Lysogenic host–virus interactions in SAR11 marine bacteria

Robert M. Morris, Kelsy R. Cain, Kelli L. Hvorecny and Justin M. Kollman

Host–virus interactions structure microbial communities, drive biogeochemical cycles and enhance genetic diversity in nature. Hypotheses proposed to explain the range of interactions that mediate these processes often invoke lysogeny, a latent infection strategy used by temperate bacterial viruses to replicate in host cells until an induction event triggers the production and lytic release of free viruses. Most cultured bacteria harbour temperate viruses in their genomes (prophage). The absence of prophages in cultures of the dominant lineages of marine bacteria has contributed to an ongoing debate over the ecological significance of lysogeny and other viral life strategies in nature. Here, we report the discovery of prophages in cultured SAR11, the ocean’s most abundant clade of heterotrophic bacteria. We show the concurrent production of cells and viruses, with enhanced virus production under carbon-limiting growth conditions. Evidence that related prophages are broadly distributed in the oceans suggests that similar interactions have contributed to the evolutionary success of SAR11 in nutrient-limited systems.

The SAR11 clade is composed of a diverse group of heterotrophic bacteria that dominate life in the oceans. They account for >10^10 cells globally and are outnumbered only by viruses that infect them or other plankton. The range of host–virus interactions that have contributed to the success of SAR11 and their viruses is largely unknown, mostly because cultured marine viruses are primarily identified by their ability to kill a host. Although about half of all marine bacteria have a prophage, these latent bacterial viruses have not been identified in the ocean’s most abundant cultured bacteria. We previously isolated SAR11 from the North Pacific Ocean to identify trophic interactions. Here we report the discovery of prophages in marine SAR11 (strains NP1 and NP2) that produce free viruses throughout the bacterial growth cycle.

Comparative analyses of the complete 1.37-megabase (Mb) Pelagibacter sp. strain NP1 genome indicate that it is most closely related to the first cultured representative from the SAR11 clade, Candidatus Pelagibacter ubique HTCC1062 (Fig. 1). The NP1 genome is 56,946 bp longer than the HTCC1062 genome and codes for 37 more proteins. The average nucleotide identity and in silico DNA–DNA hybridization values for NP1 and HTCC1062 are 90% and 41%, which are below the 94% and 70% threshold values typically used for species classifications, respectively. We therefore propose the name Candidatus Pelagibacter ubique strain NP1 in recognition of Stephen Giovannoni’s discovery of the SAR11 clade and his important contributions.

The most notable difference between strain NP1 and previously cultured SAR11 is that it contains a prophage (herein Pelagibacter phage PNP1). The prophage was first identified in an an ambiguous 35-kilobase (kb) section of the genome with >3-fold higher sequence coverage (Extended Data Fig. 1). The full 35,831-bp phage genome codes for 54 proteins, including those involved in transcriptional regulation, termination and virus assembly. The genome is flanked by 40-bp direct repeats, indicating the core sequence for site-specific recombination and formation of hybrid attL and attR sites from virus attP and host attB attachment sites (Fig. 1 and Extended Data Figs. 2 and 3). The 40-bp core sequence is located upstream of a phage tyrosine integrase and in the host is part of a transfer RNA gene (tRNA^{tyr}(CAT)). Tyrosine integrases typically recognize core sequences 6–8 bp in length and that are flanked by inverted repeats. Evidence that the same core sequence exactly matches a tRNA^{tyr}(CAT) gene in the HTCC1062 genome and sequences in published metagenomic datasets (Extended Data Fig. 4) suggests that PNP1 can integrate into the genomes of diverse SAR11.

