To: Applied and Environmental Microbiology

Genomic and phylogenetic analysis of the first myovirus isolated from *Oceanospirillaceae*, representing a novel viral cluster prevalent in polar oceans

Wenjing Zhang[a†], Yundan Liu[a†], Jinyan Xing[b†], Kaiyang Zheng[a], Qian Li[c], Chengxiang Gu[a], Ziyue Wang[a], Hongbing Shao[a,d], Cui Guo[a,d], Hui He[a,d], Hualong Wang[a,d], Yeong Yik Sung[d,e], Wen Jye Mok[d,e], Li Lian Wong[d,e], Yantao Liang[a,d#], Andrew McMinn[a,f], Min Wang[a,b,d#]

[a] College of Marine Life Sciences, Frontiers Science Center for Deep Ocean Multispheres and Earth System, Institute of Evolution and Marine Biodiversity, Ocean University of China, Qingdao 266003, China.

[b] The Affiliated Hospital of Qingdao University, Qingdao 266000, China

[c] School of Oceanography, Shanghai Jiao Tong University, Shanghai, 200030, China

[d] UMT-OUC Joint Centre for Marine Studies, Qingdao 266003, China.

[e] Institute of Marine Biotechnology, Universiti Malaysia Terengganu (UMT), 21030, Kuala Nerus, Malaysia.

[f] Institute for Marine and Antarctic Studies, University of Tasmania, Hobart, Tasmania 7001, Australia.

Running title: First myovirus infecting *Oceanospirillaceae*

22 [†]These authors contributed equally to this work

23 [#] Correspondence: liangyantao@ouc.edu.cn (Y.L.); mingwang@ouc.edu.cn (M.W.)

24

Abstract: The marine bacterial family *Oceanospirillaceae*, which is abundant in the deep-seas and polar oceans, is closely associated with algal blooms and petroleum hydrocarbons degradation. However, only a few *Oceanospirillaceae*-infecting phages have so far been reported. Here we report on a novel *Oceanospirillum* phage, vB_OsaM_PD0307, which is the first myovirus to be found that infects *Oceanospirillaceae*. vB_OsaM_PD0307 with a 44,421 bp linear dsDNA genome. Phylogenetic analysis and average nucleotide sequence identities suggest that vB_OsaM_PD0307 is different from other phage isolates and represents a novel genus-level myoviral cluster with two high-quality uncultured viral genomes, designed as *Oceanospimyovirus*. Additionally, the biogeographical distribution of the vB_OsaM_PD0307 cluster suggests that they are widespread in the oceans and abundant in polar areas. In summary, our findings expand the current understanding of the phylogenetic diversity, genomic characteristic and function of *Oceanospimyovirus* phages, and highlight the role of the vB_OsaM_PD0307 phage as a major ecological agent that can infect certain key bacterial groups associated with polar algal blooms.

Importance: Oceanospirillumphage vB_OsaM_PD0307 is the first myovirus found to infect *Oceanospirillaceae* and represents a novel viral genus, *Oceanospimyovirus*. This study provides insights into the genomic, phylogenetic, and ecological characteristics of myoviruses infecting *Oceanospirillaceae* and improves our understanding of the interactions between *Oceanospirillaceae* and their phages in the oceans.

Keywords: *Oceanospirillum*phage vB_OsaM_PD0307, *Oceanospimyovirus*, Genomic

and phylogenetic analysis, Ecological distribution

Introduction

Viruses are the most abundant and diverse "life forms" in the ocean (1, 2). They mediate fluctuations in microbial abundance and shape community structure and aggregation through the lysis of their hosts, processes referred to as the "viral shunt" and "viral shuttle" (3–5). Heterotrophic bacteria-infecting phage in particular, can cause over 50% mortality rate of their hosts through cell lysis, and this number can be higher in certain environments (6). Viral-host interactions promote marine biogeochemical cycles and marine carbon sequestration through the processes known as the "biological pump" and "microbial carbon pump". Phages also promote the active evolution and shape phylogeny of their hosts through horizontal gene transfer (7, 8). Over the past few decades, our understanding of viral diversity has significantly expanded, though mostly relying on the development of high-throughput sequencing and metagenomic technology. As a consequence, over 90% of available viral genomes remain uncultured and uncharacterized (there in after referred to as UViGs) (9). Targeted phage isolation methods are vital in terms of filling the knowledge gaps between sequences information and functionality, especially in the context of interaction dynamics and co-evolution between phage and host, as well as shedding light on their potential role in marine microbial food webs. .

The bacteria genus *Oceanospirillum,* Hylemon was separated from the genus

Spirillum based on their differing physiological properties and DNA base composition. The mean GC content of *Oceanospirillum* genomes ranges from 42 to 51 mol% (10, 11). Currently, twenty-three members have been validly published (https://www.bacterio.net/genus/Oceanospirillum) and these have been isolated from a range of marine environments, including coastal waters (10, 12), decaying seaweed (13), mangrove sediments (14, 15), putrid infusions of marine mussels (11, 16, 17) and the leaves of the seagrass (18). Bacteria within this family are well known for their capability of hydrocarbon degradation (19–22) and so are often detected at high abundances in oil-contaminated marine environments (23–25). Recently, a high abundance of *Oceanospirillaceae* was detected in the hadal zone of the Mariana Trench (21), suggesting their potential importance in these extreme marine habitats. In addition, *Oceanospirillaceae* are also prevalent in polar oceans, such as the Amundsen Sea polynya (26, 27), Ross Sea (28) and coastal waters of the Arctic (29). Members of *Oceanospirillaceae* possess a complete B12 vitamin synthesis pathway, which affects DMSP synthesis and promotes algal growth (30), suggesting that *Oceanospirillaceae* plays an important role in some algal blooms. Despite the ecological significance of *Oceanospirillaceae* in the ocean, our understanding of their co-occurring phages is still poor. Currently, only eight phages infecting *Oceanospirillaceae* have been isolated, including three autographiviruses infecting *Marinomonas*, two siphoviruses, and one corticovirus, and two unclassified phages (*Marinomonas* phage MfV and *Nitrincola* phage 1M3-16). There have been no previous reports of Myovirus infecting

Oceanospirillaceae.

In this study, the first myovirus infecting *Oceanospirillaceae*, named vB_OsaM_PD0307 was isolated and characterized its genomic and phylogenetic features. Phylogenetic analysis showed that vB_OsaM_PD0307 was distantly related to other reported viral isolates in the NCBI dataset and may represent a novel myoviral genus with two high-quality UViGs. The biogeographic distribution analysis suggests that the homologous of vB_OsaM_PD0307 are abundant in polar oceans. This study provides an insight into the genome of *Oceanospirillaceae* myovirus, shedding light on the important interactions between algal bloom-associated bacteria and myoviruses in polar oceans.

Materials and methods

Isolation and purification of host phage strain

Both *Oceanospirillum* sp. PD0307 and its phage vB_OsaM_PD0307 were isolated from a surface water sample collected from coastal waters of the Yellow Sea (120°19'32.6"E, 36°4'1.7"N) in June 2020. The host strain was isolated from and maintained in 2216E medium (peptone 5 wt.%, yeast extract 1 wt.%), at 28 ºC and 120 rpm, in a shaking incubator.

For phage isolation, 10 ml seawater sample collected from the same station was passed through a 0.22 μm membrane filter to remove any cellular organisms (Isopore™ 0.2 μm GTTP; Merck, Ireland) (31). Phage vB_OsaM_PD0307 was isolated by plaque

assay using the double-layer plating method. Briefly, 200 µl seawater filtered 0.22 µm pore-size filters was mixed with the host culture (approximately 8-hour) and incubated for 30-minute, allowing the absorption of the phages at room temperature. Then, 4 ml of the semi-solid culture (at 55 ºC) was added to the mixture, pouring it onto the plate after vortexing. Plates were cultivated at 28 ºC and monitored until visible plaques were formed in the double layer culture (usually happens within 24 h).

A single plaque was picked from the double plate and suspended in 2 ml of SM buffer, (100 mM of NaCl, 8 mM of MgSO4, 50 mM of Tris-HCl, at pH 7.5) and then purified three times via plaque assay. Culture lysates were recovered to enrich and concentrate the phages. Approximately 500 ml of exponentially growing host were challenged with 50ml purified viral stock and incubated at 28 ºC for 24 h. The lysate was first filtered through a 0.22 µm membrane filter to remove uninfected host cells, then concentrated from 500 ml to 5 ml using 30 kDa super-filters (UFC5030, Millipore). Concentrated phage lysate was treated with a second filtration by passing through 0.22 µm Supor membrane.

Preparation for phage morphology observation

Phage particles were precipitated by adding PEG 8000 and NaCl to a final concentration of 10% (w/v) and 1 M, respectively, and incubated overnight at 4°C. Phage particles were precipitated by centrifugation at 15,000 g for 30 min and then re-suspended in 5 ml of SM buffer, 20 µl of the mixture was then taken and placed on a copper net and stained with 2 wt.% phosphotungstic acids (pH 7.5) for 5 min. Its morphology was

identified using transmission electron microscope (TEM) (JEOLJEM-1200EX, Japan) at 100 KV, equipped with a diamond knife for thin sectioning, and Images were taken using GATAN INC CCD image transmission system (Gatan Inc., Pleasanton, CA, USA).

Phage DNA preparation, genome sequencing and gene annotation

The concentrated phage lysate was prepared for phage genomic DNA extraction was performed by Virus DNA Kit (OMEGA), according to the manufacturer's instructions, and quality control was subsequently carried out on the purified DNA samples. The high-quality DNA sample (OD260/280=1.8~2.0, >6ug) was used to construct the fragment library and then used for Illumina NovaSeq 6000 sequencing by Shanghai Biozeron Biotechnology Co., Ltd. (Shanghai, China.). The raw paired-end reads were trimmed and quality controlled by Trimmomatic (v. 0.3.6) with parameters: SLIDINGWINDOW:4:15, MINLEN:75 (32). ABySS was used to assemble the viral genome after the quality control processes, multiple-Kmer parameters were chosen to obtain the optimal assembly results (33). GapCloser software was subsequently applied to fill in the remaining local inner gaps and to correct the single base polymorphism for the final assembly and further analysis (34).

Gene models were identified using GeneMarkS (35). Then all gene models were blastp against non-redundant (NR) in the NCBI database, SwissProt (http://uniprot.org), KEGG (http://www.genome.jp/kegg/), and COG (http://www.ncbi.nlm.nih.gov/COG) for functional annotation by the blastp module. The genome visualization was

conducted by CLC Main Workbench (v6.8 downloaded on http://www.clcbio.com). In addition, tRNA was identified using the tRNAscan-SE (v1.23) (36) and rRNA was determined using RNAmmer (v1.2 downloaded on https://services.healthtech.dtu.dk/service.php/RNAmmer-1.2). GC skew analysis was performed on Genskew (https://genskew.csb.univie.ac.at/webskew). Average amino acid identity (AAI) between vB_OsaM_PD0307 and other viral sequences was calculated by the AAI calculator (http://enve-omics.ce.gatech.edu) to estimate the distribution of AAI between proteins from two genomic sequences.

Phylogenic tree construction of host *Oceanospirillaceae* sp.

