Microbiome

**SHORT REPORT**                                                                                      **Open Access**

# Temporal transcriptional patterns of cyanophage genes suggest synchronized infection of cyanobacteria in the oceans

Yue Chen[1] and Qinglu Zeng[1,2,3,4*]

## Abstract

**Background:** Based on the peak expression times during infection, early, middle, and late genes have been characterized in viruses (cyanophages) that infect the unicellular cyanobacterium *Prochlorococcus*. Laboratory experiments show that some cyanophages can only replicate in the light and thus exhibit diurnal infection rhythms under light-dark cycles. Field evidence also suggests synchronized infection of *Prochlorococcus* by cyanophages in the oceans, which should result in progressive expression of cyanophage early, middle, and late genes. However, distinct temporal expression patterns have not been observed in cyanophage field populations.

**Results:** In this study, we reanalyzed a previous metatranscriptomic dataset collected in the North Pacific Subtropical Gyre. In this dataset, it was previously shown that aggregate transcripts from cyanophage scaffolds display diurnal transcriptional rhythms with transcript abundances decreasing at night. By mapping metatranscriptomic reads to individual viral genes, we identified periodically expressed genes from putative viruses infecting the cyanobacteria *Prochlorococcus* and *Synechococcus*, heterotrophic bacteria, and algae. Of the 41 cyanophage genes, 35 were from cyanomyoviruses. We grouped the periodically expressed cyanomyovirus genes into early, middle, and late genes based on the conserved temporal expression patterns of their orthologs in cyanomyovirus laboratory cultures. We found that the peak expression times of late genes in cyanophage field populations were significantly later than those of early and middle genes, which were similar to the temporal expression patterns of synchronized cyanophage laboratory cultures.

**Conclusions:** The significantly later peak expression times of late genes in cyanomyovirus field populations suggest that cyanophage infection of *Prochlorococcus* is synchronized in the North Pacific Subtropical Gyre. The night-time peak expression of late genes also suggests synchronized lysis of *Prochlorococcus* at night, which might result in synchronized release of dissolved organic matter to the marine food web.

**Keywords:** Cyanobacterium, Cyanophage, Diurnal rhythm, Light-dark cycle, Metatranscriptomics

* Correspondence: zeng@ust.hk
[1]Department of Ocean Science, The Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong, China
[2]Division of Life Science, The Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong, China
Full list of author information is available at the end of the article

# Background

To adapt to the earth's daily light-dark (diel) cycle, diurnal rhythms of gene expression, metabolism, and behavior have been shown in all three domains of life [11], including cyanobacteria [7]. The unicellular picocyanobacteria *Prochlorococcus* and *Synechococcus* are the major primary producers in the oceans [13, 33, 38]. In the North Pacific Subtropical Gyre, rhythmic transcriptional patterns have been observed in *Prochlorococcus* and several co-occurring heterotrophic bacterial groups [32]. The diurnal transcriptional rhythm of marine microbial populations is suggested to be a result of metabolic coupling between *Prochlorococcus* and heterotrophic bacteria, which may influence matter and energy transformation in the sea [32].

*Prochlorococcus* and *Synechococcus* are frequently infected by lytic double-stranded DNA viruses (cyanophages) [27, 43], which rely on the photosynthetic energy of host cells to replicate [23, 25, 40]. Cyanophage strains exhibiting three classes of diel-dependent life history traits ("diel traits") have been characterized [24]: (1) no adsorption to the host cells in the dark [9, 15, 16, 24], (2) adsorption in the dark but no replication, and (3) adsorption and replication in the dark, but with a smaller burst size than in the light. Regardless of the trait, cyanophage transcripts are more abundant during infection in the light than those in the dark, resulting in a diurnal transcriptional rhythm under light-dark cycles [24]. The diurnal rhythm of cyanophage transcripts is likely caused by the photosynthetic activity of cyanobacterial host cells, since an inhibitor of photosynthetic electron flow can also reduce cyanophage transcripts in the light [24].

The three classes of diel traits indicate that cyanophages that are unable to adsorb or replicate in the dark can only initiate the infection process in the light [24]. Therefore, in marine environments, we hypothesized that the infection process of these cyanophages should be synchronized to the light-dark cycle. Indeed, there is field evidence to suggest synchronized infection of *Prochlorococcus* and *Synechococcus* in the oceans [6, 37, 51]. In the subtropical Pacific Ocean, the mortality rate of *Prochlorococcus* was found to be synchronized to the light-dark cycle [37]. The highest mortality rate appeared at night and could be caused by synchronized cyanophage burst at night; however, the synchronized mortality rate could also be explained by synchronized grazing of protists [37]. In the Indian Ocean, infectious cyanophages from surface waters were isolated using *Synechococcus* sp. WH7803 as the host and were shown to have the highest abundance at night [6]. Yoshida et al. reasoned that synchronized infection should result in an increase of viral transcripts followed by an increase of virus particles [51]. To test this, Yoshida et al. [51]

conducted metagenomic and metatranscriptomic studies using surface seawater from Osaka Bay off Japan, and observed that viruses with relatively high transcript abundances generally increased their genome abundance ranks at a later time point. However, a higher genome abundance rank does not necessarily mean a higher virus abundance. More critically, both the Indian Ocean study [6] and the Osaka Bay study [51] were conducted only for 1 day, so it is uncertain whether the patterns shown in both studies are diurnal rhythms that can repeat for several days.

