Diel cycling of the cosmopolitan abundant Pelagibacter virus 37-F6: one of the most abundant viruses in Earth

Francisco Martinez-Hernandez[1Ω], Elaine Luo[4Ω], Kento Tominaga[2], Hiroyuki Ogata[3], Takashi Yoshida[2], Edward F. DeLong[4], and Manuel Martinez-Garcia[1*]

[1]Department of Physiology, Genetics, and Microbiology, University of Alicante. Alicante, Spain.

[2]Graduate School of Agriculture, Kyoto University, Kitashirakawa-Oiwake, Sakyo-ku, Kyoto 606-8502, Japan.

[3]Institute for Chemical Research, Kyoto University, Uji 611-0011, Japan.

[4]Daniel K. Inouye Center for Microbial Oceanography: Research and Education (C-MORE), University of Hawaii, Manoa, Honolulu, HI 96822, USA

[Ω]These authors contributed equally to this work

*** Correspondence:**
Manuel Martinez-Garcia
m.martinez@ua.es

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/1758-2229.12825

Originality-Significance Statement

In this study, coastal and oceanic metatranscriptomic time-series data are used to unveil the diel cycling and temporal variations in one of the most abundant marine viruses in Earth and its putative *Pelagibacter* spp. host that altogether have a major impact on biogeochemical cycles.

Summary

The spatiotemporal dynamics for marine viral populations has only recently been explored. However, nothing is known about temporal activities of the uncultured Pelagibacter virus vSAG 37-F6, which was discovered by single-virus genomics as potentially the most abundant marine virus. Here, we investigate the diel cycling of 37-F6 virus and the putative SAR11 host using coastal and oceanic transcriptomic and viromic time-series data from Osaka Bay and North Pacific Subtropical Gyre. Virus 37-F6 and relatives displayed diel cycling of transcriptional activities synchronized with its putative host. In both virus and host, the lowest transcription rates were observed at 14:00-15:00, coinciding roughly with maximum solar irradiance, while higher transcriptional rates were detected during the night/early morning and afternoon. Diel abundance of free viruses of 37-F6 in seawater roughly mirrored the transcriptional activities of both virus and host. In Osaka Bay, among viral relatives (genus level), virus 37-F6 specifically showed the highest ratio of transcriptional activity to virome abundance, a proxy for viral transcriptional activity relative to free viral particle abundance. This high ratio suggests high infection rate efficiencies in vSAG 37-F6 virus compared to viral relatives. Thus, time-series data revealed temporal transcript activities in one of the most abundant viruses in Earth.

Viruses are paramount for structuring microbial populations and nowadays their ecological role and impact on ecosystems is unquestionable (Danovaro *et al.*, 2011). Over the last years, metagenomics have provided a wealth of information that have been extremely useful to build more accurate ecological models (Brum and Sullivan, 2015). In marine ecosystems, in part thanks to sampling efforts with geographical coverage, such as *Tara* or *Malaspina,* the biogeography and diversity of viruses are starting to be explored (Mizuno *et al.*, 2013a; Paez-Espino *et al.*, 2016; Roux *et al.*, 2016; Lara *et al.*, 2017; Gregory *et al.*, 2019).

Recently, metatranscriptomics is beginning to reveal transcriptional dynamics in a range of abundant marine viruses *in situ* (Aylward *et al.*, 2017; Yoshida *et al.*, 2018). Viral infection can be assessed by monitoring viral transcription in the host, for example by studying a marker gene for a specific virus (Kimura *et al.*, 2012). However, metatranscriptomics targeting a whole range of viruses (Moniruzzaman *et al.*, 2017; Stough *et al.*, 2018) is likely the most powerful approach to evaluate viral population dynamics in natural communities. Currently, the Osaka Bay and North Pacific Subtropical Gyre high-resolution temporal transcriptomic time-series are of great value for the viral ecology community (Aylward *et al.*, 2017; Yoshida *et al.*, 2018). These pioneer studies on marine ecosystems begin to unveil the activity and production of viruses. For instance, in Osaka Bay,