A total of 355 16S rRNA sequences of *Oceanospirillaceae* with defined taxonomy in GenBank, including the host strain Oceanospirillum sp. PD0307, were retrieved from GenBank and aligned by MAFFT (37) using G-INS-1 of strategy with 1000 iterations (mafft --globalpair --maxiterate 1000 16S_Oceanospirillaceae_pro.fasta > 16S_Oceanospirillaceae.mafft). The likelihood phylogenic tree was calculated from multiple sequence alignments using IQ-tree2 (38), applying the GTR+I+G model with 1000 bootstrap iterations (Command: iqtree -s 16S_Oceanospirillaceae.mafft -m MFP -B 1000 -T AUTO) and visualized by iTOL v4 (39).

Homologous sequence recruitment of phage vB_OsaM_PD0307

A total of 57212 isolated complete phages genomes were downloaded from NCBI GenBank to build a reference isolated-phages dataset (https://www.ncbi.nlm.nih.gov/labs/virus). The whole genome sequences of phage

vB_OsaM_PD0307 as a input to find the other similar genomes by blastn in this reference isolated-phages dataset with e-value < 1e-5, identity > 50%. At the same time, homology recruitment in uncultured virus databases was also taking place. The whole genome sequence of phage vB_OsaM_PD0307 was queried against the IMG/VR v3 (40) database using blastn to search for homologous contig sequences (threshold: e-value < 1e-5, percentage of identity > 70%) and seven UViGs were retrieved.

All the isolated and uncultured homologous sequences with vB_OsaM_PD0307, and all eight isolated *Oceanospirillaceae* phages were combined to calculate intergenomic similarity by VIRIDIC (41), as well as ANI via OAT software using the orthogonal method to determine the overall similarity (42). AAI was calculated on the website (http://enve-omics.ce.gatech.edu), which estimated the distribution of AAI between proteins from the two genomic sequences.

Phylogenetic and comparative genomic analysis

A proteomic tree based on the similarities of the whole genome was generated using VIPTree (43). Each encoding nucleic sequence as a query was searched against the Virus-Host DB using tBLASTx. All viral sequences in Virus-Host DB were selected to generate a first circular tree. The second more accurate phylogenetic tree was regenerated using 35 related phages automatically selected from the first results. 20 isolated sequences were selected as references, and seven *Mycobacterium* phages appeared as an outgroup, vB_OsaM_PD0307 and the other seven UViGs were used as queries to construct the whole-genome phylogenetic tree using VIPTree (43). The group

of seven UViGs, *Shewanella* phage SppYZU01, and vB_OsaM_PD0307 were selected to perform the multi-genomic alignments.

Global oceanic distribution of phage vB_OsaM_PD0307 and relative viral sequences

Stations in Global Ocean Viromes 2.0 (GOV 2.0) were divided into five viral ecological zones (VEZs), including the Arctic (ARC), Antarctic (ANT), temperate and tropical epipelagic (EPI), temperate and tropical mesopelagic (MES), and bathypelagic (BATHY). Five representative stations were selected for each VEZs to assess the relative abundance of vB_OsaM_PD0307 and its relative viral sequences. (ANT: ERR594377, ERR594409, ERR599352, ERR599364, ERR599384; ARC: ERR2762158, ERR2762161, ERR2762163, ERR2762165, ERR2762169; EPI: ERR594353, ERR594398, ERR594395, ERR594403, ERR594376; MES: ERR2752153, ERR2752154, ERR2752163, ERR599375, ERR599379; BATHY: msp112, msp121, msp131, msp144, msp81) The global oceanic distribution was calculated by the metagenomics tool minimap2 (parameters: -min-read-percent-identity 0.95, -min-read-aligned-percent 0.75, -m rpkm) (46) and expressed by RPKM (reads per kilobase per million mapped reads) values. Besides, pelagiphage HTVC010P, HTVC011P, cyanophages P-SSP7, P-SSM7, S-SM2, S-CBS2, roseophage SIO1, SAR116 phage HMO2011 which have significantly representative in different oceanic areas and depths as the references.

Data availability

The complete genome of *Oceanospirillum* phage vB_OsaM_PD0307 has been deposited in NCBI GeneBank under accession number OL658619, and the 16S rRNA sequence of the host also has been deposited in NCBI GeneBank under accession number OL636378.

Results and Discussion

Isolation and morphology of the plaques of vB_OsaM_PD0307

The phage vB_OsaM_PD0307 (accession: OL658619), infecting *Oceanospirillum* sp. PD0307 (accession: OL636378), was isolated from a surface seawater sample from the Yellow Sea. The morphology of vB_OsaM_PD0307 was that of a myovirus with an icosahedral head of 51 ± 2 nm in length and a 112 ± 3 nm-long contractile tail (Fig. 1A). Infection of vB_OsaM_PD0307 formed clear and round (1–2 mm diameter average) plaques in double-layer culture (Fig. 1B).

Phylogenetic analysis of the host bacterium *Oceanospirillum* sp. PD0307

The phylogenic position of the host bacterium *Oceanospirillum* sp. PD0307 was clustered with other *Oceanospirillaceae* strains in the phylogenetic tree. The monophyletic clade containing *Oceanospirillum* sp. PD0307 includes four genera (*Neptuniibacter, Profundimonas, Amphritea,* and *Oceanospirillum*), indicating its close relationship with other *Oceanospirillaceae*, especially with other *Oceanospirillum* strains (Fig. 2). Although *Oceanospirillum* sp. PD0307 is on a sister branch with *Oceanospirillum sanctuarii* AK56, which was isolated from sediment (14), the branch

with *Oceanospirillum* sp. PD0307 has a relatively distant phylogenic link with them, suggesting *Oceanospirillum* sp. PD0307 might be a novel species of *Oceanospirillum*. In addition, *Oceanospirillum* sp. PD0307 has the longest branch length in the tree (0.121), indicating that *Oceanospirillum* sp. PD0307 might have evolved from a common ancestor of *Oceanospirillum* or even *Oceanospirillaceae*. Thus, the interaction between phage vB_OsaM_PD0307 and *Oceanospirillum* sp. PD0307 might represent a novel case in the co-evolutionary history between viruses and *Oceanospirillum*.

Genomic features of phage vB_OsaM_PD0307

According to the genomic sequencing and assembly results, phage vB_OsaM_PD0307 has a 44,421-bp linear dsDNA genome with a GC content of 57.13%. No tRNA and rRNA genes were found in the genome. The genome had a 96.41% encoding rate consisting of 56 predicted open reading frames (ORFs). Thirty-seven ORFs are located on the sense strand, accounting for 66.07% of the total coding genes, and nineteen ORFs on the antisense strand (Fig. 3A, Table S1). There were 11 coding regions (19.64 % ORF angenes) that did not match any homologous sequence under the restriction of *e*-value < 1e-5 in all 56 genes. Among the remaining 45 genes that matched homologous sequences, 23 identified specific functions, and 22 matched homologous proteins containing unknown functions. The 23 functional ORFs could be classified into three different modules: twelve ORFs for DNA replication and metabolism, ten ORFs for phage structure and packing proteins, and one auxiliary metabolic gene (AMG, transcriptional regulator) (Fig. 3A).

In the genome of phage vB_OsaM_PD0307, twelve ORFs were predicted to encode genes related to DNA replication and metabolism. ORF 25 encoded the D-Ala-D-Ala carboxypeptidase family metallohydrolase gene, and also showed high homology with hedgehog signaling/DD-peptidase zinc-binding domain gene of *Vibrio* phage 1.169.O._10N.261.52.B (pident 58.6, qcovhsp 100, bitscore 154.1, evalue 4.4E-34). The structure of the N-terminal signaling domain of hedgehog proteins has been resolved and reveals a tetrahedrally coordinated zinc ion, which is structurally homologous to the zinc-binding motif in bacterial D-alanyl-D-alanine carboxypeptidases (DD-peptidases) (44–46). ORF 25 contained some amino acids of peptidase genes, commonly detected within some bacterial genomes, such as motifs HXXXXXXD and WXH, which were typical motifs for peptidase M15 subfamily A (47, 48). ORFs 44 and 45 encoded DNA-methyltransferase and DNA methylase genes, respectively. Site-specific DNA-methyltransferase, N-6 adenine-specific DNA methylase, and cytosine-N4-specific are enzymes that specifically methylate the amino group at the C-4 position of cytosines and the N-6 position of adenine in DNA. They utilize the cofactor S-adenosyl-L-methionine as the methyl donor and are active as monomeric enzymes. In prokaryotes, the major role of DNA methylation is to protect host DNA against degradation by restriction enzymes (49, 50), and DNA methylation of phage vB_OsaM_PD0307 may have a similar function to elude host immunity.

Ten ORFs encoding genes related to the structure and packaging modules are located at the front end of the genome. The putative protease gene (ORF 7) is affiliated

277 with the cl23717 superfamily, which has portal proteins upstream and capsid proteins
278 downstream. Capsid maturation in double-stranded-DNA (dsDNA) phages requires
279 proteolytic cleavage by a prohead protease (51). In the BLAST results, ClpP/crotonase-
280 like domain proteins were significantly matched to the S49 family proteins, members
281 of the large crotonase superfamily (52). One of the typical features of myotail phages
282 is the presence of tail sheath proteins. ORF 13 encoded these in Phage
283 vB_OsaM_PD0307, which had a strong homologous sequence with the myophage
284 *Shewanella* phage SppYZU01 with 99% qcov and 66.87% pident. The top 50
285 homologous sequences are all present in the myophages under the given thresholds (*e*-
286 value < 1e-50, qcov > 90%, pident > 30%). The TMhelix (ORF 17) containing gene,
287 which is sandwiched between two tail genes, is related to the transport of substances
288 across cell membranes (53) and may be related to the adsorption of host bacteria by
289 phages.

290 Only one AMG was detected in the genome of phage vB_OsaM_PD0307 and this
291 encoded a gene related to the TetR family transcriptional regulator (ORF 46). It is well
292 represented and widely distributed among bacteria with an HTH DNA-binding motif
293 (54). Members of this family are well known for their roles as regulators of antibiotic
294 efflux pumps. This gene is a phage-mediated transcriptional regulator for antimicrobial
295 resistance and can help host cells to survive in antimicrobial environments (55).

296 Cumulative GC skew analyses were performed to determine the origin and
297 terminus of replication of the phage genome (56). The minimum GC skew was at 300

298 nt and the maximum at 43900 nt, which are at the head and tail of the genome
299 respectively (Fig. 3B). Two inflection points were identified in the above regions,
300 indicating an asymmetric base composition, which were lowest at the origin and the
301 highest at the terminus.

302 **Phylogenetic and synteny analysis between phage vB_OsaM_PD0307 and its**
303 **homologous sequences**

304 To further understand the phylogenetic relationship between phage vB_OsaM_PD0307
305 and other isolated phages, 1,812 dsDNA phage genomes were selected from the Virus-
306 Host database to establish the whole-genome phylogenetic tree. Of these, phage
307 vB_OsaM_PD0307 originated from the tree root and formed a separate clade (Fig. 4A).
308 The detailed phylogenetic trees were then regenerated after adding the seven
309 homologous UViGs. Phage vB_OsaM_PD0307, *Shewanella* phage SppYZU01, and
310 the seven homologous UViGs were grouped together and formed a unique viral cluster
311 (Figs. 4B and 4C). Among them, S85_DCM_NO_526, S137 and UViG_281 were very
312 close to vB_OsaM_PD0307 and shared the same ancestral branch, indicating that they
313 had the same evolutionary characteristics and similar genetic relationships.