Temperate prophages produce virions by spontaneous prophage induction (SPI), which is a low-frequency event that occurs under ideal growth conditions, or non-spontaneously when DNA damage or stress induces a prophage at much higher rates. We identified high SPI under carbon-replete growth conditions, with a virus-to-host ratio that reached 0.84:1 (Fig. 2a). A second lysogenic strain of SAR11 (NP2) had a similar growth rate and produced a similar virus-to-host ratio (0.85:1). Attempts to induce PNP1 virus production with known inducing agents were unsuccessful and suggest that induction studies underestimate the number of prophages in the oceans (Fig. 2b). Notably, host abundance did not decline under carbon-replete conditions, even as virus abundance increased. This is in contrast to many other host–virus interactions, in which high concentrations of viruses are produced by a subset of infected cells.

The virus-to-host ratio increased substantially under carbon-deplete growth conditions, reaching 15:1 in the stationary growth phase (Fig. 2c). Virus-to-host ratios are typically 10:1 in the oceans, although deviations are common and have prompted considerable debate over the range of interactions that structure microbial communities in high- and low-nutrient marine ecosystems. We grew NP1 in a higher background of viruses to test for reinfection and confirm that carbon limitation was the cause of non-spontaneous prophage induction (Extended Data Fig. 5). There was no notable difference between the growth rate or final cell density of cultures started with a virus-to-host ratio of either 10:1 or 1:1. This suggests that lysogenic host–virus interactions enhance marine virus production under nutrient-limiting conditions. We measured virus decay to more accurately estimate virus production (Fig. 2d).

Our estimate suggests that up to 2.3% of infected NP1 cells lysed and released virions by SPI under carbon-replete growth conditions and that up to 30.6% of infected NP1 cells lysed and released virions by non-spontaneous prophage induction under carbon-deplete
growth conditions. Our estimates for the number of cells lysed by SPI are higher than for other lysogenic bacteria\(^2\). Evidence of a concurrent vertical and horizontal transmission strategy that resembles chronic infections\(^3\) supports theoretical predictions of hybrid lytic-lysogenic life strategies in nature\(^4\).

Transmission electron microscopy (TEM) images show that lysogenic SAR11 cultures produce short-tailed podoviruses that are similar in size and morphology to other *Pelagibacter* viruses\(^6\) (Fig. 3 and Extended Data Fig. 6). Strain NP1 produces viruses with 70.7 nm (±2.3 nm) capsids and strain NP2 produces viruses with 102.7 nm (±7.7 nm) capsids, suggesting that these are different prophages. We also identified elongated cell morphologies, budding cells and particles of 20–300 nm in diameter (\(n = 162\)) resembling membrane vesicles in other bacteria (Extended Data Fig. 7)\(^3\). Budding and viral lysis are known mechanisms of vesicle formation in bacteria\(^1\). The growth rate of strain NP1 is much slower than the growth rate of HTCC1062 on the same medium (0.21 ± 0.02 and 0.41 ± 0.01 \(d^{-1}\), respectively)\(^2\). This could be due to the metabolic cost of producing and releasing viruses and vesicles, for defense against a viral superinfection\(^4\), or for spreading phage susceptibility to other SAR11 (ref. \(^5\)).

We identified closely related phage genomes in a metagenomic study of marine viruses from the Mediterranean Sea\(^6\) and from the North Pacific\(^7\) (Fig. 4). Gene order between PNP1 and all phage sequences recovered from the environment is highly conserved. They have 40-bp core sequences adjacent to tyrosine integrases that match tRNA sequences in SAR11 and an uncultured Marine Group II Euryarchaeota\(^8\) (Extended Data Fig. 4). Recent evidence of prophages in freshwater SAR11 (ref. \(^9\)) and other temperate phages in marine SAR11 (ref. \(^10\)) points to a mechanism capable of mediating horizontal gene transfer in the environment.

Hypotheses proposed to explain the roles of viruses in structuring microbial communities, driving biogeochemical cycles and enhancing genetic diversity in nature have predicted a wide range of interactions\(^1\). We show that temperate phages in lysogenic SAR11 produce virions throughout the bacterial growth cycle, that virion production increases under carbon-deplete conditions and that related prophages are broadly distributed in seawater.