Similar to bacteriophages [30], cyanophage infection relies on a robust temporal transcriptional program [5, 10, 21, 22, 46]. According to the peak expression times during infection, cyanophage genes can be categorized into early, middle, and late genes that play different roles [10, 21, 22, 46]. For example, early genes might be involved in modifying the host RNA polymerase, middle genes are related to phage genome replication, and late genes mainly encode structural proteins for phage capsid assembly [10, 21]. In unsynchronized infections, early, middle, and late genes are being transcribed at any time, and should all reach their peak expression levels at around sunset due to the lack of photosynthetic activity of the host cells at night [24]. On the other hand, in synchronized infections, cyanophages initiate the infection process shortly after sunrise and the peak expression times of early, middle, and late genes should be discrete.

A metatranscriptomic study conducted by Aylward et al. in the North Pacific Subtropical Gyre revealed diurnal transcriptional rhythms of cyanophages, with peak transcript abundances at around sunset [1]. In this study, the authors sequenced viral metagenomes and metatranscriptomes (in the microbial cellular size fractions) within a coherent water mass at a depth of 15 m in the North Pacific Subtropical Gyre every 4 h over 8 days (July 27–31 and August 1–3, 2015) [1]. By mapping viral transcripts to the 170 viral scaffolds assembled from viral metagenomes, they identified 26 viral scaffolds with rhythmic aggregate transcript abundances, including 17 scaffolds similar to cyanophage genomes, one similar to *Pelagibacter* phage (pelagiphage), and several unclassified ones [1]. However, it was not shown in this study whether individual viral genes in a scaffold show diurnal transcriptional rhythms, as the combined transcripts mapped to all the genes in a scaffold were used for diel periodicity analyses [1].

Here, we explored the synchronization of cyanophage infection in marine environments on a gene-by-gene (rather than total abundance) basis. To do so, we re-examined the expression of cyanophage genes from the metatranscriptomic study conducted by Aylward et al. [1]. Building upon Aylward's finding of peaks in total transcriptional activity, we found that the early, middle,

and late genes of natural cyanophage populations show distinct peak expression times, suggesting that cyanophage infection is synchronized in the surface ocean. Since *Prochlorococcus* is the dominant primary producer in the North Pacific Subtropical Gyre [32], synchronized infection of *Prochlorococcus* may affect carbon cycling in this region.
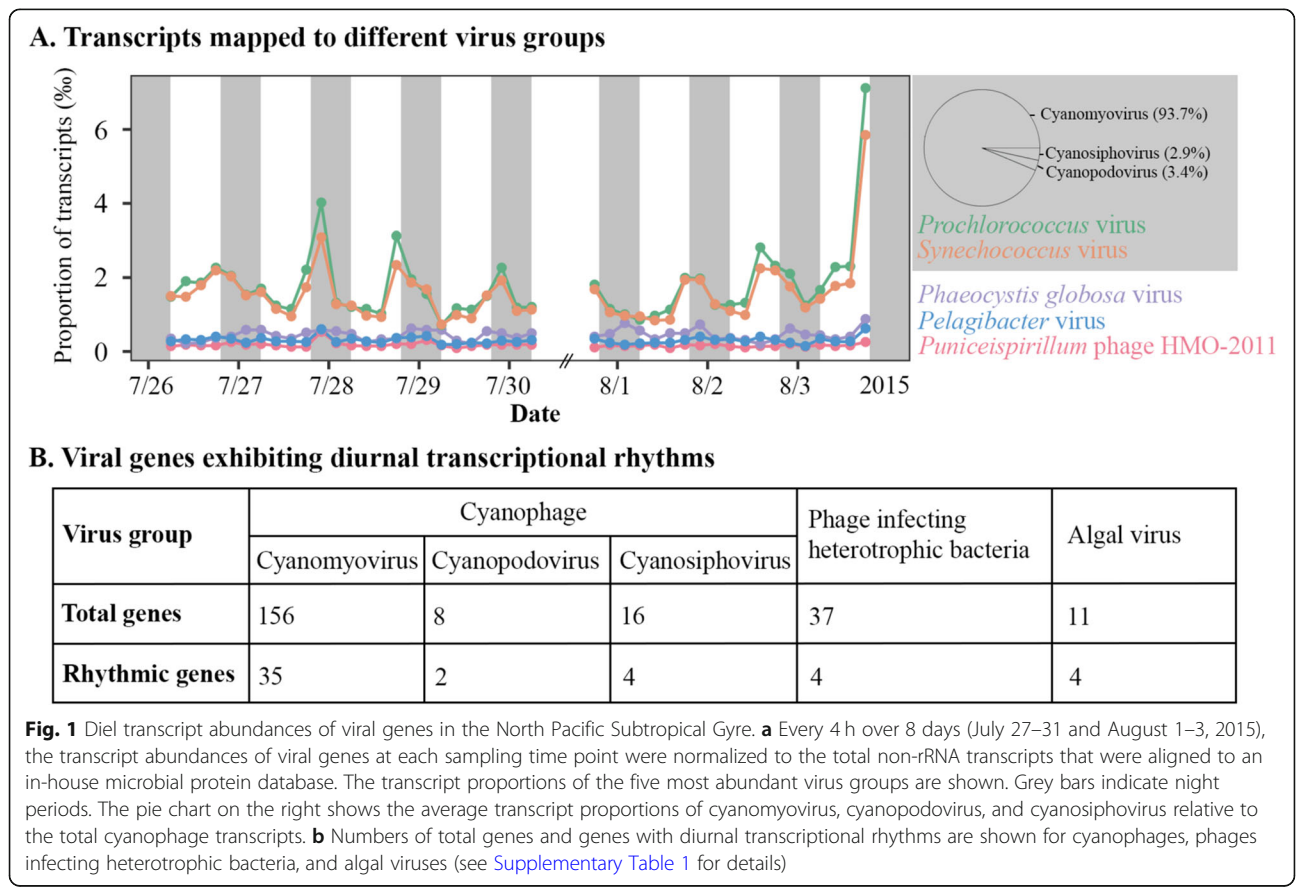
## Results