the transcriptional dynamics of ≈4,000 marine viruses available in public databases along with 878 local viral contigs assembled from Osaka Bay revealed that all local viruses were active and transcribed in the corresponding host fraction (Yoshida *et al.*, 2018). Data from the North Pacific Subtropical Gyre suggested a tightly synchronized diel coupling of viral and cellular replication cycles in both photoautotrophic and heterotrophic bacterial hosts (Aylward *et al.*, 2017).

However, none of these previously reported transcriptomic studies addressed the diel cycling of potentially the most abundant and cosmopolitan marine dsDNA virus in surface ocean, named virus vSAG 37-F6 (Martinez-Hernandez *et al.*, 2017). This Pelagibacter virus was discovered by applying a novel single-virus genomic (SVG) approach (Martinez-Hernandez *et al.*, 2017; Martinez-Hernandez *et al.*, 2019) around the same time of the publication of these reports (Aylward *et al.*, 2017; Yoshida *et al.*, 2018). The SVG based technology enabled the discovery of some of the most abundant marine viral species in surface ocean (Martinez-Hernandez *et al.*, 2017). In particular, the virus vSAG 37-F6 showed the highest recruitment virome rate in *Tara* ocean samples. Intriguingly, this virus is difficult to assemble using standard metagenomic approaches, and thus overlooked by current metagenomic techniques for many years (Martinez-Hernandez *et al.*, 2017). Virome simulation data indicated that 37-F6 vSAGs exhibited high microdiversity, which hindered its assembly and identification from metagenomes (Martinez-Hernandez *et al.*, 2017). A follow-up study showed that 37-F6 virus and relatives were present in sorted single *Pelagibacter* cells (Martinez-Hernandez *et al.*, 2019) while recent transcriptomic data provided first evidence of high *in situ* activity in coastal temperate waters of the NE Atlantic (Alonso-Sáez

et al., 2018). Recently, droplet digital PCR (dPCR) estimated that the absolute abundances of free viral particles of virus 37-F6 in the Mediterranean Sea and the Gulf of Maine ranged between hundreds to thousands (360-8,510) viruses mL-1 seawater; within the abundance range of the pelagiphage isolate HTVC010P (Martinez-Hernandez, *et al.*, 2019). dPCR data also showed that the number of infecting 37-F6 viruses in coastal bacterioplankton (i.e. genome copies in cell fraction) was 450–3480/ml with a total of ≈10–400 putative infected cells per mL (McMullen *et al.*, 2019). Thus, considering the widespread abundance of this virus and its host in all temperate and subtropical oceanic regions, as well as recent estimations of potential C release of 0.12–4.9 pg/ml through 37-F6 viral infection (McMullen *et al.*, 2019), 37-F6 viral population appears to have a major contribution to marine C cycles (Martinez-Hernandez *et al.*, 2017).

To further characterize the temporal dynamics of viral activities of this virus, we looked for 37-F6 viral transcripts in available transcriptomics time-series data (24 h time-course sampling) of Osaka Bay and North Pacific Subtropical Gyre. It is important to note here that unlike other techniques, such as digital PCR, transcriptomic abundances represent relative abundance of sequenced transcripts in a sample, though absolute abundances can be calculated using standards (Aylward *et al.*, 2017). Osaka Bay is a typical eutrophic environment owing to the inputs of nutrients from rivers, although it is also affected by an oligotrophic warm current intruding into the bay through the Kii Channel (Ozaki *et al.*, 2004). The North Pacific Subtropical Gyre is a representative habitat of oligotrophic oceans that cover ≈40% of the Earth´s surface. For instance, the chlorophyll a content in the analysed North Pacific Subtropical Gyre transcriptomic samples (0.21 ± 0.044 SD µg/L; available at

http://scope.soest.hawaii.edu/data/hoelegacy/) (Aylward *et al.*, 2017) were very similar to that of the western oligotrophic Mediterranean Sea, from where 37-F6 virus was obtained (Martinez-Hernandez *et al.*, 2017).