314 To clarity the common features of phage vB_OsaM_PD0307 and homologous
315 viral genomes, an alignment of these sequences by tBLASTx was performed.
316 Comparative genomic analysis revealed that they had a universal homology at the
317 amino acid level, especially among vB_OsaM_PD0307, S85_DNC_NO_526 and
318 S137_MES_NO_1159 (Fig. 4D). These three viral genomes showed a high degree of

319 synteny in the whole-genome arrangement, indicating correlations in the phylogenetic
320 process. Most structural and packaging gene similarities suggest that these viruses may
321 be taxonomically similar. Given that this is the first isolate of this group of viruses, it is
322 likely that this is a new and undiscovered group of viruses. Even the only AMG of
323 vB_OsaM_PD0307, the TetR family transcriptional regulator gene, can be found in a
324 similar region to the other two homologous viral genomes. This suggests that this viral
325 group may possess the ability to assist host cells survive environmental antibiotics. It is
326 worth noting that genes A, B and C of vB_OsaM_PD0307, which encode the portal
327 protein, TMhelix containing protein and hypothetical protein respectively, were
328 common but not homologous among the group; this may relate to the range of hosts of
329 these viruses.

330 **vB_OsaM_PD0307 represented a new myoviral genus**

331 Based on the results of homology search in the reference isolated-phages dataset, only
332 one genome, Shewanella phage SppYZU01, was retrieved, with a relatively low
333 average amino acid identity (AAI, 48.33%) and average nucleotide sequence identity
334 (ANI, 62.16%). Therefore, BLASTn was used to search the IMG/VR v3 (40) database
335 to find the homologous UViG sequences. In total seven UVIGs, under the thresholds
336 (min pident 70.21%, and min e-value 1.60E-08), were screened. They were all
337 assembled from marine waters or sediment samples, three of them were judged as high-
338 quality sequences (i.e. genome ⩾90% complete and non-redundant) and four contigs
339 were judged as genome fragments (i.e. genome <90% complete). The longest and

shortest sequences were 43.16 and 8.41 kb, respectively (Table S2).

All eight complete genomes of the isolated *Oceanospirillaceae*-infecting phages and the eight homologous sequences with vB_OsaM_PD0307 were combined and the intergenomic similarity of the seventeen sequences were analysed using VIRIDIC (41). Three *Marinomonas* phages assigned to the same genus, *Murciavirus* shared high intergenomic similarity (> 95%), and their aligned genome fraction and genome lengths were all 100%. In addition, the viral genomes infecting *Oceanospirillaceae* are diverse and almost all were quite different from each other (Fig. 5). Of the eight vB_OsaM_PD0307 related homologous UViGs, three sequences in the red box of the heatmap shared a high intergenomic similarity (> 50%) and had a similar aligned genome fractions and genome lengths (Fig. 5A). IMGVR_UViG_281 and vB_OsaM_PD0307 had a higher aligned genome fraction (80%), but their intergenomic similarity was low because of the differences in their genome length. The genome length of IMGVR_UViG_281 was only 8,415bp (Table S2).

To further define the similarity between vB_OsaM_PD0307 and its homologous sequence, the ANI and AAI of all eight homologous sequences were calculated and the results were found to be consistent with VIRIDIC. S85_DCM_NO_526, S137_MES_NO_1159 and IMGVR_UViG_281 were similar to vB_OsaM_PD0307 with high ANI and AAI values (> 70%) (Fig. 6, Table 1, Fig. S1, Table S3). The Bacterial and Archaeal Viruses Subcommittee (BAVS) of the International Committee on the Taxonomy of Viruses (ICTV) considers phages sharing $\geqslant$ 70% ANI as members

of the same genus (57). As vB_OsaM_PD0307 is the first isolated myovirus infecting *Oceanospirillaceae* and as it is distant from other isolated phages, based on the phylogenetic and synteny analysis, it is suggested that vB_OsaM_PD0307 represents a novel viral genus within the Myoviridae, named *Oceanospimyovirus*. S85_DCM_NO_526 and S137_MES_NO_1159 belong to this new genus together with vB_OsaM_PD0307. IMGVR_UViG_281 might also belong to the genus *Oceanospimyovirus* but having a low-quality genome sequence (Fig. 5B).

Distribution of vB_OsaM_PD0307 **in the global ocean viromes**

The biogeographical distribution of vB_OsaM_PD0307 and its closely associated viral sequences was characterized in the Global Ocean Viromes (GOV2.0) data set covering five viral ecological zones (VEZs). The reference genomes include *Shewanella* phage SppYZU01, seven homologous UViGs, and eight typical phages infecting the most abundant bacterial genera, such as *Pelagibacter*, *Prochlorococcus*, *Synechococcus, Roseobacter* and S116 cluster. The relative abundances confirmed the high abundances of pelagiphages and cyanophages, as shown in previous studies from Pacific, Indian, and Global Ocean viromes (58–60). vB_OsaM_PD0307 and its associated viral sequences have a wide and diverse distribution.

vB_OsaM_PD0307 and S137_MES_NO_1159 have similar distribution patterns, mostly in MES. Interestingly, S85_DCM_NO_526 and S82_SUR_NO_687 were relatively abundant in the ANT, while S201_DCM_NO_1678 was relatively abundant in the ARC (Fig. 6), which was significantly different from the distribution pattern of

382 vB_OsaM_PD0307. As viral abundance was mostly tightly coupled to that of their host
383 cells, it is proposed that annual marine polar algal blooms might be responsible for the
384 high abundance of *Oceanospirillaceae* (61) and thus the high abundance of
385 *Oceanospimyovirus* in polar oceans (Fig. 7). In coastal waters of the western Antarctic
386 Peninsula, Oceanospirillales and *Pelagibacteraceae* were most abundant in winter and
387 spring. In the Amundsen Sea polynya, *Oceanospirillaceae* is one of the principal
388 dominating bacteria families and in bloom events can contribute up to a 33.9% of the
389 total bacterial 16S rRNA genes (26, 62–64). In the Meade River area of the coastal
390 Arctic, OTUs related to the family *Oceanospirillaceae* comprised the largest
391 component of *Gammaproteobacteria*, approximately 22 and 8% of the bacterial
392 communities in April and August, respectively (29). Although the reasons for the high
393 abundance of host cells in the Arctic and Antarctic are not completely understood, it is
394 hypothesized that: annual algal blooms in polar regions could release large amounts of
395 organic matter that promotes increased bacterial biodiversity (65–67). In addition, algae
396 acquire vitamin B12 through a symbiotic relationship with bacteria (68). Members of
397 *Oceanospirillaceae* have been shown to support phytoplankton growth in polar waters
398 through the synthesis of cobalamin (vitamin B12) (69–72).

399 Therefore, uncultured viral sequences assembled from the polar viromes might
400 play an important but unrecognized role in regulating the polar bacterial community
401 structure associated with polar algal blooms. This study reinforces the importance and
402 power of the combination of phage isolation and metagenomics to improve our

403 knowledge of marine viral diversity and their ecological significance.

404

405 ## Conclusion

406 *Oceanospirillum* has a strong metabolic capacity and a key ecological niche; as such its
407 phage will inevitably affect its abundance, community structure, and metabolic capacity.
408 Here we isolated the first myovirus infecting *Oceanospirillaceae*, named
409 vB_OsaM_PD0307. The presence of TetR family transcriptional regulator encoded by
410 vB_OsaM_PD0307 suggests its potential contribution to the antimicrobial resistance
411 and virulence of its host. vB_OsaM_PD0307 represents a novel myoviral genus-level
412 cluster, named *Oceanospimyovirus*, with two high-quality UViGs. The relative
413 abundance and distribution of *Oceanospimyovirus* suggest that it could be prevalent in
414 polar oceans, coupling with their high-abundant and algal-associated host bacterial
415 communities. The discovery of *Oceanospimyovirus* in the Global Ocean Viromes raised
416 several questions regarding their diversity, ecology, and roles in microbial communities.
417 Here, we performed a culture-based and metagenomics-based analysis of the genomic
418 diversity and distribution of the *Oceanospimyovirus* group. The obtained
419 *Oceanospimyovirus* type genome vB_OsaM_PD0307 helps reveal the genuine extent
420 of the genetic diversity of *Oceanospimyovirus* within natural populations of marine
421 viruses. These novel insights into the diversity and ecology of *Oceanospimyovirus*
422 further expands our current understanding of these important phages. Lastly, further
423 investigation using our newly constructed virus–host models will provide additional

424 valuable insights into the influence of viruses on the interaction among algal blooms,
425 bacteria and viruses in the polar oceans.
426
427 **Acknowledgments**
428 We sincerely thank Jia Zhen, School of Computer Science and Technology, Guizhou
429 University, for his help in the data processing. We thank for the support of the high-
430 performance servers of Center for High Performance Computing and System
431 Simulation, Pilot National Laboratory for Marine Science and Technology (Qingdao),
432 the Marine Big Data Center of Institute for Advanced Ocean Study of Ocean University
433 of China, the IEMB-1, a high-performance computing cluster operated by the Institute
434 of Evolution and Marine Biodiversity, and the high-performance servers of Frontiers
435 Science Center for Deep Ocean Multispheres and Earth System.
436
437 **Funding information**
438 This study was supported by these fundings: National Natural Science Foundation of
439 China (No. 41976117, 42120104006, and 42176111), and the Fundamental Research
440 Funds for the Central Universities (202072002, 201812002, Andrew McMinn).
441
442 **Conflict of interest**
443 The authors declare that they have no conflict of interest regarding this study.
444

Ethical Approval

This article does not contain any studies with animals or human participants performed by any of the authors.

REFERENCES

1. Suttle CA. 2005. Viruses in the sea. Nature 437:356–361.
2. Wommack KE, Colwell RR. 2000. Virioplankton: Viruses in Aquatic Ecosystems. Microbiol Mol Biol Rev 64:69–114.
3. Poulton. AJ. 2021. Shunt or shuttle. Nat Geosci 14:180–181.
4. Suttle CA. 2007. Marine viruses - Major players in the global ecosystem. Nat Rev Microbiol 5:801–812.
5. Brum JR, Sullivan MB. 2015. Rising to the challenge: Accelerated pace of discovery transforms marine virology. Nat Rev Microbiol 13:147–159.
6. Fuhrman JA. 1999. Marine viruses and their biogeochemical and ecological effects. Nature 399:541–548.
7. Pál C, Papp B, Lercher MJ. 2005. Adaptive evolution of bacterial metabolic networks by horizontal gene transfer. Nat Genet 37:1372–1375.
8. Moreau H, Piganeau G, Desdevises Y, Cooke R, Derelle E, Grimsley N. 2010. Marine Prasinovirus Genomes Show Low Evolutionary Divergence and Acquisition of Protein Metabolism Genes by Horizontal Gene Transfer. J Virol 84:12555–12563.

