Biotech* **1**: 189–192.
- Stopar D, Cerne A, Zigman M, Poljalis-Prijatelj M, Turk V. (2004). Viral abundance and high proportion of lysogens suggest that viruses are important members of the microbial community in the Gulf of Trieste. *Microb Ecol* **47**: 1–8.
- Sumby P, Waldor MK. (2003). Transcription of the toxin genes present within the staphylococcal phage  $\phi$ Sa3ms is intimately linked with the phage's life cycle. *J Bacteriol* **185**: 6841–6851.
- Thingstad TF, Hagström Å, Rassoulzadegan F. (1997). Accumulation of degradable DOC in surface waters: Is it caused by a malfunctioning microbial loop? *Limnol Oceanogr* **42**: 398–404.
- Toussaint A, Gama MJ, Laachouch J, Maenhaut-Michel G, Mhammedi-Alaoui A. (1994). Regulation of bacteriophage Mu transposition. *Genetica* **93**: 27–39.
- Tyler JS, Mills MJ, Friedman DI. (2004). The operator and early promoter region of the Shiga toxin type 2-encoding bacteriophage 933W and control of toxin expression. *J Bacteriol* **186**: 7670–7679.
- Vidgen MJ, Carson M, Higgins M, Owens L. (2006). Changes to the phenotypic profile of *Vibrio harveyi* myovirus-like (VHML) bacteriophage. *J Appl Microbiol* **100**: 481–487.
- Waldor JK, Mekalanos JJ. (1996). Lysogenic conversion by a filamentous phage encoding cholera toxin. *Science* **272**: 1910–1914.
- Waterbury JB, Valois FW. (1993). Resistance to co-occurring phages enables marine *Synechococcus* to coexist with cyanophages abundant in seawater. *Appl Environ Microbiol* **59**: 3393–3399.
- Weinbauer MG, Brettar I, Hofl. (2003). Lysogeny and virus-induced mortality of bacterioplankton in surface, deep, and anoxic marine waters. *Limnol Oceanogr* **48**: 1457–1465.
- Weinbauer MG, Suttle CA. (1996). Potential significance of lysogeny to bacteriophage production and bacterial mortality in coastal waters of the Gulf of Mexico. *Appl Environ Microbiol* **62**: 4374–4380.
- Weinbauer MG, Suttle CA. (1999). Lysogeny and prophage induction in coastal and offshore bacterial communities. *Aquat Microb Ecol* **18**: 217–225.
- Wilcox RM, Fuhrman JA. (1994). Bacterial viruses in coastal seawater: lytic rather than lysogenic production. *Mar Ecol Progr Ser* **114**: 35–45.

- Williamson SJ, Paul JH. (2004). Nutrient stimulation of lytic phage production in bacterial populations of the Gulf of Mexico. *Aquat Microb Ecol* **36**: 9–17.
- Williamson S, McDaniel L, Houchin L, Paul JH. (2002). Seasonal variation in lysogeny as depicted by prophage induction in Tampa Bay, Florida. *Appl Environ Microbiol* **68**: 4307–4314.
- Wilson WH, Carr NG, Mann NH. (1996). The effect of phosphate status on the kinetics of cyanophage infection in the oceanic cyanobacterium *Synechococcus* sp. WH 7803. *J Phycol* **32**: 506–516.
- Wilson WH, Turner S, Mann NH. (1998). Population dynamics of phytoplankton and viruses in phosphate-limited mesocosm and their effect on DMSP and DMS production. *Estuar Coast Shelf Sci* **46**: 49–59.

Supplementary Information accompanies the paper on The ISME Journal website (<http://www.nature.com/ismej>)