![Fig. 1](image1.png)  
**Fig. 1** Genome alignment of Ca. *P. giovannonii* strain NP1 to *Ca. P. ubique* HTCC1062. Inner ring, the complete genome sequence of strain NP1 mapped to itself (green). Outer ring, the complete genome of HTCC1062 (blue). The red border marks the location of the prophage. The black lines mark the location of the 40-bp core nucleotide sequences (attL and attR) flanking the prophage genome.

The discovery that lysogenic SAR11 continuously produce viruses, with enhanced production under carbon-limiting conditions, raises the possibility that prophage-mediated lateral gene transfer (transduction) has contributed to the high rates of recombination underlying the evolutionary success of SAR11 in nutrient-limited systems.

**Methods**

**Genome sequencing.** Cultures for genome sequencing were grown from a single starter culture in eight 1-l polycarbonate bottles containing filtered seawater medium amended with 10 \(\mu\)M dimethylsulfoipropionate (DMSP), 25 \(\mu\)M taurine and 30 \(\mu\)M 2,3-dihydroxypropane-1-sulfonate (DHPS), at the in situ temperature of 13 °C, and in the dark under aerobic growth conditions. Cells were collected on 0.2-\(\mu\)m Supor 200 polyethersulfone filters in the late exponential growth phase (\(1 \times 10^8\) cells ml\(^{-1}\)). DNA was extracted, combined and purified using a single DNeasy PowerClean Pro cleanup column (MO Bio Laboratories). A total of 0.36 \(\mu\)g
of purified DNA was recovered and submitted to the University of Washington’s Genome Science Department’s Pacific Biosciences Services laboratory. A PacBio single-molecule real-time (SMRT) 10-kb library was constructed and sequenced using a single SMRT cell on the PacBio RS II platform. A total of 360,986 subreads were acquired, with an N50 value of 2,993 bp, where 50% of the bases are in reads longer than, or equal to, this value. Subreads shorter than 50 bp or with a quality score below 75 were removed before assembly. Initial de novo assembly of the genome from reads was conducted using the Hierarchical Genome Assembly Process 4 (HGAP4)43. Through the HGAP4 pipeline, reads were preassembled using BLASR, assembled using Celera and polished using the Quiver consensus algorithm with default settings. The HGAP4 pipeline produced 6 linear contigs with a mean coverage of 697. Another assembly was run using the Canu assembler version 1.2 (ref. 44) to produce two linear contigs with overlapping ends and a mean coverage of 696X. Contigs produced by Canu were used as a reference for HGAP4 contigs. Gaps and ambiguous regions were sequenced by PCR using custom oligonucleotide primers designed using Geneious v11.1.5 (Biomatters) and manufactured by Integrated DNA Technologies (Extended Data Fig. 2). The final assembly produced a single linear contig with overlapping ends. Genome ends were aligned and trimmed in Geneious v11.1.5 to produce a single 1.37-Mb circular chromosome. The complete, finished genome of strain NP1 was annotated using NCBI’s Prokaryotic Annotation Pipeline. The exact location and orientation of the PNP1 prophage integration sites was determined by PCR analyses (Extended Data Fig. 2) of cultures in the late stationary growth phase, which contained DNA sequences from the PNP1 prophage, free PNP1 phage and strain NP1 bacteria without the PNP1 prophage. Genome sequences were aligned using the BLAST ring image generator (BRIG) with sequences shared between strain NP1 and HTVC010P, which has a similarly sized genome, a slightly larger capsid and a burst size of 42. Growth experiments were carried out in 50-ml volumes in acid-washed and sterilized 250-ml polycarbonate bottles in biological triplicates. Culture identity and purity were verified before and after growth experiments by 16S ribosomal RNA sequence analysis. Prophage induction studies were conducted with 1 µM mitomycin C and 450 mW cm⁻² ultraviolet radiation for 30 s (Fig. 3).