Transcriptomic reads mapping to virus vSAG 37-F6 were filtered at nucleotide identities of ≥95% and ≥70%, as per previous publications, to target reads from viruses identical to or within the same species than our 37-F6 vSAGs and from the same genus or subfamily taxa level, respectively (Mizuno *et al.*, 2013b; Roux *et al.*, 2016, 2019; Martinez-Hernandez *et al.*, 2017). Fig. 1 depicts the temporal transcriptomic dynamics of 37-F6 viral species population and other viral relatives of 37-F6, coupled with diel cycling of transcriptional activities of the putative SAR11 host in North Pacific (Fig. 1A) and Osaka Bay (Fig 1B). Both time-series datasets point to a similar trend in temporal variation and uncovered a marked, synchronized diel cycling pattern of 37-F6 viruses at the species and genus levels with the SAR11 putative host (Fig. 1). In both environments, a tightly synchronized diel coupling of viral and host cellular replication cycles was observed (Fig. 1A and 1B). A sharp decrease in transcriptional activity between 14:00-15:00 in both locations coincides roughly with the maximum solar irradiance rate. In the North Pacific Subtropical Gyre, transcriptional activities increased during the night reaching a primary transcriptional peak at 6:00 (before dawn). A secondary transcriptional peak, lower than the primary early-morning peak, was observed during the afternoon. Although the general 37-F6 diel cycling trend observed in North Pacific Subtropical Gyre and Osaka Bay data series were similar and in all cases, viruses mirrored that of the putative SAR11 host, there were also some differences, observed during the afternoon. For instance, in Osaka Bay, dynamics were more

complex and a steep increase/decrease in transcriptional activities was observed (between 15:00 and 21:00). On the other hand, viral and host transcriptional activities in the North Pacific, remained more or less stable. This could be the result of subtle site-specific differences of viral populations from Osaka Bay and North Pacific Subtropical Gyre and also their contrasting trophic status; the former much more eutrophic.

Mapped transcriptomic reads covered most 37-F6 viral genes (Fig. S1). For instance, in North Pacific Subtropical Gyre samples (Fig. 1A), in which 37-F6 was more active (e.g. sample CSHLIIR00-20a-S20C001-0015), 82% of its genes were detected to be transcribed at the genus level (Fig. S1). In Osaka Bay, several assembled viral contigs obtained in a previous publication (Yoshida *et al.*, 2018) were related to virus 37-F6 (Fig. 1B) at an average amino acid identities (AAI) between 50 and 81% with very high gene synteny (Fig. S2), and in all cases, at <95% average genome nucleotide identities (gANI), approximating genus-level relatedness or even probably in some cases to subfamily level (Mizuno *et al.*, 2013b; Roux *et al.*, 2016, 2019; Martinez-Hernandez *et al.*, 2017). In addition, most of these viral contigs from Osaka Bay related to vSAG 37-F6 had the specific capsid marker gene, representing viral group 37-F6 (gene 9), which was the most abundant marine viral protein (Brum *et al.*, 2016; Martinez-Hernandez *et al.*, 2017). The amino acid similarity of that shared capsid hallmark gene ranged from 72-95%, within values obtained previously for other reported 37-F6 viral relatives found in single-cells or cloned in fosmids (Martinez-Hernandez *et al.*, 2017, 2019a). Among all assembled viral contigs from Osaka Bay, the virus OBV_N00758 was the closest one to vSAG 37-F6 (Figs. 1B and S2) showing a perfect gene synteny and a value of AAI of 81% (Fig. S2). Although sequence-based virus taxonomy is

under debate (Simmonds *et al.*, 2017; Roux *et al.*, 2019), our molecular data (both nucleotide and amino acid level), marker capsid gene of virus 37-F6 (gene 9), and gene synteny data taken together suggest that these viral contigs in Osaka Bay belonged to same genus. However, it cannot be ruled out that some of these viruses might even belong to the same sub-family, as previously discussed (Roux *et al.*, 2016; Martinez-Hernandez *et al.*, 2017).