9. Gregory AC, Zayed AA, Conceição-Neto N, Temperton B, Bolduc B, Alberti A, Ardyna M, Arkhipova K, Carmichael M, Cruaud C, Dimier C, Domínguez-Huerta G, Ferland J, Kandels S, Liu Y, Marec C, Pesant S, Picheral M, Pisarev S, Poulain J, Tremblay JÉ, Vik D, Acinas SG, Babin M, Bork P, Boss E, Bowler C, Cochrane G, de Vargas C, Follows M, Gorsky G, Grimsley N, Guidi L, Hingamp P, Iudicone D, Jaillon O, Kandels-Lewis S, Karp-Boss L, Karsenti E, Not F, Ogata H, Poulton N, Raes J, Sardet C, Speich S, Stemmann L, Sullivan MB, Sunagawa S, Wincker P, Culley AI, Dutilh BE, Roux S. 2019. Marine DNA Viral Macro- and Microdiversity from Pole to Pole. Cell 177:1109–1123.

10. Hylemon PB. 1971. A taxonomic study of the genus Spirillum ehrenberg, with special reference to nutrition and carbohydrate catabolism.

11. Pot B, Gillis M, Hoste B, Van De Velde A, Bekaert F, Kersters K, De Ley J. 1989. Intra- and intergeneric relationships of the genus Oceanospirillum. Int J Syst Bacteriol 39:23–24.

12. Trachtenberg AM, Carney JG, Linnane JD, Rheaume BA, Pitts NL, Mykles DL, MacLea KS. 2017. Draft genome sequence of the salt water bacterium Oceanospirillum linum ATCC 11336T. Genome Announc 5:49–50.

13. Krieg NR. 1981. The Genera Spirillum, Aquaspirillum, and Oceanospirillum. The Prokaryotes 595–608.

14. Sidhu C, Thakur S, Sharma G, Tanuku NRS, Pinnaka AK. 2017.

487 Oceanospirillum sanctuarii sp. Nov., Isolated from a sediment sample. Int J
488 Syst Evol Microbiol 67:3428–3434.
489 15. Krishna KK, Bhumika V, Thomas M, Anil Kumar P, Srinivas TNR. 2013.
490 Oceanospirillum nioense sp. nov., a marine bacterium isolated from sediment
491 sample of Palk bay, India. Antonie van Leeuwenhoek, Int J Gen Mol Microbiol
492 103:1015–1021.
493 16. Terasaki Y. 1979. Transfer of five species and two subspecies of Spirillum to
494 other genera (Aquaspirillum and Oceanospirillum), with emended descriptions
495 of the species and subspecies. Int J Syst Bacteriol 29:130–44.
496 17. Rosenberg E, DeLong EF, Lory S, Stackebrandt E, Thompson F. 2013. The
497 prokaryotes: GammaproteobacteriaThe Prokaryotes: Gammaproteobacteria.
498 18. Weidner S, Arnold W, Stackebrandt E, Pühler A. 2000. Phylogenetic analysis
499 of bacterial communities associated with leaves of the seagrass Halophila
500 stipulacea by a culture-independent small-subunit rRNA gene approach.
501 Microb Ecol 39:22–31.
502 19. Kleindienst S, Paul JH, Joye SB. 2015. Using dispersants after oil spills:
503 Impacts on the composition and activity of microbial communities. Nat Rev
504 Microbiol 13:388–396.
505 20. Valentine DL, Mezić I, Maćešić S, Črnjarić-Žic N, Ivić S, Hogan PJ,
506 Fonoberov VA, Loire S. 2012. Dynamic autoinoculation and the microbial
507 ecology of a deep water hydrocarbon irruption. Proc Natl Acad Sci U S A

508 109:20286–20291.

509 21. Liu J, Zheng Y, Lin H, Wang X, Li M, Liu Y, Yu MM, Zhao M, Pedentchouk
510 N, Lea-Smith DJ, Todd JD, Magill CR, Zhang WJ, Zhou S, Song D, Zhong H,
511 Xin Y, Yu MM, Tian J, Zhang XH. 2019. Proliferation of hydrocarbon-
512 degrading microbes at the bottom of the Mariana Trench. Microbiome 7:1–13.

513 22. Mason OU, Hazen TC, Borglin S, Chain PSG, Dubinsky EA, Fortney JL, Han
514 J, Holman HYN, Hultman J, Lamendella R, MacKelprang R, Malfatti S, Tom
515 LM, Tringe SG, Woyke T, Zhou J, Rubin EM, Jansson JK. 2012. Metagenome,
516 metatranscriptome and single-cell sequencing reveal microbial response to
517 Deepwater Horizon oil spill. ISME J 6:1715–1727.

518 23. Sass AM, Sass H, Coolen MJL, Cypionka H, Overmann J. 2001. Microbial
519 Communities in the Chemocline of a Hypersaline Deep-Sea Basin (Urania
520 Basin, Mediterranean Sea). Appl Environ Microbiol 67:5392–402.

521 24. Voordouw G, Armstrong SM, Reimer MF, Fouts B, Telang AJ, Shen Y,
522 Gevertz D. 1996. Characterization of 16s rRNA genes from oil field microbial
523 communities indicates the presence of a variety of sulfate-reducing,
524 fermentative, and sulfide-oxidizing bacteria. Appl Environ Microbiol 62:1623–
525 1629.

526 25. Dubinsky EA, Conrad ME, Chakraborty R, Bill M, Borglin SE, Hollibaugh JT,
527 Mason OU, M. Piceno Y, Reid FC, Stringfellow WT, Tom LM, Hazen TC,
528 Andersen GL. 2013. Succession of hydrocarbon-degrading bacteria in the

aftermath of the deepwater horizon oil spill in the gulf of Mexico. Environ Sci Technol 47:10860–10867.

26. Delmont TO, Hammar KM, Ducklow HW, Yager PL, Post AF. 2014. Phaeocystis antarctica blooms strongly influence bacterial community structures in the Amundsen Sea polynya. Front Microbiol 5:1–13.

27. Kim S-J, Kim J-G, Lee S-H, Park S-J, Gwak J-H, Jung M-Y, Chung W-H, Yang E-J, Park J, Jung J, Hahn Y, Cho J-C, Madsen EL, Rodriguez-Valera F, Hyun J-H, Rhee S-K. 2019. Genomic and metatranscriptomic analyses of carbon remineralization in an Antarctic polynya. Microbiome 7:1–15.

28. Cordone A, Errico GD, Magliulo M, Bolinesi F, Basili M, Marco R De, Saggiomo M, Rivaro P. 2021. Bacterioplankton Diversity and Distribution in Relation to Phytoplankton Community Structure in the Ross Sea surface waters. bioRxiv Microbiol.

29. Sipler RE, Kellogg CTE, Connelly TL, Roberts QN, Yager PL, Bronk DA. 2017. Microbial community response to terrestrially derived dissolved organic matter in the coastal Arctic. Front Microbiol 8:1018.

30. Delmont TO, Murat Eren A, Vineis JH, Post AF. 2015. Genome reconstructions indicate the partitioning of ecological functions inside a phytoplankton bloom in the Amundsen Sea, Antarctica. Front Microbiol 6.

31. Yang Q, Gao C, Jiang Y, Wang M, Zhou X, Shao H, Gong Z, McMinn A. 2019. Metagenomic characterization of the viral community of the South Scotia

Ridge. Viruses 11:1–19.

32. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: A flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120.

33. Simpson JT, Wong K, Jackman SD, Schein JE, Jones SJM, Birol I. 2009. ABySS: A parallel assembler for short read sequence data. Genome Res 19:1117–1123.

34. Xu M, Guo L, Gu S, Wang O, Zhang R, Peters BA, Fan G, Liu X, Xu X, Deng L, Zhang Y. 2020. TGS-GapCloser: A fast and accurate gap closer for large genomes with low coverage of error-prone long reads. Gigascience 9:1–11.

35. Besemer J, Lomsadze A, Borodovsky M. 2001. GeneMarkS: A self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. Nucleic Acids Res 29:2607–2618.

36. Lowe TM, Chan PP. 2016. tRNAscan-SE On-line: integrating search and context for analysis of transfer RNA genes. Nucleic Acids Res 44:W54–W57.

37. Katoh K, Rozewicki J, Yamada KD. 2018. MAFFT online service: Multiple sequence alignment, interactive sequence choice and visualization. Brief Bioinform 20:1160–1166.

38. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, Von Haeseler A, Lanfear R, Teeling E. 2020. IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. Mol Biol

571 Evol 37:1530–1534.

572 39. Letunic I, Bork P. 2019. Interactive Tree of Life (iTOL) v4: Recent updates
573 and new developments. Nucleic Acids Res 47:256–259.

574 40. Roux S, Páez-Espino D, Chen IMA, Palaniappan K, Ratner A, Chu K, Reddy
575 T, Nayfach S, Schulz F, Call L, Neches RY, Woyke T, Ivanova NN, Eloe-
576 Fadrosh EA, Kyrpides NC. 2021. IMG/VR v3: An integrated ecological and
577 evolutionary framework for interrogating genomes of uncultivated viruses.
578 Nucleic Acids Res 49:D764–D775.

579 41. Moraru C, Varsani A, Kropinski AM. 2020. VIRIDIC — A Novel Tool to
580 Calculate the Intergenomic Similarities of. Viruses 12:1268.

581 42. Lee I, Kim YO, Park SC, Chun J. 2016. OrthoANI: An improved algorithm and
582 software for calculating average nucleotide identity. Int J Syst Evol Microbiol
583 66:1100–1103.

584 43. Nishimura Y, Yoshida T, Kuronishi M, Uehara H, Ogata H, Goto S. 2017.
585 ViPTree: The viral proteomic tree server. Bioinformatics 33:2379–2380.

586 44. Hall TMT, Porter JA, Beachy PA, Leahy DJ. 1995. A potential catalytic site
587 revealed by the 1.7-Å crystal structure of the amino-terminal signalling domain
588 of Sonic hedgehog. Nat 1995 3786553 378:212–216.

589 45. Owens AE, Iannuzzelli JA, Gu Y, Fasan R. 2020. MOrPH-PhD: An Integrated
590 Phage Display Platform for the Discovery of Functional Genetically Encoded
591 Peptide Macrocycles. ACS Cent Sci 6:368–381.

592 46. Dawber RJ, Hebbes S, Herpers B, Docquier F, Van Den Heuvel M. 2005.
593 Differential range and activity of various forms of the Hedgehog protein. BMC
594 Dev Biol 5:1–14.
595 47. Khakhum N, Yordpratum U, Boonmee A, Tattawasart U, Rodrigues JLM,
596 Sermswan RW. 2016. Cloning, expression, and characterization of a
597 peptidoglycan hydrolase from the Burkholderia pseudomallei phage ST79.
598 AMB Express 6.
599 48. Roelink H. 2018. Sonic Hedgehog is a member of the Hh/DD-peptidase family
600 that spans the eukaryotic and bacterial domains of life. J Dev Biol 6:1–11.
601 49. Decewicz P, Radlinska M, Dziewit L. 2017. Characterization of Sinorhizobium
602 sp. LM21 prophages and virus-encoded DNA methyltransferases in the light of
603 comparative genomic analyses of the sinorhizobial virome. Viruses 9:1–19.
604 50. Dna FOF. 1995. Function of Dna. New York 293–318.
605 51. Liu J, Mushegian A. 2004. Displacements of prohead protease genes in the late
606 operons of double-stranded-DNA bacteriophages. J Bacteriol 186:4369–4375.
607 52. Zheng D, Xu Y, Yuan G, Wu X, Li Q. 2021. Bacterial ClpP Protease Is a
608 Potential Target for Methyl Gallate. Front Microbiol 11:1–10.
609 53. Kauffman KM, Hussain FA, Yang J, Arevalo P, Brown JM, Chang WK,
610 Vaninsberghe D, Elsherbini J, Sharma RS, Cutler MB, Kelly L, Polz MF. 2018.
611 A major lineage of non-tailed dsDNA viruses as unrecognized killers of marine
612 bacteria. Nature 554:118–122.