Calculations. The percentage of bacteria that lysed is calculated using $V_t/V_i = B/B_i$, where $V_t$ is the total number of viruses and $B_t$ is the total number of bacteria (Supplementary Data 1). We used 42 as an estimate of burst size because there are only minor differences in burst sizes for known SAR11 podoviruses (37, 42 and 49 viruses per cell; ref. 8). PNP1 is most closely related to HTVC010P, which has a similarly sized genome, a slightly larger capsid and a burst size of 42. We quantified virus decay using the equation $(1/t)\ln(B/B_i)$ (Fig. 2d), where $t$ is time, and $B_t$ and $B_i$ are the initial and final concentration of viruses, respectively.

Microscopy. For epifluorescence microscopy, cells and viruses were filtered onto 25-mm Anodisc filters with a pore size of 0.02 µm (Whatman), stained with SYBR Green I solution and then fixed with Citifluor AF-1 antifade mounting solution (Ted Pela) as previously described47. Counts were determined for 10–20 images at each time point and for biological triplicates. Objects were differentiated by size exclusion (Extended Data Fig. 9). Virus abundance used to calculate the decay constant was measured in cultures that were filter-sterilized through a 0.2-µm filter to remove cells. Bacterial growth was determined by quantifying cell abundance every 2 to 3 days using a flow cytometer48. Virus abundances and virus-to-host ratios were determined by counting cells and viruses weekly using a fluorescent microscope. Negative-stain TEM was used to image strain NP1 and NP2 cells and PNP1 and PNP2 phage particles. Cells and viruses were concentrated by ultracentrifugation and resuspension (64,000 RCF at 4°C for 2h) or by centrifugal filtration (Amicon Ultra15). Cultures were applied to glow-discharged continuous carbon film grids in two ways: 6 µl of culture was applied twice; or cultures were ultracentrifuged directly onto grids (62,500 RCF at 4°C for 1h). Grids were then washed 3 times in double-distilled H₂O and negatively stained using 2% uranyl formate. Samples were imaged using an FEI Morgagni operating at 100kV and a Gatan Orius CCD (charge-coupled device) camera with the software package.
Digital Micrograph. Images were prepared in Fiji and assembled in Adobe Illustrator CC. A previous study found that only a small number of membrane vesicles were visible in seawater using the SYBR staining methods (average 0.01–0.1%)\(^3\). It is unlikely that the vesicle-like particles we identified by TEM impacted our direct virus counts.

**Metagenomic sequence identification.** We identified viruses related to the PNP1 prophage in the subset of published metagenomic surveys that used long reads by searching the 40-bp direct repeat sequences identified in strain NP1 (attB site, tRNA\(^{\text{Met}}\)(CAT)) and in PNP1 (attP site) against the nucleotide collection (nr/nt) using BLAST. Only identical matches were used for further analyses. We then expanded the search using the first 40 bp of all tRNA sequences in Ca. P. giovannonii strain NP1. Identical matches were then aligned to the PNP1 prophage genome. For final confirmation, only complete phage genome sequences with a tyrosine integrase located directly adjacent to a core 40-bp attP sequence that exactly matched a tRNA sequence in SAR11 were retained for further characterization (Fig. 4 and Extended Data Fig. 4). It is likely that similar SAR11 prophages have gone undetected in short-read metagenomic surveys owing to the inherent genetic complexity of SAR11 and marine phages\(^7\). PNP1-like prophages were identified in the subset of metagenomic surveys that sequenced complete virus genomes using fosmids or long-read sequencing.

**Reporting Summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

**Data availability.** Ca. P. giovannonii strain NP1 and associated data have been deposited under GenBank accession number CP038852 and BioProject number PRJNA531001. The strain NP1 16S RNA sequence has been deposited under GenBank accession number MH192304. Original negative-stain micrographs will be made available on request. Ca. P. giovannonii strain NP1 cultures have unique growth requirements and cannot be maintained in standard culture collections. Requests for bacterial cultures or DNA from the University of Washington Culture Collection should be addressed to the corresponding author. On receiving a request for materials, we will assist the requesting institution in completing a Uniform Biological Material Transfer Agreement. There is no transmittal fee for academic institutions.