### Viral genes in the North Pacific Subtropical Gyre show diurnal transcriptional rhythms

To test whether individual viral genes showed diurnal transcriptional rhythms, we reanalyzed the metatranscriptomic data from Aylward et al. [1] by mapping transcripts to viral genes (see "Materials and methods" section). Of the total transcripts in the microbial cellular size fractions, transcripts mapped to *Prochlorococcus* and *Synechococcus* phages were the most abundant viral transcripts, followed by those from *Phaeocystis globosa* viruses, pelagiphages, and *Puniceispirillum* phage HMO-2011 (Fig. 1a). Total cyanophage transcripts in the field populations increased during the day and decreased at night (Fig. 1a), which was similar to what we have observed in laboratory cultures and likely explained by diel changes in cyanobacterial photosynthesis [24]. After diel

periodicity test (see "Materials and methods" section; Supplementary Table 1), 41 out of 180 cyanophage genes (23%) were found to have diurnal transcriptional rhythms (Fig. 1b and Supplementary Figures 1A, B, C). In addition, diurnal transcriptional rhythms were shown for 4 out of 37 genes (11%) of phages infecting heterotrophic bacteria, and 4 out of 11 genes (36%) of algal viruses (Fig. 1b and Supplementary Figures 1D, E). Therefore, our analyses revealed diurnal transcriptional rhythms of genes in various virus groups in the North Pacific Subtropical Gyre.

### Cyanophage genes with diurnal transcriptional rhythms

Cyanophage transcripts were mainly from cyanomyoviruses (93.7%), with a small portion from cyanopodoviruses and cyanosiphoviruses (Fig. 1a). In cyanomyoviruses, cyanopodoviruses, and cyanosiphoviruses, 35, 2, and 4 periodically expressed genes were identified, respectively (Fig. 1b). In cyanomyoviruses (Supplementary Table 1), the periodically expressed genes encode proteins for unknown function (16 genes), nucleic acid metabolism (9 genes), and phage capsid assembly (4 genes). Six auxiliary metabolic genes [4, 48, 49] were also found to be periodically expressed (Supplementary Table 1). These genes are involved in



**A. Transcripts mapped to different virus groups**

Cyanomyovirus (93.7%)
Cyanosiphovirus (2.9%)
Cyanopodovirus (3.4%)

*Prochlorococcus* virus
*Synechococcus* virus
*Phaeocystis globosa* virus
*Pelagibacter* virus
*Puniceispirillum* phage HMO-2011

**B. Viral genes exhibiting diurnal transcriptional rhythms**

| Virus group | Cyanophage | | | Phage infecting heterotrophic bacteria | Algal virus |
|---|---|---|---|---|---|
| | Cyanomyovirus | Cyanopodovirus | Cyanosiphovirus | | |
| **Total genes** | 156 | 8 | 16 | 37 | 11 |
| **Rhythmic genes** | 35 | 2 | 4 | 4 | 4 |

**Fig. 1** Diel transcript abundances of viral genes in the North Pacific Subtropical Gyre. **a** Every 4 h over 8 days (July 27–31 and August 1–3, 2015), the transcript abundances of viral genes at each sampling time point were normalized to the total non-rRNA transcripts that were aligned to an in-house microbial protein database. The transcript proportions of the five most abundant virus groups are shown. Grey bars indicate night periods. The pie chart on the right shows the average transcript proportions of cyanomyovirus, cyanopodovirus, and cyanosiphovirus relative to the total cyanophage transcripts. **b** Numbers of total genes and genes with diurnal transcriptional rhythms are shown for cyanophages, phages infecting heterotrophic bacteria, and algal viruses (see Supplementary Table 1 for details)

photosynthetic electron transport (plastoquinol terminal oxidase and NAD synthetase), central carbon metabolism (UDP-glucose 4-epimerase and glycerol-3-phosphate cytidylyltransferase), and stress responses (heat shock protein IbpA and P-starvation inducible protein PhoH) (Supplementary Table 1). In cyanopodoviruses, the two periodically expressed genes were both related to nucleic acid metabolism (Supplementary Table 1). In cyanosiphoviruses, two periodically expressed genes were related to nucleic acid metabolism and the other two related to capsid assembly (Supplementary Table 1).