613 54. Ramos JL, Martínez-Bueno M, Molina-Henares AJ, Terán W, Watanabe K,
614 Zhang X, Gallegos MT, Brennan R, Tobes R. 2005. The TetR Family of
615 Transcriptional Repressors. Microbiol Mol Biol Rev 69:326–356.
616 55. Colclough AL, Scadden J, Blair JMA. 2019. TetR-family transcription factors
617 in Gram-negative bacteria: Conservation, variation and implications for efflux-
618 mediated antimicrobial resistance. BMC Genomics 20:1–12.
619 56. Uchiyama J, Rashel M, Takemura I, Wakiguchi H, Matsuzaki S. 2008. In silico
620 and in vivo evaluation of bacteriophage φEF24C, a candidate for treatment of
621 Enterococcus faecalis infections. Appl Environ Microbiol 74:4149–4163.
622 57. Bin Jang H, Bolduc B, Zablocki O, Kuhn JH, Roux S, Adriaenssens EM,
623 Brister JR, Kropinski AM, Krupovic M, Lavigne R, Turner D, Sullivan MB.
624 2019. Taxonomic assignment of uncultivated prokaryotic virus genomes is
625 enabled by gene-sharing networks. Nat Biotechnol.
626 58. Kang I, Oh HM, Kang D, Cho JC. 2013. Genome of a SAR116 bacteriophage
627 shows the prevalence of this phage type in the oceans. Proc Natl Acad Sci U S
628 A 110:12343–12348.
629 59. Zhao Y, Temperton B, Thrash JC, Schwalbach MS, Vergin KL, Landry ZC,
630 Ellisman M, Deerinck T, Sullivan MB, Giovannoni SJ. 2013. Abundant SAR11
631 viruses in the ocean. Nature 494:357–360.
632 60. Hurwitz BL, Sullivan MB. 2013. The Pacific Ocean Virome (POV): A Marine
633 Viral Metagenomic Dataset and Associated Protein Clusters for Quantitative

634 Viral Ecology. PLoS One 8.

635 61. Yan Liu, Pavla Debeljak, Mathieu Rembauville, Stéphane Blain IO. 2019.
636 Diatoms shape the biogeography of heterotrophic prokaryotes in early spring in
637 the Southern Ocean. Environ Microbiol 21:1452–1465.

638 62. Kim J-G, Park S-J, Quan Z-X, Jung M-Y, Cha I-T, Kim S-J, Kim K-H, Yang
639 E-J, Kim Y-N, Lee S-H, Rhee S-K. 2013. Unveiling abundance and
640 distribution of planktonic Bacteria and Archaea in a polynya in Amundsen Sea,
641 Antarctica https://doi.org/10.1111/1462-2920.12287.

642 63. Ghiglione J-F, Galand PE, Pommier T, Pedrós-Alió C, Maas EW, Bakker K,
643 Bertilson S, Kirchman DL, Lovejoy C, Yager PL, Murray AE, Karl DM. Pole-
644 to-pole biogeography of surface and deep marine bacterial communities
645 https://doi.org/10.1073/pnas.1208160109.

646 64. Delong EF, Baumann L, Bowditch RD, Baumann P. 1984. Microbiology 9
647 170–178.

648 65. Henson SA, Cael BB, Allen SR, Dutkiewicz S. 2021. Future phytoplankton
649 diversity in a changing climate. Nat Commun 12:1–8.

650 66. Gray A, Krolikowski M, Fretwell P, Convey P, Peck LS, Mendelova M, Smith
651 AG, Davey MP. 2020. Remote sensing reveals Antarctic green snow algae as
652 important terrestrial carbon sink. Nat Commun 11.

653 67. Lutz S, Anesio AM, Raiswell R, Edwards A, Newton RJ, Gill F, Benning LG.
654 2016. The biogeography of red snow microbiomes and their role in melting

655 arctic glaciers. Nat Commun 7:1–9.

656 68. Croft MT, Lawrence AD, Raux-Deery E, Warren MJ, Smith AG. 2005. Algae
657 acquire vitamin B12 through a symbiotic relationship with bacteria. Nature
658 438:90–93.

659 69. Bertrand EM, McCrow JP, Moustafa A, Zheng H, McQuaid JB, Delmont TO,
660 Post AF, Sipler RE, Spackeen JL, Xu K, Bronk DA, Hutchins DA, Allen AE,
661 Karl DM. 2015. Phytoplankton-bacterial interactions mediate micronutrient
662 colimitation at the coastal Antarctic sea ice edge. Proc Natl Acad Sci U S A
663 112:9938–9943.

664 70. Oliver H, St-Laurent P, Sherrell RM, Yager PL. 2019. Modeling Iron and Light
665 Controls on the Summer Phaeocystis antarctica Bloom in the Amundsen Sea
666 Polynya. Global Biogeochem Cycles https://doi.org/10.1029/2018GB006168.

667 71. Mönnich J, Tebben J, Bergemann J, Case R, Wohlrab S, Harder T. 2020.
668 Niche-based assembly of bacterial consortia on the diatom Thalassiosira rotula
669 is stable and reproducible. ISME J 14:1614–1625.

670 72. Luria CM, Amaral-Zettler LA, Ducklow HW, Rich JJ. 2016. Seasonal
671 succession of free-living bacterial communities in coastal waters of the western
672 antarctic peninsula. Front Microbiol 7:1–13.

673

674

Table 1 Amino acid identity of Oceanospirillumphage vB_OsaM_PD0307 and other night homologous sequences.

Abbreviation	Identity Frequency		Bit-score Frequency	
	mean	median	mean	median
IMGVR_UViG_3300009431_000281	85.08	93.7	428.9	346.5
S137_MES_NO_1159	83.36	87.01	461.6	343
S85_DCM_NO_526	72.9	79.04	428.2	282
Shewanella phage SppYZU01	48.33	47.22	260.6	173
S201_DCM_NO_1678	48.24	47.73	260.4	144
IMGVR_UViG_3300003691_000130	46.53	46.51	247.9	125
S82_SUR_NO_687	45.45	44.27	249.1	159
IMGVR_UViG_3300024520_000016	40.97	37.32	127.7	30

675

676 # Figure legends

677 **Fig. 1** (A) Morphology and biological properties of phage vB_OsaM_PD0307, the scale
678 bar is 50 nm. (B) Phage plaques of vB_OsaM_PD0307 formed in double-layer agar
679 plates, the scale bar is 5 mm

680 **Fig. 2** The maximum likelihood phylogenic tree of *Oceanospirillum* sp. PD0307 and
681 16S rRNA sequences of 31 related *Oceanospirillaceae* species. The *Oceanospirillum*
682 sp. PD0307 was highlighted in the tree. A relatively close relationship of
683 *Oceanospirillum* sp. PD0307 with other *Oceanospirillum* was displayed, while
684 monophyly of its branch was observed.

685

686 **Fig. 3** Circularized genome map (A) and cumulative GC skew analysis (B) of the
687 genome sequence of Oceanospirillumphage vB_OsaM_PD0307. (A) The outer circle
688 represents different categories of putative functional genes, which were represented by
689 different colors. (B) the cumulative graph of minimum and maximum values of GC
690 skew were displayed and calculated by using a window size of 1,00 bp and a step size
691 of 100 bp. The GC-skew and the cumulative GC-skew were represented by blue and
692 red lines, respectively. The minimum and maximum of a GC-skew could be used to
693 predict the origin of replication (300 nt) and the terminus location (43900 nt).

694

695 **Fig. 4** Phylogenetic trees of Oceanospirillum Phage vB_OsaM_PD0307 with different
696 references. (A) The whole-genomes phylogenetic tree was constructed with 1,812

dsDNA phages genomes in the Virus-Host database as references. (B) Nine *Oceanospimyovirus* (Oceanospirillum Phage vB_OsaM_PD0307, Shewanella phage SppYZU01, and seven homologous uncultured viral genomes) as queries and other 80 related phages were used to construct a circular phylogenetic tree. (C) A rectangular phylogenetic tree was established with *Oceanospimyovirus* and seven Mycobacterium phages, which were used as outgroups for control. (D) Gene synteny of Oceanospirillum Phage vB_OsaM_PD0307, Shewanella phage SppYZU01, and seven homologous uncultured viral genomes in IMG/VR v3 database. Sequences comparison performed using tBLASTx (10 bp minimum alignment) with percent identity. Synteny was recognized when genomes featured a minimum of five consecutive syntenic genes within the same genomic area and separated by a maximum of four non-syntenic genes.

Fig. 5 (A) Heatmap of intergenomic similarity values (right half) and alignment indicators (left half and top annotation) of eight *Oceanospirillaceae* phages and eight homologous uncultured viral genomes of Oceanospirillum Phage vB_OsaM_PD0307. The numbers of intergenomic similarity values represent the similarity values for each genome pair, rounded to the first decimal. The genome length ratio for a genome pair using a black to white color gradient indicator. The aligned fraction genome was indicated by orange to white color gradient. (B) The average nucleotide sequence identity (ANI) of Oceanospirillum Phage vB_OsaM_PD0307, Shewanella phage SppYZU01, and seven homologous uncultured viral contigs based on OrthoANI values

718 calculated using OAT software.

719

720 **Fig. 6** Relative abundance of Oceanospirillum Phage vB_OsaM_PD0307, Shewanella

721 phage SppYZU01, seven homologous uncultured viral contigs and reference phage

722 genomes in the global ocean viromes (GOV 2.0) datasets.