Received: 3 February 2020; Accepted: 16 April 2020; Published online: 18 May 2020

**References**

1. Rohwer, F. & Thurber, R. V. Viruses manipulate the marine environment. *Nature* **459**, 207–212 (2009).
2. Breitbart, M., Bonnain, C., Malki, K. & Sawaya, N. A. Phage puppet masters of the marine microbial realm. *Nat. Microbiol.* **3**, 734–766 (2018).
3. Stewart, F. M. & Levin, B. R. The population biology of bacterial viruses: why they are everywhere. *Theor. Popul. Biol.* **73**, 91–117 (2008).
4. Jiang, S. & Paul, J. H. Occurrence of lysogenic bacteria in marine virus communities as determined by prophage induction. *Mar. Ecol. Prog. Ser.* **142**, 27–38 (1996).
5. Weitz, J. S., Beckett, S. J., Brum, J. R., Cael, B. & Dushoff, J. Lysis, lysogeny and virus–microbe interactions. *Nature* **494**, E1–E3 (2013).
6. Morris, R. M. et al. SAR11 clade dominates ocean surface bacterioplankton communities. *Nat. Commun.* **8**, 806–810 (2002).
7. Jiang, S. C. J. SAR11 bacteria: the most abundant plankton in the oceans. *Annu. Rev. Mar. Sci.* **9**, 231–255 (2017).
8. Jiang, S. C. & Paul, J. H. Seasonal and diet abundance of viruses and occurrence of lysogenesis/bacteriocinogenesis in the marine environment. *Mar. Ecol. Prog. Ser.* **104**, 163–172 (1994).
9. Jiang, S. C. J., Pogoreutz, C., Eberl, L., Moran, M. A. & Moran, M. A. for their help with experimental design leading to the cultivation of lysogenic strains of SAR11.
Author contributions
R.M.M. led the research effort, advised K.R.C. on all research activities, conducted direct cell and virus counts, and wrote the paper with support from K.R.C., K.L.H. and J.M.K. K.R.C. isolated SAR11 strains NP1 and NP2, sequenced the genome of strain NP1, verified the PNP1 phage integration site, maintained cultures and conducted growth experiments as part of her undergraduate research thesis. J.M.K. advised K.L.H. and contributed to the interpretation of TEM and associated data. K.L.H. prepared samples for TEM analyses, took images, and quantified virus and vesicle-like particle sizes.

Competing interests
The authors declare no competing interests.