It has been identified using cultured cyanomyoviruses that nucleic acid metabolism genes belong to middle genes and capsid assembly genes belong to late genes [10, 21]. Interestingly, in cyanomyovirus field populations, most genes related to nucleic acid metabolism showed peak transcript abundances before sunset (local time 19:11), while genes related to capsid assembly all peaked after sunset (Fig. 2). Similarly, in cyanosiphovirus field populations, the peak expression times of nucleic acid metabolism genes were also earlier than those of capsid assembly genes (Fig. 2).

### Peak expression times of early, middle, and late genes in cyanomyovirus field populations

We further tested whether the peak expression times of cyanomyovirus genes in the field populations correlated with the temporal expression patterns of their orthologs in cyanomyovirus laboratory cultures. Currently, the temporal expression patterns of cyanomyoviruses P-SSM2 [21], P-HM2 [47], P-TIM40, and Syn9 [10] have been characterized. In these laboratory experiments, cyanophage infection was synchronized so that the infecting cyanophages started the transcriptional program at the same time and the peak expression times of individual viral genes can be distinguished in bulk cultures. We were able to group 21 out of the 35 periodically expressed cyanomyovirus genes of field populations into early (3 genes), middle (11 genes), and late (7 genes) genes according to the conserved temporal expression patterns of their orthologs in the four cultured cyanomyoviruses (Fig. 2). The peak expression times of early and middle genes were not significantly different (Fig. 3). One potential explanation for this finding is that the 90% confidence interval of the calculated peak expression times of the periodically expressed genes is ∼ 38 min (Supplementary Figure 2F) and may not be able to resolve smaller differences in expression time between the early and middle genes. Nevertheless, the peak expression times of the 7 late genes were significantly later than those of early and middle genes, suggesting that the temporal expression patterns of cyanomyovirus field populations followed those of synchronized cyanophage laboratory cultures [5, 10, 21, 47].

## Discussion

In this study, we reanalyzed the metatranscriptomic data from microbial cellular size fractions collected in the North Pacific Subtropical Gyre [1]. By mapping transcripts to individual viral genes, we identified 49 viral genes with diurnal transcriptional rhythms, including 41 genes from cyanophages, four from phages infecting heterotrophic bacteria, and four from algal viruses (Fig. 1b). For the periodically expressed cyanomyovirus genes, we found that the peak expression times of late genes were significantly later than those of early and middle genes (Fig. 3). In cyanophage laboratory cultures, the peak expression times of early, middle, and late genes were distinguished when cyanobacterial cells were infected at the same time [10, 21, 22, 46]. Therefore, the much later peak expression times of late genes in cyanomyovirus field populations suggest that cyanophage infection of cyanobacteria is synchronized in the North Pacific Subtropical Gyre.

Although our results provided evidence that cyanophage infection is synchronized in the North Pacific Subtropical Gyre, we were unable to use metatranscriptomic data to estimate the percentage of cyanophages that underwent synchronized infection. Future isolation-based studies and single-cell transcriptomic analyses are needed to measure synchronized and unsynchronized infections in the field populations.