723

724 **SUPPLEMENTAL MATERIAL**

725 Supplemental material is available online only.

726 FIG S1, PDF file, 6321 KB.

727 TABLE S1, XLSX file, 17 KB.

728 TABLE S2, XLSX file, 14 KB.

729 TABLE S3, XLSX file, 28 KB.

730 **Supplemental flie 1: Fig S1** Amino acid identity and bitscore distribution between

731 vB_OsaM_PD0307, Shewanella phage SppYZU01, and seven homologous uncultured

732 viral genomes

733 **Supplemental flie 2: Table S1** Genome annotation of Oceanospirillumphage

734 vB_OsaM_PD0307

735 **Supplemental flie 3: Table S2** The result of BLASTn in IMG/VR and the uncultured

736 contigs‘ information

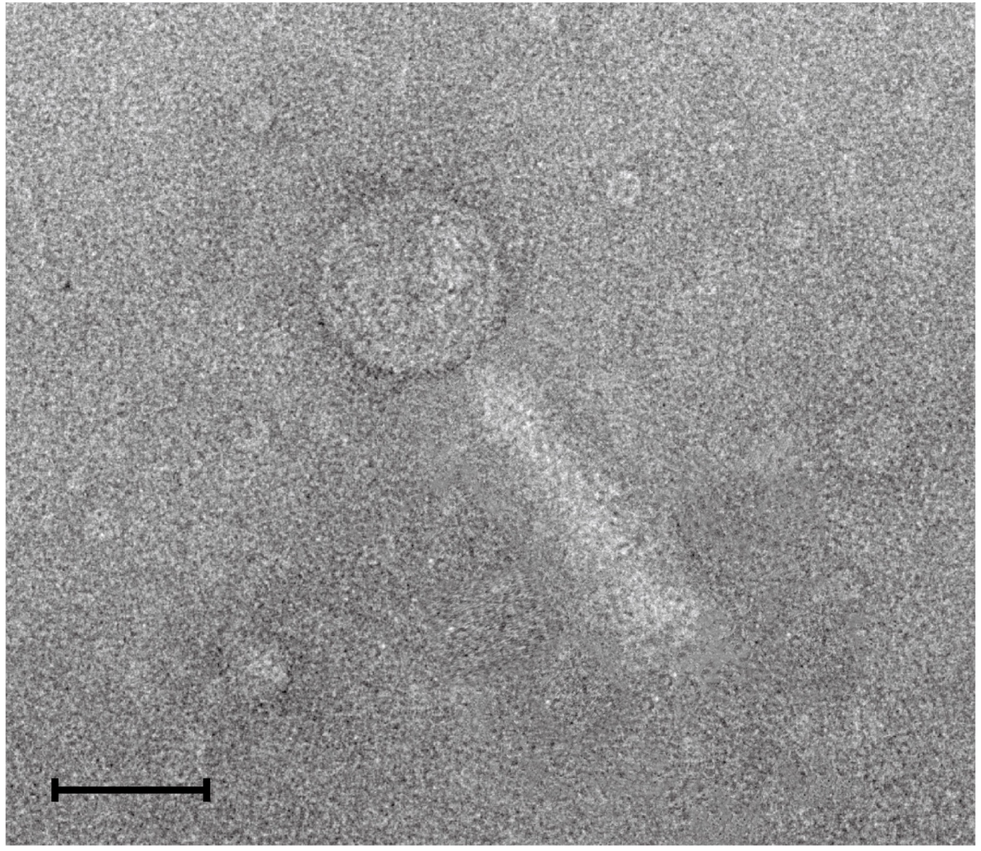
737 **Supplemental flie 4: Table S3** Average amino acid identity of Oceanospirillum Phage