Additional information
Extended data is available for this paper at https://doi.org/10.1038/s41564-020-0725-x.
Supplementary information is available for this paper at https://doi.org/10.1038/s41564-020-0725-x.
Correspondence and requests for materials should be addressed to R.M.M.
Reprints and permissions information is available at www.nature.com/reprints.
Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
© The Author(s), under exclusive licence to Springer Nature Limited 2020
Extended Data Fig. 1 | Sequence coverage across the Ca. P. Giovannonii strain NP1 genome. A total of 5346 reads. Mean coverage is 696 with 0 missing bases.
Extended Data Fig. 2 | PCR primers used to amplify NP1 and PNP1 DNA. PCR primers were designed to connect genomic contigs, to validate genomic assemblies, and to identify phage insertion sites.
Extended Data Fig. 3 | Sequences associated with phage and host attachment sites verified by PCR. a, The bacterial chromosomal sequence (attB), phage sequences (attP), and core sequences in blue, black, and red, respectively. b, PCR reactions verifying sequences associated with phage integration and excision, with bacterial (16S rRNA gene) and phage (internal) controls. All PCR products were sequenced to verify phage integration (attL and attR sites), a circular phage genome with an attachment site (attP), and phage excision from the bacterial chromosome with an attachment site (attB). PCR reactions were verified with DNA extracted from two or more cultures.
| Source sequence (phage)                                      | Match (host) | Accession |
|-------------------------------------------------------------|--------------|-----------|
| PNP1 (CP038852)                                             | NPI          | CP038852  |
| TGCGGGGTTAGCTCAGTCGTTAGACGCAGGACTCATAA                      | HTCC1062     | CP000084  |
|                                                              | Metagenomes  | 100+      |
| Mediterranean phage (AP013398, AP013399, AP013385)          | NPI          | CP038852  |
| ATTCCTCGGTAGCTCAGTTAGACGAGCAGTTGACGTTAA                     | HTCC1062     | CP000084  |
|                                                              | HIMB1321     | LT840186  |
|                                                              | HIMB5        | CP003809  |
|                                                              | RS 39        | CP020777  |
|                                                              | RS 40        | CP020778  |
|                                                              | Archaea      | KF900592  |
|                                                              | AFVG\_250C00017 | MK853003 |
|                                                              | AFVG\_250N00031 | MK853017 |
| Mediterranean phage (AP013397)                              | NPI          | CP038852  |
| TGGGGGTGAGCTCAGTTAGACGATCGCTGTCAC                           | HTCC1062     | CP000084  |
|                                                              | HIMB1321     | LT840186  |
|                                                              | HIMB5        | CP003809  |
|                                                              | RS-39        | CP020777  |
|                                                              | FZCC0015     | CP031125  |
|                                                              | IOBCBE001    | GQ234274  |
|                                                              | Metagenomes  | 100+      |
| North Pacific phage (MK853003, MK853017)                     | NPI          | CP038852  |
| ATTCCCTCGGTAGCTCAGTTAGACGAGCAGTTGACGTTAA                    | HTCC1062     | CP000084  |
|                                                              | HIMB1321     | LT840186  |
|                                                              | HIMB5        | CP003809  |
|                                                              | RS-39        | CP020777  |
|                                                              | FZCC0015     | CP031125  |
|                                                              | IOBCBE001    | GQ234274  |
|                                                              | Metagenomes  | 100+      |

Extended Data Fig. 4 | Core PNP1 integration sequences identified in public databases. 40 bp core sequences identified in PNP1 (attP) and exact matches to tRNA sequences in SAR11 (attB) or to sequences in the NCBI genomic survey database.
Extended Data Fig. 5 | Ca. P. Giovannonii strain NP1 growth initiated from cultures with PNP1 virions added at different ratios. Ratio of 1:1 (black) and 10:1 (red). Data points are the mean of n=3 biologically independent samples and the error bars are the standard deviation.
Extended Data Fig. 6 | Characteristics of PNP1 and PNP2. Size and sequence characteristics of PNP1 and PNP2 phage relative to other Pelagibacter phages. Measurements are the mean from 2 independent biological samples of NP1 and 1 sample of NP2. Images were acquired from n=2-3 distinct regions on n=2-3 grids. Errors are the standard deviation.

| Phage    | Taxon     | Capsid (nm ±) | Burst size (mean ±) | Genome size | G+C | Ref. |
|----------|-----------|---------------|---------------------|-------------|-----|------|
| HTVC011P | Podoviridae | 55 ± 2        | 49 ± 5              | 39,921      | 32.0 | 8    |
| HTVC019P | Podoviridae | 55 ± 1        | 37 ± 5              | 42,102      | 34.0 | 8    |
| HTVC010P | Podoviridae | 50 ± 3        | 42 ± 7              | 34,892      | 29.7 | 8    |
| HTVC008M | Myoviridae | 84 ± 3        | 9 ± 2               | 147,248     | 33.5 | 8    |
| PNP1 (n=15) | Podoviridae | 70.7 ± 2.3 (n=8) | ND                      | 35,831      | 32.6 | *    |
| PNP2 (n=8) | Podoviridae | 102.7 ± 7.7 (n=15) | ND                      | ND          | ND   | *    |

* This study
Extended Data Fig. 7 | Micrographs of strains NP1 and NP2, virions, and vesicle like particles. 