Besides phages infecting *Prochlorococcus* and *Synechococcus*, we also identified periodically expressed genes from viruses infecting algae and heterotrophic bacteria. It is well known that marine algae exhibit circadian rhythms of growth [14] and gene expression [31]. Thus, similar to cyanophages [24], gene expression of viruses infecting *Aureococcus anophagefferens*, *Chrysochromulina ericina*, and *Phaeocystis globosa* (Supplementary Figures 1E) might also be regulated by the photosynthetic activity of algal host cells. Consistent with previous studies [1, 51], we identified rhythmic gene expression from a SAR11 virus *Pelagibacter* phage HTVC008M, and also from bacteriophages infecting enterobacteria *Cronobacter sakazakii* and *Klebsiella* (Supplementary Figure 1D). Diurnal transcriptional rhythms of heterotrophic bacteria were observed in the North Pacific Subtropical Gyre and were thought to be caused by rhythmic release of organic compounds from *Prochlorococcus* cells [32]. Therefore, the transcriptional rhythms of viruses infecting heterotrophic bacteria may also be due to rhythmic pulses of organic compounds. It should be noted that the periodically expressed genes from algal and bacterial viruses were identified by our metatranscriptomic analysis and their expression patterns need to be verified in future laboratory experiments.
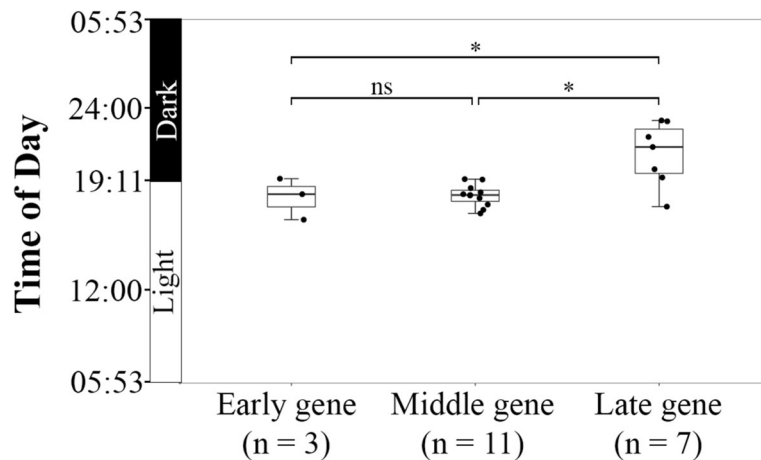
**Fig. 2** Peak expression times of periodically expressed cyanophage genes. The jackknife distributions of the estimated peak expression times of each gene are shown in violin plots. The red vertical line in each violin plot shows the peak expression time of each gene estimated from all data points and the two black vertical lines indicate the 90% confidence interval of the peak expression times estimated using the jackknife method from partial data points (Supplementary Figure 2). The color of the ortholog name denotes different cyanophage groups: blue for cyanomyovirus, green for cyanopodovirus, and orange for cyanosiphovirus. On the right of each ortholog name, E, M, L indicates early, middle, and late temporal expression classes, respectively. The double helix symbol indicates genes involve in nucleic acid metabolism and the hexagon symbol indicates genes involved in phage capsid assembly

Our metatranscriptomic analyses showed that the cyanomyovirus late genes mostly had their peak expression times in the dark period of a day and at a much later time than those of early and middle genes (Fig. 3). Our results not only suggest that cyanophage infection is synchronized to the light-dark cycle, but also hint at synchronized lysis of *Prochlorococcus* cells at night, which is consistent with the highest abundance of cyanophages at

**Fig. 3** Comparison of the peak expression times of early, middle, and late genes in cyanomyovirus field populations. As shown in Fig. 2, the periodically expressed cyanomyovirus genes are classified as early (3 genes), middle (11 genes), and late genes (7 genes) according to the temporal expression patterns of their orthologs in cultured cyanophages. The peak expression times of early, middle, and late genes in cyanomyovirus field populations are shown in box and whisker plots. The bottom and top of a box indicate the first and third quartiles, respectively. The band inside a box indicates the median and the ends of a whisker represent the upper and lower extremes. Black dots indicate the estimated peak expression times of cyanomyovirus genes. The peak expression times of early, middle, and late genes are compared using the two-tailed Student's test and the *p* values are indicated on top of the plots (*indicates $p < 0.05$ and ns indicates no significance)

night in the Indian Ocean [6]. Our results help explain the light-dark synchronized mortality of *Prochlorococcus* cells in surface waters, which also reaches a peak at night [37]. Synchronized infection of *Prochlorococcus* by cyanophages may influence carbon cycling in the world's oceans by synchronized release of dissolved organic matter to the marine food web during daily light-dark cycles.

## Conclusions

To demonstrate the synchronization of cyanophage infection in marine environments, our current study builds upon a previous metatranscriptomic dataset collected by Aylward et al. in the North Pacific Subtropical Gyre. By mapping metatranscriptomic reads to individual viral genes, we identified periodically expressed cyanophage genes. We also found that late genes (7 genes) in cyanophage field populations had a significantly later average peak expression time than those of early and middle genes, suggesting that cyanophage infection of *Prochlorococcus* is synchronized in the surface oceans. Our results are consistent with the light-dark synchronized population dynamics of *Prochlorococcus* cells in the oceans and have implications for *Prochlorococcus*-driven organic carbon production in the marine food web.

## Materials and methods

### Mapping metatranscriptomic reads to viral genes and orthologs

In the study conducted by Aylward et al. [1], microbial communities from > 0.2 μm size fractions were collected at 15 m depth in the North Pacific Subtropical Gyre

every 4 h from July 26 to July 31 and from August 1 to August 3 2015. The metatranscriptomic data generated from these samples were downloaded from the NCBI Sequence Read Archive under accession number PRJNA358725.

Following the published protocols [2], downloaded sequences were quality-filtered and aligned to an in-house protein database. In brief, adaptor sequences and low-quality regions were trimmed using Trim_Galore (https://github.com/FelixKrueger/TrimGalore). Reads with low complexity and low quality were removed using PRINSEQ [39]. SortMeRNA [18] was then used to filter out rRNA reads based on alignment to the SILVA rRNA database (release 132) [35] and the Rfam 5S rRNA database [29]. After the removal of rRNA reads, the remaining transcripts were aligned to the in-house microbial protein database using LAST [17] with a bit score cutoff of 50. The microbial protein database included RefSeq Archaea, RefSeq Bacteria, RefSeq Protozoa, RefSeq Plant, and RefSeq Viral (RefSeq release 89), together with proteins of several additional cyanophages that were not in the RefSeq database during our analyses [8, 19, 34, 42]. Ambiguous sequences were discarded without further analysis if they were aligned equally well to proteins of multiple virus groups or to both a virus and its host.