738 vB_OsaM_PD0307 between Shewanella phage SppYZU01 and seven homologous

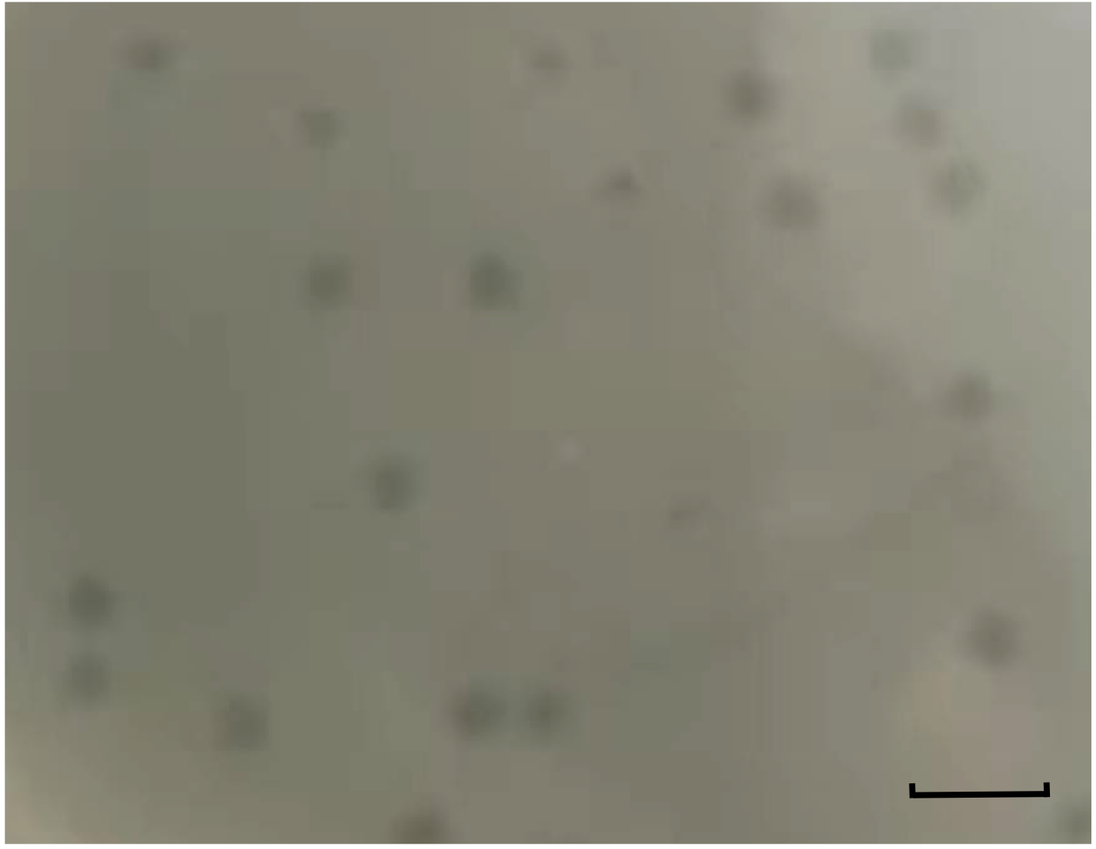
739 uncultured viral genomes

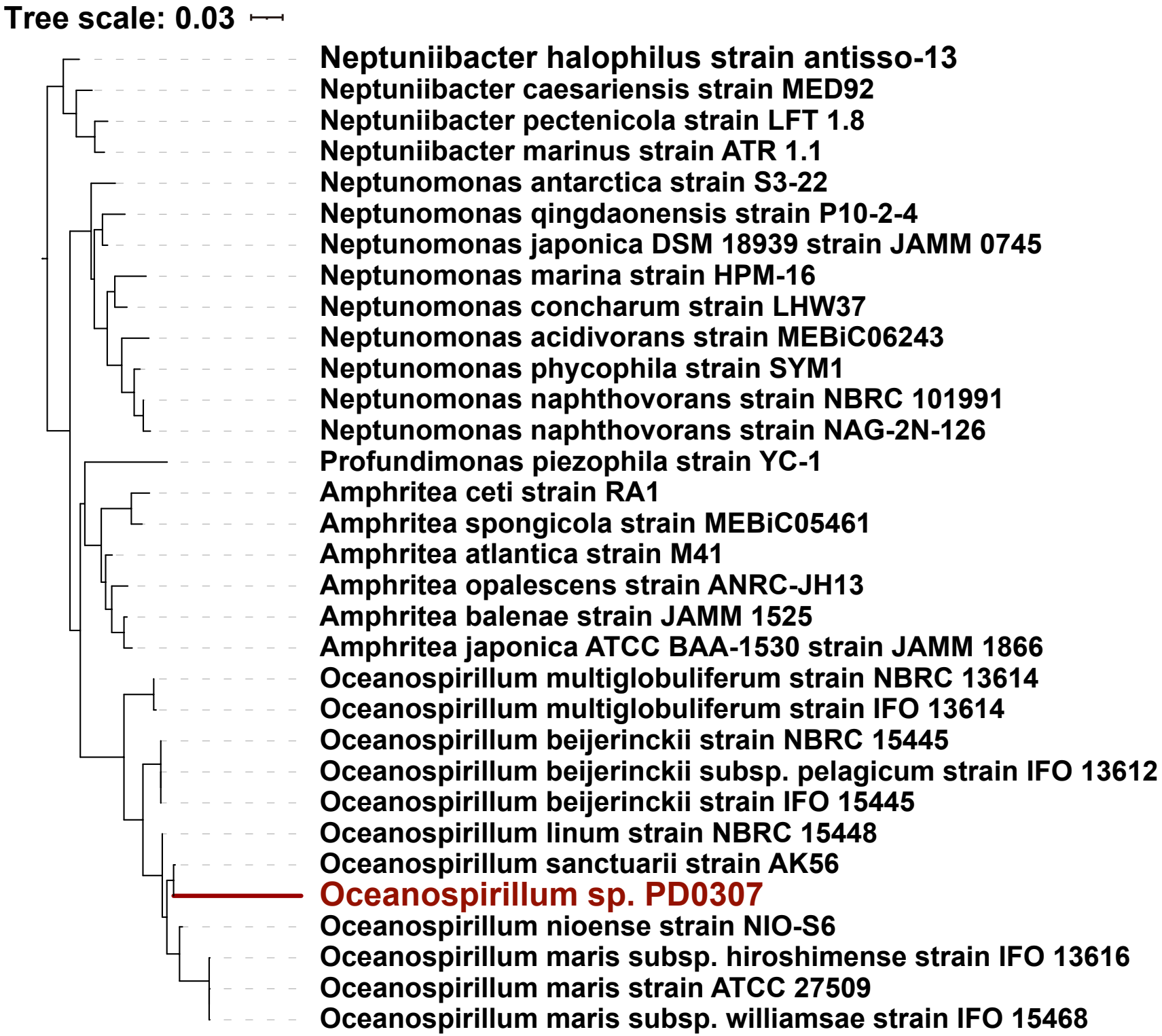
740

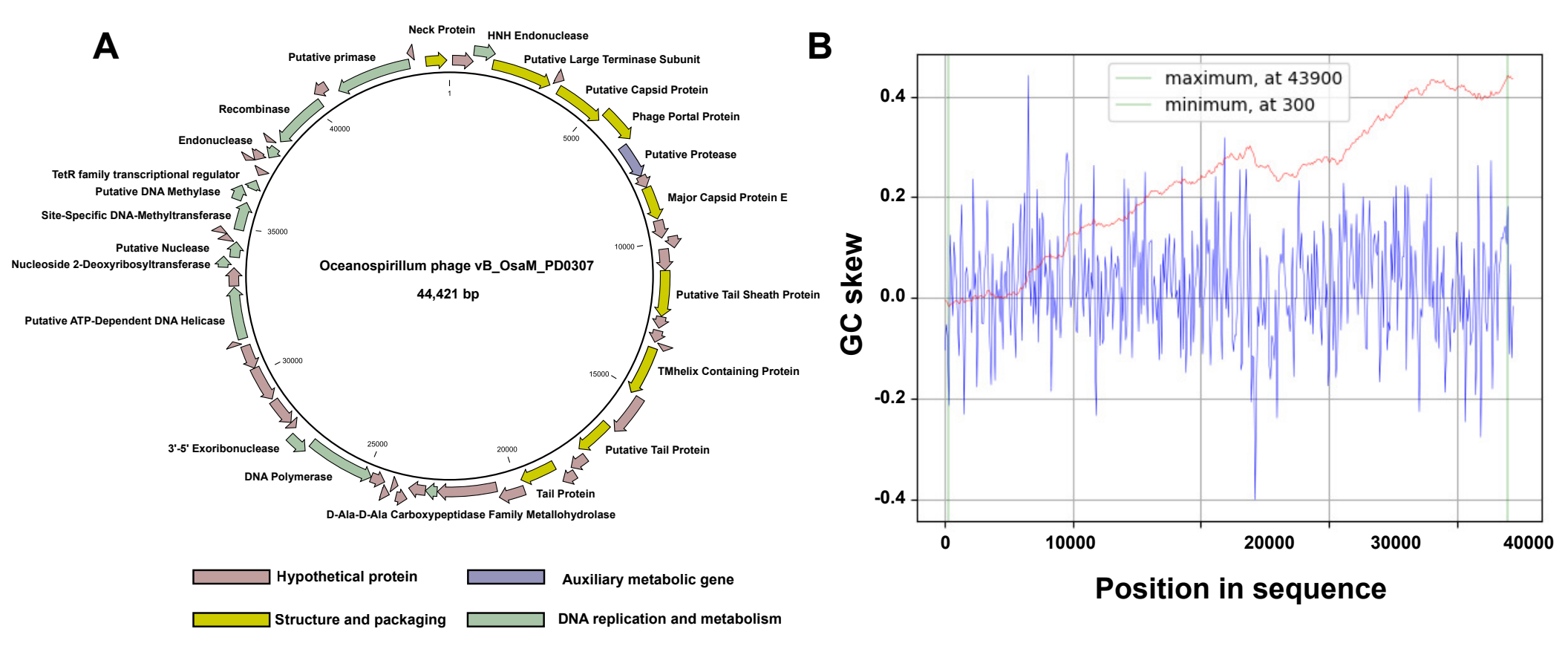
A

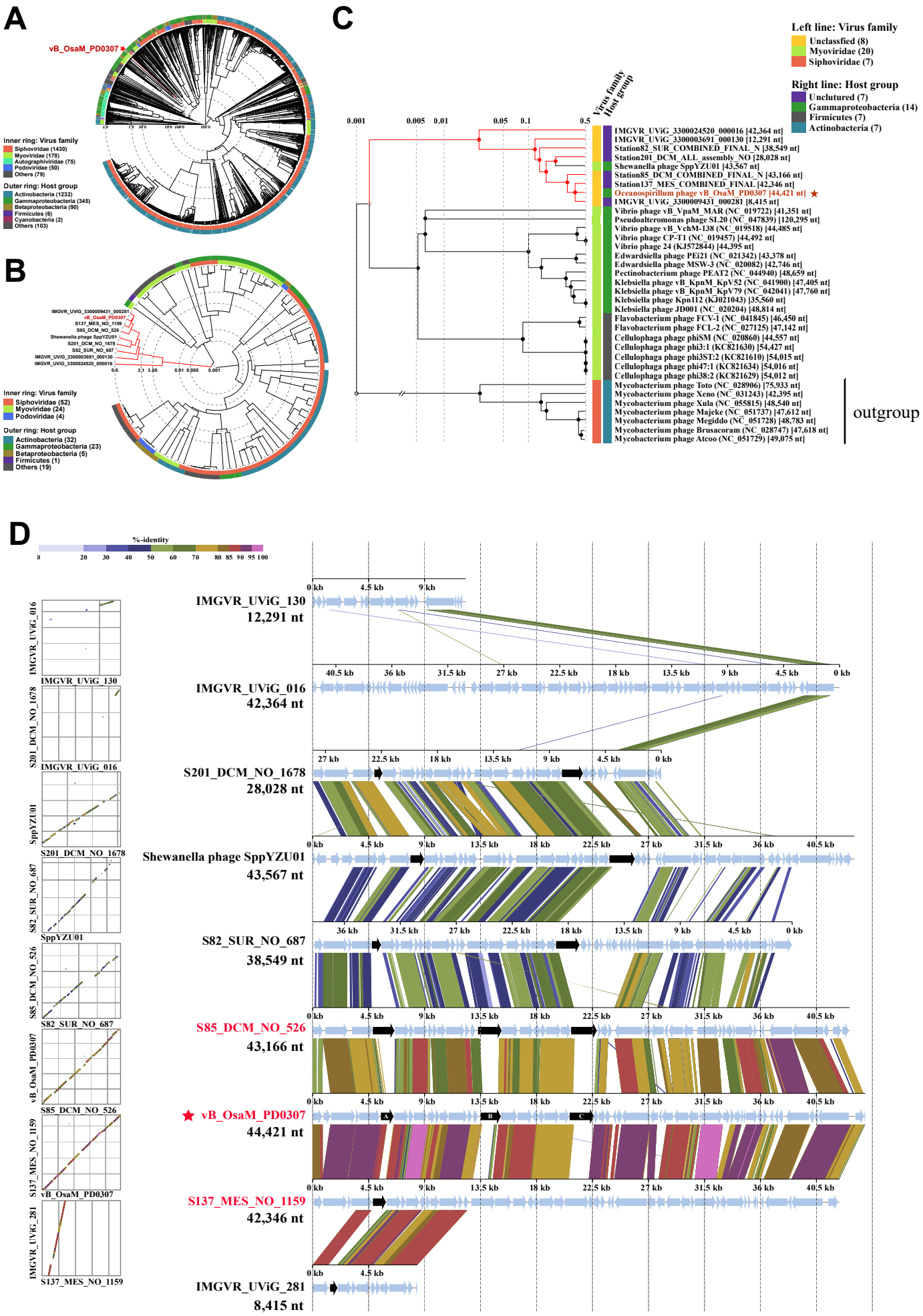


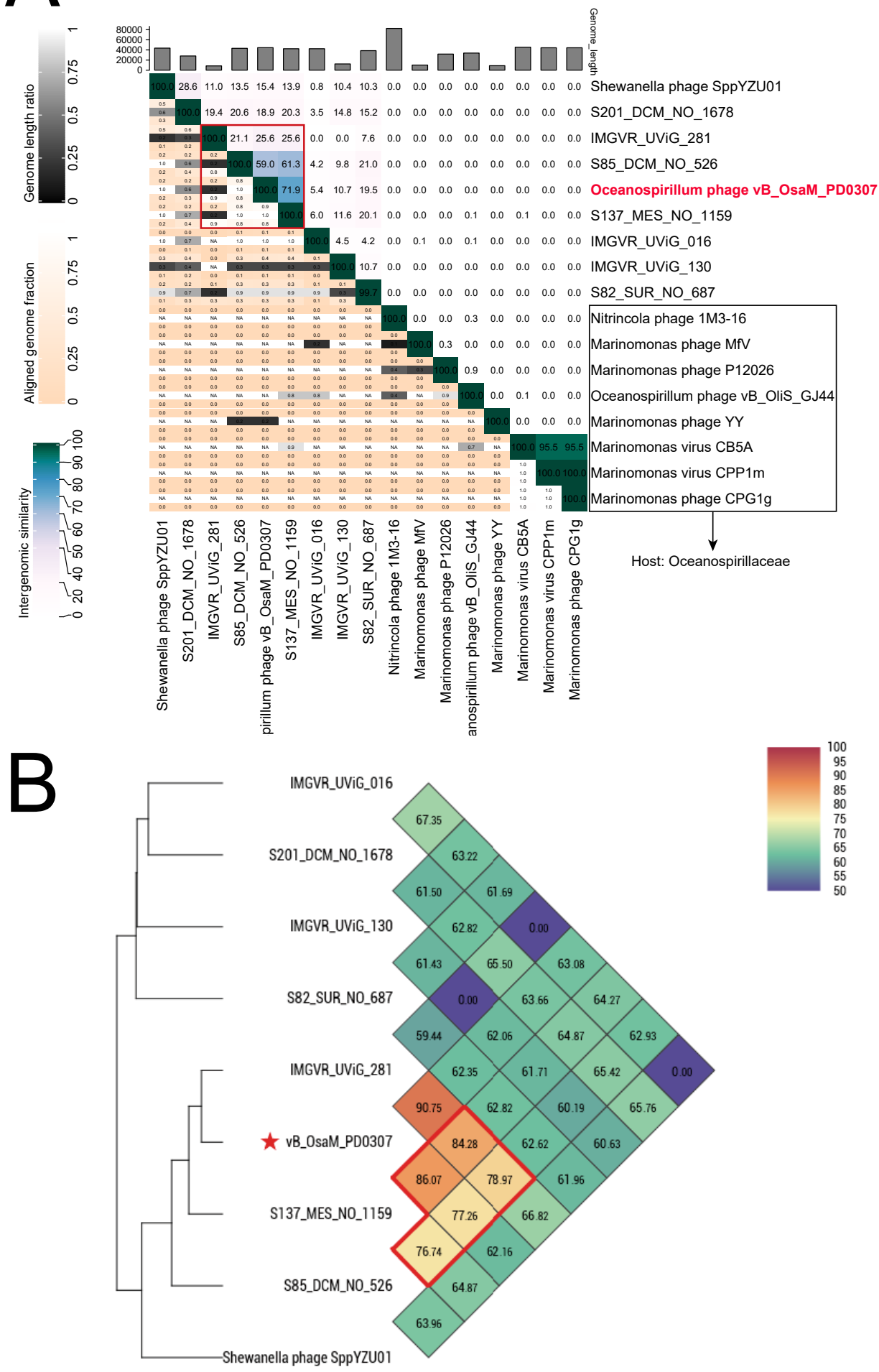
B




Tree scale: 0.03
Neptuniibacter halophilus strain antisso-13
Neptuniibacter caesariensis strain MED92
Neptuniibacter pectenicola strain LFT 1.8
Neptuniibacter marinus strain ATR 1.1
Neptunomonas antarctica strain S3-22
Neptunomonas qingdaonensis strain P10-2-4
Neptunomonas japonica DSM 18939 strain JAMM 0745
Neptunomonas marina strain HPM-16
Neptunomonas concharum strain LHW37
Neptunomonas acidivorans strain MEBiC06243
Neptunomonas phycophila strain SYM1
Neptunomonas naphthovorans strain NBRC 101991
Neptunomonas naphthovorans strain NAG-2N-126
Profundimonas piezophila strain YC-1
Amphritea ceti strain RA1
Amphritea spongicola strain MEBiC05461
Amphritea atlantica strain M41
Amphritea opalescens strain ANRC-JH13
Amphritea balenae strain JAMM 1525
Amphritea japonica ATCC BAA-1530 strain JAMM 1866
Oceanospirillum multiglobuliferum strain NBRC 13614
Oceanospirillum multiglobuliferum strain IFO 13614
Oceanospirillum beijerinckii strain NBRC 15445
Oceanospirillum beijerinckii subsp. pelagicum strain IFO 13612
Oceanospirillum beijerinckii strain IFO 15445
Oceanospirillum linum strain NBRC 15448
Oceanospirillum sanctuarii strain AK56
Oceanospirillum sp. PD0307
Oceanospirillum nioense strain NIO-S6
Oceanospirillum maris subsp. hiroshimense strain IFO 13616
Oceanospirillum maris strain ATCC 27509
Oceanospirillum maris subsp. williamsae strain IFO 15468


A
Neck Protein
HNH Endonuclease
Putative Large Terminase Subunit
Putative Capsid Protein
Phage Portal Protein
Putative Protease
Major Capsid Protein E
Putative Tail Sheath Protein
TMhelix Containing Protein
Putative Tail Protein
Tail Protein
D-Ala-D-Ala Carboxypeptidase Family Metallohydrolase
DNA Polymerase
3'-5' Exoribonuclease
Putative ATP-Dependent DNA Helicase
Nucleoside 2-Deoxyribosyltransferase
Putative Nuclease
Site-Specific DNA-Methyltransferase
Putative DNA Methylase
TetR family transcriptional regulator
Endonuclease
Recombinase
Putative primase
1
5000
10000
15000
20000
25000
30000
35000
40000
Oceanospirillum phage vB_OsaM_PD0307
44,421 bp
Hypothetical protein
Auxiliary metabolic gene
Structure and packaging
DNA replication and metabolism
B
maximum, at 43900
minimum, at 300
GC skew
0.4
0.2
0.0
-0.2
-0.4
0
10000
20000
30000
40000
Position in sequence