In Osaka Bay, in addition to transcriptional activities, virome data (i.e. abundance of free viruses in seawater) was available, representing a unique opportunity to compare the diel cycling abundance of virus 37-F6 in seawater, using virome fragment recruitment, with its transcriptional dynamics along with that of its putative host (Fig. 2). The abundance peak of free viral particles of 37-F6 coincided with that of the viral and host transcriptional activity peak, while the lowest number of free viruses were observed in tandem with decreased transcriptional activities. Thus, viral production and released of viral particles are tightly synchronized with viral and cellular replication. Furthermore, as shown in Supplementary Fig. 3, among all viral relatives of virus vSAG 37-F6 detected in Osaka Bay virome datasets (Yoshida *et al.*, 2018), specifically virus 37-F6 was not the most abundant one. However, after normalizing viral abundances by transcriptional activities, a proxy for viral transcriptional activity relative to free viral particle abundance, virus vSAG 37-F6 showed the highest fragments per kilobase per million reads mapped (FPKM) ratio amongst all viral relatives. This suggests that vSAG 37-F6 displays high infection efficiencies and/or high successful infection progression per each viral-host encounter. Finally, no correlation of transcriptional activities was obtained with chlorophyll a content (r^2<0.02) in water masses and samples (Fig. S4).

Diel cycling in heterotrophic bacteria has been observed in oligotrophic marine environments (Ruiz-González *et al.*, 2012; Ottesen *et al.*, 2014). Interestingly the diel cycling of the transcriptional activities observed here for virus 37-F6 and relatives and its putative *Pelagibacter* spp. host mirrored those diel patterns observed in marine heterotrophic bacterioplankton by bulk 3H-leucine and 3H-thymidine incorporation (see Figure 2 in (Ruiz-González *et al.*, 2012). Indeed, patterns of diel abundance of SAR11 cells (by means of CARD-FISH; see Figure 6B in (Ruiz-González *et al.*, 2012)) from the same location where virus vSAG 37-F6 was obtained (Blanes Bay Microbial Observatory), are in concordance with our transcriptomic data of 37-F6. Both datasets show highest SAR11 abundances at night, around 1:00-3:00, where viral transcription of 37-F6 is higher, and lowest SAR11 numbers during the day around 13:00-15:00, similar to 37-F6 viral activity and SAR11 host (Fig. 1). Additionally, in Osaka Bay, diel changes of relative abundances of Alphaproteobacteria group (Yoshida *et al.*, 2018) followed a similar trend observed here for SAR11.

Overall, we show here that omic data can shed light on key biological aspects of temporal activities of an uncultured virus that has a major impact on the `viral shunt´ and on biogeochemical cycles across global oceans. The use of complementary technologies applied in high-resolution, comprehensive time-series studies of marine ecosystems should help advance understanding of microbial dynamics in the ocean.

Acknowledgments

This work has been supported by Spanish Ministry of Economy and Competitiveness to MMG (ref. RTI2018-094248-B-100), Generalitat Valenciana to FMH (ACIF/2015/332),

and Gordon and Betty Moore Foundation to MMG (grant 5334). Gordon and Betty Moore Foundation to EFD (3777) and Simons Foundation Grant #329108 (to EFD).

Conflict of Interest Statement

The authors declare that they have no conflict of interest.

References

Alonso-Sáez, L., Anxelu Morán, X.G., Clokie, M.R., and Alonso-Sáez laura, L. (2018) Low activity of lytic pelagiphages in coastal marine waters. *ISME J* **12**: 2100–2102.