Transcripts mapped to cyanophage genes were binned into cyanophage orthologs. When our study was conducted, there were 93 sequenced genomes of cyanophages in public databases that infect *Prochlorococcus* and *Synechococcus*. The gene orthologs (Supplementary

Table [2]) of the 93 cyanophage genomes were constructed using ProteinOrtho5 [20]. Transcripts with best hits to multiple orthologs were discarded. For viruses other than cyanophages, transcripts were not binned into orthologs, since there were not many sequenced genomes to construct gene orthologs.

## Identification of periodically expressed genes

To identify periodically expressed genes, the nonparametric RAIN package [45] was implemented in R [36] to test whether there is a 24-h rhythm in the transcript abundances of viral genes. The periodicity test was first conducted for cyanophage orthologs with at least 44 transcripts during the sampling period (an average of one transcript per sampling time point). The periodicity test was also conducted for genes of other virus groups with at least 44 transcripts. In the periodicity test, the transcript number of a gene/ortholog at each time point was normalized to the total non-rRNA transcripts that can be aligned to the in-house microbial protein database. Following a published protocol [1], a standard FDR correction [3] was applied to the *p* values generated by the RAIN package to remove false positives. The FDR-corrected *p* values under 0.1 denote significant periodicity and are shown in Supplementary Table [1].

## Peak expression times of periodically expressed genes

The average peak expression time of each periodically expressed viral gene during the sampling period was estimated based on the regression to a sinusoidal function. The Poisson log-linear regression to Eq. [1] was conducted on the transcript number at each time point to generate a best fit. The offset of the regression is the total number of transcripts with significant hits to proteins in our database. In Eq. [1], $x_t$ represents the number of transcripts aligned to each gene at time $t$, and $\alpha$ and $\beta$ were used to calculate the amplitude ($\sqrt{\alpha^2 + \beta^2}$) and the phase shift ($\frac{12}{\pi}$ $\tan^{-1}\frac{\beta}{\alpha}$) of the sinusoidal function. The significance of the fit was assessed using the chi-squared test as described previously [31] (implemented using the anova.glm function in R) [36]. The FDR corrected [3] *p* value from the chi-squared test under 0.1 implies a significant fit, and all periodic genes show significant fits (Supplementary Table [1]). Based on the best-fit equation, the peak expression time ($t_{peak}$) of each gene was calculated from Eq. [2].

$$x_t = \alpha\cos\left(\frac{2\pi}{24}t\right) + \beta\sin\left(\frac{2\pi}{24}t\right) \qquad (1)$$

$$t_{peak} = 6 - \frac{\pi}{12}\tan^{-1}\left(\frac{\alpha}{\beta}\right) \qquad (2)$$

The confidence interval of the estimated peak time for each gene was assessed using the delete-2 jackknife method [12], which is a classic re-sampling method that has been widely used in various biological systems [26, 28, 41, 44, 50]. In the delete-2 jackknife analysis, subsamples with 42 time points (2 time points deleted) were randomly drawn from the original 44 time points (totally 946 subsamples). The peak expression time of each subsample was calculated using the regression method mentioned above. With the peak expression times of the 946 subsamples, the 5% and 95% quantiles were calculated to determine the 90% confidence interval of the estimated peak expression time for each gene (Supplementary Figure 2). The 90% confidence intervals for the 49 periodic genes range from 21 to 104 min, with a median of 38 min (Supplementary Figure 2F and Supplementary Table 1).

## Supplementary information

### Author details
[1]Department of Ocean Science, The Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong, China. [2]Division of Life Science, The Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong, China. [3]HKUST Shenzhen Research Institute, Shenzhen, China. [4]Hong Kong Branch of Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou), The Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong, China.