A
vB_OsaM_PD0307
Inner ring: Virus family
Siphoviridae (1430)
Myoviridae (178)
Autographiviridae (75)
Podoviridae (50)
Others (79)
Outer ring: Host group
Actinobacteria (1232)
Gammaproteobacteria (345)
Betaproteobacteria (90)
Firmicutes (6)
Cyanobacteria (2)
Others (103)
B
IMGVR_UViG_3300009431_000281
vB_OsaM_PD0307
S137_MES_NO_1159
S85_DCM_NO_526
Shewanella phage SppYZU01
S201_DCM_NO_1678
S82_SUR_NO_687
IMGVR_UViG_3300003691_000130
IMGVR_UViG_3300024520_000016
Inner ring: Virus family
Siphoviridae (52)
Myoviridae (24)
Podoviridae (4)
Outer ring: Host group
Actinobacteria (32)
Gammaproteobacteria (23)
Betaproteobacteria (5)
Firmicutes (1)
Others (19)
C
Left line: Virus family
Unclassfied (8)
Myoviridae (20)
Siphoviridae (7)
Right line: Host group
Unclutured (7)
Gammaproteobacteria (14)
Firmicutes (7)
Actinobacteria (7)
Virus family
Host group
0.001
0.005
0.01
0.05
0.1
0.5
IMGVR_UViG_3300024520_000016 [42,364 nt]
IMGVR_UViG_3300003691_000130 [12,291 nt]
Station82_SUR_COMBINED_FINAL_N [38,549 nt]
Station201_DCM_ALL_assembly_NO [28,028 nt]
Shewanella phage SppYZU01 [43,567 nt]
Station85_DCM_COMBINED_FINAL_N [43,166 nt]
Station137_MES_COMBINED_FINAL [42,346 nt]
Oceanospirillum phage vB_OsaM_PD0307 [44,421 nt]
IMGVR_UViG_3300009431_000281 [8,415 nt]
Vibrio phage vB_VpaM_MAR (NC_019722) [41,351 nt]
Pseudoalteromonas phage SL20 (NC_047839) [120,295 nt]
Vibrio phage vB_VchM-138 (NC_019518) [44,485 nt]
Vibrio phage CP-T1 (NC_019457) [44,492 nt]
Vibrio phage 24 (KJ572844) [44,395 nt]
Edwardsiella phage PEi21 (NC_021342) [43,378 nt]
Edwardsiella phage MSW-3 (NC_020082) [42,746 nt]
Pectinobacterium phage PEAT2 (NC_044940) [48,659 nt]
Klebsiella phage vB_KpnM_KpV52 (NC_041900) [47,405 nt]
Klebsiella phage vB_KpnM_KpV79 (NC_042041) [47,760 nt]
Klebsiella phage Kpn112 (KJ021043) [35,560 nt]
Klebsiella phage JD001 (NC_020204) [48,814 nt]
Flavobacterium phage FCV-1 (NC_041845) [46,450 nt]
Flavobacterium phage FCL-2 (NC_027125) [47,142 nt]
Cellulophaga phage phiSM (NC_020860) [44,557 nt]
Cellulophaga phage phi3:1 (KC821630) [54,427 nt]
Cellulophaga phage phi3ST:2 (KC821610) [54,015 nt]
Cellulophaga phage phi47:1 (KC821634) [54,016 nt]
Cellulophaga phage phi38:2 (KC821629) [54,012 nt]
Mycobacterium phage Toto (NC_028906) [75,933 nt]
Mycobacterium phage Xeno (NC_031243) [42,395 nt]
Mycobacterium phage Xula (NC_055815) [48,540 nt]
Mycobacterium phage Majeke (NC_051737) [47,612 nt]
Mycobacterium phage Megiddo (NC_051728) [48,783 nt]
Mycobacterium phage Brusacoram (NC_028747) [47,618 nt]
Mycobacterium phage Atcoo (NC_051729) [49,075 nt]
outgroup
D
%-identity
0
20
30
40
50
60
70
80
85
90
95
100
IMGVR_UViG_130
12,291 nt
IMGVR_UViG_016
42,364 nt
S201_DCM_NO_1678
28,028 nt
Shewanella phage SppYZU01
43,567 nt
S82_SUR_NO_687
38,549 nt
S85_DCM_NO_526
43,166 nt
vB_OsaM_PD0307
44,421 nt
S137_MES_NO_1159
42,346 nt
IMGVR_UViG_281
8,415 nt
A
B
C


A
Genome length ratio
Aligned genome fraction
Intergenomic similarity
Genome_length
Shewanella phage SppYZU01
S201_DCM_NO_1678
IMGVR_UViG_281
S85_DCM_NO_526
Oceanospirillum phage vB_OsaM_PD0307
S137_MES_NO_1159
IMGVR_UViG_016
IMGVR_UViG_130
S82_SUR_NO_687
Nitrincola phage 1M3-16
Marinomonas phage MfV
Marinomonas phage P12026
Oceanospirillum phage vB_OliS_GJ44
Marinomonas phage YY
Marinomonas virus CB5A
Marinomonas virus CPP1m
Marinomonas phage CPG1g
Host: Oceanospirillaceae
B
IMGVR_UViG_016
S201_DCM_NO_1678
IMGVR_UViG_130
S82_SUR_NO_687
IMGVR_UViG_281
vB_OsaM_PD0307
S137_MES_NO_1159
S85_DCM_NO_526
Shewanella phage SppYZU01

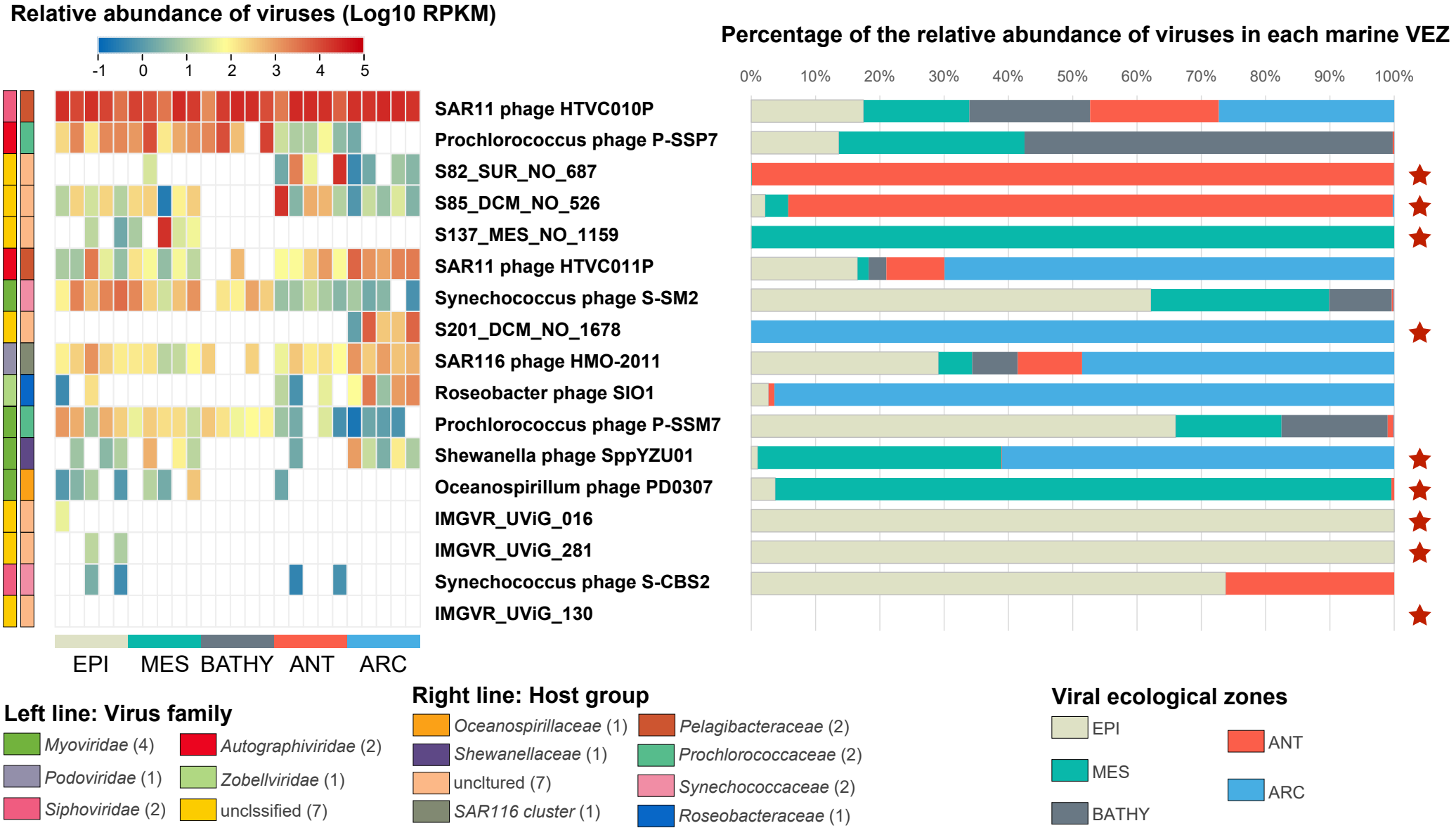

Relative abundance of viruses (Log10 RPKM)
-1
0
1
2
3
4
5
SAR11 phage HTVC010P
Prochlorococcus phage P-SSP7
S82_SUR_NO_687
S85_DCM_NO_526
S137_MES_NO_1159
SAR11 phage HTVC011P
Synechococcus phage S-SM2
S201_DCM_NO_1678
SAR116 phage HMO-2011
Roseobacter phage SIO1
Prochlorococcus phage P-SSM7
Shewanella phage SppYZU01
Oceanospirillum phage PD0307
IMGVR_UViG_016
IMGVR_UViG_281
Synechococcus phage S-CBS2
IMGVR_UViG_130
EPI
MES
BATHY
ANT
ARC
Percentage of the relative abundance of viruses in each marine VEZ
0%
10%
20%
30%
40%
50%
60%
70%
80%
90%
100%
Left line: Virus family
Myoviridae (4)
Autographiviridae (2)
Podoviridae (1)
Zobellviridae (1)
Siphoviridae (2)
unclssified (7)
Right line: Host group
Oceanospirillaceae (1)
Pelagibacteraceae (2)
Shewanellaceae (1)
Prochlorococcaceae (2)
unclture d (7)
Synechococcaceae (2)
SAR116 cluster (1)
Roseobacteraceae (1)
Viral ecological zones
EPI
ANT
MES
ARC
BATHY

Table 1 Amino acid identity of Oceanospirillum phage vB_OsaM_PD0307 and other night homologous sequences.

Abbreviation	Identity Frequency		Bit-score Frequency	
	mean	median	mean	median
IMGVR_UViG_281	85.08	93.7	428.9	346.5
S137_MES_NO_1159	83.36	87.01	461.6	343
S85_DCM_NO_526	72.9	79.04	428.2	282
Shewanella phage SppYZU01	48.33	47.22	260.6	173
S201_DCM_NO_1678	48.24	47.73	260.4	144
IMGVR_UViG_130	46.53	46.51	247.9	125
S82_SUR_NO_687	45.45	44.27	249.1	159
IMGVR_UViG_016	40.97	37.32	127.7	30