**a**, Image of a *Ca. P. giovannonii* NP1 host cell with an elongated morphology, evidence of budding, and possible virion attachment. Image acquired at 8,900x. Arrow mark budding and asterisks mark virions; not all budding and viruses have been marked. 

**b–d**, Representative images of vesicle-like particles found in strain NP1, independently acquired at 22,000x; asterisks mark virions. 

**e**, Image of strain NP2 host cell showing evidence of budding, acquired at 14,000x. Arrow marks budding. 

**f–h**, Representative images of vesicle-like particles found in strain NP2, independently acquired at 22,000x. Scale Bars: A, E = 500 nm; B–D, F–H = 100 nm. Panels A–D, images from 2 separate cultures of NP1. Panels E–H, images from 1 culture of NP2. Images for cultures were acquired from 2–3 distinct regions on 2–3 grids.
### AMS1 base media.

Defined media used to grow SAR11 strain NP1 and strain NP2.

| Compound             | Concentration |
|----------------------|---------------|
| **Salts**            |               |
| NaCl                 | 481 mM        |
| MgCl₂·6H₂O           | 27 mM         |
| CaCl₂·2H₂O           | 10 mM         |
| KCl                  | 9 mM          |
| NaHCO₃               | 6 mM          |
| MgSO₄·7H₂O           | 2.8 mM        |
| **Macronutrients**   |               |
| (NH₄)₂SO₂            | 400 μM        |
| NaH₂PO₄ (pH 7.5)     | 50 μM         |
| **Trace metals**     |               |
| FeCl₃·6H₂O           | 117 nM        |
| MnCl₂·4H₂O           | 9 nM          |
| ZnSO₄·7H₂O           | 800 pM        |
| CoCl₂·6H₂O           | 500 pM        |
| Na₂MoO₄·2H₂O         | 300 pM        |
| Na₂SeO₃              | 1 nM          |
| NiCl₂·6H₂O           | 1 nM          |
| **Vitamins**         |               |
| B₁                   | 6 μM          |
| B₃                   | 800 nm        |
| B₅                   | 425 nm        |
| B₆                   | 500 nm        |
| Biotin (B₇)          | 4 nM          |
| B₉                   | 4 nm          |
| B₁₂                  | 700 pM        |
Extended Data Fig. 9 | See next page for caption.
Extended Data Fig. 9 | Host and virus size exclusion of SYBR Green 1 stained particles. 

a, Raw image of cells and free viruses stained, mounted, and viewed by epifluorescence microscopy. 
b, Cells (purple) and viruses (green) separated and enumerated by size exclusion (for example insert). All counts were determined by taking the average number of cells and viruses from 15–20 images at each time point and in biological triplicate. Direct cell and virus counts were repeated weekly on biologically independent samples and to verify continuous virus production in transfer cultures.
Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- The statistical test(s) used AND whether they are one- or two-sided
- A description of all covariates tested
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted. Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- Estimates of effect sizes (e.g. Cohen's d, Pearson's r), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection

Negative stain micrographs were manually collected using Gatan Digital Micrograph (version 2.10.1282.0).
Fluorescence microscopy images were acquired and analyzed using HIS Elements (version 3.22).
Flow cytometry data were acquired and analyzed using CytoSoft (version 5.2).