## References

1. Aylward FO, Boeuf D, Mende DR, Wood-Charlson EM, Vislova A, Eppley JM, Romano AE, Delong EF. Diel cycling and long-term persistence of viruses in the ocean's euphotic zone. Proc Natl Acad Sci U S A. 2017;114:11446–51.
2. Aylward FO, Eppley JM, Smith JM, Chavez FP, Scholin CA, Delong EF. Microbial community transcriptional networks are conserved in three domains at ocean basin scales. Proc Natl Acad Sci U S A. 2015;112:5443–8.
3. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society. Series B (Methodological). 1995;57:289–300.
4. Breitbart M, Thompson LR, Suttle CA, Sullivan MB. Exploring the vast diversity of marine viruses. Oceanography. 2007;20:135–9.
5. Clokie MR, Shan J, Bailey S, Jia Y, Krisch HM, West S, Mann NH. Transcription of a 'photosynthetic' T4-type phage during infection of a marine cyanobacterium. Environ Microbiol. 2006a;8:827–35.
6. Clokie MRJ, Millard AD, Mehta JY, Mann NH. Virus isolation studies suggest short-term variations in abundance in natural cyanophage populations of the Indian Ocean. Journal of the Marine Biological Association of the United Kingdom. 2006b;86:499–505.
7. Cohen SE, Golden SS. Circadian rhythms in cyanobacteria. Microbiol Mol Biol Rev. 2015;79:373–85.
8. Crummett LT, Puxty RJ, Weihe C, Marston MF, Martiny JB. The genomic content and context of auxiliary metabolic genes in marine cyanomyoviruses. Virology. 2016;499:219–29.
9. Cseke CS, Farkas GL. Effect of light on the attachment of cyanophage AS-1 to *Anacystis nidulans*. J Bacteriol. 1979;137:667–9.
10. Doron S, Fedida A, Hernandez-Prieto MA, Sabehi G, Karunker I, Stazic D, Feingersch R, Steglich C, Futschik M, Lindell D, Sorek R. Transcriptome dynamics of a broad host-range cyanophage and its hosts. ISME J. 2016;10:1437–55.
11. Edgar RS, Green EW, Zhao Y, Van Ooijen G, Olmedo M, Qin X, Xu Y, Pan M, Valekunja UK, Feeney KA, Maywood ES, Hastings MH, Baliga NS, Merrow M, Millar AJ, Johnson CH, Kyriacou CP, O'neill JS, Reddy AB. Peroxiredoxins are conserved markers of circadian rhythms. Nature. 2012;485:459–64.
12. Efron B, Tibshirani R. An Introduction to the Bootstrap. New York: Chapman and Hall/CRC; 1994.
13. Flombaum P, Gallegos JL, Gordillo RA, Rincon J, Zabala LL, Jiao N, Karl DM, Li WK, Lomas MW, Veneziano D, Vera CS, Vrugt JA, Martiny AC. Present and future global distributions of the marine Cyanobacteria *Prochlorococcus* and *Synechococcus*. Proc Natl Acad Sci U S A. 2013;110:9824–9.
14. Jacquet S, Partensky F, Lennon JF, Vaulot D. Diel patterns of growth and division in marine picoplankton in culture. Journal of Phycology. 2001;37:357–69.
15. Jia Y, Shan J, Millard A, Clokie MR, Mann NH. Light-dependent adsorption of photosynthetic cyanophages to *Synechococcus* sp. WH7803. FEMS Microbiol Lett. 2010;310:120–6.
16. Kao CC, Green S, Stein B, Golden SS. Diel infection of a cyanobacterium by a contractile bacteriophage. Appl Environ Microbiol. 2005;71:4276–9.
17. Kielbasa SM, Wan R, Sato K, Horton P, Frith MC. Adaptive seeds tame genomic sequence comparison. Genome Res. 2011;21:487–93.
18. Kopylova E, Noe L, Touzet H. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. Bioinformatics. 2012;28:3211–7.
19. Labrie SJ, Frois-Moniz K, Osburne MS, Kelly L, Roggensack SE, Sullivan MB, Gearin G, Zeng Q, Fitzgerald M, Henn MR, Chisholm SW. Genomes of marine cyanopodoviruses reveal multiple origins of diversity. Environ Microbiol. 2013;15:1356–76.
20. Lechner M, Findeiss S, Steiner L, Marz M, Stadler PF, Prohaska SJ. Proteinortho: detection of (co-)orthologs in large-scale analysis. BMC Bioinformatics. 2011;12:124.
21. Lin X, Ding H, Zeng Q. Transcriptomic response during phage infection of a marine cyanobacterium under phosphorus-limited conditions. Environ Microbiol. 2016;18:450–60.
22. Lindell D, Jaffe JD, Coleman ML, Futschik ME, Axmann IM, Rector T, Kettler G, Sullivan MB, Steen R, Hess WR, Church GM, Chisholm SW. Genome-wide expression dynamics of a marine virus and host reveal features of co-evolution. Nature. 2007;449:83–6.
23. Lindell D, Jaffe JD, Johnson ZI, Church GM, Chisholm SW. Photosynthesis genes in marine viruses yield proteins during host infection. Nature. 2005;438:86–9.
24. Liu R, Liu Y, Chen Y, Zhan Y, Zeng Q. Cyanobacterial viruses exhibit diurnal rhythms during infection. Proc Natl Acad Sci U S A. 2019;116:14077–82.
25. Mackenzie JJ, Haselkorn R. Photosynthesis and the development of blue-green algal virus SM-1. Virology. 1972;49:517–21.
26. Mccauley DE. Contrasting the distribution of chloroplast DNA and allozyme polymorphism among local populations of Silene alba: implications for studies of gene flow in plants. Proc Natl Acad Sci U S A. 1994;91:8127–31.