Aylward, F.O., Boeuf, D., Mende, D.R., Wood-Charlson, E.M., Vislova, A., Eppley, J.M., et al. (2017) Diel cycling and long-term persistence of viruses in the ocean's euphotic zone. *Proc Natl Acad Sci U S A* **114**: 11446–11451.

Brum, J.R., Ignacio-Espinoza, J.C., Kim, E.-H., Trubl, G., Jones, R.M., Roux, S., et al. (2016) Illuminating structural proteins in viral "dark matter" with metaproteomics. *Proc Natl Acad Sci U S A* **113**: 2436–2441.

Brum, J.R. and Sullivan, M.B. (2015) Rising to the challenge: accelerated pace of discovery transforms marine virology. *Nat Rev Microbiol* **13**: 147–159.

Danovaro, R., Corinaldesi, C., Dell'anno, A., Fuhrman, J.A., Middelburg, J.J., Noble, R.T., and Suttle, C.A. (2011) Marine viruses and global climate change. *FEMS Microbiol Rev* **35**: 993–1034.

Gregory, A.C., Zayed, A.A., Conceição-Neto, N., Temperton, B., Bolduc, B., Alberti, A., et

al. (2019) Marine DNA Viral Macro- and Microdiversity from Pole to Pole. *Cell* **177**: 1109-1123.e14.

Kimura, S., Yoshida, T., Hosoda, N., Honda, T., Kuno, S., Kamiji, R., et al. (2012) Diurnal infection patterns and impact of Microcystis cyanophages in a Japanese pond. *Appl Environ Microbiol* **78**: 5805–11.

Lara, E., Vaqué, D., Sà, E.L., Boras, J.A., Gomes, A., Borrull, E., et al. (2017) Unveiling the role and life strategies of viruses from the surface to the dark ocean. *Sci Adv* **3**: e1602565.

Martinez-Hernandez, F., Fornas, O., Lluesma Gomez, M., Bolduc, B., de la Cruz Peña, M.J., Martínez, J.M., et al. (2017) Single-virus genomics reveals hidden cosmopolitan and abundant viruses. *Nat Commun* **8**: 15892.

Martinez-Hernandez, F., Fornas, Ò., Lluesma Gomez, M., Garcia-Heredia, I., Maestre-Carballa, L., López-Pérez, M., et al. (2019) Single-cell genomics uncover Pelagibacter as the putative host of the extremely abundant uncultured 37-F6 viral population in the ocean. *ISME J* **13**: 232–236.

Martinez-Hernandez, F., Garcia-Heredia, I., Lluesma Gomez, M., Maestre-Carballa, L., Martínez Martínez, J., and Martinez-Garcia, M. (2019) Droplet Digital PCR for Estimating Absolute Abundances of Widespread Pelagibacter Viruses. *Front Microbiol* **10**: 1226.

McMullen, A., Martinez-Hernandez, F., and Martinez- Garcia, M. (2019) Absolute

quantification of infecting viral particles by chip-based digital polymerase chain reaction. *Environ Microbiol Rep* **11**: 1758-2229.12804.

Mizuno, C.M., Rodriguez-Valera, F., Kimes, N.E., and Ghai, R. (2013b) Expanding the marine virosphere using metagenomics. *PLoS Genet* **9**: e1003987.

Moniruzzaman, M., Wurch, L.L., Alexander, H., Dyhrman, S.T., Gobler, C.J., and Wilhelm, S.W. (2017) Virus-host relationships of marine single-celled eukaryotes resolved from metatranscriptomics. *Nat Commun* **8**: 16054.

Ottesen, E.A., Young, C.R., Gifford, S.M., Eppley, J.M., Marin, R., Schuster, S.C., et al. (2014) Multispecies diel transcriptional oscillations in open ocean heterotrophic bacterial assemblages. *Science (80-)* **345**: 207–212.