Data analysis

The Candidatus Pelagibacter giovannonii strain NP1 genome was assembled using HGAP4 and Canu.
Ca. Pelagibacter spp. strain NP1 and HTCC1062 genomes were aligned using BRIG (version 1).
The Pelagibacter phage PNP1 genome and related phage contigs were aligned using MAUVE.
Virus genes were called using the NCBI automated pipeline and checked using Genemark.
PCR primers were designed and the resulting sequences were aligned with the NP1 genome using Geneious (version 11.1.5).
Growth data were plotted in Microsoft Excel (version 14.7.2) and DeltaGraph (version 7).
Negative stain micrographs were prepared in Fiji/ImageJ (version 2.0.0-rc-69/1.52p).
Figures were assembled in Adobe Illustrator CS6 or CC (versions 22.1).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Ca. P. giovannonii strain NP1 and associated data have been deposited under GenBank Accession Number CP038852 and BioProject number PRJNA531001. The strain NP1 16S rRNA sequence has been deposited under GenBank accession number: MH923014. Original negative stain micrographs will be made available on
request. Ca. P. giovannonii strain NP1 cultures have unique growth requirements and cannot be maintained in standard culture collections. Requests for bacterial cultures or DNA from the UW culture collection should be addressed to the corresponding author (RMM). Upon receiving a request for materials, we will assist the requesting institution in completing a Uniform Biological Material Transfer Agreement. There is no transmittal fee for academic institutions.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences  ☐ Behavioural & social sciences  ☒ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

| Study description | Genome sequence of a marine bacterium with a temperate phage and the initial characterization of the host-virus interaction, including an assessment of bacterial growth and virus production over time. |
|-------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Research sample   | Bacterial Culture                                                                                                                                                                                  |
| Sampling strategy | To sequence the Ca. P. giovannonii strain NP1 genome, DNA was extracted and checked for purity from 8 cultures. This was the number of samples required to obtain enough material for sequencing. To test for differences in growth rates and virus and cell abundances, cultures were grown in n=3 biological triplicates. This is standard practice and was sufficient to identify statistically significant differences (SD) between experimental treatment conditions and the controls. Flow cytometry cell counts were determined for n=3 biological triplicates. Fluorescent microscopy cell and virus counts were determined for n=3 biological triplicates by counting objects in 10-20 fields of view (images). The mean for a set of images from 1 biological sample was used to calculate the mean and standard deviation for n=3 biological triplicates. The virus to host ratios determined by direct microscopy were used to compare counting methods. TEM images used for cell and virus estimates were captured from n=1 (NP2) an n=4 (NP1) biological samples and from n=2-3 images on n=2-3 grids. |
| Data collection   | The data were collected and recorded by authors RMM, KRC, and KLH.                                                                                                                                   |
| Timing and spatial scale | Cultures were sampled daily for bacterial counts and weekly for virus counts to determine virus to host ratios. |
| Data exclusions   | No data were excluded.                                                                                                                                                                               |
| Reproducibility  | In some cases, bacterial growth curves and cell and virus abundance estimates were verified in transfer cultures and in multiple experiments. Results from all transfer and repeat experiments are consistent with published results. |
| Randomization     | DNA from cultures for sequencing were extracted randomly, combined, and purified prior to sequencing. Cultures for TEM image analyses were randomly selected from 1 of n=3 biological triplicates. Cells and viruses were enumerated sequentially from n=3 biological triplicates (1,2,3) over the bacterial growth cycle. 10-20 images were captured randomly from each sample to provide a representative distribution of cells over the entire filter area. |
| Blinding          | The Ca. P. giovannonii strain NP1 genome was sequenced at the University of Washington, Genome Sciences using Pacific Biosciences Services and de novo assembled using HGAP4 and Canu pipelines. Bling was not used for sequencing, assembly, database searches, growth experiments, and imaging because each requires specific knowledge of the system and/or experimental procedure being preformed. |

Did the study involve field work?  ☒ Yes  ☐ No

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.
### Materials & experimental systems

| n/a | Involved in the study |
|-----|-----------------------|
|     | Antibodies            |
|     | Eukaryotic cell lines |
|     | Palaeontology         |
|     | Animals and other organisms |
|     | Human research participants |
|     | Clinical data         |

### Methods

| n/a | Involved in the study |
|-----|-----------------------|
|     | ChIP-seq              |
|     | Flow cytometry        |
|     | MRI-based neuroimaging |