27. Mojica KD, Huisman J, Wilhelm SW, Brussaard CP. Latitudinal variation in virus-induced mortality of phytoplankton across the North Atlantic Ocean. ISME J. 2016;10:500–13.
28. Motta-Ribeiro GC, Vidal Melo MF, Jandre FC. A simplified 4-parameter model of volumetric capnograms improves calculations of airway dead space and slope of Phase III. J Clin Monit Comput. 2019.
29. Nawrocki EP, Burge SW, Bateman A, Daub J, Eberhardt RY, Eddy SR, Floden EW, Gardner PP, Jones TA, Tate J, Finn RD. Rfam 12.0: updates to the RNA families database. Nucleic Acids Res. 2015;43:D130–7.
30. Nechaev S, Severinov K. Bacteriophage-induced modifications of host RNA polymerase. Annu Rev Microbiol. 2003;57:301–22.
31. Ottesen EA, Young CR, Eppley JM, Ryan JP, Chavez FP, Scholin CA, Delong EF. Pattern and synchrony of gene expression among sympatric marine microbial populations. Proc Natl Acad Sci U S A. 2013;110:E488–97.
32. Ottesen EA, Young CR, Gifford SM, Eppley JM, Marin R 3rd, Schuster SC, Scholin CA, Delong EF. Multispecies diel transcriptional oscillations in open ocean heterotrophic bacterial assemblages. Science. 2014;345:207–12.
33. Partensky F, Hess WR, Vaulot D. *Prochlorococcus*, a marine photosynthetic prokaryote of global significance. Microbiol Mol Biol Rev. 1999;63:106–27.
34. Puxty RJ, Millard AD, Evans DJ, Scanlan DJ. Shedding new light on viral photosynthesis. Photosynth Res. 2015;126:71–97.
35. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glockner FO. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucleic Acids Res. 2013;41:D590–6.
36. Rdc T. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing; 2012.
37. Ribalet F, Swalwell J, Clayton S, Jimenez V, Sudek S, Lin Y, Johnson ZI, Worden AZ, Armbrust EV. Light-driven synchrony of *Prochlorococcus* growth and mortality in the subtropical Pacific gyre. Proc Natl Acad Sci U S A. 2015;112:8008–12.
38. Scanlan DJ, West NJ. Molecular ecology of the marine cyanobacterial genera *Prochlorococcus* and *Synechococcus*. FEMS Microbiol Ecol. 2002;40:1–12.
39. Schmieder R, Edwards R. Quality control and preprocessing of metagenomic datasets. Bioinformatics. 2011;27:863–4.
40. Sherman LA. Infection of *Synechococcus cedrorum* by the cyanophage AS-1 M. III. Cellular metabolism and phage development. Virology. 1976;71:199–206.
41. Singh R, Stine OC, Smith DL, Spitznagel JK Jr, Labib ME, Williams HN. Microbial diversity of biofilms in dental unit water systems. Appl Environ Microbiol. 2003;69:3412–20.
42. Sullivan MB, Huang KH, Ignacio-Espinoza JC, Berlin AM, Kelly L, Weigele PR, Defrancesco AS, Kern SE, Thompson LR, Young S, Yandava C, Fu R, Krastins B, Chase M, Sarracino D, Osburne MS, Henn MR, Chisholm SW. Genomic analysis of oceanic cyanobacterial myoviruses compared with T4-like myoviruses from diverse hosts and environments. Environ Microbiol. 2010;12:3035–56.
43. Sullivan MB, Waterbury JB, Chisholm SW. Cyanophages infecting the oceanic cyanobacterium *Prochlorococcus*. Nature. 2003;424:1047–51.
44. Szafranski SP, Wos-Oxley ML, Vilchez-Vargas R, Jauregui R, Plumeier I, Klawonn F, Tomasch J, Meisinger C, Kuhnisch J, Sztajer H, Pieper DH, Wagner-Dobler I. High-resolution taxonomic profiling of the subgingival microbiome for biomarker discovery and periodontitis diagnosis. Appl Environ Microbiol. 2015;81:1047–58.
45. Thaben PF, Westermark PO. Detecting rhythms in time series with RAIN. *J Biol Rhythms*. 2014;29:391–400.
46. Thompson AW, Huang K, Saito MA, Chisholm SW. Transcriptome response of high- and low-light-adapted *Prochlorococcus* strains to changing iron availability. ISME J. 2011a;5:1580–94.
47. Thompson LR, Zeng Q, Chisholm SW. Gene expression patterns during light and dark infection of *Prochlorococcus* by cyanophage. PLoS One. 2016;11:e0165375.
48. Thompson LR, Zeng Q, Kelly L, Huang KH, Singer AU, Stubbe J, Chisholm SW. Phage auxiliary metabolic genes and the redirection of cyanobacterial host carbon metabolism. Proc Natl Acad Sci U S A. 2011b;108:E757–64.
49. Xiao X, Zeng Q, Zhang R, Jiao N. *Prochlorococcus* viruses—from biodiversity to biogeochemical cycles. Science China Earth Sciences. 2018;61:1728–36.

50.  Xu X, Han Q, Lin J, Wang L, Wu F, Shang H. Grey matter abnormalities in Parkinson's disease: a voxel-wise meta-analysis. Eur J Neurol. 2019.
51.  Yoshida T, Nishimura Y, Watai H, Haruki N, Morimoto D, Kaneko H, Honda T, Yamamoto K, Hingamp P, Sako Y, Goto S, Ogata H. Locality and diel cycling of viral production revealed by a 24 h time course cross-omics analysis in a coastal region of Japan. ISME J. 2018;12:1287–95.

**Publisher's Note**

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.