Ozaki, K., Uye, S.-I., Kusumoto, T., and Hagino, T. (2004) Interannual variability of the ecosystem of the Kii Channel, the Inland Sea of Japan, as influenced by bottom intrusion of cold and nutrient-rich water from the Pacific Ocean, and a recent trend of warming and oligotrophication. *Fish Oceanogr* **13**: 65–79.

Paez-Espino, D., Eloe-Fadrosh, E.A., Pavlopoulos, G.A., Thomas, A.D., Huntemann, M., Mikhailova, N., et al. (2016) Uncovering Earth's virome. *Nature* **536**: 425–430.

Roux, S., Adriaenssens, E.M., Dutilh, B.E., Koonin, E. V, Kropinski, A.M., Krupovic, M., et al. (2019) Minimum Information about an Uncultivated Virus Genome (MIUViG). *Nat Biotechnol* **37**: 29–37.

Roux, S., Brum, J.R., Dutilh, B.E., Sunagawa, S., Duhaime, M.B., Loy, A., et al. (2016)

Ecogenomics and potential biogeochemical impacts of globally abundant ocean viruses. *Nature* **537**: 689–693.

Ruiz-González, C., Lefort, T., Massana, R., Simó, R., and Gasol, J.M. (2012) Diel changes in bulk and single-cell bacterial heterotrophic activity in winter surface waters of the northwestern Mediterranean sea. *Limnol Oceanogr* **57**: 29–42.

Simmonds, P., Adams, M.J., Benkő, M., Breitbart, M., Brister, J.R., Carstens, E.B., et al. (2017) Consensus statement: Virus taxonomy in the age of metagenomics. *Nat Rev Microbiol* **15**: 161–168.

Stough, J.M.A., Kolton, M., Kostka, J.E., Weston, D.J., Pelletier, D.A., and Wilhelm, S.W. (2018) Diversity of Active Viral Infections within the Sphagnum Microbiome. *Appl Environ Microbiol* **84**:.

Yoshida, T., Nishimura, Y., Watai, H., Haruki, N., Morimoto, D., Kaneko, H., et al. (2018) 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* **12**: 1287–1295.

Figure and Figure legends

Figure 1. Diel cycling of transcripts of virus vSAG 37-F6 and relatives and putative SAR11 host. (**A**) Diel timing and abundance of 37-F6 viral transcripts and SAR11 in North Pacific Subtropical Gyre, a habitat representative of oligotrophic oceans that cover ≈40% of the Earth´s surface. Data show average transcripts per litre obtained from samples taken at the same time in the day from a previous metatranscriptomic survey (Aylward *et al.*, 2017). A total of 44 different sampling points are included in the analysis. Reads were mapped with LAST program (see Methods for details) against 37-F6 genome at different coverage and identity cut-offs. (**B**) Diel timing and abundance of 37-F6 viral transcripts and viral relatives (assembled contigs), at genus level, in Osaka Bay. A total of 16 viral contigs related to virus 37-F6 at the genus level were obtained, but for convenience, 7 are displayed in the figure, although the rest followed the same trend. Transcriptomic data from Osaka Bay was

previously reported in (Yoshida *et al.*, 2018). Diel cycling of SAR11 transcripts are also shown. Transcriptional units are normalized to the standard units "fragments per kilobase per mapped million reads" (FPKM)

Figure 2. Diel transcriptional activities of SAR11 and virus vSAG 37-F6, and abundance of free viral particles in Osaka Bay. (**A)** Plot depicts diel changes of transcriptional activities of the putative SAR11 host and virus 37-F6 compared to the diel cycling abundances of free viral particles obtained in viromes expressed in FPKM units. Abundance of 37-F6 was measured by virome fragment recruitment; see method for details. These viromes were obtained from free viruses (<0.2 μm fraction) in seawater (Yoshida *et al.*, 2018). (**B**) Ratio of abundance expressed in FPKM of transcriptome vs virome in Osaka Bay.